Science

AAAS

ARCHAIC SURVIVOR

Fossil, culture, and habits
of a pre-Neanderthal *Homo*

NEWS FROM Science

Up-to-the-minute research and policy news you won't find in print

Visit us online to read all the news coverage that there just wasn't enough room to print in this issue.
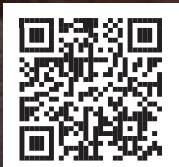
ScienceMag.org/news

Advanced technology and tomorrow's scientists drive innovation

SCIEX runs programs that support a more diverse population of future investigators and develops sophisticated analytical methods and devices for them to use.

Scientists generate innovations that change the world—or even save it. Ensuring a powerful pipeline of innovations requires an ongoing flow of young scientists and advances in technology that support exciting experiments.

The Extraordinary Grace project, developed by SCIEX, a biotech company based in Framingham, Massachusetts, which specializes in mass spectrometry and capillary electrophoresis, promotes diversity in tomorrow's innovators. This project encourages scientists to guide a young woman, "Grace," through her journey toward a career in science. "Extraordinary Grace tells the world that science is accessible to everybody—no matter what gender, race, or skin color," says Naomi Diaz, senior product manager at SCIEX. "It shines a light on opportunities and inclusion." Scientists can take a few minutes and participate in the project survey (https://extraordinarygrace.com/questions).

Those opportunities should include research experiences in education. "One of the most important things in the very early days is to get this amazing feeling of discovering something new," says Christoph Messner, training fellow at the Francis Crick Institute in London, United Kingdom. "That's what drives us as scientists."

To make those new discoveries, Messner says "we need to have the right technology, and that's why technical innovations are key."

Improving acquisition

Collecting more information is one area of innovation that scientists seek to improve. For example, SWATH acquisition from SCIEX enhances data collection for analysis with mass spectrometry (MS). Messner says that SWATH acquisition is "a great technology because you can measure a lot of molecules in parallel in an unbiased way, and you can do that quantitatively." He calls SWATH "a game changer in terms of applications."

In one project, Messner and his colleagues used SWATH acquisition to analyze 5,000 different knockout strains. "That was unimaginable to do before," he says. In the future, he expects to increase the sample size to as many as hundreds of thousands of samples a day.

The broad dynamic range enabled by SWATH acquisition also benefits Messner's work. In plasma samples, for instance, SWATH can even analyze low-level proteins among many high-level ones.

The EAD edge

In tandem MS, a sample is fragmented before analysis. Traditionally, scientists performed fragmentation with collision-induced dissociation (CID), or with complex, specialized electron-based methods, but a new technique, electron-activated dissociation (EAD), provides many more options.

"EAD is a very flexible and reagent-free fragmentation technique that can be modulated to optimize the fragmentation patterns and obtain the information that you're looking for," says Diaz. "It's an orthogonal fragmentation tool [that complements] CID, and it improves sensitivity at the MS/MS level by providing permutation information that you might not be able to see with CID."

Zooming ahead with Zeno

The Zeno trap is a novel trap included in the SCIEX ZenoTOF 7600 system that releases trapped ions based on their potential energy, so that a wider range of ions are analyzed in each TOF cycle. EAD fragmentation can be paired with Zeno trap pulsing, and this combination "provides an even greater boost on MS/MS sensitivity to levels we really haven't seen before in an accurate mass system," Diaz explains. "It allows you to identify and quantify compounds at very low levels without compromising your resolution and speed."

Messner gained many advantages from this technology. For one thing, he found about a 10-fold increase in sensitivity with the ZenoTOF platform. "So, you can really reduce the sample amount," he says. Plus, Messner points out that Zeno enables reduced accumulation times, which allows more MS/MS acquisitions in less time.

When asked how this technology will impact future research, Messner mentions posttranslational modifications (PTMs), such as glycosylation patterns. "We don't understand much of the roles of some PTMs, because we lack methods that can quantify and characterize them," he says. With the increased sensitivity of Zeno technology combined with the efficiency of EAD, Messner hopes to better characterize PTMs, even in single cells.

Having already used this technology for high samples sizes, Messner envisions using tomorrow's even higher sample sizes to explore new areas. "We can measure many more samples, so we can increase the sample size in different projects, which of course allows us to go to population scale and measure, for example, how proteins in blood change in response to disease, lifestyle or therapies," he says.

For drug screening in cell lines, scientists want to analyze the proteome response to thousands of different drugs in different concentrations, maybe even over time. To do that, Messner says, "We need higher throughput, and I believe Zeno can potentially give us that."

The Zeno revolution is now. The Zeno trap and EAD are a powerful combination of unparalleled MS/MS sensitivity and a step change in fragmentation technology. Learn more at sciex.com/zenorevolution.

Sponsored by

SCIEX

The Power of Precision

CONTENTS

1382

1402

ON THE COVER

Digitally reconstructed skull of a hominin from 130,000 years ago. This previously unrecognized *Homo* group, identified at Nesher Ramla, Israel, sheds light on Neanderthal origins and the evolutionary history of the *Homo* lineage. Late members of this group inhabited the Levant contemporaneously with *Homo sapiens* and interacted with them. See pages 1395, 1424, and 1429. *Illustration: Ariel Pokhojaev, Rachel Sarig, Hila May, Israel Hershkovitz*

Work undone—HIV/AIDS in the USA

This month marks a somber anniversary—40 years since the first reported cases in the United States of what would later become known as AIDS. There have been tremendous basic, clinical, and prevention advances in HIV science over the past four decades. Yet, despite widespread messaging that the United States is on track to "end AIDS," the latest trends in infection tell a different story. One fundamental reality underlies the country's failure to achieve control of the HIV epidemic and could undermine efforts to end AIDS—the lack of access to health care for all Americans.

The 2019 (pre–COVID-19) map of AIDS in America has shifted. According to a May 2021 report from the US Centers for Disease Control and Prevention (CDC), the South—from Texas through Louisiana, Mississippi, Georgia, Florida, and the Carolinas to Washington, DC—is now the most affected region. And the epidemic is markedly concentrated, with 66% of new diagnoses reported among gay, bisexual, and other men who have sex with men (MSM).

Although there was a modest overall decline in new diagnoses of roughly 8% nationally, much of this was attributed to a 33% decline in MSM aged 13 to 24. This heartening trend, attributed to the increasing use of pre-exposure prophylaxis (PrEP) in this age and risk group, masks a challenging reality. PrEP has had a substantial impact on HIV acquisition risks among white MSM, but rates of access and use are sharply lower among Black and Latinx MSM. The lifetime risk acquisition for a young Black gay man in America is an astonishing 50%. The 2019 data suggest that this health disparity is widening.

The geographic and racial disparities are just as marked for women. Black women accounted for 13% of US women in 2019, but for 55% of all new infections in women. This reality is summarized starkly in the CDC report: "African Americans continue to face rates of infection that are more than 8 times as high as whites, and Hispanics/Latinos face rates that are almost 4 times as high, in large part because they experience the greatest barriers to accessing prevention and care services."

What are the drivers of these trends and how can we do better? Numerous studies make clear that individual-level risks for HIV cannot account for these disparities. It was reported a decade ago that Black MSM had much lower individual-level risks for HIV acquisition compared to white MSM. Black MSM, however, were more likely to have undetected HIV infection or untreated HIV disease, were poorer, were less likely to have health insurance, and were more likely to have an untreated sexually transmitted infection, notably syphilis. These attributes are markers of lack of access to health care and of the social determinants of health. Indeed, the current geography of HIV in the United States reflects those states that refused to expand Medicaid benefits through the Affordable Care Act. What used to be called the cotton belt, and then the syphilis belt, is now the HIV belt. It should surprise no one that this is the region with the lowest rates of COVID-19 immunization nationally—nor that racial and ethnic minority burdens for COVID-19 so swiftly replicated those long known for HIV/AIDS.

Tools exist to address these marked health disparities, but they have to be implemented where they will matter most—where new diagnoses are occurring now. Access to effective prevention, including PrEP, must be expanded to those at risk in contexts of culturally competent care. And it is necessary to implement the suite of tools to prevent transmission among people who inject drugs, including needle and syringe exchanges, substance use treatment on demand, and antiviral therapy for those living with HIV. These are all evidence-based measures, which should have been taken to scale decades ago. For these interventions to work, the community of people affected by HIV must be front and center in our shared efforts.

If we are serious about ending AIDS—and we must be—we must extend the health care franchise to all Americans. Whether we have the political will to achieve this long-held goal is uncertain, but the Biden administration has already expanded health care access through the Affordable Care Act and is committed to addressing the systemic racial, ethnic, sexual, gender, and minority inequalities that are at the heart of AIDS in America. Entering the fifth decade of this epidemic, we must finally address the root causes of health disparities. The country must finally accept that health care is a human right from which no one should be excluded.

–Chris Beyrer

> "…we must extend the health care franchise to all Americans."

Chris Beyrer is the Desmond M. Tutu Professor in Public Health and Human Rights and a professor in the Department of Epidemiology at the Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA. cbeyrer@jhu.edu

Science

AAAS

Science serves as a forum for discussion of important issues related to the advancement of science by publishing material on which a consensus has been reached as well as including the presentation of minority or conflicting points of view. Accordingly, all articles published in *Science*—including editorials, news and comment, and book reviews—are signed and reflect the individual views of the authors and not official points of view adopted by AAAS or the institutions with which the authors are affiliated.

FDA's green light, science's red light

Alzheimer's disease (AD) afflicts some 6 million Americans with progressive cognitive impairment and personal anguish while imposing a huge economic burden on society. Everyone wants to find a way to help slow or even halt this disease. But there will be no quick fix. Responding to mounting pressure, the United States Food and Drug Administration (FDA) jumped the gun by granting accelerated approval this month to Biogen's pricey, questionably effective, and possibly harmful new drug aducanumab—a decision supported by not one of the 11 members of the agency's Expert Advisory Committee. Even worse, the approval may divert funding into a therapeutic dead end and away from approaches that might actually work.

As a member of this advisory committee, who resigned in protest over the decision to grant approval, I am still trying to fathom how this happened. No doubt, the FDA faced a difficult decision. The public pressure must have been immense, and the influence of industry on the FDA in general has been a growing concern. Any trickle of hope about this drug has been magnified far beyond the facts. The Alzheimer's Association has pushed this bandwagon and already stated that aducanumab "demonstrates that removing amyloid from the brain may delay clinical decline in people living with Alzheimer's." The science does not support such a delay.

There is a scientific basis for trying to develop this type of drug. Rare genetic forms of AD caused by mutations result in early-onset dementia related to substantial brain deposition of amyloid-β or tau protein. Yet, most people with AD do not suffer from a genetic form. Rather, sporadic forms feature early deposition of abnormal amyloid-β long before cognitive problems arise and later develop abnormal deposition of tau that more closely coincides with cognitive impairment. Many studies investigating drugs designed to clear abnormal brain amyloid-β in people with sporadic AD failed to demonstrate clinical benefit (the alleviation of symptoms).

Despite these bad odds, a drug targeting amyloid-β was presented for approval to the FDA. On 6 November 2020, our FDA advisory committee reviewed Biogen's application for aducanumab, primarily on the basis of a two-part study that had been stopped early because of futility—the chances of clinical benefit were very small if the study continued to the planned conclusion.

But then, the data were reanalyzed and Biogen proposed that because one part of the study was positive, though the other was not, that was sufficient for FDA approval. Never mind that the side effects of the proposed dose included localized brain swelling in 35% of clinical trial participants and microhemorrhages in 20%.

When all this was put to a vote by the advisory committee, 10 voted no, 1 voted uncertain, and no one voted yes.

And yet, the FDA granted accelerated approval of aducanumab for treatment of AD, merely requiring Biogen to do a prospective study over the next 9 years to confirm if there is some clinical benefit. Even worse, the FDA changed the standard for determining this benefit from clinical evidence that the drug actually helps to evidence that the drug simply reduced brain amyloid-β.

Although all of this may be well and good for Biogen with a potential $56 billion dollars for the first year of treatment in 1 million people with AD, this decision may impair future research into better treatments for AD. Studies may be required to compare a new drug with aducanumab rather than placebo, which could potentially bias the research. Furthermore, enthusiasm from potential volunteer participants or funders for new treatments may wane owing to the false belief that effective treatment already exists. And, the matter of economics cannot be overlooked. The billions of dollars spent on aducanumab may be better invested in developing stronger evidence for aducanumab or alternative therapies. These potentially serious issues could delay investigation and implementation of a truly effective therapy for AD.

The FDA and the advisory committee have a responsibility to help protect vulnerable patients and their families, not just from sketchy drugs but also from false hopes. That can mean making hard decisions that disappoint them in the short term to increase the chances of ultimately finding drugs that work.

–Joel S. Perlmutter

> "...the approval may divert funding...from approaches that might actually work."

Joel S. Perlmutter is a professor of Neurology, Radiology, Neuroscience, Physical Therapy, and Occupational Therapy at Washington University in St. Louis, St. Louis, MO, USA. perlmutterjoel@wustl.edu

NEWS

IN BRIEF

Edited by **Jeffrey Brainard**

Workers clean seats in Japan's Yokohama Stadium, which will host baseball games during the Olympics.

COVID-19

Tokyo Olympics to limit spectators

Tokyo Olympics officials this week decided to limit attendance at the games to 50% of venue capacity, with a maximum of 10,000 people per event, brushing aside experts' recommendation days earlier that completely barring spectators would be the best way to limit COVID-19 infections. Organizers decided in March to close the games to overseas visitors but had not ruled out domestic spectators. An unofficial group of medical and public health scientists that has been meeting outside the government's formal COVID-19 advisory framework also urged the government to be ready to reimpose a state of emergency if there are signs that cases are rebounding or medical facilities are being overloaded.

COVID-19 drug studies get boost

BIOMEDICINE | President Joe Biden's administration said last week it will allot $3 billion in COVID-19 relief funds toward developing new antiviral medications to treat pandemic virus infections. Such medications promise to protect people who are reluctant to get a shot, are unlikely to benefit fully from vaccines because they are immunocompromised, or live in countries with little access to them. The money includes $1.3 billion for preclinical and clinical research and nearly $700 million for development and manufacturing. As much as $1.2 billion would go toward developing novel treatments against viruses other than SARS-CoV-2 that could cause pandemics.

Guinea Ebola outbreak ends

INFECTIOUS DISEASES | Guinea's Ebola outbreak, first detected in February, officially ended on 19 June, after 42 days had elapsed since the last identified patient recovered. Sixteen cases were confirmed in the outbreak, with an additional seven suspected. Twelve people died. The outbreak was the first in West Africa since the devastating outbreak between 2013 and 2016 that killed more than 11,000 people and infected more than 28,000 across six countries in the region. Both outbreaks started in the Nzérékoré region, in southeastern Guinea. Roughly 11,000 people were vaccinated in the effort to contain the new outbreak.

Mistrial in China ties case

RESEARCH COLLABORATION | The first criminal trial involving the China Initiative, a 3-year effort that has led to the prosecution of several U.S. scientists, ended last week when the judge declared a mistrial. A Tennessee jury failed to reach a verdict on federal charges that Anming Hu, a former bioengineering researcher at the University of Tennessee, Knoxville, had lied about his ties to Chinese entities. The China Initiative is meant to keep federally funded research from being stolen, but some civil rights groups have accused the U.S. government of improperly targeting scientists of Chinese ancestry. Hu was charged in connection with a 2011 law prohibiting NASA from funding any work involving the Chinese government. The U.S. government must now decide whether to retry or drop the case. Hu, a Canadian citizen with a U.S. work visa, has been out of work since the university fired him after his arrest in February 2020.

NOAA chief confirmed

LEADERSHIP | The U.S. Senate last week confirmed Rick Spinrad as administrator of the National Oceanic and Atmospheric Administration (NOAA), one of the country's premier climate agencies. NOAA went without a confirmed chief for the entire Trump administration, the longest drought in its history. Spinrad, an oceanographer at Oregon State University, Corvallis,
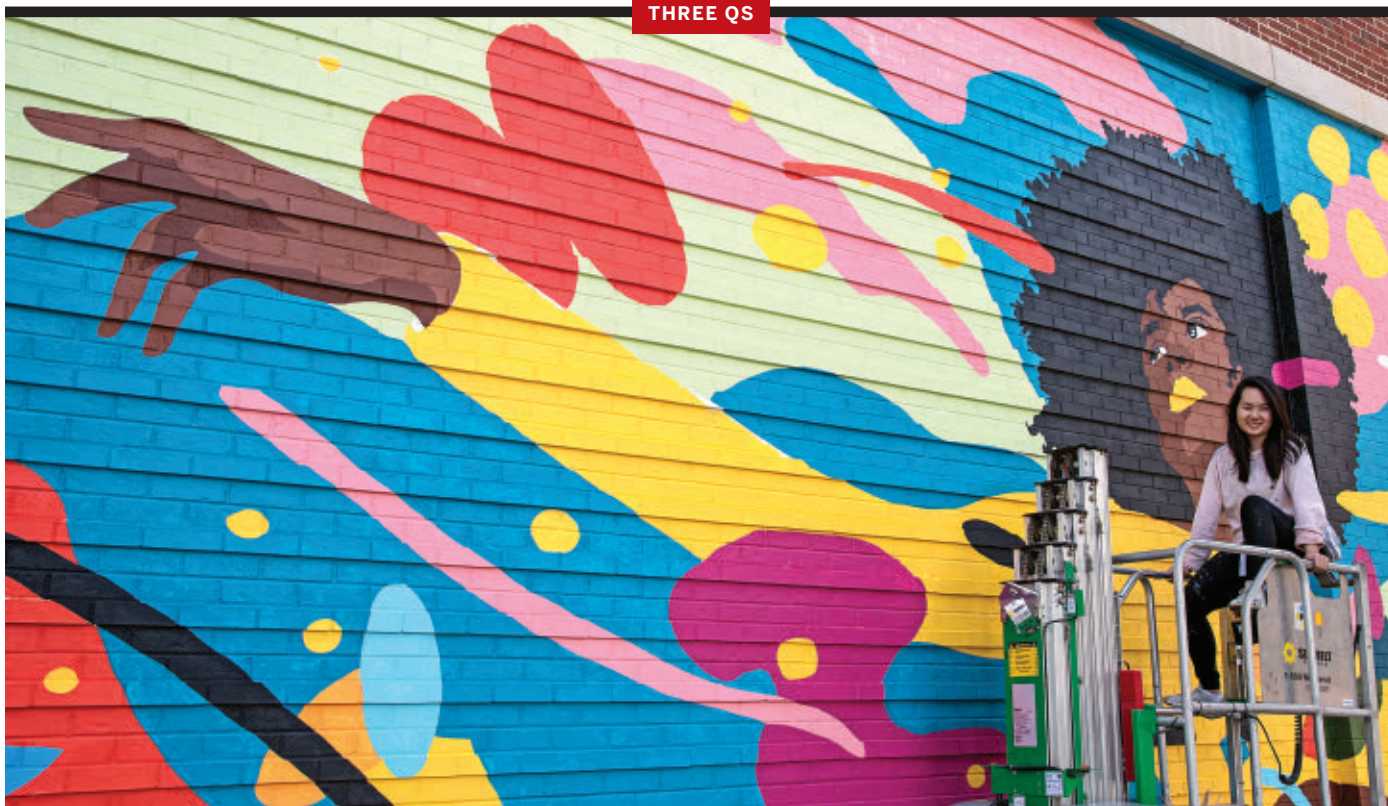
Muralist celebrates female scientists

Amanda Phingbodhipakkiya learned about the importance of good communication from her father, who enticed wary diners to try new dishes in his Thai restaurants in Atlanta. While working in an Alzheimer's disease lab after college, the neuroscience major (above) became keenly aware of how poorly scientists—herself included—communicate with the public. So she abandoned her Ph.D. ambitions and obtained a master's degree in fine arts. That launched a career in science-focused art and design that led her to projects like Beyond Curie, which, she says, "highlights badass women in STEM."

Now, she is creating murals in public spaces in 10 U.S. cities spotlighting female physicists, astronomers, glaciologists, and oceanographers, in a project sponsored by the Heising-Simons Foundation. She finished the latest installation, in Washington, D.C., this month. As a nearby plaque explains, the 5-by-23-meter painting, inspired by the work of Duke University particle physicist Ayana Arce, imagines women building bridges to each other like quarks that find each other after intense proton-proton collisions. A longer version of this interview is at **https://scim.ag/murals**.

Q: How do you see your career transformation now?
A: Art and science are simply different facets of creativity for me. Each one is an iterative process that requires tremendous focus, resourcefulness, and perseverance. That spark of insight I felt doing research in a neuroscience lab is the same one I feel when I'm working on a new installation or mural.

Q: What led you to propose murals?
A: The power of murals is that they are impossible to ignore. Public art doesn't sit behind gallery walls. It's out there in the wild for anyone to enjoy. That's really the approach we need to take with science literacy and communication. Because so much of the time, we're in an echo chamber.

Q: What mural comes next?
A: The next will be in San Carlos. It's based on the research of astrophysicist Chung-Pei Ma at the University of California, Berkeley. She studies supermassive black holes. I am depicting women as illuminators in this mural. There is enormous sexism and racism in science and technology. The murals uplift the voices and stories that have typically gone unseen.

previously served as NOAA's chief scientist and led its research office and the National Ocean Service.

NSF aims to 'democratize' NMR

FACILITIES | The U.S. National Science Foundation (NSF) said last week it will spend $40 million to set up a national network for high-field nuclear magnetic resonance (NMR) spectroscopy, which researchers use to determine the structure of complex molecules and materials. The network is intended to "democratize" these studies by increasing access to the most powerful instruments, a priority set out in 2013 by the U.S. National Research Council, says Jeffrey Hoch, a structural biologist;

he will head the network's central hub at UConn Health, which already houses an NMR data processing center. The UConn center will help investigators around the country schedule access to an existing collection of 26 NMR machines up to 900 megahertz and, eventually, to two new, NSF-funded, 1.1-gigahertz NMRs scheduled to start up within 3 years—one at

the University of Georgia, the other at the University of Wisconsin, Madison. Because a molecule's structure determines how it functions, NMR has become a mainstay for studying everything from enzymes to catalysts and battery materials.

Research money goes to refinery

MEXICO | Researchers in Mexico are outraged by the government's plan, announced this month, to use money gained from slashing science programs to buy an oil refinery in Texas. In 2020, the government enacted a law that terminated 109 trust funds, some of which supported student scholarships, international collaborations, and equipment in climate science and other fields. The country's ruling party, Morena, said then that it would use the almost 68 billion pesos ($3 billion) in funds to help manage the coronavirus pandemic. But on 4 June, President Andrés Manuel López Obrador changed course, announcing that the government instead would use part of the money to purchase the Deer Park Refinery to enhance Mexico's energy independence (despite its location in a different country). Scientists who depended on the trust fund money argue that the diversion of nonfederal money would be illegal. They also complain that the move would conflict with Mexico's pledge to join other countries in reducing reliance on fossil fuels to control carbon emissions and global warming.

Registered reports rate high

PUBLISHING | "Registered reports"—papers that journals review and accept in principle based on methods alone, before results are known—measure up on quality, a study reports. The findings, published this week in *Nature Human Behaviour*, offer some of the first evidence for the purported benefits of the publishing format, now offered by 295 journals. Supporters say it promises to improve scientific rigor and reduce publication bias, but some researchers worry the format could stifle creativity. Asked to review 29 registered reports and 57 comparison papers published in the standard format, 353 researchers—blinded to each paper's publishing format—rated the registered reports as more rigorous and the methods as higher in quality. Scores for creativity were similar for both sets.

Africa gets a shot at mRNA skills

PUBLIC HEALTH | A South African consortium working with the World Health Organization (WHO) will soon launch a program to teach scientists from all over the continent how to make COVID-19 vaccines with messenger RNA (mRNA) at a training plant. Two South African companies, Biogen and Afrigen Biologics and Vaccines, will work with academic scientists who have offered to share their mRNA intellectual property and know-how

to create a "technology transfer hub." WHO hopes the makers of the two proven, highly effective mRNA COVID-19 vaccines, the Pfizer-BioNTech collaboration and Moderna, will join the consortium; if they do, plants in South Africa could start to produce doses in as soon as 9 months. The project could help narrow the wide gap between rich and poor countries in availability of mRNA COVID-19 vaccines, WHO said this week. The longer term vision is to build plants in several African countries that can make mRNA vaccines against several diseases.

Virologist's stalker found dead

COVID-19 | Prominent Belgian virologist Marc Van Ranst of KU Leuven returned home with his family on 20 June after prosecutors confirmed the death of a man who had threatened to kill him for advising the government to take strict measures to control COVID-19. Van Ranst, a top adviser to the Belgian government, had spent more than 5 weeks in a safe house while police searched for Jürgen Conings, 46, a former soldier with extreme-right sympathies who had taken weapons from a military barracks and written farewell letters before disappearing on 17 May. Conings's body was found in a forest near Maaseik, Belgium, a village close to the Dutch border; he had died by suicide using a firearm, prosecutors said on 21 June.
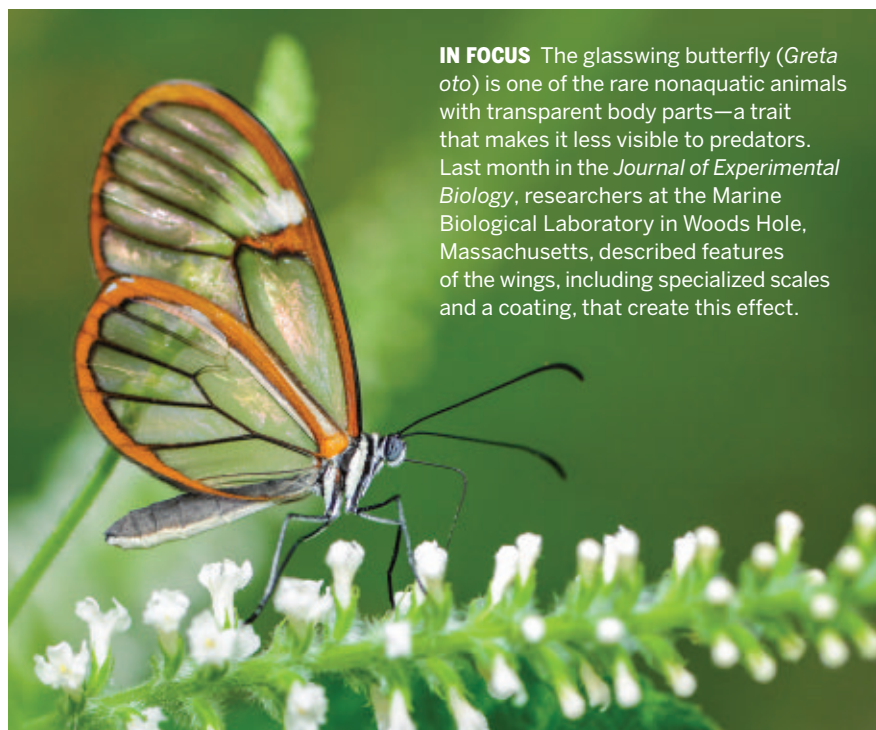
IN FOCUS The glasswing butterfly (*Greta oto*) is one of the rare nonaquatic animals with transparent body parts—a trait that makes it less visible to predators. Last month in the *Journal of Experimental Biology*, researchers at the Marine Biological Laboratory in Woods Hole, Massachusetts, described features of the wings, including specialized scales and a coating, that create this effect.

CORRECTIONS

The 18 June article titled "Accusations of colonial science fly after eruption" (p. 1248), about the Goma Volcano Observatory in the Democratic Republic of the Congo, mistakenly implied that a known number of GVO's staff members supported allegations by their union against GVO's Congolese leadership and European partners, and also that the European Center for Geodynamics and Seismology in Luxembourg processed seismic and GPS data before sending them to GVO. In addition, the article used imprecise wording to describe a dispute over whether the 22 May eruption of nearby Mount Nyiragongo could have been foreseen.

The images accompanying the 18 June story titled "Marshes on the move" (p. 1254) omitted the photographer's affiliation. The correct credit is: Michael O. Snyder is a Bertha Climate Journalism Fellow based in Charlottesville, VA.

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

Police officers stop cars at a checkpoint to prevent unauthorized travel into and out of Lisbon, Portugal, on 18 June.

IN DEPTH

COVID-19

Delta variant triggers new phase in the pandemic

Scientists are probing why a fresh set of viral mutations is taking the world by storm

By **Kai Kupferschmidt** *and* **Meredith Wadman**

When the coronavirus variant now called Delta first appeared in December 2020, in the Indian state of Maharashtra, it did not seem all that remarkable. But when it descended on New Delhi a few months later, its impact was devastating, with almost 30,000 cases reported daily in late April. "Suddenly … it is dominant and completely sweeps away Alpha," which until then was most prevalent in the city, says Anurag Agrawal, who leads the Institute of Genomics and Integrative Biology in New Delhi.

New Delhi seemed unlikely to suffer a big new outbreak because so many of its residents had already been infected or vaccinated, Agrawal says. But those protections seemed to barely slow Delta, which is more transmissible and may evade immunity, he says: "It went from a 10-foot wall around the city to a 2-foot wall you could just walk over."

From New Delhi, the variant has quickly spread, and it now looks set to sweep the globe in what could be a devastating new wave. In the United Kingdom, Delta already makes up more than 90% of all infections; it has driven COVID-19 case numbers up again after a dramatic decline and led the government last week to postpone the final stage of its reopening plan. A Delta-driven resurgence

in Lisbon prompted the Portuguese government to enact a 3-day travel ban between the city and the rest of the country. Delta also appears also to be causing surges in Russia, Indonesia, and many other countries. In the United States, where its prevalence is now estimated to be at least 14%, the Centers for Disease Control and Prevention declared Delta a "variant of concern" on 15 June.

The surge has set off a frenzy of research to understand why Delta appears to spread so much faster than the three other variants of concern, whether it is more dangerous in other ways, and how its unique pattern of mutations, which cause subtle changes in its proteins, can wreak havoc. Delta's arrival has

Evasive maneuver

The spike protein's N-terminal domain (left) includes a "supersite" where powerful antibodies latch on to the virus (middle). Mutations there (right) can prevent them from binding.

Spike protein N-terminal domain — Neutralizing antibody — N-terminal domain with mutations

also brought fresh attention to the potential of SARS-CoV-2 to evolve and adapt in the months and years ahead. For the moment, Delta is a particular threat to the poorest countries with little or no access to vaccines, says Soumya Swaminathan, chief science adviser at the World Health Organization. "My immediate worry is what will happen when Delta is introduced into Africa," she says. "You could end up with explosive outbreaks."

Research by Public Health England underscores Delta's ability to spread. Compared with Alpha, which appeared in the United Kingdom in 2020, "You're getting estimates of 50% or 100% more transmission," says Adam Kucharski, a modeler at the London School of Hygiene & Tropical Medicine.

But Kucharski says reduced protection from vaccines may play a role as well. Data from England and Scotland indicate that both the Pfizer-BioNTech and AstraZeneca vaccines offer slightly less protection against symptomatic infections from the new variant than from Alpha. People who have received just one shot of vaccine—as many U.K. residents have—are especially vulnerable. (Two doses of either vaccine still offer the same high level of protection from hospitalization from Delta.) How well the many other vaccines now in use around the world protect against it is unclear, and there are few data on the protection resulting from a prior bout of COVID-19.

The two effects—increased transmissibility and immune evasion—are hard to disentangle, but "I would argue the Delta variant has been driven by its transmissibility, not its ability to escape immunity," says Jeremy Farrar, head of the Wellcome Trust. "If Alpha really is approximately 50% more transmissible than the wild type strain, and Delta is 50% more transmissible again than Alpha, we're talking about a virus that's more than twice as transmissible as the initial strain," adds Aris Katzourakis, an evolutionary virologist at the University of Oxford. That would mean countries and populations with low vaccination rates are likely to see big new outbreaks. Indeed, if the faster spread "is entirely down to the fundamentals of the virus, that's catastrophic news for the rest of the world," Kucharski says.

On top of this, Delta may be more likely to put unvaccinated people in the hospital than Alpha. Early data from the United Kingdom suggest the risk of hospitalization may be twice as high. Together these characteristics could cause huge problems in Africa, Swaminathan says. "There isn't going to be enough oxygen there, there aren't going to be enough hospital beds. And we already know that outcomes for people that are hospitalized in Africa are worse than in other countries," she says. "So this could really lead to higher mortality, even among younger people."

SCIENTISTS ARE just beginning to probe what makes Delta so dangerous. They're concentrating on a suite of nine mutations in the gene encoding spike, the protein that studs the virus' surface and allows it to invade human cells. One important mutation, called P681R, changes an amino acid at a spot directly beside the furin cleavage site, where a human enzyme cuts the protein, a key step enabling the virus to invade human cells. In the Alpha variant, a mutation at that site made cleavage more efficient; a preprint published in late May showed Delta's different change makes furin cleavage even easier. The researchers suggest this could make the virus more transmissible.

Japanese researchers who made pseudoviruses carrying the mutation have not found it to confer increased infectivity in the lab, however, and in India, other coronavirus variants that include the same mutation have been far less successful than Delta, says evolutionary virologist Andrew Rambaut of the University of Edinburgh. "So it must be an interaction with something else in the genome."

Other mutations in Delta could help it thwart immunity. Some alter the spike's N-

terminal domain (NTD), which protrudes from the protein's surface. A recent *Cell* paper identified one spot in the NTD as a "supersite," unfailingly targeted by "ultra-potent" neutralizing antibodies from recovered patients. Delta's unique mutations delete the amino acids at positions 156 and 157 in the supersite and changes the 158th amino acid from arginine to glycine; the latter eliminates a direct contact point for antibody binding, says David Ostrov, a structural biologist at the University of Florida. "We think the 157/158 mutation is one of the hallmark mutations in Delta that has given it this more immune-evasion phenotype," concurs Trevor Bedford, a computational biologist at the Fred Hutchinson Cancer Research Center.

Another mutation in the NTD supersite may also help rebuff antibodies. And scientists should start to examine the role of changes in other Delta variant proteins, says Nevan Krogan, a molecular biologist at the University of California, San Francisco. "There is so much we don't know about these variants on every level. We are so in the dark." Delta has several mutations in the nucleocapsid protein, for example, which has many jobs, "like a Swiss Army Knife protein," says virologist David Bauer of the Francis Crick Institute. The experiments to bring clarity will take months, however.

IN THE MEANTIME, scientists agree urgent action is needed to stop the spread of the new variant. "Worries about Delta should galvanize us to really ramp up vaccination efforts and surge vaccines to places where Delta is ticking up," says virologist Angela Rasmussen of the University of Saskatchewan. U.S. President Joe Biden on 18 June urged young people to get fully vaccinated to protect themselves from Delta. Countries with little access to vaccine need to resort again to interventions such as physical distancing and masks, Rasmussen says.

The goal is not just to save lives immediately, but also to give the virus less room to evolve further. Delta's success has shown scientists aren't able to identify dangerous new variants in time to stop them spreading, says Emma Hodcroft, a virologist at the University of Basel, despite an unprecedented global effort to track its evolution in real time (*Science*, 21 May, p. 773). And it would be dangerous to assume that SARS-CoV-2 can't do much better, Katzourakis says. "Anything that's happened at least twice in evolution is part of a pattern," he says. "I would be very unsurprised if we saw equivalent changes over the coming year or two." ∎

> "There is so much we don't know about these variants on every level."
>
> **Nevan Krogan**,
> University of California,
> San Francisco

For WHO leader, a 'feeling that we're failing'

Tedros Adhanom Ghebreyesus had expected more social responsibility in the face of a global crisis

By **Jon Cohen**, *in Geneva*

Few have spoken out as forcefully against the global disparity in COVID-19 vaccine distribution as Tedros Adhanom Ghebreyesus, director-general of the World Health Organization (WHO). Tedros, as he prefers to be called, has labeled the inequity "vaccine apartheid" and a "catastrophic moral failure" that has led to a "two-track pandemic." A global procurement scheme by WHO and other parties to supply vaccines to poorer countries, the COVID-19 Vaccines Global Access (COVAX) Facility, has not had much impact so far.

Tedros served as Ethiopia's minister of health from 2005 to 2012—and then as minister of foreign affairs—before taking the helm at WHO in 2017. *Science* spoke to him on 12 June, hours before he addressed a summit of G7 nations underway in the United Kingdom that resulted in a pledge to donate 870 million additional vaccine doses to COVAX by the end of 2022. The interview has been edited for brevity and clarity. A longer version is online at https://scim.ag/TedrosQA.

Q: The conversation about inequity in COVID-19 vaccination started before there was anything that worked. Were you trying to get ahead of it, by doing the advocacy early?
A: When I was Ethiopia's minister of health, I saw two global failures. One is HIV. Those who needed [antiretroviral treatment] in low-income countries didn't get it until 10 years after the discovery. That's very, very disappointing. Not only that, there was an H1N1 influenza pandemic in 2009. Vaccines were developed, but they didn't reach low-income countries, especially in Africa. These are the

---

Tedros Adhanom Ghebreyesus waited to get vaccinated in Geneva until it would have been his turn in Africa.

things I remembered when we were confronted with COVID. And from the start, we were saying: We shouldn't repeat the same mistake. It's unprecedented actually to have eight vaccines [authorized] in just a year. But where we're failing is the vaccine equity for distribution.

Q: You've had a voice with some anger and frustration, mixed with advocacy. Why do you think so few other world leaders are willing to be blunt?
A: I don't think I know the answer.

Q: Has anyone become very ill or even died from COVID-19 who was close to you?
A: Yes. The closest is the stepmother of my wife [in Ethiopia]. And this is before vaccination started, and even now she probably wouldn't have had access. She died. And there are others, a bit more at a distance.

Q: Operation Warp Speed wasn't designed to vaccinate the world. It was designed to deliver vaccines for the American people and protect the country. What if there had been a different administration that had a more global vision?
A: The problem of just focusing on one country is we will not use the full potential of the whole world. I would turn to the G7 first, and [then] I would take G20, which controls 80% of the [global economy]. They can influence the whole world. What we have been saying from the start—solidarity, everybody contributes to the basket. That [conversation] should happen, and that's what we have been

pushing for. This is not just the responsibility of the U.S., but of all the major countries that have capacity for financial [help] or even production capacity.

Q: What do you think of the Biden administration?
A: This person has changed everything, turned it upside down. It's day and night what happened. From day one, when Biden reversed the decision [by former President Donald Trump to defund] the WHO, he has shown his commitment. And the U.S. has made the largest contribution to donating vaccine doses. We appreciate the help and hope other leaders will follow suit and make a significant contribution. We cannot ask the U.S. to do it alone.

Q: You've been talking recently about the idea of a pandemic treaty. Would it include waivers of intellectual property (IP) for vaccines and other medicines?
A: The IP waiver is one option, and it is key, especially in emergency conditions. Even TRIPS [a World Trade Organization agreement] has a provision to use an IP waiver in emergency conditions. In unprecedented situations, you need to take unprecedented measures. If we're not going to use it now, then when do we use it? Why do we even have it in the first place? I'm a strong supporter of IP. I'm very grateful to the private sector. And I don't believe in just taking IP away from them. I actually proposed giving incentives, some compensations [in exchange for an IP waiver]. At the same time, the companies also have a social responsibility. This

is about the whole world. Who would like the world to be hostage of a virus? Companies can make profit on 99% of their medicines. This is 1% of things.

Q: You tried to launch an IP sharing program called the COVID-19 Technology Access Pool. Are you surprised that no vaccine companies participated?
A: It's a mixed feeling: surprised, not surprised. Social responsibility is just something you would expect from a decent human being when the whole world is burning.

Q: You've been heavily criticized for a report that came from WHO, with its imprimatur, saying the lab origin hypothesis of the pandemic virus is extremely unlikely.
A: That's where there is a misunderstanding. The group came from different institutions and different countries, and they're independent. There were only two WHO staff who joined them. They came up with their study. At the end of the day, we said: "OK, there is progress, but there are these challenges." And we said it openly. We treated China the same way we treat every country.

Q: When SARS-CoV-3 or whatever is the next pandemic virus shows up, and you're sitting where you're sitting now, what are you going to do differently on day one?
A: If you're asking for a magic bullet, there's no magic bullet. I appointed a very independent, high-level commission to assess the situation so far and make recommendations. They did a very impartial independent assessment, which we are really proud of. We have to calm down and chart a clear road map on how to respond to the next pandemic. But we have to be very, very serious in learning from this and do things that are as ambitious as possible, and that will be impactful.

Q: What was the date you got your first shot?
A: May 12.

Q: You're the head of WHO. You could have said in December 2020, "I'm ready."
A: I was protesting. I wanted to wait until Africa and other countries in other regions, low-income countries, started vaccination. I have a background as a health worker and I'm in one of the risk groups. They were beginning to vaccinate health workers and risk groups [in Africa] around that time, so I thought that was my turn.

Q: How did it feel once you got vaccinated?
A: I'm still feeling that we're failing. I was having my shot with disappointment. ∎

PSYCHIATRY

Botox depression treatment raises eyebrows

Researchers cast doubt on study claiming strong evidence for the unusual treatment

By **Cathleen O'Grady**

By using a bacterial neurotoxin to paralyze facial muscles, Botox treatments get rid of forehead wrinkle lines. They can also make it hard to frown. That has led some clinicians to the unusual idea that by eliminating negative emotional feedback that frowns feed the brain, Botox can relieve depression. Botox offers "another option for severely depressed people, especially those that can't tolerate medication," says Michelle Magid, a psychiatrist at the University of Texas, Austin, who has prescribed Botox to dozens in her own practice. In a March meta-analysis combining data from five trials, Magid and colleagues found what they say is top-quality evidence that should "pave the way" for Botox's widespread use in psychiatry.

But many researchers remain skeptical. Last month, Nicholas Coles, a social psychologist at Stanford University, published a response to the meta-analysis, saying it doesn't address problems that he and his colleagues flagged in a 2019 meta-analysis that used much of the same trial data—but came to an opposite conclusion about the value of the treatment. If doctors push ahead without a firm foundation of evidence, he says, "people could be harmed."

The notion that Botox could help depression rests on the facial feedback hypothesis: the idea that physical expression of emotion—like smiling or frowning—provides feedback to the brain that reinforces, or even sparks, an emotional experience. The idea can be traced back to the 1800s; in 1872, Charles Darwin wrote that "the free expression by outward signs of an emotion intensifies it."

Social psychology experiments in the 1970s found evidence that even fake smiles seemed to boost a person's mood. If the brain recognizes something emotional happening in your body, "the emotion behind it gets enriched," says Tillmann Krüger, a psychiatrist at Hannover Medical School who led the new meta-analysis, published in the *Journal of Psychiatric Research*. Blocking frowning with Botox "breaks the feedback loop," he says.

In the new meta-analysis, Krüger and colleagues combined data from five trials, conducted between 2012 and 2020, in which people suffering from depression had their forehead muscles injected with either Botox or saline. They were evaluated psychologically weeks later. The meta-analysis results were startling: The power of Botox to alleviate depression was more than twice as strong as the best approved oral antidepressants. Botox "has an excellent tolerability and safety profile, and the patients don't have to think about taking a pill every day," Krüger says.

Botox injections make frowning hard. Could that break a feedback loop?

But in their own 2019 analysis, which drew from four of the same five studies, missing only one recent published trial, Coles and his colleagues saw reasons for concern. Each trial had fewer than 100 patients, and participants might be able to tell whether they were given a placebo or the toxin because the effects of Botox are so obvious. The researchers also saw the huge effect size, but it was a cause for concern, not celebration. Coles thought it strange that Botox's effect was four times larger than in tests of the facial feedback hypothesis, in which people's emotional states are measured after they hold a facial expression for a few seconds or minutes. And Coles points out that four of the five trials were conducted by researchers who had declared payments from Botox's maker, Allergan.

The fact that the antidepressant effect is so much larger than in facial feedback experiments is not cause for concern, Krüger says. For one thing, he says, using Botox to prevent frowns day in, day out for weeks could have a more powerful influence than a few minutes of frowning or smiling in a lab. He and colleagues also suggest Botox could affect the nervous system directly, improve a patient's self-image, or change how the world responds to them.

Krüger argues that Botox should be used for depression even if it's unclear how it helps. He points out that the mood-altering mechanism behind the broadest class of antidepressants—which allow the neurotransmitter serotonin to stick around longer in the brain—is still hotly contested. Magid says not using Botox risks "withholding a valuable treatment from patients that might benefit from it."

But others share Coles's concerns. Large effect sizes seen in small studies tend to disappear in larger trials—a trend seen even for today's commonly used antidepressant medications, says Eiko Fried, a depression researcher at Leiden University. He worries psychiatrists who use Botox for depression are giving their patients false hope and neglecting established treatments instead.

In 2018, Allergan said it had postponed a phase 3 trial of Botox for depression. (Allergan did not respond to a request for comment.) Magid and Krüger, who have received compensation for advising Allergan on the treatment, said they do not know of any plans for a large, high-quality trial. "We're hoping that larger institutions, or somebody with a huge NIH [National Institutes of Health] grant, can take on something this big," Magid says.

The question that still needs to be answered, says Jonathan Kimmelman, a bioethicist at McGill University, is whether Botox works well enough on depression to justify the risk of its rare but worrying side effects. Botox injections into the face can occasionally produce difficulty breathing and swallowing. Kimmelman says, "Nothing is safe unless there's some reason to use it." ∎

WORKFORCE

At DOE, efforts to address climate and diversity dovetail

Proposed urban integrated field laboratories would reach out to minority communities

By **Adrian Cho**

Taking aim at two goals at once, the Department of Energy (DOE) wants to launch an initiative both to address the climate crisis and increase diversity in the U.S. scientific workforce. In its 2022 budget request to Congress, DOE requests funds to create urban integrated field laboratories (IFLs) that would gather climate data in cities and build bridges to urban communities, including by collaborating with minority-serving universities, such as historically Black colleges and universities (HBCUs).

"I was surprised but thrilled to see the IFL language," says Lucy Hutyra, a biogeochemist at Boston University. "Urban areas are radically understudied." David Padgett, a geoscientist at Tennessee State University, an HBCU in Nashville, says, "This sounds like something I might want to collaborate on with my colleagues at TSU or Spelman" College, an HBCU in Atlanta.

The effort is timely, scientists say, as evidence suggests the impacts of climate change will often fall hardest on poorer urban communities. But collecting climate data in cities poses major challenges, and Black researchers stress that to really boost diversity, DOE will have to help minority institutions grow their research capacity.

The IFL concept originated with DOE's biological and environmental research advisory committee, which in 2015 urged DOE to build such labs in cities especially sensitive to climate change: those in arid, mountainous, coastal, and agricultural environments. The DOE budget request adds a social dimension, noting that the labs will "incorporate environmental justice as a key tenet of research." That language "is unusual," says Bruce Hungate, an ecosystems scientist at Northern Arizona University, Flagstaff. "However, it absolutely reflects the direction the conversation is turning in environmental science. So, kudos to DOE."

Other federal agencies already gather climate-related data in urban areas. The National Science Foundation supports 28 Long-Term Ecological Research sites, including stations in Baltimore and Phoenix, with one to be added in Minneapolis. The Environmental Protection Agency and the National Oceanic and Atmospheric Administration sample air in various cities, and NOAA has long funded an atmospheric science lab at Howard University, an HBCU in Washington, D.C. (p. 1382). But such efforts have often focused on biology and local phenomena, scientists say.

In contrast, the IFLs would collect comprehensive data on the flows of energy, water, and airborne chemicals through complex urban environments and feed them into the department's global climate models, says Sharlene Weatherwax, DOE's associate director for biological and environmental research. Ultimately, she says,

Hotter summers could hit city dwellers especially hard.

researchers hope to predict how climate variables such as temperature and precipitation will change on spatial scales of kilometers and temporal scales from 2 to 100 years. "We really care what will happen to climate where people live, and a big swath of that is urban."

The agency plans to call for proposals and select sites for more than one IFL in 2022. Each would likely cost between $1 million and $10 million and involve a collaboration between some of DOE's 17 national laboratories and local universities. The agency aims to engage with urban institutions it may not have worked with before, Weatherwax says. In a separate effort, DOE wants to begin planning for a national climate laboratory or center, to be sited at an HBCU or other minority-serving institution.

Methods used to collect climate data in environments such as forests and grasslands may not work well in urban areas, warns Hank Loescher, a biogeochemist and director of strategic development at Battelle. Tracking water in cities can be tricky because much of it flows through sewers, he notes, and monitoring gases such as methane requires extremely tall towers that cities may not allow. Researchers also lack a comprehensive theory of the urban environment, Loescher says. "They're still figuring out what to measure and how to measure it."

DOE must also recognize that urban institutions serving minority communities often lack the staff and infrastructure of wealthier universities, says Beverly Wright, a sociologist at Dillard University (an HBCU) and executive director of the Deep South Center for Environmental Justice in New Orleans. "Part of the program should be a focus on building capacity at these universities."

If done poorly, IFLs could even perpetuate a "colonialist" relationship between majority and minority institutions, warns Everette Joseph, an atmospheric physicist at the National Center for Atmospheric Research. For example, he worries HBCUs might be relegated to a purely educational role in IFLs. "There'll be an emphasis on sending your students to the lab and no thinking about how you build capacity until it's exploitative," he says.

Scientists acknowledge that DOE is taking a political risk in linking climate research with environmental justice. DOE research enjoys bipartisan support in Congress, but that's because legislators view it as an engine of economic competitiveness. Weatherwax says she's not worried about a political backlash. "We're hoping that people will understand that [environmental justice] means: 'This could be you.'" ∎

Some researchers fear a future in which smart speakers play ads to unsuspecting sleepers.

NEUROSCIENCE

Advertisers could come for your dreams, researchers warn

Inserting ads into dreams may one day be feasible

By **Sofia Moutinho**

I f you've ever crammed for an exam before bedtime, you may have tried something dream researchers have been attempting for decades: coaxing knowledge into dreams. Such efforts have had glimmers of success in the lab. Now, brands from Xbox to Coors are teaming up with a few scientists to attempt something similar: "Engineer" advertisements into consumers' dreams, via video and audio clips. This month, 40 sleep and dream researchers have pushed back in an online letter, calling for the regulation of commercial dream manipulation.

"Dream incubation advertising is not some fun gimmick, but a slippery slope with real consequences," they write on the op-ed website EOS.

Dream incubation, in which images, sounds, or other sensory cues are used to shape nighttime visions, has a long history. Greeks who fell ill in the fourth century B.C.E., for example, would sleep on earthen beds in the temples of the god Asclepius, to prompt a state of dreaming in which their cure would be revealed.

Modern science has opened a whole new world of possibilities. Researchers can now identify sleep stages when most people dream by monitoring brain waves, eye movements, and even snoring. They have

also shown that external stimuli such as sounds, smells, and lights can alter dreams' content. This year, researchers communicated directly with lucid dreamers—people who are aware while they are dreaming—getting them to answer questions and solve math problems as they slept.

"People are particularly vulnerable [to suggestion] when asleep," says Adam Haar, a cognitive scientist and Ph.D. student at the Massachusetts Institute of Technology who co-authored the letter. Haar invented a glove that tracks sleep patterns and guides wearers to dream about specific subjects by playing audio cues during susceptible sleep stages. He has been contacted by three companies in the past 2 years, including Microsoft and two airlines, asking for his help on dream incubation projects. He helped with one game-related project, but wasn't comfortable participating in advertising campaigns.

Work by Harvard University dream researcher Deirdre Barrett has also attracted corporate attention. In a 1993 study, she asked 66 students to select an academic or personal problem, write it down, and think about it each night for a week before bed. At the end of the study, nearly half reported dreaming about the problem. Similar work published in 2000 in *Science*, in which Harvard neuroscientists asked people to play several hours of the computer game *Tetris*,

found that slightly more than 60% of the players reported dreaming about it.

This year, Barrett consulted with the Molson Coors Beverage Company on an online advertising campaign that ran during the Super Bowl. Following her instructions, Coors, which has mountains and waterfalls on its logo, asked 18 people (12 of them paid actors) to watch a 90-second video featuring images of mountains and Coors beer right before falling asleep. When the participants awoke, five reported dreaming about the beer, according to a YouTube video documenting the effort. (The result remains unpublished.)

Barrett says advertising strategies like these can get the public's attention, but will likely have little practical impact. "Of course you can play ads to someone as they are sleeping, but as far as having much effect, there is little evidence."

That doesn't mean that future attempts couldn't do better, says Antonio Zadra, a dream researcher at the University of Montreal who signed the statement. "We can see the waves forming a tsunami that will come, but most people are just sleeping on a beach unaware," he says.

The letter writers fear that because there are no specific regulations for in-dream advertising, companies might one day use smart speakers to detect people's sleep stages and play back sounds to influence their dreams and behaviors. "It is easy to envision a world in which smart speakers—40 million Americans currently have them in their bedrooms—become instruments of passive, unconscious overnight advertising, with or without our permission," says the letter, which the writers have sent to U.S. Senator Elizabeth Warren (D–MA).

Such a world is worth preparing for, says Dennis Hirsch, a professor of law and a privacy expert at Ohio State University, Columbus. But he adds that the U.S. Federal Trade Commission Act, which prohibits "unfair or deceptive" business acts, likely already applies to using smart speakers for in-dream advertising.

Tore Nielsen, a dream researcher at the University of Montreal who did not sign the statement, agrees that his colleagues have aired a "legitimate concern." But he says interventions like this won't work unless the dreamer is aware of the manipulation—and willing to participate. "I am not overly concerned, just as I am not concerned that people can be hypnotized against their will," he says. "[But] if it does indeed happen and no regulatory actions are taken to prevent it … whether or not our dreams can be modified would likely be the least of our worries." ∎

COVID-19

What went wrong with CureVac's mRNA vaccine?

Company blames variants of the coronavirus, but type of RNA may have been key

By **Jon Cohen**

The startlingly poor performance revealed last week for a COVID-19 vaccine made by the German company CureVac isn't just a disappointment, it's a scientific puzzle. The company blames the rapidly changing pandemic virus. But several outside researchers suspect the vaccine's design is at fault.

Many scientists and investors alike had expected CureVac's candidate, which uses messenger RNA (mRNA) to code for the spike surface protein of SARS-CoV-2, had a good chance of becoming one of the most powerful new weapons against the pandemic. It relies on essentially the same novel mRNA technology as vaccines from the Pfizer-BioNTech collaboration and Moderna, which demonstrated more than 90% efficacy in their trials, and it holds some practical transportation and storage advantages over those rival shots.

But preliminary data from a trial enrolling some 40,000 people, about 75% in Latin America and 25% in Europe, suggest the efficacy of the CureVac vaccine is a lackluster 47%—low enough that, if further data are equally disappointing, health regulators likely won't authorize it for emergency use. The topline finding came from an interim analysis evaluating 134 participants who developed at least one COVID-19 symptom. Although the company did not give a breakdown, the reported 47% efficacy translates to roughly 88 COVID-19 cases in the placebo group and 46 among the vaccinated. "The results are sobering," said Franz-Werner Haas, CureVac's CEO.

The vaccine's mRNA was designed for a version of spike that was dominant among the SARS-CoV-2 circulating early in the pandemic, but since then the virus has evolved into many variants. Only 1% of infected trial participants had a virus with the original spike protein; the others harbored a total of 12 different variants. "We are virtually fighting a different virus, different pandemic over the last 6 months," Haas said. "Demonstrating high efficacy in this unprecedented broad diversity of variants is quite challenging."

Other efficacy trials have found that certain mutant strains of the coronavirus can compromise the ability of COVID-19 vaccines to protect against mild disease, but

the variant that has most powerfully undermined other vaccines, Beta, was not seen in the CureVac study. In contrast, Alpha, first seen in the United Kingdom and one of the earliest variants of concern, caused 41% of the 124 cases overall and 91% of the 44 cases that occurred in Europe.

Kathleen Neuzil of the University of Maryland School of Medicine doubts variants fully explain the poor performance of CureVac's vaccine. Unlike CureVac's mRNA shot, she says, the Pfizer-BioNTech and Moderna vaccines "work very well against Alpha." She cautions that it's difficult to compare trials of different vaccines, but says, "It's just hard for me to believe that the variants could have this degree of effect."

CureVac did not provide any data about how many of the people infected in the trial developed severe disease. Other vaccines

> ### "It's certainly a good possibility that the vaccine is just not immunogenic enough."
> **John Moore**, Weill Cornell Medicine

continue to prevent most hospitalizations and deaths even when variants reduce their protection against mild COVID-19.

Some scientists trying to make sense of the CureVac result point to an earlier, phase 1 study of the vaccine. It showed that serum levels of so-called neutralizing antibodies, which prevent the virus from binding to cells, were relatively low in vaccine recipients compared with people naturally infected with the coronavirus. "It's certainly a good possibility that the vaccine is just not immunogenic enough," says immunologist John Moore of Weill Cornell Medicine.

The type of mRNA used by CureVac may undermine antibody formation, contends Drew Weissman of the University of Pennsylvania's Perelman School of Medicine, who helped pioneer certain mRNA modifications used in the Pfizer-BioNTech and Moderna vaccines. (Those companies license the technology, which may financially benefit the university and Weissman.) CureVac's vaccine used an unmodified form of mRNA. When natural

mRNA is injected into the body, it triggers the production of interferons, signaling molecules that can rev up the immune system. CureVac touted that as an advantage of its formulation. But Weissman notes interferons can also block the generation of T helper cells that, in turn, direct B cells to make antibodies.

The other companies' mRNA vaccines, in contrast, use chemically altered uracils, one of the four nucleotides that make up RNA. Weissman's group had shown in 2018 that uracil-modified mRNA triggered potent neutralizing antibodies and other protective immune responses in animal models. He notes that a BioNTech study comparing modified and natural mRNA vaccines also found that the modifications boosted the antibody response.

Peter Kremsner of University Hospital Tübingen, who helped run the CureVac efficacy study, suggests the company may have given people too low a dose of its vaccine. "I am uncertain what it was finally, natural uracil or only dose or both," he says.

CureVac's phase 1 study had compared the safety and immune responses generated by doses between 2 and 20 micrograms, but because of side effects at the highest doses, the company settled on a standard dose of 12 micrograms. (The Pfizer-BioNTech vaccine uses a 30-microgram dose and Moderna's is 100 micrograms.)

Other data presented by CureVac last week suggest, however, that the design of the vaccine is more important than the dose. It reported data from a monkey study that compared the current vaccine with a next-generation version, whose mRNA is more stable inside of cells: Even when the dose was the same, the new candidate produced higher levels of the spike protein, triggering a 10-fold higher level of neutralizing antibodies. Dose-ranging studies of the Pfizer and Moderna vaccines have also found that higher mRNA doses offer relatively modest gains in antibody levels.

Still, CureVac says it must wait for the final analysis of the current efficacy trial, expected to amass more than 200 COVID-19 cases, before it decides whether to make a "strategic shift" to the second-generation vaccine. "For now, we are going full speed exactly where we are," Haas said. "We are expecting the data to come within the next 3 weeks." ∎

A CHANGE IN THE AIR

Atmospheric science is overwhelmingly white.
These Black scientists ignited a change

By **Paul Voosen**

Everything depended on finding the gigantic dust cloud.

Seventeen years ago, a scientific crew set out on the *Ron Brown*, a gleaming U.S. research ship, in search of a plume of desert dust that had sheared off the Sahara and wafted over the Atlantic Ocean. Sailing from Barbados with day after day of clear sky, the crew was anxious about finding the dust storm. But that wasn't the only thing weighing on Vernon Morris, the atmospheric chemist who'd organized the cruise, the first of its kind run largely by Black and Hispanic scientists.

Failure could make life complicated for the atmospheric sciences program Morris had recently co-founded at Howard University. And, he says, "They would have never entrusted the ship to a bunch of Puerto Ricans and Blacks again."

Finally, on the fifth morning, the crew awoke to an eclipse of dust. It settled on railings, instruments, and weather bal-

At Howard University, Vernon Morris helped train more than half of recent U.S. Black atmospheric science Ph.D.s.

loons. "We were just all giddy, the happiest scientists," says Michelle Hawkins, a Howard graduate student at the time who now leads severe weather forecasts at the National Weather Service. The schedule was intense, as the crew launched weather balloons around the clock and took measurements in the ocean and at the surface to capture a picture of what turned out to be one of the largest Saharan plumes ever observed. And the cruise began a trend: The Aerosol and Ocean Science Expeditions (AEROSEs) have set out nearly every year since then, with the latest cruise defying the pandemic to sail in January.

Although AEROSE cruises are still full of black and brown faces, most research cruises—indeed, most of the geosciences—are not. Nearly 90% of geoscience doctorate holders are white, and by Morris's count, there are no more than seven Black atmospheric scientists in tenured positions in the country's major academic departments. Many of the communities most at risk from climate-related extremes, such as heat waves or hurricanes, are those least represented in the science that explores those threats. "It's not only a Black man's challenge," says Belay Demoz, an atmospheric physicist originally from Eritrea who is at the University of Maryland, Baltimore County. "It's the nation's problem."

Amid this dire portrait, however, the Howard University Graduate Program in Atmospheric Science has been a notable success. From 2006 to 2018, the Howard program, one of AEROSE's sponsors, produced 17 Black doctorate holders in atmospheric science—more than half the country's total—and 30% of its Latina doctorate holders. Seventy-five percent of graduates came from low-income backgrounds. Nearly all hold scientific jobs and some are now carving out prominent positions in the country's premier climate science agencies, especially the National Oceanic and Atmospheric Administration (NOAA). "I feel like 'successful' is not strong enough," Hawkins says. "It's been such a guiding light for all of us."

With the Biden administration promising to increase its support for historically Black colleges and universities (HBCUs) like Howard, including a potential new climate research lab based at one, this should be a time of hope and expansion for the Howard program. But its future is in doubt. For years, school leadership neglected the program, according to former faculty, and most of its founders, including Morris himself, have left. "I don't think Howard ever appreciated what they had on their hands," says Marshall Shepherd, a Black atmospheric scientist at the University of Georgia, Athens, and former president of the American Meteorological Society (AMS). "They had a game-changing program."

As the geosciences seek to improve their diversity, Howard's success holds lessons for leaders and institutions. How can it be replicated? And can it even be sustained at Howard itself?
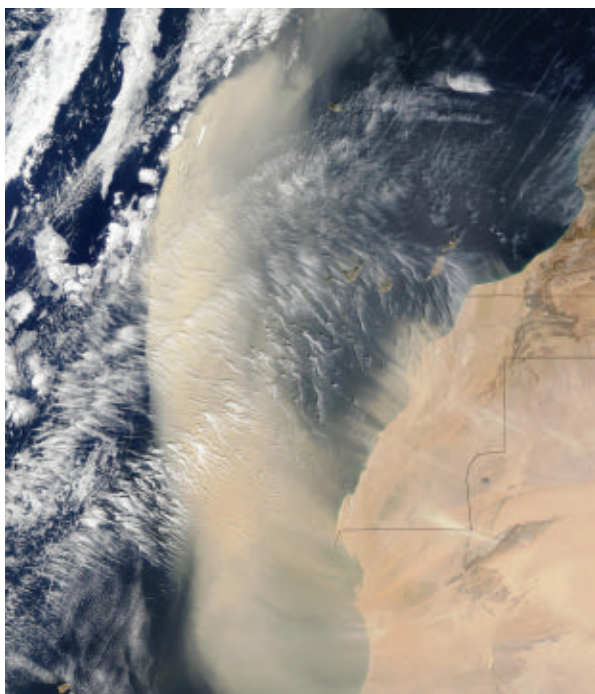
Vernon Morris (top, right) led students on research cruises in the Pacific and Atlantic oceans, where they charted Saharan dust storms (bottom).

MORRIS "BACKED" into college," as he puts it, attending Morehouse College, an HBCU in Atlanta, after deciding not to follow his father into the Air Force. But he felt adrift, far from his family in Washington, spending above his means to keep up with the rich kids. Thinking about dropping out, Morris ran into Henry McBay, a legendary Black chemist who offered to buy him books as long as he majored in chemistry and math. That led to a mentorship with John Hall, the pioneering Black atmospheric chemist who worked on ozone depletion. "I started doing work on chlorine nitrate chemistry and matrices," Morris says. "I loved it." In 1990 he became the first Black person to earn a Ph.D. from the Georgia Institute of Technology's earth science program.

As a postdoc, he became intrigued by cosmic chemistry and the atmosphere of Venus. He wanted to get back to an HBCU, and he noticed that Howard had recently launched a small, NASA-funded Center for the Study of Terrestrial and Extraterrestrial Atmospheres. Morris wrangled an invite and sold himself as someone who could help focus the center's research and ensure its continuation—at the time, external NASA reviewers were questioning its lack of relevant expertise and threatening to claw back the grant. He was hired as deputy director in 1994.

At the time, Morris was one of a few people of color who regularly appeared at national atmospheric science meetings. He and a handful of others would meet at a nearby hotel lobby bar to unwind. "Those meetings used to be superstressful, not seeing anyone who looks like you," says Gregory Jenkins, a climatologist who now works at Pennsylvania State University, University Park. "You wanted to give a paper, check out other papers, and go back to the hotel room."

The stress did not stop there. In an essay this year in *AGU Advances*, Morris wrote of being harassed by police. "I have been that figure sitting on the curb with my hands bound behind my back," he wrote. "I have been struck so hard with the butt of a police flashlight that the manufacturer label could be deciphered from the imprint on my chest days later."

Morris and his peers decided to create a community of Black atmospheric scientists at Howard by founding a new graduate program. Sonya Smith, an atmospheric engineer, came the year after Morris. Jenkins came, too, for a time. Everette Joseph, a code-savvy atmospheric scientist, joined soon after, as did Anthony Reynolds, a white physicist now at Embry-Riddle Aeronautical University. They started with a single course in atmospheric chemistry and gradually added more to build a case for a degree-granting program, ironing out their talking points over drinks and wings at a local bar. They won approval for the program from the Board of Trustees in 1997.

None of them realized what they were getting into. The board had allocated little funding, so they would have to support the program by winning external grants. "We had to worry about building a program and building our careers at the same time," says Joseph, who in 2019 became the first Black director of the National Center for

Atmospheric Research, one of the country's leading climate labs. "We were stupid and brash not to realize the monumental nature of the challenge."

They also needed to recruit students. Some they found at Howard, including DaNa Carlis, a first-generation college student who initially wanted to become a doctor. He ran into Morris, who pitched the program and offered him a scholarship. "I was like, 'What, you're going to pay me to go to school!'" says Carlis, who is now deputy director at NOAA's Global Systems Laboratory. Soon enough, he was doing research on ozone chemistry at NASA's Goddard Space Flight Center.

Casting a wider net, the Howard faculty began to lure chemistry, physics, and math undergraduates at five HBCUs to Howard for NASA-sponsored summer programs. They also scouted graduate programs where students often stopped at a master's degree. "We recruited those students hard," Morris says, by touting their connections to NASA and NOAA. They cobbled together money for a conference that attracted nearly 200 students of color to Washington, D.C., a gambit that gave them enough recruits to fill the program for 5 years.

In 2001, the Howard team scored a huge victory, winning a $2.5 million annual grant from NOAA to establish one of four new scientific centers at minority-serving institutions. The NOAA grant, which also supports research and students elsewhere, enabled Howard to bring on several new faculty, including Demoz, an atmospheric physicist. Demoz belonged to the majority ethnic group in his home country of Eritrea, and it wasn't until he worked at Goddard, where the only other Black scientist in his division was Shepherd, that he became truly aware of U.S. racial dynamics. Training minority scientists became a cause equal to his research. "I'm Eritrean and I have a guilty feeling that I didn't fight for independence," he says. "This is my independence."

Meanwhile, the first generation of Howard atmospheric science students was advancing toward doctorates. It was a tight group. One day, Joseph hauled them all into his office, ready to breathe fire because they had all gotten the same test question wrong in the same way. Then it became clear what happened: They had spent days together studying the problem in the wrong way. The tests required your best, says Isha M. Renta López, now a program analyst at NOAA's Office of Oceanic and Atmospheric Research. Once, Morris asked students to imagine a pyramid-shaped planet that doesn't rotate,

then describe the chemical reactions in its atmosphere. "He would put in the craziest stuff," López says.

The professors pushed the students to present their research at scientific meetings, paying their way if they couldn't get a scholarship. "I remember our very first meeting, Vernon giving us the talk," Hawkins says. "This is what you wear. This is how you introduce yourself. This is how you give a presentation." The students, in turn, helped recruit their own successors.

The program let the students be themselves and focus on their research without facing the tension and microaggressions that could come with being the sole Black student in a program, like passing comments about how "intelligent" or "well spoken" they were, or a critique that seemed excessively harsh for no reason. At Howard, Jenkins says, "They knew that it wasn't going to be easy, but one thing they didn't have to worry about were questions about whether they were qualified."

> "We can move the needle. We can provide equity where we can."
>
> **Michelle Hawkins,**
> National Weather Service

KNOWING FIELD RESEARCH would be critical to the students' success, the Howard faculty began to build opportunities. Many revolved around a small wooded campus owned by Howard in Beltsville, Maryland. Its haphazard assortment of squat red-brick buildings had previously hosted an observatory, animal experiments, and an argon beamline for particle physics research. Joseph and others saw huge potential in the site.

At the time, Joseph says, simulations of the lower atmosphere tended to rely heavily on data from simple environments, like "pristine homogeneous fields in Kansas." The Beltsville campus was right at the urban-rural divide, with countryside to the north and the sprawl of Washington, D.C., to the south, and the overlying air pushed and pulled from continental and ocean sources. Students took a lead role in launching weather balloons to study this complex environment. In 2006, they won a NASA grant to use their balloons to validate water vapor and ozone mea-

surements from its newly launched Aura satellite—work that might have traditionally gone to a NASA center. "I'm very proud of that," Demoz says. "Somebody took a risk and said let's do this."

Today, the Beltsville site is one of the few U.S. locations certified to make measurements of the upper atmosphere to the exacting standards of the World Meteorological Organization. It hosts a 30-meter tower for measuring winds, carbon dioxide, and air pollution; instruments for studying dust, soot, and similar particles; and sensors for measuring the precise amount of sunlight reaching the surface. Data gathered by Howard students help feed the numerical models used in National Weather Service forecasts.

Many students interned at NOAA and NASA branches around the country, and some opportunities took them even farther afield. In the late 2000s, Jenkins flew students to West Africa for NASA-sponsored research on the birth of Atlantic hurricanes and their interaction with Saharan dust. His students flew on a DC-8 into tropical storms, running the radar instruments to capture these rare data. Follow-up flights chased dust storms across the Atlantic and studied ozone variability in the tropics. "The most important thing was to send students way outside their comfort zone," Jenkins says.

Perhaps nothing was more uncomfortable for some Howard students than spending weeks at sea, as the AEROSE cruises have done now for nearly 2 decades. "Being a girl from the South Side of Chicago, that was really intimidating," Hawkins says. Mayra Oyola-Merced, now an atmospheric physicist at NASA's Jet Propulsion Laboratory, remembers when they got a new lidar scanner on board for tracking dust in the air. Joseph asked whether she wanted to run it. "This was a $250,000 instrument," she says, and she was just an undergraduate. But Joseph patiently explained the physics behind it, and that stuck with her. "Feeling someone trusted me and someone could see my potential to do things—it changed my complete perspective about going to graduate school."

FOR ALL ITS SUCCESS, the Howard program's future remains cloudy. For decades it was just one failed grant application away from collapse. The strain on faculty prompted many to leave, their vacant positions often going unfilled—an exodus that has many former faculty and alumni fearing for the program. Morris left early last year when Arizona State University offered him a chance to lead the school of science at its small liberal arts col-

lege. Jenkins, Joseph, and Demoz have also left, though Demoz still helps at Beltsville. "I'm the last one left," Smith says. She and Terri Adams, a Howard sociologist who led the NOAA center on an interim basis, have been tasked with rebuilding the program, even though Smith's work has veered toward hypersonic weapons and Adams studies emergency management of climate disasters, not geoscience itself.

For the university, Morris's departure served as a wake-up call. The university is now creating a department of earth, environment, and equity based on the program, giving it a lasting home that will include undergraduate students. (Previously the program was jointly administered by three departments, diluting its clout with the administration.) Charles Ichoku, a former senior NASA scientist originally from Nigeria, joined the faculty just before the pandemic and is pursuing grants for upgrades at Beltsville. Howard has hired a new director for its NOAA center and three new faculty in atmospheric sciences, two junior and one senior.

Atmospheric sciences is essential to the university's future, says Bruce Jones, who joined Howard in 2018 as its first ever vice president for research and is charged with increasing its research ambitions. "While we acknowledge we lost some of the great minds in the field, the recent investment is just the beginning of the rebirth of the program," he says. The recommitment is good to see, says Jack Kaye, associate director of NASA's earth science division, who is white. "Howard is special place in American education," he says. "Having a program focused on atmospheric sciences is important to the community. It's important to the nation."

Whatever the future of the program, its mark will be lasting. Morris, his colleagues, and former students have a tradition of getting together at the AMS annual meeting. Their once-informal gathering—online during the pandemic but scheduled to resume next year—is now one of the meeting's primary social events, attended by a diverse crowd that includes agency leaders and senior scientists. Called the "Colour of Weather," it regularly fills a ballroom.

Many former Howard students now occupy leadership positions, especially at NOAA, which remains the primary sponsor of the Howard graduate program. "They're smart, pushing, building communities— that's a pay back, there's no dollar sign on that," Jenkins says.

And they continue to try to make atmospheric science more welcoming. In 2014, Hawkins and Carlis started a diversity working group at NOAA, which hosts monthly seminars and encourages minority scientists scattered across the agency's many divisions to connect and support each other. They have also advised the agency to implement steps like evaluating supervisors on their support for inclusion and increasing the diversity of hiring panels. "We can move the needle," Hawkins says. "We can provide equity where we can."

At a research station in Beltsville, Maryland, Everette Joseph (left) taught students to gather data with weather balloons.

One disappointment, though, is that few Howard graduates have ended up in academia. Universities hire faculty from the same few elite academic programs, which continue to churn out predominately white graduates—and when few faculty are people of color, potential students are put off, Marshall says. "Students need to see professors of color," he says. "We just don't see enough people at the university level, teaching," Jenkins adds. "That's still the struggle."

In the wake of the murder of George Floyd in May 2020, Morris and several peers published a letter online urging geoscientists to end their silence on the racism in their field. "In place of police brutality, careers are killed through forced attrition and under-investment," they wrote. New Black professors are subject to biased reviews and limited mentorship. "The time has passed for predominantly white scientific organizations to post photos of happy, multiracial groups on their websites in place of actually diversifying their leadership and members," they wrote.

Since that letter, professional organizations have taken some real steps, Morris says. At the AMS meeting in January, the presidential forum focused on building a culture of antiracism. (Morris was among the speakers.) AMS and the American Geophysical Union have created new senior positions focused on improving diversity, featured talks on diversity more prominently at meetings, and gathered more data about the race and ethnicity of members.

But academic programs have been much slower to react, Morris adds. Overall, he sees little evidence that the toxic environment often faced by students and faculty of color has changed. New hires still face an evaluation and tenure process that reinforces white privilege, he says. "Many of the current cluster hires are like experiments that introduce new fish into a polluted pond to see if they survive," he says.

So much more can be done, Morris says. Investing more federal research dollars in HBCUs and minority-serving institutions is an obvious start. "A lot of HBCU programs are funded at levels that aren't going to enable their success," Morris says.

But it's not just money that made the Howard program so successful. It had federal agencies invested in its success and benefited from their material support in the form of adjunct professors and lab space. It recruited students from places other programs hadn't looked. And most of all, it had a cohort of young Black professors hungry for the program to succeed—who saw its success as their own.

Jamese Sims, an alumnus who's now NOAA's senior science adviser for artificial intelligence, recalls that passion. Howard's faculty asked how they could speak up for those without a voice, she says. They asked how their scientific expertise could make this a better country for all. "Because, if our goal is ultimately to save lives and property," she says, "that includes everyone." ∎

DIVERSITY

Inclusion in citizen science: The conundrum of rebranding

Does replacing the term "citizen science" do more harm than good?

By **Caren B. Cooper, Chris L. Hawn, Lincoln R. Larson, Julia K. Parrish, Gillian Bowser, Darlene Cavalier, Robert R. Dunn, Mordechai (Muki) Haklay, Kaberi Kar Gupta, Na'Taki Osborne Jelks, Valerie A. Johnson, Madhusudan Katti, Zakiya Leggett, Omega R. Wilson, Sacoby Wilson**

As the scientific community, like society more broadly, reckons with long-standing challenges around accessibility, justice, equity, diversity, and inclusion, we would be wise to pay attention to issues and lessons emerging in debates around citizen science. When practitioners first placed the modifier "citizen" on science, they intended to signify an inclusive variant within the scientific enter-

The list of author affiliations is available in the supplementary materials. Email: cbcoope3@ncsu.edu

prise that enables those without formal scientific credentials to engage in authoritative knowledge production (*1*). Given that participants are overwhelmingly white adults, above median income, with a college degree (*2*, *3*), it is clear that citizen science is typically not truly an egalitarian variant of science, open and available to all members of society, particularly those underrepresented in the scientific enterprise. Some question whether the term "citizen" itself is a barrier to inclusion, with many organizations rebranding their programs as "community science." But this co-opts a term that has long referred to distinct, grassroots practices of those underserved by science and is thus not synonymous with citizen science. Swapping the terms is not a benign action. Our goal is not to defend the term citizen science, nor provide a singular name for the field. Rather, we aim to explore what the field, and the

multiple publics it serves, might gain or lose by replacing the term citizen science and the potential repercussions of adopting alternative terminology (including whether a simple name change alone would do much to improve inclusion).

A more fruitful way forward, rather than focusing on name changes, is to focus on approaches that increase inclusion—that is, to enable all people to feel that the identity they hold belongs and authentically influences the culture, values, and future of the field. To lend weight to those approaches, we recommend increases in funding for community science and the subset of citizen science and science more generally that address the interests, concerns, and needs of members of society historically and currently underserved by science.

CO-OPTING LANGUAGE

The term citizen science has come to have two intertwined meanings. The original, narrower definition, coined in the mid-1990s, refers to projects led by institutions guiding decentralized data collection by volunteers often unknown to each other yet sharing the common goal of advancing scientific research (*1*). These projects number in the thousands, and even a single project can engage millions of people (*4*). The second definition arose later as a kind of "big tent" concept to refer to highly varied projects across many disciplines

Brittany Carson prepares sound recording equipment to distribute to volunteers in Sound Around Town, a citizen science project based in Raleigh, NC, that measures acoustic environments in residential settings to improve maps of noise pollution.

with public-inclusive approaches—regardless of the leadership, size, or design—and balancing multiple goals: science, engagement, education, policy, and/or empowerment (*1*). We begin with the narrower meaning.

There are dozens of terms used to describe participants in citizen science, including phrases in different languages as well as terms within English that hold different meanings in different cultures. Terms that might offend in one culture (such as "amateur") may be perfectly suitable to others, underscoring terminology challenges (*5*). Much of the debate about the use of the term citizen science has been in the United States. People born in the United States to currently or historically oppressed groups (such as by race, ethnicity, religion, gender, or sexual orientation) could perceive the term citizen as a source of power inasmuch as all these groups have struggled to obtain the rights of democratic citizenship. Although the term citizen also refers to people who reside in a place or are citizens of the world (*6*), many people contest the term because they perceive it to exclude, or even convey hostility toward, those without citizenship status within a given nation (*7*). Consequently, an increasing number of organizations in the United States, such as the National Audubon Society and others (*8*), have adopted the term community science to rebrand their citizen science programs as open to all publics. Other institutions have selected alternative terms such as "civic science" (by the American Association for the Advancement of Science, the publisher of *Science*) and "neighborhood science" (Los Angeles Public Libraries). In our personal experience, we have seen those in the sphere of public engagement in science call on others to use the term "community science" to describe citizen science activities.

In the United States, the urgent social pressures to relabel citizen science as community science pose a conundrum. Those using the term community science to replace the term citizen science hope to engage a wider range of demographic groups. However, the unintended impacts could be counter to inclusion. Although most science, including citizen science, aims to produce new knowledge, the term community science describes a very specific, formalized, and long-standing research paradigm. Distinct from that of citizen science, community science is linked to social action with aims including protection of human rights and measurable improvements for communities who face environmental

injustices and public health challenges (*9*). Community science includes community-based participatory research (CBPR), community-engaged research, community-owned and managed research (COMR), street science, and other participatory methods aimed to bring social change, with roots in the critical pedagogy of Paulo Freire and the social psychology of Kurt Lewin. Community science elevates local experts and place-based issues above academic experts and publication-driven research agendas (*10*).

The circumstances in which community science occurs are varied with regard to social context and topic. Community science may arise, irrespective of race or income levels, when groups need scientific evidence that is not part of typical scientific agendas. For example, the Silent Spring Institute, formed by community members on Cape Cod, Massachusetts, uses CBPR to prioritize cancer-prevention research on environmental causes of cancer to complement government and industry focus on cures for cancer. In other cases, community-based organizations, such as the West End Revitalization Association in central North Carolina, use COMR to address inequities in environmental protections and basic amenities within historically marginalized communities (*11*). The COMR principles set expectations between formal institutions and community-based organizations to achieve funding equity, management parity, and science to support enforcement of regulatory compliance and other legal venues to protect human rights (*10*). What unites such diverse projects is that the authority, power, and funding rests with communities (*12*). In this way, community science represents a fundamental departure from institution-based science, including citizen science.

The basis of citizen science, in strong contrast to that of community science, is typically volunteerism within the realm of mainstream science, in which funds flow to academic, government agency, or nongovernmental organizations; credentialed individuals at those institutions make decisions, partially or wholly, about research directions; and projects can be geographically large, vastly exceeding the community scale. Relabeling citizen science as community science without consideration of these fundamental and structural differences may actually impede social justice efforts being carried out in the context of existing community science projects. We believe that switching the words citizen and community without regard to the traditions and norms associated with these well-established and quite different approaches to science is at least misleading and disingenuous and at most directly harmful because larger citizen science organizations could dilute the goals

of, and potentially siphon donor funds away from, authentic community-driven efforts. Because community science is already underfunded, a clear distinction in terminology is necessary for establishing sources of support for authentic community-driven efforts.

The term community science should be reserved for projects that focus on local priorities and local perspectives and are able to maintain the locus of power in the community. A hallmark of individuals and organizations behind these efforts has been commitment to social action and antiracist, decolonizing research praxis aimed at elevating multiple ways of knowing, engendering trust, and sharing power (*9*). A name change alone for citizen science, not accompanied by altering underlying practices so that projects bring about structural change (*12*), is akin to false marketing.

CITIZEN SCIENCE IN POLICY

Adding complexity to the conundrum, the term citizen science has a second meaning: a "big tent" encompassing the blurry continuum from the narrower meaning of the term through to community science and beyond. The broader meaning also includes other forms of public engagement and aspects of both formal and informal education.

In the race to rebrand, a cost of abandoning citizen science as the name of the big tent is the loss of tremendous global momentum in professional practice and in policy that have unified support for highly diverse participatory practices under this well-recognized moniker. Since 2014, scholars and practitioners around the globe who are focused on public engagement in science have formed professional organizations using this term (*13*). Citizen science is also included in a range of laws and regulations in different countries (*14*). In the United States, the reauthorization of America Creating Opportunities to Meaningfully Promote Excellence in Technology, Education, and Science Act in 2016, which included the Crowdsourcing and Citizen Science Act of 2015, codified and defined citizen science in federal law. The Act authorizes the federal government to carry out a wide variety of scientific activities with the inclusion of people irrespective of professional scientific credentials and irrespective of citizenship status. As a big tent, citizen science provides legal protections for community science.

INCLUSIVE CITIZEN SCIENCE

The motivation of those who aim to change the name of citizen science is to make the field more inclusive. This is an extraordinarily important goal. To accomplish this, we argue that the most important change must focus on ways to actually broaden participa-

tion in, and enlarge the number of beneficiaries of, citizen science regardless of project nomenclature. We call for strategic planning to advance accessibility, justice, equity, diversity, and inclusion in citizen science, both narrow and broad definitions. Although the terminology debate is mostly a US phenomenon, these issues are universal.

Strategies may or may not require a shift in terminology, but decisions about what to name a research enterprise—or a movement—need to happen within a broader portfolio of strategies designed to advance inclusive practices. However, considerations of terminology should (i) avoid exporting from the global North limitations on meanings of words such as "citizen" into other areas of the world, (ii) be reflective to prevent harm from well-intentioned virtue signaling that can unintentionally undermine social justice efforts, and (iii) identify different terms for the big tent and the narrower field of institution-led projects. Discussions must occur beyond the narrow domain of scholars. A strategy should include perspectives from many sectors that have a stake in the outcomes, including government (such as tribal, federal, state, and local), nongovernmental and community-based organizations, academia, and corporations. It should also include those unaffiliated with institutions (written here with full recognition that our own author roster includes the voice of only one person not affiliated with an institution). It should include representations of Black, Indigenous, people of color, and other underrepresented groups as well as individuals living in countries in which they do not hold citizenship.

In recommending changes to enhance inclusivity, we are not suggesting that all projects should become community science (in name or in practice). We applaud that the content focus and decentralized design of institution-led citizen science has intentionally facilitated discovery science by expanding spatial and temporal scales of data collection through the engagement of hundreds to millions of participants. Field-based citizen science has allowed fine-grain, continental-scale documentation of shifts in species occurrence, abundance, and phenology and of precipitation, extreme weather, and earthquakes. Digital crowdsourcing projects have discovered new astronomical structures, organized genomic data, and solved puzzles of protein structure. Given that success of large-scale projects depends on reaching many people, enhanced inclusion could translate into massive broadening of participation.

CENTERING IN THE MARGINS
We suggest that citizen science projects will only become inclusive through action. Whether realigning existing projects and programs with inclusive practices or designing new projects, we recommend centering in the margins (15): If a project is accessible to the marginalized, it will be accessible to all. Although implementation will vary in its details, the broad approach is general and generalizable. For some projects, the best strategy may be to elevate culturally relevant perspectives (emphasizing diversity and inclusion). In others, the best strategy may be a focus on racial and economic disparities in environmental conditions (emphasizing justice and equity), aiming for sustained efforts to produce tangible outcomes beneficial to underserved groups. For institutions that house citizen science, attention to diverse representation in project leadership can assist in fostering accessibility, as will addressing structural barriers, such as economics (for example, costs of transportation and gear). Inclusion can be advanced by making a clear, honest linkage between project outcomes and the lives, livelihoods, values, and cultures of the participants. Prioritizing research funding to address the needs and interests of those historically and currently underserved by science will be a major step in providing the foundation for inclusive citizen science.

The impetus of practitioners to relabel citizen science as community science is evidence of their recognition that citizen science is not serving all people. We applaud this momentum and hope to refocus it on deeper work to create inclusive citizen science. In addition to project-specific actions, we advocate for research across the big tent on real and perceived barriers to citizen science volunteerism, including public perceptions of alternative terminology. We urge critical reflection to identify project design principles for citizen science that can answer relevant research questions and lead to positive social change. We recommend that evaluation of projects and assessment of outcomes include measures of participant diversity and encourage practitioners to publish participant demographics where feasible.

Citizen science has opened the doors and put out a welcome mat to create a bridge between science and society. Yet the result has been homophily; the overwhelming majority of participants in citizen science are similar in many respects to those overrepresented in the science professions. The challenges of inclusion in citizen science reveal that words—no matter what the terminology—and intentions—no matter how good—are not enough. Recognizing the distinct practices of community science and their necessity to those underrepresented in science highlights the reality that science has failed to serve all segments of society equitably, thus creating hollow invitations to participate. The choices of science agendas, what to study and how to affect that work, are not neutral and will serve some segments of society more than others. It is important, therefore, to be mindful of where the locus of power and decision-making lies within a given project and which sectors of society benefit the most.

We believe that new edges of scientific discovery and actionable science lie in inclusion: the addition, without assimilation, of diverse voices, values, perspectives, lived experiences, and identities. Because citizen science has multiple goals—research, education, policy, and empowerment—this braided path provides multiple inroads to inclusive practices. Citizen science will achieve its egalitarian aspirations when individual projects actively engage in inclusive praxis and the big tent collectively engages across diverse publics. As the boundaries of inclusive citizen science expand so that no segments of society remain underserved, so too will the face, and the foci, of science. ∎

REFERENCES AND NOTES
1. J. L. Shirk et al., Ecol. Soc. **17**, 29 (2012).
2. R. Pandya, K. A. Dibner, Eds., Learning Through Citizen Science: Enhancing Opportunities by Design (National Academies of Sciences, Engineering, and Medicine, 2018).
3. R. Pateman, A. Dyke, S. West, Citiz. Sci. **6**, 9 (2021).
4. J. Waldispühl, A. Szantner, R. Knight, S. Caisse, R. Pitchford, Nat. Biotechnol. **38**, 1124 (2020).
5. M. V. Eitzel et al., Citiz. Sci. **2**, 1 (2017).
6. R. Bonney, Bioscience **71**, 448 (2021).
7. C. S. Lowry, K. F. Stepenuck, Environ. Sci. Technol. **55**, 4194 (2021).
8. Other examples include the Great Smoky Mountains Institute at Tremont, Monarch Joint Venture, the Urban Ecology Center, the Natural History Museums of Los Angeles and San Diego, and the Xerces Society for Invertebrate Conservation.
9. S. Wilson, D. Campbell, L. Dalemarre, H. Fraser-Rahim, E. Williams, Int. J. Environ. Res. Public Health **11**, 12817 (2014).
10. C. Heaney et al., Progr. Commun. Health Partner. Res. Ed. Action **1**, 339 (2007).
11. www.niehs.nih.gov/research/supported/translational/community/west-end/index.cfm.
12. L. F. Davis, M. D. Ramírez-Andreotta, Environ. Health Perspect. **129**, 026001 (2021).
13. Examples of professional associations include the US-based Citizen Science Association, the European Citizen Science Association, the Australian Citizen Science Association, Citizen Science Asia, and the Iberoamerican Network of Participatory Science, the Brazilian Citizen Science Network.
14. Citizen science informed the Global Environmental Outlook 6 assessment of the United Nations (UN) Environment Programme and the UN Cities and Biodiversity Outlook. It is an integral part of the funding strategy of the EU Horizon Europe program, recognized within the Aarhus guidelines of the UN Economic Commission for Europe and for monitoring progress toward the UN's Sustainable Development Goals.
15. C. L. Ford, C. O. Airhihenbuwa, Am. J. Public Health **100** (Suppl 1), S30 (2010).

FERROELECTRICS

Two-dimensional ferroelectricity by design

A synthetic ferroelectric is made from a van der Waals assembly of boron nitride

By **Evgeny Y. Tsymbal**

The discovery of ferroelectricity marks its 100th anniversary this year (*1*), and this phenomenon continues to enrich our understanding of many fields of physics and material science, as well as creating subfields on its own. All of the ferroelectrics discovered have been limited to those exhibiting a polar space group of the bulk crystal that supports two or more topologically equivalent variants with different orientations of electric polarization. On pages 1458 and 1462 of this issue, Yasuda *et al*. (*2*) and Vizner Stern *et al*. (*3*), respectively, show that ferroelectricity can be engineered by artificially stacking a nonpolar in bulk, two-dimensional (2D) material, boron nitride (BN). A relatively weak van der Waals (vdW) coupling between the adjacent BN monolayers allows their parallel alignment in a metastable non-centrosymmetric coordination supporting 2D ferroelectricity with an out-of-plane electric polarization. These findings open opportunities to design 2D ferroelectrics out of parent nonpolar compounds.

Owing to the switchable electric polarization, ferroelectric materials are useful for various technological applications (*4*). A continuing quest for miniaturizing electronic devices inspires the search for appropriate materials exhibiting a reversible polarization at reduced dimensions. Although perovskite oxides have been widely used to explore ferroelectricity in thin-film structures and, in certain cases, showed promising results, they suffer from several limitations (*5*). Among them is the strong sensitivity of thin-film ferroelectricity to boundary conditions, largely resulting from the dangling bonds inherited from the perovskite structure, and oxygen stoichiometry, which is often not easy to control.

The rise of vDW materials (*6*) opened opportunities for materials science, including studies of collective phenomena such as 2D ferroelectricity (*7*). 2D vdW ferroelectrics are appealing because of their uniform atomic thickness, absence of dangling bonds, and the ability to be integrated with other vdW materials, which allows for useful functionalities. For example, in combination with high-mobility materials such as graphene, the 2D vdW ferroelectrics can be employed in ferroelectric field-effect transistors or can be used as ultrathin barriers in ferroelectric tunnel junctions (*8*).

Similar to perovskite oxides, however, the emergence of ferroelectricity in 2D vdW materials is limited by the requirement of the polar space group of the bulk crystal. In many-layered crystals, like hexagonal boron nitride (*h*-BN) and transition-metal dichalcogenides, the polarization is prohibited by a centrosymmetric vdW structure that has lower energy than other stacking configurations. Yasuda *et al*. and Vizner Stern *et al*. remedy this limitation by showing that an intrinsically centrosymmetric vdW material, *h*-BN, can be engineered on the atomic scale to become ferroelectric.

A bulk *h*-BN crystal has a layered structure where each BN monolayer displays a honeycomb coordination, similar to graphene, but with two interleaving sublattices composed of boron (B) and nitrogen (N). Bulk *h*-BN exhibits

Polarization in hexagonal boron nitride (*h*-BN)

Nonpolar BN

AA'

AA' layer stacking, the bulk form of *h*-BN. Layers are stacked at a 180° angle with respect to each other, making *h*-BN nonpolar.

AA

AA layer stacking has layers related by mirror symmetry, but this configuration is not stable.

Polar BN

AB

AB or BA stacking is more stable than AA. The vertical alignment of the B and N atoms distorts the N-2p_z orbital.

BA

The AB or BA stacking creates an out-of-plane electric dipole pointing down (−P) or up (+P).

Twisted bilayer BN

Adding a small-angle (Y) twist between *h*-BN sheets creates relatively large AA regions separated by AB and BA domains.

Atomic relaxation of the twisted bilayer creates relatively large AB and BA domains separated by domain walls and AA regions.

Department of Physics and Astronomy and Nebraska Center for Materials and Nanoscience, University of Nebraska, Lincoln, NE 68588, USA. Email: tsymbal@unl.edu

AA′ stacking (see the figure), where the top BN monolayer is 180° rotated with respect to the bottom so that the B (N) atoms are positioned atop the N (B) atoms, which makes the *h*-BN crystal centrosymmetric. A different BN coordination involves AA stacking, where two BN monolayers are stacked mirror symmetrically without rotation. This stacking is, however, unstable with respect to a B-N bond length displacement along the bond, which forms either AB or BA stacking configuration with half of the atoms lying atop each other and the other half facing empty centers of the hexagons (*9*). The AB and BA structures have broken inversion symmetry and exhibit an out-of-plane polarization, which can be qualitatively understood by the distortion of the $2p_z$ orbital of the N atoms sitting atop B atoms (*10*). The AB and BA stackings are related by mirror reflection and hence have opposite polarizations.

To realize AB (BA) stacked BN bilayers, an exfoliated BN monolayer was ripped off or few-monolayer *h*-BN flake were separated into two pieces, and then stamped them on top of each other with a precisely controlled twist angle. A zero-twist angle supports the AB (BA) stacking, whereas a small finite twist imposes interlayer translations forming a moiré pattern that consists of the three nearly commensurate stacking configurations AB, BA, and AA. Structural relaxation reconstructs this moiré pattern into large triangular domains of commensurate AB and BA stackings separated by abrupt incommensurate domain walls and topological AA defects that accommodate the global twist.

The staggered polarization of the twisted BN was demonstrated by using piezoelectric force microscopy (PFM) and Kelvin-probe force microscopy (KPFM). Both scanning probe methods revealed an array of triangular areas extending over several square micrometers and exhibiting an alternating contrast either in the piezoelectric response or in the electrostatic potential, thus indicating an opposite out-of-plane polarization of the AB and BA domains. The triangular contrast disappeared in a BN monolayer and AA′ stacked regions, confirming that the polarization is driven by the AB (BA) interlayer stacking.

Notably, the authors showed that the spontaneous polarization can be switched by applied bias voltage. Yasuda *et al.* probed the polarization switching by measuring resistance of a graphene sheet deposited over a single AB (BA) domain to detect extra charge carriers induced by the BN-bilayer polarization. The forward and backward scans of the resistance versus gate voltage showed a pronounced hysteresis, indicating bistability driven by the polarization switching. In parallel, Vizner Stern *et al.* detected the polarization switching by scanning a biased tip above an individual AB (BA) domain. They observed redistribution of domain walls to orient the local polarization with the electric field under the biased tip. Scanning the same area with the tip of opposite polarity reversed the KPFM contrast, indicating that polarization switching was fully reversible.

The polarization switching was found to occur through the domain nucleation and growth mechanism reminiscent of that known to occur in perovskites oxides. However, whereas in perovskites the two polarization states are distinguished by the minute displacements of tightly bonded atoms, in the AB (BA) stacked BN, they differ by a double B-N bond length translation along a weakly coupled interface. As a result, the polarization reversal could be regarded as a lateral sliding of one BN monolayer with respect to the other occurring through domain-wall motion.

Yasuda *et al.* and Vizner Stern *et al.*, along with recent work by Woods *et al.* (*11*) and Zheng *et al.* (*12*), identify synthetic ferroelectricity as an emergent field of research. These findings provide new opportunities for the fundamental studies of 2D ferroelectricity, as well as paving the way for innovative device applications. The proposed engineering of 2D ferroelectrics can be extended beyond BN to other vdW materials, such as transition metal dichalcogenides. The synthetic 2D ferroelectrics can be combined with other vdW materials to functionalize their electronic, spintronic, and optical responses. Overall, there are no doubts that this emergent field is rich in opportunities, and exciting new developments are to come. ∎

REFERENCES AND NOTES

1. J. Valasek, *Phys. Rev.* **17**, 475 (1921).
2. K. Yasuda, X. Wang, K. Watanabe, T. Taniguchi, P. Jarillo-Herrero, *Science* **372**, 1458 (2021).
3. M. Vizner Stern *et al.*, *Science* **372**, 1462 (2021).
4. J. F. Scott, *Science* **315**, 954 (2007).
5. L. W. Martin, A. M. Rappe, *Nat. Rev. Mater.* **2**, 16087 (2016).
6. A. K. Geim, I. V. Grigorieva, *Nature* **499**, 419 (2013).
7. C. Cui, F. Xue, W.-J. Hu, L.-J. Li, *npj 2D Mater. Appl.* **2**, 18 (2018).
8. E. Y. Tsymbal, H. Kohlstedt, *Science* **313**, 181 (2006).
9. G. Constantinescu, A. Kuc, T. Heine, *Phys. Rev. Lett.* **111**, 036104 (2013).
10. L. Li, M. Wu, *ACS Nano* **11**, 6382 (2017).
11. C. R. Woods *et al.*, *Nat. Commun.* **12**, 347 (2021).
12. Z. Zheng *et al.*, *Nature* **588**, 71 (2020).

Repeat after Me(CP2)!

A motif of dinucleotide repeats in the genome may be associated with Rett syndrome

By **Jian Zhou**[1,2] *and* **Huda Zoghbi**[1,2,3,4]

Rett syndrome (RTT) is a devastating neurodevelopmental disease caused primarily by loss-of-function mutations in *methyl-CpG-binding protein 2* (*MECP2*) (*1*). MeCP2 is a DNA binding protein (*2*) that controls gene expression, but the precise molecular mechanism by which MeCP2 loss drives RTT pathology remains unclear, partially because a distinct DNA motif that specifies MeCP2-DNA interactions is lacking. On page 1411 of this issue, Ibrahim *et al.* (*3*) demonstrate that MeCP2 binds modified cytosine in cytosine-adenine (CA) dinucleotide repeats, providing a new signature DNA motif for MeCP2 binding. MeCP2 protects CA repeats from high nucleosome occupancy, raising questions about the consequence of this binding on maintaining chromatin structure in neurons.

MeCP2 was first characterized as binding to methylated cytosine residues in the context of cytosine-guanine (CG) dinucleotides. Many mutations in the methyl-CpG binding domain (MBD) of MeCP2 cause the most severe RTT phenotypes in patients and mouse models of the condition, indicating that DNA binding is essential to MeCP2 function (*4–6*). Later studies revealed that MeCP2 also binds methylated CH [mCH, where H is adenine (A), cytosine (C), or thymine (T)], hydroxymethylated CA (hmCA), and methylated or hydroxymethylated CAC (*7–9*). Unlike canonical transcription factors, the characterization of MeCP2 binding sites did not identify a signature motif. Because of this featureless binding pattern and high abundance of the protein—MeCP2 broadly coats the genome (*10*)—it has been challenging to associate MeCP2 DNA binding with specific gene expression changes. Identifying the DNA sequences and modifications that

[1]Jan and Dan Duncan Neurological Research Institute, Houston, TX, USA. [2]Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA. [3]Department of Neuroscience, Baylor College of Medicine, Houston, TX, USA. [4]Howard Hughes Medical Institute, Baylor College of Medicine, Houston, TX, USA. Email: hzoghbi@bcm.edu

specify MeCP2-DNA association is therefore vital to better understand its function and connection to disease pathogenesis.

CA repeats are repetitive microsatellite DNA elements distributed throughout the genome whose modifications may potentially constitute an epigenetic code with an as-yet-unknown function (11). In a search of CA repeat "readers," Ibrahim et al. unexpectedly identified MeCP2 as a binder in biochemical assays. The authors subsequently found MeCP2 to be the only MBD family member to bind CA repeats. Using genome-wide chromatin and DNA immunoprecipitation sequencing (ChIP-seq and DIP-seq), Ibrahim et al. further discovered that MeCP2 also

resolution imaging and electron microscopy reveal that changes in MeCP2 abundance substantially alters heterochromatin structure and compaction in neurons (12, 13). The study of Ibrahim et al. is consistent with MeCP2's role in chromatin organization because they demonstrated a new function of MeCP2 as a long-range nucleosome organizer. The absence of MeCP2 resulted in an intriguing dual effect in a chromatin context-dependent manner. Within lamina-associated domains (LADs), the nucleosome density is increased both within and around CA repeats. By contrast, outside the LADs, the nucleosome density is increased within CA repeats but decreased around them. LADs

MeCP2 loss, expression of the majority is down-regulated in LADs.

Ibrahim et al. further crystalized the MeCP2 MBD in contact with hydroxymethylated CA repeat DNA and identified R133 as the residue that recognizes the hydroxymethyl group through direct interaction. The RTT-causing mutation R133C severely impaired this interaction, suggesting that binding to these repeats may be relevant to RTT pathogenesis. Human mutations have been extremely helpful in pinpointing the contributions of MeCP2 protein domains to RTT pathogenesis. Mutations in the MBD that abolish mCG and mCH binding cause severe patient phenotypes similar to complete loss of the gene (null), whereas other RTT-causing mutations that disrupt the transcriptional repression domain, such as R306C and R273X (where X represents a truncating mutation), do not reproduce the severe null phenotype (6, 14). Similarly, R133C has a milder effect than other MBD mutations (6, 15), probably because this mutant still binds to methylated CA repeats, which may constitute ~80% of all modified CA repeats in neurons. The loss of MeCP2 binding to hydroxymethylated CA repeats could be one of the many factors that contribute to RTT pathogenesis.

It will be important to determine the true abundance and genomic distribution of methylated and hydroxymethylated CA repeats, as well as the proportion of these DNA elements bound by MeCP2 in neurons. Also, the molecular basis of nucleosome density alteration inside and outside LADs upon MeCP2 loss is unclear. And it will be interesting to determine the effects of MeCP2 binding to modified CA repeats on gene expression within or outside LADs in the brain. The discovery that MeCP2 binds a new CA repeat motif that may influence chromatin structure in specific LAD contexts provides opportunities to investigate the consequence of this binding on maintaining healthy brain function. ∎

A new binding motif for MeCP2

MeCP2 binds to methylated or hydroxymethylated cytosine within cytosine-adenine (CA) repeats in addition to its known binding sites. A Rett syndrome-causing *MECP2* mutation, R133C, disrupts the MeCP2 binding to hydroxymethylated CA repeats. This leads to more condensed nucleosomes both within and flanking the CA repeats within lamina-associated domains (LADs); however, outside LADs, the nucleosomes are more condensed only within CA repeats.
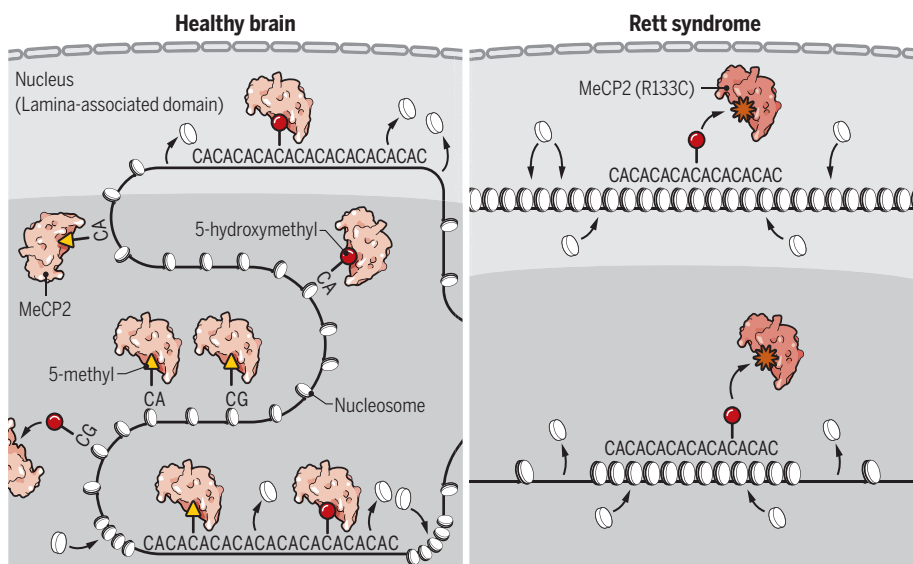
binds both methylated and hydroxymethylated CA repeats in fibroblast cells, with hydroxymethyl being the strongest target. This binding feature was overlooked for decades, probably because of their repetitive nature and because asymmetric modification of CA repeats require an unusually high sequencing coverage to be detected. By reanalyzing whole-genome bisulfite sequencing (WGBS) and ChIP-seq data in neuronal cells and in the mouse brain, the authors found strong MeCP2-enrichment on modified cytosines within CA repeats, implying a physiological relevance of this interaction in the brain.

Although MeCP2 has been described as a transcriptional repressor, growing evidence suggests that it also organizes three-dimensional (3D) chromatin architecture. High-

are chromatin domains associated with the nuclear lamina (area at the inner face of the nuclear membrane) that harbor hundreds of genes that are mostly transcriptionally inhibited (~5 to 10% are highly expressed). Therefore, the presence of MeCP2 results in more "open" chromatin that favors gene activation inside LADs and more "closed" chromatin around CA repeats, which tends to inhibit gene expression outside LADs (see the figure). This observation raises an interesting possibility that in addition to linear sequence features and modifications, the 3D subnuclear localization of MeCP2 binding sites also contributes to chromatin remodeling and gene expression changes. In line with this notion, Ibrahim et al. found that among the genes dysregulated in fibroblasts upon

REFERENCES AND NOTES

1. R. E. Amir et al., Nat. Genet. 23, 185 (1999).
2. J. D. Lewis et al., Cell 69, 905 (1992).
3. A. Ibrahim et al., Science 372, eabd5581 (2021).
4. D. Goffin et al., Nat. Neurosci. 15, 274 (2011).
5. L. D. Heckman et al., eLife 3, e02676 (2014).
6. K. Brown et al., Hum. Mol. Genet. 25, 558 (2016).
7. J. U. Guo et al., Nat. Neurosci. 17, 215 (2014).
8. S. Lagger et al., PLOS Genet. 13, e1006793 (2017).
9. M. Mellén, P. Ayata, N. Heintz, Proc. Natl. Acad. Sci. U.S.A. 114, E7812 (2017).
10. P. J. Skene et al., Mol. Cell 37, 457 (2010).
11. C. Papin et al., Genome Res. 27, 934 (2017).
12. M. W. Linhoff, S. K. Garg, G. Mandel, Cell 163, 246 (2015).
13. A. Ito-Ishida et al., J. Neurosci. 40, 8746 (2020).
14. S. A. Baker et al., Cell 152, 984 (2013).
15. H. Leonard et al., J. Med Genet. 40, e52 (2003).

CORONAVIRUS

Hybrid immunity

COVID-19 vaccine responses provide insights into how the immune system perceives threats

By **Shane Crotty[1,2]**

Immunity to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection is a vital issue for global society. Determining the quality and duration of that immunity is therefore key. But the adaptive immune system is complex, and these factors may differ between natural immunity (obtained by infection) and vaccine-generated immunity (1). Additionally, there is the question of the combination: What kind of immunity develops in people with natural immunity who are subsequently vaccinated? Such "hybrid immunity" is particularly interesting because of the notable finding that people with previous SARS-CoV-2 infection mount unusually potent immune responses to COVID-19 vaccines (2). This is exemplified in two studies in this issue on pages 1413 and 1418 by Stamatatos *et al.* (3) and Reynolds *et al.* (4), respectively, which also highlight natural and vaccine-induced immune responses to variants.

Immunological memory is the source of protective immunity. Natural immunity and vaccine-generated immunity to SARS-CoV-2 are two different paths to protection. The adaptive immune system consists of three major branches: B cells (the source of antibodies), CD4[+] T cells, and CD8[+] T cells. For natural immunity, immunological memory to SARS-CoV-2 has been observed for more than 8 months for CD4[+] T cells, CD8[+] T cells, memory B cells, and antibodies (5), with a relatively gradual decline that appears to partially stabilize within a year (5–7). Levels of immunity can be placed on a spectrum, and natural immunity against symptomatic infection (COVID-19) has been found to be between 93 and 100% over 7 to 8 months in large studies, including locations where the SARS-CoV-2 variant of concern (VOC) B.1.1.7 (alpha) was widespread.

Natural immunity against variants with changes that substantially reduce antibody recognition [e.g., B.1.351 (beta), P.1 (gamma), B.1.526 (iota), and B.1.617] is less clear; there is evidence of more reinfections with such variants (8). Neutralizing antibody activity against most VOCs is reduced for natural immunity and vaccine-generated immunity. That most VOCs have mutations engendering partial antibody escape is evidence of selection pressure to evade natural immunity. The biological relevance of the reductions in neutralizing antibody potency against variants is most clearly evident from vaccine clinical trials and observational studies. Among current COVID-19 vaccines in use, ChAdOx1 nCoV-19 (AstraZeneca) vaccine efficacy against symptomatic cases dropped from 75% to 11% against B.1.351 (9). By contrast, BNT162b2 (Pfizer/BioNTech) vaccine efficacy against symptomatic cases dropped from ~95% to 75% against B.1.351, and protection against severe disease remained at 97% (10). Initial reports suggest that both vaccines retain most of their efficacy against B.1.617.2 (delta).

What happens when previously infected individuals are vaccinated? The observations in several studies, including those by Stamatatos *et al.* and Reynolds *et al.*, are that an impressive synergy occurs—a "hybrid vigor immunity" resulting from a combination of natural immunity and vaccine-generated immunity (see the figure). When natural immunity to SARS-CoV-2 is combined with vaccine-generated immunity, a larger-than-expected immune response arises.

There appear to be both B cell and T cell components to hybrid immunity. An important question about antibody-mediated immunity against VOCs has been whether neutralizing antibody reductions are due to intrinsically low antigenicity of the VOCs. That is, is it intrinsically challenging for B cells to recognize the variants' mutated spike proteins? The answer is no. Studies of natural infection with B.1.351 showed that neutralizing antibody responses were robust against that variant and the ancestral strain (11). Moreover, neutralizing antibodies against B.1.351 after vaccination of individuals previously infected with non-B.1.351 SARS-CoV-2 were ~100 times higher than after infection alone and 25 times higher than after vaccination alone—even though neither the vaccine nor infection involved the B.1.351 spike. This enhanced neutralizing breadth was first reported by Stamatatos *et al.* and then confirmed by multiple groups (4, 12). Overall, the strength and breadth of the antibody responses after vaccination of previously SARS-CoV-2–infected persons was unanticipated.

Why does this pronounced neutralizing breadth occur? Memory B cells are a primary reason. They have two major functions: one is to produce identical antibodies upon reinfection with the same virus, and the other is to encode a library of antibody mutations, a stockpile of immunological variants. These diverse memory B cells, created in response to the original infection, appear to be pre-emptive guesses by the immune system as to what viral variants may emerge in the future. This brilliant evolutionary strategy is observed clearly for immunity to SARS-CoV-2: A substantial proportion of memory B cells encode antibodies that are capable of binding or neutralizing VOCs, and the quality of those memory B cells increases over time (7). Thus, the increase in variant-neutralizing antibodies after vaccination of previously SARS-CoV-2–infected persons reflects recall of diverse and high-quality memory B cells generated after the original infection (7, 12).

T cells are required for the generation of diverse memory B cells. The evolution of B cells in response to infection, or vaccination, is powered by immunological

Hybrid vigor immunity with COVID-19 vaccines

Hybrid vigor can occur when different plant lines are bred together and the hybrid is a much stronger plant. Something similar happens when natural immunity is combined with vaccine-generated immunity, resulting in 25 to 100 times higher antibody responses, driven by memory B cells and CD4[+] T cells and broader cross-protection from variants.

- Memory B cells
- Antibodies
- CD4[+] T cells
- CD8[+] T cells

Natural immunity   **Vaccine immunity**   **Hybrid immunity**

[1]Center for Infectious Disease and Vaccine Research, La Jolla Institute for Immunology (LJI), La Jolla, CA, USA. [2]Department of Medicine, Division of Infectious Diseases and Global Public Health, University of California, San Diego (UCSD), La Jolla, CA, USA. Email: shane@lji.org

GRAPHIC: N. DESAI/*SCIENCE*

microanatomical structures called germinal centers, which are T cell–dependent, instructed by T follicular helper (T_{FH}) CD4+ T cells. Thus, T cells and B cells work together to generate antibody breadth against variants. Additionally, T cells appear to be important at the recall stage. Memory B cells do not actively produce antibodies; they are quiescent cells that only synthesize antibodies upon reinfection or subsequent vaccination. Memory B cells are increased 5- to 10-fold in hybrid immunity compared with natural infection or vaccination alone (*3, 12*). Virus-specific CD4+ T cells and T_{FH} cells appear to be key drivers of the recall and expansion of those SARS-CoV-2 memory B cells and the impressive antibody titers observed (*13, 14*).

Antibodies are clearly involved in protection against SARS-CoV-2 reinfection, but evidence also points to contributions from T cells (*1*). T cell responses against SARS-CoV-2 in natural infection are quite broad (*1*), and most T cell epitopes are not mutated in VOCs, indicating that the contributions of T cells to protective immunity are likely to be retained (*4, 15*). Most of the COVID-19 vaccines in use consist of a single antigen, spike, whereas 25 different viral proteins are present in SARS-CoV-2. Thus, the epitope breadth of both the CD4+ and CD8+ T cell responses is more restricted in current COVID-19 vaccines than in natural infection (*1*), whereas hybrid immunity consists of both spike and non-spike T cell memory. Notably, the Pfizer/BioNTech and Moderna COVID-19 messenger RNA (mRNA) vaccines can substantially boost spike CD4+ T cell responses in previously infected persons after one immunization (*3, 4, 13, 14*). Differences in T cell responses after two doses of vaccine are more variable in those individuals (*3, 13*).

The immune system treats any new exposure—be it infection or vaccination—with a cost-benefit threat analysis for the magnitude of immunological memory to generate and maintain. There are resource-commitment decisions: more cells and more protein throughout the body, potentially for decades. Although all of the calculus involved in these immunological cost-benefit analyses is not understood, a long-standing rule of thumb is that repeated exposures are recognized as an increased threat. Hence the success of vaccine regimens split into two or three immunizations. Heightened response to repeated exposure is clearly at play in hybrid immunity, but it is not so simple, because the magnitude of the response to the second exposure (vaccination after infection) was much larger than after the second dose of vaccine in uninfected individuals. Additionally, the response to the second vaccine dose was minimal for previously infected persons,

indicating an immunity plateau that is not simple to predict. Moreover, previously infected people in some SARS-CoV-2 vaccine studies included both asymptomatic and symptomatic COVID-19 cases. Enhanced vaccine immune responses were observed in both groups, indicating that the magnitude of hybrid immunity is not directly proportional to previous COVID-19 severity.

Overall, hybrid immunity to SARS-CoV-2 appears to be impressively potent. The synergy is primarily observed for the antibody response more so than the T cell response after vaccination, although the enhanced antibody response depends on memory T cells. This discordance needs to be better understood. Will hybrid natural/vaccine-immunity approaches be a reproducible way to enhance immunity? The Shingrix vaccine to prevent shingles, which is given to people previously infected with the varicella zoster virus, is impressively effective (~97% efficacy), and elicits much higher antibody responses than viral infection alone. These principles also apply to combinations of vaccine modalities. It has long been observed that combining two different kinds of vaccines in a heterologous prime-boost regimen can elicit substantially stronger immune responses than either modality alone—depending on the order in which they are used and on which vaccine modalities are combined—for reasons that are not well understood. This may occur with combinations of COVID-19 vaccines, such as mRNA and adenoviral vectors, or mRNA and recombinant protein vaccines. These recent findings about SARS-CoV-2 immunology are pleasant surprises and can potentially be leveraged to generate better immunity to COVID-19 and other diseases. ■

REFERENCES AND NOTES

1. A. Sette, S. Crotty, *Cell* **184**, 861 (2021).
2. M. Frieman *et al.*, *EBioMedicine* **68**, 103401 (2021).
3. L. Stamatatos *et al.*, *Science* **372**, 1413 (2021).
4. C. J. Reynolds *et al.*, *Science* **372**, 1418 (2021).
5. J. M. Dan *et al.*, *Science* **371**, eabf4063 (2021).
6. K. W. Cohen *et al.*, medRxiv 10.1101/2021.04.19.21255739 (2021).
7. Z. Wang *et al.*, bioRxiv 10.1101/2021.05.07.443175 (2021).
8. N. R. Faria *et al.*, *Science* **372**, 815 (2021).
9. S. A. Madhi *et al.*, *N. Engl. J. Med.* **384**, 1885 (2021).
10. L. J. Abu-Raddad, H. Chemaitelly, A. A. Butt, *N. Engl. J. Med.* 10.1056/NEJMc2104974 (2021).
11. T. Moyo-Gwete *et al.*, *N. Engl. J. Med.* **384**, 2161 (2021).
12. R. R. Goel *et al.*, *Sci. Immunol.* **6**, eabi6950 (2021).
13. M. M. Painter *et al.*, bioRxiv 10.1101/2021.04.21.440862 (2021).
14. A. Tauzin *et al.*, *Cell Host Microbe* 10.1016/j.chom.2021.06.001 (2021).
15. A. Grifoni *et al.*, *Cell Host Microbe* 10.1016/j.chom.2021.05.010 (2021).

CELL BIOLOGY

The chains of stress recovery

Ubiquitination primes the cell for recovery from heat stress

By **Dorothee Dormann**

Cells often encounter stressful situations and respond to them with a stereotypical program to ensure survival. These responses involve increased expression of stress response factors, formation of stress granules (SGs), and shutting down of essential cellular processes, including RNA splicing, global translation, and nucleocytoplasmic transport. These adaptive changes are protective in the short term but need to be reversed once the stress has subsided so that cells can return to their normal activities. The molecular mechanisms involved in this reversal are only poorly understood. On pages 1409 and 1410 of this issue, Maxwell *et al.* (*1*) and Gwon *et al.* (*2*), respectively, reveal that recovery from heat stress requires tagging proteins with polyubiquitin chains. Together, these studies establish that ubiquitination has surprising regulatory and context-specific roles in the heat-stress response and emphasize that more attention should be paid to the stress recovery phase.

A global increase in polyubiquitin conjugation in response to stress has long been noticed, but this was mainly considered to induce the degradation of damaged or misfolded proteins that arise during cellular stress. Maxwell *et al.* examined which proteins become ubiquitinated upon different types of stress, including heat shock (42°C), arsenite treatment, osmotic stress, ultraviolet irradiation, and proteasome inhibition. They found many stress-specific changes, revealing that certain proteins became ubiquitinated exclusively in response to heat shock, whereas a different set of proteins was ubiquitinated upon arsenite exposure. This indicated that stress-specific patterns of ubiquitination represent distinct adaptive responses that could have important roles in the stress response.

Maxwell *et al.* dug deeper into the heat-stress response. Using deep quantitative proteomic analyses, they identified ~4900

Faculty of Biology, Johannes Gutenberg University Mainz, and Institute of Molecular Biology (IMB), Mainz, Germany. Email: ddormann@uni-mainz.de

proteins that showed an increase in ubiquitination following heat stress. They found that this "heat shock ubiquitinome" is highly enriched in proteins that function in processes that are down-regulated or shut down during cellular stress—e.g., translation, RNA splicing, nucleocytoplasmic transport, DNA damage repair, or the cell cycle. Among the heat shock ubiquitinome were also many proteins involved in the formation of SGs. These form when inhibition of translation leads to a rapid increase in untranslated cytoplasmic messenger RNA (mRNA). This triggers multicomponent liquid-liquid phase separation

the reversal of heat shock–induced nuclear transport impairments, and SG dissolution. Collectively, these findings establish an essential role of ubiquitination in the recovery of cellular activities following heat stress.

The authors furthermore noted that one of the proteins that became polyubiquitinated after heat stress is the central SG protein G3BP1 (Ras GTPase-activating protein–binding protein 1). The protein caught their attention because G3BP1 and its paralog G3BP2 are central nodes in the protein-RNA interaction network that leads to LLPS and SG assembly (3). Elimination of G3BP1

the dynamic disassembly of heat shock–induced SGs (see the figure).

Additionally, Gwon et al. demonstrated that upon heat shock, polyubiquitinated G3BP1 interacts with the protein segregase VCP (valosin-containing protein), which is coupled to both proteasome- and autophagy-dependent degradation. They found that this interaction occurs through a cofactor with a ubiquitin-binding domain, FAF2 (FAS-associated factor 2), which is an endoplasmic reticulum (ER)–associated protein. They observed recruitment of G3BP1-, VCP-, and FAF2-positive SGs to the ER upon heat shock, which suggests that heat shock–induced SG disassembly occurs on the ER membrane. This underpins the idea that the cytosolic surface of the ER is an important nexus for coordinating the heat-stress response, including the integrated stress response, the unfolded protein response, ER-associated degradation, and, as now discovered by Gwon et al., SG disassembly.

Ubiquitin controls stress granule disassembly after heat stress

Upon heat stress, RNA binding proteins (RBPs) and messenger RNAs (mRNAs) coalesce to form stress granules, and the central stress granule protein Ras GTPase-activating protein–binding protein 1 (G3BP1) is tagged with polyubiquitin chains. In the subsequent stress recovery phase, polyubiquitinated G3BP1 is degraded by the proteasome, involving valosin-containing protein (VCP) and an endoplasmic reticulum–associated adaptor protein, leading to disassembly of stress granules.

The studies of Maxwell et al. and Gwon et al. illustrate that not all stresses are created equal. This has been recognized in other contexts—for example, SGs have distinct morphologies and composition depending on the type of stress (5, 6), and posttranslational modifications are often introduced in a stress-specific manner (7). The study by Maxwell et al. underscores this theme by demonstrating that different stressors generate distinct ubiquitination patterns. This explains why an earlier study using arsenite stress found ubiquitination to be dispensable for SG dynamics (8), whereas others have found an important role of small ubiquitin-like modifier (SUMO)–primed ubiquitination in the disassembly of heat stress–induced SGs (9). Gwon et al. found that even the duration of stress can have profoundly different consequences: Short (30 min) heat stress induces the formation of SGs that are fully disassembled during stress recovery, whereas prolonged (90 min) heat stress leads to persistent SGs that are cleared by autophagic degradation. This finding explains apparent discrepancies in the literature on the role of autophagy in SG clearance (10, 11). It also emphasizes that researchers should pay attention to disease-relevant contexts, including the use of specific cell types and exact type of stress.

(LLPS) of mRNA with RNA binding proteins and gives rise to dynamic, reversible ribonucleoprotein (RNP) droplets in the cytoplasm. Upon removal of stress, SGs disassemble and translation resumes. Maxwell et al. found that numerous RNA binding proteins in SGs are polyubiquitinated after heat stress but not after arsenite stress. Thus, proteins involved in various stress-regulated pathways become specifically polyubiquitinated upon heat shock, suggesting that polyubiquitination could regulate these pathways.

Maxwell et al. suspected that ubiquitination could be particularly important in the stress recovery phase, because they observed that most of the ubiquitinated proteins were not degraded during heat stress but rather during subsequent recovery at 37°C. Using an inhibitor of the ubiquitin conjugation reaction, they demonstrated that ubiquitination is required for the reinitiation of translation,

interferes with SG assembly, and G3BP1 overexpression or optogenetic induction of G3BP1 dimerization is sufficient to trigger SG formation in the absence of stress (3, 4). Could ubiquitination of G3BP1 followed by its degradation underlie SG disassembly during recovery from heat shock?

Gwon et al. confirmed that G3BP1 undergoes K63-linked polyubiquitination after heat shock and is degraded by the proteasome during heat-stress recovery but not after other forms of stress. Using a mutational analysis, they were able to map the ubiquitin conjugation sites to the amino-terminal dimerization domain, which is crucial for RNA-dependent LLPS of G3BP1 (3). Cells expressing G3BP1 mutants that cannot be ubiquitinated showed severely delayed SG disassembly and featured less-dynamic SGs, which supports the idea that ubiquitination of G3BP1 is required for

Dynamic posttranslational modifications have recently been appreciated as key regulators of LLPS and RNP granules—e.g., they can alter biomolecular interactions and thus determine the threshold at which a protein phase separates (12). Gwon et al. identify an additional mechanism to control the dynamics of RNP droplets in cells. Polyubiquitin-mediated degradation of key constituents of cellular RNP granules could be a general

mechanism controlling the dynamics of such granules—e.g., some of the known stress-induced nuclear bodies. Moreover, Gwon *et al.* found that ubiquitination alters G3BP1 mobility within SGs, hence an interesting question to address in the future is whether polyubiquitin chains can directly affect the phase separation behavior of the modified proteins and thus alter RNP granule dynamics independent of proteasomal degradation. A role for polyubiquitin in protein phase separation and aggregation has been previously demonstrated (*13*, *14*), so it seems possible that stress-induced polyubiquitination regulates RNP droplets also through degradation-independent mechanisms.

Regulation of condensate dynamics is a particularly important topic in the context of neurodegenerative diseases, where impaired dynamics of RNP granules (e.g., SGs) are believed to promote aberrant LLPS and aggregation of disease-linked RNA binding proteins. Some of the aggregating proteins found in neurodegenerative disorders [e.g., TAR DNA binding protein of 43 kDa (TDP-43) and fused in sarcoma (FUS)] are highly polyubiquitinated in disease (*15*) and also were found by Maxwell *et al.* to become polyubiquitinated in the heat-stress response. This could prompt future studies on the role of ubiquitination of these disease-linked proteins and its relevance in the disease process.

The studies of Maxwell *et al.* and Gwon *et al.* reveal interesting paradigms that require further research. For example, it will be interesting to determine how ubiquitination regulates the reversal of other heat shock–regulated processes, such as translation, splicing, or nuclear transport. Another topic to address is whether other SG proteins that become ubiquitinated upon heat shock contribute to the regulation of SGs and, if so, through which mechanisms. Finally, the role of ubiquitination in response to other types of stress—e.g., oxidative or genotoxic stress—remains to be resolved. The heat is on to find out. ∎

REFERENCES AND NOTES

1. B. A. Maxwell *et al.*, *Science* **372**, eabc3593 (2021).
2. Y. Gwon *et al.*, *Science* **372**, eabf6548 (2021).
3. P. Yang *et al.*, *Cell* **181**, 325 (2020).
4. P. Zhang *et al.*, *eLife* **8**, e39578 (2019).
5. A. Aulas *et al.*, *J. Cell Sci.* **130**, 927 (2017).
6. S. Markmiller *et al.*, *Cell* **172**, 590 (2018).
7. S. Jungmichel *et al.*, *Mol. Cell* **52**, 272 (2013).
8. S. Markmiller *et al.*, *Cell Rep.* **27**, 1356 (2019).
9. J. Keiten-Schmitz *et al.*, *Mol. Cell* **79**, 54 (2020).
10. J. R. Buchan, R.-M. Kolaitis, J. P. Taylor, R. Parker, *Cell* **153**, 1461 (2013).
11. B. Wang *et al.*, *Mol. Cell* **74**, 742 (2019).
12. M. Hofweber, D. Dormann, *J. Biol. Chem.* **294**, 7137 (2019).
13. S. Yasuda *et al.*, *Nature* **578**, 296 (2020).
14. D. Morimoto *et al.*, *Nat. Commun.* **6**, 6116 (2015).
15. I. R. Mackenzie, R. Rademakers, M. Neumann, *Lancet Neurol.* **9**, 995 (2010).

PALEOANTHROPOLOGY

The complex landscape of recent human evolution

Archaic hominins in the Middle East underscore local demographic diversity in the last half million years

By **Marta Mirazón Lahr**

New genomic studies (*1*), new fossils (*2*, *3*), and new dates of existing ones (*4*) suggest that our African origin has a deeper history—one that took place in the context of high population and lineage diversity and which was intermeshed by periods of contact with Eurasian hominins using the Middle East as a geographical bridge. On pages 1424 and 1429 of this issue, Herskovitz *et al.* (*5*) and Zaidner *et al.* (*6*), respectively, report new archaic *Homo* fossils and stone tools in Nesher Ramla, Israel, that date to about 126 thousand years (ka) ago. This discovery, at the crossroads of Africa and Eurasia, adds substantial complexity to our reconstruction of those potential interactions, raising questions about the coexistence of different hominin populations in this region and complex population dynamics in the Late Pleistocene.

Perhaps the most insightful aspect of this emerging picture relates to the demography of the populations involved. The rich hominin genomic archive from Siberia reveals that Neanderthal populations experienced localized events of inbreeding suggestive of collapsed social networks as populations contracted (*7*), and repeated moments of expansion that repopulated areas they had previously inhabited (*8*). The poor preservation of ancient DNA in Africa precludes similar insights into our African demographic history. However, the recent discovery of modern human fossils in Greece and Israel dating to about 210 to 177 ka ago (*9*, *10*) and ancient European genomes show that there were multiple out-of-Africa dispersals in the last 400,000 years, during which early humans and Neanderthals interbred (*11*, *12*). Unlike what happened 60 ka ago (*13*), the offspring grew up within Neanderthal communities. The findings of Hershkovitz *et al.* and Zaidner *et al.* add a new level of complexity.

It has long been known that early modern human populations were in the Levant ~130 to 100 ka ago. The hominin fossils from Nesher Ramla now suggest that a different population, with anatomical features more archaic than those of both humans and Neanderthals, lived in this region at broadly the same time. Hershkovitz *et al.* argue that this represents the late survivorship of an archaic population of *Homo* that can be traced to earlier Levantine fossils (Tabun Cave, Zuttiyeh Cave, and Qesem Cave in Israel), whose taxonomy has been much disputed over the years. At a larger scale, this lineage shows affinities to Middle Pleistocene (789 to 130 ka ago) hominins and, more distantly, to Neanderthals, adding one more Eurasian group to those of Neanderthals and Denisovans.

The stone tools found with the new hominins belong to the group of prehistoric industries known as Mode 3, characterized by traits such as preparation of the core prior to flake removal and the production of points that could later be hafted onto shafts (see the illustration). These industries are patterned geographically. Western Eurasian Mode 3 industries, known as Middle

Flint tools (rendering shown) in the Nesher Ramla site indicate that the archaic hominin group shared their lithic technology with modern humans in the area.

Department of Archaeology, University of Cambridge, Cambridge, UK. Email: mbml1@cam.ac.uk

Paleolithic, are commonly associated with Neanderthals; those in Africa, known as the Middle Stone Age, are associated with both anatomically modern and nonmodern fossils. Within these broad categories, more local or temporal traditions exist, largely shaped by differences in the predominant way of manufacturing the stone tools and/or the type of stone tools made. The technological tradition at Nesher Ramla is one such variant, and, surprisingly, in contrast with the anatomical evidence, the lithics found with the archaic hominins at Nesher Ramla and the contemporaneous Levantine modern humans are the same.

The interpretation of the Nesher Ramla fossils and stone tools will meet with different reactions among paleoanthropologists. Notwithstanding, the age of the Nesher Ramla material, the mismatched morphological and archaeological affinities, and the location of the site at the crossroads of Africa and Eurasia make this a major discovery. But how does this discovery affect the hominin evolutionary landscape of the last half million years (see the figure)? Beyond the particulars of historical reconstructions, the finds add another piece to the puzzle of how late, cultural-bearing hominin species formed, survived, and disappeared. Hershkovitz *et al.* refer to the fossils as the "Nesher Ramla *Homo,*" avoiding formal taxonomic attribution. Their argument is that the taxon into which most researchers classify European and African Middle Pleistocene hominins, *Homo heidelbergensis,* is too variable and the Nesher Ramla mandible shows clear differences to its type specimen. One could also ask, to what extent is a species-based nomenclature useful when dealing with demographically and spatially dynamic populations that experienced some level of cultural and/or biological interaction? Answers will vary as much as species concepts vary, but the relatively rich resolution of recent hominin evolution offers a stage for testing, paleontologically and genomically, broader ideas in evolutionary biology.

From bonobos (*14*) to Caribbean pupfishes (*15*), genomics is changing our understanding of the process of speciation and species formation. Closely related species can differ substantially in the extent to which they experienced adaptive introgression—events of interbreeding with another species that generate new combinations that may result in phenotypes that overcome specific selective challenges. Given the complex and dynamic history of expansions, interactions, extinctions, and sometimes private histories in the genus *Homo,* similar differences should be expected in the processes that generated diversity. One of these differences will be in the outcome of intergroup contact. We are only too painfully aware of the disparate long-term consequences of contact to colonizers and colonized. It should, therefore, not be surprising to find that the cultural and genetic legacy of interpopulation contact in our deep past also differed according to social, economic, demographic, and technological circumstances. For disciplines that build historical narratives on the basis of shared similarities, this poses major challenges.

Demographic and spatial instability is the emerging picture. Paleontologically, the genus *Homo* is a palimpsest of populations, lineages, and species—*erectus, georgicus, antecessor, naledi, floresiensis, luzonensis, heidelbergensis, helmei,* Denisovans, *neanderthalensis, sapiens*; genomically, there is evidence of multiple expansions, contractions, and local extinctions. This fluid and population-specific demographic pattern is consistent with the nature of climate change in the last million years. This makes the difference between our recent selves and all the hominins that came before, including our earlier modern human ancestors, all the more salient. The new remains from Nesher Ramla add to the emerging complexity in the hominin evolutionary landscape of the last half million years. ∎

Hominin populations in the Middle East over the last half million years

A schematic representation of recent phylogenetic relationships shows the estimated time of divergence of African (red lines) and Eurasian (blue lines) lineages and the crucial role played by the Middle East (gray box), where hominins of African (red box) and Eurasian (teal box) lineages overlapped with a local population—"Nesher Ramla *Homo*" (purple box).

★ Nesher Ramla fossils and tools ----- Inferred relationships from morphological and genomic affinities
● *Homo sapiens* ● *H. sapiens* ancestral lineage ● Neanderthal ● Denisovan
● Neanderthal/Denisovan ancestral lineage
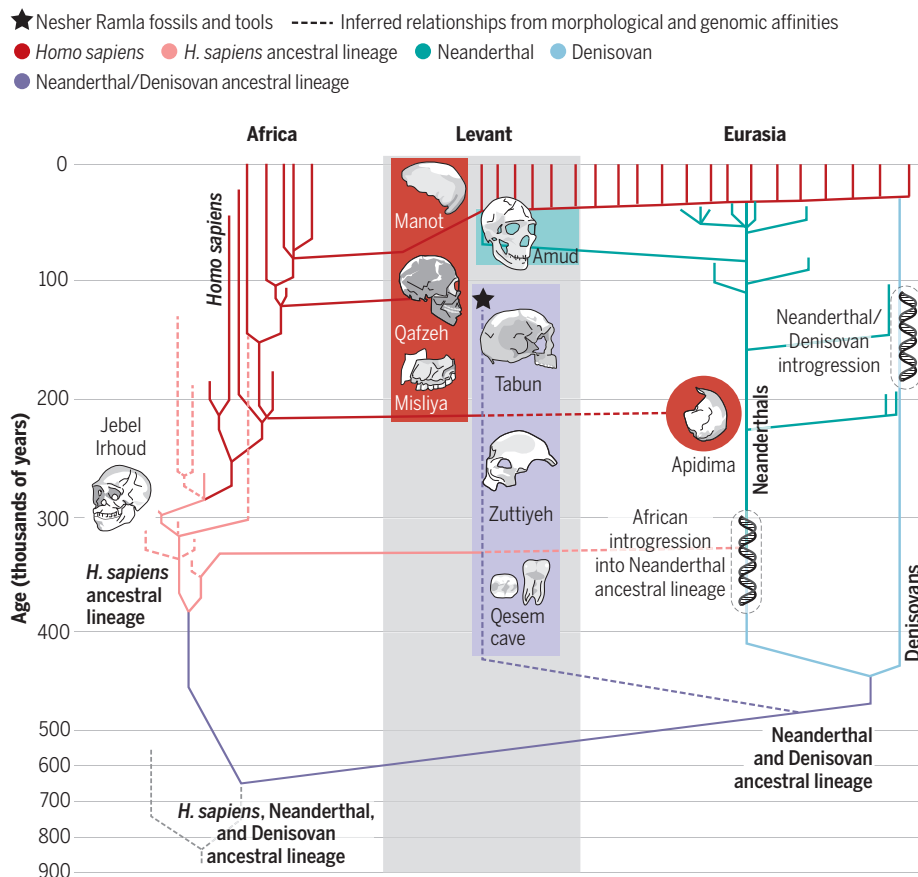
REFERENCES AND NOTES

1. C. M. Schlebusch *et al.*, *Science* **358**, 652 (2017).
2. L. R. Berger *et al.*, *eLife* **4**, e09560 (2015).
3. J.-J. Hublin *et al.*, *Nature* **546**, 289 (2017).
4. R. Grün *et al.*, *Nature* **580**, 372 (2020).
5. I. Hershkovitz *et al.*, *Science* **372**, 1424 (2021).
6. Y. Zaidner *et al.*, *Science* **372**, 1429 (2021).
7. K. Prüfer *et al.*, *Nature* **505**, 43 (2014).
8. F. Mafessoni *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 15132 (2020).
9. K. Harvati *et al.*, *Nature* **571**, 500 (2019).
10. I. Hershkovitz *et al.*, *Science* **359**, 456 (2018).
11. C. Posth *et al.*, *Nat. Commun.* **8**, 16046 (2017).
12. M. Petr *et al.*, *Science* **369**, 1653 (2020).
13. M. Meyer *et al.*, *Science* **338**, 222 (2012).
14. M. Kuhlwilm, S. Han, V. C. Sousa, L. Excoffier, T. Marques-Bonet, *Nat. Ecol. Evol.* **3**, 957 (2019).
15. E. J. Richards *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2011811118 (2021).

10.1126/science.abj3077

Amping up HIV antibodies

High serum titers of neutralizing antibody can protect humans against HIV

By **Dennis R. Burton**[1,2]

The recent results (1) from the Antibody Mediated Prevention (AMP) study are a landmark in AIDS research. They show that a broadly neutralizing antibody (bnAb) can protect humans against infection with sensitive strains of HIV. BnAbs, antibodies that can neutralize a large fraction of globally circulating HIV strains, are the focus of many HIV vaccine efforts and of strategies to prevent or treat HIV by passive immunization (infusion with antibodies). This focus results from the enormous strain variability of HIV—only bnAbs can hope to counter this variability. The results have profound implications for HIV vaccine design and for the use of passively administered bnAbs. Within the context of earlier animal model studies, the AMP study results provide insight into the factors that are important for antibody protection against HIV.

The AMP study involved two trials that enrolled nearly 3000 HIV-negative men and transgender men who have sex with men in the United States, South America, and Europe and nearly 2000 HIV-negative women in Africa to receive the bnAb VRC01 at two doses, 10 or 30 mg/kg, every 8 weeks for a total of 10 intravenous (IV) infusions. VRC01 is a bnAb, isolated in 2010 from an HIV-positive donor (2), that potently neutralizes a large proportion of circulating strains, although it is not as potent as some bnAbs isolated more recently. The dosing regime maintained serum bnAbs within defined limits. Overall, the study failed to show efficacy because the numbers of infected individuals in the treated groups were not significantly different from those in the placebo groups.

However, if the study was examined in terms of the neutralization sensitivity of viruses isolated from trial participants who became infected, then there was an important difference between treated and placebo groups. The frequency of infection by VRC01-sensitive viruses [defined as those that are 80% neutralized in vitro (80% inhibitory concentration, or IC_{80}) by less than 1 µg/ml

of VRC01] was significantly lower in treated individuals than in placebos. In other words, there is evidence that the antibody offers protection against neutralization-sensitive viruses. The data can be translated into a protective efficacy against HIV infection of ~50% at very approximate serum neutralizing titers of 1:100, ~75% at 1:250, and ~90% at 1:500. A serum neutralizing titer of 1:100, denoted as 50% inhibitory dose (ID_{50}) = 100, indicates that the serum of the donor could be diluted 100-fold and it would produce 50% neutralization in an in vitro neutralization assay; higher ID_{50}'s (more potent neutralizing sera) are required to achieve higher levels of protection. In the AMP study, neutralization was measured in a high-throughput pseudovirus assay (3), and an important conclusion of the study was that, with some caveats, this assay could be used to predict bnAb-based protection against HIV infection in humans. The lack of overall efficacy in the study arose from an underestimate of the serum bnAb titers required for protection in humans. Estimates were made based in part on an incorrect assumption that bnAb titers that were lower than those required in studies of nonhuman primate (NHP) viral infection would suffice.

Many studies have investigated the ability of antibodies to protect against HIV infection in animal models. Passive immunization with IV transfusion of a monoclonal antibody was shown to protect a single chimpanzee against challenge with a laboratory-adapted neutralization-sensitive virus as early as 1992 (4). Antibody protection titration studies were initially carried out in severe combined immunodeficiency mice populated with human peripheral blood lymphocytes (hu-PBL-SCID mice) that could be infected with HIV. Such studies showed that passive immunization with a first-generation bnAb could protect against both laboratory-adapted, neutralization-sensitive and neutralization-resistant HIV challenge (5). The latter are much more representative of global circulating viruses. The concentrations of bnAbs required for protection in this crude mouse model were high; serum ID_{50}'s were typically on the order of 100. Similar ID_{50}'s were generally noted in a second animal model: high-dose mucosal challenge of NHPs with chimeric simian–human immunodeficiency viruses (SHIVs) that have the surface envelope (Env) glycoprotein of HIV (so are neutralized by HIV bnAbs) but

the remaining gene segments of SIV (and thus replicate in NHPs).

A comprehensive examination (6) of available bnAb NHP SHIV protection data showed, in a logistic model that adjusts for bnAb epitopes and challenge viruses, that the serum ID_{50}'s to achieve 50, 75, and 95% protection were 91, 219, and 685, respectively. These numbers are similar to those estimated from the AMP study for protection in humans (see the figure). Excluding a small number of outliers, the NHP study suggested that protection was largely dependent on serum neutralizing antibody titer alone and was independent of antibody specificity (dose is adjusted to give similar neutralizing titers for antibodies of differing potency against the virus) and independent of challenge route.

Recent HIV Env vaccine-induced protection in NHPs was consistent with passive bnAb studies and showed ~90% protection at ID_{50}'s greater than ~500 (7). In that study, because bnAbs have not yet been induced through vaccination, the NHPs were immunized with an Env protein of a given strain and then challenged with the virus of the same strain, and the nAbs ("autologous" nAbs) provided protection. A common caveat that has been attached to animal model protection studies is the use of relatively high-dose viral challenge to ensure that all control animals become infected, which keeps the number of animals per study manageable. Consequently, the viral challenge dose typically used in NHPs was thought to be much higher than the average dose of HIV to which humans were exposed in sexual transmission. Thus, it had been assumed that lower bnAb titers than those derived from animal model studies would provide protection against HIV in humans, but the AMP study results do not support this assumption. In particular, protective titers (at least in terms of ID_{50}'s) from high-dose SHIV challenge in NHPs correlate well with those from passive bnAb VRC01 protection against HIV in humans.

Is a requirement for high bnAb concentrations for protection against HIV unusual? It is likely that bnAbs used in passive immunization to protect against HIV must provide sterilizing immunity, which is defined here as either neutralizing every transmitted virion directly or containing virus replication so that infection can be rapidly aborted without a primary infection. A requirement for sterilizing immunity is expected for a retrovirus that can readily establish latency. Thus, even a small number of virus particles could infect and become integrated into host chromosomes to be later activated and establish a full-blown infection. However, the requirement for high protective nAb concentrations in passive transfer studies is not unusual

[1]Department of Immunology and Microbiology, International AIDS Vaccine Initiative (IAVI) Neutralizing Antibody Center, Consortium for HIV/AIDS Vaccine Development, The Scripps Research Institute, La Jolla, CA 92037, USA. [2]Ragon Institute of MGH, MIT and Harvard, Cambridge, MA, USA. Email: burton@scripps.edu

for viral infections. For example, this is the case in the favored cotton rat model of respiratory syncytial virus (RSV) challenge, which was used to inform the dose of the RSV nAb palivizumab to protect at-risk infants from RSV lung disease (8). Of note, in passive immunization, nAbs are the sole immune mechanism of protection, whereas in vaccine-induced immunity, cellular immune responses may also be important and lower nAb titers may be highly beneficial.

Why are such high serum bnAb concentrations in humans relative to in vitro neutralization titer required for protection against HIV? Several potential contributors can be considered. Higher nAb concentrations may be required to prevent viral entry into cells in vivo than with in vitro assays. Indeed, in the AMP study, neutralization was measured using a pseudovirus and a target cell line, compared with the natural situation of a replicating virus targeting primary $CD4^+$ T cells or mononuclear phagocytes. However, although there are well-established differences in nAb sensitivity between the two virus target cell systems, they are not sufficient to explain the high serum bnAb concentrations needed for HIV protection. Another factor could be that serum may not be appropriate for nAb measurement; mucosal tissue has much lower nAb concentrations but may be more critical. An argument against this interpretation is that similar serum nAb concentrations are required for protection in hu-PBL-SCID mice and in IV challenge in NHPs (9). Alternatively, high nAb concentrations may be required in vivo to mediate critical functions other than neutralization, such as inhibition of cell-cell spread of virus or Fc-dependent effector functions of bnAbs. Neither of these mechanisms can currently be ruled out. Perhaps most notably, the neutralization curve is asymptotic, meaning that relatively high serum nAb concentrations are required to neutralize every virion if multiple viruses are transmitted and/or if some replication occurs after transmission. This could help explain the general nature of the observation that high serum nAb concentrations are required to completely protect against viruses from different families under widely varying conditions such as animal model and route of challenge and with antibodies of differing specificity.

The AMP study results have important implications for the use of bnAbs for the prevention of HIV infection and for HIV vaccine design. In terms of prevention, a first question is how extendable are the results to bnAbs other than VRC01 likely to be? The animal model studies suggest that most other bnAbs can be expected to offer protection in humans at similar neutralizing titers, although this remains to be determined and

there may be outliers, as there are in NHP studies. In addition, a neutralization-resistant fraction of virus may reduce the efficacy of some antibodies against some HIV strains even though the virus is defined as sensitive in terms of an IC_{50} or an IC_{80} (10).

The results also suggest that, to achieve and sustain the proposed protective titers, attention should be focused on the most potent bnAbs, likely as cocktails and likely with half-life extension mutations. For example, the half-life of VRC01 in the AMP study was around 2 weeks, but this could be increased ~fivefold with antibody engineering. Should one incorporate antibody effector function–enhancing mutations into HIV prophylactic bnAbs? Current animal model studies suggest that the effect of such mutations is an-

Correlates of protection

The graph shows serum neutralization titers (ID_{50}, bootstrap confidence intervals are shown in parentheses) aggregated from studies of simian–human immunodeficiency virus (SHIV) in nonhuman primates (NHPs) [reproduced from (6)]. The red dots show the approximate serum neutralization titers for humans from (1). This suggests that similar titers are required to protect humans and NHPs against virus.
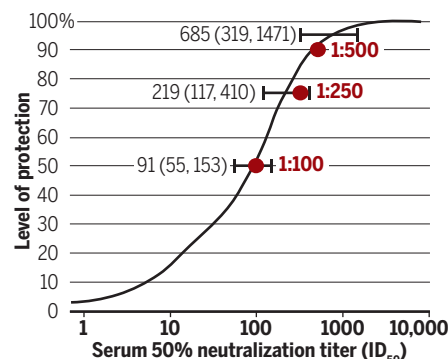
The figures for human 50% inhibitory dose (ID_{50}) were deduced from 80% inhibitory concentration (IC_{80}) and mean antibody serum concentration and multiplied by 3.

tibody dependent, and further research is required before such a step is taken (11). Major issues for the use of HIV prophylactic bnAbs include whether sufficient potency can be achieved to administer antibodies subcutaneously rather than intravenously, whether multivalent platforms such as bispecific or trispecific antibodies can be generated as effective reagents to replace single-specificity antibody cocktails (12), whether unit costs can be brought down to make bnAbs a feasible option for prevention in low- and middle-income countries (13), and how bnAbs compare with long-acting small-molecule drugs in terms of cost, efficacy, longevity of action, and other factors (14).

The AMP trial results pose a challenge for

HIV vaccine design; the induction of sustained high serum bnAb titers will be difficult to achieve. If protection is simply related to serum antibody neutralizing titer irrespective of the antibody specificity involved, it may be easier to induce moderate titers against multiple sites than high titers against a single site given current difficulties in inducing bnAbs through immunization. Targeting multiple sites should also restrict virus neutralization escape. An alternative approach that might ameliorate a requirement for high serum bnAb concentrations is to try to induce protective concentrations of antibodies at mucosal surfaces through mucosal vaccination, although it is not clear that mucosal immunity alone could prevent HIV infection in humans. Lower serum bnAb titers may also be beneficial in the presence of effective cellular immunity. Indeed, a recent study in NHPs showed protection against SHIV infection for a vaccine that induced both cellular immunity and nAbs at lower concentrations than those associated with protection induced by a vaccine that induced only nAbs (15).

Overall, it is likely that data from the AMP study will be mined for some time and will doubtless provide new insights into HIV infection and the role of neutralizing antibody in containing virus infection. For example, infected donors in the AMP study provide a large number of contemporaneous viruses that have been sexually transmitted and their sensitivities to a range of bnAbs, thereby potentially improving existing virus panels used for assessing bnAbs. Notably, the study results are the most important development yet in establishing the goalposts for passive immunity through antibody transfusion as a prevention strategy and for bnAb-based HIV vaccine design. ∎

REFERENCES AND NOTES

1. L. Corey et al., N. Engl. J. Med. 384, 1003 (2021).
2. X. Wu et al., Science 329, 856 (2010).
3. M. Li et al., J. Virol. 79, 10108 (2005).
4. E. A. Emini et al., Nature 355, 728 (1992).
5. M. C. Gauduin et al., Nat. Med. 3, 1389 (1997).
6. A. Pegu et al., Cell Host Microbe 26, 336 (2019).
7. M. G. Pauthner et al., Immunity 50, 241 (2019).
8. S. Johnson et al., J. Infect. Dis. 176, 1215 (1997).
9. Y. Nishimura et al., J. Virol. 76, 2123 (2002).
10. L. E. McCoy et al., PLOS Pathog. 11, e1005110 (2015).
11. L. Hangartner et al., Sci. Transl. Med. 13, eabe3349 (2021).
12. N. N. Padte et al., Retrovirology 15, 60 (2018).
13. IAVI/Wellcome, "Expanding access to monoclonal antibody-based products: A global call to action" (2020); https://bit.ly/3vbFnDN.
14. R. M. Gulick, C. Flexner, Annu. Rev. Med. 70, 137 (2019).
15. P. S. Arunachalam et al., Nat. Med. 26, 932 (2020).

David B. Wake (1936–2021)

Pioneering evolutionary biologist and amphibian advocate

By **Kelly R. Zamudio**

David B. Wake, a pioneer in the fields of evolutionary morphology, evolutionary developmental biology (evo-devo), and organismal diversification, died on 29 April. He was 84. Wake was a career-long visionary in organismal biology who led evolutionary biologists to examine not only how organisms are different but also how they become different. As a graduate student, he set the framework for his career by detailing the evolutionary relationships and morphological diversity of salamanders. He then delved into functional morphology (how organismal structures work), evolutionary development (how developmental pathways influence diversification of form), and speciation (how species come to be). One of the most influential and integrative biodiversity scientists of his era, Dave was boundlessly curious about all aspects of evolution and unusually open-minded about new techniques and analyses.

Dave was born on 8 June 1936 and raised in South Dakota. He attended Pacific Lutheran College in Tacoma, Washington, where he became fascinated by salamanders after uncovering some while looking for insects for an entomology course. After receiving his BA in biology in 1958, he joined the lab of herpetologist Jay Savage at the University of Southern California (USC). At USC, he met Marvalee Hendricks, a fellow graduate student and scholar of caecilians, another understudied group of amphibians. Dave and Marvalee married in 1962 and became collaborators in life and in science. Dave completed his MS in 1960 and his PhD in 1964, both in biology at USC. He joined the faculty at the University of Chicago for 5 years and then moved to the University of California, Berkeley (UC Berkeley), where he was director of the Museum of Vertebrate Zoology from 1971 to 1998 and professor of integrative biology until his retirement in 2003. Marvalee joined the faculty at UC Berkeley as a tenure-track professor soon after Dave. From the time I first heard of them, they were known as "the Wakes," and for many of us, they were an early model for successful dual careers in academia.

Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, NY 14853, USA.
Email: kelly.zamudio@cornell.edu

An unapologetic organismal biologist, Dave used salamanders as a model taxon (as opposed to a model organism) to ask questions, answer some, and rework others, creating a cycle of ever-deepening inquiry into their evolution. Rather than focusing on a single model organism or particular evolutionary mechanisms, Dave developed extensive knowledge of many salamander species, enough to use the whole taxon as a model platform to inform his many research foci. In doing so, Dave achieved an unprecedented level of integration across approaches to address evolutionary mechanisms and their

consequences for diversification. His exemplary integration inspired a series of papers by James Griesemer in the field of history and philosophy of science. Dave's salamander research transcended boundaries of methodologies, specialties, lines of inquiry, and disciplines. He developed a distinctive form of scientific problem-solving and iterative questioning that synergistically increased our general understanding of evolution.

Dave developed expertise in phylogenetics, morphology, development, ecology, neurobiology, behavior, and physiology, and his discoveries were groundbreaking in many fields. Coupling knowledge of adult morphology, ontogeny, embryology, function, and selection, he developed predictions about the retention and/or loss of morphological structures during development and became one of the first to frame these findings as part of the nascent field of evo-devo. By combining detailed spatial knowledge of morpho-

logical variation, biogeography, behavior, and genomics, he contributed a classic example of speciation in action with his exploration of the salamander ring species *Ensatina*. These discoveries, none of which could have been made without integrating approaches, span micro- to macroevolution and are now classics in evolutionary biology. His accomplishments led to many accolades, including election to the American Philosophical Society, the American Academy of Arts and Sciences, and the National Academy of Sciences.

In the late 1980s, Dave was an early proponent of action in response to the alarming global decline in amphibians. He chaired the first Declining Amphibian Populations Task Force and raised awareness of the predicament posed to amphibians by anthropogenic changes in climate and landscapes. As with his own research, he promoted diverse approaches in finding the causes of this large-scale biodiversity loss, to the benefit of both the scientific community and amphibians.

I joined the Wake lab in 1996 as a postdoc. It was an exciting time, with many field trips and countless discussions of species concepts, salamander tongues, and amphibian declines. Dave had a work ethic that amazed everyone and often left us challenging our expectations of ourselves. His leadership style emphasized showing, not telling. That work ethic resulted in some humorous "Wakeisms"—when lab members took vacations that he perceived as just a bit long, he began lab meetings by naming those who were absent and stating that "they must be gallivanting around the world." To this day, former Wake lab members refer to vacationing as doing just that.

Dave set high standards for research and expected us to meet them. He was honest in his delivery of criticism but somehow made it feel like it was for our own good. He was an unwavering supporter of those who worked with him and incredibly loyal to his students and colleagues. No matter how busy, he eagerly welcomed visiting students and early-career scientists, always showing interest in their stories and offering advice about their research and professional development.

Nearly 15 years ago, at a symposium organized around the potential of new genomic approaches in herpetology, Dave regaled the audience with what he saw as the biggest questions still to be answered by integrating this new approach. He genuinely reveled in witnessing the advancement of science, not only through his own work but also through that of his lab members and anyone else who stepped up to the plate. He concluded with a typically positive outlook on science: "I only wish I had another 50 years to live, just to see what you are all going to discover!" ■

BOOKS *et al.*

SCIENCE AND SOCIETY

Cultivating discerning citizens

A pair of psychologists confront the obstacles that threaten to undermine scientific literacy

By **Stephen M. Casner**

"Are vaccines safe for my baby?" wonders a new mother. After reading a few articles online that seem authoritative, she steps away from the computer and decides that there is not enough evidence to answer her question definitively. This scenario appears in the first of many vignettes in *Science Denial* that educational psychologists Gale Sinatra and Barbara Hofer use to confront a worrisome problem that extends beyond ideological science denial itself: the denial of science to those who seek credible information and who are often in great need of it.

Most people who search for information online favor trusted, easy-to-find sources. What they encounter is a forum that offers a platform to anyone with an online marketing strategy. Sinatra and Hofer point to a rise in "the sophistication of those who wish to portray fiction as fact." Herculean efforts are

Science Denial: Why It Happens and What to Do About It
Gale M. Sinatra and Barbara K. Hofer
Oxford University Press, 2021. 208 pp.

being made by determined lobbies to counter scientific sources and undermine public confidence in science itself, they note, and even websites run by government agencies can sometimes stray from scientific consensus. As intelligent virtual assistants become more widespread and the number of online information searches performed daily continues to rise, we become more and more tethered to an information source that can be as misleading as it is valuable.

Sinatra and Hofer remind us that we are more vulnerable to misinformation than we may think. Those who craft messages that run counter to accepted science know that the layperson's understanding of science is limited. They know that people are quick to use simple heuristics and the opinions of those around them as substitutes for deeper investigations. Hearing the same message repeatedly and seeing a few friends nod their heads in agreement with it can make it seem more credible. Appeals to remain "fair and balanced" are sometimes used to convince people to give equal consideration to messages that fly in the face of scientific consensus.

The reviewer is at the Ames Research Center, NASA, Moffett Field, CA 94035, USA. Email: stephen.casner@nasa.gov

Some students will go on to do science, but all students must learn to be good consumers of science.

The authors join other psychologists who remind us that our own biases can prompt even the most prudent among us to dismiss scientific findings when they conflict with what we think we already know about the world. For example, drivers are known to remain confident in their ability to safely multitask behind the wheel, even when that ability has been measured and confirmed to be poor (*1*). And once our beliefs have been formulated, we often come to personally identify with them and regard negations of them as personal criticism.

Sinatra and Hofer argue that the path toward a better future starts in schools, at the loftiest conceptual levels. Educators should strive to help kids form what they call a "science attitude"—one that places value on the truth, on hypotheses and theories that have a fair chance to be right or wrong, and most of all on evidence. A science attitude needs to be accompanied by science knowledge, including a familiarity with where good research can be found and a basic understanding of the methods used to evaluate scientific hypotheses, including practices such as peer review and replicability. School curricula, they argue, need to prepare some students to become producers of science and all students to become good consumers of science.

But what can be done about the grown-ups who already hold beliefs that run counter to scientific consensus? The authors offer hope that reason can prevail. Experimental evidence suggests that strongly held beliefs in unsupported theories can be moderated or even overturned using refutational techniques that identify specific misconceptions, state that they are incorrect, and detail the reasons why. The authors describe their success with using these techniques to change, or at least moderate, strong negative opinions about genetically modified foods.

Falling somewhere between academic and trade writing, *Science Denial* is filled with relatable scenarios, research studies, and helpful advice for individuals, educators, science communicators, and policy-makers. As social media discussions of science topics continue to proliferate and carefully reported coverage of science continues to decline, the authors warn readers to ready themselves for a future in which separating fact and fiction may be more difficult than ever. Their book offers abundant practical guidance to help us meet the challenge. ∎

REFERENCES AND NOTES

1. D. M. Sanbonmatsu, D. L. Strayer, N. Medeiros-Ward, J. M. Watson, *PLOS ONE* **8**, e54402 (2013).

10.1126/science.abj0816

PHOTO: DAVID GILDER/SHUTTERSTOCK

Envisioning the emotive mind

A neuroscientist recounts a career spent searching for insight into our psychic struggles

By **Stephen T. Casper**

Karl Deisseroth's *Projections: A Story of Human Emotions* is a memoir about his life in psychiatry, the patients he has met, and his efforts as a scientist to solve the mysteries of mental illness. His imaginative narrative flows effortlessly through his many passions. There is a first love of reading and writing and hints of a literary imagination that draws on James Joyce and Toni Morrison. And there is the light of optogenetics, a trick of nature pioneered in Deisseroth's laboratory at Stanford University that makes brain cells responsive to photonic pulses.

Like many doctors, scientists, and caregivers before him, Deisseroth endeavors to find an emotive language for the psychiatric patients he meets. He accomplishes this by weaving together what he describes as three strands: psychiatric knowledge, technology, and imagination. These together free him to write poetically, seeking to discover loss and grief, fracture from reality, and disruptions that threaten selfhood. The patients who wander through *Projections* thus encapsulate a variety of heartbreaking tragedies and shocking breaks from reality. Along with individuals suffering from anorexia and bulimia, Deisseroth describes cases of profound depression, mania, autism, self-harm, stroke, dementia, and schizophrenia.

His narratives are always sensitive, but they come to the reader as an admixture of fact and fiction, reality and imagination, damage and desire. Deisseroth is, in essence, proposing that we must use our imaginations to understand those in psychiatric crisis.

At one point, he projects himself into the mind of one of the patients he encounters and subsequently diagnoses as experiencing psychosis. He asks us to think with her about a neighbor's satellite dish that seemed "to download her thoughts." A trick then occurred to her: "She dug out a heavy black knit cap...pulled it down tight over her ears...there was no doubt—the satellite signal felt less likely to get in, or her own

thoughts to leak out." The reader at once comprehends the perceived threat posed by the satellite dish and the patient's desire to preempt the mind reader's damaging effects on her psyche. If an entity is trying to read our mind, shouldn't we stop it? And shouldn't we stop the psychiatrist who seeks to control us?

Our expectations of such control prove illusory in all ways. Deisseroth is disturbed to discover that people around him, a student in particular, have seemingly hidden mental disorders in plain sight. "What had I missed?" he whispers in the prose, just as we might ask ourselves were we placed in a similar situation. He ruminates on his own

Optogenetics pioneer Karl Deisseroth combines science and personal reflections in a new memoir.

identity as a single father, leaving unsaid but obvious to his readers that the mysteries psychiatry confronts may strike any of us at any time and we may be more blameworthy than we know.

Against these imaginative strands, Deisseroth situates his optogenetics research as a technological source of illumination. Optogenetics helps us understand the way ancient anatomical pathways or deeper evolved drives can project into our behaviors. Each time, however, the pull of the science, and indeed what the science shows, only takes him back to his suffering patients.

In a long analysis of dementia, Deisseroth contemplates the way optogenetics has revealed that valence—the sense

Projections: A Story of Human Emotions
Karl Deisseroth
Random House, 2021.
256 pp.

of value that accompanies brain states—is made across long-range connections in the brain. He notes that a memory must be accompanied by feelings in order for it to have meaning and salience. In dementia, memory falls away, leaving only feelings of helplessness. Optogenetics can show that those feelings exist and suggest that they might have once been tied to memories of a life with meaning, but it cannot restore that which is lost when memory fails nor can it alleviate the burdens borne by individuals with dementia or those who care for them.

In a chapter describing eating disorders, Deisseroth editorializes the explicit ambiguity his science creates in his book. There is, he writes, "a key difference between experiments with optogenetics and realities of the disease." Optogenetics allows scientists to turn cells off and on. But neurons do this themselves all the time without the help of a scientist, and no one knows how. There is something hollow in our struggle to make truth from mind.

Deisseroth never establishes balance between the cold rationality of his technology and the heroic and emotive struggles his prose otherwise records. Indeed, for all of their literal brilliance, the optogenetic experiments he discusses ultimately render our psychic struggles harder rather than easier to understand.

These tensions are familiar in books of this genre that foreground case histories of mental disease to explore humanity's ever-shifting natures and boundaries. However, *Projections* sits awkwardly in this tradition because it breaks from the authoritative objectivity that typically guides readers. Deisseroth instead embraces the fictional and poetic and leaves it up to us to decide what and where authority on such matters resides. The result is a trenchant story about human emotions that may leave some readers wanting answers where there are none. ■

The reviewer is associate director of the Clarkson Honors Program, Clarkson University, Potsdam, NY 13699, USA. Email: scasper@clarkson.edu

Afghanistan's unexplored mountains could hold hidden biodiversity.

Edited by **Jennifer Sills**

Biodiversity research in a changing Afghanistan

In the second half of the 20th century, 90% of major armed conflicts took place in countries that fall in biodiversity hotspots (*1*). Afghanistan is not considered a biodiversity hotspot (*2*), but this designation may be inaccurate. Biodiversity research is thus desperately needed to add to our knowledge of the region.

Afghanistan has suffered from more than 40 years of political instability and civil war, hindering far-reaching conservation research activities (*3*). The golden age of Afghanistan's biodiversity research took place decades ago (*4*, *5*), before the advent of DNA techniques used to evaluate biodiversity today. Afghanistan's biodiversity research is mostly based on museum collections assembled in the 1970s or before (*6–8*). Given the country's unique position, which is influenced by the Palearctic and Oriental biogeographical realms, and its habitat diversity, Afghanistan may hold a high level of hidden species and genetic diversity that are crucial for understanding the historical biogeography of Asia. Mountains have

driven past speciation events worldwide (*9*), yet the mostly mountainous Afghanistan has remained unexplored (*6*).

The first Afghan national park was established only 12 years ago, and others have followed, providing hope that wildlife research and conservation would contribute to the stability of the country (*3*, *10*). Although these national parks are a step forward, we still have little information about the distribution of the country's diverse biota or the threats that they face. Many species are likely in circumstances similar to the endemic, critically endangered mountain salamander, *Paradactylodon* (*Afghanodon*) *mustersi*, which is virtually unprotected and at increased risk due to human activities (*11*).

Despite the ongoing unrest in Afghanistan (*12*), the Afghan government must prioritize biodiversity research. Taking each local security situation into account, the government should work together with local universities and conservation organizations to bridge gaps in biodiversity research and seek support from the international scientific community. Protected areas should be expanded, and local communities should be supported and empowered to safeguard them. Scientists should collaborate to reestablish and update natural history museum

collections, identify species-rich areas, assemble comprehensive checklists of biota, and create national distribution atlases for all known species, especially those that are endangered.

Daniel Jablonski[1]*, **Abdul Basit[2]**, **Javeed Farooqi[3]**, **Rafaqat Masroor[4]**, **Wolfgang Böhme[5]**
[1]Department of Zoology, Comenius University in Bratislava, Bratislava, Slovakia. [2]Wildlife Conservation Society, Kabul, Afghanistan. [3]Department of Zoology, University of Chitral, Chitral, Pakistan. [4]Zoological Science Division, Pakistan Museum of Natural History, Islamabad, Pakistan. [5]Department of Vertebrata, Zoological Research Museum Alexander Koenig, Bonn, Germany.
*Corresponding author.
Email: daniel.jablonski@uniba.sk

REFERENCES AND NOTES

1. T. Hanson *et al.*, *Conserv. Biol.* **23**, 578 (2009).
2. C. C. Shank *et al.*, *Biodiversity Profile of Afghanistan* (United Nations Environment Programme, Kabul, 2008).
3. P. Smallwood *et al.*, *BioScience* **61**, 506 (2011).
4. C. N. Naumann, G. Nogge, *GAWI Newsletter* **73**, 25 (1973).
5. A. A. Nahif, *Bonn Zool. Bull.* **37**, 311 (1986).
6. P. Wagner *et al.*, *Proc. Calif. Acad. Sci.* **63**, 457 (2016).
7. D. Jablonski *et al.*, *ZooKeys* **843**, 129 (2019).
8. D. Jablonski, R. Masroor, *Bonn Zool. Bull.* **70**, 63 (2021).
9. C. Rahbek *et al.*, *Science* **365**, 1114 (2019).
10. E. Onion, "In Afghanistan, a new national park carries hopes for conservation and peace." *Mongabay* (2020).
11. S. N. Stuart *et al.*, Eds., *Threatened Amphibians of the World* (International Union for the Conservation of Nature, Gland, Switzerland, 2008).
12. S. Worden *et al.*, "U.S. Withdrawal from Afghanistan: End to an Endless War?", *United States Institute of Peace* (2021).

Developing countries must fund local research

Conversations on the imperialistic undertones in global research funding overlook the lack of support for researchers in developing countries from their own governments (1). Research spending by developing countries is often meager to nonexistent (2), resulting in an overreliance on foreign grants and international collaborations. The lack of alignment between research priorities of foreign funders and developing nations can lead to local research activities that do not benefit local communities. For that reason, member states of the African Union recently agreed to spend 1% of their gross domestic product (GDP) on research and development (R&D) to promote responsibility and ownership of their own economic growth (3). This positive commitment is an encouraging start, but developing nations will not succeed with good science policy alone; they also need a supporting research funding infrastructure and strong financial commitments.

Zambia's past efforts demonstrate how good intentions can fall short. Despite launching an ambitious science policy in 1996 that established the National Science and Technology Council and mandated an R&D spending of 3% of GDP, Zambia lags in innovation and spent less than 0.6% on R&D in 2014 (4). In 2016, a Zambian parliamentary committee attributed the lack of success to a lack of legislation, harmonized planning and coordination, and interaction between scientists and policy-makers, as well as disjointed funding infrastructure for R&D (4).

Zimbabwe's President Emmerson Mnangagwa recently signed the promising Manpower Planning and Development Amendment Act of 2020, which provides a funding infrastructure plan, into law (5). The legislation will establish a Manpower Development Fund and stimulate the business sector to support the Innovation and Industrialization Fund. It also outlines plans for several academies of science that are similar to the Chinese Academy of Sciences and South Korea's Institute for Basic Science. However, because the Act does not include any strong commitments to R&D spending, the infrastructure could stall without financial support.

Rwanda has shown how strong financial commitments underpin R&D success. To ensure that it meets its target of 1% of the GDP by 2024, the country has established two national research and innovation funds, integrated R&D spending into the national budget, and stated target funds for basic and applied research in priority areas. Rwanda has also promoted partnerships between the public and the business sector and offered incentives for the business sector to invest in R&D (6). Such strong commitment to R&D has already yielded results: Rwanda developed innovative strategies to combat the COVID-19 pandemic (7).

For science and technology to drive economic growth, developing countries need ambitious science policies, concrete R&D funding commitments, and a supporting funding infrastructure. African countries should study the challenges faced by Zambia, the gaps in Zimbabwe's laws, and the model set by Rwanda as they formulate their own R&D policies.

Edmond Sanganyado
Guangdong Provincial Laboratory of Marine Biotechnology, Institute of Marine Science, Shantou University, Shantou, Guangdong, China, and Research Excellence Initiative, Zimbabwe Young Academy of Sciences, Harare, Zimbabwe.
Email: esang001@ucr.edu

REFERENCES AND NOTES

1. N. A. Erondu *et al.*, *Nat. Med.* **27**, 742 (2021).
2. K. P. Acharya, S. Pathak, Applied Research in Low-Income Countries: Why and How? *Front. Res. Metrics Anal.* **4**, 10.3389/frma.2019.00003 (2019).
3. African Union Commision, *Science, Technology and Innovation Strategy for Africa 2024* (Ethiopia, 2020); https://au.int/sites/default/files/documents/38756-doc-stisa_science_tech_innovation_strategy.pdf.
4. National Assembly of Zambia, *Report of the Committee on Education, Science and Technology for the First Session of the Twelfth National Assembly* (Zambia, 2016); www.parliament.gov.zm/sites/default/files/documents/committee_reports/Educational%20Science%20and%20Technology.pdf.
5. Parliament of Zimbabwe, *Manpower Planning and Development Amendment Bill, 2020* (Zimbabwe, 2020; http://veritaszim.net/sites/veritas_d/files/Manpower%20Planning%20and%20Development%20Amendment%20Bill%202020.pdf.
6. National Council for Science and Technology, *Science, Technology, and Innovation policy* (Rwanda, 2020); http://ncst.gov.rw/sites/default/files/documents/official/STI_POLICY_2020.pdf.
7. N. Karim *et al.*, *Ann. Glob. Health* **87**, 23 (2021).

Science transcends national borders

The US Endless Frontier Act, which recently passed in the Senate as part of the "US Innovation and Competition Act" (1), has been lauded by the academic community [e.g., (2, 3)] for its massive increase in science and technology funding—a proposed total of US$250 billion over the next 5 years (4). However, this proposed legislation is not solely about funding science and technology. As J. Mervis explains in his News story "Senate panel backs funding ban on US researchers in Chinese talent programs" (13 May, https://scim.ag/3vCoxOF), Section 303 of the bill prohibits scholars who participate in China's national talent program from receiving US national funding, serving as the primary applicant in US federal grants, and, "to the extent possible," being listed on US federal grant applications. Even beyond Section 303, the bill is unequivocally framed as a response to the "China threat," building on the Cold War rhetoric of space races and iron curtains. The Act intends to set the United States on a path to out-innovate, out-produce, and out-compete China in strategic emerging industries. Healthy competition between prosperous nation-states should be welcomed on the global stage, but the latent Cold War rhetoric is at best misplaced and at worst counterproductive.

An extreme competitive focus, in addition to increasing anti-Asian sentiment within the United States (5), could stifle scientific and policy collaboration between the United States and China in the face of much larger existential threats, such as global climate change. China and the United States share a multitude of common causes, such as helping the world end poverty, halting biodiversity loss, and promoting peace and development. Rousing nationalistic sentiment at the expense of scientific cooperation may be a fleeting strategy for bipartisanism within the halls of Congress, but in the long run, it is a lose-lose proposition for the planet. The world has witnessed the consequences of politicization associated with the COVID-19 pandemic, with soaring cases and fatalities alongside anti-scientific denialism and increasing anti-Asian sentiment. The politicization of science inherent in the Endless Frontier Act is not a productive path forward. The so-called "endless frontier" that science is said to enable can only be realized if it does not stop at national borders.

Annah Zhu[1], George Zhu[2], Ruishan Chen[3]*
[1]Wageningen University, Wageningen 6706KN, Netherlands. [2]Double Bind Media, San Francisco, CA 94116, USA. [3]East China Normal University, Shanghai 200042, China.
*Corresponding author.
Email: chenrsh04@gmail.com

REFERENCES AND NOTES

1. US Congress, S.1260— United States Innovation and Competition Act of 2021 (2021); www.congress.gov/bill/117th-congress/senate-bill/1260.
2. "AAU President welcomes reintroduction of 'Endless Frontier Act'" (2021); www.aau.edu/newsroom/press-releases/aau-president-welcomes-reintroduction-endless-frontier-act.
3. A. DeMarco, "Achieving the endless frontier," *Inside Higher Ed* (2021); www.insidehighered.com/views/2021/03/30/congress-should-support-science-through-proposed-endless-frontier-act-opinion.
4. D. E. Sanger, C. Edmondson, D. McCabe, T. Kaplan, "Senate poised to pass huge industrial policy bill to counter China," *New York Times* (2021).
5. "Cold War with China is a dangerous and self-defeating strategy" (2021); https://winwithoutwar.org/wp-content/uploads/2021/05/60-Orgs-No-Cold-War-with-China-joint-statement-May-2021.pdf.

Fellowship highlights need for science communicators

Mass Media internships provide new careers and skills

By **Becky Ham**

The 47th class of Mass Media Science & Technology Fellows heading out to newsrooms around the country this summer is part of one of the oldest ongoing science communication initiatives by the American Association for the Advancement of Science (AAAS).

Since 1975, more than 750 scientists, engineers, and mathematicians in the fellowship have spent 10 weeks reporting and writing news stories at media outlets around the US. For the participants, the program is a way to explore a new career or improve their science communication skills. But alumni and AAAS staff say the fellowship also expands the public audience for scientific news.

"Journalism, writing, communicating about science—these are all critical for shaping public understanding of the world around us," said AAAS CEO Sudip Parikh to the newest fellowship class on 2 June.

The program has several alumni who made the switch from science to journalism after the fellowship. David Kestenbaum, a former correspondent for National Public Radio and now senior editor at This American Life, got his start in radio as a 1997 Mass Media fellow. At the time, he was a physics researcher at Fermilab, but the writing life beckoned.

"I did have this feeling, looking at press coverage of particle physics that I was doing, that it seemed very sterile and disconnected from the real lives of what it was like to actually do the science…this isn't actually getting at the craziness of it," he recalled at the orientation for the 2021 fellows. "And so I had some idea that I could bring something different to it."

Other fellows continued their academic careers. Johanna Varner, an assistant professor of biology at Colorado Mesa University, had considered switching to journalism during her 2015 summer working at KQED Science but took a teaching position instead. During the fellowship, she learned film production and storytelling skills that she uses in her teaching, especially with students who aren't science majors.

"I think having done the Mass Media fellowship is one of the most valuable experiences that I have ever had," said Varner, who received AAAS's 2018 Early Career Award for Public Engagement with Science for her citizen science initiatives and serves as a AAAS IF/THEN Ambassador.

"I think that it made me a better teacher, it made me better as a scientist trying to portray my own research in the media in a way that I will be happy with, and it gave me the skills to teach students to be more informed consumers of news and media," she said.

Screeners needed for journalism awards

Scientists from the United States and abroad are needed to review the scientific accuracy of entries in the prestigious AAAS Kavli Science Journalism Awards competition. The screening sessions in late August and September will be online this year, opening them to participation by scientists beyond the Washington, DC, area. We need additional screeners with expertise in virology, epidemiology, and public health. If you can volunteer, please contact Emily Hughes at ehughes@aaas.org.

The fellowship program has expanded into online video sites and Spanish-language media and has made a concerted effort to place fellows in newsrooms outside of the East and West Coasts, said program director Kristin Lewis.

"Local journalism is one of the places where this fellowship really shines," Lewis said, since these smaller media outlets may not have dedicated science writers on staff or resources to cover science stories.

She recalled a story from 2019 fellow Jerald Pinson at *The Austin American-Statesman*, who covered a toxic outbreak of blue-green algae in Lady Bird Lake. "With him, they had an ecologist in-house who could understand what was going on, speak to the experts, and really be able to translate that story for the local community," said Lewis.

The fellowship brings science content to new audiences in this way, she added. "The people who are reading *Scientific American* or *Wired* are seeking out science news, generally. The people who are reading your local paper might not be, they might just be looking for the news that is local to them. And so if we can find ways to make those science stories relevant to local audiences, the better."

Relevant storytelling is one of the things that fellows learn during their summer—"how to write stories with science in them as opposed to 'science stories,'" said Bill Manny, a former community engagement editor with the *Idaho Statesman*, who supervised 2018 fellow Kevin Davenport at the newspaper.

Davenport wrote about everything from bird migration to interstate accidents, figuring out "how you can use a science perspective and a science background on just good basic journalism," Manny said.

Throwing off the mantle of "scientist" can be difficult in unexpected ways, said Katherine Wu, a 2018 alumna who is now a staff writer at *The Atlantic*. "I think one thing that wasn't immediately intuitive to me was that when I became a journalist, I no longer wanted to identify as a science communicator," she told the 2021 fellows. "I didn't want to position myself as a scientist championing the institution of science, I wanted to portray reality as it was [and] try to put a more objective lens on the way that I was approaching all of these things."

Neal Baer, a 1983 alumnus and pediatrician who spent his internship at WEWS-TV, has continued to work in both media and academia. His career includes stints as a television writer and producer on ER, Law & Order: Special Victims Unit, and Designated Survivor, among other shows. He also lectures at Harvard Medical School and contributes to scientific journals. Baer is sponsoring the 2021 AAAS Diverse Voices in Science Journalism Internship.

"I think I'm able to apply those two worlds to my work, the academic world that I was in for so long, and also this world where people have dialogue, and so it's not, to me, that wholly different," said Baer at the orientation. "It's always about telling stories grounded in science."

Last year, AAAS partnered with science journalism nonprofit The Open Notebook to create a 6-month mentorship program for the 2020 fellows, who missed out on some of the usual newsroom networking and guidance due to COVID-19 restrictions. Lewis said that the fellowship would like to continue the program if financial support is available.

"The pandemic of the past year has only served to heighten the importance of accurate science journalism," Lewis said, "so this is a particularly relevant fellowship that AAAS is funding at the moment."

RESEARCH

EVOLUTIONARY BIOLOGY
Thriving in the benthic zone

Lobsters are among the most successfully adapted organisms of marine benthic ecosystems. Despite lacking an adaptive immune response, lobsters display a long life-span of up to 100 years. To help uncover the secret of their longevity, Polinski *et al.* sequenced the American lobster genome and identified genome-wide innovations in genes related to chemosensory machinery, innate immunity, and cellular defense. These include a new class of chimeric receptors with the potential to integrate neuronal and immune sensory systems. The results provide insights on the mechanisms mediating the ecological success of this benthic predator. —DGL

Sci. Adv. 10.1126/sciadv.abe8290 (2021).

The genome of the American lobster has uncovered distinctive adaptations in this benthic invertebrate.

SUPERCONDUCTIVITY
Imaging an exotic state

Among the most intriguing of the many phases of cuprate superconductors is the so-called pair density wave (PDW) state. PDW is characterized by a spatially modulated density of Cooper pairs and can be detected with a scanning tunneling microscope equipped with a superconducting tip. Liu *et al.* used Josephson tunneling microscopy, modified for the task, to detect PDW in niobium diselenide, a superconductor with a layered hexagonal structure. The PDW state is expected to appear in other transition metal dichalcogenides as well. —JS

Science, abd4607, this issue p. 1447

EPIGENETICS
Methyl readers that repress transcription

DNA methylation is a conserved epigenetic mark required for gene silencing in many different organisms. However, how the methyl mark is able to silence genes is still largely unknown. Ichino *et al.* discovered two *Arabidopsis* proteins named MBD5 and MBD6 that are recruited to DNA by direct recognition of methylation. These methyl readers recruit the class C J-domain protein SILENZIO to chromatin to silence methylated genes and transposons. SILENZIO likely acts through its interaction with heat shock chaperone proteins. —DJ

Science, abg6130, this issue p. 1434

ORGANIC CHEMISTRY
Easing oxygen into arenes

Although oxygen is all around us, it is often surprisingly difficult to use it for selective chemical oxidations, necessitating more expensive, wasteful alternatives. Li *et al.* report that careful ligand optimization produces palladium catalysts that can efficiently activate oxygen to hydroxylate a variety of aryl and heteroaromatic rings adjacent to a carboxylic acid substituent. The ligand binds to palladium through pyridine and pyridone components, and the authors posit that tautomerization between dative and anionic coordination modes plays a role in its effectiveness. —JSY

Science, abg2362, this issue p. 1452

T CELLS
Rewiring aged T cells

Age-associated decline in T cell function contributes to impaired immune responses to infection and vaccination. Effector versus memory T cell differentiation is controlled in part by signaling and metabolic reprogramming mediated by mechanistic target of rapamycin complex 1 (mTORC1), which is typically activated at amino acid–producing

lysosomes. Jin *et al.* demonstrate that in naïve CD4+ T cells from older individuals, mTORC1 activation instead occurs at late endosomes and depends on the amino acid transporter SLC7A5. Late endosomal mTORC1 impairs T cell lysosomal function, reducing the degradation of PD-1 and proliferative responses. Silencing VPS39, a gene that promotes late endosome formation, was able to increase the proliferation of aged human T cells and memory responses of lysosome-defective T cells in a mouse viral infection model, demonstrating that targeting late endosomal mTORC1 activity may improve T cell function. —CO

Sci. Immunol. **6**, eabg0791 (2021).

CATALYSIS
Rhodium atoms for alkane dehydrogenation

Nanoparticles of rhodium dispersed on metal oxides are generally poor catalysts for alkane dehydrogenation because the reactants bind too strongly to the metal. Hannagan *et al.* performed first-principle calculations indicating that single rhodium atoms in a copper surface should be stable and selective for conversion of propane to propene and hydrogen. Model studies of single rhodium atoms embedded in a copper (111) surface revealed a very high selectivity to propene

An artist's rendition of the electron density redistribution occurring as a single methane molecule is activated by a rhodium atom within a copper surface lattice

and high resistance to the formation of surface carbon that would deactivate the catalyst. —PDS

Science, abg8389, this issue p. 1444

MATERIALS SCIENCE
The making of a monolith

Amorphous calcium carbonate is a hard material that is difficult to make into large, clear blocks. It is also sensitive to heating, and compacting the starting nanoparticles too much tends to lead to crystallization. Mu *et al.* determined the optimal amount of water in amorphous calcium carbonate to create clear, solid monoliths through compression. The key is to regulate the amount of diffusion in the system so that particle boundaries fuse without triggering sample-wide crystallization. This fusion strategy may also work for similar amorphous inorganic ionic compounds. —BG

Science, abg1915, this issue p. 1466

CORONAVIRUS
Masking out air sharing

The effectiveness of masks in preventing the transmission of severe acute respiratory syndrome coronavirus 2 has been debated since the beginning of the COVID-19 pandemic. One important question is whether masks are effective despite the forceful expulsion of respiratory matter during coughing and sneezing. Cheng *et al.* convincingly show that most people live in conditions in which the airborne virus load is low. The probability of infection changes nonlinearly with the amount of respiratory matter to which a person is exposed. If most people in the wider community wear even simple surgical masks, then the probability of an encounter with a virus particle is even further limited. In indoor settings, it is impossible to avoid breathing in air that someone else has exhaled, and in hospital situations where the virus concentration is the highest, even the best-performing masks used without other protective gear such as hazmat suits will not provide adequate protection. —CA

Science, abg6296, this issue p. 1439

IN OTHER JOURNALS
Edited by **Caroline Ash** *and* **Jesse Smith**

MEDICINE
Acceptable algorithms for radiotherapy

Machine-learning applications in medicine have so far promised more than they have delivered. McIntosh *et al.* evaluated an algorithm that was integrated into the clinical workflow to plan curative-intent radiation therapy for prostate cancer. Human- and algorithm-generated treatment plans were compared in a blinded manner by physicians and one plan was selected. The machine-learning plans were generated faster than the human-generated plans and were selected by physicians for 72% of patients. However, when it came to treating patients, implementation of the machine-learning–generated plans decreased, likely because of the perception and preferences of the treating physicians and their experience to ensure patient care. Thus, such real-world variables need to be accounted for in studies of medical applications for machine learning to increase its utility and acceptance in the clinical setting. —GKA

Nat. Med. **27**, 999 (2021).

STEM WORKFORCE
Grassroots effort for change

Efforts to increase diversity in STEM (science, technology, engineering, and math) will not be truly successful until the systemic issues within academia that have historically prevented marginalized populations from persisting and succeeding there are transformed. Stachl *et al.* present the details of a collaborative effort to improve the academic climate of an R1 (very high research activity) STEM department: that

FOOD SECURITY

Beef and more than two veg

To grow beef requires many more resources and emits 11-fold more greenhouse gases than raising poultry. It is easy to say eat less meat to mitigate environmental degradation and climate change, but how to do this? Eshel calculated how adopting nitrogen-sparing agriculture (NSA) in the United States could feed the country nutritiously, enhance carbon sequestration, and reduce nitrogen leakage into water supplies. The proposal is to shift to small, mixed agricultural enterprises in which the nitrogen budget is kept closed (nitrogen and carbon resources are biogeochemically tightly linked in agriculture). The core of a 1.43–hectare unit NSA farm is an intensive cattle facility from which manure production supports a variety of rain-fed and rotated crops of plant-based foods for humans, as well as livestock fodder. Less than half of today's current cropland in the United States is suitable for NSA. Nevertheless, this area has the potential to feed 330 million Americans and still have some left over for export. Arid rangelands could thus be spared or left to low-intensity grazing. —CA

PLoS Biol. 10.1371/journal.pbio.3001264 (2021).

In a rain-fed nitrogen-sparing agricultural system, intensive cattle farming can still play a role.

boundary, can be seen clearly in the geological record on time scales of hundreds of thousands to millions of years. However, how this affected the marine biosphere in the decades and centuries after the strike is difficult to tell from proxy data because of their low temporal resolution. Brugger *et al.* modeled the immediate aftermath of sulfate aerosols, carbon dioxide, and dust caused by the impact, and found a short-lived algal bloom caused by upwelling of nutrients from the deep ocean and nutrient input from the impactor, a strong temperature decrease, and moderate surface ocean acidification. These results help to fill in a gap that proxies have left blank. —HJS

Geophys. Res. Lett.
10.1029/2020GL092260 (2021).

CELL BIOLOGY

Revealing hotspots for interaction

Biomolecular condensates are compartments in eukaryotic cells that are not membrane bound but can still concentrate specific functions. These phase-separated entities are increasingly recognized for their role in regulating a range of cellular processes. Interactions between intrinsically disordered protein regions are key to regulating phase separation, but experimentally obtaining atomic-resolution information remains challenging. Kim *et al.* have developed a suite of nuclear magnetic resonance (NMR) techniques to obtain quantitative and site-specific data on interactions that govern the phase separation of the RNA-binding protein CAPRIN 1. Aromatic and arginine residues are known to be important, but this work also shows a major role for backbone interactions. The NMR experiments also explore how interactions are modulated by posttranslational modifications or interactions with ATP. —VV

Proc. Natl. Acad. Sci. U.S.A. **118**,
e2104897118 (2021).

of the University of California, Berkeley. Longitudinal assessment of the academic climate by annual department-wide surveys indicated that these interventions have succeeded in shifting perceptions. These results support the idea that implementing practical, sustainable, and data-driven frameworks for effecting change can improve the climate within a departmental community. —MMc

ACS Omega **6**, 14410 (2021).

GRAPHENE GROWTH

Graphene growth on molten copper

Although graphene can be grown on solid copper surfaces, higher precursor pressures that lead to faster growth can be used on molten metals, and smoother graphene films can be made on liquid surfaces.

However, investigating chemical vapor deposition growth on a liquid surface can be challenging. Jankowski *et al.* used a combination of x-ray diffraction and reflectivity, Raman spectroscopy, and optical microscopy methods to characterize and control graphene growth on molten copper. By changing the ratios of methane to hydrogen, the authors were able to produce highly ordered flake assemblies, and by using a single nucleation seed, they could grow millimeter-scale, single-crystalline graphene sheets. —PDS

ACS Nano 10.1021/acsnano.0c10377 (2021).

DEVELOPMENT

Cavernous lesions

Blood vessels within the brain supply oxygen but also allow circulation of other molecules, ions, and cells between

the blood and the brain. About one in 200 people form cerebral cavernous malformations (CCMs), which are irregular collections of small cerebral vessels that alter the flow of blood and can result in headaches, seizure, hemorrhage, paralysis, or stroke. Li *et al.* used the zebrafish model system and CRISPR-Cas9 mutagenesis to inactivate the CCM gene. As a result, the caudal venous plexus became dilated, "honeycombing" of the vein lumen disrupted blood flow, and multiple blood-filled chambers with a risk of hemorrhaging formed. —BAP

eLife 10.7554/eLife.62155 (2021).

EXTINCTION EVENTS

Immediate impact

The environmental effects of the Chicxulub impact, the trigger of the mass extinction event at the Cretaceous–Paleogene

ALSO IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

CORONAVIRUS
Boosterism could save lives

Postinfection immune protection against severe acute respiratory syndrome coronavirus 2 reinfection is not fully understood. It will be devastating if waves of new variants emerge that undermine natural immune protection. Stamatatos *et al*. investigated immune responsiveness 4 to 8 months after previously infected individuals were given a messenger RNA–based vaccine developed for the original Wuhan variant (see the Perspective by Crotty). Before vaccination, postinfection serum antibody neutralization responses to virus variants were variable and weak. Vaccination elevated postinfection serum-neutralizing capacity approximately 1000-fold against Wuhan-Hu-1 and other strains, and serum neutralization against the variant B.1.351 was enhanced. Although responses were relatively muted against the variant, they still showed characteristic memory responses. Vaccination with the Wuhan-Hu-1 variant may thus offer a valuable boost to protective responses against subsequent infection with variant viruses. —CA

Science, abg9175, this issue p. 1413; see also abj2258, p. 1392

URBAN MANAGEMENT
Adapting to the new normal

Successfully responding to the impacts of climate change will be a challenge for many communities, especially cities. Considering the situation in the United States, Shi and Moser examine how stakeholders can help to build urban resilience even in the absence of federal leadership. They discuss how local and state governments, private industry, and civil society can engage to adapt to the extreme weather events and other consequences of changing climate that are expected in the future. Preparing for these looming issues requires coherent, cohesive, and collective responses across all scales and sectors of society. —HJS

Science, abc8054, this issue p. 1408

CELL BIOLOGY
Tailoring stress responses

When faced with environmental stress, cells respond by shutting down cellular processes such as translation and nucleocytoplasmic transport. At the same time, cells preserve cytoplasmic messenger RNAs in structures known as stress granules, and many cellular proteins are modified by the covalent addition of ubiquitin, which has long been presumed to reflect degradation of stress-damaged proteins (see the Perspective by Dormann). Maxwell *et al*. show that cells generate distinct patterns of ubiquitination in response to different stressors. Rather than reflecting the degradation of stress-damaged proteins, this ubiquitination primes cells to dismantle stress granules and reinitiate normal cellular activities once the stress is removed. Gwon *et al*. show that persistent stress granules are degraded by autophagy, whereas short-lived granules undergo a process of disassembly that is autophagy independent. The mechanism of this disassembly depends on the initiating stress. —SMH

Science, abc3593 and abf6548, this issue p. 1409 and p. 1410; see also abj2400, p. 1393

EPIGENOME
MeCP2 binds hydroxymethylated CA repeats

Despite of decades of research on the Rett syndrome protein MeCP2, its function remains unclear. Ibrahim *et al*. show that MeCP2 is a hydroxymethylated cytosine-adenosine (CA) repeat-binding protein that modulates chromatin architecture at a distance from the transcription start site (see the Perspective by Zhou and Zoghbi). MeCP2 accumulates and spreads around modified CA repeats and competes for nucleosome occupancy. Loss of MeCP2 results in a widespread increase in nucleosome density inside lamina-associated domains and transcriptional dysregulation of genes enriched in CA repeats. These results shed light on the underlying molecular mechanism of Rett syndrome, a severe disease associated with mutations in MeCP2. —DJ

Science, abd5581, this issue p. 1411; see also abj5027, p. 1390

MICROBIOLOGY
Cell death limits pathogens

During infection, *Yersinia* inhibition of the protein kinase TAK1 triggers cleavage of the pore-forming protein gasdermin D (GSDMD), which leads to a kind of inflammatory cell death called pyroptosis. In a genome-wide screen, Zheng *et al*. identified a lysosome-tethered regulatory supercomplex as being a critical driver of *Yersinia*-induced pyroptosis. The activity of the GTPase Rag and lysosomal tethering of Rag-Ragulator were required to recruit and activate the kinase RIPK1 and protease caspase-8 to cleave GSDMD, which causes cell death and limits infection. By contrast, Rag-Ragulator was not required for inflammasome-mediated pyroptosis. Thus, metabolic signaling on lysosomes can regulate cell death during pathogenic bacterial infection. —SMH

Science, abg0269, this issue p. 1412

FERROELECTRICS
Stacking a ferroelectric

Properties of layered van der Waals structures can depend sensitively on the stacking arrangement of constituent layers. This phenomenon has been exploited to engineer superconducting, correlated insulator, and magnetic states. Two groups now show that ferroelectricity can also be engineered through stacking: Parallel-stacked bilayers of hexagonal boron nitride exhibit ferroelectric switching even though the bulk material is not ferroelectric (see the Perspective by Tsymbal). To explore these phenomena, Yasuda *et al*. used transport measurements, whereas Vizner Stern *et al*. focused on atomic force microscopy. —JS

Science, abd3230 and abe8177, this issue p. 1458 and p. 1462; see also abi7296, p. 1389

PALEOANTHROPOLOGY
Middle Pleistocene *Homo* in the Levant

Our understanding of the origin, distribution, and evolution of early humans and their close relatives has been greatly refined by recent new information. Adding to this trend, Hershkovitz *et al*. have uncovered evidence of a previously unknown archaic *Homo* population, the "Nesher Ramla *Homo*" (see the Perspective by Mirazon Lahr). The authors present comprehensive qualitative and quantitative analyses of fossilized remains from a site in Israel dated to 140,000 to 120,000 years ago indicating the presence of a previously unrecognized group of hominins representing the last surviving populations of Middle Pleistocene *Homo* in Europe, southwest Asia, and Africa. In a companion paper, Zaidner *et al*. present the radiometric ages, stone tool assemblages, faunal assemblages, and other behavioral and environmental data associated with these fossils. This evidence shows that these hominins had fully mastered technology that until only recently was linked to either *Homo sapiens* or Neanderthals. Nesher Ramla *Homo* was an efficient hunter of large and

small game, used wood for fuel, cooked or roasted meat, and maintained fires. These findings provide archaeological support for cultural interactions between different human lineages during the Middle Paleolithic, suggesting that admixture between Middle Pleistocene *Homo* and *H. sapiens* had already occurred by this time. —AMS

Science, abh3169 and abh3020, this issue p. 1424 and p. 1429; see also abj3077, p. 1395

MALARIA
A role for IgA in malaria

Immunoglobulin A (IgA) is known to play a protective role against pathogens at mucosal surfaces. However, the protective effects of IgA in the serum are less well understood, particularly in the context of pathogens such as *Plasmodium falciparum*. Tan *et al.* isolated and characterized serum IgA from three independent cohorts of humans exposed to *P. falciparum*. The authors also studied IgA antibodies isolated from individuals who were consistently resistant to malaria and found that these antibodies bound to a conserved site on sporozoites and were protective in mouse models in vivo. These results establish a role for serum IgA in the context of malaria and suggest a region of the circumsporozoite protein as a target for protective antibodies. —CSM

Sci. Transl. Med. **13**, eabg2344 (2021).

CELL BIOLOGY
Delayed repression of a tumor suppressor

The rapid induction or repression of gene expression in response to growth factors such as epidermal growth factor (EGF) has been intensively investigated. Uribe *et al.* examined genes they called delayed down-regulated genes (DDGs) that are repressed hours after EGF stimulation in mammary cells. One DDG encoded the transcription factor TSHZ2, which interacted with and inhibited mitotic proteins through multiple mechanisms. Overexpression

of TSHZ2 inhibited mammary tumor growth and progression in mice. In human breast tumors, decreased *TSHZ2* expression often correlated with increased methylation of its promoter. —LKF

Sci. Signal. **14**, eabe6156 (2021).

MEDICINE
HIV antibody treatments

A recent clinical trial reported that treatment with a broadly neutralizing antibody failed to show efficacy in preventing HIV infection. However, further analysis of the data provides a glimmer of hope that could better inform both future trials aimed at achieving such passive immunity and vaccine development. In a Perspective, Burton discusses the data in the trial indicating that protection from HIV may be achievable with higher doses of infused therapeutic antibody. This trial highlights how inferences from animal studies could be improved for future human trials and helps to explain why higher serum antibody titers are needed for immune protection from HIV. —GKA

Science, abf5376, this issue p. 1397

CORONAVIRUS
A boost from infection

During clinical trials of severe acute respiratory syndrome coronavirus 2 vaccines, no one who had survived infection with the virus was tested. A year after the pandemic was declared, vaccination of previously infected persons is a reality. Reynolds *et al.* address the knowledge gap in a cohort of UK health care workers given the Pfizer/BioNTech vaccine in which half of the participants had experienced natural virus infections early in the pandemic (see the Perspective by Crotty). Genotyping indicated that a genetic component underlies heterogeneity in immune responses to vaccine and to natural infection. After vaccination, naïve individuals developed antibody responses similar to those seen in naturally infected

persons, but T cell responses were more limited and sometimes absent. However, antibody and memory responses in individuals vaccinated after infection were substantially boosted to the extent that a single vaccine dose is likely to protect against the more aggressive B.1.1.7 variant. It is possible that the messenger RNA vaccine has an adjuvant effect, biasing responses toward antibody generation. —CA

Science, abh1282, this issue p. 1418; see also abj2258, p. 1392

URBAN MANAGEMENT

Transformative climate adaptation in the United States: Trends and prospects

Linda Shi* and Susanne Moser

BACKGROUND: Climate disasters are on the rise, with devastating effects on communities, built infrastructure, and ecosystems. Between 2015 and 2020, the United States has seen an average of 14 disasters per year that cost at least $1 billion, compared with 6.5 such events annually between 1980 and 2019. The COVID-19 pandemic and attendant socioeconomic crisis have laid bare the United States' systemic vulnerabilities and difficulties in launching large-scale, coordinated, just, and effective responses to external shocks, resulting in short-term disruptions and prolonged crises. These same challenges inhibit societal adaptation to climate change, which is already being felt in communities nationwide as the number of disaster events and their geographic reach and intensity increase. The current financial downturn reduces resources for near-term resilience planning, further exposing cities to the next hazard event and driving vicious cycles of fiscal and environmental shocks. More than ever, preparing and adapting to climate impacts require a coherent, cohesive, and collective response across localities, sectors of society, and scales of governance.

ADVANCES: In this Review, we distill three major trends in federal government, industry, and civil society that shape how local communities adapt to extreme weather events and other climate change impacts. First, inconsistent federal leadership on climate adaptation has done little to address drivers of climate injustice and uneven development. The Obama administration initiated various policies to mainstream climate considerations into federal properties and investments that the Trump administration overturned. But even the Obama administration narrowly framed adaptation as disaster resilience, infrastructure investment, and national security issues rather than addressing the drivers of vulnerability, such as social, land, and income inequality. To date, the Biden administration appears to merely reinstate Obama-era approaches to adaptation.

Second, design, engineering, and legal professions are considering systemic amendments to building codes and standards that could force federal and state governments' hands in addressing climate risk. Financial industry decisions over how to rate credit worthiness, where to issue mortgages, and when to raise or rescind insurance have the power to send mortgage markets and municipal revenues tumbling and increase the costs of infrastructure investment. At present, decisions in these arenas reflect professional and board room concerns for industry risk and liability rather than justice- and community well-being–oriented outcomes.

Third, grassroots and academic advocates increasingly call on leaders to redress exclusionary and environmentally exploitative development and avoid using climate resilience to rehash racialized capitalist development. The groups emphasize the need for an ethics of care, restoring urban and rural communities' relationship to land ownership and stewardship, and deepening democratic engagement.

The growing divergence in how public, private, and civil society actors are responding to climate impacts contributes to maladaptive investment in climate-blind infrastructure, justice-blind reforms to financial and professional sectors, and ultimately, greater societal vulnerability to climate impacts.

OUTLOOK: A shift in presidential leadership alone will not alter the politics, power dynamics, and paradigms that shape US adaptation. Large-scale change such as infrastructure investments and managed retreat can preserve the status quo, unless reforms change the underlying social relationships and power dynamics and center a different set of values and beliefs about humans and human-nature relations. If public, private, and civil society actors are to take up the transformation imperative in a proactive and deliberate way, they need to address the material, relational, and normative factors that hold the current systems in place. Ambitious civil society and private sector leadership invite federal government to respond with bold, integrated, and holistic policies. Community movement strategies for coalition building, including ones across urban-rural divides, demonstrate how to build a political movement for just adaptation. Private-sector reforms of financial instruments pinpoint the places where community organizations and the federal government need to push for just and equitable adaptation. Despite substantial contestation, divergences also point to opportunities to better engage with and learn from one another to advance toward more transformative adaptation. We conclude with examples of possible directions for transformative practice and research in support of these efforts. ∎

Larger local jurisdictions, such as Oakland, and smaller neighborhoods around San Francisco Bay, California, struggle to advance urgent and equitable adaptation action. Leadership, coordination, capacity, regulatory alignment, and adequate funding are needed for coherent, cohesive, and collective preparedness and response. COVID-19 magnifies these needs and demonstrates the risks of complex crises.

S READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abc8054

CREDIT: PI.1415926535/WIKIMEDIA COMMONS/CC BY-SA 3.0

URBAN MANAGEMENT

Transformative climate adaptation in the United States: Trends and prospects

Linda Shi[1]* and Susanne Moser[2,3]

As climate change intensifies, civil society is increasingly calling for transformative adaptation that redresses drivers of climate vulnerability. We review trends in how US federal government, private industry, and civil society are planning for climate adaptation. We find growing divergence in their approaches and impacts. This incoherence increases maladaptive investment in climate-blind infrastructure, justice-blind reforms in financial and professional sectors, and greater societal vulnerability to climate impacts. If these actors were to proactively and deliberately engage in transformative adaptation, they would need to address the material, relational, and normative factors that hold current systems in place. Drawing on a review of transformation and collective impact literatures, we conclude with directions for research and policy engagement to support more transformative adaptation moving forward.

Climate disasters are on the rise, with devastating effects on communities, built infrastructure, and ecosystems. From wildfires in California to unprecedented floods in the Midwest and hurricanes affecting the Carolinas, Florida, Texas, Puerto Rico, and more, no corner of the United States is left unscathed (*1*). These events and their costly impacts are blind to political affiliation and jurisdictional boundaries, although they have the greatest consequences for disadvantaged groups who already struggle with poverty and marginalization (*1*). Even before the COVID-19 pandemic and resulting socioeconomic crisis, communities were inadequately prepared for emerging climate impacts. Recent events, however, have laid bare the United States' systemic vulnerabilities and constraints in launching large-scale, coordinated, equitable, and effective responses to external shocks, resulting in severe disruptions and prolonged crises (*2*). Existing challenges, chronic underfunding, subsequent short-sightedness, and ineffective government coordination—exacerbated by partisan politics and lack of consistent federal leadership—have hobbled state and local governments' ability to mount effective responses (*3*, *4*).

The Biden administration has taken bold steps toward leadership on climate change, raising hopes again that the logjam of action may finally be broken, bringing renewed focus on climate mitigation and adaptation. Biden's early climate initiatives indicate a transformative push for decarbonization, but vision-

ing documents and plans say little about the strategy for climate resilience. Proposals for a multitrillion-dollar infrastructure package would provide an influx of cash for local and state governments, but the desire for an "infrastructure fix" may be misguided, illusory, or simply inadequate. To date, no individual has been named to be the leading force on adaptation in the new administration, and it is reasonable to ask whether federal climate adaptation leadership in the intellectual and political footsteps of the Obama administration adequately addresses the challenges facing US communities.

Here, we review major federal policy changes, industry actions, and civil discourses on climate adaptation and how divergent societal trends inhibit efficient, effective, and fair adaptation to climate impacts on the ground. We examine these trends through a transformation lens and—more specifically—through insights from the collective impact literature on factors that hold systems in place (*5*–*7*). Drawing on the growing body of literature that critiques existing adaptation practices as inadequate, uncoordinated, and unjust, we examine the extent to which the sectoral trends reflect changes in policies, practices, resource flows, social relations, power dynamics, and underlying mindsets. We find that although there have been substantial shifts since the end of the Obama administration, they do not converge toward deliberative, purposive, and just adaptation. As the 2020 US election underscores, a change in the White House clearly establishes a new direction, but congressional partisanship, societal divisions, and structural and normative barriers challenge fundamental shifts in societal adaptation across the United States. This suggests that transformative adaptation requires tackling historical legacies and societal engagement.

The imperative for transformative climate adaptation

The frequency, severity, and costs of climate change are rapidly increasing. Between 2015 and 2020, the United States has seen an average of 14 disasters per year that cost at least $1 billion, compared to 6.5 such events annually between 1980 and 2019 (*8*). The Fourth National Climate Assessment (2018) details how climate change has contributed to and will exacerbate the frequency and intensity of acute and long-term disaster trends (*1*). For example, more than $1 trillion of coastal real estate and 13.1 million people are threatened by rising sea levels, higher storm surges, and higher tidal flooding (*1*). Under current emissions trends, $66 billion to $106 billion of real estate will highly likely be below sea level by 2050, one mortgage cycle away (*1*). Left unaddressed, failing infrastructure systems will reduce gross domestic product (GDP) by $3.9 trillion within just 5 years, roughly 20% of current GDP (*9*). Currently debated infrastructure solutions are costly and already beyond reach for many communities, incentivize continued development in hazardous locations, protect against only low to moderate levels of climate impacts, and are inequitably distributed across communities (*10*, *11*). Practitioners and the media now openly discuss large-scale relocation due to climate change, even while recognizing that climate is one among many social, economic, political, and environmental factors driving migration choices (*12*–*15*).

Past experiences with climate impacts and the COVID-19 crisis illustrate how hazard impacts fall disproportionately on the most vulnerable populations, those who are structurally disadvantaged because of race, income, or other identity markers in access to housing, jobs, basic services, and political voice in planning and decision-making (*1*). Moreover, teleconnected, compounding, and cascading risks increase exposure and vulnerability (*16*), reduce adaptive capacity (by eroding funding, diminishing staff capacity, and leaving fewer options), and contribute to erosive cycles of social trauma and economic decline (*3*, *17*, *18*). Without funding to shore up protective infrastructure or investments in water, food, energy, healthcare, and other systems, communities will experience unmitigated climate impacts sooner than expected, with knock-on impacts to local economies and community well-being (*19*). Without systemic policies, communities will experience—and be left on their own to navigate—displacement, unaffordable and insufficient housing markets, and declining economies and regions (*20*).

Growing civic discourse and bodies of research advocate for transformative adaptation that redresses the underlying drivers of societal vulnerability to climate change (*21*–*24*). The complexity and magnitude of the current

[1]Department of City and Regional Planning, Cornell University, 213 West Sibley Hall, Ithaca, NY 14853, USA. [2]Susanne Moser Research and Consulting, Hadley, MA 01035, USA. [3]Department of Landscape Architecture and Regional Planning, University of Massachusetts–Amherst, Amherst, MA 01002, USA.
*Corresponding author. Email: lindashi@cornell.edu

multistressor crisis (health, economic, climate, and racism) illustrate that incremental efforts "aimed at accommodating change, rather than contesting it" (*25*) no longer suffice. Whereas dominant approaches to adaptation address "end-point vulnerability" (the most visible symptoms of past development patterns, societal relations, and human-environment interactions), transformative adaptation responds to "starting-point vulnerability" (the deep roots and conditions producing vulnerability in the first place) (*26*). Transformative adaptation, as a relatively new concept, is still being explored and awaits widespread application in practice. It may include changing economic paradigms and development patterns away from those predicated on the exploitation of nature without limits; redressing systemic racism, imperialism, and misogyny; decolonizing of knowledge systems; reforming governance institutions that operationalize these developmental logics; and reckoning with underlying worldviews and values that legitimize dominance and exceptionalism.

In short, transformative adaptation is not just about "climate-proofing" existing structures and systems but about deliberately and fundamentally changing systems to achieve more just and equitable adaptation outcomes (*27*). This implies investigating the factors that maintain the status quo and strategically addressing them to intentionally shift systems in new directions.

Key factors in changing systems: A theoretical lens

Six fundamental factors that keep systems in place are shown in Fig. 1: At the surface level, policies, practices, and resource flows are determined at a deeper level by the processes, relationships, and power dynamics among those making decisions, which in turn is a reflection of the mindsets, values, and beliefs of those involved. The depiction implies the type and depth of change (Fig. 1, notations at left of triangle) and points to the needs and requirements to deeply transform systems, linked to a series of leverage points (Fig. 1, notations at right of triangle) to address them. Large, multiscalar, and multisystem coordination and reform to adaptation policies, practices, and resource flows would change the structural and material aspects of systems. However, as critics of the "resilience dividend" (*28*) and "disaster capitalism" and climate-centered movements for a Green New Deal (GND) and Extinction Rebellion (in the UK) have argued (*29, 30*), large-scale change can seek to preserve the status quo unless reforms change the underlying social relationships and power dynamics and center a different set of values and beliefs about humans and human-nature relations.

This conceptual framework of transformation (*31*) provides a basis for assessing the implications of how the major sectors of US society—public, private, and civil—are reacting to climate impacts. What changes at the level of policies, practices, and resource flows do they espouse with respect to climate adaptation? What power dynamics and relationships do these actions reflect? What mindsets and values inform these relationships and power dynamics? Who and what is valued more or less? How do they need to change to affect decisions and the direction of actions taken toward more just adaptation? What deeply held assumptions underlie adaptation decision-making processes, and what prevents them from changing?

Below, we review three trends that illustrate how and why emerging adaptation practices neither respond to the transformation imperative nor converge toward the necessary changes in the six factors that hold systems in place. The result is what can be observed across the United States at present: delayed, uncoordinated, and potentially ineffective or even maladaptive response to growing climate risks. Aligning the key actors for positive transformative change will require attending not only to the top-level structural and material aspects of adaptation (itself no easy task) but to the social relations, processes, and mindset changes necessary to achieve deeper transformational outcomes (*32*).

Recent trends in US adaptation
Trend 1: Federal restraint in and retreat from climate adaptation policy

Federal engagement with climate adaptation has been mixed. The Obama administration initiated adaptation policies in its second term, which the Trump administration actively revoked. The Biden administration has promised to embed climate resilience into diverse administrative efforts, to be carried out by mission agencies (for example, reinstating Obama's flood standards). To date, however, there has been no consistent or comprehensive approach to adaptation policymaking and limited federal funding support across all three administrations, resulting in a widening gap between levels of government and fragmented responses across jurisdictions. The tasks and costs of recovering from disasters have strained the
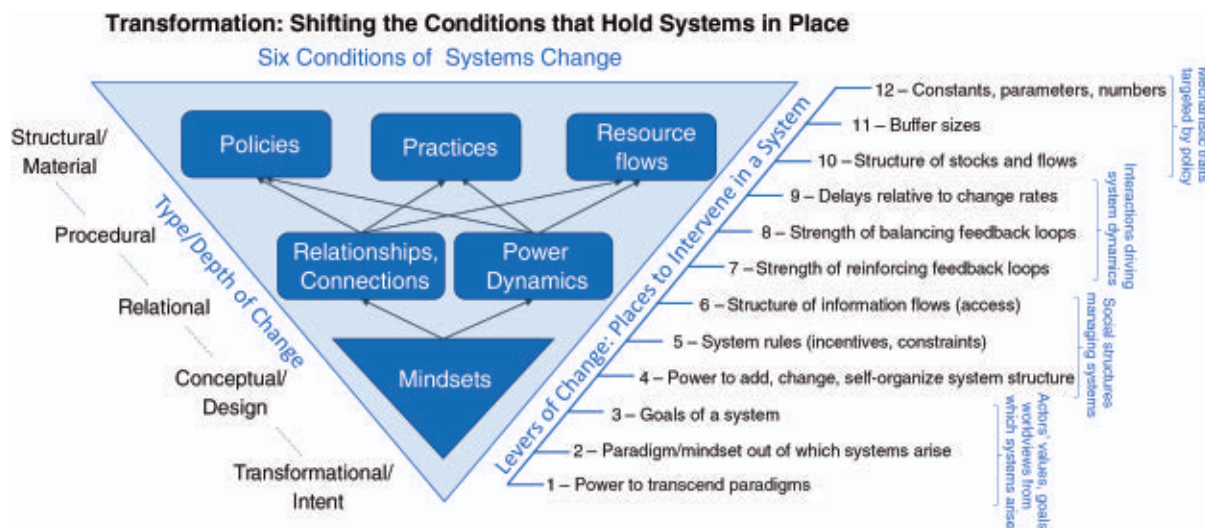
Fig. 1. Six conditions of transformative systems change. Systems are shaped by surface level, policies, practices, and resource flows, which in turn are driven by relationships and power dynamics among those making decisions, which in turn reflect stakeholders' mindsets, values, and beliefs (triangles). (Left) Transformative change of material outcomes stems from changing the intent and conceptual, relational, and procedural conditions driving systems. (Right) Levers to create change range from altering the structure of stocks and flows at more superficial levels, to changing incentives and constraints at deeper levels, to changing system goals and mindsets. [Reprinted with permission from (*31*)]

fiscal, administrative, and response capacities of the Federal Emergency Management Agency (FEMA) and state and local governments, whose existing programs and funds are far outstripped by the scale of climate impacts and needs (*33*, *34*). Moreover, beyond surface-level policy retreat or inconsistency, the federal government has consistently shown little support for national-level visioning on climate adaptation, justice-oriented adaptation planning, and governance reform, while increasing partisanship has hampered its ability to address societal vulnerability to changing land, infrastructure, and health conditions under climate change.

The federal government recognizes its climate risks as a major owner and operator of infrastructure, facilities, military installations, natural and working lands, and insurer of properties and crops (*35*, *36*), yet for years it has lacked a comprehensive approach to assessing and disclosing the vulnerability of federal assets and programs, much less planning for their adaptation, leaving states, localities, and individual agencies to make their own plans (*37*, *38*). Recognizing that leadership came from outside the federal government, the Obama administration formed a State, Local, and Tribal Leaders Task Force on Climate Preparedness and Resilience in 2014 (*39*). Following their recommendations, the administration issued a series of executive orders (Table 1) that required federal agencies to consider climate risks in their policies, practices, investments, and programs (using the federally mandated National Climate Assessments); to coordinate among themselves and with other levels of government; to issue plans; and to train staff (*36*, *40*).

The Trump administration rolled back these efforts by revoking each of Obama's executive orders (Table 1). Budget cuts to the US Environmental Protection Agency (EPA) and National Oceanographic and Atmospheric Administration (NOAA), as well as erasure of climate data on NOAA's portal, have reduced the federal government's capacity to meet state and local governments' technical data needs. A 2019 US Government Accountability Office (GAO) report found that the federal government has done nothing to reduce its high fiscal risk exposure due to climate change (*41*). Moreover, federal responses are inconsistent across agencies: The 2019 Department of Homeland Security's *National Preparedness Report* does not mention "climate change" or "sea-level rise" (or pandemics, for that matter) (*42*), yet the US Army Corps of Engineers (USACE) has begun to require localities that receive federal disaster assistance to use eminent domain to force reluctant homeowners from repeatedly flooded homes. This controversial application of eminent domain signals USACE's and FEMA's awareness of the urgent and worsening nature of flood impacts (*43*). Although reducing federal disaster recovery costs, such measures overlook how flood insurance and local land-use fiscalization incentivize floodplain development and how relocation without accompanying housing and fiscal reforms can harm households and municipal coffers.

Where federal agencies do plan for adaptation, the focus has been on "climate proofing" physical infrastructure without considering land-use policies and development paradigms that drive inequitable and unsustainable development. Under the Obama administration, federal action on adaptation accelerated after Hurricane Sandy in 2012. Federal, state, and local alignment on climate-resilient reconstruction, especially in New York State, enabled more effective recovery (*44*). However, to secure congressional support for the region's reconstruction and other climate projects, the Obama administration framed adaptation as disaster resilience, infrastructure investment, and national security rather than addressing the drivers of vulnerability and the need for systemic transformation (*40*). Politically sensitive issues such as housing and tax policy, food security, land inequality, rural and indigenous resource extraction, or long-term national and regional spatial planning have not been part of the national adaptation conversation. The Biden administration has demonstrated commitment to environmental justice and investment in wide-ranging social welfare programs, but specific climate policies have focused primarily on decarbonization. Attention to resilience is taking the form of infrastructure funding and connecting climate change to international migration—evidence that the administration's approach echoes policies of the Obama era.

Although the Trump administration's retreat on adaptation leadership has cost precious time, the possibly even more concerning legacy is the political polarization at the federal level that challenges science-based problem-solving and policymaking across partisan divides. The proposed GND connects decarbonization and clean energy to social justice issues of health care, wages, and housing (*45*, *46*). The approach remains embedded within the current economic development model but stresses redistribution to enable equitable development. Although attempts are underway to implement

---

Table 1. Shifts in federal policy regarding climate adaptation under the Obama and Trump administrations. Although the Biden administration has taken steps to revoke several Trump climate policies, domestic policy proposals for adaptation have not yet coalesced into a coherent approach.

Obama administration	Trump administration
Preparing the United States for the Impacts of Climate Change (Executive Order 13563, 2013) and *Climate Action Plan* (2013) required federal agencies to submit resilience plans to the White House Council on Environmental Quality (CEQ) and Office of Management and Budget (OMB).	Revoked by *Promoting Energy Independence and Economic Growth* (Executive Order 13783, 2017).
Planning for Federal Sustainability in the Next Decade (Executive Order 13693, 2015) required interagency working groups to form and coordinate resilience planning with other levels of government and stakeholders; OMB to review agencies' sustainability performance; and Office of Personnel Management to train federal staff on resilience.	Revoked by *Efficient Federal Operations* (Executive Order 13834, 2018).
Establishing a Federal Flood Risk Management Standard and a Process for Further Soliciting and Considering Stakeholder Input (Executive Order 13690, 2015) tasked federal agencies to consider climate science as part of all federal planning for facilities and programs for floods.	Revoked by *Establishing Discipline and Accountability in the Environmental Review and Permitting Process for Infrastructure Projects* (Executive Order 13807, 2017).
FEMA's *Climate Change Adaptation Policy* required federal (starting in 2012) and state (2015) programs and policies to account for climate impacts; General Services Administration's 2016 Guiding Principles Checklist (P100) for new construction or major renovation requires consideration of climate change.	2020 amendment to the *National Environmental Policy Act* absolved federal agencies from accounting for climate impacts in infrastructure proposals.

elements of the GND piecemeal, continued congressional partisanship raises questions about whether cooperative policy engagement will be possible (*47*), even if younger Republicans recognize the need for climate solutions and support some of the policies endorsed by the GND (*48*). Private industry shifts may open political doors and offer opportunities for pragmatic advances on climate action.

Trend 2: Industry reforms in practices and standards governing risk and liability

Standard-setting organizations overseeing engineering, design, insurance, and lending practices are moving toward incorporating climate science into their benchmarks, requirements, and guidelines. This reflects a growing concern with managing their own legal and financial liability risks, including the failure to anticipate climate impacts within project life cycles under "standards of care" expectations (*49–51*). Industry-led adjustments place growing pressures on state and local governments to act on climate adaptation. Against the backdrop of recent federal reversals, these shifts create substantial confusion of competing standards. Governments and their employees could themselves become liable for failing to anticipate climate impacts during the life span of projects in this era of nebulous and uncharted legal waters (*51*). Moreover, state and local fallout from these policy shifts may also increase demands on federal investments and aid (*52*).

Housing markets also are beginning to account for climate impacts, although there continues to be a spatial mismatch between real estate values and property risk (*53*). Recent research has found that nationally, exposure to sea-level rise, flooding, and higher insurance premiums is beginning to have a negative impact on property values (*54–56*). Calls to reform the National Flood Insurance Program—such as through better maps of future flood risk under climate change, disclosure agreements during property transactions, streamlined and expanded buyout programs, shifting from individual to community flood insurance, actuarial risk pricing, and addressing affordability and retention issues (*57, 58*)—accelerate market internalization of climate risk. In the absence of federal leadership on risk disclosure, private consulting firms (and some nongovernmental organizations) are growing their in-house technical expertise to map forward-looking flood risks, informing not only individual homeowner purchasing decisions but, implicitly, the real estate market (*58, 59*). These private-sector shifts affect property values, which can negatively affect individual households' primary asset as well as municipal tax rolls and locally funded services (*11*).

Globally, financial institutions—including public and private banks, insurance companies,

and bond-rating agencies—understand the shifting landscapes of market risk and are engaged in an "intelligence arms race" to measure climate impacts on investments and steer them to new speculative sites and cities (*60*). A network of central and public banks outside the United States is defining assets as "green" or "brown" according to their carbon emissions and climate risks. In tandem, the global Task Force on Climate-Related Financial Disclosures has tasked its 800-plus member public and private organizations holding $118 trillion in assets to disclose their climate risks (*61*). Domestically, bond-rating agencies such as Standard & Poor's and Moody's are starting to incorporate climate mitigation and preparedness measures into municipal bond ratings (*62, 63*).

Last, international and national associations of engineers are working to provide guidance on how to integrate forward-looking climate science into design guidelines, engineering standards, and standards of care. In 2015, the World Federation of Engineering Organizations issued a report, "Model Code of Practice: Principles of Climate Change Adaptation for Engineers," as guidance for voluntary adoption by individuals and nations (*64*). Canada is incorporating climate impacts into its national building code, and in 2018, the American Society of Civil Engineers Committee on Adaptation to a Changing Climate released a manual on incorporating forward-looking climate science into engineering design (*65*) and is working on a standard to further codify climate-sensitive engineering.

These reforms highlight the private sector's recognition of the reality of climate change. They also point to industry's agility in large-scale responses, even if they are not (yet) deviating from historic processes of unequal development or the underlying worldviews and paradigmatic commitments. Industry is far more rapidly learning about climate impacts; buying or investing in new modeling techniques to identify risk, liability, and potential "climate oases" (http://futuremap.io/climateoases) (*60*); and forming committees and task forces to systematize standards across large networks. As such, they represent a welcome, constructive shift in who joins and what is being discussed at the climate action table. However, these efforts still prioritize corporate investments and liability (that is, profitability) rather than human well-being or the inherent rights of nature (*66*). Contemporary observers see the financialization of climate risks presenting new sacrifice zones predicated on race (*66*)—for example, by funding speculative investment in a Resilient Rustbelt or by refusing to lend in "bluelined" low-lying areas just like banks once refused to lend in "redlined" black neighborhoods (*67*). Although disasters threaten to destabilize the insurance industry, insurers

also thrive on disasters, which swell demand for new policies (*68*). Municipalities that lose tax revenues from property devaluation may face growing difficulty maintaining bond ratings and/or repaying more expensive bonds, further destabilizing local service provision in a vicious cycle (*11, 69*). Whereas some cities may use climate resilience discourse to market their competitiveness (*70*) or benefit from shifting standards, the decline of "climate slums" in other municipalities contributes to new intra- and intermunicipal inequality (*71*). In other words, the private sector's engagement on adaptation is not oriented toward societal well-being or reparative justice for the historically disadvantaged.

Trend 3: Academic and community-based advocacy of transformative adaptation

Against the backdrop of lagging federal leadership, corporatization of climate action, and the increasingly pressing threats of climate change affecting communities, mobilization from the bottom up is a third notable trend. Civil society groups increasingly contest mainstream adaptation efforts as inadequately protecting historically marginalized communities—whether through "acts of commission" or "omission" (*72*). For example, the neighborhoods where Philadelphia expanded household green infrastructure to absorb stormwater saw the highest levels of gentrification (*73*). In Houston, historically Black communities that have long been neglected for drainage improvements were targeted for FEMA property buyouts after Hurricane Harvey using eminent domain, whereas similarly affected but wealthier and whiter communities were offered FEMA home renovation and elevation funding (*74*). In south Florida, historically disadvantaged Black and Haitian communities on higher elevations are being targeted for speculative development as markets slowly begin to penalize coastal condos (*75*). Additionally, whereas 5 years ago practitioners noted that "re" words (such as retreat, relocation, and resettlement) were forbidden, now there are conferences such as Columbia University's "At What Point Managed Retreat?," planning for whole-community (albeit not yet whole-city) resettlement (for example, in Louisiana and Alaska), and websites and initiatives devoted to "Climigration" (*76, 77*). Critical discourses of past redlining, buyout, and relocation approaches are emerging, questioning not only administrative inadequacies but the fundamental inequities besetting conventional technocratic approaches to relocation, both for "sending communities" and "receiving communities" in near-coastal and interior cities (*13, 14*).

In response, community-based groups are proposing alternative pathways forward (*78*). They have articulated new proposals for equitable and just adaptation (*79*) that emphasize

recognition of historic racism and oppression, demand solidarity and diverse and inclusive leadership, and uplift holistic visions of what it takes to create a safer, thriving society (*78*). They center opposition to racism and structural drivers of inequality and as such focus first on deep mindset and relational and power issues as foundations of policy, practice, and resource flow.

In parallel, a growing body of academic research demonstrates the oppressiveness and ineffectiveness of ongoing adaptation and so-called resiliency strategies that recapitulate historic exclusionary, extractive, and racist or classist policies (*79*–*82*). Research shows that vulnerability to natural hazards stem not only from the probability of risk or even individual sensitivity due to disability, age, poverty, or linguistic isolation. Rather, vulnerability results from historic centralization of political power, land, and resources; dispossession of communities from rural self-sufficiency and concentration of people in urban areas as productive workers and consumers; unequal allocation of land, housing, environmental goods (such as clean air, water, and parks); and—through voter disenfranchisement and divisive politics—sustaining a disadvantaged class that enables the social reproduction of this political economic system (*83*, *84*). Vulnerability assessments and resource allocations that consider only present-day variations in flood risk deny the impact of historic oppression and trauma, such as redlining and urban renewal; forced relocation of Native Americans and denial of their kinship ties and lifeways (*85*); exposure to toxic pollutants, particulate matter, and environmental extremes; and infrastructure and service provision (*86*, *87*).

Alternative visions of societal well-being have emerged from different groups, finding solidarity across race, gender, and urban-rural divides and building coalitions across issues. This is changing social relations and helps bring greater power against entrenched interests. Both researchers and community activists in this way have pushed local governments to recognize the relational and historic nature of present-day uneven vulnerability as a first step toward more just practices (*78*, *88*). Although some local adaptation plans now include language on participation and inclusion of disadvantaged communities in planning processes (*89*), community organizers go much farther by calling for building community leadership capacity and social movement infrastructure, forming coalitions and partnerships, and gaining elected office or decision-making power (*78*, *90*).

In an attempt to scale up such local efforts, Climate Justice Alliance, a group of 65-plus frontline and movement-building organizations and networks across the country, has articulated a framework for "Just Transition"

(*91*). This framework advocates a paradigm shift away from an extractive economy predicated on "global plunder, the profit-driven industrial economy rooted in patriarchy and white supremacy," to a regenerative economy that focuses on "redressing past harms and creating new relationships of power for the future through reparations" (*91*). This non-Eurocentric leadership model emphasizes alternative values and development possibilities, such as buen vivir, consent, trust, accountability, respect, reciprocity, and responsibility (*85*, *92*, *93*).

Prospects of transformative adaptation

These three trends can be distilled as federal retreat or stalling on comprehensive adaptation policy and support; private-sector engagement reinforcing existing economic paradigms and interests; and civil society critique, resistance, and movement building for transformative adaptation. The divergence among these actors makes clear that a shift in presidential leadership alone cannot shift the politics, power dynamics, and paradigms that shape adaptation in the United States. If public, private, and civil society actors were to take up the transformation imperative in a proactive and deliberate way, the six factors that hold systems in place would need to be addressed. Both scholarship and practice suggest how this might occur at the level of the material, relational, and normative (Fig. 1).

The following sections reflect on ways to redress underlying drivers of vulnerability and align conflicting discourses at each of these three levels of transformation. These ideas are visualized in Fig. 2. The diagram recognizes the need to adapt to changing geophysical climate conditions (Fig. 2, yellow arrow) but roots the causal conditions in historic social, political, and economic drivers of vulnerability (Fig. 2, gray arrow). The middle ring of Fig. 2 comprises strategies to transform adaptation normatively (by asserting values of recognition, justice, equity, and inclusion, in red), procedurally (through empowered and deliberative practices, in green), and materially (through large-scale, systemic thinking, in orange). Figure 2's inner wheel points out the need to embed adaptation's normative values, processes, and policies throughout the sectors that shape the country's physical and social well-being. This approach seeks to break through partisan divides by creating opportunities to reflect on shared traumas, shared values, and mutual interdependence of communities across the urban and rural landscape in grappling with the challenge of climate change.

Large-scale, systemic thinking in shaping policies, practices, and resource flows

Effective change must address large geographic scales and connections across scales and account for individual localities as embedded

units within territorial landscapes of regions, urban-rural gradients, and natural ecosystems (*94*, *95*). Traditional hazard assessments and adaptation planning focus narrowly on single cities, hazards, or sectors (*16*, *96*, *97*), despite the interconnectedness of landscapes, infrastructure, markets, and political systems. Instead, all governments need to assess and reduce their structural and material vulnerabilities to long-distance risks, compound impacts, and cascading failures as well as engage in multilevel coordination (*3*, *98*, *99*). Ambitious strategies such as a GND would put money toward infrastructure and new technologies, accelerate industrial and market shifts, and expand the sites where such technologies are available and adopted. New emphasis on nature-based solutions for flood risk reduction may also transform engineering and infrastructure operating practices (*100*–*102*). However, the large-scale adoption of new technologies alone does not challenge historic drivers of income, housing, and land inequality; the extractive nature of industrial development and urbanization; or the need for more dynamic and agile solutions in place of static infrastructure projects. Systemic policies such as the GND need to explicitly address these to change high-level policy guidance, practices, and resource flows (*103*, *104*).

Empowerment and deliberative processes

Any effort to transform the structural and material dimensions of systems necessitates transforming the relationships and power dynamics among actors as well as their ability to lead inclusively during turbulent times (*21*, *25*, *105*). Within the adaptation context, scholars and activists have called for changes in public engagement processes, as well as more transparent and accessible decision-making procedures, so that stakeholders' and rightsholders' concerns and needs have a better chance of being heard (*72*, *106*, *107*). They claim that more democratic and inclusive processes are more likely to address stakeholders' concerns, particularly around equity and justice (*108*, *109*). In addition, changing power structures by changing who is at the table of deliberative processes, whose voices get heard, who holds leadership positions, and who has the real political power to make adaptation decisions is essential to changing course and addressing deep-seated vulnerabilities (*78*, *110*).

As the election of 2020 demonstrates, winning elections alone does not eliminate the necessity of standing the difficult ground of principle while engaging in civil argument and education with those across the line of factual reality about racism, climate change, and so on, to find a path toward collective problem-solving. Deliberative and communicative planning research advocates processes that enable individual and community empowerment

Fig. 2. Framework for transformative adaptation.

Transformative adaptation needs to respond to the magnitude of climate risks (yellow arrow) by addressing root drivers of vulnerability (gray arrow). Large-scale, systemic thinking is necessary to coordinate adaptation across scales, sectors, and hazards (orange arrow). Such societal mobilization requires both deep deliberation across silos (green arrow) and an assertion of normative values of justice and equity (red arrow) so that large-scale actions do not repeat racist, inequitable, and unsustainable outcomes. Transformative thinking at all three levels (material, relational, and mindset) is needed in all areas that shape societal well-being and across urban-rural landscapes. [Original graphic by the authors]
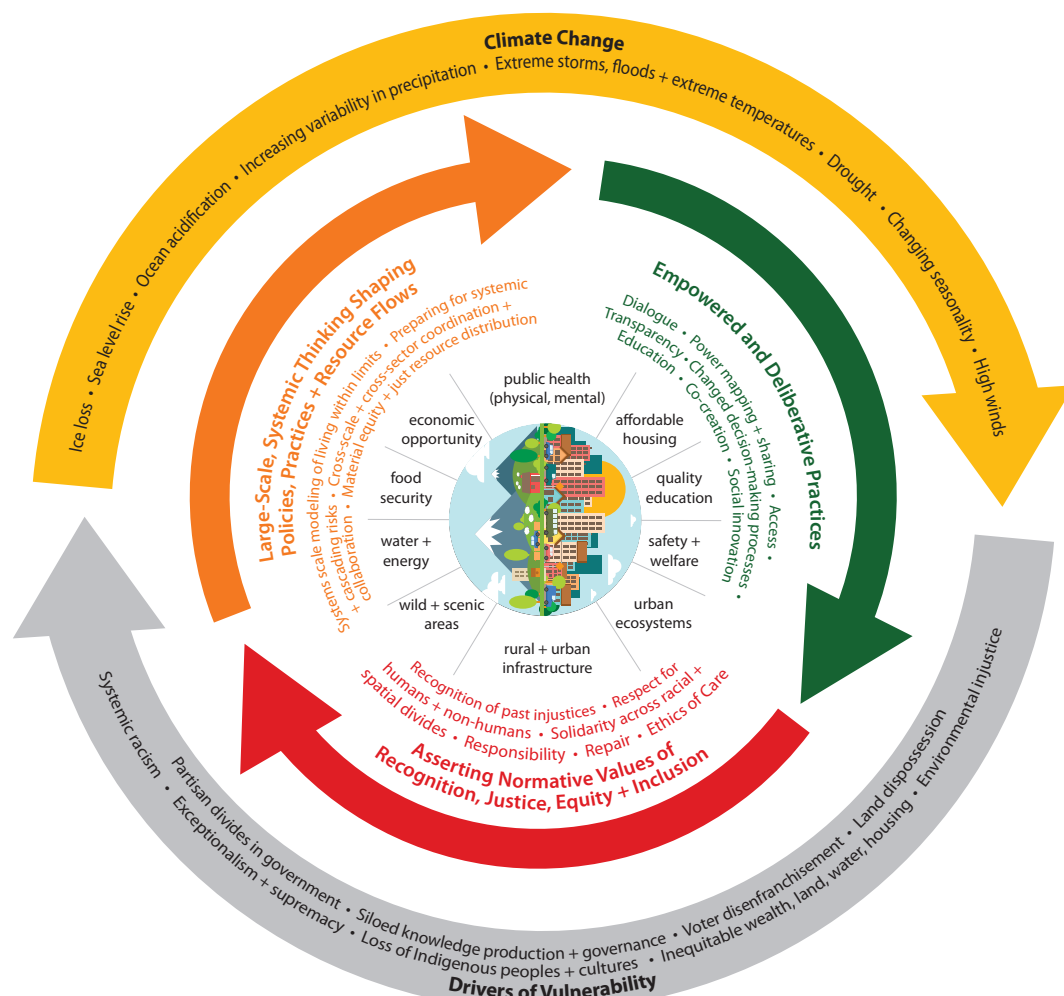
through learning, knowledge coproduction, and personal and collective change (*23*). Deliberative, positively transformative adaptation—often community-led, better capacitated, and politically empowered—would focus squarely on the procedural aspects and center the normative dimensions of adaptation.

Asserting normative values of equity, justice, and inclusion

Because large-scale, systemic change alone does not guarantee a desirable direction of change, scholars argue that transformative adaptation must normatively center justice in planning, infrastructure, and governance reforms (*24, 27*). There is growing recognition that adaptation efforts without such a normative commitment can be maladaptive and worsen inequality (*72, 111*). Building a seawall in one area may increase flooding downcoast and gentrify existing neighborhoods, offering some measure of protection to one enclave while transferring vulnerability to others (*72, 101, 112*). Climate or green gentrification is already occurring where higher land elevations, climate-smart infrastructure upgrades, or nature-based

solutions result in increased property values, taxes, or rents (*20, 72*). This can displace lower-income residents to more affordable places that confront new environmental risks (*20, 72, 113*). Adaptive interventions can create other spillover effects that worsen conditions elsewhere [for example, for low-income urban, suburban, or rural areas in the face of increasing competition over water, food, and land resources (*72*)]. Transformative adaptation reorients "climate actions around addressing entrenched equity and climate justice challenges. It [should focus] on systemic changes to development processes and [aim] at improving people's quality of life, enhancing the social and economic vibrancy of cities and ensuring sustainable, resilient, and inclusive urban futures" (*27*).

Science in support of transformative adaptation

Whether it is a new administration, a new Congress, relatively recent arrivals at the climate action table, or long-engaged implementers and observers of adaptation, critical questions must be asked at every turn regarding the material, relational, and normative dimensions of transformative adaptation: Who do

the changes in policies, practice, or resource flows benefit? Will a strategy lead to a more equitable and just society, repairing historical systemic disadvantaging? Does a project, program, or policy account not only for climate impacts but societal responses to these impacts in other sectors and geographies? Who is at the table? Whose claims and interests are heard, and which kinds of knowledge count?

Conflicting notions of adaptation and the appropriate paths forward underscore the persistent disconnect between academic research and policy and practice. Almost all government funding on climate change to date has been in the natural and physical sciences (*114*). Unsurprisingly, there is limited knowledge still about basic aspects of societal adaptation, including how to support contentious communication and engagement about transformative adaptation, how to integrate Indigenous or community-based knowledges with scientific knowledge and how this differs from integrating scientific knowledge with private-sector and business knowledge, how to navigate and change the power dynamics underlying the politics of decision-making, how to

measure adaptation progress and "success" (www.resiliencemetrics.org for a bottom-up example), and now, how to identify and track transformative adaptation and related social tipping points through meaningful indicators (*104*). The need to fund and support research on the intersections of coupled social-ecological-political systems, cascading impacts, and vicious and virtuous cycles also persists.

At the level of power and deliberative processes, grassroots movements have focused on finding and building solidarity across racial minorities, middle-class predominantly white environmentalists, and blue-collar workers. Some now also reach across the rural-urban divide (for example, Alliance for Appalachia, Farmworker Association of Florida, and Another Gulf is Possible). Community organizing strategies for solidarity and alliance-building offer powerful counterpoints to current partisan divisions at the federal and, increasingly, state levels (*115*). As researchers and pundits wring their hands about partisan divides and how to achieve aspirations for unity, these practices demonstrate an existing path toward solidarity, one that requires long years of sitting down, reaching across divides, and building a different politics (including knowledge politics) from the ground up. Some of the critical questions here include, under what conditions have large-scale reforms empowered democratic transformation? How can big-data science help the social sciences in an ethical way to synthesize robust policy-relevant knowledge from the rich case study research base available? For example, what are the necessary conditions that enable coalition building, social movements, and effective institutionalization of knowledge? Conversations have begun about the skills, competencies, and personal resilience of those working day in and day out on responding to climate change (*97*, *116–118*), but what canon of leadership capacities, including psychosocial skills, is necessary for transformative change? What historic experiences inform how to rapidly develop and grow this new generation of leadership (*117*, *119*)?

On mindsets and values, although sectors fundamentally disagree on the normative values and purposes underlying societal adaptation, both science (*120–122*) and grassroots organizations have mounted powerful normative proposals for human flourishing on a resource-limited Earth. However, how to advance the necessary mindset and values shifts and how to affect deep cultural change, especially in increasingly existentially threatening conditions, is far from clear. If a regenerative economy is what people support, but the extractive economy is what we have, how can governance systems be destabilized and transitioned at a national scale, not just at a community and site scale? What are the necessary conditions connecting these changes across

scales? How should existing spatial, fiscal, and financial systems be reformed to promote justice? What engagement platforms are needed to enable reconciliation between community and industry viewpoints? What kinds of learning, dialogue, and engagement foster transformative mindsets?

Last, the slow processes of learning, values changes, and trust-building that are implied in each of these areas of research ultimately raise a cross-cutting question for the scientific enterprise. A necessary fundamental change in the dynamics between researchers and practitioners, between white privileged and far-too-long marginalized communities (and their knowledge systems), points to the difficult-to-overcome tension between the urgency of climate impacts and proposed, large-scale solutions on the one hand and the time needed for deliberative processes, coalition, and trust building and the careful ethical considerations and reckoning with historical legacies required for transformation on the other (*123*). Although the growing literature on transdisciplinarity has established how to work across traditional disciplinary and sectoral silos, the actual practice of engaged research is still not the dominant mode of scientific practice (*124*). Moreover, the literature has not grappled with the question of how to do "slow" engaged science that remains relevant to decision-making amid accelerating environmental and social changes. This raises critical scientific, pragmatic, and institutional questions about how to accelerate the necessary scientific work, the decision- and policy-making processes, and the interaction between both.

Conclusion

Forces of globalization, urbanization, and climate change combined with the societal divisions that have always existed are bringing rapid and disorienting social and spatial changes. The persistent level of political stalemate in the United States reflects the scale of political, economic, social, and geographic dislocation in recent decades. Adaptation responses to date have largely focused on maintaining existing systems, which have contributed to inequitable and unsustainable development, without addressing underlying drivers of vulnerability to climate impacts. Continuing with modestly adapting business-as-usual practices enables current holders of wealth and power to reduce their risks but leaves most residents ill-equipped for the far more dramatic changes ahead. Any effort to challenge long-established systems in favor of transformative shifts almost inevitably encounters politics of opposition, division, and othering.

Transformative adaptation demands not only renewed executive branch commitment to climate action but poses profound challenges to deliberative democracy, collective action,

distributive justice, and the science to support it. More than ever, a deeper understanding of transformation, illustrative examples, and courageous leadership at all levels are necessary to change the pace, scale, and depth of climate adaptation and the drivers of vulnerability that would move society to more just adaptation. This Review points to how adaptation actors can stay at the table and constructively help to shape what's on the table for transformative adaptation. It remains possible for all relevant actors to creatively and constructively engage more deeply on adaptation. The opportunity for demonstrating effective adaptive leadership under rapidly changing and ever more difficult circumstances has maybe never been greater.

REFERENCES AND NOTES

1. D. Easterling, K. Kunkel, K. Lewis, D. Reidmiller, B. Stewart, Eds., *Impacts, Risks, and Adaptation in the United States: Fourth National Climate Assessment, Volume II: Report-in-Brief* (U.S. Global Change Research Program, 2018); doi: 10.7930/NCA4.2018.RiB
2. C. Phillips *et al.*, Compound climate risks in the COVID-19 pandemic. *Nat. Clim. Chang.* **10**, 586–588 (2020). doi: 10.1038/s41558-020-0804-2
3. S. C. Moser, J. A. F. Hart, The long arm of climate change: Societal teleconnections and the future of climate change impacts studies. *Clim. Change* **129**, 13–26 (2015). doi: 10.1007/s10584-015-1328-z; pmid: 32214560
4. M. Centeno, M. Nag, T. Patterson, A. Shaver, A. J. Windawi, The emergence of global systemic risk. *Annu. Rev. Sociol.* **41**, 65–85 (2015). doi: 10.1146/annurev-soc-073014-112317
5. J. Kania, M. Kramer, P. Senge, "The Water of Systems Change" (FSG, 2018); www.fsg.org/publications/water_of_systems_change.
6. M. Cabaj, L. Weaver, "Collective Impact 3.0: An Evolving Framework for Community Change," *Community Change Series* (The Tamarack Institute, 2016), p. 14.
7. J. Fischer, M. Riechers, A leverage points perspective on sustainability. *People Nat.* **1**, 115–120 (2019). doi: 10.1002/pan3.13
8. A. Smith, *2010-2019: A Landmark Decade of U.S. Billion-Dollar Weather and Climate Disasters* (National Oceanographic and Atmospheric Administration, 2020); www.climate.gov/news-features/blogs/beyond-data/2010-2019-landmark-decade-us-billion-dollar-weather-and-climate.
9. A.S.C.E., *Infrastructure Report Card: A Comprehensive Assessment of America's Infrastructure* (American Society of Civil Engineers, 2017); www.infrastructurereportcard.org.
10. C.S.I.W.G., "Paying It Forward: The Path Toward Climate-Safe Infrastructure in California" (A report by the Climate-Safe Infrastructure Working Group to the California State Legislature and Strategic Growth Council, 2018).
11. L. Shi, A. M. Varuzzo, Surging seas, rising fiscal stress: Exploring municipal fiscal vulnerability to climate change. *Cities* **100**, 102658 (2020). doi: 10.1016/j.cities.2020.102658
12. I. Boas *et al.*, Climate migration myths. *Nat. Clim. Chang.* **9**, 901–903 (2019). doi: 10.1038/s41558-019-0633-3
13. C. Robinson, B. Dilkina, J. Moreno-Cruz, Modeling migration patterns in the USA under sea level rise. *PLOS ONE* **15**, e0227436 (2020). doi: 10.1371/journal.pone.0227436; pmid: 31968017
14. M. Hauer, Migration induced by sea-level rise could reshape the US population landscape. *Nat. Clim. Chang.* **7**, 321–325 (2017). doi: 10.1038/nclimate3271
15. K. Curtis, J. DeWaard, E. Fussell, R. Rosenfeld, Differential recovery migration across the rural–urban gradient: Minimal and short-term population gains for rural disaster-affected Gulf Coast counties. *Rural Sociol.* **85**, 856–898 (2019). doi: 10.1111/ruso.12305
16. A. Cavallo, V. Ireland, Preparing for complex interdependent risks: A System of Systems approach to building disaster resilience. *Int. J. Disaster Risk Reduct.* **9**, 181–193 (2014). doi: 10.1016/j.ijdrr.2014.05.001
17. A. R. Siders, Adaptive capacity to climate change: A synthesis of concepts, methods, and findings in a fragmented field. *Wiley Interdiscip. Rev. Clim. Change* **10**, e573 (2019). doi: 10.1002/wcc.573

18. J. E. Cinner et al., Building adaptive capacity to climate change in tropical coastal communities. Nat. Clim. Chang. 8, 117–123 (2018). doi: 10.1038/s41558-017-0065-x

19. B. Pfefferbaum, C. S. North, Mental health and the Covid-19 pandemic. N. Engl. J. Med. 383, 510–512 (2020). doi: 10.1056/NEJMp2008017; pmid: 32283003

20. K. T. Aune, D. Gesch, G. S. Smith, A spatial analysis of climate gentrification in Orleans Parish, Louisiana post-Hurricane Katrina. Environ. Res. 185, 109384 (2020). doi: 10.1016/j.envres.2020.109384; pmid: 32240840

21. K. O'Brien, Global environmental change II: From adaptation to deliberate transformation. Prog. Hum. Geogr. 36, 667–676 (2012). doi: 10.1177/0309132511425767

22. J. Patterson et al., Exploring the governance and politics of transformations towards sustainability. Environ. Innov. Soc. Transit. 24, 1–16 (2017). doi: 10.1016/j.eist.2016.09.001

23. I. Scoones et al., Transformations to sustainability: Combining structural, systemic and enabling approaches. Curr. Opin. Environ. Sustain. 42, 65–75 (2020). doi: 10.1016/j.cosust.2019.12.004

24. K. P. Henrique, P. Tschakert, Pathways to urban transformation: From dispossession to climate justice. Prog. Hum. Geogr. 10.1177/0309132520962856 (2020). doi: 10.1177/0309132520962856

25. M. Pelling, Adaptation to Climate Change: From Resilience to Transformation (Routledge, 2010).

26. K. L. O'Brien, S. Eriksen, A. Schjolden, L. Nygaard, What's in a word? Conflicting interpretations of vulnerability in climate change research (2009); http://dspace.cigilibrary.org/jspui/handle/123456789/7304.

27. E. Chu et al., "Unlocking the Potential for Transformative Climate Adaptation in Cities," Background Paper prepared for the Global Commission on Adaptation (World Resources Institute, 2019); www.gca.org.

28. J. Rodin, The Resilience Dividend: Being Strong in a World Where Things Go Wrong (Public Affairs, 2014).

29. N. Klein, The Shock Doctrine: The Rise of Disaster Capitalism (Alfred A. Knopf Canada, 2007).

30. Wretched of the Earth, An open letter to Extinction Rebellion. Red Pepper (2019); www.redpepper.org.uk/an-open-letter-to-extinction-rebellion.

31. S. C. Moser, P. Aldunce, A. Rudnick, M. Rojas, L. Muñoz, "Transformation from science to decision-making," Policy Brief to COP25, written with support of the COP25 Scientific Committee and the Ministry of Science, Technology, Knowledge and Innovation of Chile, the Economic Commission for Latin America and the Caribbean, the European Union, the Inter-American Development Bank, and the International Science Council (2019); https://t2sresearch.org/2019/12/19/transformation-from-science-to-decision-making.

32. C. L. Berzonsky, S. C. Moser, Becoming homo sapiens sapiens: Mapping the psycho-cultural transformation in the anthropocene. Anthropocene 20, 15–23 (2017). doi: 10.1016/j.ancene.2017.11.002

33. C. Ratcliffe, W. J. Congdon, A. Stanczyk, D. Teles, C. Martín, B. Kotapati, Insult to Injury: Natural Disasters and Residents (Urban Insititute, 2019).

34. K. J. Mach et al., Managed retreat through voluntary buyouts of flood-prone properties. Sci. Adv. 5, eaax8995 (2019). doi: 10.1126/sciadv.aax8995; pmid: 31633030

35. C. N. A. Corporation, "National Security and the Threat of Climate Change" (CNA Corporation, 2007).

36. GAO, in Report to the Chairman, Select Committee on Energy Independence and Global Warming, House of Representatives (GAO, 2009); www.gao.gov/assets/gao-10-113.pdf.

37. V. Kalesnikaite, Keeping cities afloat: Climate change adaptation and collaborative governance at the local level. Public Perform. Manag. Rev. 42, 864–888 (2019). doi: 10.1080/15309576.2018.1526091

38. A. Schulz, A. Zia, C. Koliba, Adapting bridge infrastructure to climate change: Institutionalizing resilience in intergovernmental transportation planning processes in the Northeastern USA. Mitig. Adapt. Strategies Glob. Change 22, 175–198 (2017). doi: 10.1007/s11027-015-9672-x

39. Obama Task Force, President's State, Local, and Tribal Leaders Task Force on Climate Preparedness and Response: Recommendations to the President (White House, 2014); https://obamawhitehouse.archives.gov/sites/default/files/docs/task_force_report_0.pdf.

40. J. M. Keenan, From climate change to national security: Analysis of the Obama administration's federal resilience mandates and measures. Nat. Hazards Rev. 19, 04017022 (2018). doi: 10.1061/(ASCE)NH.1527-6996.0000273

41. G.A.O., in Statement of J. Alfredo Gómez, Director, Natural Resources and Environment before the Subcommittee on Environment, Committee on Oversight and Reform, House of Representatives (GAO, 2019); www.gao.gov/products/gao-20-338t.

42. US Department of Homeland Security (DHS), National Preparedness Report (US Department of Homeland Security, 2019).

43. C. Flavelle, Trump administration presses cities to evict homeowners from flood zones, The New York Times 11 March 2020).

44. D. Finn, D. Chandrasekhar, Y. Xiao, A region recovers: Planning for resilience after superstorm Sandy. J. Plann. Educ. Res. 0739456X1986414 (2019). doi: 10.1177/0739456X19864145

45. A. Gustafson et al., The development of partisan polarization over the Green New Deal. Nat. Clim. Chang. 9, 940–944 (2019). doi: 10.1038/s41558-019-0621-7

46. R. Gunn-Wright, R. Hockett, "The Green New Deal," Legal Studies Research Paper Series, Cornell Law School research paper 19–09,(New Consensus and Cornell University Law School, 2019).

47. J. W. Straka, B. C. Straka, Reframe policymaking dysfunction through bipartisan-inclusion leadership. Policy Sci. 53, 1–24 (2020). doi: 10.1007/s11077-020-09383-2; pmid: 32351255

48. C. Funk, M. Hefferon, "U.S. Public Views on Climate and Energy" (Pew Research Center Science & Society, Washington, D.C., 2019; www.pewresearch.org/science/2019/11/25/u-s-public-views-on-climate-and-energy.

49. S. Andrews, A. Selman, in Excerpted from the Proceedings of the 58th Annual Meeting of Invited Attorneys (Victor O. Schinnerer & Company, 2019).

50. E. Kaplan-Leiserson, High winds of change. PE Magazine (2019); www.nspe.org/resources/pe-magazine/may-2019/high-winds-change.

51. C.L.F., G.R.C., Climate Adaptation and Liability: A Legal Primer and Workshop Summary Report (Conservation Law Foundation and Boston Green Ribbon Commission, 2018); www.clf.org/wp-content/uploads/2018/01/GRC_CLF_Report_R8.pdf.

52. Q. Miao, Y. Hou, M. Abrigo, Measuring the financial shocks of natural disasters: A panel study of U.S. states. Natl. Tax J. 71, 11–44 (2018). doi: 10.17310/ntj.2018.1.01

53. Z. A. Conyers, R. Grant, S. Roy, Sea level rise in Miami Beach: Vulnerability and real estate exposure. Prof. Geogr. 71, 278–291 (2019). doi: 10.1080/00330124.2018.1531037

54. A. Bernstein, M. Gustafson, R. Lewis, Disaster on the horizon: The price effect of sea level Rise. J. Financ. Econ. 134, 253–272 (2019). doi: 10.1016/j.jfineco.2019.03.013

55. S. A. McAlpine, J. R. Porter, Estimating recent local impacts of sea-level rise on current real-estate losses: A housing market case study in Miami-Dade, Florida. Popul. Res. Policy Rev. 37, 871–895 (2018). doi: 10.1007/s11113-018-9473-5; pmid: 30546178

56. P. Walsh, C. Griffiths, D. Guignet, H. Klemick, Adaptation, sea level rise, and property prices in the Chesapeake Bay watershed. Land Econ. 95, 19–34 (2019). doi: 10.3368/le.95.1.19; pmid: 30799882

57. C. Kousky, Financing Flood Losses: A Discussion of the National Flood Insurance Program (Resources for the Future, 2017).

58. UCS, "Underwater: Rising Seas, Chronic Floods, and the Implications for US Coastal Real Estate" (Union of Concerned Scientists, 2018).

59. C. Flavelle, D. Lu, V. Penney, N. Popovich, J. Schwartz, New Data Reveals Hidden Flood Risk Across America, The New York Times 29 June 2020.

60. J. M. Keenan, A climate intelligence arms race in financial markets. Science 365, 1240–1243 (2019). doi: 10.1126/science.aay8442; pmid: 31604223

61. N.G.F.S., A Call for Action: Climate Change as a Source for Financial Risk (Network for Greening the Financial System Secretariat/Banque de France, 2019).

62. B. Grayson, Moody's: Climate change adaptation and mitigation could affect cities' bond ratings. Urban Land (2018); https://urbanland.uli.org/news/moodys-climate-change-adaption-mitigation-affect-cities-bond-ratings.

63. Standard's &Poor, in Global Ratings, Poor, Ed. (Standard, 2019); www.spratings.com/en_US/products/-/product-detail/our-approach-toesg-in-ratings.

64. World Federation of Engineering Organizations (WFEO), Model Code of Practice: Principles of Climate Change Adaptation for Engineers (WFEO, 2015); www.wfeo.org/wp-content/uploads/code-of-practice/WFEO_Model_Code_of_Practice_Principles_Climate_Change_Adaptation_Engineers.pdf.

65. B. Ayyub, Ed., in Committee on Adaptation to a Changing Climate (American Society of Civil Engineers, 2018); https://ascelibrary.org/doi/book/10.1061/9780784415191.

66. P. Bigger, N. Millington, Getting soaked? Climate crisis, adaptation finance, and racialized austerity. Environ. Planning E Nat. Space 3, 601–623 (2019). doi: 10.1177/2514848619876539

67. J. M. Keenan, J. T. Bradt, Underwaterwriting: From theory to empiricism in regional mortgage markets in the U.S. Clim. Change 162, 2043–2067 (2020). doi: 10.1007/s10584-020-02734-1

68. Z. J. Taylor, J. L. Weinkle, The riskscapes of re/insurance. Camb. J. Regions Econ. Soc. 13, 405–422 (2020). doi: 10.1093/cjres/rsaa015

69. S. Woodruff, T. BenDor, A. Strong, Fighting the inevitable: Infrastructure investment and coastal community adaptation to sea level rise. Syst. Dyn. Rev. 34, 48–77 (2018). doi: 10.1002/sdr.1597

70. M. Garcia-Lamarca et al., Urban green boosterism and city affordability: For whom is the 'branded' green city? Urban Stud. 58, 90–112 (2019). doi: 10.1177/0042098019885330

71. V. Eubanks, My drowning city is a harbinger of climate slums to come. The Nation (2016); www.thenation.com/article/low-water-mark.

72. I. Anguelovski et al., Equity impacts of urban land use planning for climate adaptation: Critical perspectives from the global north and south. J. Plann. Educ. Res. 36, 333–348 (2016). doi: 10.1177/0739456X16645166

73. G. Shokry, J. J. Connolly, I. Anguelovski, Understanding climate gentrification and shifting landscapes of protection and vulnerability in green resilient Philadelphia. Urban Climate 31, 100539 (2020). doi: 10.1016/j.uclim.2019.100539

74. E. Nance, "Flood buyout inequity in greater Houston," presented at the Association of Collegiate Schools of Planning Conference, Toronto, 5–8 November 2020.

75. J. M. Keenan, T. Hill, A. Gumber, Climate gentrification: From theory to empiricism in Miami-Dade County, Florida. Environ. Res. Lett. 13, 054001 (2018). doi: 10.1088/1748-9326/aabb32

76. J. K. Maldonado, B. Colombi, R. Pandya, Eds., Climate Change and Indigenous Peoples in the United States: Impacts, Experiences and Actions (Springer International Publishing, 2014); http://link.springer.com/10.1007/978-3-319-05266-3.

77. R. Bronen, F. S. Chapin 3rd, Adaptive governance and institutional strategies for climate-induced community relocations in Alaska. Proc. Natl. Acad. Sci. U.S.A. 110, 9320–9325 (2013). doi: 10.1073/pnas.1210508110; pmid: 23690592

78. L. Shi, From progressive cities to resilient cities: Lessons from history for new debates in equitable adaptation to climate change. Urban Aff. Rev. 10.1177/1078087419910827 (2020). doi: 10.1177/1078087419910827

79. A. K. Magnan et al., Addressing the risk of maladaptation to climate change. Wiley Interdiscip. Rev. Clim. Change 7, 646–665 (2016). doi: 10.1002/wcc.409

80. E. Marino, Adaptation privilege and voluntary buyouts: Perspectives on ethnocentrism in sea level rise relocation and retreat policies in the US. Glob. Environ. Change 49, 10–13 (2018). doi: 10.1016/j.gloenvcha.2018.01.002

81. E. Torabi, A. Dedekorkut-Howes, M. Howes, Adapting or maladapting: Building resilience to climate-related disasters in coastal cities. Cities 72, 295–309 (2018). doi: 10.1016/j.cities.2017.09.008

82. V. C. Broto, E. Robin, A. While, Eds., Climate Urbanism: Towards a Critical Research Agenda (Palgrave Macmillan, 2020); www.palgrave.com/gp/book/9783030533854.

83. B. P. Warner, C. P. Kuzdas, The role of political economy in framing and producing transformative adaptation. Curr. Opin. Environ. Sustain. 29, 69–74 (2017). doi: 10.1016/j.cosust.2017.12.012

84. F.-A. Hoover, T. C. Lim, Examining privilege and power in US urban parks and open space during the double crises of antiblack racism and COVID-19. Socio. Ecol. Pract. Res. 10.1007/s42532-020-00070-3 (2020).

85. K. P. Whyte, J. P. Brewer II, J. T. Johnson, Weaving Indigenous science, protocols and sustainability science. Sustain. Sci. 11, 25–32 (2016). doi: 10.1007/s11625-015-0296-6

86. NAACP, "Equity in building resilience in adaptation planning" (National Association for the Advancement of Colored People, 2015).

87. APEN, "Mapping resilience: A blueprint for thriving in the face of climate disasters" (Asian Pacific Environmental Network, 2019).

88. N. Fraser, Recognition or redistribution? A critical reading of Iris Young's Justice and the Politics of Difference. *J. Polit. Philos.* **3**, 166–180 (1995). doi: 10.1111/j.1467-9760.1995.tb00033.x

89. S. Woodruff, M. Stults, Numerous strategies but limited implementation guidance in US local adaptation plans. *Nat. Clim. Chang.* **6**, 796–802 (2016). doi: 10.1038/nclimate3012

90. E. D. Mondainé, M. Lee, Beyond Theory: Climate Justice in Practice. *Buildings and Cities* (2020); www.buildingsandcities.org/insights/commentaries/beyond-theory-climate-justice-practice.html.

91. Climate Justice Alliance; https://climatejusticealliance.org.

92. A. Kothari, F. Demaria, A. Acosta, Buen Vivir, Degrowth and ecological Swaraj: Alternatives to sustainable development and the green economy. *Development* **57**, 362–375 (2014). doi: 10.1057/dev.2015.24

93. M. Ranganathan, E. Bratman, *Antipode* 10.1111/anti.12555 (2019).

94. S. C. Moser, J. Coffee, A. Seville, "Rising to the Challenge, Together: A Review and Critical Assessment of the State of the US Climate Adaptation Field" (Kresge Foundation, 2017).

95. R. W. Kates, W. R. Travis, T. J. Wilbanks, Transformational adaptation when incremental adaptations to climate change are insufficient. *Proc. Natl. Acad. Sci. U.S.A.* **109**, 7156–7161 (2012). doi: 10.1073/pnas.1115521109; pmid: 22509036

96. V. Gallina *et al.*, A review of multi-risk methodologies for natural hazards: Consequences and challenges for a climate change impact assessment. *J. Environ. Manage.* **168**, 123–132 (2016). doi: 10.1016/j.jenvman.2015.11.011; pmid: 26704454

97. L. Shi, Promise and paradox of metropolitan regional climate adaptation. *Environ. Sci. Policy* **92**, 262–274 (2019). doi: 10.1016/j.envsci.2018.11.002

98. J. Zscheischler *et al.*, Future climate risk from compound events. *Nat. Clim. Chang.* **8**, 469–477 (2018). doi: 10.1038/s41558-018-0156-3

99. H. R. Moftakhari, G. Salvadori, A. AghaKouchak, B. F. Sanders, R. A. Matthew, Compounding effects of sea level rise and fluvial flooding. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 9785–9790 (2017). doi: 10.1073/pnas.1620325114; pmid: 28847932

100. FEMA, "Building Community Resilience with Nature-Based Solutions: A Guide for Local Communities" (Federal Emergency Management Agency, 2020); www.fema.gov/sites/default/files/2020-07/fema_bric_nature-based-solutions-guide_2020.pdf.

101. J. J. Opperman, "A Flood of Benefits: Using Green Infrastructure to Reduce Flood Risks" (The Nature Conservancy, 2014); www.conservationgateway.org/ConservationPractices/Freshwater/HabitatProtectionandRestoration/Documents/A%20Flood%20of%20Benefits%20-%20J.Opperman%20-%20May%202014.pdf.

102. L. Shi, Beyond flood risk reduction: How can green infrastructure advance both social justice and regional impact? *Socio. Ecol. Pract. Res.* **2**, 311–320 (2020). doi: 10.1007/s42532-020-00065-0

103. J. L. MacArthur, C. E. Hoicka, H. Castleden, R. Das, J. Lieu, Canada's Green New Deal: Forging the socio-political foundations of climate resilient infrastructure? *Energy Res. Soc. Sci.* **65**, 101442 (2020). doi: 10.1016/j.erss.2020.101442

104. R. Galvin, N. Healy, The Green New Deal in the United States: What it is and how to pay for it. *Energy Res. Soc. Sci.* **67**, 101529 (2020). doi: 10.1016/j.erss.2020.101529

105. P. Tschakert, B. van Oort, A. L. St. Clair, A. LaMadrid, Inequality and transformation analyses: A complementary lens for addressing vulnerability to climate change. *Clim. Dev.* **5**, 340–350 (2013). doi: 10.1080/17565529.2013.828583

106. A. R. Siders, Social justice implications of US managed retreat buyout programs. *Clim. Change* **152**, 239–257 (2019). doi: 10.1007/s10584-018-2272-5

107. G. Owen, What makes climate change adaptation effective? A systematic review of the literature. *Glob. Environ. Change* **62**, 102071 (2020). doi: 10.1016/j.gloenvcha.2020.102071

108. J. T. Malloy, C. M. Ashcraft, A framework for implementing socially just climate adaptation. *Clim. Change* **160**, 1–14 (2020). doi: 10.1007/s10584-020-02705-6

109. N. Suckall, E. L. Tompkins, K. Vincent, A framework to analyse the implications of coastal transformation on inclusive development. *Environ. Sci. Policy* **96**, 64–69 (2019). doi: 10.1016/j.envsci.2019.03.003

110. J. Klein *et al.*, The role of the private sector and citizens in urban climate change adaptation: Evidence from a global assessment of large cities. *Glob. Environ. Change* **53**, 127–136 (2018). doi: 10.1016/j.gloenvcha.2018.09.012

111. S. Juhola, E. Glaas, B.-O. Linnér, T.-S. Neset, Redefining maladaptation. *Environ. Sci. Policy* **55**, 135–140 (2016). doi: 10.1016/j.envsci.2015.09.014

112. G. Zuccaro, D. De Gregorio, M. Leone, Theoretical model for cascading effects analyses. *Int. J. Disaster Risk Reduct.* **30**, 199–215 (2018). doi: 10.1016/j.ijdrr.2018.04.019

113. I. Anguelovski *et al.*, Opinion: Why green "climate gentrification" threatens poor and vulnerable populations. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 26139–26143 (2019). doi: 10.1073/pnas.1920490117

114. I. Overland, B. Sovacool, The misallocation of climate research funding. *Energy Res. Soc. Sci.* **62**, 101349 (2020). doi: 10.1016/j.erss.2019.101349

115. R. Bromley-Trujillo, M. R. Holman, Climate change policymaking in the states: A view at 2020. *Publius* **50**, 446–472 (2020). doi: 10.1093/publius/pjaa008

116. R. E. Boyatzis, K. Thiel, K. Rochford, A. Black, Emotional and social intelligence competencies of incident team commanders fighting wildfires. *J. Appl. Behav. Sci.* **53**, 498–516 (2017). doi: 10.1177/0021886317731575

117. D. Gilford, S. C. Moser, B. DePodwin, R. Moulton, S. Watson, The emotional toll of climate change on science professionals. *Eos* **100**, 6 (2019). doi: 10.1029/2019EO137460

118. D. Pruneau *et al.*, Competencies demonstrated by municipal employees during adaptation to climate change: A pilot study. *J. Environ. Educ.* **44**, 217–231 (2013). doi: 10.1080/00958964.2012.759521

119. P. Cianconi, S. Betrò, L. Janiri, The impact of climate change on mental health: A systematic descriptive review. *Front. Psychiatry* **11**, 74 (2020). doi: 10.3389/fpsyt.2020.00074; pmid: 32210846

120. K. Raworth, A doughnut for the Anthropocene: Humanity's compass in the 21st century. *Lancet Planet. Health* **1**, e48–e49 (2017). doi: 10.1016/S2542-5196(17)30028-1; pmid: 29851576

121. G. Kallis *et al.*, Research on degrowth. *Annu. Rev. Environ. Resour.* **43**, 291–316 (2018). doi: 10.1146/annurev-environ-102017-025941

122. I. M. Algunaibet *et al.*, Powering sustainable development within planetary boundaries. *Energy Environ. Sci.* **12**, 1890–1900 (2019). doi: 10.1039/C8EE03423K; pmid: 31303892

123. S. C. Moser, Science for the speed-up: Quo vadis science-society interactions in a climate changed world? (AGU Annual Meeting, 2019).

124. I. Fazey *et al.*, Transforming knowledge systems for life on Earth: Visions of future systems and how to get there. *Energy Res. Soc. Sci.* **70**, 101724 (2020). doi: 10.1016/j.erss.2020.101724

CELL BIOLOGY

Ubiquitination is essential for recovery of cellular activities after heat shock

Brian A. Maxwell, Youngdae Gwon, Ashutosh Mishra, Junmin Peng, Haruko Nakamura, Ke Zhang, Hong Joo Kim, J. Paul Taylor*

INTRODUCTION: In response to many types of stress, eukaryotic cells initiate an adaptive and reversible response that includes down-regulation of key cellular activities along with sequestration of cytoplasmic mRNAs into structures called stress granules. Accompanying these stress responses is a global increase in ubiquitination that has been conventionally ascribed to the need for degradation of misfolded or damaged proteins. However, detailed characterization of how the ubiquitinome is reshaped in response to stress is lacking. Furthermore, it is unclear whether stress-dependent ubiquitination plays a more complex role in the larger stress response beyond its known protective function in targeting hazardous proteins for proteasomal degradation.

RATIONALE: To explore the role of ubiquitination in the stress response, we used tandem ubiquitin binding entity (TUBE) proteomics to investigate changes to the ubiquitination landscape in response to five different types of stress in cultured mammalian cells, including human induced pluripotent stem cell (iPSC)–derived neurons. The discovery of unanticipated patterns of ubiquitination prompted a detailed analysis of the ubiquitination pattern specifically induced by heat shock by using diGly ubiquitin remnant profiling along with tandem mass tag quantitative proteomics in combination with additional total proteome and transcriptome analyses. Insights from this newly defined "heat shock ubiquitinome" guided subsequent investigation of the functional importance of this posttranslational modification in the cellular response to heat shock.

RESULTS: Each of the five different types of stress induced a distinctive pattern of ubiquitination. The heat shock ubiquitinome in human embryonic kidney 293T cells was defined by ubiquitination of specific proteins that function within cellular activities that are down-regulated during stress (e.g., translation and nucleocytoplasmic transport), and this pattern was similar in U2OS cells, primary mouse neurons, and human iPSC-derived neurons. The heat shock ubiquitinome was also enriched in protein constituents of stress granules. Suprisingly, this stress-induced ubiquitination was dispensable for the formation of stress granules and shutdown of cellular pathways; rather, heat shock–induced ubiquitination was a prerequisite for p97/valosin-containing protein (VCP)–mediated stress granule disassembly and for resumption of normal cellular activities, including nucleocytoplasmic transport and translation, upon recovery from stress. Many ubiquitination events were specific to one or another stress. For example, ubiquitination was required for disassembly of stress granules induced by heat stress but dispensable for disassembly for stress granules induced by oxidative (arsenite) stress.

CONCLUSION: Ubiquitination patterns are specific to different types of stress and indicate additional regulatory functions for stress-induced ubiquitination beyond the removal of misfolded or damaged proteins. Specifically, heat shock–induced ubiquitination primes the cell for recovery from stress by targeting specific proteins involved several pathways down-regulated during stress. Furthermore, some key stress granule constituents are ubiquitinated in response to heat stress but not arsenite stress, thus engaging a mechanism of VCP mediated–disassembly of heat shock–induced granules that is not shared by arsenite stress–induced granules. Finally, our deep proteomics datasets provide a rich community resource illuminating additional aspects of the roles of ubiquitination in response to stress. ∎

S READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abc3593

Heat stress–specific ubiquitination primes cells for recovery. (**A**) Proteomics-based ubiquitinome profiling reveals that different stresses induce distinct patterns of ubiquitin conjugation. TEV, tobacco etch virus protease cleavage site; Ub, ubiquitin; UBA, ubiquitin-associated domain; UV, ultraviolet. (**B**) Heat stress–induced ubiquitination targets proteins associated with cellular activities down-regulated during stress, including nucleocytoplasmic (NC) transport and translation, as well as stress granule constituents. This ubiquitination is required for the timely resumption of biological activity and stress granule disassembly after the removal of stress. mRNP, messenger ribonucleoprotein.

A Distinct stress-induced ubiquitinomes

TUBE pulldown — Ubiquitinome — Heat shock / Oxidative stress / Osmotic stress / UV stress / Proteasome inhibition

Heat stress–induced ubiquitination
- mRNA splicing
- Protein and RNA transport
- Cell cycle
- Metabolism
- Cytoskeleton
- DNA damage and repair
- Stress granule factors
- Protein biosynthesis and translation

B Strategic ubiquitination primes cells for recovery

Normal cellular activity — Heat shock–dependent shutdown of activity coincident with ubiquitination — Ubiquitination-dependent recovery — Normal cellular activity

CELL BIOLOGY

Ubiquitination is essential for recovery of cellular activities after heat shock

Brian A. Maxwell[1], Youngdae Gwon[1], Ashutosh Mishra[2], Junmin Peng[2], Haruko Nakamura[1], Ke Zhang[3], Hong Joo Kim[1], J. Paul Taylor[1,4]*

Eukaryotic cells respond to stress through adaptive programs that include reversible shutdown of key cellular processes, the formation of stress granules, and a global increase in ubiquitination. The primary function of this ubiquitination is thought to be for tagging damaged or misfolded proteins for degradation. Here, working in mammalian cultured cells, we found that different stresses elicited distinct ubiquitination patterns. For heat stress, ubiquitination targeted specific proteins associated with cellular activities that are down-regulated during stress, including nucleocytoplasmic transport and translation, as well as stress granule constituents. Ubiquitination was not required for the shutdown of these processes or for stress granule formation but was essential for the resumption of cellular activities and for stress granule disassembly. Thus, stress-induced ubiquitination primes the cell for recovery after heat stress.

ukaryotic cells have a stereotypical response to a variety of stresses that aids in their survival until normal growth conditions are restored (*1*, *2*). This stress response includes the inhibition of global translation along with the up-regulation of expression of select stress response factors (*1–7*). Other biological pathways are also shut down or perturbed, including nucleocytoplasmic transport (*8–13*), RNA splicing (*14*, *15*), and cell cycle activities (*16–18*). Inhibition of translation and subsequent polysome disassembly lead to a rise in the cytoplasmic concentration of ribosome-free mRNA that culminates in the formation of cytoplasmic condensates known as stress granules (*19*). These stress-induced cellular changes are adaptive in the short term but require coordinated reversal after stress is removed in order to resume cellular activities and reestablish homeostasis. Accordingly, upon the removal of stress, stress granules disassemble while translation and other biological pathways return to normal activity on a time scale of minutes to a few hours (*2*, *20–22*). However, the mechanisms that facilitate this recovery are not well understood.

One hallmark of the cellular response to many types of stress is a global increase in polyubiquitin conjugation, which has largely been attributed to increased protein quality control (PQC) activity in response to stress-induced protein damage and translation arrest (*1–3*). This increased activity is required for the ubiquitin-dependent degradation of both defective ribosomal products (DRiPs), the primary source of misfolded proteins in cells under normal conditions (*4*), and proteins susceptible to stress-induced misfolding, such as thermolabile proteins or proteins with intrinsically disordered domains (*2*, *5–7*). Ubiquitin and some ubiquitin pathway proteins are found in stress granules (*23–28*), and several deubiquitinating enzymes can help to regulate stress granule dynamics (*23*, *27*). Furthermore, when PQC function is disrupted, ubiquitinated DRiPs can accumulate in stress granules and impair stress granule function (*29–31*).

The ubiquitin-selective segregase p97/valosin-containing protein (VCP) facilitates stress granule removal (*29*, *32*, *33*), which may suggest a role for ubiquitination in this process. However, recent studies investigating ubiquitin conjugation have produced conflicting results. Whereas one study found that nuclear small ubiquitin-like modifier (SUMO)–primed ubiquitination was important for stress granule disassembly (*34*), another group reported minimal ubiquitination of stress granule proteins and showed that ubiquitin conjugation was entirely dispensable for stress granule dynamics (*35*). Thus, the role of ubiquitination in regulating stress granule dynamics remains unresolved, as does any potential role of ubiquitination in regulating other cellular activities that are altered by stress.

Analysis of stress-dependent changes to the global protein ubiquitin modification landscape, or "ubiquitinome," can provide unbiased insights into the stress response by connecting differentially ubiquitinated proteins with specific biological pathways. To this end, several proteomic techniques allow for quantitative comparisons of ubiquitinomes from differentially treated cell or tissue samples (*36–39*). Analysis of the ubiquitinome in yeast has suggested that heat shock–induced ubiquitination primarily targets cytosolic misfolded proteins (*40*, *41*), consistent with the presumption that ubiquitination in this setting relates mainly to PQC. However, comprehensive analyses in human cells or directly comparing changes in ubiquitination in response to various stresses have not been undertaken, and it remains unclear to what extent ubiquitination patterns may differ in these different contexts.

Results

Different cellular stresses induce distinct patterns of ubiquitination

To investigate whether cells deploy different patterns of ubiquitination to cope with different types of stress, we performed proteomic analysis after tandem ubiquitin binding entity (TUBE)–based capture of the ubiquitinome in human embryonic kidney (HEK) 293T cells exposed to heat stress (42°C), oxidative stress (sodium arsenite), osmotic stress (sorbitol), ultraviolet (UV) stress, or proteasomal inhibition (bortezomib) as well as unstressed cells (Fig. 1, A and B). After exposure to each treatment, lysates were incubated with Halo-linked TUBE resin to capture ubiquitin conjugates, followed by label-free liquid chromatography with tandem mass spectrometry (LC-MS/MS) (Fig. 1, A and B). Immunoblotting indicated that all stress types except UV led to an increase in total ubiquitin conjugates (input blot in Fig. 1C). Although the TUBEs likely have higher baseline affinity for binding polyubiquitinated over monoubiquitinated proteins, conditions were optimized such that the TUBEs captured nearly all ubiquitinated material (compare bound and unbound blots in Fig. 1C). We used stringent buffer conditions (1% NP-40) and a high salt wash to minimize capture of nonubiquitinated proteins by the TUBEs. We detected ~200 to 300 ubiquitinated proteins per stress condition and 500 unique proteins across all five stress conditions (Fig. 1D and table S1).

Unsupervised hierarchical clustering revealed that each type of stress was associated with a distinct pattern of ubiquitination relative to control cells (Fig. 1E). Indeed, most of the proteins with increased or decreased ubiquitination were associated with a single type of stress (Fig. 1F). Several patterns emerged from this analysis. Some proteins showed increased or decreased ubiquitination in response to all five stresses, suggesting that certain ubiquitination events were part of a nonspecific, generalized stress response (e.g., groups a and b in Fig. 1E).

[1]Department of Cell and Molecular Biology, St. Jude Children's Research Hospital, Memphis, TN 38105, USA. [2]Department of Structural Biology, St. Jude Children's Research Hospital, Memphis, TN 38105, USA. [3]Department of Neuroscience, Mayo Clinic Florida, Jacksonville, FL 32224, USA. [4]Howard Hughes Medical Institute, St. Jude Children's Research Hospital, Memphis, TN 38105, USA.
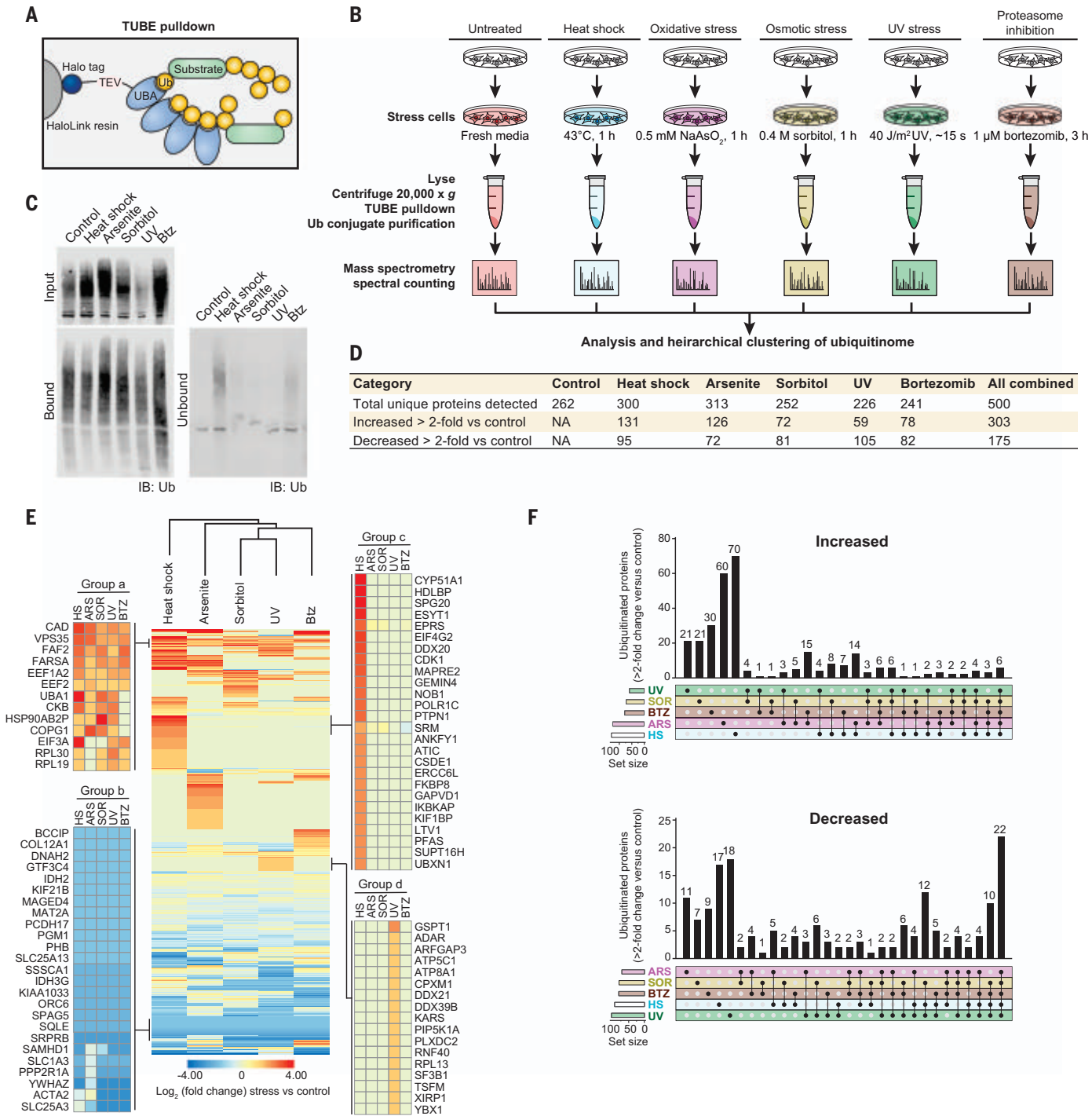*Corresponding author. Email: jpaul.taylor@stjude.org

Fig. 1. Different cellular stresses induce different patterns of ubiquitination.
(**A**) Schematic illustrating HALO-linked TUBE resin and ubiquitinated protein capture from HEK293T cells. TEV, tobacco etch virus protease cleavage site; Ub, ubiquitin; UBA, ubiquitin-associated domain. (**B**) Workflow for sample preparation and TUBE proteomics. HEK293T cells were treated with the indicated stress or incubated with fresh media for 1 hour as a control before lysis. For UV treatment, media was removed and cells were briefly exposed to UV light (40 J/m^2) followed by the addition of fresh media 1 hour before lysis. (**C**) Immunoblot (IB) for ubiquitin (P4D1 antibody) showing TUBE capture of ubiquitinated proteins from cells stressed as in (B). Btz, bortezomib. (**D**) Summary of statistics from proteomic analysis. (**E**) Heatmap illustrating global stress-induced changes to the ubiquitinome. Colors indicate log$_2$ fold change in spectral counts of stressed samples versus control samples. Example groups of proteins with a shared pattern are shown in the blow-ups. (**F**) UpSet plot indicating proteins that changed more than twofold versus control for each stress.

Other proteins showed altered ubiquitination in a stress-specific manner. For example, ubiquitination of some proteins was increased exclusively in response to heat shock (group c in Fig. 1E) or UV exposure (group d in Fig. 1E). The emergence of these distinct protein groups suggested stress-specific patterns of ubiquitination that represented distinct adaptive responses to stressors. To test this hypothesis, we investigated ubiquitination in response to heat and oxidative stress more deeply.

Heat stress leads to rapid increases in total ubiquitination and a time-dependent shift in the solubility of ubiquitin conjugates

We first examined the kinetics of ubiquitin conjugate accumulation during heat shock and recovery. Heat stress–induced ubiquitination occurred rapidly, with increased ubiquitin conjugates apparent as early as 15 min and peaking within 30 min after heat shock (Fig. 2A). The accumulated ubiquitin conjugates remained at a constant increased level during prolonged stress for at least 150 min (Fig. 2A). Immunoblotting using antibodies specific for polyubiquitin (FK2) and K48-linked or K63-linked polyubiquitin chains showed similar stress-dependent increases in signal as those observed using the P4D1 antibody, which recognizes unconjugated ubiquitin as well as monoubiquitinated and polyubiquitinated proteins (fig. S1A). Thus, the observed stress-dependent increase in ubiquitin conjugation consisted largely of polyubiquitination and was not highly chain type–specific.

Prolonged heat shock results in a progressive loss of solubility in a fraction of the proteome (42–45), although the relationship to ubiquitination is not clear. Here, we observed a time-dependent dynamic shift in the ubiquitinome between the soluble and insoluble fractions. Accumulation of polyubiquitin conjugates in the soluble fraction peaked at about 30 min and then declined over the course of the experiment. Polyubiquitin conjugates accumulated more slowly in the insoluble fraction, plateauing after 60 min (Fig. 2A). The accumulation of polyubiquitin conjugates after heat stress was temporary, returning to baseline within 3 hours after a 30-min heat shock (Fig. 2B).

With this kinetic profile in mind, we next performed further proteomic analysis of the heat stress–associated ubiquitinome by TUBE–LC-MS/MS. We examined this ubiquitinome in the soluble fraction at an early time point (15 min) corresponding to the initial increase in ubiquitination and at a later time point (60 min) (Fig. 2C). To recover ubiquitinated proteins accumulated in the pellet fraction at a later time point (Fig. 2A), we resolubilized the pellet fractions from samples at the 60-min time point in a urea buffer before incubation with TUBEs. Combined data from three replicates for both fractions yielded quantification of >1000 ubiquitinated proteins in the TUBE-recovered samples, indicating good depth of coverage and reproducibility (fig. S1, B and C, and table S2) and that our TUBE-based method can detect ubiquitin conjugates at least as efficiently other methods of ubiquitin capture (35). We also verified a heat stress–induced increase in polyubiquitination of seven representative proteins by immunoblotting (fig. S1D). Replicating the 60-min time point experiment in U2OS cells gave results very similar to those in HEK293T cells (fig. S1E and table S3). Similarly,

heat shock led to an accumulation of ubiquitin conjugation in both induced pluripotent stem cell (iPSC)–derived neurons and mouse cortical neurons (fig. S1F), and TUBE proteomic analysis revealed substantial overlap between the heat shock ubiquitinome of HEK293T cells and these neuronal cell types (fig. S1, G to I, and table S3). Indeed, many of the proteins with the largest increase in abundance after heat stress were shared between these three cell types (fig. S1, G to I). Thus, stress-dependent changes to the ubiquitinome were not highly cell type–specific.

The temporal profile of changes in the ubiquitinome in the soluble fraction revealed only modest changes at 15 min but more substantial changes at 60 min after heat stress (Fig. 2, D to F). Thus, despite an increase in a polyubiquitin smear by immunoblotting at 15 min (Fig. 2A), the global shift in the ubiquitinated protein landscape was better reflected later in the stress response. We therefore directed most of our subsequent analyses to examining changes to the ubiquitinome after 60 min of heat shock.

We also detected substantial changes to the ubiquitinome in the pellet fraction. Whereas there was no loss of ubiquitinated proteins in the pellet fraction, 172 ubiquitinated proteins were significantly increased after 60 min of heat stress (Fig. 2, G and H). Roughly 80% of these proteins had also been detected in the soluble fraction, where these proteins were enriched, depleted, or unchanged in response to heat stress (Fig. 2, I and J). Although some ubiquitinated proteins did accumulate in the pellet fraction after heat shock, an individual protein's ubiquitination status was not predictive of its solubility. Because detailed descriptions of heat-dependent changes in solubility have been published elsewhere (42–45), in subsequent analyses we assessed the significance of heat stress–induced ubiquitination of all proteins whether they appeared in the soluble fraction, pellet fraction, or both. Heat stress induced a global shift in the ubiquitin landscape, resulting in a "heat shock ubiquitinome" of 381 proteins, defined as all proteins showing a predicted net total increase in ubiquitination in response to heat stress (Fig. 2J and table S4).

Heat and arsenite stress induce different changes to the ubiquitinome

The pattern of ubiquitination after arsenite stress was most similar to that generated by heat stress (Fig. 1E). To examine this more deeply, we characterized arsenite stress–induced ubiquitination at 60 min following the same approach that we used to define heat stress–induced ubiquitination. After identifying proteins whose ubiquitination was significantly altered relative to control (tables S4 and S5), we compared the responses to arsenite stress versus heat stress. This analysis confirmed substantial overlap in the arsenite

stress– and heat stress–induced ubiquitinomes (fig. S1J). This overlap was evident both for proteins that showed increased ubiquitination and for those that showed decreased ubiquitination (fig. S1K). Nevertheless, this analysis also confirmed stress-specific patterns in ubiquitination. Indeed, more than 400 significant changes in ubiquitination were specific to either arsenite or heat stress (fig. S1, J and K, and tables S2 and S5).

Stress-dependent ubiquitination is not reflected by altered abundance

We next sought to determine the extent to which changes in the stress-induced ubiquitinomes reflected changes in total protein levels owing to altered rates of synthesis or ubiquitin-dependent degradation. To investigate changes in expression at the protein level, we performed multiplexed tandem mass tagging (TMT) MS quantification of the whole proteome from unstressed, heat-stressed, or arsenite-stressed cells in the presence or absence of proteasome inhibition (0.5 μM bortezomib) (Fig. 3A). Lysates were digested and labeled by TMT in 11-plex mode, followed by nanoscale LC–high-resolution MS/MS. Of the 12,586 proteins quantified, very few proteins showed significant stress-dependent change in total abundance using 1.5-fold and $P \leq 0.05$ as a cutoff (Fig. 3B, fig. S2A, and table S6). Nevertheless, we detected several expected changes in response to both stress types, including the accumulation of stress-response transcription factors ATF4 and XBP1 and depletion of BTG1, a regulator of ATF4 (46) (fig. S2B and table S6). No proteins with a stress-dependent increase in ubiquitination showed a significant increase in total protein levels. Thus, increased capture in the TUBE experiments reflected increased ubiquitination rather than increased levels of total protein. Additionally, only eight proteins with heat shock–dependent increases in ubiquitination showed a significant decrease of ≥1.25-fold in total protein levels in response to stress, of which only three proteins (EPPK1, GCN1, and CYP51A1) decreased ≥1.5-fold (blue dots in left panel of Fig. 3B and fig. S2B). The levels of these three proteins were stabilized during heat shock by the addition of bortezomib, indicating that they were targets of heat stress–dependent proteasomal degradation (right panel of Fig. 3B and fig. S2B). Nonetheless, for nearly all cases with heat shock (Fig. 3B) and all cases with arsenite (fig. S2A), a stress-dependent increase in ubiquitination did not result in an overall down-regulation of total protein levels. Treatment with bortezomib during the 1-hour heat shock did not significantly alter ubiquitinated protein enrichment based on TUBE pulldown (fig. S2, C and D), indicating that proteasomal degradation did not significantly alter the composition of the heat shock ubiquitinome during this time frame. Thus, either
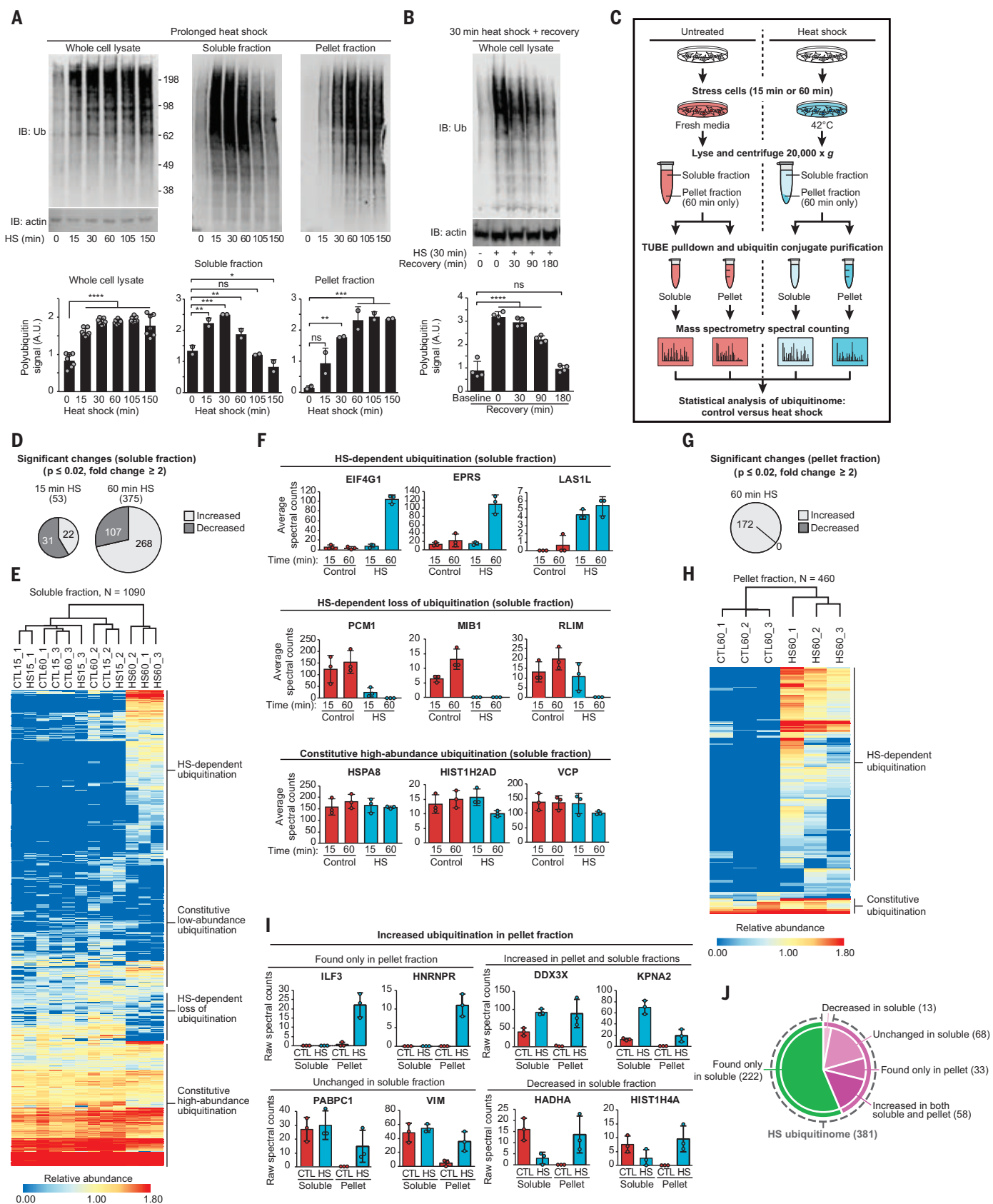
Fig. 2. Heat shock induces substantial changes to the global ubiquitinome.
(**A** and **B**) Immunoblots showing ubiquitin conjugation levels at time points after heat shock (A) and during recovery from 30-min heat shock (B) in HEK293T whole-cell lysates, soluble fractions, and pellet fractions after centrifugation

(14,000 × g). Bar graphs indicate mean + SD and individual values for quantification of Western blots from ≥3 replicates for each experiment. ns, not significant, *P < 0.05, **P < 0.01, ***P < 0.001, and ****P < 0.0001 by analysis of variance (ANOVA) with Dunnett's test. A.U., arbitrary units; HS, heat shock. (C) Workflow for samples from HEK293T cells. (D) Charts indicating the number of significantly changed proteins in soluble fractions for heat-shocked samples (15 min or 60 min) compared with control. (E) Heatmap illustrating heat shock–induced changes to the global ubiquitinome in soluble fractions. Colors indicate relative abundance as quantified by spectral counting. (F) Spectral count values for representative proteins of indicated categories; individual values and mean (±SD) are shown from three replicates of each sample condition in soluble fractions. (G) Chart indicating the number of significantly changed proteins in pellet fractions for heat-shocked samples compared with control. (H) Heatmap illustrating heat shock–induced changes to the global ubiquitinome in pellet fractions. (I) Spectral count values for representative proteins of indicated categories; individual values and mean (±SD) are shown from three replicates of each sample condition. CTL, control; HS, 60-min heat shock. (J) Chart indicating the number of proteins with a heat shock–dependent increase in ubiquitination. Proteins detected in the pellet fraction are shown in purple, and proteins detected only in the soluble fraction are shown in green. The dashed line indicates categories included in the heat shock ubiquitinome.

the observed ubiquitination was partially non-degradative or, more likely, only a fraction of each protein was ubiquitinated in response to stress, such that ubiquitin-dependent degradation did not substantially reduce the overall abundance of most proteins. Similarly, RNA sequencing (RNA-seq) analysis revealed that, although we observed many significant heat shock–dependent changes in transcription, including heat shock proteins (HSPs) and other stress response genes (47), most heat shock ubiquitinome genes were not regulated at the level of transcription, with a few exceptions (Fig. 3C, fig. S2E, and table S7). Furthermore, up-regulation of heat shock response genes at the mRNA level (e.g., *HSPA6* and *DNAJB1*; table S7) was not yet apparent at the protein level at the 1-hour time point used here (table S6). This difference in protein level versus transcript level changes is consistent with the expected lag time between stress-dependent induction at the transcriptional and translational levels (48–50). Thus, although it is likely that the transcriptional response to heat shock has a substantial contribution to recovery from stress over longer time scales, it does not appear to be directly involved in the short-term reversal of stress-response activities investigated here.

Paired di-GLY and TMT quantitative proteomics indicate that heat stress–induced ubiquitination leads to degradation during the recovery phase

Di-GLY proteomics, which uses enrichment of the Lys-ε-Gly-Gly (di-GLY) remnant generated on ubiquitinated peptides after trypsin digestion, is an alternative approach to identify and quantify the ubiquitin-modified proteome (51, 52). To complement and validate our TUBE proteomics, we paired di-GLY proteomics with TMT labeling in the presence or absence of proteasome inhibition to quantify changes in ubiquitination in response to a 60-min heat stress and after 2 hours of recovery. Across all samples, we quantified 16,525 unique ubiquitin-modified peptides, corresponding to 4892 unique proteins (Fig. 3D and table S8). The heat stress–induced ubiquitinomes as determined by di-GLY–TMT proteomics were highly correlated with our TUBE-based proteomics, detecting at least one ubiquitin-modified peptide for most proteins identified by TUBE pull-

down (Fig. 3, E and F, and tables S2 and S8). Of the 381 proteins in the heat shock ubiquitinome identified by TUBE pulldown, 335 had at least one peptide increased by ≥1.5-fold in response to heat shock in the di-GLY analysis (Fig. 3G and tables S4 and S8). Of the 46 that did not, 35 were not detected in the di-GLY experiment, 8 had overall reduced ubiquitinated peptide levels in response to heat shock [heat shock/control (HS/CTL) = 0.53 to 0.95], and 3 had modestly increased levels (HS/CTL = 1.29 to 1.33) (Fig. 3G and tables S4 and S8). The amplitudes of the observed changes were often lower in the di-GLY profiling experiment, likely owing to ratio suppression commonly observed in TMT-based quantification (tables S2 and S8) (53). The heat shock ubiquitinome had much higher numbers of ubiquitination sites per protein (8.6) than the global ubiquitinome (3.4) (Fig. 3, F and H), suggesting concerted ubiquitination of a subset of proteins that define the heat shock ubiquitinome. We also identified 3722 proteins that were not detected in the TUBE spectral counting for which ubiquitin-modified peptide levels were increased by at least 1.5-fold in response to heat shock (Fig. 3G). Comparing these results with the total proteome (table S6), >8000 proteins had no detectable increase in heat shock–dependent ubiquitination, representing most of the expressed proteome in HEK293T cells. Thus, this ubiquitination does not represent a global increase in ubiquitination across the entire proteome. Although the additional heat shock–dependent ubiquitination events detected only by di-GLY profiling represent interesting biology that warrants further investigation, the TUBE-based experiments identified a critical subset of the heat shock ubiquitinome that allowed for comparison between more sample conditions (e.g., soluble versus pellet and arsenite versus other stresses) owing to the reduced time and resources required. We thus adjusted the definition of the "heat shock ubiquitinome" as those proteins detected in the TUBE experiments but excluding the eight proteins that showed a heat shock–dependent decrease in the di-GLY experiment (373 proteins; table S4).

Analysis of the di-GLY data and quantification of modification sites on the internal lysines of ubiquitin revealed a ~1.5- to 2-fold

heat shock–dependent increase in all poly-ubiquitin linkage types except K6 (Fig. 3I). This finding was consistent with our immunoblotting results (fig. S1A) that suggested that the heat shock–dependent increase in polyubiquitination was not highly chain type–specific. K48 was the most abundant linkage type in control conditions and exhibited a further twofold increase after heat shock, suggesting that a substantial portion of the heat shock ubiquitinome is targeted for proteasomal degradation. We therefore analyzed the di-GLY data to determine the effect of proteasome inhibition on the levels of ubiquitin-modified peptides (Fig. 3J). In accordance with our TUBE experiments and whole-proteome analysis, treatment with bortezomib had a modest effect on peptide abundance in heat-shocked samples for most heat shock ubiquitinome proteins. Although most heat shock–dependent ubiquitination recovered to baseline levels after 2 hours in the absence of proteasomal inhibition, the levels remained increased during recovery in the presence of bortezomib. This result suggests that many proteins ubiquitinated during continued heat stress are ultimately targeted for proteasomal degradation upon recovery.

The heat shock ubiquitinome is not defined by the most abundant, highly translated, or thermolabile proteins

We next examined general features of the heat shock ubiquitinome by analyzing whole-proteome TMT and RNA-seq datasets along with several proteome-wide datasets. We note that the label-free MS approach used in our TUBE experiments can be biased toward identifying abundant proteins. However, proteins detected in the TUBE pulldown experiment spanned five orders of magnitude in TMT intensity and included proteins represented in the bottom fifth percentile of abundance in the proteome (tables S2 and S6), indicating that the heat shock ubiquitinome included low-abundance proteins. The heat shock ubiquitinome was not biased by protein or mRNA abundance, and low-abundance proteins were well represented (Fig. 3, K and L). This observation, along with the high degree of correlation with the results of the di-GLY profiling,
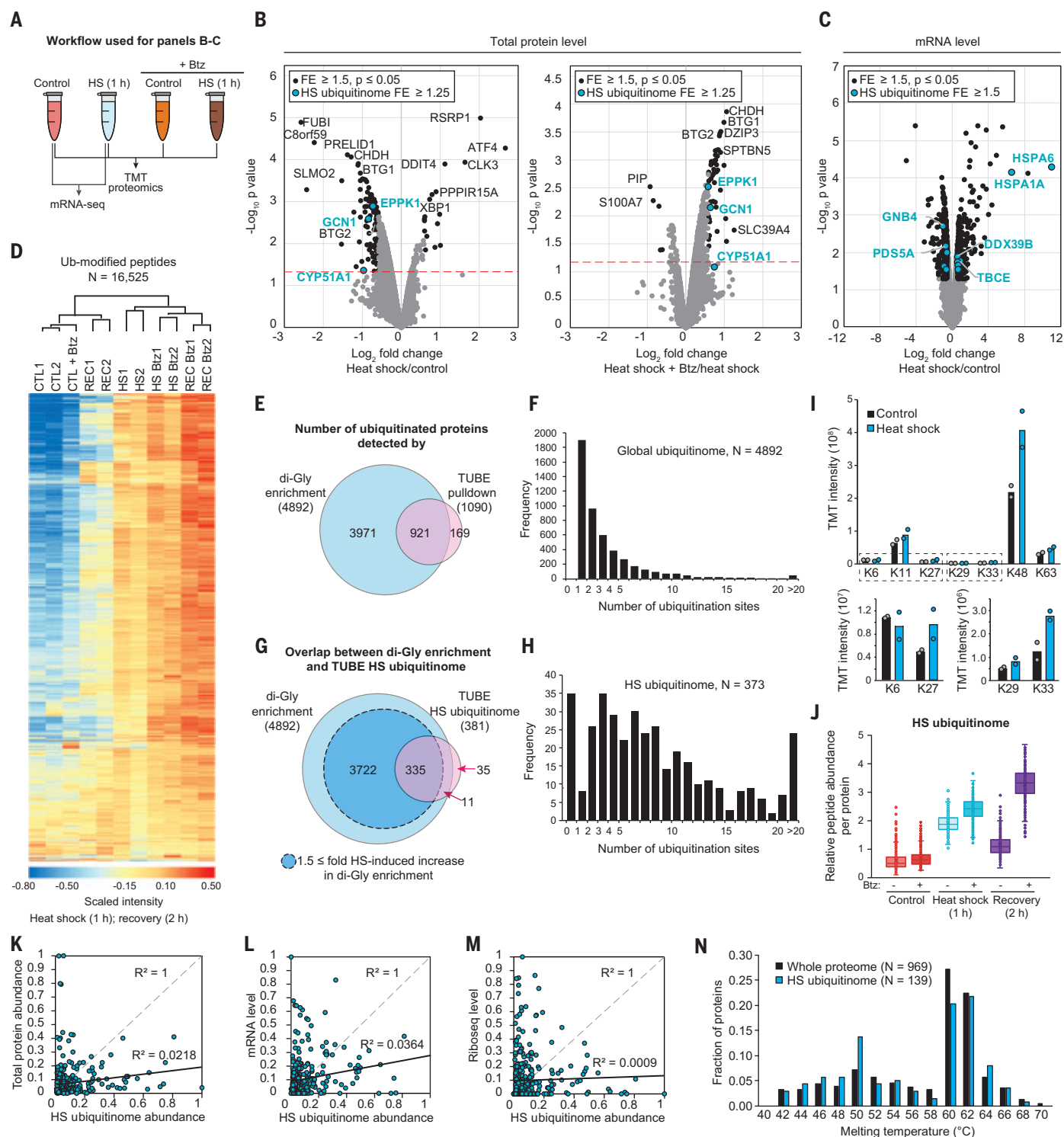
Fig. 3. Proteomic and transcriptomic analyses reveal additional details of the heat shock ubiquitinome. (**A**) Workflow for analysis in HEK293T cells. (**B** and **C**) Volcano plots indicating changes in protein abundance for heat-shocked versus control samples (left) and versus heat-shocked samples with proteasome inhibitor bortezomib (Btz) (right) (B) and changes in mRNA levels in response to heat shock (C). Statistically significant changes are shown for the whole proteome or transcriptome (black dots) and heat shock ubiquitinome (blue dots). Red dashed lines indicate P = 0.05. FE, fold enrichment. (**D**) Heatmap of changes to the ubiquitinome in HEK293T cells after 1-hour

heat shock (HS) and 1-hour heat shock followed by a 2-hour recovery (REC) in the presence or absence of bortezomib as determined by di-GLY profiling. Color indicates scaled TMT intensity. (**E**) Venn diagram showing overlap of ubiquitinated proteins detected in di-GLY and TUBE experiments. (**F**) Histogram showing the number of ubiquitination sites per protein detected for all proteins in the di-GLY experiments. (**G**) Venn diagram showing overlap of heat shock–dependent increases in ubiquitination as determined by di-GLY and TUBE experiments. (**H**) Histogram showing the number of ubiquitination sites per protein for the heat shock ubiquitinome. (**I**) Bar graphs indicating TMT

intensity for each polyubiquitin linkage type. Mean and individual values are shown. (**J**) Box-and-whisker plots showing relative abundance of ubiquitinated proteins for heat shock ubiquitinome proteins. Boxes represent first and third quartiles and whiskers represent minimum and maximum excluding outliers. (**K** to **M**) Correlation between abundance of heat shock ubiquitinome proteins in TUBE experiments and total protein abundance (K), transcript abundance (L), and ribosome occupancy (M) as determined by ribosome profiling (Ribo-seq), with coefficient of determination (R^2) values displayed. (**N**) Histogram showing melting temperatures as determined in (*54*) for all proteins measured (black bars) and measured heat shock ubiquitinome proteins (blue bars).

indicated that the quantification of TUBE pull-down samples reflected levels of ubiquitinated protein rather than total protein or transcript levels. Similarly, based on an analysis of ribosome profiling data from unstressed cells, we found no correlation between the TUBE-based spectral counting quantification of heat shock ubiquitinome proteins and density of ribosome occupancy for the corresponding mRNA (Fig. 3M). Thus, the heat shock ubiquitinome was not composed of the most actively translated proteins. Finally, we compared the heat shock ubiquitinome to a previous proteome-wide investigation of protein thermostability (*54*). This published dataset included 139 proteins from the heat shock ubiquitinome, which showed a similar bimodal distribution of melting temperatures as compared to all 969 proteins analyzed in the study (Fig. 3N). Thus, it is unlikely that the heat shock ubiquitinome represents the proteins most likely to misfold as a result of temperature increase. Taken together, it seemed that the heat shock ubiquitinome was not primarily defined by total protein abundance, translation activity, or thermostability.

The functional landscape of the heat shock ubiquitinome reflects cellular activities perturbed by stress

To investigate the functional importance of the heat shock ubiquitinome, we performed gene ontology (GO) analysis [DAVID (*55*) and PANTHER (*56*)] along with literature curation (Fig. 4, A to C). The heat shock ubiquitinome was highly enriched in biological pathways that were down-regulated or shut down during cellular stress. Furthermore, we detected enrichment of proteins involved in stress granule formation, including translation initiation factors and other components of the translational machinery (Fig. 4C). The heat shock ubiquitinome also included nucleocytoplasmic transport proteins (e.g., 11 of the 18 importin α/β superfamily members, nuclear pore complex components, and RNA transporters). Indeed, stress disrupts nucleocytoplasmic transport through localization of transport factors into stress granules (*8*), further linking ubiquitination with down-regulated cellular activities and stress granules (Fig. 4C).

Under nonstress conditions, PQC proteins, such as HSP chaperones and VCP, are engaged with misfolded proteins and nascent translation products that are targets for proteasomal degradation, and in some cases HSPs themselves can be directly ubiquitinated as a result (*57–59*). These PQC-related proteins were not over-represented in the GO analysis but were captured by TUBEs at high levels in both control and heat shock conditions in many cases (bottom row of Fig. 2F and table S2). Thus, heat shock–induced ubiquitination both occurs as a response to increased PQC demands and plays a role in regulating activities that comprise the cellular stress response, including stress granule assembly. Furthermore, ubiquitination may contribute to underexplored aspects of the heat-stress response such as folate and cholesterol metabolism (Fig. 4C).

Comparisons of the GO analysis of the heat shock– and arsenite-induced ubiquitinomes showed overlap of many GO terms (e.g., splicing and nucleocytoplasmic transport) and others that were specific to heat shock (e.g., translation and metabolic process) (Fig. 4, A and B), suggesting that ubiquitination may play distinct roles in the response to each stress type. As expected, owing to the greater number of unique proteins identified in the heat shock ubiquitinome than for the arsenite ubiquitinome, the heat shock ubiquitinome generally included both a higher number of proteins and higher levels of enrichment over control samples within the GO categories shared by both ubiquitinomes (Fig. 4A and tables S2, S4, and S5). However, the stress-dependent increase in total ubiquitin conjugation was actually higher for arsenite than for heat shock (Fig. 1C and fig. S3A). Thus, the more prominent GO pattern observed for the heat shock ubiquitinome was not simply because of a stronger induction of stress responsive ubiquitination compared with arsenite stress. Rather, different stress types could lead to different patterns of ubiquitination.

mRNPs are ubiquitinated and undergo compositional remodeling upon heat shock

A large fraction of the heat shock ubiquitinome was mRNA-binding proteins (Fig. 4), suggesting that some heat shock–dependent ubiquitination might occur in the context of messenger ribonucleoprotein (mRNP) complexes. To examine this possibility, we first captured mRNPs using magnetic beads conjugated with oligo(dT)$_{25}$, which binds to the poly(A) tail of mRNAs, and performed immunoblotting to assess the presence of mRNA bound to polyubiquitinated proteins (Fig. 5, A and B). Ubiquitin conjugates copurified with polyadenylated mRNA to a much greater extent after heat stress. Poly(A) binding protein (PABP)–mediated 3′ end retrieval is an orthologous method for capturing polyadenylated mRNAs (*60*). When we isolated mRNPs by immunoprecipitation (IP) of endogenous PABPC1, we similarly observed heat stress–induced ubiquitin conjugates that copurified with mRNPs (Fig. 5, A and B). Arsenite stress did not produce a similar level of increased ubiquitin conjugates associated with PABPC1 nor did treatment with the translation inhibitors cycloheximide, puromycin, or cephaeline in the absence of heat stress, indicating that this phenomenon was specific to heat stress (fig. S3A).

The observed heat shock–dependent increase in ubiquitination associated with captured mRNPs suggests that ubiquitination may occur on assembled complexes, rather than on individual RNA-binding proteins. To examine this possibility, we used proteomics to investigate the link between heat shock–dependent changes to the composition of mRNPs and the heat shock ubiquitinome. Direct analysis of the PABPC1 protein-protein interaction network from IP of the endogenous protein revealed substantial remodeling in response to heat stress. Indeed, 105 proteins showed increased interaction with PABPC1 after heat shock, whereas 18 proteins showed decreased interaction ($P ≤ 0.02$, fold enrichment or depletion ≥2, detected in both replicates) (Fig. 5C and table S9). More than half of the proteins that exhibited a heat stress–dependent increase in co-IP with PABPC1 (59 of 103) were also components of the heat shock ubiquitinome (Fig. 5C). Moreover, the relative amount of protein retrieved in PABPC1 IP after heat stress correlated with abundance in the TUBE pull-down experiment (Fig. 5, D and E, and tables S2 and S9). Consistent results were obtained with four additional bait proteins and with oligo(dT) capture, substantiating the idea that heat shock–induced ubiquitination of mRNPs coincides with a remodeling of their composition (Fig. 5E and table S9). Some of the interactions in the co-IP experiments were observed at low levels in control samples but increased with heat shock (e.g., GCN1 with EIF4G), whereas others were only detected after heat shock (e.g., TTLL12 with PABP and EIF4G) (Fig. 5E and table S9). One interpretation of these experiments is that PABPC1 complexes were stabilized through direct, polyubiquitin-mediated interactions. However, when we repeated the PABPC1 IP with the addition of the nonspecific deubiquitinating enzyme USP2, which was sufficient to eliminate polyubiquitin, we saw no change in the PABPC1 interactome (fig. S3B). Thus, heat shock–induced ubiquitination correlated with
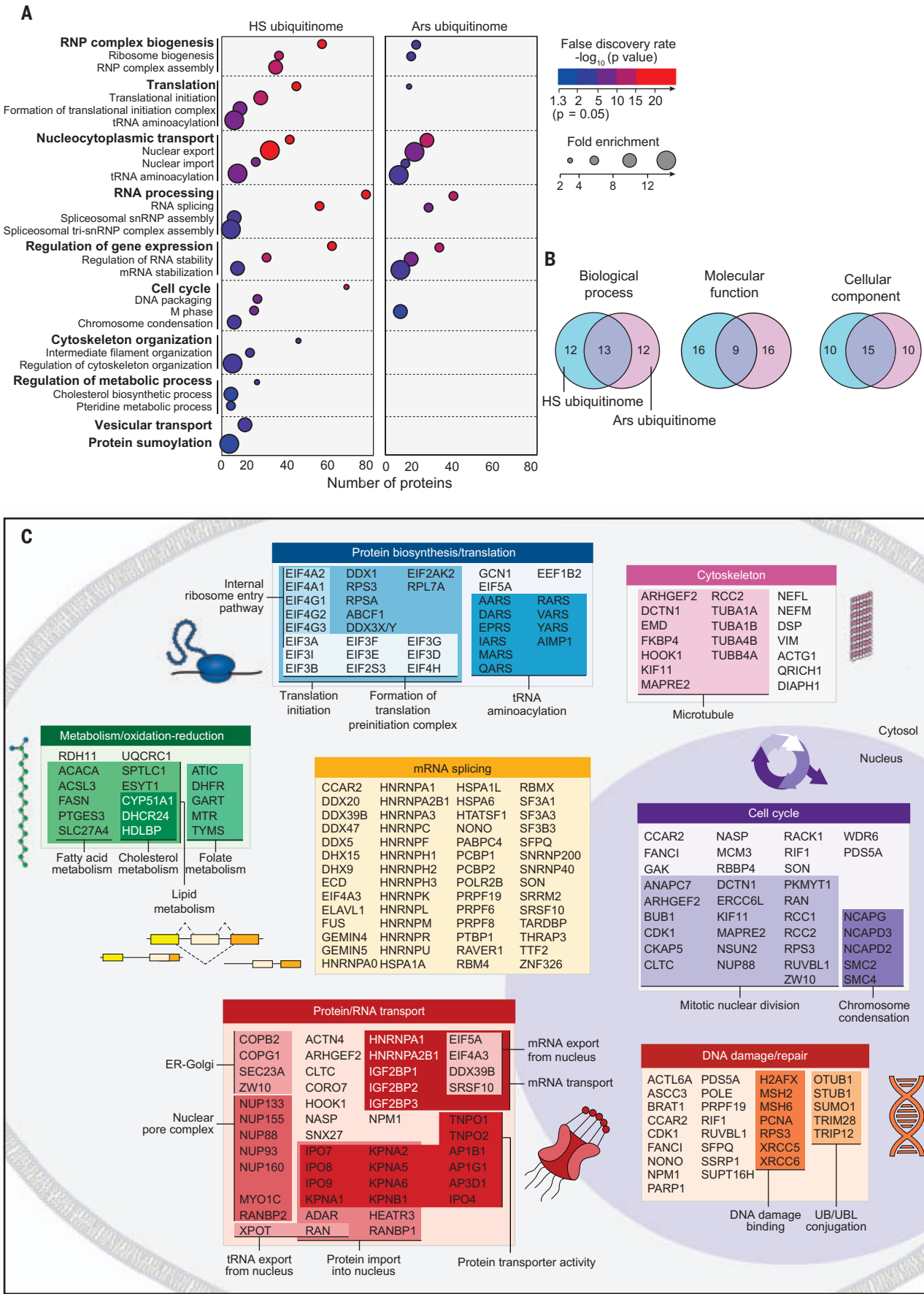
A

B

Biological process, Molecular function, Cellular component Venn diagrams

C

Fig. 4. The heat shock ubiquitinome consists of proteins from biological pathways associated with the stress response. (**A**) Dot plot of GO enrichment showing significantly overrepresented GO terms for heat shock and arsenite (Ars) ubiquitinomes in HEK293T cells. Color indicates FDR *P* value, and dot size indicates overrepresentation fold enrichment compared to the whole genome. Dots are not shown for terms with no statistically significant (*P* < 0.05) enrichment. (**B**) Overlap of top 25 GO as terms ranked by statistical significance (FDR *P* value) for heat shock and arsenite ubiquitinomes. (**C**) DAVID functional clustering of the heat shock ubiquitinome along with literature curation was used to identify pathways targeted by heat shock–induced ubiquitination.
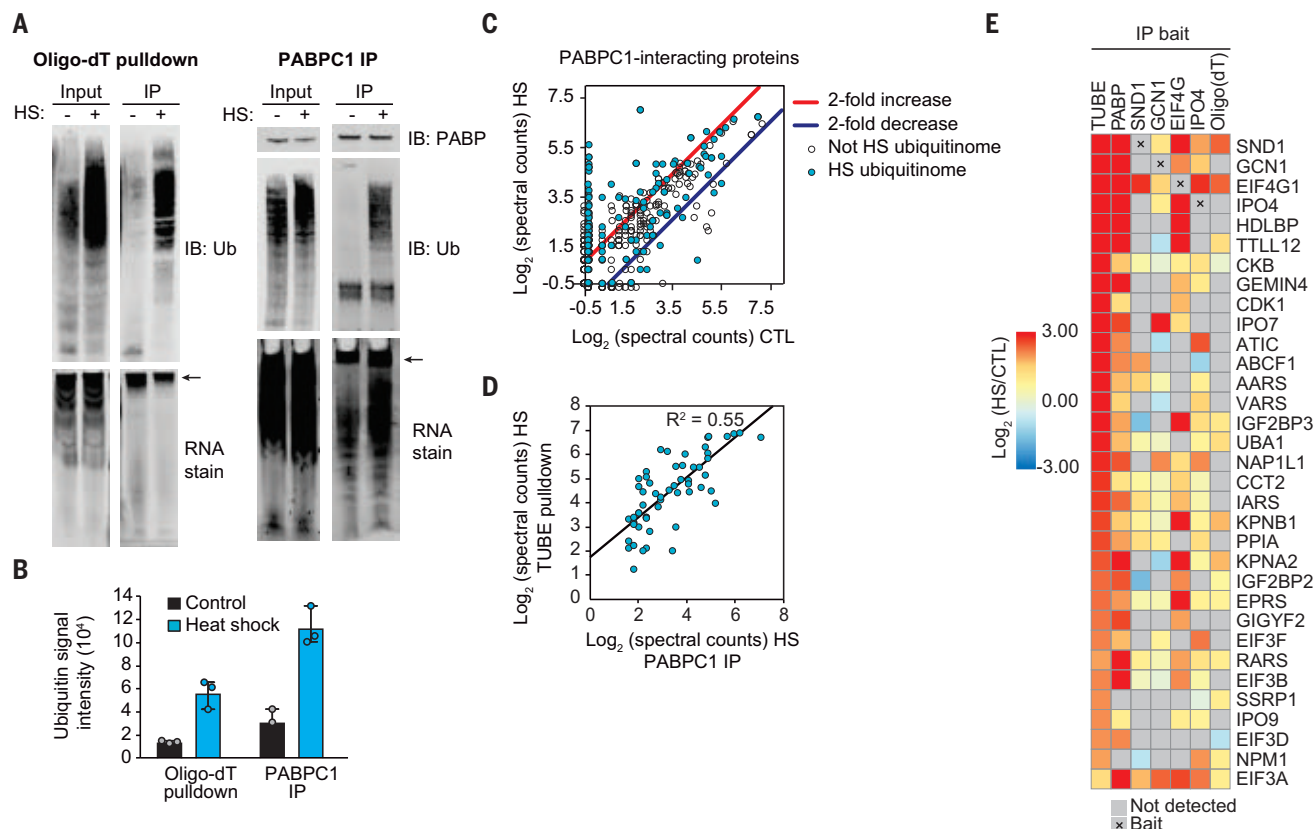

Fig. 5. Heat shock induces ubiquitination of mRNP complexes. (**A**) Immuno-blots and RNA gels showing polyadenylated mRNA isolated by oligo(dT) resin (left) or PABPC1 IP (right) pulling down polyubiquitinated proteins from HEK293T cells after 60-min heat shock. Isolated mRNA was visualized by SYBR Green RNA stain; the arrows indicate high–molecular weight RNA. (**B**) Quantification of immunoblot analysis for ubiquitin conjugation shown in (A). Results represent mean and individual values of three replicate experiments. Error bars indicate data range. (**C**) Scatter plot showing abundance [log$_2$ (spectral counts)] of PABPC1-interacting proteins in heat shock versus control conditions. Results represent averages of two replicates. Lines indicate twofold increase (red) or decrease (blue) with heat shock; blue dots indicate heat shock ubiquitinome proteins. (**D**) Correlation between abundance of heat shock ubiquitinome proteins detected in TUBE pulldown and PABPC1 IP in 60-min heat-shocked samples. (**E**) Heatmap illustrating changes in abundance in 60-min heat-shocked versus control samples for proteins detected in TUBE pulldown, endogenous protein IP (PABP, SND1, GCN1, EIF4G, IPO4), or oligo(dT) pulldown.

a remodeling of the protein components of mRNPs, with some preexisting interactions that became stabilized upon heat shock as well as some newly formed heat shock-dependent interactions.

Ubiquitination is required for the disassembly of heat shock–induced stress granules

We next pursued our observation that the heat shock ubiquitinome was strongly enriched for protein constituents of stress granules. We began by assembling a list of 726 protein constituents of stress granules as reported in three studies (*25*, *26*, *28*) (table S10). Many more stress granule proteins were represented in the heat shock ubiquitinome than in the arsenite ubiquitinome (Fig. 6, A and B). For a high-confidence stress granule proteome (those detected in at least two of the three studies mentioned above) (tables S2, S5, and S10), ubiquitination of stress granule proteins clearly increased in response to heat shock (Fig. 6C). As expected (*35*), this increase in ubiquitination was not apparent in response to arsenite stress (Fig. 6D). Indeed, polyubiquitin robustly colocalized with stress granule markers in nearly all cells exposed to heat shock but not in those exposed to arsenite (Fig. 6, E to G). The group of stress granule proteins with increased ubiquitination in response specifically to heat stress

included the key stress granule regulator G3BP1, the central node in the network of interactions that underlie stress granule assembly (*61*). In our companion study, we define the lysine residues in G3BP1 that are ubiquitinated in response to heat shock and the mechanism whereby G3BP1 ubiquitination mediates stress granule disassembly during recovery from heat stress (*62*).

To further investigate the relationship between polyubiquitin conjugation and stress granule dynamics, we used the E1 ubiquitin activating enzyme (UBA1) inhibitor TAK243 (*63*). First, we verified that TAK243 treatment led to dose-dependent and time-dependent
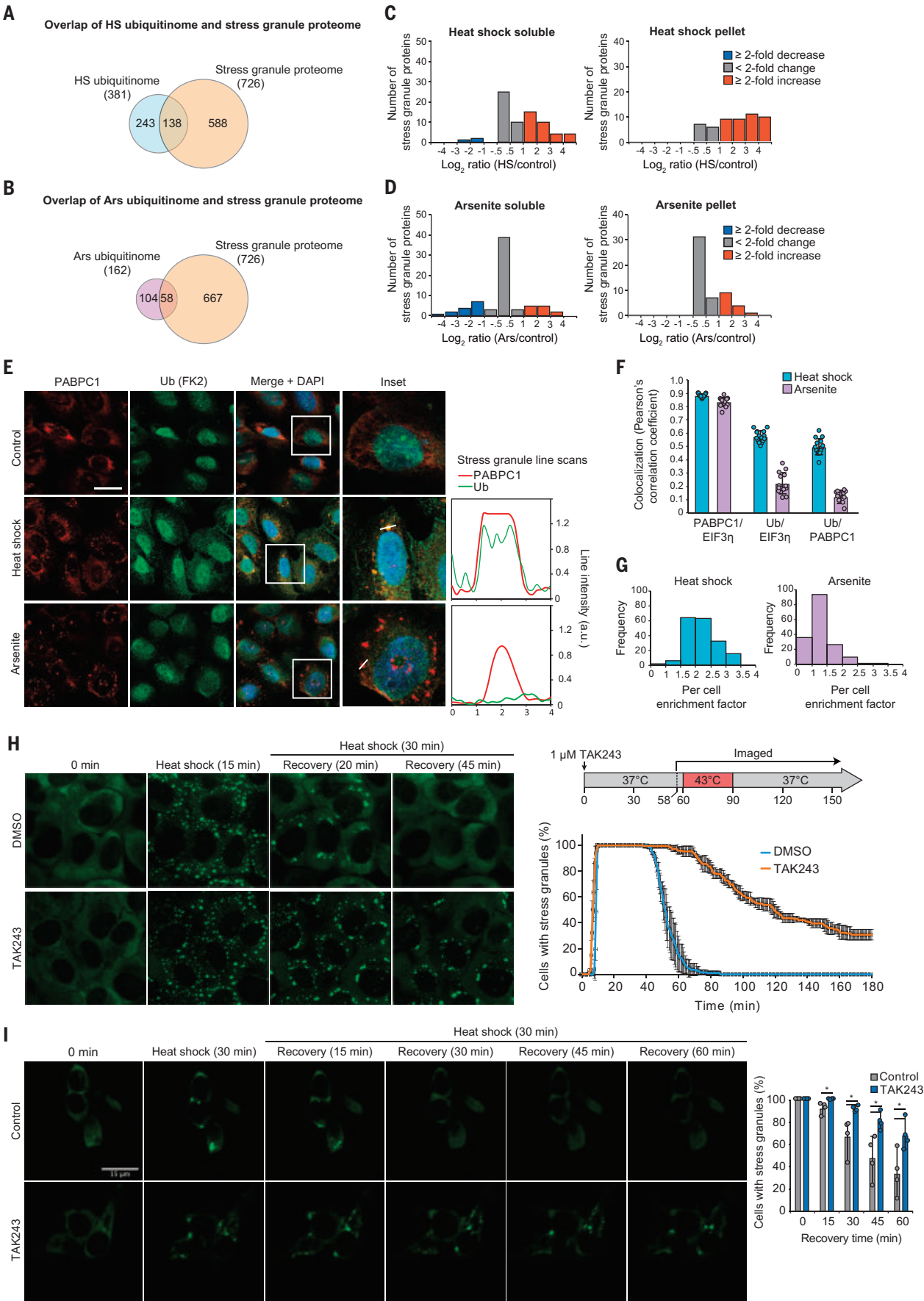
Fig. 6. Heat shock–induced stress granules contain ubiquitinated proteins and require active ubiquitination for disassembly. (**A** and **B**) Overlap of heat shock (A) and arsenite (B) ubiquitinomes of HEK293T cells with the stress granule proteome. (**C** and **D**) Histograms showing changes in ubiquitinated protein abundance for high-confidence stress granule proteins in response to 60-min heat shock (C) or 60-min arsenite stress (D) as determined by TUBE (tables S2 and S5). (**E**) U2OS cells were exposed to 90 min of sodium arsenite or heat shock, fixed, and stained for stress granule markers PABPC1 (red), EIF3η (not shown), and polyubiquitin (FK2, green). Graphs represent a line scan of signal intensity for PABPC1 and FK2 channels across the indicated stress granule. Scale bar, 50 μm. DAPI, 4′,6-diamidino-2-phenylindole. (**F**) Colocalization between immunofluorescent signal for PABPC1 and EIF3η and polyubiquitin as determined images collected as in (E). Bar graphs represent mean (±SD) and individual Pearson's correlation coefficient values for >10 images (total of >100 cells for each condition). (**G**) Histogram of enrichment factor of ubiquitin in stress granules (ratio of the mean intensity of FK2-ubiquitin signal in granules to the mean intensity of the total cell area, excluding granules) for >100 cells for each condition. (**H** and **I**) Live-cell imaging of cells stably expressing green fluorescent protein (GFP)–G3BP1 during heat stress and recovery in the presence or absence of TAK243 for U2OS cells (H) or iPSC-derived neurons (I). Plots show the percentage of cells with ≥2 stress granules at the indicated time, with [(H), line graph] solid lines and error bars representing average values and SD for three biological replicates with 30 to 50 cells each and [(I), bar graph] mean, data range [error bars in (I)], and individual values shown for four biological replicates with 48 to 120 neurons each. DMSO, dimethyl sulfoxide.

depletion of cellular ubiquitin conjugates (fig. S4A). Whereas high concentrations or prolonged exposure to TAK243 was toxic to cells (fig. S4B), incubation with 1 μM TAK243 was sufficient to block most cellular ubiquitination within 20 min while causing no detectable loss of viability for up to 4 hours (fig. S4, A and B). Furthermore, ubiquitination did not affect the dynamics of stress granules induced by arsenite (fig. S4C and movies S1 and S2), as expected (*35*). However, given the differences between the ubiquitin response to heat shock and arsenite, we next assessed the role of ubiquitination in regulating the assembly and disassembly of stress granules induced by heat shock. Surprisingly, elimination of ubiquitination with TAK243 did not prevent stress granule formation in response to heat shock but impaired stress granule disassembly after the removal of stress (Fig. 6H and movies S3 and S4). We confirmed this observation with imaging of 10 additional stress granule markers, which showed persistence in stress granules for up to 2 hours after restoration of normal growth temperature (fig. S4, D to F). We also observed a similar stress granule disassembly defect in cells genetically depleted for *UBA1* (fig. S4, G to I). We next repeated the live-cell imaging analysis of stress granule dynamics in iPSC-derived neurons. Here, too, we observed a significant delay in stress granule disassembly during recovery from heat shock in response to TAK243 treatment (Fig. 6I). Thus, across a wide variety of cell types, ubiquitination is essential for normal disassembly of stress granules during recovery from heat stress.

Given that VCP is involved in stress granule disassembly (*29, 32, 33*), we next sought to compare the effects of inhibition of UBA1 and VCP on stress granule disassembly kinetics. Treatment with the VCP inhibitor CB5083 before heat shock led to delayed stress granule disassembly with nearly identical kinetics as those observed for cells treated with TAK243 (fig. S4J). Thus, ubiquitination and VCP likely act on stress granule disassembly through the same pathway.

We also examined how mitigating stress granule assembly affected the heat shock ubiquitinome. Blocking stress granule assembly either pharmacologically [using cycloheximide (*64*)] or genetically [using *G3BP1/2* double-knockout cells (*61*)] had no detectable effect on the heat shock ubiquitinome, consistent with the idea that ubiquitination of mRNPs occurs upstream and independent from stress granule formation (fig. S5, A to D, and table S3). Thus, whereas stress-dependent ubiquitination can occur in the absence of heat shock–induced stress granules and is not required for their assembly, it is essential for their rapid disassembly during recovery.

Reversal of mRNP remodeling after heat shock requires ubiquitination

Given our finding that disassembly of heat shock–induced stress granules was dependent on ubiquitination, we next investigated the importance of ubiquitination in the reversibility of the heat shock–dependent interactions we observed in the remodeled mRNPs. We added 1 μM TAK243 to the media before and during heat shock and confirmed elimination of the polyubiquitin signal associated with either PABPC1 or oligo(dT) retrieval (Fig. 7A and fig. S6). TAK243 treatment did not block the heat shock–dependent remodeling of the mRNP interactome; however, similar to what we observed with stress granules, elimination of ubiquitination prevented the reversal of this stress-dependent remodeling during recovery (Fig. 7A and fig. S6). Thus, ubiquitination is required for disassembly of stress granules and for the reversibility of mRNP remodeling after restoration of normal growth conditions after heat shock.

Ubiquitination is required for recovery of heat shock–induced impairment of nucleocytoplasmic transport

To investigate the role of ubiquitination in stress-induced perturbation of nucleocytoplasmic transport, we used a Shuttle-tdTomato nucleocytoplasmic reporter assay in which tdTomato was conjugated to both a nuclear localization signal (NLS) and nuclear export signal (NES) that mediate its shuttling between the nucleus and cytoplasm (*8*). Upon heat shock, we observed a significant redistribution of the reporter from the nucleus [ratio of nuclear to total fluorescence (N/W) ~ 0.75 at steady state] to the cytoplasm (N/W ~ 0.5 after heat shock) (Fig. 7, B and C). This result is consistent with a previous report that stress-dependent relocalization of this reporter was due to a stress-induced defect in exportin-1–mediated nuclear export (*8*). Under normal growth conditions, inhibition of ubiquitination with TAK243 had no significant impact on the ratio of nuclear versus cytoplasmic reporter signal, indicating that ubiquitination was not a critical determinant of canonical nuclear import or export. Moreover, after elimination of ubiquitination with TAK243, cells exposed to heat shock still showed a defect in nuclear import activity, indicating that ubiquitination was not required for this inhibition. However, cells treated with TAK243 failed to recover from impaired nucleocytoplasmic shuttling even upon restoration of normal growth conditions for 2 hours. Thus, similar to disassembly of stress granules and reversal of mRNP remodeling, ubiquitination was essential for recovery of nucleocytoplasmic shuttling after heat stress (Fig. 7, B and C).

Ubiquitination is required for reinitiating translation after heat shock

Protein synthesis is also strongly perturbed by heat stress (*65*) owing to broad shutdown of most translation, with the exception of HSPs whose translation is preserved (*66*). To investigate the role of ubiquitination in stress-induced perturbation of translation, we used a metabolic labeling assay (*67*). As expected, upon heat shock, translation activity was dramatically decreased as measured by puromycin incorporation into nascent translation products, followed by a time-dependent recovery after restoration of normal growth conditions (Fig. 7, D and E). Whereas TAK243 treatment had no effect on basal translation levels, heat-shocked cells treated with TAK243 showed a profound failure to recover protein synthesis after restoration of normal growth conditions (Fig. 7, D and E). Thus, ubiquitination was essential for the recovery of translation during the recovery phase from heat shock.
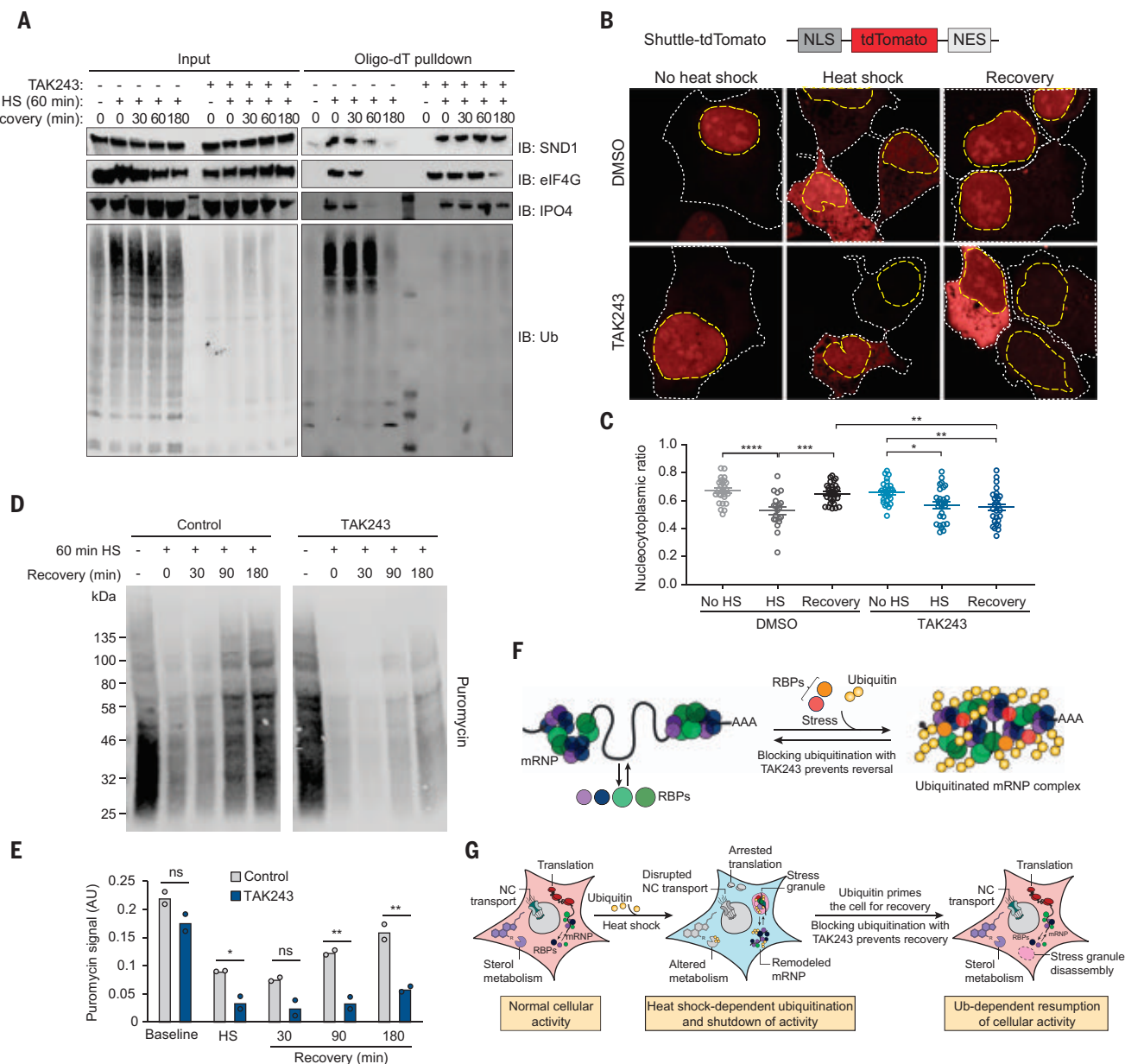
Fig. 7. Ubiquitination is required for the recovery of nucleocytoplasmic transport and translation after heat shock. (**A**) Immunoblot showing formation and dissolution of polyubiquitinated protein–mRNA complexes during heat shock and isolated by oligo(dT) resin. Where indicated, TAK243 was added to media 30 min before heat shock and maintained for the duration of the experiment. In drug-treated, unstressed samples, HEK293T cells were incubated with TAK243 for 180 min before lysis. (**B**) HEK293T cells expressing nucleocytoplasmic transport reporter NLS-tdTomato-NES were fixed and imaged after no stress, 60-min heat shock, or 60-min heat shock and 120-min recovery. Cells were treated with TAK243 or DMSO for 30 min before heat shock or for a total of 120 min in non–heat-shocked cells. Nuclear and cytoplasmic boundaries are indicated with dashed yellow and white lines, respectively. (**C**) Quantification of nucleocytoplasmic ratio of tdTomato intensity from cell images described in (B) from 20 to 30 cells for each condition. Error bars indicate SEM. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$, and ****$P < 0.0001$ by ANOVA with Tukey's test. (**D**) Immunoblotting of HEK293T cells treated with puromycin to label nascent transcripts for 30 min before heat shock and recovery in the presence or absence of TAK243. Cells were lysed at indicated times, and translational activity was analyzed by immunoblotting for puromycin. (**E**) Quantification of immunoblots shown in (D). Average and individual values for puromycin signals are shown for two replicate experiments. ns is not significant, *$P < 0.05$, and **$P < 0.01$ by Student's t test. (**F** and **G**) Proposed model illustrating heat shock–induced ubiquitination associated with mRNP remodeling (F) and stress granule formation and shutdown of cellular activities and the requirement for active ubiquitination for reversal of these processes during recovery (G). NC, nucleocytoplasmic; RBPs, RNA-binding proteins.

Together, our results demonstrate that heat stress induces the targeted ubiquitination of a specific set of proteins and that this ubiquitination is required for recovery from multiple heat stress–induced cellular perturbations, including reversal of mRNP remodeling, disassembly of stress granules, and resumption of nucleocytoplasmic transport and protein synthesis (Fig. 7, F and G). Furthermore, the role of stress-induced ubiquitination is context dependent; indeed, several of these observations did not hold true for arsenite stress. Our extensive

proteomics datasets describing the heat shock ubiquitinome also suggest that ubiquitination facilitates other aspects of the stress response that have yet to be elucidated. Indeed, in preliminary studies exploring the role of ubiquitination on sterol metabolism, we observed that TAK243 treatment had considerable effects on the cholesterol biosynthesis pathway during both heat shock and recovery in a pattern entirely different than that observed for the other cellular activities investigated above (Fig. 7G and fig. S7).

Discussion

The results presented here provide several insights into the mechanisms by which ubiquitination facilitates recovery from heat stress. Our proteomic and biochemical results reveal a heat shock–dependent increase in ubiquitination of mRNPs and stress granule resident proteins, and we propose that the presence of these ubiquitin conjugates within stress granules facilitates an interaction with VCP to promote their timely disassembly during recovery (62). This disassembly paradigm is apparently not used during recovery from arsenite stress (35), despite previous evidence suggesting the involvement of VCP in this process. VCP can function in a ubiquitin-independent manner to disassemble some protein complexes (68), which may explain this apparent contradiction. Alternatively, VCP may not actively disassemble arsenite-induced stress granules but may act to prevent the localization of ubiquitinated DRiPs into aberrant stress granules. These DRiPs would then accumulate when VCP is inhibited, resulting in a loss of dynamism and resistance to disassembly or clearance by a ubiquitin-independent mechanism. This scenario would account for the observation of a disassembly defect upon VCP inhibition but not UBA1 inhibition.

Inhibited ubiquitination also prevented the recovery of nucleocytoplasmic transport and translation activity. Perhaps critical transport and translation factors are sequestered in the persistent stress granules and thus unable to perform their normal functions upon the removal of stress. This has been proposed as a mechanism by which cellular stress disrupts nucleocytoplasmic transport (8), and our results further suggest a role for ubiquitination in liberating proteins from stress granules after the removal of heat stress.

Although the ubiquitin conjugation substrates identified here provide insight into the stress response, identification of the E3 ligases responsible for this heat shock–dependent ubiquitination is an important next step. Our total proteome analysis did not reveal any statistically significant stress-dependent increase in E3 ligases, suggesting that the increased ubiquitination was not likely influenced by up-regulation of specific E3s. However, several E3 ligases were highly enriched in the TUBE pulldown experiments, including TRIP12, TRIM25, and TRIM28, which warrant future investigation as candidate E3s. Nedd4, an E3 ligase critical for heat shock–dependent ubiquitination in HeLa cells and mouse embryonic fibroblasts, was highly enriched after heat shock in our TUBE pulldown from mouse cortical neurons but not in any of the human cells used here. Thus, despite a commonality of ubiquitination substrates between cell lines, the E3 ligases involved may have some cell type specificity. Finally, in accord with a recent investigation identifying a role for SUMO-targeted ubiquitination in the stress response (34), we found that SUMO was among the most enriched heat shock ubiquitinome proteins in both our TUBE and di-GLY analyses, suggesting that SUMO-targeted E3 ligases are likely to contribute to heat shock–dependent ubiquitin conjugation.

Certain aspects of the adaptive nature of stress-induced ubiquitination have long been recognized. Particularly with proteotoxic insults such as heat shock, stress induces several types of potentially hazardous proteins (e.g., DRiPs, thermolabile misfolded proteins, and other damaged proteins) that must be cleared to prevent the formation of toxic species (29–31, 42–44). This need is largely met by ubiquitination and subsequent proteasomal degradation of the offending species. A substantial portion of the proteome becomes transiently insoluble in response to heat stress (42–45), which suggested a second mode whereby the cell can temporarily reshape the proteome to survive a stressful period. Our results add a third dimension to these concepts, indicating that changes to the ubiquitinome are tailored to specific stressors and that, in the case of heat shock stress, ubiquitination serves to prime the cell for recovery. Furthermore, our direct comparisons of the relationship between ubiquitination and heat shock– or arsenite-induced stress granules highlight important differences in context-dependent mechanisms regulating what otherwise appear to be very similar assemblies. Finally, our combined proteome and ubiquitinome datasets represent a vast repository of potentially interesting findings. Our di-GLY-TMT analysis is especially rich, representing a deep and comprehensive ubiquitinome, including many heat shock–dependent ubiquitination events that were not reflected in our TUBE results. These datasets therefore provide a valuable community resource providing insight into many unexplored aspects of the roles of ubiquitination in response to stress.

Methods summary

All experiments were performed in HEK293T and/or U2OS cells unless otherwise indicated. Primary cultures of cortical neurons were prepared from embryo day 15 ICR mice (Charles River Laboratories) as previously described (69). iPSC neurons were differentiated with a two-step protocol (predifferentiation and maturation) as previously described (70). For heat shock treatment, cell culture media was replaced with fresh media prewarmed to 42°C and placed at 42°C for the indicated time. In recovery experiments, the cells were returned to the 37°C incubator for the indicated time. For matched control-treated cells, media was replaced with fresh media warmed to 37°C and maintained at 37°C for the indicated time before harvest. For other stresses, media was replaced with fresh media containing 0.5 mM sodium arsenite, 0.4 M sorbitol, or 1 μM bortezomib for the indicated time. For UV stress, media was removed and replaced with 2 ml of phosphate-buffered saline (PBS) and 40 J/m^2 UV radiation was delivered. When indicated, 1 μM TAK243 was added directly to cell culture media 30 min before stress treatment and maintained in the media until lysis.

Cell pellets were collected by centrifugation followed by removal of media by aspiration, and the cell pellet was washed twice with PBS. Cell pellets were then lysed in the indicated buffers supplemented with deubiquitinase and protease inhibitors. Lysates were centrifuged, and the supernatant fraction was collected as the "soluble fraction" for TUBE or IP lysis buffers or "whole-cell lysate" for urea lysis buffer. To collect pellet fractions, the insoluble pellet from TUBE lysis buffer was washed and resolubilized in urea lysis buffer.

For TUBE pulldowns, soluble or pellet cell lysate fractions were incubated with His-Halo-TUBE beads overnight at 4°C. Beads with captured proteins were then washed, and captured proteins were then released from the beads by denaturation in 1X Laemmli sample buffer with reducing agent and heating at 75°C for 5 min.

For IP and oligo(dT) pulldown, cells were lysed in IP lysis buffer and lysates were incubated overnight at 4°C with Protein G Dynabeads (Invitrogen) conjugated to antibodies against endogenous proteins according to the manufacturer's protocol. For oligo(dT) pulldown, lysate was incubated with Oligo d(T)$_{25}$ magnetic bead resin (New England Biolabs) overnight at 4°C. After overnight incubation, beads were washed and proteins were eluted in 1X lithium dodecyl sulfate (LDS) sample buffer with reducing agent for analysis by immunoblot and MS.

MS analysis was performed according to an optimized platform as previously reported (71). Eluates from the TUBE pulldown or IP experiments were run on an SDS gel and visualized by Coomassie staining (Thermo Fisher Scientific). The whole gel lanes were excised into gel bands. The digested peptides were extracted from the gel bands, dried, and reconstituted in loading buffer. Peptides were

then fractionated by a high-performance LC (HPLC) system and detected by an in-line mass spectrometer. MS/MS spectra were searched against the UniProt human protein database and filtered by mass accuracy and matching scores to reduce the protein false discovery rate (FDR) to about 1%.

Proteomic profiling of the whole proteome was carried out with a previously reported protocol (72). Briefly, for each sample, HEK293T cells from one 10-cm dish were harvested, washed with ice-cold PBS, and collected by centrifugation. About 100 μg of protein per sample of cell lysate was digested, followed by Cys reduction, alkylation, and quenching. The samples were further digested with trypsin and differentially labeled with 11-plex TMT reagents (Thermo Fisher Scientific) according to the manufacturer's protocol.

The TMT-labeled samples were mixed equally, desalted, and fractionated on an off-line HPLC (Agilent 1220) using basic-pH reverse-phase LC. The fractions were dried, reconstituted and analyzed by acidic-pH reverse-phase LC-MS/MS on an Ultimate 3000 UPLC system (Thermo Fisher Scientific). Peptides were eluted, ionized by electrospray ionization, and detected by an in-line Orbitrap Fusion mass spectrometer. MS/MS spectra were searched against the UniProt human protein database and filtered by mass accuracy and matching scores to reduce protein FDR to about 1%.

For di-GLY TMT ubiquitinome analysis, proteomic profiling of the ubiquitinome was performed with a previously reported protocol (52) with modification. Briefly, for each sample, HEK293T cells from five 15-cm dishes were harvested, washed with ice-cold PBS, and collected by centrifugation. Cells were lysed, immediately digested with LysC (Wako), quenched with dithiothreitol (DTT) (30 mM for 30 min), and subjected to trypsin digestion and desalting. The desalted peptides were resuspended, centrifuged, and incubated with di-GLY antibody beads (Cell Signaling Technology) for 2 hours at 4°C. The di-GLY peptides were eluted at room temperature, dried, and resuspended for 11-plex TMT labeling. The TMT-labeled peptides were combined equally among replicates, desalted, and analyzed by multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) using a similar protocol descried above.

For proteomics statistics, the summed spectral counts of replicate experiments were compared among different treatments with missing values replaced by a constant low value to calculate fold change and P values. Stress-induced changes were considered statistically significant with $P \leq 0.02$ and fold change ≥2, excluding proteins detected in only one of three replicates for experiments performed in triplicate. In TMT-based quantification experiments (total proteome and di-GLY profiling), stress-induced changes were considered statistically significant with $P \leq 0.05$ and fold change ≥1.5. The heat shock ubiquitinome was defined from the TUBE experiment as proteins with a statistically significant increase in the soluble fraction and proteins with a statistically significant increase in the pellet fraction but not decreased in the soluble fraction (Fig. 2J), excluding eight proteins for which peptides were measured to have decreased abundance in the di-GLY TMT experiment.

REFERENCES AND NOTES

1. H. P. Harding et al., An integrated stress response regulates amino acid metabolism and resistance to oxidative stress. Mol. Cell 11, 619–633 (2003). doi: 10.1016/S1097-2765(03)00105-9; pmid: 12667446

2. K. Pakos-Zebrucka et al., The integrated stress response. EMBO Rep. 17, 1374–1395 (2016). doi: 10.15252/embr.201642195; pmid: 27629041

3. C. O. Brostrom, C. R. Prostko, R. J. Kaufman, M. A. Brostrom, Inhibition of translational initiation by activators of the glucose-regulated stress protein and heat shock protein stress response systems. Role of the interferon-inducible double-stranded RNA-activated eukaryotic initiation factor 2α kinase. J. Biol. Chem. 271, 24995–25002 (1996). doi: 10.1074/jbc.271.40.24995; pmid: 8798781

4. T. E. Dever et al., Phosphorylation of initiation factor 2α by protein kinase GCN2 mediates gene-specific translational control of GCN4 in yeast. Cell 68, 585–596 (1992). doi: 10.1016/0092-8674(92)90193-G; pmid: 1739968

5. D. Ron, Translational control in the endoplasmic reticulum stress response. J. Clin. Invest. 110, 1383–1388 (2002). doi: 10.1172/JCI0216784; pmid: 12438433

6. R. C. Wek, H. Y. Jiang, T. G. Anthony, Coping with stress: eIF2 kinases and translational control. Biochem. Soc. Trans. 34, 7–11 (2006). doi: 10.1042/BST0340007; pmid: 16246168

7. P. D. Lu, H. P. Harding, D. Ron, Translation reinitiation at alternative open reading frames regulates gene expression in an integrated stress response. J. Cell Biol. 167, 27–33 (2004). doi: 10.1083/jcb.200408003; pmid: 15479734

8. K. Zhang et al., Stress granule assembly disrupts nucleocytoplasmic transport. Cell 173, 958–971.e17 (2018). doi: 10.1016/j.cell.2018.03.025; pmid: 29628143

9. M. Furuta et al., Heat-shock induced nuclear retention and recycling inhibition of importin alpha. Genes Cells 9, 429–441 (2004). doi: 10.1111/j.1356-9597.2004.00734.x; pmid: 15147272

10. J. B. Kelley, B. M. Paschal, Hyperosmotic stress signaling to the nucleus disrupts the Ran gradient and the production of RanGTP. Mol. Biol. Cell 18, 4365–4376 (2007). doi: 10.1091/mbc.e07-01-0089; pmid: 17761537

11. M. Kodiha, A. Chu, N. Matusiewicz, U. Stochaj, Multiple mechanisms promote the inhibition of classical nuclear import upon exposure to severe oxidative stress. Cell Death Differ. 11, 862–874 (2004). doi: 10.1038/sj.cdd.4401432; pmid: 15088071

12. S. Kose, N. Imamoto, Nucleocytoplasmic transport under stress conditions and its role in HSP70 chaperone systems. Biochim. Biophys. Acta 1840, 2953–2960 (2014). doi: 10.1016/j.bbagen.2014.04.022; pmid: 24797038

13. Y. Miyamoto et al., Cellular stresses induce the nuclear accumulation of importin α and cause a conventional nuclear import block. J. Cell Biol. 165, 617–623 (2004). doi: 10.1083/jcb.200312008; pmid: 15184398

14. R. Shalgi, J. A. Hurt, S. Lindquist, C. B. Burge, Widespread inhibition of posttranscriptional splicing shapes the cellular transcriptome following heat shock. Cell Rep. 7, 1362–1370 (2014). doi: 10.1016/j.celrep.2014.04.044; pmid: 24857664

15. K. Yamamoto et al., Control of the heat stress-induced alternative splicing of a subset of genes by hnRNP K. Genes Cells 21, 1006–1014 (2016). doi: 10.1111/gtc.12400; pmid: 27491955

16. A. K. Pearce, T. C. Humphrey, Integrating stress-response and cell-cycle checkpoint pathways. Trends Cell Biol. 11, 426–433 (2001). doi: 10.1016/S0962-8924(01)02119-5; pmid: 11567876

17. E. Radmaneshfar et al., From START to FINISH: The influence of osmotic stress on the cell cycle. PLOS ONE 8, e68067 (2013). doi: 10.1371/journal.pone.0068067; pmid: 23874495

18. N. M. Kühl, L. Rensing, Heat shock effects on cell cycle progression. Cell. Mol. Life Sci. 57, 450–463 (2000). doi: 10.1007/PL00000707; pmid: 10823246

19. V. Balagopal, R. Parker, Polysomes, P bodies and stress granules: States and fates of eukaryotic mRNAs. Curr. Opin. Cell Biol. 21, 403–408 (2009). doi: 10.1016/j.ceb.2009.03.005; pmid: 19394210

20. C. Jousse et al., Inhibition of a constitutive translation initiation factor 2α phosphatase, CReP, promotes survival of stressed cells. J. Cell Biol. 163, 767–775 (2003). doi: 10.1083/jcb.200308075; pmid: 14638860

21. I. Novoa, H. Zeng, H. P. Harding, D. Ron, Feedback inhibition of the unfolded protein response by GADD34-mediated dephosphorylation of eIF2α. J. Cell Biol. 153, 1011–1022 (2001). doi: 10.1083/jcb.153.5.1011; pmid: 11381086

22. J. R. Wheeler, T. Matheny, S. Jain, R. Abrisch, R. Parker, Distinct stages in stress granule assembly and disassembly. eLife 5, e18413 (2016). doi: 10.7554/eLife.18413; pmid: 27602576

23. X. Xie et al., Deubiquitylases USP5 and USP13 are recruited to and regulate heat-induced stress granules through their deubiquitylating activities. J. Cell Sci. 131, jcs210856 (2018). doi: 10.1242/jcs.210856; pmid: 29567855

24. S. Kwon, Y. Zhang, P. Matthias, The deacetylase HDAC6 is a novel critical component of stress granules involved in the stress response. Genes Dev. 21, 3381–3394 (2007). doi: 10.1101/gad.461107; pmid: 18079183

25. S. Markmiller et al., Context-dependent and disease-specific diversity in protein interactions within stress granules. Cell 172, 590–604.e13 (2018). doi: 10.1016/j.cell.2017.12.032; pmid: 29373831

26. J. Y. Youn et al., High-density proximity mapping reveals the subcellular organization of mRNA-associated granules and bodies. Mol. Cell 69, 517–532.e11 (2018). doi: 10.1016/j.molcel.2017.12.020; pmid: 29395067

27. N. Kedersha et al., G3BP–Caprin1–USP10 complexes mediate stress granule condensation and associate with 40S subunits. J. Cell Biol. 212, e201508028 (2016). doi: 10.1083/jcb.201508028; pmid: 27022092

28. S. Jain et al., ATPase-modulated stress granules contain a diverse proteome and substructure. Cell 164, 487–498 (2016). doi: 10.1016/j.cell.2015.12.038; pmid: 26777405

29. A. Turakhiya et al., ZFAND1 recruits p97 and the 26S proteasome to promote the clearance of arsenite-induced stress granules. Mol. Cell 70, 906–919.e7 (2018). doi: 10.1016/j.molcel.2018.04.021; pmid: 29804830

30. S. Alberti, D. Mateju, L. Mediani, S. Carra, Granulostasis: Protein quality control of RNP granules. Front. Mol. Neurosci. 10, 84 (2017). doi: 10.3389/fnmol.2017.00084; pmid: 28396624

31. M. Ganassi et al., A surveillance function of the HSPB8-BAG3-HSP70 chaperone complex ensures stress granule integrity and dynamism. Mol. Cell 63, 796–810 (2016). doi: 10.1016/j.molcel.2016.07.021; pmid: 27570075

32. J. R. Buchan, R. M. Kolaitis, J. P. Taylor, R. Parker, Eukaryotic stress granules are cleared by autophagy and Cdc48/VCP function. Cell 153, 1461–1474 (2013). doi: 10.1016/j.cell.2013.05.037; pmid: 23791177

33. B. Wang et al., ULK1 and ULK2 regulate stress granule disassembly through phosphorylation and activation of VCP/p97. Mol. Cell 74, 742–757.e8 (2019). doi: 10.1016/j.molcel.2019.03.027; pmid: 30975686

34. J. Keiten-Schmitz et al., The nuclear SUMO-targeted ubiquitin quality control network regulates the dynamics of cytoplasmic stress granules. Mol. Cell 79, 54–67.e7 (2020). doi: 10.1016/j.molcel.2020.05.017; pmid: 32521226

35. S. Markmiller et al., Active protein neddylation or ubiquitylation is dispensable for stress granule dynamics. Cell Rep. 27, 1356–1363.e3 (2019). doi: 10.1016/j.celrep.2019.04.015; pmid: 31042464

36. P. Beaudette, O. Popp, G. Dittmar, Proteomic techniques to probe the ubiquitin landscape. Proteomics 16, 273–287 (2016). doi: 10.1002/pmic.201500290; pmid: 26460060

37. A. Ordureau, C. Münch, J. W. Harper, Quantifying ubiquitin signaling. Mol. Cell 58, 660–676 (2015). doi: 10.1016/j.molcel.2015.02.020; pmid: 26000850

38. F. Lopitz-Otsoa et al., Integrative analysis of the ubiquitin proteome isolated using Tandem Ubiquitin Binding Entities (TUBEs). J. Proteomics 75, 2998–3014 (2012). doi: 10.1016/j.jprot.2011.12.001; pmid: 22178446

39. A. Fulzele, E. J. Bennett, Ubiquitin diGLY proteomics as an approach to identify and quantify the ubiquitin-modified proteome. Methods Mol. Biol. 1844, 363–384 (2018). doi: 10.1007/978-1-4939-8706-1_23; pmid: 30242721

40. N. N. Fang *et al.*, Rsp5/Nedd4 is the main ubiquitin ligase that targets cytosolic misfolded proteins following heat stress. *Nat. Cell Biol.* **16**, 1227–1237 (2014). doi: 10.1038/ncb3054; pmid: 25344756

41. N. N. Fang, M. Zhu, A. Rose, K. P. Wu, T. Mayor, Deubiquitinase activity is required for the proteasomal degradation of misfolded cytosolic proteins upon heat-stress. *Nat. Commun.* **7**, 12907 (2016). doi: 10.1038/ncomms12907; pmid: 27698423

42. G. Xu *et al.*, Identification of proteins sensitive to thermal stress in human neuroblastoma and glioma cell lines. *PLOS ONE* **7**, e49021 (2012). doi: 10.1371/journal.pone.0049021; pmid: 23145051

43. E. W. Wallace *et al.*, Reversible, specific, active aggregates of endogenous proteins assemble upon heat stress. *Cell* **162**, 1286–1298 (2015). doi: 10.1016/j.cell.2015.08.041; pmid: 26359986

44. X. Sui *et al.*, Widespread remodeling of proteome solubility in response to different protein homeostasis stresses. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 2422–2431 (2020). doi: 10.1073/pnas.1912897117; pmid: 31964829

45. A. H. Ng, N. N. Fang, S. A. Comyn, J. Gsponer, T. Mayor, System-wide analysis reveals intrinsically disordered proteins are prone to ubiquitylation after misfolding stress. *Mol. Cell. Proteomics* **12**, 2456–2467 (2013). doi: 10.1074/mcp.M112.023416; pmid: 23716602

46. L. Yuniati *et al.*, Tumor suppressor BTG1 promotes PRMT1-mediated ATF4 function in response to cellular stress. *Oncotarget* **7**, 3128–3143 (2016). doi: 10.18632/oncotarget.6519; pmid: 26657730

47. A. Vihervaara, F. M. Duarte, J. T. Lis, Molecular mechanisms driving transcriptional stress responses. *Nat. Rev. Genet.* **19**, 385–397 (2018). doi: 10.1038/s41576-018-0001-6; pmid: 29556092

48. K. R. Diller, Stress protein expression kinetics. *Annu. Rev. Biomed. Eng.* **8**, 403–424 (2006). doi: 10.1146/annurev.bioeng.7.060804.100449; pmid: 16834562

49. J. R. Subjeck, J. J. Sciandra, R. J. Johnson, Heat shock proteins and thermotolerance; a comparison of induction kinetics. *Br. J. Radiol.* **55**, 579–584 (1982). doi: 10.1259/0007-1285-55-656-579; pmid: 7116088

50. Y. Zhang *et al.*, The role of heat shock factors in stress-induced transcription. *Methods Mol. Biol.* **787**, 21–32 (2011). doi: 10.1007/978-1-61779-295-3_2; pmid: 21898224

51. G. Xu, J. S. Paige, S. R. Jaffrey, Global analysis of lysine ubiquitination by ubiquitin remnant immunoaffinity profiling. *Nat. Biotechnol.* **28**, 868–873 (2010). doi: 10.1038/nbt.1654; pmid: 20639865

52. C. M. Rose *et al.*, Highly multiplexed quantitative mass spectrometry analysis of ubiquitylomes. *Cell Syst.* **3**, 395–403.e4 (2016). doi: 10.1016/j.cels.2016.08.009; pmid: 27667366

53. N. Rauniyar, J. R. Yates III, Isobaric labeling-based relative quantification in shotgun proteomics. *J. Proteome Res.* **13**, 5293–5309 (2014). doi: 10.1021/pr500880b; pmid: 25337643

54. P. Leuenberger *et al.*, Cell-wide analysis of protein thermal unfolding reveals determinants of thermostability. *Science* **355**, eaai7825 (2017). doi: 10.1126/science.aai7825; pmid: 28232526

55. W. Huang, B. T. Sherman, R. A. Lempicki, Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009). doi: 10.1038/nprot.2008.211; pmid: 19131956

56. H. Mi *et al.*, PANTHER version 11: Expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res.* **45**, D183–D189 (2017). doi: 10.1093/nar/gkw1138; pmid: 27899595

57. J. Jiang *et al.*, CHIP is a U-box-dependent E3 ubiquitin ligase: Identification of Hsc70 as a target for ubiquitylation. *J. Biol. Chem.* **276**, 42938–42944 (2001). doi: 10.1074/jbc.M101968200; pmid: 11557750

58. M. Taipale *et al.*, A quantitative chaperone interaction network reveals the architecture of cellular protein homeostasis pathways. *Cell* **158**, 434–448 (2014). doi: 10.1016/j.cell.2014.05.039; pmid: 25036637

59. I. Paul, M. K. Ghosh, A CHIPotle in physiology and disease. *Int. J. Biochem. Cell Biol.* **58**, 37–52 (2015). doi: 10.1016/j.biocel.2014.10.027; pmid: 25448416

60. H. W. Hwang *et al.*, PAPERCLIP identifies microRNA targets and a role of CstF64/64tau in promoting non-canonical poly(A) site usage. *Cell Rep.* **15**, 423–435 (2016). doi: 10.1016/j.celrep.2016.03.023; pmid: 27050522

61. P. Yang *et al.*, G3BP1 is a tunable switch that triggers phase separation to assemble stress granules. *Cell* **181**, 325–345.e28 (2020). doi: 10.1016/j.cell.2020.03.046; pmid: 32302571

62. Y. Gwon *et al.*, Ubiquitination of G3BP1 mediates stress granule disassembly in a context-specific manner. *Science* **372**, eabf6548 (2021). doi: 10.1126/science.abf6548

63. M. L. Hyer *et al.*, A small-molecule inhibitor of the ubiquitin activating enzyme for cancer treatment. *Nat. Med.* **24**, 186–193 (2018). doi: 10.1038/nm.4474; pmid: 29334375

64. N. Kedersha *et al.*, Dynamic shuttling of TIA-1 accompanies the recruitment of mRNA to mammalian stress granules. *J. Cell Biol.* **151**, 1257–1268 (2000). doi: 10.1083/jcb.151.6.1257; pmid: 11121440

65. R. Panniers, Translational control during heat shock. *Biochimie* **76**, 737–747 (1994). doi: 10.1016/0300-9084(94)90078-7; pmid: 7893824

66. R. Panniers, E. C. Henshaw, Mechanism of inhibition of polypeptide chain initiation in heat-shocked Ehrlich ascites tumour cells. *Eur. J. Biochem.* **140**, 209–214 (1984). doi: 10.1111/j.1432-1033.1984.tb08088.x; pmid: 6705795

67. E. K. Schmidt, G. Clavarino, M. Ceppi, P. Pierre, SUnSET, a nonradioactive method to monitor protein synthesis. *Nat. Methods* **6**, 275–277 (2009). doi: 10.1038/nmeth.1314; pmid: 19305406

68. M. Weith *et al.*, Ubiquitin-independent disassembly by a p97 AAA-ATPase complex drives PP1 holoenzyme formation. *Mol. Cell* **72**, 766–777.e6 (2018). doi: 10.1016/j.molcel.2018.09.020; pmid: 30344098

69. N. Yamashita *et al.*, Regulation of spine development by semaphorin3A through cyclin-dependent kinase 5 phosphorylation of collapsin response mediator protein 1. *J. Neurosci.* **27**, 12546–12554 (2007). doi: 10.1523/JNEUROSCI.3463-07.2007; pmid: 18003833

70. C. Wang *et al.*, Scalable production of iPSC-derived human neurons to identify tau-lowering compounds by high-content screening. *Stem Cell Reports* **9**, 1221–1233 (2017). doi: 10.1016/j.stemcr.2017.08.019; pmid: 28966121

71. P. Xu, D. M. Duong, J. Peng, Systematical optimization of reverse-phase chromatography for shotgun proteomics. *J. Proteome Res.* **8**, 3944–3950 (2009). doi: 10.1021/pr900251d; pmid: 19566079

72. B. Bai *et al.*, Deep profiling of proteome and phosphoproteome by isobaric labeling, extensive liquid chromatography, and mass spectrometry. *Methods Enzymol.* **585**, 377–395 (2017). doi: 10.1016/bs.mie.2016.10.007; pmid: 28109439

CELL BIOLOGY

Ubiquitination of G3BP1 mediates stress granule disassembly in a context-specific manner

Youngdae Gwon, Brian A. Maxwell, Regina-Maria Kolaitis, Peipei Zhang, Hong Joo Kim, J. Paul Taylor*

INTRODUCTION: Stress granules are dynamic structures composed of RNA and protein that arise in the cytoplasm in response to a variety of stressors. These structures form via liquid-liquid phase separation and are usually promptly disassembled after the initiating stress is relieved. Stress granules assemble when the sum of protein-protein, protein-RNA, and RNA-RNA interactions breaches a particular threshold known as the percolation threshold. When this threshold is crossed, individual protein and RNA molecules form a system-spanning network that separates itself from its milieu to form a distinct granule. When the network of interactions falls below the percolation threshold, the granule disassembles. For stress granules, G3BP1 and G3BP2 are the proteins that pro-vide the largest contribution to establishing the percolation threshold for granule assembly.

RATIONALE: Whereas great progress has been made in understanding the molecular basis of stress granule assembly, little is known about the mechanisms that govern their elimination. This process is of particular interest given that many mutations that cause neurodegeneration lead to impaired clearance of stress granules. The present investigation into the mechanisms of stress granule disassembly was prompted by findings in an accompanying report that G3BP1 is ubiquitinated when stress granules assemble in response to heat shock, but not when cells are exposed to other types of granule-inducing stress.

RESULTS: We found that stress granule elimination in cultured human cells is accomplished via one of two possible pathways: autophagy-independent disassembly or autophagy-dependent degradation. The fate of stress granules depended on how long they remained assembled. Persistent stress granules, such as those that arise through chronic stress or disease mutations, were eliminated by autophagy-dependent degradation. In contrast, short-lived granules were rapidly disassembled in an autophagy-independent manner, which permits recycling of constituents. Moreover, the mechanism whereby stress granules are disassembled depended upon the nature of the initiating stress. We identified the molecular mechanism of disassembly of stress granules induced by heat stress and found that this ubiquitin-dependent mechanism is specific to heat shock and not other types of stress. In response to heat shock, polyubiquitination of G3BP1 enabled binding by FAF2, an endoplasmic reticulum (ER)–associated adaptor for the ubiquitin-dependent segregase p97/VCP. Subsequent extraction of G3BP1 from stress granules by VCP leads to their disassembly. Disease-causing mutations in VCP impaired this disassembly mechanism.

CONCLUSION: We found that the fate of stress granules and the mechanism of their elimination depends on the context in which they were formed and the duration of their assembly. In the setting of heat shock, ubiquitination and subsequent removal of G3BP1 reduce the driving forces within the stress granule network below the percolation threshold for phase separation. Furthermore, the localization of FAF2 in the ER membrane indicates that stress granules that arise in response to heat stress are disassembled at the ER, consistent with other heat shock–dependent stress responses such as the unfolded protein response and ER-associated degradation. Disease-causing mutations in VCP not only impair autophagy-dependent stress granule degradation, as previously reported, but also impair autophagy-independent disassembly. Thus, mutations in VCP represent a m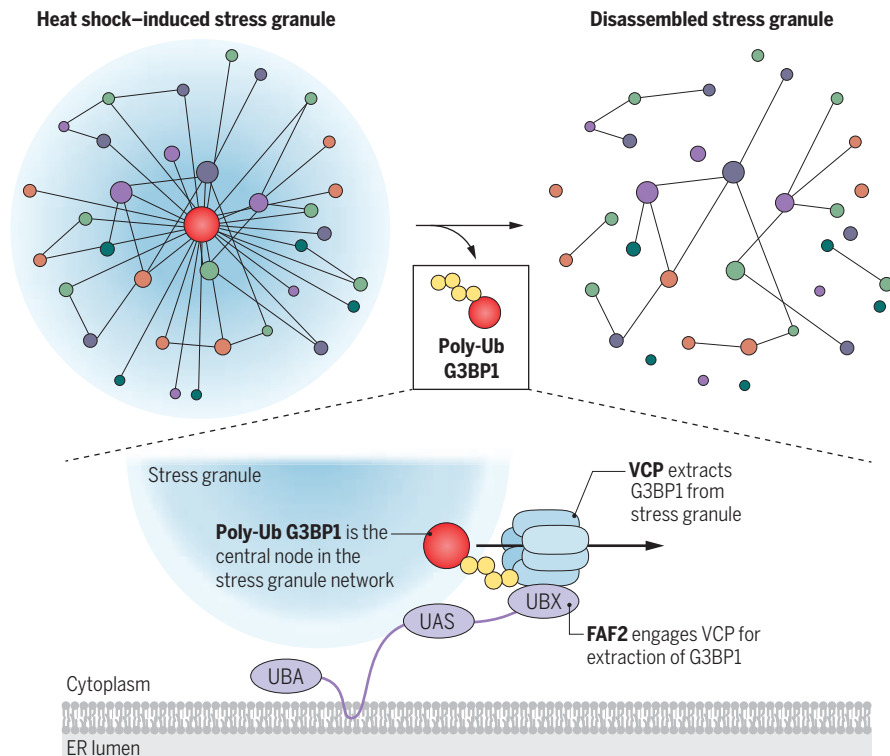echanistic link between neurodegeneration and aberrant phase transitions. This finding aligns with other disease-causing mutations (e.g., mutations in intrinsically disordered regions of RNA-binding proteins, hexanucleotide repeat expansions in *C9ORF72*) that similarly impinge on intracellular phase transitions. ■

Heat shock–induced stress granules disassemble when G3BP1 is polyubiquitinated. Ubiquitinated G3BP1 interacts with the ER-associated protein FAF2, which engages the ubiquitin-dependent segregase VCP. Targeting G3BP1 enables the stress granule–specific interaction network to fall below the percolation threshold for phase separation, which causes disassembly.

S READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abf6548

CELL BIOLOGY

Ubiquitination of G3BP1 mediates stress granule disassembly in a context-specific manner

Youngdae Gwon[1], Brian A. Maxwell[1], Regina-Maria Kolaitis[1], Peipei Zhang[1], Hong Joo Kim[1], J. Paul Taylor[1,2]*

Stress granules are dynamic, reversible condensates composed of RNA and protein that assemble in eukaryotic cells in response to a variety of stressors and are normally disassembled after stress is removed. The composition and assembly of stress granules is well understood, but little is known about the mechanisms that govern disassembly. Impaired disassembly has been implicated in some diseases including amyotrophic lateral sclerosis, frontotemporal dementia, and multisystem proteinopathy. Using cultured human cells, we found that stress granule disassembly was context-dependent: Specifically in the setting of heat shock, disassembly required ubiquitination of G3BP1, the central protein within the stress granule RNA-protein network. We found that ubiquitinated G3BP1 interacted with the endoplasmic reticulum–associated protein FAF2, which engaged the ubiquitin-dependent segregase p97/VCP (valosin-containing protein). Thus, targeting of G3BP1 weakened the stress granule–specific interaction network, resulting in granule disassembly.

iomolecular condensation is a vital strategy of cellular organization that regulates a variety of biological functions. Ribonucleoprotein (RNP) granules are a highly conserved class of biomolecular condensates that govern many aspects of RNA metabolism (*1*, *2*). One prominent type of RNP granule is the stress granule, a dynamic and reversible cytoplasmic assembly formed in eukaryotic cells in response to a variety of stressors. Formation of stress granules typically follows upon inhibited translation initiation and polysome disassembly, which causes a rapid increase in the cytoplasmic concentration of uncoated mRNA. This rise in cytoplasmic mRNA triggers multicomponent liquid-liquid phase separation (LLPS) with RNA-binding proteins, creating a condensed liquid compartment that remains in dynamic equilibrium with, yet distinct from, the surrounding cytosol (*3*–*5*). Impaired dynamics of RNP granules such as stress granules are implicated as an important contributor to certain pathological conditions, including neurodegenerative diseases (*6*).

Recent investigations of stress granule assembly have yielded insight into how multicomponent networks self-organize to form a compartment that is distinct from its surroundings. Stress granules are composed of RNA and protein that interact via protein-protein, protein-RNA, and RNA-RNA interactions, many of which may be weak and transient. However, once the sum of these interactions breaches a particular threshold, known as the percolation threshold (*6*), the individual molecules form a system-spanning network that separates itself from its milieu—in other words, LLPS ensues (*3*–*5*). The concept of a percolation threshold is generalizable to other biomolecular condensates, each of which encodes a system-specific threshold for LLPS.

In U2OS cell stress granules, there are ~36 proteins that, together with RNA, provide the majority of the interactions that set the percolation threshold for RNA-dependent LLPS (*3*). Whereas each constituent node of this network contributes toward the sum of interactions required to reach the percolation threshold, a small subset of constituents of high centrality within the interaction network predominately establish this threshold (*3*, *7*). The most important proteins in the stress granule interaction network are G3BP1 and G3BP2, which provide the largest contribution to establishing the percolation threshold for stress granule assembly (*3*–*5*). Indeed, elimination of G3BP1/2 proteins can preclude stress granule assembly (*3*, *8*) whereas their enforced activation by optogenetic induction of dimerization can initiate stress granule assembly (*3*, *9*). Higher-order regulation over stress granule assembly may also be afforded by posttranslational modifications of stress granule proteins that affect the interaction network (*10*).

In healthy cells, stress granules are transient, dynamic structures whose presence generally corresponds to the duration of time during which a stress is applied. However, exactly how stress granules are removed from cells remains unknown. Specifically, the role of autophagy in the elimination of stress granules is unclear:

Whereas some studies have shown that stress granule removal is autophagy-dependent (*11*), autophagy-independent elimination has also been reported (*12*). Autophagy independence is consistent with evidence suggesting that stress granule constituents are recycled, with mRNPs reentering the translational pool after removal of stress (*13*).

Insight into reconciling these apparent contradictions may come from studies examining the role of context in defining the mechanisms of stress granule assembly and disassembly. As shown in our companion paper (*14*), stress granules formed by heat stress are disassembled via ubiquitin-dependent mechanisms, whereas those formed in response to arsenite are not. Thus, the key nodes and cellular processes engaged in disassembly of stress granules are specific to the context of each stress. This finding may explain why ubiquitination is dispensable for arsenite-induced stress granule dynamics (*15*), whereas the ubiquitin-binding adaptor protein ZFAND1 is required (*16*). Here, we wanted to ascertain the extent to which the mechanisms involved in stress granule elimination depend on the nature of the initiating stress.

G3BP1 undergoes K63-linked ubiquitination upon heat shock

In our companion paper, we found that G3BP1 is ubiquitinated when stress granules are assembled in response to heat shock, but not when cells are exposed to other types of granule-inducing stress (*14*). Furthermore, we found that ubiquitination was not required for heat shock–induced stress granule assembly but was essential for their rapid disassembly after removal of stress. However, the mechanism whereby ubiquitination was required for stress granule disassembly remained unclear (*14*).

In addition to its essential role in stress granule formation, G3BP1 is required to maintain the assembly of stress granules, as pharmacological disruption of G3BP1 dimers leads to rapid disassembly of stress granules (*3*). Thus, we hypothesized that ubiquitination and targeted degradation of G3BP1 may have a key role in disassembly.

To confirm that G3BP1 is ubiquitinated in response to heat shock, we performed tandem-ubiquitin binding entity (TUBE) pulldown of ubiquitinated proteins from U2OS cells exposed to heat shock (43°C) followed by immunoblotting (Fig. 1A), using *G3BP1/2* double knockout (dKO) cells as controls. We detected increased levels of polyubiquitin-conjugated G3BP1 as early as 15 min after exposure to heat stress; with continuing stress, these levels peaked at ~60 min and remained detectable in the soluble fraction for at least 150 min (Fig. 1B). After stress was removed, levels of polyubiquitinated G3BP1 decreased, returning to baseline within 3 hours after a 60-min heat

[1]Department of Cell and Molecular Biology, St. Jude Children's Research Hospital, Memphis, TN, USA. [2]Howard Hughes Medical Institute, St. Jude Children's Research Hospital, Memphis, TN, USA.
*Corresponding author. Email: jpaul.taylor@stjude.org
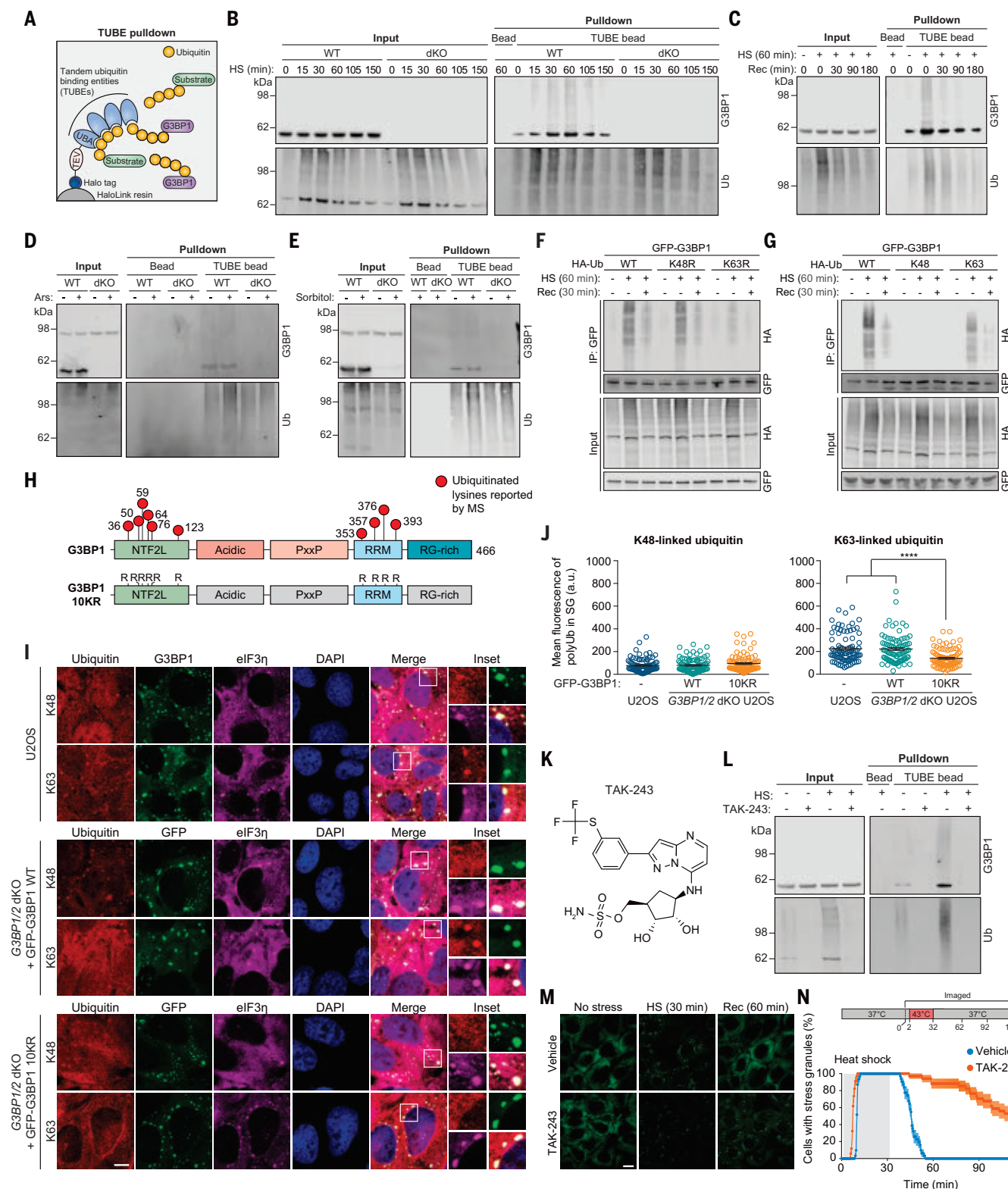
Fig. 1. G3BP1 undergoes K63-linked ubiquitination in response to heat stress. (**A**) Illustration showing TUBE capture of ubiquitinated G3BP1. TEV, tobacco etch virus protease cleavage site. (**B** and **C**) Immunoblots of TUBE-captured cell extracts showing levels of ubiquitinated G3BP1 in U2OS cells after different durations of 43°C heat shock, using U2OS *G3BP1/2* dKO cells as controls (B), and during 37°C recovery (C). (**D** and **E**) Immunoblots of TUBE-captured cell extracts showing levels of ubiquitinated G3BP1 in response to

oxidative stress (0.5 mM $NaAsO_2$, 1 hour) (D) or osmotic stress (0.4 M sorbitol, 1 hour) (E). (**F** and **G**) Immunoblots of cell extracts captured with antibody to GFP, showing K63-linked ubiquitination of G3BP1. Transfected HEK293T cells were exposed to heat shock (43°C, 1 hour) and recovery (37°C, 30 min). K48R and K63R prevent the formation of K48-linked (K48R) or K63-linked (K63R) chains; K48 and K63 permit K48-linked or K63-linked chains exclusively. HA, hemagglutinin; IP, immunoprecipitate. (**H**) G3BP1 domain labeled with lysines on

which ubiquitination has been reported. In the G3BP1 10KR mutant, six lysines in NTF2L and four lysines in RRM are mutated to arginine. (**I**) Immunofluorescent staining of fixed U2OS WT and U2OS *G3BP1/2* dKO cells stably expressing GFP-G3BP WT and 10KR. Scale bar, 10 μm. (**J**) Fluorescence intensities of K48- and K63-linked ubiquitin in eIF3η-positive stress granules from three technical replicates are plotted in (n = 90). Error bars indicate SEM. ****P < 0.0001 (ANOVA with Tukey's test). (**K**) Structure of TAK-243. (**L**) Immunoblot of TUBE-captured cell extracts showing block of heat shock–induced G3BP1

ubiquitination by TAK-243. U2OS cells were treated with DMSO or TAK-243 for 1 hour prior to heat shock. (**M** and **N**) Fluorescent imaging of U2OS cells stably expressing GFP-G3BP1 were treated with DMSO or TAK-243 (1 hour) prior to imaging. Representative images are shown in (M). In (N), GFP signals were monitored at 30-s intervals to count cells with two or more stress granules from three technical replicates (vehicle n = 32, TAK-243 n = 35). Scale bar, 20 μm. Error bars indicate SEM. ****P < 0.0001 (Mantel-Cox test).

---

stress (Fig. 1C). Polyubiquitin conjugation of G3BP1 was specific to heat stress: Despite an increase in total polyubiquitin conjugates in response to oxidative or osmotic stress, cells exposed to these stresses did not accumulate polyubiquitinated G3BP1 (Fig. 1, D and E).

Polyubiquitin chains can be formed by conjugating ubiquitin to one of seven lysine (K) residues of another ubiquitin molecule, giving rise to K6-, K11-, K27-, K29-, K33-, K48-, or K63-linked polyubiquitin chains (17). Among these, K48- and K63-linked chains are the most abundant and functionally well-characterized linkage types (18). To determine which of these linkage types were present in polyubiquitinated G3BP1, we used ubiquitin mutants that prevent the formation of K48-linked (K48R) or K63-linked (K63R) chains or that permit K48-linked or K63-linked chains exclusively (K48 or K63 being the only available lysines). Expression of K63R inhibited accumulation of polyubiquitin-conjugated G3BP1 upon heat shock, whereas expression of K48R had no impact (Fig. 1F). Conversely, cells expressing K63 ubiquitin (in which the only available lysine is K63) accumulated polyubiquitin-conjugated G3BP1, whereas cells expressing K48 ubiquitin did not (Fig. 1G), which suggests that heat shock promotes K63-linked polyubiquitination of G3BP1. The levels of K63-linked polyubiquitinated G3BP1 were lower than the levels of polyubiquitinated G3BP1 using wild-type ubiquitin; therefore, other linkage types may also contribute to this polyubiquitination.

Heat shock–induced stress granules accumulate K63-linked polyubiquitin chains that are required for stress granule disassembly

Examination of public databases suggests 10 lysine residues in G3BP1 that are potentially subject to ubiquitination (19). Of these, six residues are located within the N-terminal NTF2-like (NTF2L) domain and four in the C-terminal RNA-recognition motif (RRM) (Fig. 1H). When we mutated these 10 residues (G3BP1 10KR) and expressed this construct in *G3BP1/2* dKO cells, we found no ubiquitination of G3BP1 in response to heat shock, which suggests that G3BP1 10KR functions as a ubiquitination null mutant (fig. S1A).

Stress granules induced by heat stress (but not arsenite stress) exhibit a robust accumu-

lation of polyubiquitin signal (14). To examine the contributions of K48- versus K63-linked chains to this ubiquitin signal within stress granules, we used linkage-specific antibodies. Whereas K48- and K63-linked polyubiquitin signals were both uniformly distributed throughout the cell under basal conditions (fig. S1B), only K63-linked polyubiquitin signals accumulated in stress granules upon heat shock in wild-type U2OS cells (Fig. 1I). Furthermore, K63-linked polyubiquitin signal was significantly diminished in *G3BP1/2* dKO cells expressing G3BP1 10KR fused to green fluorescent protein (GFP-G3BP1 10KR), whereas *G3BP1/2* dKO cells expressing wild-type GFP-G3BP1 (GFP-G3BP1 WT) accumulated K63-linked polyubiquitin signal in stress granules at levels comparable to those observed in wild-type U2OS cells (Fig. 1, I and J). Thus, a substantial portion of heat shock–induced polyubiquitin signal in stress granules arises from K63-linked polyubiquitination of G3BP1. Polyubiquitin conjugation of G3BP1 was eliminated by treatment with TAK-243, a small-molecule inhibitor of UBA1, an E1 ubiquitin-activating enzyme (Fig. 1, K and L). Consistent with our previous observations (14), blocking ubiquitination by TAK-243 significantly delayed stress granule disassembly during recovery from heat stress (Fig. 1, M and N, and movie S1). Thus, ubiquitination is required for stress granule disassembly after heat shock. Similar results were obtained upon small interfering RNA (siRNA)–mediated knockdown of *UBA1* (14). We next sought to investigate the relationship between polyubiquitin conjugation of G3BP1 and stress granule dynamics.

Ubiquitination within G3BP1 NTF2L is required for disassembly of heat shock–induced stress granules

To investigate the effect of G3BP1 ubiquitination on stress granule dynamics, we first sought to identify the heat shock–dependent ubiquitination site(s) within G3BP1. To this end, we generated constructs in which six lysine residues in the NTF2L domain (6KR) or four lysine residues in the RRM domain (4KR) were mutated to arginines (fig. S2A). When expressed in *G3BP1/2* dKO cells, G3BP1 6KR showed greatly reduced ubiquitination in response to heat shock, whereas G3BP1 4KR showed no appar-

ent change in polyubiquitin conjugation; this result implies that the NTF2L domain is preferentially ubiquitinated upon heat shock (Fig. 2A). The crystal structure of the NTF2L dimer (PDB: 5FW5) indicates that all six lysine residues are surface-exposed. K36, K50, K59, and K64 are clustered to form a positively charged surface on the lateral sides of homodimeric NTF2L domains (Fig. 2B); K50 and K76 from each monomer lie adjacent to the dimeric interface; and K123 is located near the hydrophobic groove between α helices, where the nsP3 protein of Old World alphaviruses binds to inhibit stress granule assembly (fig. S2B). To gain insight into which lysines in G3BP1 are ubiquitinated, we generated an additional series of Lys → Arg (K-to-R) mutations. Mutation of the four clustered lysine residues (K36/50/59/64R) nearly abolished heat shock–dependent ubiquitination of G3BP1 in *G3BP1/2* dKO cells, whereas other combinations of mutations to NTF2L lysines had no strong effect (Fig. 2C). However, ubiquitination levels further decreased in the 6KR mutant, and we therefore used G3BP1 6KR as a ubiquitination-deficient mutant. Ubiquitination of G3BP1 in this region is consistent with results from an independent approach (14). Specifically, paired di-Gly and tandem mass tag (TMT) analysis revealed enrichment of ubiquitinated G3BP1 peptide containing K50 [K.NSSYVHGGLDSNGK^{50}PADAVYGQK.E (20)] upon heat shock and subsequent decrease in this ubiquitinated species during recovery from stress (Fig. 2D). The decrease in ubiquitinated G3BP1 during recovery appears to reflect proteasome-dependent degradation, because treatment with the proteasome inhibitor bortezomib (Btz) led to accumulation of ubiquitinated G3BP1 (Fig. 2D).

We next sought to determine the role of G3BP1 ubiquitination in stress granule dynamics, including assembly, disassembly, and dynamic exchange between the stress granule and surrounding cytoplasm. To this end, we first generated stable cell lines in which G3BP1 6KR, 4KR, or 10KR were reintroduced into *G3BP1/2* dKO cells at levels comparable to endogenous G3BP1 levels in wild-type cells (fig. S2, C and D). To rule out the possibility that the K-to-R mutations impair intrinsic properties of G3BP1, we performed a series of control experiments. First, we assessed the
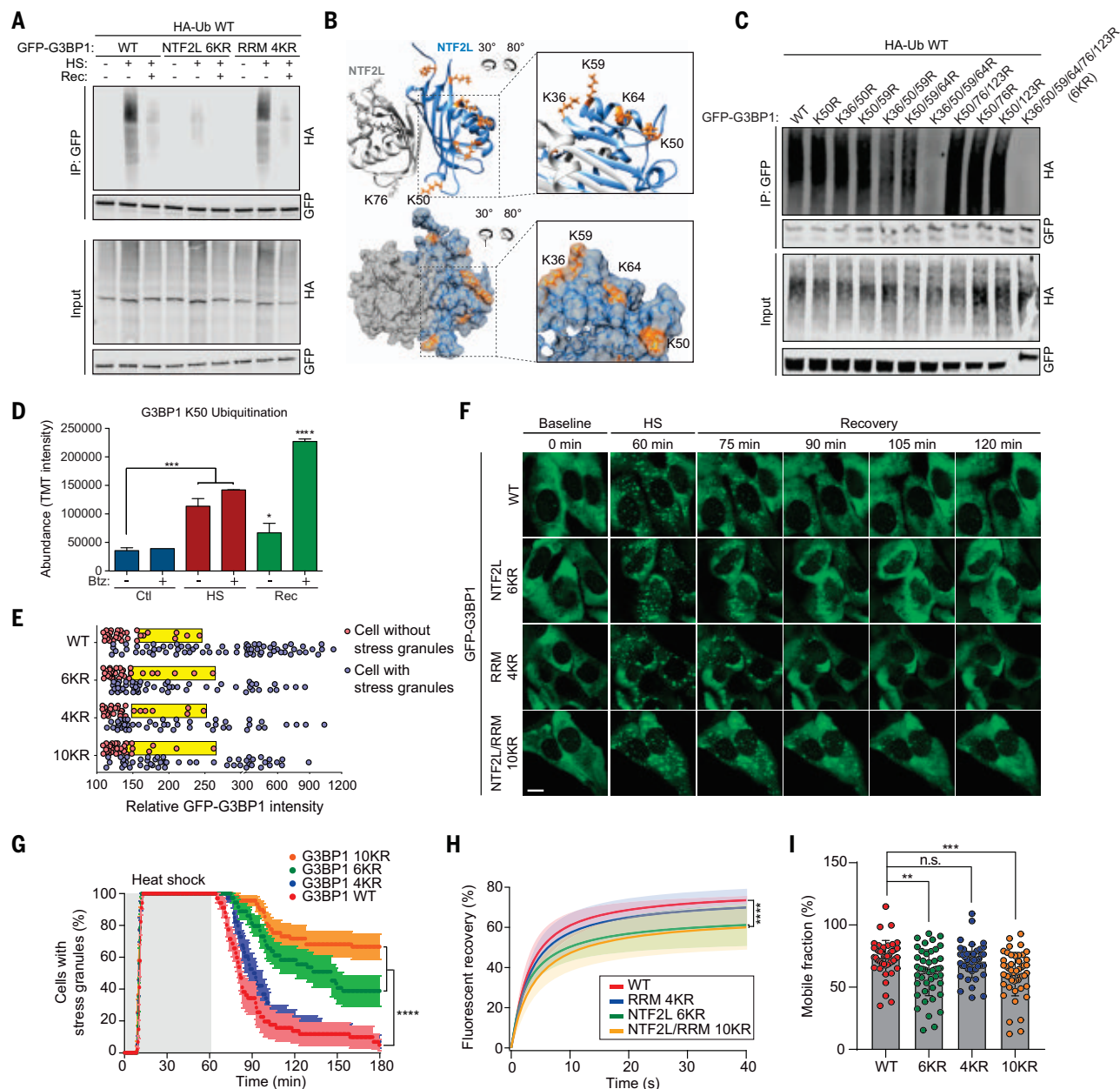
Fig. 2. Ubiquitination of the NTF2L domain of G3BP1 is required for the disassembly of stress granules. (**A**) Immunoblot of transfected HEK293T cell extracts captured with antibody to GFP after exposure to no stress, heat shock for 1 hour, or 37°C recovery for 30 min. (**B**) Dimeric structure of G3BP1 NTF2L domain. K36, K50, K59, and K64 are highlighted in enlarged images (orange). (**C**) Immunoblot of transfected HEK293T cell extract captured with antibody to GFP after exposure to heat shock for 1 hour. (**D**) TMT intensity of ubiquitinated G3BP1 peptides during heat shock and recovery with or without bortezomib (Btz). Error bars indicate SD. *$P < 0.05$, ***$P < 0.001$, ****$P < 0.0001$ (ANOVA with Dunnett's test). (**E**) Percolation threshold of GFP-G3BP1 WT, 6KR, 4KR, and 10KR exposed to heat shock for 1 hour. U2OS *G3BP1/2* dKO cells without stress granules (red) and with stress granules (blue) are plotted from three biological replicates according to GFP intensities (WT $n = 109$, 6KR $n = 86$, 4KR $n = 78$, 10KR $n = 85$). Yellow boxes indicate the cells with 25%

highest levels of GFP within the stress granule–null group. (**F**) Stress granule disassembly in U2OS *G3BP1/2* dKO cells expressing G3BP1 WT, 6KR, 4KR, or 10KR. Scale bar, 20 μm. (**G**) GFP signals were monitored at 30-s intervals at 37°C for 2 min, 43°C for 60 min, and 37°C for 118 min to count cells with two or more stress granules from three biological replicates (WT $n = 38$, 6KR $n = 36$, 4KR $n = 39$, 10KR $n = 39$). Error bars indicate SEM. ****$P < 0.0001$ (Mantel-Cox test). (**H**) FRAP of GFP-positive puncta in U2OS *G3BP1/2* dKO cells transfected with GFP-G3BP1 WT, 6KR, 4KR, or 10KR and exposed to heat shock (1 hour) followed by recovery (10 min) from three biological replicates is plotted as mean ± SEM (WT $n = 48$, 6KR $n = 48$, 4KR $n = 45$, 10KR $n = 49$). ****$P < 0.0001$ (ANOVA with Tukey's test). (**I**) Quantification of mobile fraction at 40 s of recovery from three biological replicates (WT $n = 48$, 6KR $n = 48$, 4KR $n = 45$, 10KR $n = 49$). Error bars indicate SD. **$P < 0.01$, ***$P < 0.001$ (ANOVA with Dunnett's test); n.s., not significant.

Fig. 3. VCP interacts with ubiquitinated G3BP1 in response to heat shock.
(**A**) Fluorescent imaging of U2OS *G3BP1/2* dKO cells stably expressing GFP-G3BP WT, 6KR, 4KR, or 10KR exposed to heat shock for 1 hour. Scale bar, 20 μm. (**B**) Fluorescence intensities of VCP in eIF3η-positive stress granules from three technical replicates (*n* = 90). Error bars indicate SEM. ****$P < 0.0001$ (ANOVA with Tukey's test). (**C**) Immunoblot of U2OS cell extracts exposed to heat shock for the indicated times. Cell extracts were captured with ma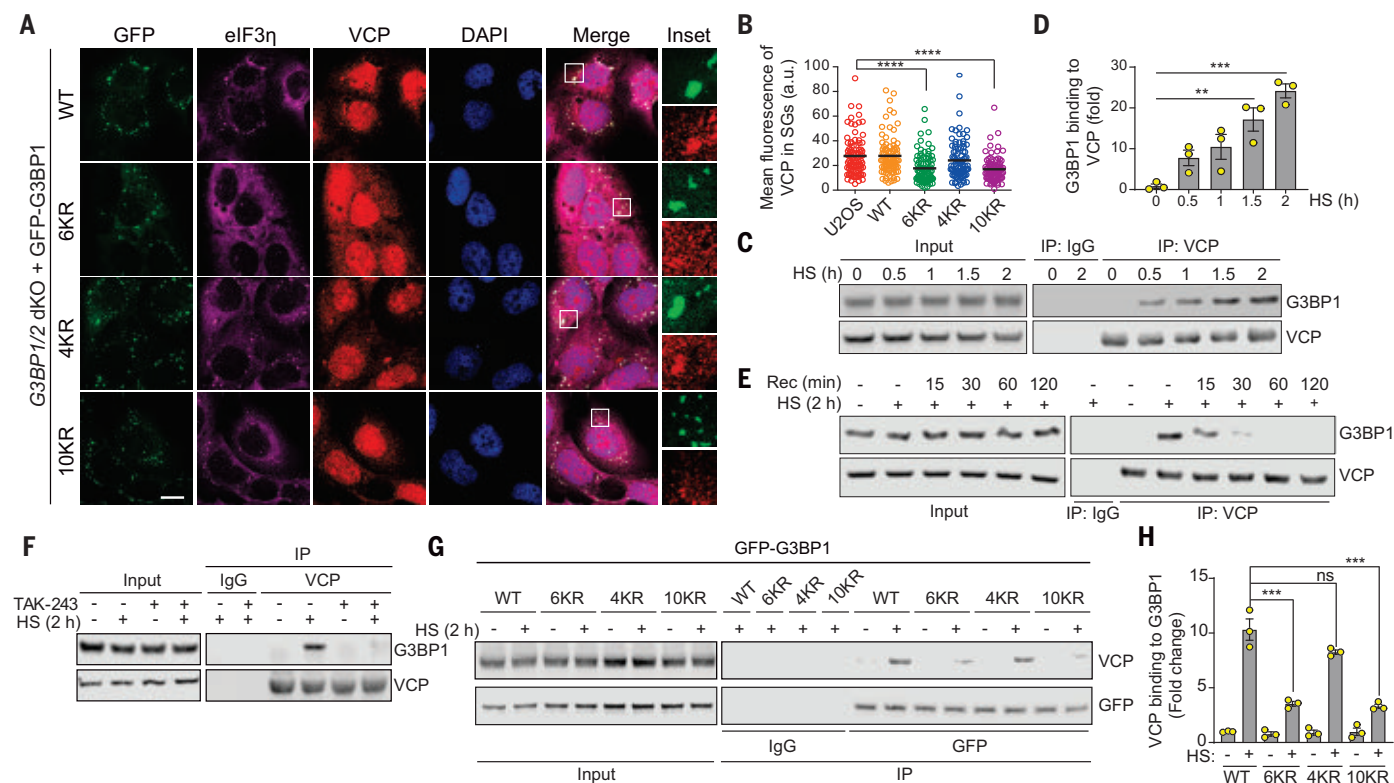gnetic beads conjugated with VCP antibody for IP, and resulting beads were analyzed by immunoblot. (**D**) Quantification of immunoblots from three biological replicates. Error bars indicate SEM. **$P < 0.01$, ***$P < 0.001$ (ANOVA with Tukey's test). (**E**) Immunoblot showing G3BP1-VCP interaction in U2OS cells based on duration of recovery after heat shock for 2 hours. Cell extracts were captured with magnetic beads conjugated with antibody to VCP for IP, and the resulting beads were analyzed by immunoblot. (**F**) Immunoblot of U2OS cells treated with DMSO or TAK-243 (60 min) and exposed to heat shock for 2 hours. Cell extracts were captured with magnetic beads conjugated with antibody to VCP for IP. Blots show an attenuated G3BP1-VCP interaction with inhibition of ubiquitination. (**G**) Immunoblot of U2OS *G3BP1/2* dKO cells stably expressing G3BP1 WT, 6KR, 4KR, or 10KR and exposed to heat shock for 2 hours. Cell extracts were captured with magnetic beads conjugated with antibody to GFP for IP. (**H**) Quantification of immunoblots from three biological replicates. Error bars indicate SEM. ***$P < 0.001$ (ANOVA with Tukey's test).

impact of K-to-R mutations on G3BP1 dimerization and binding to interaction partners. By coimmunoprecipitation, we confirmed that GFP-G3BP1 WT, 6KR, 4KR, and 10KR mutants formed dimers with endogenous G3BP1 WT (fig. S2E). Moreover, 6KR mutations in the NTF2L domain did not alter G3BP's ability to interact with caprin 1 and USP10 through this domain (fig. S2F), nor did 4KR mutations in the RRM domain alter its ability to bind RNA (fig. S2G), which suggested that the intrinsic properties of G3BP1 were not compromised by K-to-R substitution.

The most sensitive test of perturbation to G3BP1 function is the concentration threshold needed to trigger stress granule assembly in cells—the percolation threshold. To test this, we transiently transfected G3BP1 mutants into *G3BP1/2* dKO cells and measured the G3BP1 concentration threshold at which stress granule assembly was initiated (*3*). The concentration threshold for stress granule assembly for the 6KR, 4KR, and 10KR mutants was unchanged from G3BP1 WT (Fig. 2E). Thus, the K-to-R mutants block G3BP1 ubiquitination but do not otherwise impair intrinsic phase separation properties or interactions that support stress granule assembly.

We next assessed how preventing ubiquitination of G3BP1 affected stress granule dynamics. Upon heat shock, all three mutant forms of G3BP1 assembled stress granules with kinetics identical to those of G3BP1 WT (Fig. 2, F and G, and movie S2), indicating that the assembly of stress granules is independent of ubiquitination of G3BP1. This result is consistent with our findings that global inhibition of ubiquitination with TAK-243 did not alter the dynamics of stress granule assembly in response to heat shock (Fig. 1, M and N) (*14*). However, 6KR and 10KR mutations resulted in significantly prolonged stress granule disassembly (Fig. 2, F and G, and movie S2). G3BP1 10KR (KR mutations in NTF2L and RRM domains) had a more severe impact on stress granule disassembly than did G3BP1 6KR (KR mutations in NTF2L only); hence, the four lysine residues in the RRM domain may also contribute to stress granule disassembly, although the levels of polyubiquitination in the RRM domain were below the detection limit by immunoblotting. When we assessed G3BP1 dynamics within stress granules during heat shock by fluorescent recovery after photobleaching (FRAP) analysis, when stress granules were fully assembled, all three mutant forms of G3BP1 showed fluorescence recovery at a rate identical to that of G3BP1 WT (fig. S2H). In contrast, after heat stress was removed and the disassembly phase had begun, the G3BP1 6KR and 10KR mutants showed significantly slower recovery rates and greater immobile fractions relative to wild-type and 4KR proteins (Fig. 2, H and I). Thus, mutations impeding ubiquitination of NTF2L caused a

Fig. 4. VCP regulates disassembly of stress granules. (**A** to **C**) Top: U2OS/GFP-G3BP1 cells were treated with DMSO or bafilomycin A1 (BafA1) for 18 hours before live imaging. Scale bars, 20 μm. Bottom: GFP signals were monitored at 30-s intervals at 37°C for 2 min, 43°C for 30 min (A), 60 min (B), or 90 min (C), and 37°C for 88 min (A), 118 min (B), or 148 min (C) to count cells with two or more stress granules from three technical replicates [DMSO n = 56, BafA1 n = 58 in (A); DMSO n = 56, BafA1 n = 58 in (B); DMSO n = 52, BafA1 n = 59 in (C)]. Error bars indicate SEM. ****P < 0.0001 (Mantel-Cox test). (**D** and **E**) U2OS/GFP-G3BP1 cells were treated with DMSO or CB-5083 (1 hour) (D) or transfected with nontargeting (NT) or VCP siRNA (E) before live imaging. GFP signals were monitored at 30-s intervals at 37°C for 2 min, 43°C for 60 min, and 37°C for 118 min to count the cells with two or more stress granules [DMSO n = 40, CB-5083 n = 47 in (D); NT siRNA n = 44, VCP siRNA n = 53 in (E)]. Scale bars, 50 μm (D), 20 μm (E). Error bars indicate SEM. ****P < 0.0001 (Mantel-Cox test). (**F**) Immunoblot of U2OS cells transfected with NT or VCP siRNA from three biological replicates. Error bars indicate SEM. **P < 0.01 (Student's t test). (**G**) U2OS/GFP-G3BP1 cells were transfected with pmCherry-N1, VCP WT-mCherry, VCP R155H-mCherry, or VCP A232E-mCherry. Scale bar, 10 μm. (**H**) GFP signals of mCherry-positive cells were monitored at 60-s intervals at 37°C for 2 min, 43°C for 60 min, and 37°C for

88 min to count cells with two or more stress granules from three biological replicates (control n = 23, VCP WT n = 29, VCP R155H n = 26, VCP A232E n = 39). Error bars indicate SEM. **P < 0.01 (Mantel-Cox test). (**I**) U2OS cells expressing photoactivatable G3BP1 (G3BP1-PAGFP) were transfected with NT or VCP siRNA. (**J**) Intensities of GFP signals within activated ROIs were monitored at 200-ms intervals from three technical replicates. ****P < 0.0001 (ANOVA with Sidak's test). (**K**) Immunoblots of U2OS cells expressing G3BP1-PAGFP transfected with NT or VCP siRNA from three biological replicates. ***P < 0.001 (Student's t test).
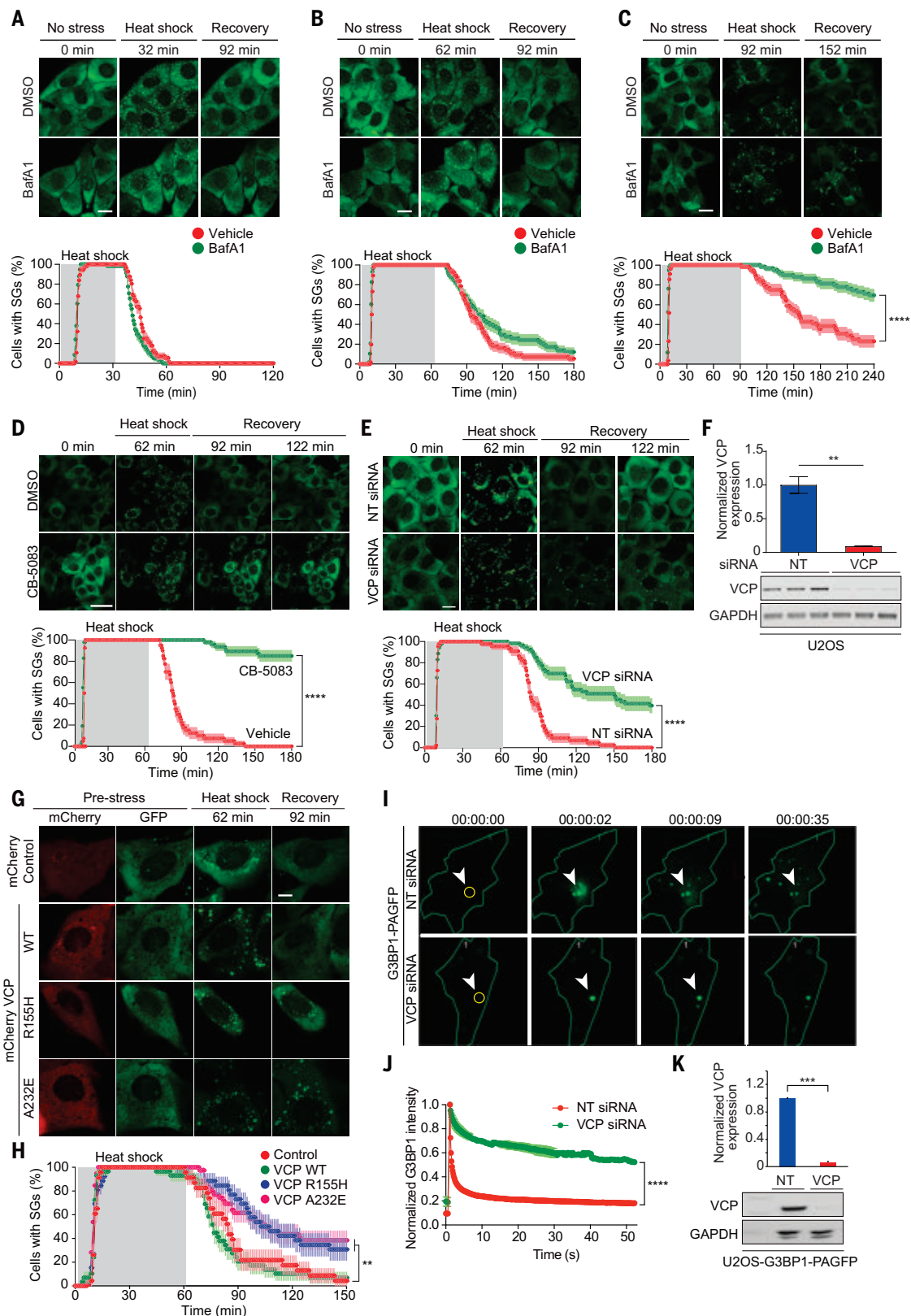
relative reduction in G3BP1 mobility that was specific to the disassembly phase.

Heat shock induces interaction of VCP with ubiquitinated G3BP1

Depletion of cellular adenosine triphosphate (ATP) by the addition of 2-deoxyglucose (2-DG), an inhibitor of the glycolytic pathway, impairs stress granule assembly and subsequent dynamic behavior, implicating energy-dependent processes in assembly dynamics (21). However, the requirement of cellular ATP for the disassembly of stress granules has not been explored. Thus, we examined the contribution of ATP to stress granule disassembly. Under basal conditions, addition of 2-DG led to a ~40% reduction in cellular ATP levels after 30 min, as assessed by a luminescent signal–based ATP sensor (fig. S3A). In contrast, addition of 2-DG after 60 min of heat shock and immediately before recovery led to a ~70% reduction in cellular ATP levels within 10 min (fig. S3A). Concurrently, we observed defects in disassembly of stress granules induced by heat shock or arsenite stresses (fig. S3, B to E, and movie S3), which suggests that ATP is hydrolyzed during the recovery phase and that stress granules are disassembled via energy-dependent processes irrespective of stress type (i.e., heat shock or arsenite).

The stress granule proteome contains a large group of proteins with adenosine triphosphatase (ATPase) activities (21, 22), and the ATPase VCP is recruited to stress granules and contributes to their clearance (11, 16). VCP is a ubiquitin-dependent protein segregase coupled to both proteasome-dependent and autophagy-dependent degradation (11, 12, 16). Because disassembly of heat shock–induced stress granules depends on the ubiquitination of G3BP1, we hypothesized that VCP acts on ubiquitinated G3BP1 to promote disassembly of stress granules. Thus, we first tested whether recruitment of VCP to stress granules is contingent upon ubiquitination of G3BP1. We first confirmed that VCP was recruited to stress granules upon heat shock in wild-type U2OS cells (fig. S4A). Although *G3BP1/2* dKO cells do not form stress granules (3, 8), reintroduction of exogenous GFP-G3BP1 WT or 4KR in these cells restored VCP recruitment to stress granules to levels similar to those observed in wild-type U2OS cells (Fig. 3, A and B, and fig. S4A). However, VCP recruitment to stress granules was significantly reduced in *G3BP1/2* dKO cells stably expressing GFP-G3BP1 6KR or 10KR (Fig. 3, A and B). Thus, ubiquitination of G3BP1 contributes to the recruitment of VCP to stress granules. Although G3BP1 is not the only stress granule protein that is ubiquitinated (14), these data are consistent with the pronounced abundance of G3BP1 among the protein constituents of stress granules (21).

Consistent with these findings, endogenous G3BP1 coimmunoprecipitated with VCP upon heat shock (Fig. 3, C and D). This interaction gradually increased during a 2-hour heat shock period and decreased to baseline levels within 1 hour of recovery (Fig. 3, C to E). TAK-243 treatment abolished the VCP-G3BP1 interaction upon heat shock (Fig. 3F), confirming that the interaction is ubiquitin-dependent. To further assess the role of G3BP1 ubiquitination in this interaction, we expressed WT, 6KR, 4KR, and 10KR mutants in *G3BP1/2* dKO cells and assessed their binding to VCP. 6KR and 10KR mutants showed significantly reduced binding to VCP compared to G3BP1 WT upon heat shock, whereas VCP interaction was not significantly affected in the 4KR mutant (Fig. 3, G and H). This observation is consistent with ubiquitination of the G3BP1 NTF2L domain, but not that of the RRM domain, being primarily responsible for interaction with VCP. Both 6KR and 10KR mutants still showed modest interaction with VCP upon heat shock, although the levels were substantially lower than for G3BP1 WT. Because we did not observe VCP-G3BP1 interaction in cells treated with TAK-243, this finding may be attributable to indirect interaction of VCP with other ubiquitinated proteins bound to G3BP1. Finally, the levels of ubiquitinated G3BP1 were increased by the addition of a VCP inhibitor (CB-5083) and stabilized by the addition of the proteasome inhibitor bortezomib; these findings suggest that ubiquitinated G3BP1 is targeted by VCP for proteasomal degradation during stress granule disassembly (fig. S4, B and C).

The mechanism of stress granule clearance shifts from disassembly to autophagy-dependent degradation during prolonged stress

VCP is essential for autophagy-dependent clearance of persistent stress granules, such as those arising from prolonged stress or disease mutations (11, 23–25). A recent report corroborated the importance of VCP in stress granule clearance but found the process to be autophagy-independent (12). This latter report is consistent with evidence indicating that stress granule constituents are often recycled, with mRNPs reentering the translational pool after removal of stress; however, it was unclear how to reconcile these results with the earlier reports (13). We reasoned that the mechanism of stress granule clearance might be influenced by the chronicity of the initiating stress. To test this hypothesis, we compared the kinetics of stress granule assembly and disassembly in the presence and absence of bafilomycin A1 (BafA1), a vacuolar H$^+$-ATPase inhibitor that blocks autophagy (26). With a short (30-min) exposure to heat shock, stress granules rapidly assembled and then completely disassembled, with all stress granules resolved by 30 min after removal

from heat shock (Fig. 4A and movie S4). With an intermediate (60-min) exposure to heat shock, stress granules also rapidly assembled and completely disassembled, although the disassembly phase was protracted to 60 min in proportion to the longer duration of stress (Fig. 4B and movie S5). BafA1 treatment did not have a significant impact on the kinetics of disassembly after a 30- or 60-min heat shock, indicating that disassembly was taking place independent of any autophagic degradation. With more prolonged heat shock (90 min), the kinetics of stress granule assembly were again unchanged, but disassembly became even more protracted and highly sensitive to BafA1 treatment, indicating a transition to autophagy-dependent clearance of stress granules (Fig. 4C and movie S6). These results are consistent with prior observations of autophagy-dependent clearance of persistent stress granules such as those initiated by disease mutations (11). In such instances, some stress granule disassembly is observed, but autophagy is required for complete removal of stress granules.

VCP is required for disassembly of stress granules

Heat shock–dependent ubiquitination of G3BP1 and subsequent interaction with VCP prompted us to examine whether VCP activity is also required for disassembly of more typical heat shock–induced stress granules. Chemical inhibition of VCP using two different inhibitors (Fig. 4D, fig. S4D, and movie S7) and siRNA knockdown of VCP (Fig. 4, E and F, and movie S8) each led to significant delay in the disassembly of heat shock–induced stress granules. Expression of disease-causing mutant forms of VCP (A232E and R155H) causes the accumulation of poorly dynamic stress granules (11). In this prior study it was unclear whether VCP mutations generally impaired autophagy-dependent clearance, including stress granules, or whether VCP played a more direct role in controlling some aspect of stress granule dynamics that was also impaired by disease mutations. The appreciation that VCP is important for stress granule disassembly afforded the opportunity to determine whether this activity of VCP is also impaired by disease mutations. Thus, we tested the impact of mutant VCP on stress granule dynamics. Whereas exogenous expression of wild-type VCP did not alter rates of stress granule assembly or disassembly, expression of VCP A232E or R155H significantly delayed disassembly of stress granules upon removal of heat stress (Fig. 4, G and H, and movie S9). Thus, mutant VCP proteins have a dominant negative effect on stress granule disassembly similar to that observed in the presence of VCP inhibitors or VCP knockdown.

To further assess the impact of VCP on G3BP1 dynamics in stress granules, we generated a
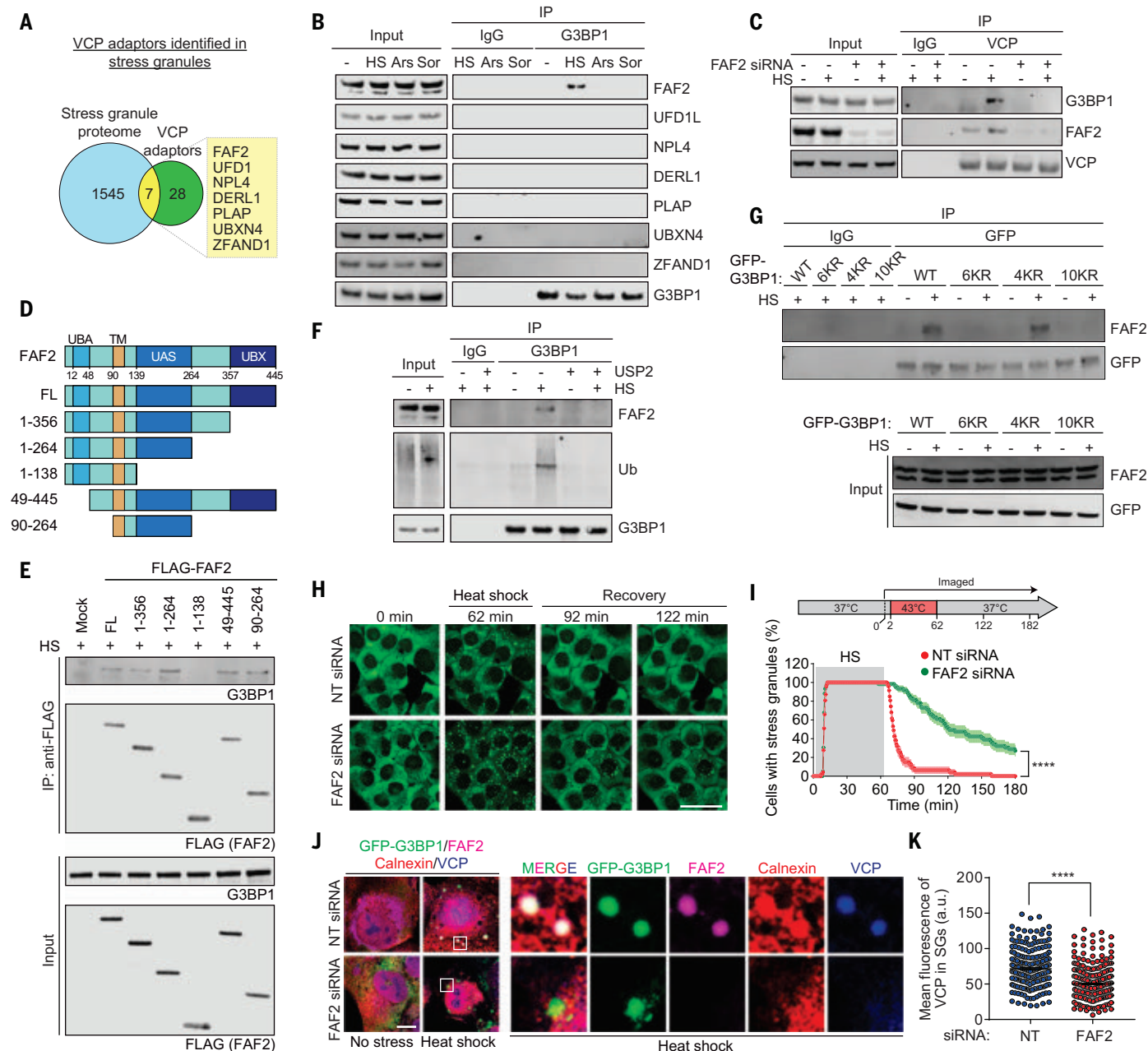
Fig. 5. FAF2 links ubiquitinated G3BP1 to VCP. (**A**) Venn diagram showing overlapping proteins among the stress granule proteome and known VCP adaptors. (**B**) Immunoblot of U2OS cells extracts captured with antibody to GFP exposed to no stress, heat shock (1.5 hours), oxidative stress (sodium arsenite, 1.5 hours), or osmotic stress (sorbitol, 1.5 hours). (**C**) Immunoblot of U2OS cell extracts captured with antibody to VCP after transfection with NT or FAF2 siRNA and exposure to heat shock for 1.5 hours. (**D**) Domain structure of human FAF2 protein and deletion constructs used to investigate the function of individual domains of FAF2. UBA, ubiquitin-associated domain; TM, transmembrane domain; UAS, upstream activation sequence domain; UBX, ubiquitin regulatory X domain. (**E**) Immunoblot of U2OS cells captured with antibody to FLAG exposed to heat shock for 1.5 hours after transfection of FLAG-FAF2 full length (FL) or deletion mutants. (**F**) Immunoblot of U2OS cells exposed to no stress or heat shock for 1.5 hours. Cell extracts were incubated

with or without purified USP2 and captured with magnetic beads conjugated with antibody to G3BP1 for IP, and the resulting beads were analyzed by immunoblot. (**G**) Immunoblots of U2OS *G3BP1/2* dKO cells stably expressing G3BP1 WT, 6KR, 4KR, or 10KR mutants and exposed to heat shock for 2 hours. Cell extracts were captured with magnetic beads conjugated with antibody to GFP for IP. (**H**) U2OS/GFP-G3BP1 cells were transfected with NT or FAF2 siRNA. Scale bar, 50 μm. (**I**) GFP signals were monitored at 30-s intervals at 37°C for 2 min, 43°C for 60 min, and 37°C for 118 min to count cells with two or more stress granules from three technical replicates (NT siRNA n = 47, FAF siRNA n = 59). Error bars indicate SEM. ****P < 0.0001 (Mantel-Cox test). (**J**) Fluorescent imaging of U2OS/GFP-G3BP1 cells exposed to heat shock for 1 hour. Scale bar, 10 μm. (**K**) Fluorescence intensities of VCP in stress granules from three biological replicates are plotted as mean ± SEM (NT siRNA n = 272, FAF2 siRNA n = 349). ****P < 0.0001 (Student's t test).

stable U2OS cell line expressing G3BP1 conjugated to photoactivatable GFP (G3BP1-PAGFP) (27) that allowed us to explore intracellular G3BP1 dynamics by tracking photoactivated molecules. Regions of interest (ROIs) within cells expressing G3BP1-PAGFP were excited briefly with a ~400-nm laser to activate selected pools of G3BP1, followed by time-lapse 488-nm imaging. Stress granule–localized G3BP1 proteins were successfully visualized after PAGFP photoconversion, which enabled monitoring of their mobility (Fig. 4I). In control cells transfected with nontargeting siRNA, the fluorescence intensity of stress granules diminished over time within ROIs and redistributed to nearby stress granules (Fig. 4, I to K). Photoactivated G3BP1-PAGFP in cells depleted of VCP displayed significantly longer residence time in stress granules and limited redistribution of signal to neighboring stress granules; this suggests that VCP is needed to facilitate G3BP1 exit from stress granules (Fig. 4, I to K, and movies S10 and S11).

FAF2, an endoplasmic reticulum–associated VCP adaptor, recruits G3BP1 to promote disassembly of stress granules

VCP associates with a large number of interacting partners and cofactors that regulate its activities in a variety of cellular pathways (28–31). VCP binds to ubiquitinated substrates largely through complexes that include cofactor proteins with ubiquitin-binding domains that function as ubiquitin adaptors for VCP (32). Approximately 35 VCP cofactor proteins have been identified in mammalian cells (16, 31, 33, 34), although in most cases the pathways in which these cofactors function, and the substrates that they recognize, remain poorly defined. Thus, we sought to identify the cofactor(s) that link VCP to stress granules by binding ubiquitinated G3BP1 and thereby influencing the disassembly of heat shock–induced stress granules.

To establish a candidate list of relevant VCP interactors, we integrated a list of 35 VCP cofactors with 1552 proteins previously identified in stress granules initiated by multiple stressors (table S1) (3, 21, 35, 36). The stress granule proteome and VCP cofactors had seven proteins in common (Fig. 5A), including ZFAND1, a protein that promotes the clearance of arsenite-induced stress granules by recruiting VCP and the 26S proteasome (16). We confirmed the interaction of five of these adaptors with endogenous VCP (fig. S5A). Of the seven proteins in common between the stress granule proteome and known VCP cofactors, only the endoplasmic reticulum (ER)–associated protein FAF2 (FAS-associated factor 2; also known as UBXD8) demonstrated interaction with G3BP1 upon heat shock (Fig. 5B). FAF2 interacts with VCP via its C-terminal UBX domain (37, 38). Thus, we hypothesized

that FAF2 binding to VCP might mediate the interaction between VCP and G3BP1. Indeed, siRNA-mediated knockdown of FAF2 eliminated the VCP-G3BP1 interaction (Fig. 5C and fig. S5B). We next performed systematic deletion analysis of FAF2 to identify the domain needed for interaction with G3BP1. FAF2-G3BP1 interaction did not require the UBA domain, but rather the UAS domain (Fig. 5, D and E). The same strategy of domain interactions of FAF2 (i.e., UBX domain–mediated interaction with VCP and UAS domain–mediated interaction with substrate) is also found in an entirely different context, enabling VCP-dependent degradation of adipose triglyceride lipase (ATGL) in lipid metabolism (37). Nonetheless, we found that formation of a complex between FAF2 and G3BP1 was ubiquitin-dependent, because it was prevented by treatment of G3BP1 with the deubiquitinase USP2 (Fig. 5F) and ubiquitination-deficient G3BP1 6KR and 10KR mutants failed to interact with FAF2 (Fig. 5G). Because FAF2-G3BP1 interaction is ubiquitin-dependent but does not require the UBA domain of FAF2, these results may suggest the existence of an additional ubiquitin-binding entity within the G3BP1-FAF2-VCP complex.

Disassembly of heat shock–induced stress granules is FAF2-dependent

We next examined the importance of FAF2 in stress granule disassembly. Depletion of FAF2 by siRNA delayed the disassembly of heat shock–induced stress granules upon removal of heat stress, phenocopying our earlier observations using chemical inhibitors of VCP (Fig. 5, H and I, and movie S12). In contrast, depletion of six other adaptors, including ZFAND1, did not significantly alter the disassembly of heat shock–induced stress granules upon removal of heat stress (fig. S5C), which suggests that VCP specifically engages FAF2 to regulate the dynamics of stress granules induced by heat shock. Depletion of ZFAND1 delayed disassembly of arsenite-induced stress granules upon removal of arsenite as expected (16), whereas depletion of FAF2 failed to delay disassembly of arsenite-induced stress granules, demonstrating a context-dependent control of stress granule dynamics by ZFAND1 and FAF2 (fig. S5, D to G).

The FAF2-G3BP1-VCP pathway of stress granule disassembly occurs at the ER membrane

FAF2 is an ER membrane–associated protein that contains a stretch of 28 hydrophobic amino acid residues that are inserted into the cytosolic surface of the ER membrane (39). FAF2 is perhaps best known for its role in ubiquitin-dependent degradation of misfolded ER proteins in the context of ER-associated protein degradation (ERAD) (40, 41). FAF2 is

also essential for VCP-dependent degradation of ATGL in ER-associated lipid droplet metabolism (37). The VCP-FAF2 complex disassembles mRNPs by promoting the extraction and degradation of ubiquitinated HuR, an RNA-binding protein, from mRNA (42). Thus, there is precedent for FAF2 involvement in ubiquitin-dependent degradation of diverse substrates, including RNP granule constituents, at the ER membrane.

To assess whether the FAF2-G3BP1-VCP pathway of stress granule disassembly occurs at the ER membrane, we began by examining the spatiotemporal relationship among the ER, FAF2, G3BP1, and VCP upon heat shock. At baseline, FAF2 was diffusely distributed with the ER, as indicated by colocalization with the ER marker calnexin (fig. S6A). Upon heat shock, calnexin and FAF2 signals appeared as bulges within ER tubules that colocalized with stress granules marked by GFP-G3BP1 (Fig. 5J). VCP showed a similar colocalization with G3BP1, FAF2, and calnexin upon heat shock (Fig. 5J). In contrast, arsenite stress did not induce appreciable FAF2 localization with stress granules (fig. S6B). VCP accumulation in heat stress–induced stress granules was significantly decreased in cells depleted of FAF2, whereas the more modest VCP accumulation in arsenite-induced stress granules was not influenced by depletion of FAF2 (Fig. 5K and fig. S6, C and D). These findings support our hypothesis that FAF2 functions as a VCP adaptor that connects the ER to heat shock–induced stress granules by interacting with ubiquitinated G3BP1.

Discussion

Whereas stress granule assembly is evidently a universal process across cell types and in response to a wide variety of stresses, it has recently become apparent that the composition of stress granules is context-dependent with respect to cell type and the initiating stress (35). Our study builds on this perspective by demonstrating context-dependent mechanisms of stress granule elimination that depend on both the type and duration of stress. This context is particularly important when considering how we investigate the relationship of stress granules to disease: We must be attentive to disease-relevant contexts, including the use of specific cell types or the administration of different types of stress.

Indeed, deconvolving the complexity of the existing literature on mechanisms of stress granule elimination requires careful consideration of the type and duration of the initiating stress. As demonstrated here, persistent stress granules, such as those that arise through chronic stress or disease mutations, are eliminated by autophagy-dependent degradation. In contrast, short-lived granules are rapidly disassembled, which permits recycling

of constituents. Thus, we propose that prolonged stress or disease mutations result in persistence of stress granules that become gradually less dynamic and are eventually cleared through an autophagy-dependent degradative process, consistent with previous reports that clearance of such granules requires VCP and autophagy machinery such as ATG7 (6, 11, 43).

Autophagy-dependent clearance of stress granules results in degradation and permanent loss of constituent proteins and RNAs. In contrast, if cellular stress is removed relatively quickly while stress granules remain dynamic, they are disassembled by decondensation—a reversal of LLPS—that occurs when sum of protein-protein, protein-RNA, and RNA-RNA interactions falls below the percolation threshold. This decrease in the sum total of interactions within the stress granule network could occur through several mechanisms, including a decrease in the concentration of individual constituents or altered posttranslational modifications of constituent proteins that weaken the interaction network. In contrast to stress granule clearance, disassembly is a nondegradative process wherein the individual constituents disassociate from one another and are recycled, including liberation of the mRNAs to rejoin the translating pool (13).

The importance of context also helps to resolve earlier apparent inconsistencies regarding the role of ubiquitination in stress granule dynamics, because ubiquitination had previously been reported to be dispensable for the formation and disassembly of arsenite-induced stress granules (15). A striking example of the importance of context can also be seen in the use of VCP adaptors in the process of stress granule clearance. Elimination of arsenite-induced stress granules requires the adaptor ZFAND1 but not FAF2, whereas elimination of heat shock–induced stress granules requires FAF2 but not ZFAND1.

The cytosolic surface of the ER is a nexus for coordinating responses to heat stress, including the integrated stress response (44), the UPR (45), ERAD (46) and, now, disassembly of stress granules. The basis for disassembly of heat shock–induced stress granules at the ER surface appears to involve the ER-associated protein FAF2, which is essential for VCP to engage ubiquitinated G3BP1. Interestingly, FAF2 is not only an important factor in stress granule disassembly, but also in ERAD (33) and ER-associated metabolism of lipid droplets (37, 47). The role of ER localization in dismantling stress granules in response to other stresses remains unclear. A recent report describes "fission" of arsenite-induced stress granules occurring at contact sites with ER, interpreted as ER membrane–mediated mechanical cleavage of stress granules, which may be somehow related to disassembly (48).

Much remains to be learned about the distinct mechanisms of stress granule disassembly, and in particular how these mechanisms may be affected by disease-causing mutations. Indeed, although disease mutations in VCP impair autophagy-dependent stress granule clearance (11), here we found that these mutations also impair autophagy-independent disassembly. Similar TAR DNA-binding protein 43 (TDP-43) laden, spontaneously arising, poorly dynamic stress granules are found in cells expressing disease mutant forms of RNA-binding proteins, such as FUS, hnRNPA1, hnRNPA2, and TIA-1 (43). Thus, mutations in VCP may impair its ability to disassemble stress granules, representing a mechanistic intersection for multiple, distinct disease-causing mutations.

Stress granule disassembly is mediated by extraction of ubiquitinated G3BP, a finding that underscores the central role for G3BP in maintaining the stress granule interaction network. Moreover, because cells produce distinct patterns of ubiquitination in response to different stressors (14), our study illustrates how even a single protein within a stress-specific ubiquitinome can be pursued experimentally to demonstrate meaningful consequences for cellular function.

Materials and methods
Plasmid, siRNA, and transfection

Synthetic G3BP1 fragments with K36/50/59/64/76/123R (NTF2L 6KR), K353/357/376/393R (RRM 4KR), and K36/50/59/64/76/123/353/357/376/393R (NTF2L/RRM 10KR) mutations were inserted into HindIII and BamHI sites of pEGFP-C3 (Clontech) by Bio Basic Inc. pEGFP-C3 G3BP1 K50R, K36/50R, K50/59R, K36/50/59R, K50/59/64R, K36/50/59/64R, K50/76R, K50/123R, and K50/76/123R mutants were generated by site-directed mutagenesis using a Q5 Site-Directed Mutagenesis kit (New England Biolabs E0054S). pRK5-HA-ubiquitin-WT, pRK5-HA-ubiquitin-K48R, pRK5-HA-ubiquitin-K63R, pRK5-HA-ubiquitin-K48, and pRK5-HA-ubiquitin-K63 (Addgene 17608, 17604, 17606, 17605, and 17606, respectively) were kindly provided by T. Dawson. VCP WT, VCP R155H, and VCP A232E were inserted into BamHI and HindIII sites of pmCherry-N1 (Clontech) (49). pCLi40w-MND-G3BP1-GFP and pCLiw40-MND-G3BP1-PAGFP plasmids, used to generate G3BP-GFP stable U2Os cells via lentiviral transduction, were constructed by releasing dsRedEX2-EIF1α-GFP from pCLEG-MND-dsRedEX2 (Vector Development and Production, SJCRH) and inserting G3BP1-GFP or G3BP1-PAGFP into the EcoRI and BsrGI positions. G3BP1-GFP and G3BP1-PAGFP inserts were released from peGFP-N1-G3BP and pePAGFP-N1-G3BP by EcoRI and dBsrGI digestion. G3BP1 was inserted in peGFP-N1 or pePAGFP-N1 (Clontech) with primers containing EcoRI and BamHI sites in the 5′ and

3′ sites of G3BP. Similarly, the pCLi40w-MND-TIAL1-PAGFP construct was generated using EcoRI and KpnI. pRK5-FLAG-FAF2 (Addgene 53777) was kindly provided by Y. Ye. pRK5-FLAG-FAF2 1-356, pRK5-FLAG-FAF2 1-264, pRK5-FLAG-FAF2 1-138, pRK5-FLAG-FAF2 49-445, and pRK5-FLAG-FAF2 90-264 were inserted into SalI and NotI sites of pRK5-FLAG-FAF2. For knockdown experiments, the following siRNA constructs were purchased from Horizon Discovery: pooled nontargeting siRNA (D-001810-10), VCP (L-008728-00), FAF2 (L-010649-02), UFD1L (L-017918-00), NPL4 (L-020796-01), DERL1 (L-010733-02), PLAP (L-016215-00), UBXN4 (L-014184-01), and ZFAND1 (L-009638-02). Transporter 5 (Polysciences 26008) and FuGENE 6 (Promega E2691) were used for transient transfections of cDNA into HEK293T and U2OS cells, respectively, according to the manufacturer's instructions. Lipofectamine RNAiMax (Thermo Fisher Scientific 13778150) was used for transfection of siRNA according to the manufacturer's instructions.

Cell culture

HEK293T (CRL-3216) and U2OS (HTB-96) cells were purchased from ATCC, cultured in Dulbecco's modified Eagle's medium (HyClone) supplemented with 10% fetal bovine serum (HyClone SH30396.03), 1× GlutaMAX (Thermo Fisher Scientific 35050061), and penicillin (50 U/ml)–streptomycin (50 μg/ml) (Gibco 15140–122), and maintained at 37°C in a humidified incubator with 5% CO_2. U2OS *G3BP1/2* KO cells were previously described in (50). U2OS cells stably expressing GFP-G3BP1 were previously described in (51).

Generation of G3BP1 add-back cell lines

U2OS *G3BP1/2* dKO cells were transfected with pEGFP-C3 G3BP1 WT, NTF2L 6KR, RRM 4KR, or NTF2L/RRM 10KR constructs using FUGENE 6 (Promega); 48 hours after transfection, G418 sulfate (500 μg/ml; Thermo Fisher Scientific 10131035) was added to culture media without penicillin and streptomycin for selection. After pharmacological selection, GFP-positive cells were purified using cell sorting to produce stable cell lines.

Heat shock and drug treatments

For heat shock, cells were transferred to a 43°C humidified incubator with 5% CO_2. Chemicals dissolved in DMSO were prepared and added to cells as follows: sodium arsenite (0.5 mM, Millipore Sigma 1062771000), TAK-243 (1 μM, ChemieTek CT-M7243), CB-5083 (1 μM, Cayman Chemical 19311), bafilomycin A1 (10 nM, Cayman Chemical 11038), cycloheximide (100 μg/ml, Sigma-Aldrich C7698), and bortezomib (1 μM, Millipore Sigma 5043140001). D-Sorbitol (0.4 M, Millipore Sigma S1876) was dissolved in culture media and prewarmed before treatment.

Eeyarestatin I was used at a final concentration of 56 μM (30 min pretreatment).

Immunoblotting

Cells were washed twice with PBS and lysed with RIPA buffer (25 mM Tris-HCl, pH 7.6, 150 mM NaCl, 1% NP-40, 1% sodium deoxycholate, 0.1% SDS; Thermo Scientific 89901) supplemented with 1 mM EDTA (Invitrogen 15575020) and proteinase inhibitor cocktail (Roche 1183617001). Lysates were centrifuged for 15 min at 4°C at 20,000g. 4× NuPAGE LDS sample buffer (Thermo Fisher Scientific NP0007) was added to the supernatant and samples were boiled at 90°C for 5 min. Samples were separated in 4 to 12% NuPAGE Bis-Tris gels (Thermo Fisher Scientific NP0336BOX or NP0321BOX) and transferred to PVDF membranes (Thermo Fisher Scientific IB24001) using an iBlot 2 transfer device (Thermo Fisher Scientific). Membranes were blocked with Odyssey blocking buffer (LI-COR Biosciences 927-50000) and then incubated with primary antibodies at 4°C overnight: G3BP1 (Proteintech 13057-2-AP), ubiquitin (Santa Cruz Biotechnology sc-8017), HA tag (Invitrogen 715500), GFP (Invitrogen A11122), FLAG (Proteintech 20543-1-AP), VCP (Santa Cruz Biotechnology sc-20799 or Thermo Fisher Scientific MA3-004), FAF2 (Proteintech 16251-1-AP), UFD1L (Proteintech 10615-1-AP), NPL4 (Proteintech 11638-1-AP), DERL1 (Thermo Scientific PA553444), PLAP (Santa Cruz Biotechnology sc-390454), UBXN4 (Thermo Fisher Scientific PA5577611), ZFAND1 (Sigma-Aldrich HPA023383), β-actin (Santa Cruz Biotechnology sc-47778 or Sigma-Aldrich A5316), and GAPDH (Santa Cruz Biotechnology sc-32233 or Sigma-Aldrich G9545). Membranes were washed three times with TBS-T (0.05% Tween) and further incubated with IRDye 680RD/800CW-labeled secondary antibodies (LI-COR Biosciences 926-68073 or 926-32212) at a dilution of 1:10,000. Membranes were visualized with an Odyssey Fc imaging system (LI-COR Biosciences) and quantified using ImageJ software (NIH).

Immunoprecipitation

Cells were washed twice with PBS and lysed with IP lysis buffer (25 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1% NP-40, 1 mM EDTA, 5% glycerol; Thermo Scientific 87787) supplemented with 20 mM N-ethylmaleimide (NEM) (Sigma-Aldrich E3876), 50 μM PR-619 (Sigma-Aldrich, 662141) and proteinase inhibitor cocktail (Roche 1183617001). Lysates were centrifuged at 4°C for 15 min at 20,000g. The supernatants were incubated with normal mouse IgG (Santa Cruz Biotechnology sc-2025), GFP antibody (Santa Cruz Biotechnology sc-9996), VCP antibody (Santa Cruz Biotechnology sc-57492), FLAG antibody (Santa Cruz Biotechnology sc-166355), or G3BP1 antibody (BD 611127) conjugated with protein A/G magnetic beads (Thermo Fisher

Scientific 88803) at 4°C overnight. Beads were washed three times with wash buffer (50 mM Tris-HCl, pH 7.5, 500 mM NaCl, 0.05% Tween-20, 20 mM NEM and 50 μM PR-619) and treated with 0.1 M glycine pH 2.7 (Teknova G4527) at room temperature for 10 min to elute proteins. The resulting samples were analyzed by immunoblotting.

TUBE pulldown

Lysates were prepared using the same procedures as when performing immunoprecipitation. Lysates were incubated with Halo-beads or Halo-4xUBA[UBQLN1]-beads at 4°C overnight. Beads were washed three times with wash buffer, mixed with 4× NuPAGE LDS sample buffer, and boiled at 90°C for 5 min. The resulting samples were analyzed by immunoblotting.

Deubiquitination assay

Lysates were prepared using the same procedures as when performing immunoprecipitation, except that NEM and PR-619 were excluded from the tubes where the deubiquitination reaction was performed. Lysates were incubated with or without 2.72 μg of USP2 (LifeSensors DB501) overnight and further incubated with normal mouse IgG or G3BP1 antibody conjugated with protein A/G magnetic beads for 2 hours at 4°C. Beads were washed three times with wash buffer and treated with 0.1 M glycine pH 2.7 at room temperature for 10 min to elute proteins. The resulting samples were analyzed by immunoblotting.

Di-GLY TMT ubiquitinome analysis

Proteomic profiling of the ubiquitinome was performed with a previously reported protocol (52) with modification. Briefly, for each sample, HEK293T cells from five 15-cm dishes were harvested, washed with ice-cold PBS, and collected by centrifugation. Cells were lysed in a buffer (50 mM HEPES, pH 8.5, 8 M urea, and 0.5% sodium deoxycholate) with deubiquitinating enzyme (DUB) inhibitors (50 μM PR-619 and 10 mM iodoacetamide) using probe sonication (15 s × 3 cycles). Iodoacetamide-induced pseudo–di-GLY peptides were not detected at room temperature (53) and were not recognized by di-GLY antibodies during the enrichment (54). To further minimize the impact of DUB activities, the cell lysis protocol was modified by immediately digesting the lysates with LysC (Wako) at an enzyme/substrate ratio of 1:100 (w/w) for 2 hours at room temperature, to achieve rapid DUB deactivation through proteolysis. Following this, the samples were quenched with DTT (30 mM for 30 min) and subjected to trypsin digestion and desalting. The desalted peptides (~4 mg) were resuspended in 400 μl of ice-cold IAP buffer (50 mM MOPS, pH 7.2, 10 mM sodium phosphate, and 50 mM NaCl)

and centrifuged at 21,000g for 10 min at 4°C to remove any insoluble material. The peptides were then incubated with di-GLY antibody beads (Cell Signaling Technology) at an antibody-to-peptide ratio of 1:20 (w/w, optimized through a pilot experiment) for 2 hours at 4°C with gentle end-over-end rotation. The antibody beads were then collected by a brief centrifugation and washed three times with 1 ml of ice-cold IAP buffer and twice with 1 ml of ice-cold PBS, while the supernatants were carefully removed and saved. The di-GLY peptides were eluted at room temperature twice with 50 μl of 0.15% TFA, dried, and resuspended in 50 mM HEPES (pH 8.5) for 11-plex TMT labeling. The TMT-labeled peptides were combined equally among replicates, desalted, and analyzed by LC/LC-MS/MS using a similar protocol as described above except with the following differences: (i) a 60-min gradient used in the acidic pH LC-MS/MS with Orbitrap QE HF mass spectrometer (Thermo Fisher Scientific), (ii) database search using the COMET algorithm [v2018.013 (55)] with 0.02-Da mass tolerance for MS/MS ions, full trypticity with maximal five missed cleavages, maximal five dynamic modifications per peptide, and di-GLY modification on Lys (+114.04293 Da) as a dynamic modification.

Immunofluorescence

Cells were grown in 8-well chamber slides (Millipore PEZGS0816). Cells were fixed with 4% paraformaldehyde (Alfa Aesar J61899) in PBS for 10 min, permeabilized with 0.2% Triton X-100 in PBS for 5 min, and blocked with 3% BSA for 1 hour. Samples were further incubated with primary antibodies as the following targets in blocking buffer at 4°C overnight: Lys[48] linkage–specific ubiquitin (Millipore Sigma 05-1307), Lys[63] linkage–specific ubiquitin (Millipore Sigma 05-1308), G3BP1 (BD 611127), G3BP1 (Proteintech 13057-2-AP), eIF3η (Santa Cruz Biotechnology sc-16377), VCP (BD Biosciences 612183), FAF2 (Proteintech 16251-1-AP), and calnexin (Thermo Fisher Scientific PA5-19169). Samples were washed three times with PBS and incubated with host-specific Alexa Fluor 488/555/647 secondary antibodies (Thermo Fisher Scientific) for 1 hour at room temperature. For microscopic imaging, slides were mounted with ProLong Gold Antifade reagent with DAPI (Thermo Fisher Scientific P36931). Images were captured using a Leica TCS SP8 STED 3× confocal microscope with a 63× oil objective. To stain ER-resident calnexin, cells were grown in fibronectin-coated coverslips (neuVitro GG18FIBRONECTIN) and permeabilized with 0.2% Triton X-100 in PBS for 1 min. Slides were mounted with ProLong Glass Antifade reagent (Thermo Fisher Scientific P36984).

Disassembly of stress granules, intracellular phase diagram, and FRAP using time-lapse live-cell microscopy

A Yokogawa CSU W1 spinning disk attached to a Nikon Ti2 eclipse with a Photometrics Prime 95B camera using Nikon Elements software was used in time-lapse live-cell imaging and FRAP. The light path was split between the port for the spinning disk/ acquisition laser and the FRAP lasers, enabling FRAP to occur simultaneously while imaging. Imaging was taken using a 60× Plan Apo 1.4NA oil objective and Perfect Focus 2.0 (Nikon) engaged for the duration of the capture. During imaging, cells were maintained at 37°C and supplied with 5% CO_2 using a Bold Line Cage Incubator (Okolabs) and an objective heater (Bioptechs).

To monitor the disassembly of heat shock–induced stress granules, multipoint images over 5 xy fields for each condition per one replicate were taken with the 555-nm laser. At 2 min into imaging, the objective temperatures were raised to 43°C for 60 min. After heat shock, the temperature was lowered back to 37°C to alleviate the stress, and cells were imaged after 2 hours had passed. Images were taken at each xy position every 30 s. Cells containing two or more stress granules larger than 0.54 μm were counted as stress granule–positive cells.

For intracellular phase diagrams, multipoint images over 25 xy fields for each condition per one replicate were taken with the 555-nm laser. At 2 min into imaging, the objective temperatures were raised to 43°C for 60 min. Images were taken at each xy position every 1 min. Phase diagrams were constructed by measuring GFP fluorescence intensity in each cell and assessing the presence of stress granules using Fiji software.

For fluorescence recovery after photobleaching, time lapses were acquired every 100 ms over the course of 45 s for stress granules with photobleaching with the 488-nm FRAP laser occurring 2 s into capture. Data were taken from at least n =10 different cells or lysate granules for each condition. In Nikon Elements, ROIs were generated in the photobleached region, a nonphotobleached cell, and the background for each time lapse, and the mean intensity of each was extracted. For photobleached regions, a 2.5-μm-diameter circle was used. Data were repeated in triplicate for each condition, with each replicate having at least n = 10 cells. The values retrieved from the ROIs were exported into Igor Pro 7.0 (WaveMetrics) and fit curves were generated after photobleach and background values were corrected.

Photoactivation of G3BP1

For live imaging of photoactivatable GFP, U2OS cells expressing G3BP-GFP or G3BP1-

PAGFP were plated on 40-mm #1.5 thick coverslips (Bioptechs) or chambered coverglass (Millipore). Cells were observed with a Marianas confocal microscope (Leica) with a 63× objective. For heat shock live imaging experiments 48 hours after transfection, the coverslip was transferred to a FCS2 chamber assembled according to the manufacturer's instructions (Bioptechs). Media was perfused through the chamber, and then the chamber was placed into a Marianas spinning disk confocal system with a stage-top incubator and 63× objective with an objective heater (Bioptechs), both preheated to 37°C. The Microaqueduct Slide heater (FCS2 system) and the heated objective with 37°C immersion oil (Zeiss) were used to control the temperature. Movies were collected on a Marianas confocal microscope with 63× objective, with 40-s intervals for the live imaging experiments, 200-ms intervals for the FLAP experiments of the time period. Average and standard errors were calculated from three independent experiments measuring at least 35 cells.

ATP measurement

Cellular contents of ATP were measured using CellTiter-Glo 2.0 assay kit (Promega G9242) according to the manufacturer's instructions; 200 mM 2-deoxy-D-gluose (Millipore Sigma D6134) was used to inhibit the glycolysis pathway.

Statistical analysis

Statistical analysis was performed in Graph-Pad Prism. Comparisons between two means were performed by two-tailed t test. Comparisons among multiple means over 2 were performed by one-way analysis of variance (ANOVA) with Tukey's test. Mantel-Cox tests were used to compare the dissociation curves showing cells with stress granules in live-cell imaging.

REFERENCES AND NOTES

1. J. R. Buchan, mRNP granules: Assembly, function, and connections with disease. *RNA Biol.* **11**, 1019–1030 (2014). doi: 10.4161/15476286.2014.972208; pmid: 25531407

2. E. Gomes, J. Shorter, The molecular language of membraneless organelles. *J. Biol. Chem.* **294**, 7115–7127 (2019). doi: 10.1074/jbc.TM118.001192; pmid: 30045872

3. P. Yang et al., G3BP1 Is a Tunable Switch that Triggers Phase Separation to Assemble Stress Granules. *Cell* **181**, 325–345. e28 (2020). doi: 10.1016/j.cell.2020.03.046; pmid: 32302571

4. D. W. Sanders et al., Competing Protein-RNA Interaction Networks Control Multiphase Intracellular Organization. *Cell* **181**, 306–324.e28 (2020). doi: 10.1016/j.cell.2020.03.050; pmid: 32302570

5. J. Guillén-Boixet et al., RNA-Induced Conformational Switching and Clustering of G3BP Drive Stress Granule Assembly by Condensation. *Cell* **181**, 346–361.e17 (2020). doi: 10.1016/j.cell.2020.03.049; pmid: 32302572

6. C. Mathieu, R. V. Pappu, J. P. Taylor, Beyond aggregation: Pathological phase transitions in neurodegenerative disease. *Science* **370**, 56–60 (2020). doi: 10.1126/science.abb8032; pmid: 33004511

7. S. F. Banani, H. O. Lee, A. A. Hyman, M. K. Rosen, Biomolecular condensates: Organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285–298 (2017). doi: 10.1038/nrm.2017.7; pmid: 28225081

8. N. Kedersha et al., G3BP-Caprin1-USP10 complexes mediate stress granule condensation and associate with 40S subunits. *J. Cell Biol.* **212**, 845–860 (2016). doi: 10.1083/jcb.201508028; pmid: 27022092

9. P. Zhang et al., Chronic optogenetic induction of stress granules is cytotoxic and reveals the evolution of ALS-FTD pathology. *eLife* **8**, e39578 (2019). doi: 10.7554/eLife.39578; pmid: 30893049

10. D. Tauber, G. Tauber, R. Parker, Mechanisms and Regulation of RNA Condensation in RNP Granule Formation. *Trends Biochem. Sci.* **45**, 764–778 (2020). doi: 10.1016/j.tibs.2020.05.002; pmid: 32475683

11. J. R. Buchan, R. M. Kolaitis, J. P. Taylor, R. Parker, Eukaryotic stress granules are cleared by autophagy and Cdc48/VCP function. *Cell* **153**, 1461–1474 (2013). doi: 10.1016/j.cell.2013.05.037; pmid: 23791177

12. B. Wang et al., ULK1 and ULK2 Regulate Stress Granule Disassembly Through Phosphorylation and Activation of VCP/p97. *Mol. Cell* **74**, 742–757.e8 (2019). doi: 10.1016/j.molcel.2019.03.027; pmid: 30979586

13. P. Anderson, N. Kedersha, RNA granules: Post-transcriptional and epigenetic modulators of gene expression. *Nat. Rev. Mol. Cell Biol.* **10**, 430–436 (2009). doi: 10.1038/nrm2694; pmid: 19461665

14. B. A. Maxwell et al., Ubiquitination is essential for recovery of cellular activities following heat shock. *Science* **372**, eabc3593 (2021).

15. S. Markmiller et al., Active Protein Neddylation or Ubiquitylation Is Dispensable for Stress Granule Dynamics. *Cell Rep.* **27**, 1356–1363.e3 (2019). doi: 10.1016/j.celrep.2019.04.015; pmid: 31042464

16. A. Turakhiya et al., ZFAND1 Recruits p97 and the 26S Proteasome to Promote the Clearance of Arsenite-Induced Stress Granules. *Mol. Cell* **70**, 906–919.e7 (2018). doi: 10.1016/j.molcel.2018.04.021; pmid: 29804830

17. K. N. Swatek, D. Komander, Ubiquitin modifications. *Cell Res.* **26**, 399–422 (2016). doi: 10.1038/cr.2016.39; pmid: 27012465

18. F. Ohtake, Y. Saeki, S. Ishido, J. Kanno, K. Tanaka, The K48-K63 Branched Ubiquitin Chain Regulates NF-κB Signaling. *Mol. Cell* **64**, 251–266 (2016). doi: 10.1016/j.molcel.2016.09.014; pmid: 27746020

19. P. V. Hornbeck et al., PhosphoSitePlus, 2014: Mutations, PTMs and recalibrations. *Nucleic Acids Res.* **43**, D512–D520 (2015). doi: 10.1093/nar/gku1267; pmid: 25514926

20. Amino acid abbreviations: A, Ala; D, Asp; E, Glu; G, Gly; H, His; K, Lys; L, Leu; N, Asn; P, Pro; Q, Gln; S, Ser; V, Val; Y, Tyr.

21. S. Jain et al., ATPase-Modulated Stress Granules Contain a Diverse Proteome and Substructure. *Cell* **220**, 487–498 (2016). doi: 10.1016/j.cell.2015.12.038; pmid: 26777405

22. B. D. Freibaum, J. Messing, P. Yang, H. J. Kim, J. P. Taylor, High-fidelity reconstitution of stress granules and nucleoli in mammalian cellular lysate. *J. Cell Biol.* **290**, e202009079 (2021). doi: 10.1016/j.neurobiolaging.2014.08.019; pmid: 25444610

23. S. Matus, D. A. Bosco, C. Hetz, Autophagy meets fused in sarcoma-positive stress granules. *Neurobiol. Aging* **35**, 2832–2835 (2014). doi: 10.1016/j.neurobiolaging.2014.08.019; pmid: 25444610

24. M. O. Krisenko et al., Syk Is Recruited to Stress Granules and Promotes Their Clearance through Autophagy. *J. Biol. Chem.* **290**, 27803–27815 (2015). doi: 10.1074/jbc.M115.642900; pmid: 26429917

25. L. Marrone et al., Isogenic FUS-eGFP iPSC Reporter Lines Enable Quantification of FUS Stress Granule Pathology that Is Rescued by Drugs Inducing Autophagy. *Stem Cell Rep.* **10**, 375–389 (2018). doi: 10.1016/j.stemcr.2017.12.018; pmid: 29358088

26. A. Yamamoto et al., Bafilomycin A1 prevents maturation of autophagic vacuoles by inhibiting fusion between autophagosomes and lysosomes in rat hepatoma cell line, H-4-II-E cells. *Cell Struct. Funct.* **23**, 33–42 (1998). doi: 10.1247/csf.23.33; pmid: 9639028

27. G. H. Patterson, J. Lippincott-Schwartz, A photoactivatable GFP for selective photolabeling of proteins and cells. *Science* **297**, 1873–1877 (2002). doi: 10.1126/science.1074952; pmid: 12228718

28. H. Meyer, M. Bug, S. Bremer, Emerging functions of the VCP/p97 AAA-ATPase in the ubiquitin system. *Nat. Cell Biol.* **14**, 117–123 (2012). doi: 10.1038/ncb2407; pmid: 22298039

29. A. Stolz, W. Hilt, A. Buchberger, D. H. Wolf, Cdc48: A power machine in protein degradation. *Trends Biochem. Sci.* **36**, 515–523 (2011). doi: 10.1016/j.tibs.2011.06.001; pmid: 21741246

30. H. O. Yeung et al., Insights into adaptor binding to the AAA protein p97. Biochem. Soc. Trans. **36**, 62–67 (2008). doi: 10.1042/BST0360062; pmid: 18208387

31. Y. Ye, W. K. Tang, T. Zhang, D. Xia, A Mighty "Protein Extractor" of the Cell: Structure and Function of the p97/CDC48 ATPase. Front. Mol. Biosci. **4**, 39 (2017). doi: 10.3389/fmolb.2017.00039; pmid: 28660197

32. Y. Ye, Diverse functions with a common regulator: Ubiquitin takes command of an AAA ATPase. J. Struct. Biol. **156**, 29–40 (2006). doi: 10.1016/j.jsb.2006.01.005; pmid: 16529947

33. H. Meyer, C. C. Weihl, The VCP/p97 system at a glance: Connecting cellular function to disease pathogenesis. J. Cell Sci. **127**, 3877–3883 (2014). doi: 10.1242/jcs.093831; pmid: 25146396

34. D. Burana et al., The Ankrd13 Family of Ubiquitin-interacting Motif-bearing Proteins Regulates Valosin-containing Protein/p97 Protein-mediated Lysosomal Trafficking of Caveolin 1. J. Biol. Chem. **291**, 6218–6231 (2016). doi: 10.1074/jbc.M115.710707; pmid: 26797118

35. S. Markmiller et al., Context-Dependent and Disease-Specific Diversity in Protein Interactions within Stress Granules. Cell **172**, 590–604.e13 (2018). doi: 10.1016/j.cell.2017.12.032; pmid: 29373831

36. J. Y. Youn et al., High-Density Proximity Mapping Reveals the Subcellular Organization of mRNA-Associated Granules and Bodies. Mol. Cell **69**, 517–532.e11 (2018). doi: 10.1016/j.molcel.2017.12.020; pmid: 29395067

37. J. A. Olzmann, C. M. Richter, R. R. Kopito, Spatial regulation of UBXD8 and p97/VCP controls ATGL-mediated lipid droplet turnover. Proc. Natl. Acad. Sci. U.S.A. **110**, 1345–1350 (2013). doi: 10.1073/pnas.1213738110; pmid: 23297223

38. P. Kloppsteck, C. A. Ewens, A. Förster, X. Zhang, P. S. Freemont, Regulation of p97 in the ubiquitin-proteasome system by the UBX protein-family. Biochim. Biophys. Acta **1823**, 125–129 (2012). doi: 10.1016/j.bbamcr.2011.09.006; pmid: 21963883

39. J. N. Lee et al., Identification of Ubxd8 protein as a sensor for unsaturated fatty acids and regulator of triglyceride synthesis. Proc. Natl. Acad. Sci. U.S.A. **107**, 21424–21429 (2010). doi: 10.1073/pnas.1011859107; pmid: 21115839

40. B. Mueller, E. J. Klemm, E. Spooner, J. H. Claessen, H. L. Ploegh, SEL1L nucleates a protein complex required for dislocation of misfolded glycoproteins. Proc. Natl. Acad. Sci. U.S.A. **105**, 12325–12330 (2008). doi: 10.1073/pnas.0805371105; pmid: 18711132

41. Y. Xia et al., Pathogenic mutation of UBQLN2 impairs its interaction with UBXD8 and disrupts endoplasmic reticulum-associated protein degradation. J. Neurochem. **129**, 99–106 (2014). doi: 10.1111/jnc.12606; pmid: 24215460

42. H. L. Zhou, C. Geng, G. Luo, H. Lou, The p97-UBXD8 complex destabilizes mRNA by promoting release of ubiquitinated HuR from mRNP. Genes Dev. **27**, 1046–1058 (2013). doi: 10.1101/gad.215681.113; pmid: 23618873

43. N. B. Nedelsky, J. P. Taylor, Bridging biophysics and neurology: Aberrant phase transitions in neurodegenerative disease. Nat. Rev. Neurol. **15**, 272–286 (2019). doi: 10.1038/s41582-019-0157-5; pmid: 30890779

44. M. Costa-Mattioli, P. Walter, The integrated stress response: From mechanism to disease. Science **368**, eaat5314 (2020). doi: 10.1126/science.aat5314; pmid: 32327570

45. C. J. Adams, M. C. Kopp, N. Larburu, P. R. Nowak, M. M. U. Ali, Structure and Molecular Mechanism of ER Stress Signaling by the Unfolded Protein Response Signal Activator IRE1. Front. Mol. Biosci. **6**, 11 (2019). doi: 10.3389/fmolb.2019.00011; pmid: 30931312

46. A. Lopata, A. Kniss, F. Löhr, V. V. Rogov, V. Dötsch, Ubiquitination in the ERAD Process. Int. J. Mol. Sci. **21**, 5369 (2020). doi: 10.3390/ijms21155369; pmid: 32731622

47. J. K. Zehmer et al., Targeting sequences of UBXD8 and AAM-B reveal that the ER has a direct role in the emergence and regression of lipid droplets. J. Cell Sci. **122**, 3694–3702 (2009). doi: 10.1242/jcs.054700; pmid: 19773358

48. J. E. Lee, P. I. Cathey, H. Wu, R. Parker, G. K. Voeltz, Endoplasmic reticulum contact sites regulate the dynamics of membraneless organelles. Science **367**, eaay7108 (2020). doi: 10.1126/science.aay7108; pmid: 32001628

49. E. Tresse et al., VCP/p97 is essential for maturation of ubiquitin-containing autophagosomes and this function is impaired by mutations that cause IBMPFD. Autophagy **6**, 217–227 (2010). doi: 10.4161/auto.6.2.11014; pmid: 20104022

50. K. Zhang et al., Stress Granule Assembly Disrupts Nucleocytoplasmic Transport. Cell **173**, 958–971.e17 (2018). doi: 10.1016/j.cell.2018.03.025; pmid: 29628143

51. M. D. Figley, G. Bieri, R. M. Kolaitis, J. P. Taylor, A. D. Gitler, Profilin 1 associates with stress granules and ALS-linked mutations alter stress granule dynamics. J. Neurosci. **34**, 8083–8097 (2014). doi: 10.1523/JNEUROSCI.0543-14.2014; pmid: 24920614

52. C. M. Rose et al., Highly Multiplexed Quantitative Mass Spectrometry Analysis of Ubiquitylomes. Cell Syst. **3**, 395–403.e4 (2016). doi: 10.1016/j.cels.2016.08.009; pmid: 27667366

53. P. Xu et al., Quantitative proteomics reveals the function of unconventional ubiquitin chains in proteasomal degradation. Cell **137**, 133–145 (2009). doi: 10.1016/j.cell.2009.01.041; pmid: 19345192

54. D. Bustos, C. E. Bakalarski, Y. Yang, J. Peng, D. S. Kirkpatrick, Characterizing ubiquitination sites by peptide-based immunoaffinity enrichment. Mol. Cell. Proteomics **11**, 1529–1540 (2012). doi: 10.1074/mcp.R112.019117; pmid: 22729469

55. J. K. Eng et al., A deeper look into Comet—Implementation and features. J. Am. Soc. Mass Spectrom. **26**, 1865–1874 (2015). doi: 10.1007/s13361-015-1179-x; pmid: 26115965

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/372/6549/eabf6548/suppl/DC1
Figs. S1 to S6
Table S1
Movies S1 to S12

EPIGENOME

MeCP2 is a microsatellite binding protein that protects CA repeats from nucleosome invasion

Abdulkhaleg Ibrahim†, Christophe Papin†, Kareem Mohideen-Abdul†, Stéphanie Le Gras, Isabelle Stoll, Christian Bronner, Stefan Dimitrov*, Bruno P. Klaholz*, Ali Hamiche*

INTRODUCTION: Rett syndrome is a severe neurodevelopmental disorder that is mainly caused by mutations in the methyl-CpG-binding protein 2 gene (*MeCP2*). Initially, MeCP2 was identified as an essential brain protein that binds to methylated CpG (mCG) via its methyl-binding domain (MBD) and acts as transcriptional repressor. However, during early brain development, the postnatal accumulation of MeCP2 parallels the genome-wide high-level accumulation of hydroxymethylcytosine (hmC) and methylated CpA (mCA), suggesting that MeCP2 may also recognize and bind to DNA sequences that contain these modified nucleotides.

The ability of MeCP2 to recognize both mCA and hmC as well as mCG has led to conflicting conclusions regarding its function in transcriptional regulation, because these cytosine modifications are associated with either repression (mCA and mCG) or activation (hmC) of transcription. The unambiguous identification of the MeCP2 target sequence(s) would help to clarify this issue.

RATIONALE: CA repeats (CAn) represent ~1% of the mouse genome and belong to the microsatellite family. They are widely distributed throughout the genome and have been shown to affect transcription of nearby genes. Our recent data reveal that CAn are methylated (mCAn) or hydroxymethylated (hmCAn) in various cell types. In a search for proteins that could specifically recognize and bind these CA repeats, we identified MeCP2 as a specific reader of CA repeats. We hypothesized that the methylation status of CAn is essential for the recognition and binding of MeCP2, possibly through recognition of the modified nucleotides in CA repeats with distinct affinities, relevant for its neuronal function in transcriptional regulation.

RESULTS: Here we show that within the MBD family, MeCP2 is the only protein that specifically recognizes and binds to CA repeats, with much stronger affinity than mCG and mCA. MeCP2 selectively recognizes CA repeat DNA in a strand-specific manner and requires at least five consecutive CA dinucleotides to optimally bind DNA. While MeCP2 can bind in vitro to modified and nonmodified CA repeats, it exhibits impressive selectivity toward hydroxymethylated CA repeats, which are modified by DNA (cytosine-5)-methyltransferase 3A. The modified cytosine, only when located within a CA repeat, serves as a nucleation point for both MeCP2 accumulation and spreading around the repeat, which, in turn, correlates with nu-cleosome exclusion. In addition, loss of MeCP2 results in widespread increase in nucleosome density within lamina-associated domains (LADs) and transcriptional dysregulation of CA repeat–enriched genes located outside LADs.

We have also dissected the molecular basis of the MeCP2 hydroxymethylated CA repeat recognition by solving the crystal structure of MeCP2 in complex with hmCAn. The CA repeat creates a well-defined DNA shape, with a considerably modified geometry, including a widened major groove and negative roll parameters, located precisely at the modification site. We show that the molecular recognition of the hydroxymethylated CA repeat specifically occurs through Arg[133], a key MeCP2 residue whose mutation causes Rett syndrome.

CONCLUSION: Our work identifies MeCP2 as a hydroxymethylated CA repeat DNA binding protein that targets the 5hmC-CA-rich sequence, which are specifically located on one strand. Our data provide insights into the origin of Rett syndrome at the molecular level and suggest that this neurodevelopmental disorder could be viewed as a chromatin disease, originating from the inability of mutant MeCP2 to bind and protect the CA repeats from nucleosome invasion. Our results open a previously unexplored area of research focused on understanding the role of specific protein binding to microsatellites and other repeats in neurological diseases of unknown etiology. ∎

S READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abd5581

WT cells **KO cells**
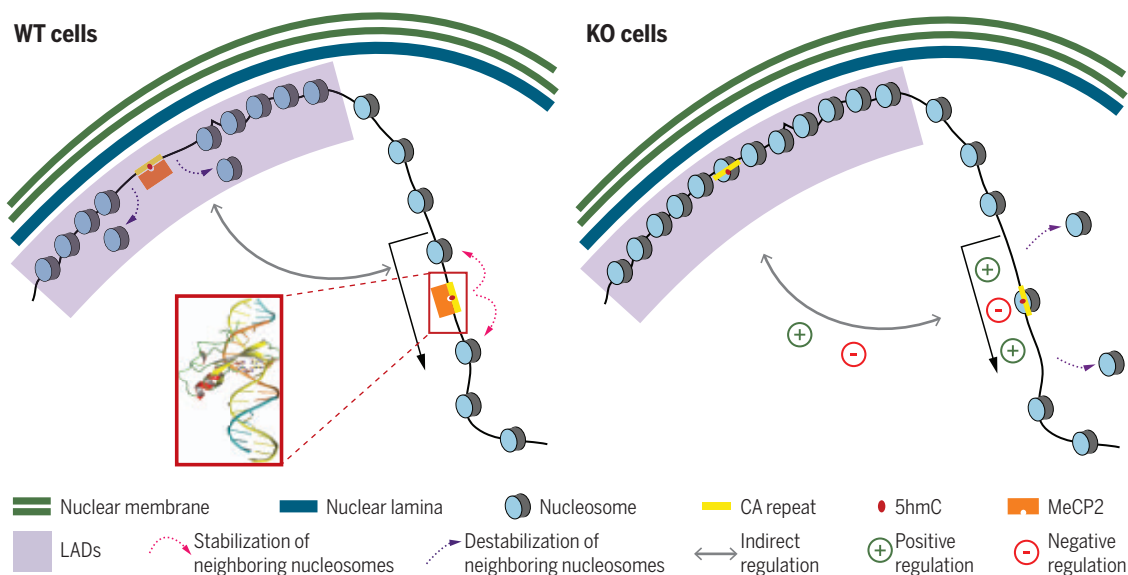
The Rett syndrome protein MeCP2 is a microsatellite CA repeat binding protein regulating chromatin architecture. MeCP2 is a microsatellite binding protein that specifically recognizes hydroxymethylated CA repeats. Depletion of MeCP2 alters the chromatin organization of CA repeats and LADs and results in nucleosome accumulation on CA repeats and genome-wide transcriptional dysregulation. WT, wild-type; KO, knockout.

Nuclear membrane · Nuclear lamina · Nucleosome · CA repeat · 5hmC · MeCP2
LADs · Stabilization of neighboring nucleosomes · Destabilization of neighboring nucleosomes · Indirect regulation · Positive regulation · Negative regulation

EPIGENOME

MeCP2 is a microsatellite binding protein that protects CA repeats from nucleosome invasion

Abdulkhaleg Ibrahim[1,2,3,4,5,6]†, Christophe Papin[1,2,3,4,5]†, Kareem Mohideen-Abdul[1,3,4,5,7]†, Stéphanie Le Gras[1,3,4,5], Isabelle Stoll[1,2,3,4,5], Christian Bronner[1,2,3,4,5], Stefan Dimitrov[8,9]*, Bruno P. Klaholz[1,3,4,5,7]*, Ali Hamiche[1,2,3,4,5,7,10]*

The Rett syndrome protein MeCP2 was described as a methyl-CpG-binding protein, but its exact function remains unknown. Here we show that mouse MeCP2 is a microsatellite binding protein that specifically recognizes hydroxymethylated CA repeats. Depletion of MeCP2 alters chromatin organization of CA repeats and lamina-associated domains and results in nucleosome accumulation on CA repeats and genome-wide transcriptional dysregulation. The structure of MeCP2 in complex with a hydroxymethylated CA repeat reveals a characteristic DNA shape, with considerably modified geometry at the 5-hydroxymethylcytosine, which is recognized specifically by Arg[133], a key residue whose mutation causes Rett syndrome. Our work identifies MeCP2 as a microsatellite DNA binding protein that targets the 5hmC-modified CA-rich strand and maintains genome regions nucleosome-free, suggesting a role for MeCP2 dysfunction in Rett syndrome.

R ett syndrome (RTT) is a severe neurological disorder that occurs almost exclusively in girls (1, 2). RTT is caused mainly by single amino acid mutations (~45%) or C-terminal truncations (~40%) in methyl-CpG-binding protein 2 (MeCP2) (3–5). MeCP2 is essential for normal brain function (6–8). MeCP2 consists of several domains, including the methyl-CpG-binding domain (MBD), the transcriptional repression domain (TRD), the NCoR-SMRT interaction domain (NID), and the AT-hook motifs (AT) (9). MBD and NID are both important for MeCP2 function (10–12). Initially, MeCP2 was identified as a factor that would interact specifically via MBD with DNA bearing methylated CpG dinucleotides (mCG) (11–14). However, recent data show that MeCP2 interacts with methylated CpA (mCA) and hydroxymethylated CpA (hmCA) dinucleotides (15–17), while its ability to bind

hydroxymethylated CpG (hmCG) is controversial (18–20), raising the question of which DNA represents the specific target involved in regulation and whether the chemical modification is specific for a particular function of MeCP2 in neurons [reviewed in (21)].

The sites that harbor mCA modification are frequently found in long genes that are preferentially expressed in neurons (15). mCA sites are concentrated within gene bodies, which indicates that MeCP2 may obstruct transcription within transcribed regions (22). A recent study suggests that mCA sites are methylated during early postnatal life by DNA (cytosine-5)-methyltransferase 3A (DNMT3A) (23).

We have recently found cytosine modification enrichment at CA repeats [also known as (CA)n microsatellite repeats] and have shown that these repetitive elements are preferentially methylated in mouse embryonic stem cells (mESCs) but hydroxymethylated in mouse embryonic fibroblasts (MEFs) (24). CA repeats, which vary in size from 10 to 2000 base pairs (bp), represent ~1% of the mouse genome and belong to the microsatellite family.

In a search for proteins that could specifically recognize and bind the CA repeats, we identified MeCP2 as a specific reader of CA repeats. We dissected the molecular basis of hydroxymethylated CA repeat recognition by solving the structure of MeCP2 in complex with a hydroxymethylated CA repeat. Genomics and transcriptomics data reveal that MeCP2 regulates CA repeat–enriched genes by maintaining their microsatellites nucleosome-free. We show that MeCP2 is essential for the organization of lamina-associated domain (LAD) chromatin and consequently for both LAD function and transcription at a genome-wide level. Our study

sheds light on the molecular mechanism underlying RTT and the essential features of MeCP2 by identifying it as CA repeat–binding protein that modulates chromatin architecture at a distance from the transcription start site (TSS).

MeCP2 binds to hydroxymethylated CA repeats in fibroblasts

To search for proteins that recognize and bind to CA repeats, we used pulldown assays with biotinylated primers having at least one helical turn of DNA and containing seven consecutive CpA dinucleotides [(CA)7], or non-CA sequences as a control. Affinity purification using biotinylated DNA and SDS–polyacrylamide gel electrophoresis (SDS-PAGE) analysis of DNA-bound complexes purified from HeLa nuclear extracts reveals a distinct protein band—absent in the control—migrating below 70 kD (fig. S1A). Mass spectrometry analysis identifies this band as MeCP2, a complex multidomain protein proposed to bind methylcytosine DNA (11–17) (Fig. 1A).

This unexpected finding prompted us to analyze the genomic distribution of MeCP2 and its relationship with the methylation and oxidation pattern of DNA. We created knockout (KO) MeCP2 cells upon transfection of $Mecp2^{fl/fl}$ MEFs [derived from homozygous floxed mice (6)], with vectors expressing Cre recombinase (fig. S1, B and C); $Mecp2^{-/-}$ MEFs served as a negative control. The MeCP2 chromatin immunoprecipitation sequencing (ChIP-seq) identified 4113 wild type (WT)–specific peaks. Notably, two-thirds of these peaks localize to CA repeats (Fig. 1B), which have an average CA density of 22% (fig. S1D). The remaining 30% of the WT-specific peaks, although they were not annotated in the RepeatMasker as CA repeats (see Materials and methods), show a significant enrichment in CpA dinucleotide repeats of variable length (Fig. 1C and fig. S1D). A statistically dominant motif of these repeats is presented in fig. S1E, which is a CA repeat. Unexpectedly, CpG dinucleotides are under-represented in these regions (Fig. 1C and fig. S1D). These data suggest that it is the local CA density, rather than CG density, that is responsible for recruiting MeCP2 at these regions in MEFs.

To correlate the distribution of MeCP2 to the methylation status of CA repeats, we developed a novel strand-oriented (hydroxy)methylated DNA immunoprecipitation sequencing (MeDIP-seq) protocol in WT and KO MeCP2 MEFs (see Materials and methods), which enabled mapping the strand-specific genome-wide patterns of 5-methylcytosine and 5-hydroxymethylcytosine (5mC and 5hmC, respectively). As expected, because CpG dinucleotides are symmetrical, 5mC modifications were found to accumulate in an equivalent manner between the two complementary DNA strands (Fig. 1D and fig. S1F).

[1]Institute of Genetics and of Molecular and Cellular Biology (IGBMC), 67400 Illkirch, France. [2]Department of Functional Genomics and Cancer, IGBMC, CNRS, INSERM, Université de Strasbourg, 67404 Illkirch, France. [3]Centre National de la Recherche Scientifique (CNRS), UMR 7104, 67404 Illkirch, France. [4]Institut National de la Santé et de la Recherche Médicale (INSERM), U964, 67404 Illkirch, France. [5]Université de Strasbourg, 67404 Illkirch, France. [6]Biotechnology Research Center (BTRC), 30303 Tripoli, Libya. [7]Centre for Integrative Biology (CBI), Department of Integrated Structural Biology, IGBMC, CNRS, INSERM, Université de Strasbourg, 67404 Illkirch, France. [8]Université Grenoble Alpes, CNRS UMR 5309, INSERM U1209, Institute for Advanced Biosciences (IAB), Site Santé - Allée des Alpes, 38700 La Tronche, France. [9]Roumen Tsanev Institute of Molecular Biology, Bulgarian Academy of Sciences, Sofia 1113, Bulgaria. [10]Center of Excellence in Bionanoscience Research, King Abdulaziz University (KAU), Jeddah 21589, Saudi Arabia.
*Corresponding author. Email: hamiche@igbmc.fr (A.H.); klaholz@igbmc.fr (B.P.K.); stefan.dimitrov@univ-grenoble-alpes.fr (S.D.)
†These authors contributed equally to this work.

Fig. 1. MeCP2 binds to hydroxymethylated CA repeats in mouse embryonic fibroblasts. (**A**) Functional domains of MeCP2. (**B**) Pie charts of the genomic location distribution of MeCP2, 5mC, and 5hmC peaks in MEFs. More than two-thirds of the MeCP2 peaks (*n* = 4113) and more than half of the 5hmC peaks (*n* = 2371) were found at CA repeats. 5'UTR, 5' untranslated region; CGI, CpG island; 3'UTR, 3' untranslated region; TTS, transcription termination site. (**C**) Heatmaps of MeCP2, 5mC, 5hmC, CA, and CG densities at MeCP2 peaks (±1 kb) in WT and KO cells. MeCP2-bound CA repeats are specifically hydroxymethylated in fibroblasts. mAb, monoclonal antibody; pAb, polyclonal antibody. (**D** and **E**) Heatmaps of CA, TG, and CG densities (left panels) and modified cytosine densities (right panels) around 5mC (D) and 5hmC peaks (E) in WT cells. Peaks were annotated using Homer bioinformatic tools and sorted according to their genomic location. Strand-specific analysis of DIP-seq reads reveals that 5mC is found at both strands of CG-rich sites, whereas 5hmC specifically targets the CA-rich strands of CA repeats. (**F**) Genome browser view at representative locus highlighting the strong

correlation between 5hmC, MeCP2, and CA repeats patterns in MEFs (highlighted in yellow). (**G**) Average 5mC, 5hmC, and MeCP2 levels at CA repeats as a function of their CpA dinucleotide density in WT and KO cells. CA repeats were sorted by their number of CpA dinucleotides and binned into 30 equal-sized bins (~10,000 CA repeats per bin), and the density values for each bin were averaged and plotted. A positive correlation between CpA, 5hmC, and MeCP2 densities was observed.

However, CA- or TG-rich regions exhibited a very clear asymmetrical 5hmC distribution targeting exclusively the CA-rich strand (Fig. 1E and fig. S1F). Accordingly, CA repeats concentrated more than half of the 5hmC peaks (Fig. 1B).

Integrative analyses of the MeCP2 ChIP-seq and DIP-seq data reveal that all the MeCP2-bound CA repeats are hydroxymethylated in MEFs (Fig. 1, C and F, and fig. S1G). The steady-state level of 5hmC at CA repeats was not affected by the absence of MeCP2, suggesting that MeCP2 is a reader of 5hmC (Fig. 1, C and F, and fig. S1G). A positive correlation between

CA, 5hmC, and MeCP2 densities was also observed at CA repeats, an observation that further underlines the genome-wide targeting of MeCP2 toward 5hm-modified CA repeats (Fig. 1G and fig. S1, H and I).

MeCP2 is the only protein from the MBD family that selectively binds to CA repeats

CA repeats are methylated in mESCs and oxidized during differentiation (*24*). To test whether other members of the methyl-CpG binding protein family, namely MBD1, MBD2, MBD3, and MBD4, are able to specifically interact with methylated CA repeats, we analyzed their genome-

wide distributions using available ChIP-seq data (*25*) and compared them with both the 5mC and 5hmC genome-wide pattern distributions in mESCs (*26*). Normalized densities of the different MBD proteins and their corresponding mutants at 5mC and 5hmC peaks show that, with the exception of MBD3, all MBD proteins are enriched at both 5mC and 5hmC peaks. However, only MeCP2 was found enriched at methylated CA repeats (mESCs do not show hydroxymethylation on CA repeats) (fig. S2A). Notably, MeCP2 was depleted from CA repeats in mESCs deficient for the three active DNA methyltransferases Dnmt1, Dnmt3a,
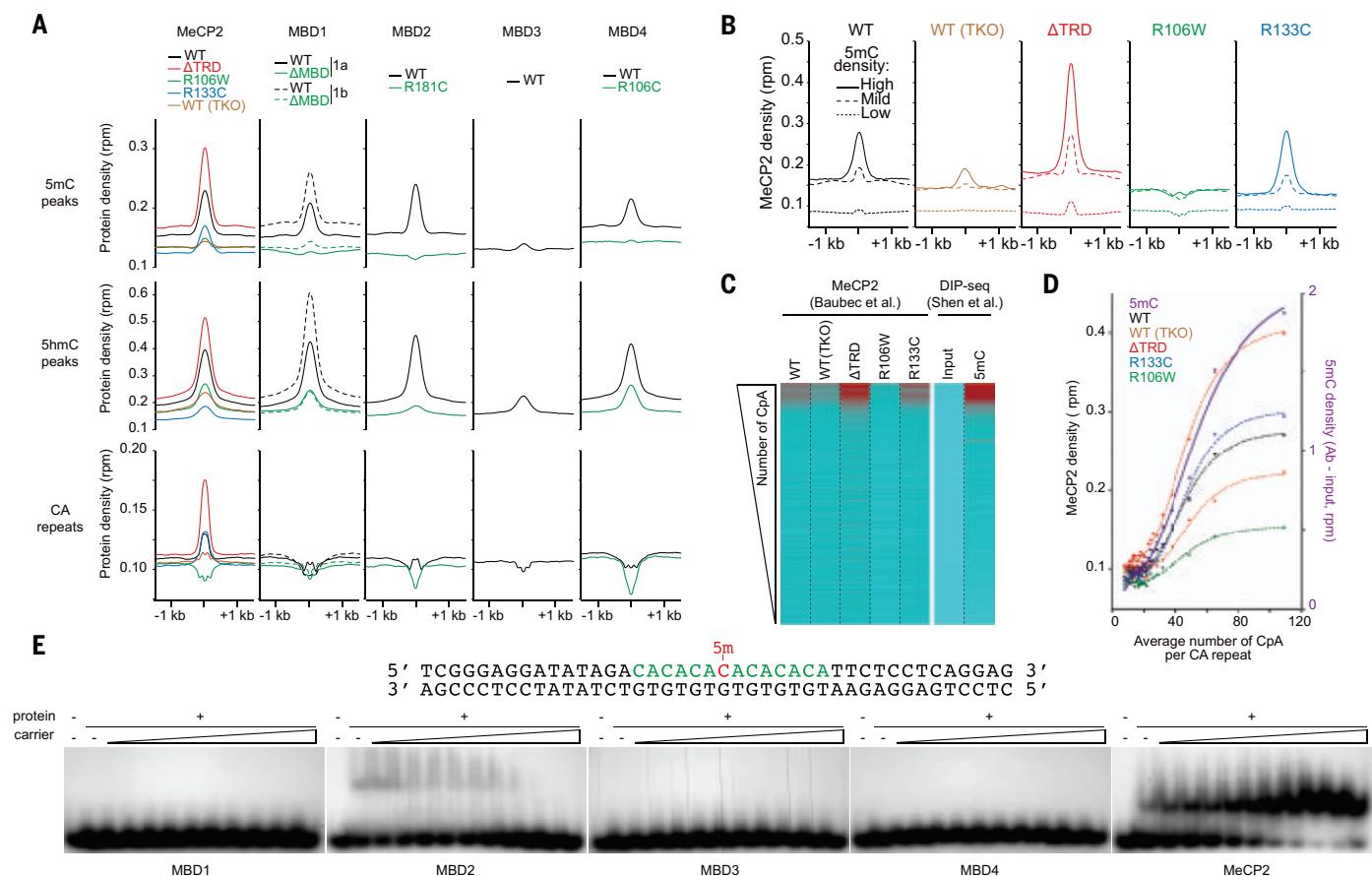
Fig. 2. Highly specific and methylation-dependent localization of MeCP2 at CA repeats in embryonic stem cells. (**A**) DIP-seq datasets [from (*26*)] and ChIP-seq datasets [from (*25*)] obtained in ESCs were reanalyzed and cross-compared. Normalized densities of MBD proteins and their corresponding isoforms or disease-related mutants at 5mC peaks, 5hmC peaks, and CA repeats. TKO, triple knockout. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. (**B**) Average MeCP2 densities of the WT or mutant proteins at CA repeats, clustered according to their 5mC enrichment (high, mild, or low).

(**C** and **D**) Heatmaps (C) and quantification (D) of MeCP2 and its corresponding mutants and 5mC levels at CA repeats ranked in descending order on the basis of their number of CpA dinucleotides. (D) CA repeats were sorted by their number of CpA dinucleotides and binned into 30 equal-sized bins, and the density values for each bin were averaged and plotted. A positive correlation between CA, 5mC, and MeCP2 densities was observed. (**E**) EMSA analyses of the MBD domain of MBD1, MBD2, MBD3, MBD4, and MeCP2 (400 nM) bound to ^{32}P-end-labeled 5mCA repeat DNA probe incubated with an increasing excess of competitive DNA (from 5- to 2500-fold excess). Of note, a minimal concentration of carrier DNA is required for the MBD domain to bind efficiently to radiolabeled DNA.

and Dnmt3b (*Dnmt*-TKO) (Fig. 2, A to D, and fig. S2, A and B), revealing that methylation is required for its recruitment and/or stable association with CA repeats. In line with these results, Arg^{106}→Trp (R106W), the DNA binding dead mutant of MeCP2, was not found enriched at methylated CA repeats (Fig. 2, A to D, and fig. S2, A and B). However, the R133C mutant was found enriched at methylated CA repeats (Fig. 2, A to D, and fig. S2, A and B). In addition, the deletion mutant lacking the entire TRD domain (ΔTRD), which is known to contribute to nonspecific binding of MeCP2 to DNA (*27*), exhibits clear enrichment on CA repeats (Fig. 2, A to D, and fig. S2, A and B), a result that further confirms the specific CA repeat reader activity of the MeCP2-MBD identified in this work, consistent with the super-shifts

observed using electrophoretic mobility shift assay (EMSA) (Fig. 2E).

MeCP2 regulates transcription of CA repeat–enriched genes

CA repeats are widely distributed throughout the genome where they could affect gene expression of nearby genes (*24*, *28*, *29*). To gain more insights into their function, we asked whether the binding of MeCP2 to hydroxymethylated CA repeats could be involved in gene regulation. Differential gene expression and gene ontology analyses in the WT versus KO MEFs reveal a large panel of dysregulated genes, including synaptic genes, in the absence of MeCP2 (Fig. 3, A and B). These genes were found enriched in CA repeats, regardless of whether they were up- or down-regulated

(Fig. 3C). The impact of MeCP2 on their misregulation levels was found to be directly proportional to their CA repeat content (Fig. 3D).

MeCP2 mediates its neuronal function through modified CA repeats

The data gathered from MEF cells suggest that MeCP2 could regulate expression of neuronal genes, which are highly expressed in the brain, through binding to their CA repeats. To address this possibility, we analyzed the distribution of MeCP2 in the brain using available ChIP-seq data obtained from both mouse brain tissues (whole brain, forebrain, and cerebellum) and cultured cortical neurons (*15*, *30*, *31*). In agreement with the MEF cell data, MeCP2 was found specifically enriched at, and associated with, CA repeats in brain
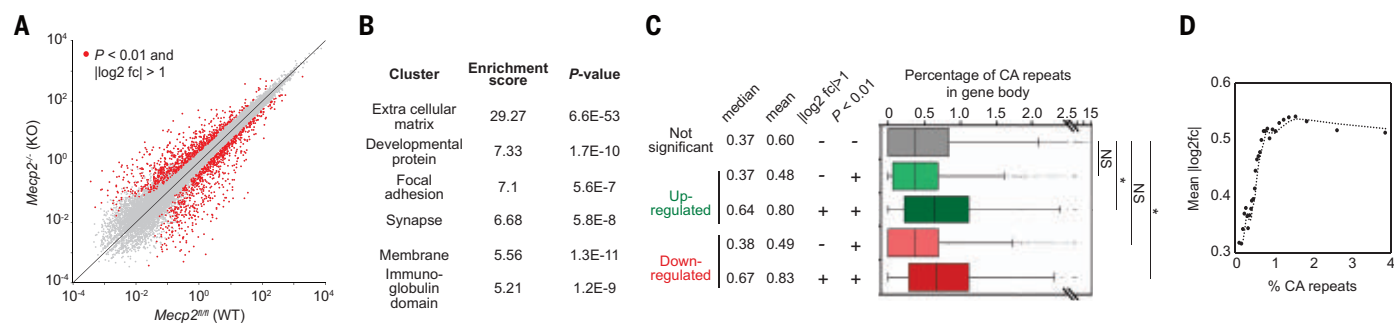
Fig. 3. MeCP2 controls the expression of CA repeat–rich genes in fibroblasts. (**A**) Scatterplot comparing global gene expression levels between WT and KO cells. In the absence of MeCP2, 775 genes are overexpressed and 829 genes are down-regulated ($|\log_2$ fold change$| > 1$ and $P < 0.01$). (**B**) Functional annotation clustering of genes differentially expressed in the absence of MeCP2 in MEFs. (**C**) Misregulated genes are enriched in CA repeats. Distribution of the percentage of CA repeats per gene (CA repeat length divided by gene body length multiplied by 100) for genes grouped according to their mRNA expression change in the absence of MeCP2 (*$P < 10^{-4}$, Mann-Whitney-Wilcoxon test). (**D**) Mouse coding genes were sorted according to their percentage of CA repeats and binned into 30 equal-sized bins. The absolute value of the mean fold change was averaged and plotted for each bin.

(Fig. 4A and fig. S3A), suggesting that MeCP2 could function by binding to CA repeats in neurons. Both the analysis of independent gene expression datasets (fig. S3B) and the integrative analysis of 38 different transcriptomic studies (Fig. 4B) performed on a wide range of neural tissues or cells from RTT mouse models further confirmed an enrichment in CA repeats of MeCP2-regulated genes in the brain, which contain, on average, five times more CA repeats than the rest of the genes.

We next examined the modification status of CA repeats in the brain by reanalyzing the previously published whole-genome bisulfite sequencing (WGBS) dataset derived from mouse adult frontal cortex (*32*). We focused our analysis on 6-week-old mice in which the 5mCA modification is known to have reached its maximum level (*32*). Among the 381 million CpA dinucleotides found in the mouse genome, only 8 million fall within CA repeats. We found that CA dinucleotides located within CA repeats are globally more enriched in cytosine modifications than are CA dinucleotides located elsewhere in the genome (11.9% versus 3.6%) (Fig. 4C). Although some CA repeats are found to be highly modified (Fig. 4D), the average modification level of modified CA dinucleotides (modCA) is ~10% at CA repeats (fig. S3C), suggesting that high sequencing coverage is required to efficiently detect modCA. While this dataset gives an average coverage of ~20× (10× per strand), it decreases to 5× per CA-strand for CA repeats (fig. S3D). Only 3% of CA dinucleotides located within CA repeats (CAn) are covered by more than 10 reads (compared with ~40% for CA dinucleotides located elsewhere in the genome) (fig. S3E). CA repeats are poorly covered by WGBS, yet a large fraction of the covered CA repeats are found modified, suggesting that the number of modCA identified in the adult brain is largely underestimated. We next compared the relative enrichment of MeCP2 around nonmodified

and modified CG dinucleotides, CA dinucleotides, and CA repeats using WGBS (*32*) and ChIP-seq (*15*) datasets from matching brain regions (6-week-old mouse frontal cortex). In agreement with previous studies (*15*, *17*), MeCP2 was detected at similar levels at modified CG and CA dinucleotides as well as at unmodified CA located inside or outside CA repeats. Importantly, MeCP2 preferentially accumulates at modified cytosines inside CA repeats, where the modification serves as a nucleation point for further MeCP2 molecule binding within a 2-kb region around the CA repeat (Fig. 4E).

We then asked whether CA repeats are methylated and/or hydroxymethylated in neurons. Because WGBS does not distinguish between 5mC and 5hmC, we reanalyzed bisulfite sequencing (BS-seq; which detects both 5mC and 5hmC) and oxidative bisulfite sequencing (oxBS-seq; which only detects 5mC) datasets derived from isolated granule cells (*16*) and performed a comparative analysis to estimate the 5mC/5hmC ratio at CA repeats. The BS-seq dataset reveals that CA dinucleotides found in CA repeats are strongly modified in comparison with CA dinucleotides located elsewhere (31% versus 1%) (Fig. 4F), as shown in frontal cortex. In oxBS-seq datasets, the percentage of modCA in CA repeats decreases to 25%, suggesting that, in granule cells, modified CA repeats are ~80% methylated and ~20% hydroxymethylated. Of note, the average modification level of modCA within CA repeats increases to ~80% in granule cells (versus 10% in frontal cortex) (fig. S3C). This difference could be explained by the fact that (i) isolated granule cells are essentially neurons, whereas frontal cortex contains both neuronal and glial cells, and (ii) neurons contain five times more 5mCA than do glia in adult brain (*32*).

We also reanalyzed the WGBS dataset derived from DNMT3A$^{-/WT}$ mice cortex and found that the heterozygous disruption of DNMT3A leads to a global disruption of CA

methylation both outside and inside CA repeats in mouse cortex, supporting the implication of DNMT3A in CA repeat methylation (*33*). The high level of CA repeat modifications in neurons is in line with the function of MeCP2 exercised through its recognition and binding to hydroxymethylated CA.

MeCP2 is a long-range chromatin organizer that protects CA repeats from nucleosome invasion

MeCP2 has long been viewed as a global repressor of transcription. Recent evidence suggests that MeCP2 could regulate transcription through a more complex mechanism that involves global chromatin organization [reviewed in (*34*, *35*)]. To shed light on how genome-wide nucleosome positioning, a key feature of chromatin architecture, is related to the presence of MeCP2, we performed genome-wide paired-end micrococcal nuclease sequencing (MNase-seq) analysis comparing chromatin samples derived from WT and KO MeCP2 MEFs. We found similar nucleosomal DNA length distributions between WT and KO cells (fig. S4A) as well as very high nucleosome occupancy at microsatellites in knockout cells (fig. S4B). A detailed analysis of nucleosome occupancy patterns around microsatellites reveals that these repetitive elements are nucleosome-free in WT cells and that their boundaries define nucleosome positioning upstream and downstream of the repeats (Fig. 5, A to C). Loss of MeCP2 results in accumulation of specific nucleosomes within CA repeats but not on other microsatellite repeats (Fig. 5, A to C), suggesting that MeCP2 protects CA repeats from nucleosome binding. Moreover, knockout of MeCP2 leads to a widespread increase of nucleosome density and occupancy inside LADs, while outside LADs, the nucleosome density is decreased (Fig. 5D).

We next studied the effect of MeCP2 on nucleosome organization around the CA
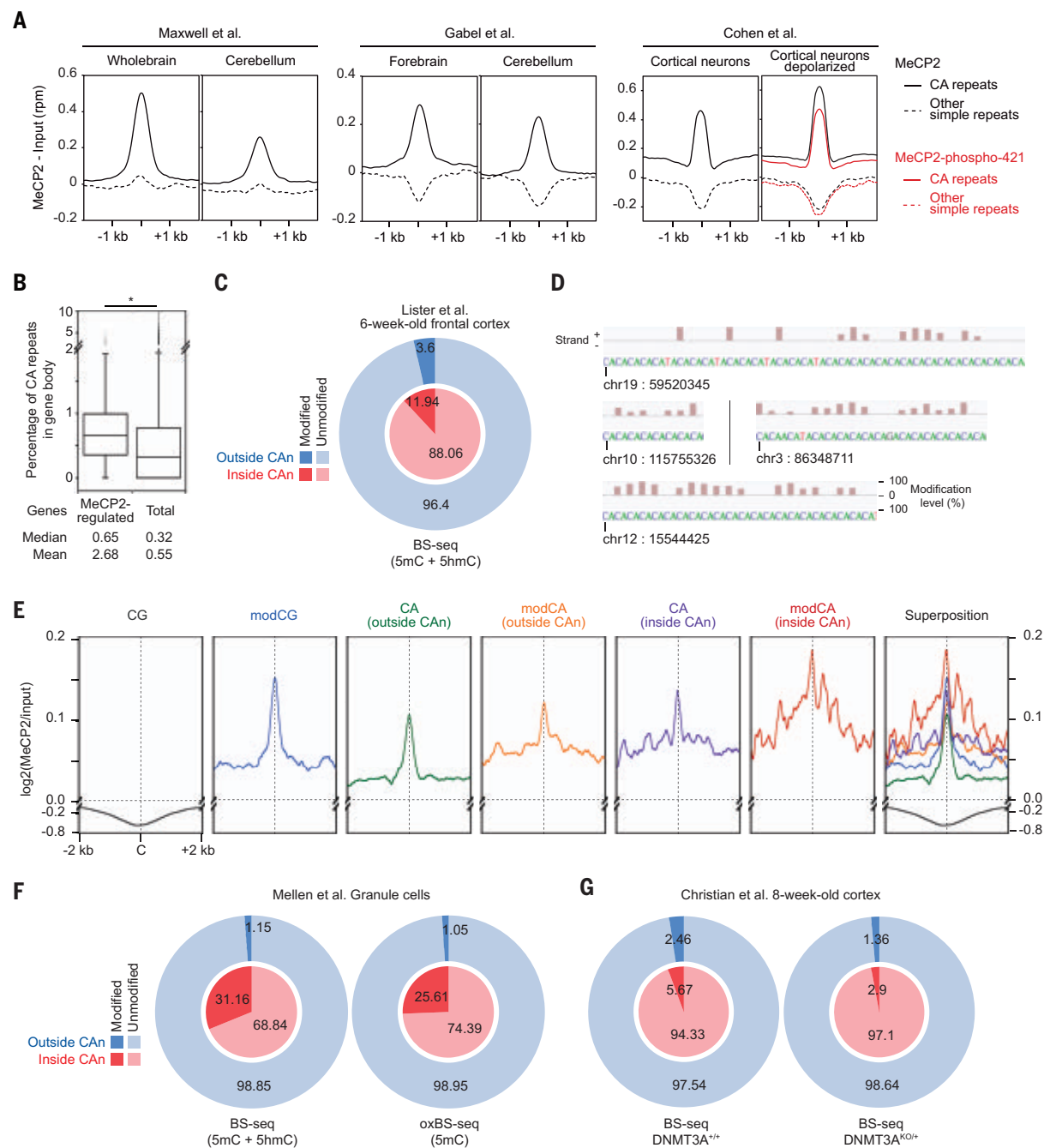
Fig. 4. MeCP2 mediates its neuronal function through CA repeats.
(**A**) MeCP2 distribution in brain tissues. Reanalysis of seven published MeCP2 ChIP-seq datasets performed in five different brain tissues (*15, 30, 31*). Average densities of MeCP2 at microsatellite regions (CA repeats or other simple repeats) reveal that MeCP2 binds specifically to CA repeats in brain tissues. (**B**) Misregulated genes in MeCP2-deficient brain are enriched in CA repeats. Distribution of the percentage of CA repeats per gene body for mouse protein-coding genes (total) and MeCP2-regulated genes. Genes regulated by MeCP2 were identified after an integrative analysis of 38 different transcriptomic studies performed in a wide range of neural tissue and cells from RTT mouse models (*62*) (*$P < 10^{-4}$, Mann-Whitney-Wilcoxon test). (**C**) CA dinucleotides found in CA repeat context are strongly modified in mouse cortex. Global modCA/CA level (%) inside or outside CA repeat in mouse 6-week-old frontal cortex based on the WGBS dataset derived from (*32*).

(**D**) Browser representation showing the modification level at selected highly modified CA repeats in 6-week-old mouse frontal cortex [from (*32*)]. chr, chromosome. (**E**) MeCP2 preferentially accumulates around modified CA found in CA repeat context in mouse cortex. MeCP2 distribution around (±2 kb) modified or unmodified CG or CA using ChIP-seq (*15*) and WGBS (*32*) datasets from matching brain regions (6-week-old mouse frontal cortex). Only cytosines that were covered by at least 10 reads were included. (**F**) CA repeats are found both methylated and hydroxymethylated in neurons. Global modCA/CA level (%) inside or outside of CA repeats based on BS-seq and oxBS-seq performed in mouse granule cells (*16*). (**G**) Heterozygous disruption of DNMT3A leads to global disruption of CA methylation both outside and inside CA repeats in mouse cortex. Global modCA/CA level (%) inside or outside of CA repeats based on BS-seq performed in 8-week-old cortex of DNMT3A$^{KO/+}$ and WT mice (*33*).
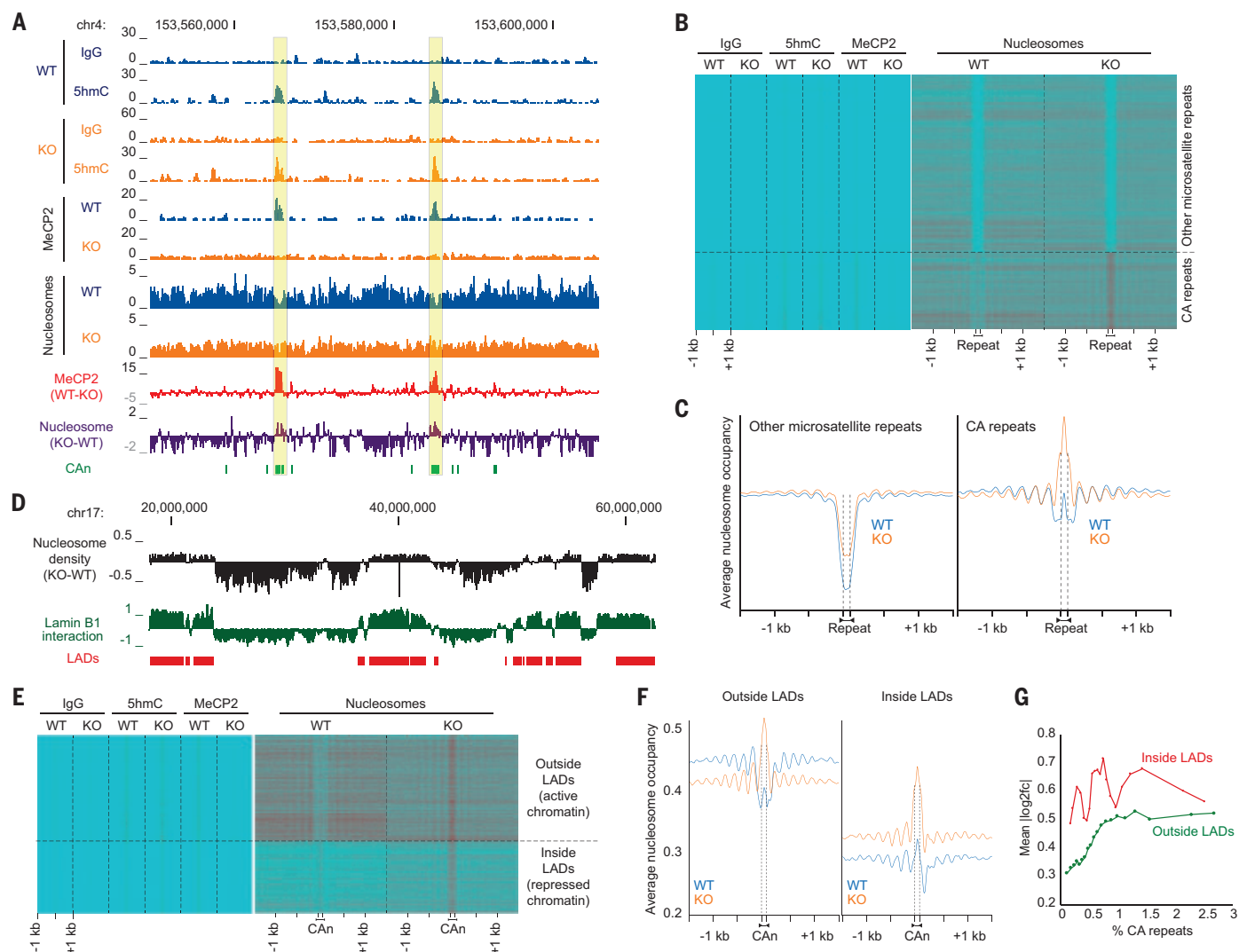
Fig. 5. MeCP2 functions as a chromatin architecture modulator through CA repeats. (**A** to **C**) MeCP2 protects CA repeats from nucleosome invasion. (A) Nucleosomes are incorporated at CA repeats in the absence of MeCP2. 5hmC and MeCP2 profiles (reads per million), nucleosome occupancies (normalized dyad frequency), and CA repeat positions are shown at a representative locus in WT and KO cells. (B) Heatmaps of IgG, 5hmC, and MeCP2 densities (left panels) and nucleosome densities (right panels) around microsatellites in both WT and KO cells. (C) Nucleosome occupancy profiles around CA repeats (right panel) or around other microsatellites (left panel) in both WT and KO cells. Nucleosome dyad densities were collected in 50-bp windows 1.5 kb upstream (divided into 75 bins) and 1.5 kb downstream (divided into 75 bins) of the length-normalized microsatellites (divided into five bins). (**D** to **F**) The absence of MeCP2 leads to widespread increase in nucleosome accessibility inside LADs. (D) Wiggle tracks representing the difference in nucleosome density between WT and KO cells and the log$_2$ ratio of Lamin B1 enrichment in MEFs [from (*41*)] along half of chromosome 17. Genomic regions contacting the nuclear lamina, defined as lamina-associated domains (LADs), are also shown in red. (E) Heatmaps of IgG, 5hmC, and MeCP2 densities (left panels) and nucleosome densities (right panels) around CA repeats in both WT and KO cells. CA repeats were sorted according to their genomic location inside (repressed chromatin) or outside (active chromatin) LADs. (F) Nucleosome occupancy profiles around CA repeats located outside (left panel) or inside (right panel) LADs. (**G**) MeCP2 regulates transcription of CA repeat-rich genes located outside LADs. Mouse coding-genes located inside or outside LADs were sorted according to their CA repeat content (%) and binned in 20 equal-sized bins. The absolute value of the mean log$_2$ fold change was averaged and plotted for each bin.

repeats both outside and inside LADs. We found that the presence of MeCP2 excludes nucleosomes from CA repeats residing outside LADs, while its absence resulted in a global increase in nucleosome density around CA repeats located within LADs (Fig. 5, E and F). Accordingly, the positive correlation between CA repeats content and transcriptional dysregulation is only observed for genes located outside LADs (Fig.

5G). The trend that the majority of the dys-regulated genes are down-regulated inside LADs is consistent with the global accumulation of nucleosomes observed in these regions (Fig. 5D and fig. S4C). We also found altered nucleosome occupancies around TSSs and CTCF binding sites in KO cells (fig. S4D), revealing an essential role of MeCP2 in maintaining chromatin architecture.

Specificity of DNA sequence and chemical modification status for MeCP2 binding

Our in vivo ChIP-seq data show that MeCP2 is bound to hydroxymethylated cytosines within CA repeats. To biochemically analyze how MeCP2 specifically recognizes and interacts with these repeats, we carried out a series of in vitro experiments using EMSA. We assessed the binding of the MeCP2-MBD domain (Fig. 6)
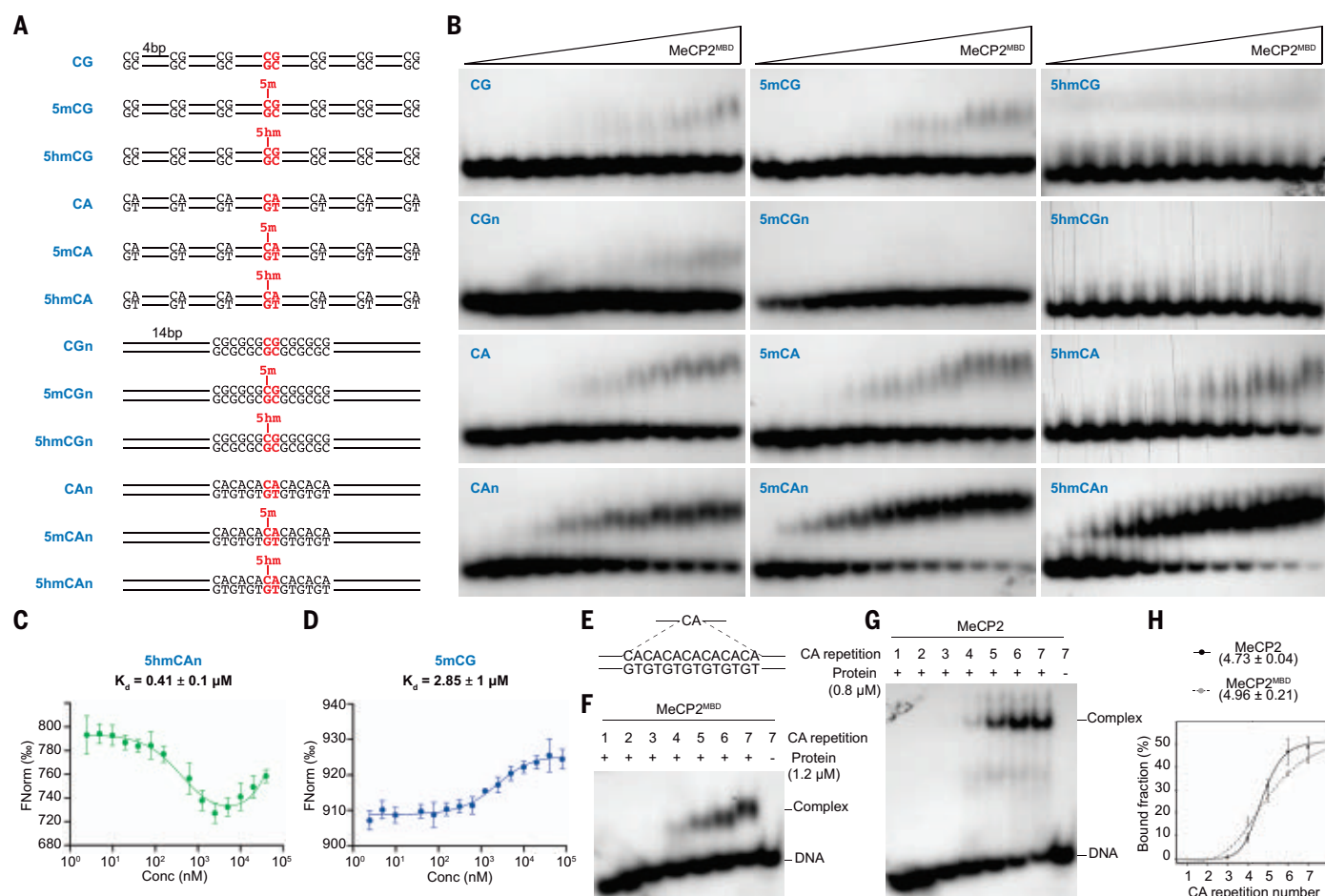
Fig. 6. High affinity of MeCP2 to CA repeat–containing substrates in electrophoretic mobility shift assays. (**A**) Schematic presentation of the different substrates used in this study. Note that each substrate is strictly composed of seven CpA or CpG dinucleotides in a dispersed or repeated context. (**B**) EMSA analysis of the MeCP2-MBD (from 0 to 2.4 μM) bound to 2 nM of ^{32}P-end-labeled substrates [presented in (A)] in the presence of 2000-fold excess of unlabeled competitor. (**C** and **D**) Microscale thermophoresis assay of MeCP2-MBD with

5hmCAn (C) and 5mCG (D) DNA substrates. Differences in the direction of the curves are due to change in the hydration shell. FNorm, normalized fluorescence; Conc, concentration. (**E** to **H**) Binding of MeCP2-MBD (F) or MeCP2 full-length (G) to different DNA substrates containing one to seven CA consecutive dinucleotides [presented in (E)]. The minimal number of CpA repetitions (H) required for an optimal binding of MeCP2 was estimated to be about five CAs (corresponding to 50% of the maximum bound fraction, indicated in brackets).

or full-length protein (fig. S5) to various forms of methylated or hydroxymethylated DNA containing either seven interspersed or grouped CpG dinucleotides or seven CpA dinucleotides (Fig. 6A). We found that both the MBD domain of MeCP2 and the full-length protein have a very weak affinity for DNA containing CpG or CpA dinucleotides, irrespective of their distribution or methylation status (Fig. 6B and fig. S5). In contrast, MeCP2 has a very strong affinity for grouped CpA dinucleotides (Fig. 6B and fig. S5). Methylation enhances the affinity for seven CpA dinucleotide stretches, and hydroxymethylation further enhances that affinity (Fig. 6B and fig. S5). The binding measurement determined by microscale thermophoresis (MST) indicates that the MeCP2-MBD has stronger interactions with 5hmCAn substrate (dissociation constant, K_d, of ~0.41 μM), which is about seven-fold higher than the interaction with 5mCG (K_d ~ 2.85 μM).

We next sought to determine the minimal CA dinucleotide repeats required for MeCP2 binding using EMSA (Fig. 6E). We found that the affinity of MeCP2 for DNA increased linearly with the number of CA repeats, approaching a plateau at about five CA repeats (Fig. 6, F to H). These data reveal that MeCP2 requires at least five consecutive CA dinucleotides to efficiently bind DNA and that hydroxymethylated CA repeats represent the specific DNA target for MeCP2.

Structural basis of CA repeat DNA and hydroxymethyl recognition by MeCP2

To address the molecular basis of DNA recognition by MeCP2 and of the underlying specific interactions with hydroxymethylated CA repeats, we determined the crystal structure of MeCP2 bound to a CA repeat–containing DNA with one hmCA in the middle. Crystals of the complex formed by the MeCP2-MBD domain

and the hmCA-modified DNA (Fig. 7A) diffracted to 2.1-Å resolution. The structure of the complex was determined by molecular replacement using the structure of the MeCP2-CpG complex as a starting model [Protein Data Bank (PDB) ID 3C2I] (36), and the atomic model was refined iteratively to obtain good geometry parameters (see Materials and methods; table S1). MeCP2 adopts a tertiary structure in which its β-sheet region binds asymmetrically to the major groove region of the DNA where the 5hmC is located (Fig. 7A). The overall structure of the MeCP2-MBD protein part in the complex is very similar to that of the protein alone, as determined by nuclear magnetic resonance spectroscopy in solution (37), and its overall shape is similar to that of MeCP2 bound to a methylated CpG (fig. S6A) (36). However, the recognition and binding modes of MeCP2-MBD toward the hydroxymethylated CA repeat and the neighbor bases (G/C versus A/T) and
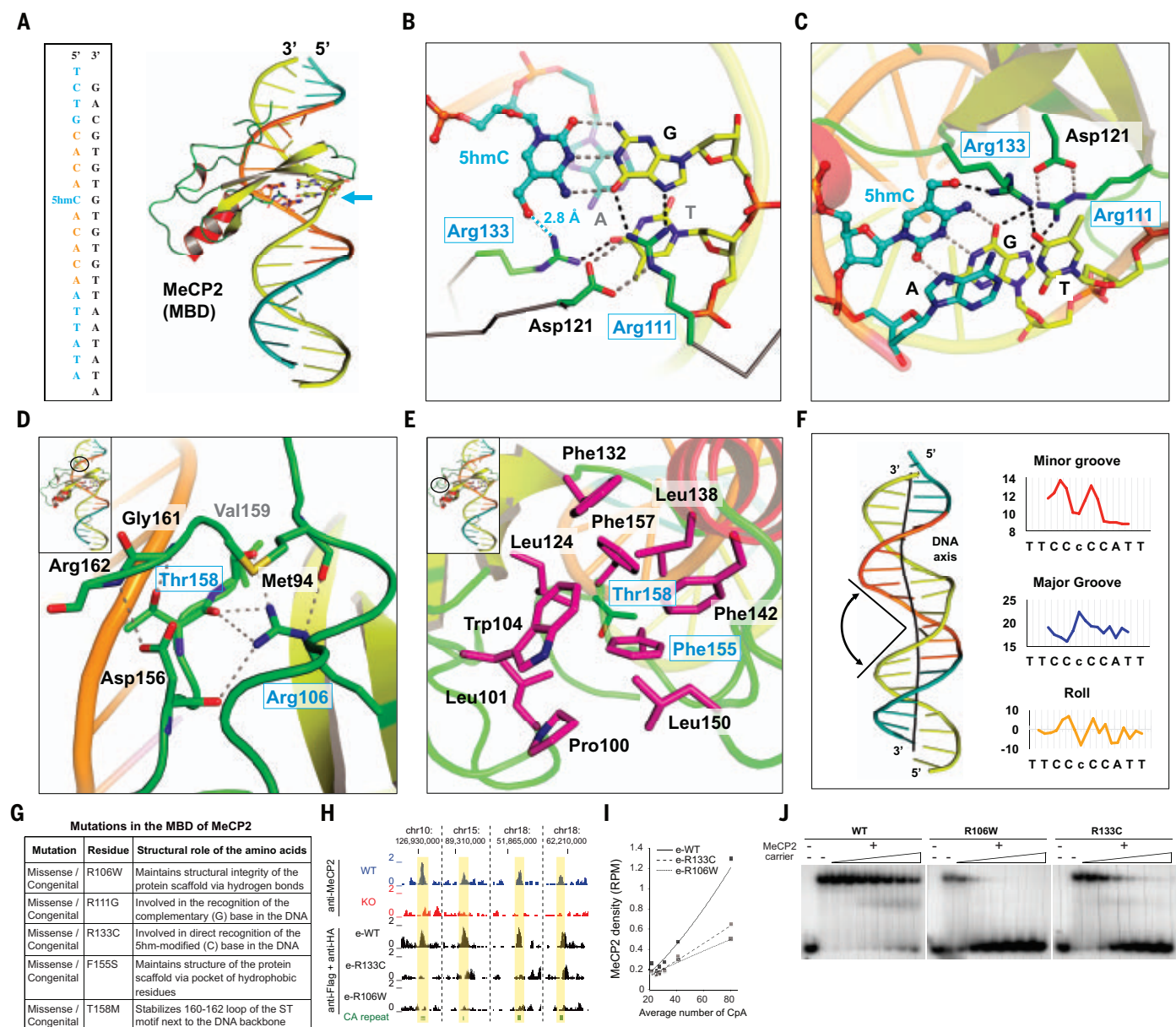
Fig. 7. Crystal structure of MeCP2-MBD complexed with five CA repeat DNA containing a central hydroxymethylated cytosine base. (**A**) Overall structure of the DNA−MeCP2-MBD complex; the leading strand is indicated in cyan with the CA repeat, bearing the central hydroxymethylated cytosine, in orange (arrow). (**B** and **C**) Details of the DNA−MeCP2-MBD molecular recognition (top and side views, respectively). Arg[111] serves as a primary DNA recognition site of the G base, while Arg[133] provides specificity toward the hydroxymethyl group of the modified C base and also recognizes the neighbor T (on the complementary strand) of the CA repeat [color coded as in (A); hydrogen bond interactions are shown as dashed lines]. R133C is a natural mutant of MeCP2 in Rett syndrome (RTT residues are highlighted by blue boxes). (**D** and **E**) Structural environment of sites where other Rett syndrome mutants exist, including Arg[106], Phe[155], and Thr[158], which are involved either in hydrogen bonding or hydrophobic contacts to stabilize the MeCP2-MBD fold and nonspecific interactions with the DNA. Overall structure of DNA−MeCP2-MBD shown as insets with regions of molecular interactions highlighted (black circles). (**F**) DNA base pair parameters (DNA alone shown as cartoon representation). At the hydroxymethylated site, the CA repeat shows a widening of the major groove (black arrow). A slight kink of the central DNA axis is visible on the CA repeat (central black line). At the modified site, major and minor grooves exhibit widening (~22 Å) and narrowing (~10 Å), respectively, and the roll base pair step parameters show a strong negative roll (~−8°). (**G**) List of mutations involved in the MeCP2-MBD and their structural role. (**H**) Genome browser view at selected CA repeats showing the distribution of MeCP2 in MEFs *Mecp2^{fl/fl}* (WT), *Mecp2^{−/−}* (KO), *Mecp2^{−/−} + e-Mecp2 WT* (e-WT), *Mecp2^{−/−} + e-Mecp2 R133C* (e-R133C), and *Mecp2^{−/−} + e-Mecp2 R106W* (e-R106W). (**I**) Average MeCP2 density at CA repeats as a function of their CpA dinucleotides density in e-WT, e-R133C, and e-R106W cells. (**J**) EMSA analyses of the full-length MeCP2 WT, MeCP2 R133C, or MeCP2 R106W mutant proteins (600 nM) to ^{32}P-end-labeled 5hmCA repeat-containing DNA probe incubated with an increasing excess of competitive DNA (from 100- to 800-fold excess).

the DNA geometry are different. The MBD domain of MeCP2 recognizes and interacts with the hydroxymethylated CA repeat through two key amino acids, Arg[111] and Arg[133] (Fig. 7, B and C). Arg[111] is on the β-sheet 1 element and forms a hydrogen bond with the N7 and O6 positions of the guanine base that is located on the opposite strand of the DNA and base pairs with the hydroxymethyl-cytosine; in this network, Asp[121] helps position the Arg[111] side chain via two hydrogen bonds (Fig. 7C). Arg[111] thus serves as a primary DNA recognition site (Fig. 7C). In turn, Arg[133] recognizes the modified C and the T on the complementary strand of the CA repeat. Arg[133] is oriented in a perpendicular manner to the DNA major groove of the CA repeat and forms a hydrogen bond with the hydroxyl group of the 5hmC moiety and the oxygen atoms of the neighboring T base (Fig. 7, B and C). This reveals that Arg[133] provides the specificity for the recognition of the hydroxymethyl group via a direct interaction through a single hydrogen bond (Fig. 7C).

The structural data also inform on possible changes in MeCP2-DNA interaction in RTT. A cysteine residue at position 133 (as R133C mutation in RTT) would disrupt both the 5hmC and the neighboring T recognition, whereas a glycine residue (R111G) would induce loss of G recognition. In contrast, Arg[106], one of the most frequently mutated residues in RTT, is not directly involved in DNA interaction but instead stabilizes the MBD domain by forming hydrogen bonds with the carbonyl backbone of Met[94], Asp[156], Thr[158], and Val[159] (Fig. 7D). A tryptophan residue in this position (R106W in RTT) would disrupt the entire hydrogen bonding network. Mutation of Thr[158] to a methionine (T158M in RTT) would disrupt the observed hydrogen bond with the backbone of the 160–162 loop in which Arg[162] interacts with the DNA phosphate backbone in a sequence-independent manner (Fig. 7D). Reversely, mutation of Phe[155] to a serine (F155S in RTT) located in the direct vicinity would introduce a hydrophilic residue in the middle of a patch of hydrophobic residues around Phe[155], which would have a destabilizing effect (Fig. 7E). The impact of the different RTT mutations on the MeCP2-MBD structure is summarized in Fig. 7G.

Notably, the hydroxymethylation of the centrally located cytosine within the five CA repeats leads to a narrowing of the minor groove to 10.0 Å (i.e., reduced by 1.8 Å compared with classical B-form DNA) and to a widening of the major groove to 22.3 Å (i.e., increased by ~5.2 Å) (Fig. 7F). This change in DNA geometry in the CA repeat helps the MBD domain insert into its target DNA via its two arginine residues, Arg[111] and Arg[133]. In addition, there is a stark contrast in the roll base pair step parameter of the CA repeat DNA (Fig. 7F), which changes from a strong positive roll (6.7°) to a strong

negative roll (−8.2°) at the hydroxymethylcytosine step (Fig. 7F). This suggests that the specific shape of the CA repeat is important for the recognition of the CA repeat by MeCP2, whereas Arg[133] determines the specificity toward hydroxymethylated cytosines (Fig. 7, B and C).

On the basis of the structure, we hypothesized that a mutation of Arg[133] should affect both DNA binding and the function of MeCP2. We found that the RTT MeCP2 R133C mutant ectopically expressed in $Mecp2^{-/-}$ MEF cells had an altered distribution compared with WT (fig. S6, B and C). Moreover, the RTT MeCP2 R106W mutant also exhibited a different staining pattern (fig. S6, B and C). This suggests that the R106W mutation of MeCP2 interferes with DNA binding (13), possibly owing to alterations in the hydrogen bonding network of MeCP2 (see Fig. 7). Next, we carried out a comprehensive study on the genome-wide distribution pattern of the WT and two MeCP2 mutants using ChIP-seq. The data demonstrate that, in contrast to WT MeCP2, the binding to all CA repeat sites is severely impaired in the R133C mutant and in the case of the R106W mutant is entirely lost (Fig. 7H and quantification in Fig. 7I). This result was confirmed in vitro, by testing the affinity of the full-length R133C and R106W mutant proteins for hydroxymethylated CA repeat DNA by EMSA (Fig. 7J).

Discussion

The function of the essential brain protein MeCP2 has remained elusive for a long time. The present comprehensive study using genome-wide approaches and structural analysis reveals its molecular function both in vitro and in vivo. We show that MeCP2 selectively recognizes CA repeat DNA in a strand-specific manner and that it requires at least five consecutive CA dinucleotides to optimally bind DNA. The affinity of MeCP2 to modified CA repeats is one order of magnitude higher compared to previously identified substrates such as 5mCG and 5mCA dinucleotides (11, 17, 21). Accordingly, chimeric MeCP2 protein containing the MBD2 DNA binding domain that only recognizes mCG failed to rescue the RTT phenotype in mice, thus highlighting the importance of modified CA recognition by MeCP2 (38). Within the MBD family, MeCP2 is the only one that recognizes and specifically binds to CA repeats. Furthermore, while MeCP2 can bind in vitro modified and nonmodified CA repeats, with hydroxymethyl being the strongest, in vivo it shows impressive selectivity toward modified CA repeats, as shown by genome-wide ChIP-seq and DIP-seq analysis in MEF cells and pluripotent cells, and WGBS in neuronal cells and in the mouse brain, thus identifying the natural DNA target sequence. We also show that CA repeats are modified by DNMT3A.

The modified cytosine, only when located within a CA repeat, serves as a nucleation point for further MeCP2 accumulation and spreading around the repeat where it could compete for nucleosome binding. However, despite the impressive selectivity of MeCP2 for modified CA repeats, we cannot rule out a partial contribution of DNA methylation at CpG and CpA dinucleotides. The association of MeCP2 with modified CA repeats was missed by previous research studies using WGBS (16, 17, 39), which tend to strongly underestimate modified sites owing to the repetitive nature and asymmetric modification (specifically on one DNA strand) of CA repeats, requiring an unusually high sequencing coverage to be efficiently detected.

Our structural data provide the molecular details of the MeCP2-hydroxymethylated DNA interactions. The CA repeat creates a well-defined DNA shape, with a substantially modified geometry, including a widened major groove and negative roll parameters, located precisely at the modification site. This particular DNA geometry is important, but not sufficient, to ensure the specificity of MeCP2 binding. The CA repeat has to be hydroxymethylated to enable specific interactions with MeCP2. The hydroxyl moiety of the hydroxymethyl group forms a hydrogen bond with Arg[133] of MeCP2, that is, Arg[133] serves to probe the presence of the hydroxymethyl modification. This recognition is direct and does not involve a water molecule, which explains the high specificity of MeCP2 toward hydroxymethylated but not to methylated CA repeats and the deleterious effect of the R133C mutation in RTT patients. The interaction of Arg[133] with the hmCA is critical because it provides the specificity toward the chemically modified site. This key interaction does not exclude that other mutants contribute toward Rett pathogenesis in a somehow modulated manner, consistent with the existence of various phenotypes. Of note, the fact that the R133C mutation in MeCP2 results in a mild RTT phenotype suggests that this mutant may in part still bind to the methylated CA repeats in the brain (as observed in stem cells) (Fig. 2) but probably to a lesser extent and with specificity loss as compared to hydroxymethylated CA repeats. Our structure also offers the molecular basis for understanding the altered interactions of other MeCP2-MBD mutants with DNA in RTT. The R111G mutation affects the direct interaction of Arg[111] with DNA, while R106W, F155S, and T158M would act indirectly through changes of the hydrogen bonding network inside the MBD protein fold.

The identification of MeCP2 as a CA microsatellite binding protein establishes a previously unknown function for MeCP2. MeCP2 appears to serve as a long-range chromatin architecture organizer across the genome to

protect CA repeats from nucleosome invasion, consistent with the idea that MeCP2-mediated CA repeat protection involves global chromatin reorganization [reviewed in (*34, 35*)]. Loss of MeCP2 results in a widespread increase in nucleosome density inside LADs, which play a crucial role in three-dimensional genome architecture organization and transcription regulation (*40*). The MeCP2 depletion-mediated LAD chromatin structural perturbations affect LAD function and result in alterations in transcription at a genome-wide level, in agreement with our and other available data (*40–42*). Furthermore, numerous other studies have shown that CA repeats are involved in transcription and splicing regulation (*43–46*), and the presence of several thousands of simple repeats (including CA repeats) has been shown to affect the gene expression of nearby genes (*28, 29*). How CA repeat distribution and orientation affect gene expression remains to be determined.

Taken together, our data provide insights into the origin of RTT at the molecular level and suggest that RTT could be viewed as a chromatin disease, originating from the inability of mutant MeCP2 to bind and protect CA repeat–containing genes from nucleosome invasion. Our results may open a new area of research that considers the roles of specific proteins binding to microsatellites and other repeats in neurological diseases of unknown etiology.

Materials and methods
Pull-down assay

One milliliter of MEF nuclear extracts, corresponding to ~800 million cells, were incubated with 5 μg of biotinylated double-stranded DNA with or without a (CA)7 repeat motif in TGEN buffer (20 mM Tris pH 7.65, 3 mM MgCl₂, 0.1 mM EDTA, 10% glycerol, 0.01% NP40) containing 1.5 M NaCl and dialyzed overnight at 4°C in TGEN buffer containing 300 mM NaCl. Fifty microliters of streptavidin-coupled magnetic beads (Dynabeads M-280 Streptavidin, Invitrogen) pre-blocked with bovine serum albumin (BSA; 1 μg/μl) and yeast tRNA (1 μg/μl) were added to the dialyzed mixture for 2 hours at 4°C. DNA-protein complexes were washed five times with TGEN buffer containing 300 mM NaCl and then eluted with 100 μl of SDS loading buffer. Ten microliters of each eluate fraction was loaded on 10% Tris-Glycine gel and analyzed by silver staining and mass spectrometry.

Cell lines

Primary MEFs were isolated from Mecp2^tm1Bird mice (JAX stock #007177) (*6*). These mice possess two functional *loxP* sites flanking exons 3 and 4 of *Mecp2* (*Mecp2^fl/fl*). MEFs were derived from embryonic day 13.5 (E13.5) homozygous female embryo, as previously described (*47*). *Mecp2^fl/f* MEFs were transfected with pCMV-CRE vector using FuGENE 6 Transfection Reagent (Promega) according to the manufacturer's instructions to generate KO cells (*Mecp2^−/−*). After selection, targeted clones were identified by real-time quantitative PCR (RT-qPCR) and further confirmed by Western blot analysis.

For rescue experiments, MeCP2 WT, MeCP2 R133C, or MeCP2 R106W mutant proteins fused to N-terminal Flag- and hemagglutinin (HA)–epitope tags were stably expressed in *Mecp2^−/−* MEFs by retroviral transduction (*48*).

ChIP-seq
MeCP2 ChIP

MeCP2 ChIP experiments were performed as previously described (*48*) in MEFs WT or KO for *Mecp2*. Briefly, 50 μg of sonicated chromatin isolated from subconfluent MEFs WT or KO for *Mecp2* were immunoprecipitated using 2 μg of anti-MeCP2 antibody (Proteintech, 10861-1-AP). Six independent chromatin immunoprecipitations were pooled before sequencing analysis.

Flag-HA ChIP

Subconfluent MEFs expressing FLAG-HA–tagged MeCP2 were cross-linked with 0.4% formaldehyde for 10 min and quenched with 150 mM glycine for 10 min at room temperature. Cells were lysed in hypotonic buffer (10 mM Tris-HCl pH 7.65, 1.5 mM MgCl₂, 10 mM KCl) and disrupted with a Dounce homogenizer. The cytosolic fraction was separated from the nuclei fraction by centrifugation at 4°C. Nuclei were resuspended in sucrose buffer (20 mM Tris-HCl at pH 7.65, 15 mM KCl, 60 mM NaCl, 0.15 mM spermine, 0.5 mM spermidine) adjusted with high-salt buffer (20 mM Tris-HCl pH 7.65, 25% glycerol, 1.5 mM MgCl₂, 0.2 mM EDTA, 900 mM NaCl) to reach a final NaCl concentration of 300 mM. The nuclear-soluble fraction was recovered by centrifugation at 4°C. The pellet containing chromatin fraction was incubated in sucrose buffer containing 1 mM CaCl₂ and MNase (2.5 U/g of cells) for 10 min at 37°C. Digestion was stopped with 4 mM EDTA and soluble chromatin was recovered by centrifugation at 4°C. Tagged MeCP2 was immunoprecipitated with anti-Flag M2 antibody conjugated to agarose beads (Sigma), eluted with Flag peptide (0.5 mg/ml), further affinity-purified with anti-HA antibody conjugated to agarose beads, and eluted with HA peptide (1 mg/ml). The HA and Flag peptides were first buffered with 50 mM Tris-Cl (pH 8.5) then diluted to 4 mg/ml in TGEN 150 buffer (20 mM Tris at pH 7.65, 150 mM NaCl, 3 mM MgCl₂, 0.1 mM EDTA, 10% glycerol, 0.01% NP40) and stored at −20°C until use. Between each step, beads were washed in TGEN 150 buffer. MeCP2-bound chromatin was submitted to RNase and proteinase K digestion, and DNA was purified by phenol-chloroform.

Libraries construction and sequencing

Libraries were prepared using the Diagenode MicroPlex Library Preparation kit v2 and sequenced on Illumina HiSeq 4000 sequencer as single-end 50-bp reads following Illumina's instructions. Image analysis and base calling were performed using RTA 2.7.3 and bcl2fastq 2.17.1.14. Adapter dimer reads were removed using DimerRemover v0.9.2. Reads were mapped to the mouse genome (mm9) using Bowtie (*49*) v1.0.0 with the following arguments: -m 1 –strata –best -y -S –l 40 -p 2.

MeDIP-seq and hMeDIP-seq
DNA immunoprecipitation

Twenty micrograms of DNA was used as input, then 2 μl 5mC monoclonal antibody (Active Motif, 39649), 2 μl monoclonal immunoglobulin G (IgG) control antibody (Abcam, ab81032), 4 μl 5hmC polyclonal antibody (Active Motif, 39791) or 4 μl IgG polyclonal antibody (Abcam) were used to immunoprecipitate DNA. DNA and antibodies were incubated at 4°C overnight in a final volume of 1 ml DIP buffer (10 mM sodium phosphate pH 7.0, 140 mM NaCl, 0.05% Triton X-100). The bound material was recovered after incubation with 60 μl blocked protein G Dynabeads (beads washed three times with 1 ml DIP buffer and incubated for 4 hours minimum with BSA 1 mg/ml and yeast tRNA 0.5 mg/ml). The beads were washed three times with 1 ml DIP buffer, then treated overnight with RNase A at 65°C in the presence of 300 mM NaCl and then treated for 4 hours with proteinase K at 55°C. Immunoprecipitated DNA was purified by phenol-chloroform extraction followed by ethanol precipitation. Two independent DNA immunoprecipitations were pooled for each condition before sequencing analysis. For each condition, 5mC, 5hmC, and IgG control DIP assays were performed using the same batches of genomic DNA.

Libraries construction and sequencing

Libraries were prepared using the SMART cDNA Library Construction Kit and sequenced on Illumina HiSeq 4000 sequencer as single-end 50-bp reads following Illumina's instructions. Image analysis and base calling were performed using RTA 2.7.3 and bcl2fastq 2.17.1.14. Adapter dimer reads were removed using DimerRemover. Data were preprocessed with Cutadapt (*50*) v1.13 to remove the first nine nucleotides and to remove sequences with a trailing polyT of at least 10 Ts. Cutadapt was used with the following parameters: "-u 9 -a T [10] –discard-trimmed". Reads were mapped to the mouse genome (mm9) using Bowtie v1.0.0 with default parameters except for "-p 3 -m 1 – strata –best –chunkmbs 128."

RNA-seq

Total RNAs were purified from subconfluent MEFs using standard methods. Libraries of

Table 1. MeCP2 ChIP-seq datasets.

Mark	Cell or tissue	Type	GEO accession	Reference
MeCP2	Forebrain	ChIP-seq	GSE60071	(15)
MeCP2	Cerebellum	ChIP-seq	GSE60071	(15)
MeCP2	Whole brain	ChIP-seq	GSE38324	(31)
MeCP2	6-week-old cortex	ChIP-seq	GSE67293	(15)
MeCP2	Cerebellum	ChIP-seq	GSE38324	(31)
MeCP2	Cortical neurons	ChIP-seq	GSE31951	(30)

template molecules suitable for strand-specific high-throughput DNA sequencing were created using a TruSeq Stranded Total RNA with Ribo-Zero Gold Prep Kit (RS-122-2301; Illumina), as previously described (47). The libraries were sequenced on Illumina HiSeq 4000 sequencer as single-end 50-bp reads following Illumina's instructions. Image analysis and base calling were performed using RTA 2.7.3 and bcl2fastq 2.17.1.14. Adapter dimer reads were removed using DimerRemover v0.9.2. Reads were mapped to the mouse genome (mm9) using Tophat v2.0.14 and Bowtie v2-2.1.0. Quantification of gene expression was performed using HTSeq v0.6.1 and gene annotations from Ensembl release 67. Read counts have been normalized across WT and KO libraries with the method proposed in (51). Comparisons of interest were performed using the test for differential expression proposed in (52) and implemented in the Bioconductor package DESeq2 (v1.16.1). Resulting P values were adjusted for multiple testing using the Benjamini and Hochberg method (53).

MNase-seq
Mononucleosomes isolation
Subconfluent MEFs were lysed in hypotonic buffer (10 mM Tris-HCl at pH 7.65, 1.5 mM MgCl$_2$, 10 mM KCl) and disrupted by Dounce homogenizer. The cytosolic fraction was separated from the nuclei by centrifugation at 4°C. Nuclei were resuspended in sucrose buffer (20 mM Tris-HCl at pH 7.65, 15 mM KCl, 60 mM NaCl, 0.15 mM spermine, 0.5 mM spermidine) adjusted with high-salt buffer (20 mM Tris-HCl at pH 7.65, 25% glycerol, 1.5 mM MgCl$_2$, 0.2 mM EDTA, 900 mM NaCl) to get a final NaCl concentration of 300 mM. The nuclear-soluble fraction was recovered by centrifugation at 4°C. The pellet containing chromatin fraction was incubated in sucrose buffer containing 1 mM CaCl$_2$ and MNase (2.5 U/g of cells) for 10 min at 37°C. Digestion was stopped with 4 mM EDTA and mononucleosomes were recovered by centrifugation at 4°C. Mononucleosomes were submitted to RNase and proteinase K digestion, and DNA was purified by phenol-chloroform extraction.

Library construction and sequencing
Libraries were prepared using the Diagenode MicroPlex Library Preparation kit v2 and sequenced on Illumina HiSeq 4000 sequencer as paired-end 100-bp reads following Illumina's instructions. Image analysis and base calling were performed using RTA 2.7.3 and bcl2fastq 2.17.1.14. Adapter dimer reads were removed using DimerRemover v0.9.2 (https://source-forge.net/projects/dimerremover/). Reads were mapped to the mouse genome (mm9) using Bowtie (49) v1.0.0 with the following arguments: -m 1 –strata –best -y -S –l 40 -p 2. To generate the nucleosome position data track using MNase-seq data, we extracted read pairs which fragment sizes ranged between 100 and 200 bases. Then, we extracted the genomic coordinates of the midpoint (nucleosome dyads) between the highest and lowest coordinates of the two read pairs.

Whole-genome bisulfite sequencing datasets
The following previously published independent datasets were used. Raw reads from adult mouse frontal cortex (6-week-old) were downloaded from the Gene Expression Omnibus (GEO) under the accession number GSM1173783 (BS-seq) (32). Reads from granule cells were downloaded under the accession numbers GSM2517045 (BS-seq) and GSM2517048 (oxBS-seq) (16). Reads from 8-week-old cortex were downloaded under the accession numbers GSM448427 (DNMT3A$^{KO/+}$ mouse) and GSM4448426 (WT mouse) (33).

Data was processed using the Bismark v0.22.3 (54) pipeline. Reads were aligned to the mouse reference genome (mm9) using bismark with default parameters for Lister et al.'s data (32) and with the default parameters and nondirectional configuration for Mellén et al.'s data (16). Duplicated reads were removed using the tool deduplicate_bismark with default parameters and methylated sites were called using the tool bismark_methylation_extractor.

Heatmaps of read depth coverage were generated using Deeptools v3.5 (55) using the tools bamCoverage to generate bigwigs files with a step of 10 nucleotides, computeMatrix to generate a count matrix at the positions of

CA repeats and finally plotHeatmap and plotProfile to generate heatmaps and mean profile plots. The x axis represents the center of CA repeats ±2 kb. Pie charts and boxplots are drawn using R scripts done with R v4.0.2 and ggplot2 v3.3.2 (56). To calculate the percentage of the modification level of 5(h)mCA, we restricted analysis to CA sites of high coverage (≥10 reads).

MeCP2 ChIP-seq datasets
The published datasets shown in Table 1 were downloaded from GEO in SRA format and converted to FASTQ format using the fastq-dump program in the sratoolkit (version 2.1.9). Reads were mapped to the mouse genome (mm9) using Bowtie v1.0.0 with the following setting: -p 3 -m 1 –strata –best –chunkmbs 128 -S.

Analysis of MeCP2-regulated genes in brain tissues of RTT mouse models
Genes lists whose expression was shown to be up-regulated in hypothalamus (57), cerebellum (58), amygdala (59), hippocampus (60), stratium (61), and visual cortex (15) of MeCP2 KO mouse were downloaded from Gabel et al. (15). The 404 genes shown to be significantly misregulated across 38 different transcriptomic studies performed in a wide range of neural tissue and cells of RTT mouse models was identified by (62).

Computational analyses
Heatmaps and quantifications of the ChIP-seq, DIP-seq, and MNase-seq data were performed running seqMINER (63), using datasets normalized to 10 million uniquely mapped reads (for ChIP-seq and DIP-seq) or 50 million nucleosome dyads coordinates (for MNase-seq). For DIP-seq and ChIP-seq, peak detection was performed running MACS (64) using datasets normalized to 10 million uniquely mapped reads under settings in which the corresponding IgG (MeDIP-seq) or the KO cells (MeCP2 ChIP-seq) datasets were used as negative control. Detected peak summits were annotated using HOMER v4.7.2. (65). Homer provides repeats annotation obtained from the rmsk table of UCSC. UCSC repeats data are generated using the xRepeatMasker program (www.repeatmasker.org) with the -s and -m flags. RepeatMasker uses the comprehensive database of repetitive element consensus sequences Repbase (66), which allows imperfections for short tandem repeats detection. As reference coordinates, we used the RepeatMasker database (for DNA repeats) or the Ensembl 67 database (for coding genes) of the mouse genome (mm9). CpA/CpG heatmaps were generated using custom perl scripts wrapping BEDtools v2.26.0 utilities (67) and developed with custom functions. The scripts generated matrices of dinucleotide counts

per bin in regions of interest. TreeView v3.0 (68) was then used to plot heatmaps from generated matrices.

MBD proteins purification for electromobility shift assay

MeCP2 WT, MeCP2 R133C, or MeCP2 R106W mutants fused to N-terminal His-tag were cloned into pFastBac vector (Invitrogen). The cloned vectors were transformed into bacterial DH10Bac competent cells for making recombinant bacmid. The recombinant bacmids was then extracted and transfected into Sf9 cells by Cellfectin II Reagent (Invitrogen). After viral amplification, SF9 cells were infected (10^6 cells/ml) with baculoviruses expressing His-MeCP2 for 2 days at 27°C. Cells were harvested and resuspended in 25 ml lysis buffer containing 20 mM Tris-HCl pH 7.65, 400 mM NaCl, 10% glycerol, 0.01% NP40, 30 mM Imidazole, 0.2 mM PMSF and protease inhibitor cocktail tablets (Roche). The lysate was dounced 30 times, sonicated, and centrifuged for 10 min at 12,000 rpm. The clarified supernatant was applied to cOmplete His-Tag Purification Resin (Roche), washed with 50 mM imidazole, and eluted with 300 mM imidazole using buffer containing 20 mM Tris-HCl pH 7.65, 400 mM NaCl, 10% glycerol, 0.01% NP40. The eluate was diluted twice with sodium phosphate buffer [50 mM sodium phosphate pH 7, 1 mM dithiothreitol (DTT), 1 mM EDTA], incubated with SP sepharose fast flow bead (GE Healthcare), extensively washed with sodium phosphate buffer containing 100 mM NaCl, and eluted with sodium phosphate buffer containing 500 mM NaCl. The eluate was desalted with PD-10 Sephadex G-25 columns (GE Healthcare) previously equilibrated with TGEN 150 buffer (20 mM Tris-HCl pH 7.65, 10% glycerol, 150 mM NaCl, 3 mM MgCl₂, 0.1 mM EDTA, 0.01% NP40).

Preparation of MBD domains

The MBD domains of MBD1 (1–75), MBD2 (146–214), MBD3 (1–69), MBD4 (69–136), and MeCP2 (77–165) were cloned in a pET15b vector and expressed in *Escherichia coli* BL21-CodonPlus-RIL-pLysS (Stratagene). A 1-liter culture was grown in LB medium at 37°C until OD_{600} of 0.5 was reached before induction with 1 mM IPTG for 3 hours at 37°C. Cells were lysed in 15-ml buffer containing 20 mM Tris-HCl pH 7.65, 400 mM NaCl, 10% glycerol, 0.01% NP40, 30 mM imidazole, 0.2 mM PMSF and protease inhibitor cocktail tablets (Roche) in the presence of lysozyme (1 mg/ml) and sonicated on ice. The clarified supernatant was applied to cOmplete His-Tag Purification Resin (Roche), washed with 50 mM imidazole, and eluted with 300 mM imidazole using buffer containing 20 mM Tris-HCl pH 7.65, 400 mM NaCl, 10% glycerol, 0.01% NP40. The eluate was desalted with PD-10 Sephadex G-25 columns (GE Healthcare) previously

equilibrated with TGEN 150 buffer (20 mM Tris-HCl pH 7.65, 10% glycerol, 150 mM NaCl, 3 mM MgCl₂, 0.1 mM EDTA, 0.01% NP40).

Electromobility shift assay

The DNA substrates were prepared by annealing equimolar amounts of the corresponding synthetic oligonucleotides in a buffer containing 10 mM Tris-HCl pH 7.5, 1 mM EDTA, and 100 mM NaCl. DNA substrates were 5′-end labeled with [γ-³²P]ATP and T4 polynucleotide kinase. Reaction mixtures (10 μl) containing 20 mM Tris-HCl pH 7.6, 125 mM NaCl, 5% glycerol, 0.1 μg/μl BSA, and 2 nM of end-labeled substrates were incubated for 8 min at 4°C with the indicated concentration of tested protein, then an excess (as indicated) of unlabeled competitor DNA (sheared salmon sperm DNA) was added to reaction mixtures for 8 min. The binding reactions were loaded on a 6% nondenaturing polyacrylamide gel (pre-run for 20 min at 60 V in 0.5X TBE, run at 60V for 2 hours at 4°C).

Microscale thermophoresis

Binding of MeCP2-MBD with 5hmCAn or 5mCG DNA substrates (presented in Fig. 6A) were determined by microscale thermophoresis (MST). The purified MeCP2-MBD were coupled to a fluorophore (RED-NHS 2nd generation) in the HEPES buffer (10 mM HEPES, pH 7.6, 150 mM NaCl) with the protein Labeling Kit (Nanotemper, Munich, Germany) according to the manufacturer's instructions. Serial dilutions of DNA were made by 16 successive 1:1 dilutions of the highest concentration (80 μM) sample into the HEPES buffer. Each of these solutions was mixed 1:1 with the labeled MBD protein to a final concentration of 2 μM in all samples. The samples were incubated in the dark at room temperature for a minimum of 1 hour. All 16 samples were loaded into premium capillary tubes before the final measurements were taken in a Monolith NT.115 instrument (Nanotemper, Munich, Germany). The instrument's light-emitting diode (LED) power and the MST laser power were set to 40%. Measurements were performed at room temperature by turning on the LED followed by fluorescence monitoring as a function of time using a 5 s waiting period before the MST laser was switched on. The MST laser remained on for 30 s followed by 5-s monitoring of the recovery period. The resulting fluorescence time traces were examined as previously described (69). Differences in the direction of the curves is due to change in the hydration shell (70). All the measurements were repeated three times with triplicate samples. All the statistics calculations, curve fitting, and K_d calculation were performed with Prism Software (GraphPad).

Preparation of MeCP2-MBD/DNA complex for crystallization

The MBD domain (77–177 amino acids) of MeCP2 was cloned in pET28 vector with C-terminal 6XHis-tag. The protein was expressed in *E. coli* BL21 DE3 pLYSs (Promega, US) and induced at 0.6 OD with IPTG (0.4 mM) for 3 hours at 37°C. The bacterial cells were harvested by centrifugation and dissolved in lysis buffer (20 mM Tris-HCl pH 7.5, 300 mM NaCl, 2 mM β-mercaptoethanol, 0.01% NP40, and protease inhibitor cocktail). The cells were flash frozen in liquid nitrogen and stored at −80°C for further processing. The recombinant protein was purified by passing the crude lysate on a Ni sepharose column (GE) via affinity chromatography. The protein was washed thoroughly with buffer A (20 mM Tris-HCl pH 7.5, 400 mM NaCl, 1 mM β-mercaptoethanol, 1 mM EDTA, and 5 mM imidazole) and was eluted with increasing concentration of buffer B (20 mM Tris-HCl pH 7.5, 400 mM NaCl, 1 mM β-mercaptoethanol, 1 mM EDTA, and 250 mM imidazole). Further, the protein was purified via size-exclusion chromatography using a Superdex S 200 (GE) column in buffer C (20 mM Tris-HCl pH 7.5, 400 mM NaCl, 1 mM β-mercaptoethanol, 1 mM EDTA) followed by another round of purification using ion-exchange chromatography with HP S Sepharose (GE) column by increasing the salt concentration using buffer D (20 mM Tris-HCl pH 7.5, 1 M NaCl, 1 mM β-mercaptoethanol, 1 mM EDTA). Finally, the protein was buffer-exchanged with buffer E (10 mM HEPES pH 7.6, 300 mM NaCl, 1 mM DTT, and 0.1 mM EDTA) and immediately concentrated using Amicon column (Millipore) to 30 mg/ml.

DNA oligonucleotides were synthesized (Keck Biotech Lab, Yale University) and purified by HPLC chromatography. The single-stranded DNA was dissolved in a DNA annealing buffer (HEPES pH 7.9, 50 mM KCl, and 1 mM EDTA). Both the DNA strands were mixed to obtain the final 1 mM concentration. DNA was annealed in a thermal block (BIORAD) by heating to 96°C for 5 min and gradually cooling down to 4°C by decreasing by 4°C for every 2 min. The annealing of the DNA was verified on 10% DNA PAGE, and the concentration was checked using a Nanodrop device (Thomas Scientific).

DNA and the MBD protein were mixed in a ratio of 1:1.5 and the final concentration of the complex (2 μg/μl) was adjusted with buffer F (10 mM HEPES pH 7.6, 100 mM NaCl, 1 mM DTT, and 0.1 mM EDTA). The complex was incubated overnight at 4°C before starting the crystallization screening. The complex formation was also verified on 7% DNA PAGE.

Crystallization and structure determination

MeCP2-MBD/DNA complex crystals were obtained by mixing the complex with the reservoir solution (50 mM HEPES 7.5, 300 mM NH₄Cl,

30% PEG 2000, 1% dioxane, and 1 mM CaCl$_2$) in a ratio of 1:0.5 at 20°C. The reservoir solution was overlayed with Paratone (Hampton Research) to delay the nucleation and get better diffracting crystals, which were obtained after 3 months. Crystals were cryo-cooled with the solution containing 50 mM HEPES 7.5, 150 mM NH$_4$Cl, 35% PEG 2000, 1% dioxane, and 1 mM CaCl$_2$ overnight before cryo-cooling into liquid nitrogen.

Data were collected at SLS, Villigen, Switzerland, on the PXII beamline equipped with a Pilatus detector. Data were processed with the autoPROC suite (*71*) and Staraniso software (*72*). The space group of the crystal was C2 (a = 64.38 Å, b = 48.99 Å, c = 66.19 Å, β = 90.53°), with one molecule in the asymmetric unit and a solvent content of 49.5%. The structure was determined by molecular replacement using PDB ID 3C2I (*36*) as initial model using the Phaser program (*73*) inside the Phenix software (*74*). A primary search was carried out by using the protein part alone, followed by a search for the DNA coordinates by fixing the initial solution. The structure was refined iteratively through model building with COOT (*75*), and the refinement was carried out using Phenix and Buster programs (*76*). DNA parameters were analyzed with 3DNA (*77*) and CURVES (*78*). All the structural figures were prepared using the software PyMol (*79*).

REFERENCES AND NOTES

1. A. Rett, [On a unusual brain atrophy syndrome in hyperammonemia in childhood]. *Wien. Med. Wochenschr.* **116**, 723–726 (1966). pmid: 5300597
2. B. Hagberg, J. Aicardi, K. Dias, O. Ramos, A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases. *Ann. Neurol.* **14**, 471–479 (1983). doi: 10.1002/ana.410140412; pmid: 6638958
3. R. E. Amir et al., Rett syndrome is caused by mutations in X-linked *MECP2*, encoding methyl-CpG-binding protein 2. *Nat. Genet.* **23**, 185–188 (1999). doi: 10.1038/13810; pmid: 10508514
4. R. E. Amir et al., Influence of mutation type and X chromosome inactivation on Rett syndrome phenotypes. *Ann. Neurol.* **47**, 670–679 (2000). doi: 10.1002/1531-8249(200005)47:5<670::AID-ANA20>3.0.CO;2-F; pmid: 10805343
5. C. Philippe et al., Spectrum and distribution of *MECP2* mutations in 424 Rett syndrome patients: A molecular update. *Eur. J. Med. Genet.* **49**, 9–18 (2006). doi: 10.1016/j.ejmg.2005.04.003; pmid: 16473305
6. R. J. Guy, B. Hendrich, M. Holmes, J. E. Martin, A. Bird, A mouse *Mecp2*-null mutation causes neurological symptoms that mimic Rett syndrome. *Nat. Genet.* **27**, 322–326 (2001). doi: 10.1038/85899; pmid: 11242117
7. R. Z. Chen, S. Akbarian, M. Tudor, R. Jaenisch, Deficiency of methyl-CpG binding protein-2 in CNS neurons results in a Rett-like phenotype in mice. *Nat. Genet.* **27**, 327–331 (2001). doi: 10.1038/85906; pmid: 11242118
8. M. C. Marchetto et al., A model for neural development and treatment of Rett syndrome using human induced pluripotent stem cells. *Cell* **143**, 527–539 (2010). doi: 10.1016/j.cell.2010.10.016; pmid: 21074045
9. M. J. Lyst, A. Bird, Rett syndrome: A complex disorder with simple roots. *Nat. Rev. Genet.* **16**, 261–275 (2015). doi: 10.1038/nrg3897; pmid: 25732612
10. M. J. Lyst et al., Rett syndrome mutations abolish the interaction of MeCP2 with the NCoR/SMRT co-repressor. *Nat. Neurosci.* **16**, 898–902 (2013). doi: 10.1038/nn.3434; pmid: 23770565
11. J. D. Lewis et al., Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell* **69**, 905–914 (1992). doi: 10.1016/0092-8674(92)90610-O; pmid: 1606614
12. R. R. Meehan, J. D. Lewis, A. P. Bird, Characterization of MeCP2, a vertebrate DNA binding protein with affinity for methylated DNA. *Nucleic Acids Res.* **20**, 5085–5092 (1992). doi: 10.1093/nar/20.19.5085; pmid: 1408825
13. E. Ballestar, T. M. Yusufzai, A. P. Wolffe, Effects of Rett syndrome mutations of the methyl-CpG binding domain of the transcriptional repressor MeCP2 on selectivity for association with methylated DNA. *Biochemistry* **39**, 7100–7106 (2000). doi: 10.1021/bi0001271; pmid: 10852707
14. T. M. Yusufzai, A. P. Wolffe, Functional consequences of Rett syndrome mutations on human MeCP2. *Nucleic Acids Res.* **28**, 4172–4179 (2000). doi: 10.1093/nar/28.21.4172; pmid: 11058114
15. H. W. Gabel et al., Disruption of DNA-methylation-dependent long gene repression in Rett syndrome. *Nature* **522**, 89–93 (2015). doi: 10.1038/nature14319; pmid: 25762136
16. M. Mellén, P. Ayata, N. Heintz, 5-hydroxymethylcytosine accumulation in postmitotic neurons results in functional demethylation of expressed genes. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E7812–E7821 (2017). doi: 10.1073/pnas.1708044114; pmid: 28847947
17. S. Lagger et al., MeCP2 recognizes cytosine methylated tri-nucleotide and di-nucleotide sequences to tune transcription in the mammalian brain. *PLOS Genet.* **13**, e1006793 (2017). doi: 10.1371/journal.pgen.1006793; pmid: 28498846
18. M. Mellén, P. Ayata, S. Dewell, S. Kriaucionis, N. Heintz, MeCP2 binds to 5hmC enriched within active genes and accessible chromatin in the nervous system. *Cell* **151**, 1417–1430 (2012). doi: 10.1016/j.cell.2012.11.022; pmid: 23260135
19. C. G. Spruijt et al., Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives. *Cell* **152**, 1146–1159 (2013). doi: 10.1016/j.cell.2013.02.004; pmid: 23434322
20. S. Khrapunov et al., Unusual characteristics of the DNA binding domain of epigenetic regulatory protein MeCP2 determine its binding specificity. *Biochemistry* **53**, 3379–3391 (2014). doi: 10.1021/bi500424z; pmid: 24828757
21. B. Kinde, H. W. Gabel, C. S. Gilbert, E. C. Griffith, M. E. Greenberg, Reading the unique DNA methylation landscape of the brain: Non-CpG methylation, hydroxymethylation, and MeCP2. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 6800–6806 (2015). doi: 10.1073/pnas.1411269112; pmid: 25739960
22. B. Kinde, D. Y. Wu, M. E. Greenberg, H. W. Gabel, DNA methylation in the gene body influences MeCP2-mediated gene repression. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 15114–15119 (2016). doi: 10.1073/pnas.1618737114; pmid: 27965390
23. H. Stroud et al., Early-life gene expression in neurons modulates lasting epigenetic states. *Cell* **171**, 1151–1164.e16 (2017). doi: 10.1016/j.cell.2017.09.047; pmid: 29056337
24. C. Papin et al., Combinatorial DNA methylation codes at repetitive elements. *Genome Res.* **27**, 934–946 (2017). doi: 10.1101/gr.213983.116; pmid: 28348165
25. T. Baubec, R. Ivánek, F. Lienert, D. Schübeler, Methylation-dependent and -independent genomic targeting principles of the MBD protein family. *Cell* **153**, 480–492 (2013). doi: 10.1016/j.cell.2013.03.011; pmid: 23582333
26. L. Shen et al., Genome-wide analysis reveals TET- and TDG-dependent 5-methylcytosine oxidation dynamics. *Cell* **153**, 692–706 (2013). doi: 10.1016/j.cell.2013.04.002; pmid: 23602152
27. V. H. Adams, S. J. McBryant, P. A. Wade, C. L. Woodcock, J. C. Hansen, Intrinsic disorder and autonomous domain function in the multifunctional nuclear protein, MeCP2. *J. Biol. Chem.* **282**, 15057–15064 (2007). doi: 10.1074/jbc.M700855200; pmid: 17371874
28. S. F. Fotsing et al., The impact of short tandem repeat variation on gene expression. *Nat. Genet.* **51**, 1652–1659 (2019). doi: 10.1038/s41588-019-0521-9; pmid: 31676866
29. M. Gymrek et al., Abundant contribution of short tandem repeats to gene expression variation in humans. *Nat. Genet.* **48**, 22–29 (2016). doi: 10.1038/ng.3461; pmid: 26642241
30. S. Cohen et al., Genome-wide activity-dependent MeCP2 phosphorylation regulates nervous system development and function. *Neuron* **72**, 72–85 (2011). doi: 10.1016/j.neuron.2011.08.022; pmid: 21982370
31. S. S. Maxwell, G. J. Pelka, P. P. Tam, A. El-Osta, Chromatin context and ncRNA highlight targets of MeCP2 in brain. *RNA Biol.* **10**, 1741–1757 (2013). doi: 10.4161/rna.26921; pmid: 24270455
32. R. Lister et al., Global epigenomic reconfiguration during mammalian brain development. *Science* **341**, 1237905 (2013). doi: 10.1126/science.1237905; pmid: 23828890
33. D. L. Christian et al., DNMT3A haploinsufficiency results in behavioral deficits and global epigenomic dysregulation shared across neurodevelopmental disorders. *Cell Rep.* **33**, 108416 (2020). doi: 10.1016/j.celrep.2020.108416; pmid: 33238114
34. K. Gulmez Karaca, D. V. C. Brito, A. M. M. Oliveira, MeCP2: A critical regulator of chromatin in neurodevelopment and adult brain function. *Int. J. Mol. Sci.* **20**, 4577 (2019). doi: 10.3390/ijms20184577; pmid: 31527487
35. D. R. Connolly, Z. Zhou, Genomic insights into MeCP2 function: A role for the maintenance of chromatin architecture. *Curr. Opin. Neurobiol.* **59**, 174–179 (2019). doi: 10.1016/j.conb.2019.07.002; pmid: 31430649
36. K. L. Ho et al., MeCP2 binding to DNA depends upon hydration at methyl-CpG. *Mol. Cell* **29**, 525–531 (2008). doi: 10.1016/j.molcel.2007.12.028; pmid: 18313390
37. R. I. Wakefield et al., The solution structure of the domain from MeCP2 that binds to methylated DNA. *J. Mol. Biol.* **291**, 1055–1065 (1999). doi: 10.1006/jmbi.1999.3023; pmid: 10518942
38. R. Tillotson et al., Neuronal non-CG methylation is an essential target for MeCP2 function. *Mol. Cell* **81**, 1260–1275.e12 (2021). doi: 10.1016/j.molcel.2021.01.011; pmid: 33561390
39. R. Lister et al., Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315–322 (2009). doi: 10.1038/nature08514; pmid: 19829295
40. B. van Steensel, A. S. Belmont, Lamina-associated domains: Links with chromosome architecture, heterochromatin, and gene repression. *Cell* **169**, 780–791 (2017). doi: 10.1016/j.cell.2017.04.022; pmid: 28525751
41. D. Peric-Hupkes et al., Molecular maps of the reorganization of genome-nuclear lamina interactions during differentiation. *Mol. Cell* **38**, 603–613 (2010). doi: 10.1016/j.molcel.2010.03.016; pmid: 20513434
42. C. Leemans et al., Promoter-intrinsic and local chromatin features determine gene repression in LADs. *Cell* **177**, 852–864.e14 (2019). doi: 10.1016/j.cell.2019.03.009; pmid: 30982597
43. L. H. Naylor, E. M. Clark, d(TG)$_n$·d(CA)$_n$ sequences upstream of the rat prolactin gene form Z-DNA and inhibit gene transcription. *Nucleic Acids Res.* **18**, 1595–1601 (1990). doi: 10.1093/nar/18.6.1595; pmid: 2158081
44. F. Gebhardt, K. S. Zänker, B. Brandt, Modulation of epidermal growth factor receptor gene transcription by a polymorphic dinucleotide repeat in intron 1. *J. Biol. Chem.* **274**, 13176–13180 (1999). doi: 10.1074/jbc.274.19.13176; pmid: 10224073
45. N. Gabellini, A polymorphic GT repeat from the human cardiac Na$^+$Ca^{2+} exchanger intron 2 activates splicing. *Eur. J. Biochem.* **268**, 1076–1083 (2001). doi: 10.1046/j.1432-1327.2001.01974.x; pmid: 11179974
46. J. Hui, G. Reither, A. Bindereif, Novel functional role of CA repeats and hnRNP L in RNA stability. *RNA* **9**, 931–936 (2003). doi: 10.1261/rna.5660803; pmid: 12869704
47. A. Ors et al., Histone H3.3 regulates mitotic progression in mouse embryonic fibroblasts. *Biochem. Cell Biol.* **95**, 491–499 (2017). doi: 10.1139/bcb-2016-0190; pmid: 28177753
48. A. Obri et al., ANP32E is a histone chaperone that removes H2A.Z from chromatin. *Nature* **505**, 648–653 (2014). doi: 10.1038/nature12922; pmid: 24463511
49. B. Langmead, C. Trapnell, M. Pop, S. L. Salzberg, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009). doi: 10.1186/gb-2009-10-3-r25; pmid: 19261174
50. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **17**, 10–12 (2011). doi: 10.14806/ej.17.1.200
51. S. Anders, W. Huber, Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010). doi: 10.1186/gb-2010-11-10-r106; pmid: 20979621
52. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014). doi: 10.1186/s13059-014-0550-8; pmid: 25516281
53. Y. Hochberg, Y. Benjamini, More powerful procedures for multiple significance testing. *Stat. Med.* **9**, 811–818 (1990). doi: 10.1002/sim.4780090710; pmid: 2218183
54. F. Krueger, S. R. Andrews, Bismark: A flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**, 1571–1572 (2011). doi: 10.1093/bioinformatics/btr167; pmid: 21493656

55. F. Ramírez et al., deepTools2: A next generation web server for deep-sequencing data analysis. Nucleic Acids Res. 44, W160–W165 (2016). doi: 10.1093/nar/gkw257; pmid: 27079975

56. H. Wickham, ggplot2: Elegant Graphics for Data Analysis (Springer-Verlag, 2016).

57. M. Chahrour et al., MeCP2, a key contributor to neurological disease, activates and represses transcription. Science 320, 1224–1229 (2008). doi: 10.1126/science.1153252; pmid: 18511691

58. S. Ben-Shachar, M. Chahrour, C. Thaller, C. A. Shaw, H. Y. Zoghbi, Mouse models of MeCP2 disorders share gene expression changes in the cerebellum and hypothalamus. Hum. Mol. Genet. 18, 2431–2442 (2009). doi: 10.1093/hmg/ddp181; pmid: 19369296

59. R. C. Samaco et al., Crh and Oprm1 mediate anxiety-related behavior and social approach in a mouse model of MECP2 duplication syndrome. Nat. Genet. 44, 206–211 (2012). doi: 10.1038/ng.1066; pmid: 22231481

60. S. A. Baker et al., An AT-hook domain in MeCP2 determines the clinical course of Rett syndrome and related disorders. Cell 152, 984–996 (2013). doi: 10.1016/j.cell.2013.01.038; pmid: 23452848

61. Y. T. Zhao, D. Goffin, B. S. Johnson, Z. Zhou, Loss of MeCP2 function is associated with distinct gene expression changes in the striatum. Neurobiol. Dis. 59, 257–266 (2013). doi: 10.1016/j.nbd.2013.08.001; pmid: 23948639

62. A. Sanfeliu, W. E. Kaufmann, M. Gill, P. Guasoni, D. Tropea, Transcriptomic studies in mouse models of Rett syndrome: A review. Neuroscience 413, 183–205 (2019). doi: 10.1016/j.neuroscience.2019.06.013; pmid: 31229631

63. T. Ye et al., seqMINER: An integrated ChIP-seq data interpretation platform. Nucleic Acids Res. 39, e35 (2011). doi: 10.1093/nar/gkq1287; pmid: 21177645

64. Y. Zhang et al., Model-based analysis of ChIP-Seq (MACS). Genome Biol. 9, R137 (2008). doi: 10.1186/gb-2008-9-9-r137; pmid: 18798982

65. S. Heinz et al., Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell 38, 576–589 (2010). doi: 10.1016/j.molcel.2010.05.004; pmid: 20513432

66. J. Jurka et al., Repbase Update, a database of eukaryotic repetitive elements. Cytogenet. Genome Res. 110, 462–467 (2005). doi: 10.1159/000084979; pmid: 16093699

67. A. R. Quinlan, I. M. Hall, BEDTools: A flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842 (2010). doi: 10.1093/bioinformatics/btq033; pmid: 20110278

68. A. J. Saldanha, Java Treeview—extensible visualization of microarray data. Bioinformatics 20, 3246–3248 (2004). doi: 10.1093/bioinformatics/bth349; pmid: 15180930

69. C. J. Wienken, P. Baaske, U. Rothbauer, D. Braun, S. Duhr, Protein-binding assays in biological liquids using microscale thermophoresis. Nat. Commun. 1, 100 (2010). doi: 10.1038/ncomms1093; pmid: 20981028

70. P. Linke et al., An automated microscale thermophoresis screening approach for fragment-based lead discovery. J. Biomol. Screen. 21, 414–421 (2016). doi: 10.1177/1087057115618347; pmid: 26637553

71. C. Vonrhein et al., Data processing and analysis with the autoPROC toolbox. Acta Crystallogr. D 67, 293–302 (2011). doi: 10.1107/S0907444911007773; pmid: 21460447

72. I. J. Tickle et al., STARANISO (Global Phasing Ltd., 2017); http://staraniso.globalphasing.org/cgi-bin/staraniso.cgi.

73. A. J. McCoy et al., Phaser crystallographic software. J. Appl. Crystallogr. 40, 658–674 (2007). doi: 10.1107/S0021889807021206; pmid: 19461840

74. P. D. Adams et al., PHENIX: Building new software for automated crystallographic structure determination. Acta Crystallogr. D 58, 1948–1954 (2002). doi: 10.1107/S0907444902016657; pmid: 12393927

75. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. Acta Crystallogr. D 66, 486–501 (2010). doi: 10.1107/S0907444910007493; pmid: 20383002

76. G. Bricogne et al., BUSTER version 2.10.3 (Global Phasing Ltd., 2017).

77. X. J. Lu, W. K. Olson, 3DNA: A software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. Nucleic Acids Res. 31, 5108–5121 (2003). doi: 10.1093/nar/gkg680; pmid: 12930962

78. C. Blanchet, M. Pasi, K. Zakrzewska, R. Lavery, CURVES+ web server for analyzing and visualizing the helical, backbone and groove parameters of nucleic acid structures. Nucleic Acids Res. 39 (suppl. 2), W68–W73 (2011). doi: 10.1093/nar/gkr316; pmid: 21558323

79. W. L. DeLano, The PyMOL Molecular Graphics System, Version 1.7.0.1 (Schrödinger, LLC, 2002).

SUPPLEMENTARY MATERIALS

science.sciencemag.org/content/372/6549/eabd5581/suppl/DC1
Figs. S1 to S6
Table S1
References
MDAR Reproducibility Checklist

MICROBIOLOGY

The lysosomal Rag-Ragulator complex licenses RIPK1– and caspase-8–mediated pyroptosis by *Yersinia*

Zengzhang Zheng†, Wanyan Deng†, Yang Bai†, Rui Miao†, Shenglin Mei, Zhibin Zhang, Youdong Pan, Yi Wang, Rui Min, Fan Deng, Zeyu Wu, Wu Li, Pengcheng Chen, Tianchi Ma, Xiwen Lou, Judy Lieberman*, Xing Liu*

INTRODUCTION: An inflammatory innate immune response is a first line of host defense against invading pathogens. Inflammation recruits immune cells to the infection site and activates protective adaptive immune responses. Invasive bacteria have evolved multiple ways to interfere with host innate immune signaling to facilitate infection. The *Yersinia* effector protein YopJ suppresses proinflammatory cytokine production by inhibiting transforming growth factor–β–activated kinase 1 (TAK1) and nuclear factor κB (NF-κB) activation. To counteract this virulence factor, host cells initiate receptor-interacting serine-threonine protein kinase 1 (RIPK1)–dependent caspase-8–directed gasdermin D (GSDMD) cleavage and inflammatory cell death (pyroptosis) when TAK1 is inhibited. However, how the RIPK1–caspase-8–GSDMD axis is instructed during *Yersinia* infection is unclear.

RATIONALE: The best-studied mechanism by which pathogens stimulate inflammatory cell death involves triggering cytosolic sensors, called inflammasomes, which activate inflammatory caspases (1/4/5/11) to process proinflammatory cytokines and cause pyroptosis. Inflammatory caspase cleavage of GSDMD causes cell membrane pores that mediate both pyroptosis and proinflammatory cytokine secretion. An alternate pyroptotic pathway, mediated by activation of RIPK1 and caspase-8, is triggered when the YopJ virulence factor secreted during pathogenic *Yersinia* infection blocks TAK1 activation. To determine the molecular mechanisms underlying *Yersinia* activation of RIPK1–caspase-8–dependent pyroptosis, we performed a genome-wide CRISPR screen using Cas9-expressing immortalized mouse bone marrow–derived macrophages infected with a genome-wide library of single-guide RNA–encoding lentiviruses. The genomes of cells resistant to caspase-8– or caspase-11–dependent pyroptosis were sequenced to identify the knocked-out genes required for pyroptosis.

RESULTS: The screen identified multiple genes in the lysosomal membrane–anchored Folliculin (Flcn)–Folliculin-interacting protein 2 (Fnip2)–Rag-Ragulator complex as necessary for caspase-8– but not caspase-11–mediated pyroptosis. Deficiency of Rag-Ragulator complex genes rendered cells highly resistant to TAK1 inhibition–triggered pyroptosis, indicating a critical and unexpected role of the lysosomal membrane–tethered Rag-Ragulator supercomplex in RIPK1-dependent caspase-8–directed pyroptosis. In response to pathogenic *Yersinia* or its mimic [lipo-polysaccharide (LPS) plus TAK1 inhibitor], a Fas-associated death domain (FADD)–RIPK1–caspase-8–containing complex was recruited to lysosomal membrane–tethered Rag-Ragulator. Activation of RIPK1 phosphorylation, caspase-8 activation, and pyroptosis depended on Rag guanosine triphosphatase (GTPase) activity and Rag-Ragulator lysosomal binding but was independent of the mechanistic target of rapamycin complex 1 (mTORC1), a well-known Rag-Ragulator–regulated complex. By contrast, Rag-Ragulator did not regulate canonical or noncanonical inflammasome-triggered pyroptosis.

CONCLUSION: Our study revealed an instructive role of metabolic signaling in directing TAK1 inhibition–induced pyroptosis during a pathogenic bacterial infection. Rag-Ragulator is a well-known critical regulator of cellular responses to changes in nutrient availability and metabolism. Here, Rag-Ragulator served as a platform for activating a FADD–RIPK1–caspase-8 complex formed in response to Toll-like receptor (TLR) or tumor necrosis factor receptor (TNFR) ligation. Rag GTPase activity was critical for triggering the pathway. The role found here for Rag-Ragulator in pyroptosis expands its known roles in metabolic regulation to include regulation of the response to pathogenic infection. Rag-Ragulator monitors both metabolism and infection to serve as a central hub for helping to decide whether available nutrients are adequate for cell proliferation and if an infected cell should die and send out inflammatory danger signals. Future studies can further explore the conditions that stimulate caspase-8–mediated pyroptosis and provide more mechanistic details of how it is regulated, as well as investigate whether manipulating this pathway could have therapeutic benefit. ∎
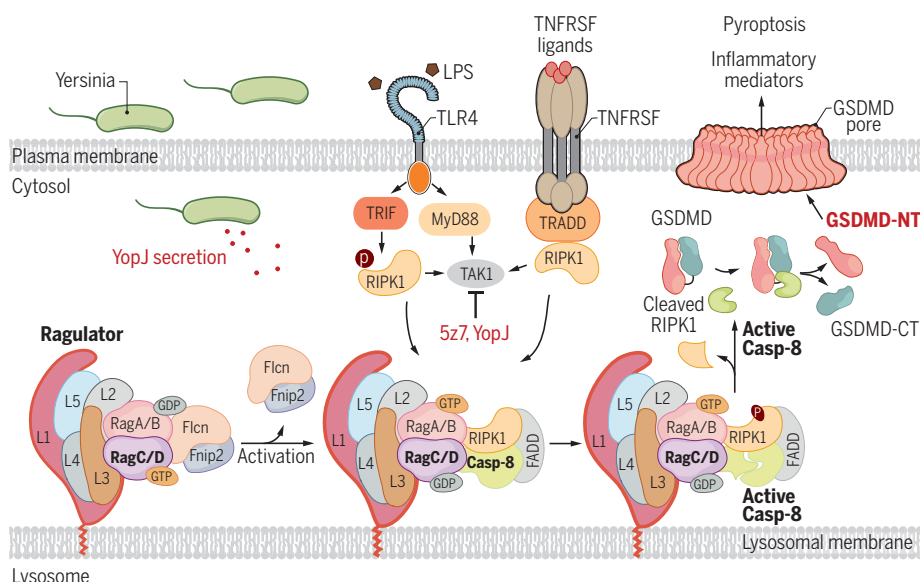
Rag-Ragulator serves as a platform for activating a FADD–RIPK1–caspase-8 complex to trigger pyroptosis in response to TLR and TNFR ligation. When TAK1 is inhibited by the *Yersinia* effector YopJ or its mimic 5z-7-oxozeaenol (5z), a FADD–RIPK1–caspase-8–containing complex is recruited to lysosome-tethered Rag-Ragulator, which activates caspase-8– and GSDMD-dependent pyroptosis. TNFRSF, TNFR superfamily; L1-5, Lamtor1-5.

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abg0269

MICROBIOLOGY

The lysosomal Rag-Ragulator complex licenses RIPK1– and caspase-8–mediated pyroptosis by *Yersinia*

Zengzhang Zheng[1,2]†, Wanyan Deng[1,2]†, Yang Bai[2,3]†, Rui Miao[4,5]†, Shenglin Mei[6], Zhibin Zhang[4,5], Youdong Pan[7], Yi Wang[8], Rui Min[2,3], Fan Deng[2,3], Zeyu Wu[2,3], Wu Li[1,2], Pengcheng Chen[2], Tianchi Ma[2], Xiwen Lou[2], Judy Lieberman[4,5]*, Xing Liu[1,2]*

Host cells initiate cell death programs to limit pathogen infection. Inhibition of transforming growth factor–β–activated kinase 1 (TAK1) by pathogenic *Yersinia* in macrophages triggers receptor-interacting serine-threonine protein kinase 1 (RIPK1)–dependent caspase-8 cleavage of gasdermin D (GSDMD) and inflammatory cell death (pyroptosis). A genome-wide CRISPR screen to uncover mediators of caspase-8–dependent pyroptosis identified an unexpected role of the lysosomal folliculin (FLCN)–folliculin-interacting protein 2 (FNIP2)–Rag-Ragulator supercomplex, which regulates metabolic signaling and the mechanistic target of rapamycin complex 1 (mTORC1). In response to *Yersinia* infection, Fas-associated death domain (FADD), RIPK1, and caspase-8 were recruited to Rag-Ragulator, causing RIPK1 phosphorylation and caspase-8 activation. Pyroptosis activation depended on Rag guanosine triphosphatase activity and lysosomal tethering of Rag-Ragulator but not mTORC1. Thus, the lysosomal metabolic regulator Rag-Ragulator instructs the inflammatory response to *Yersinia*.

When mucosal and immune sentinel cells sense invasive infection or other danger signals, they activate the assembly of large multiprotein complexes called inflammasomes that recruit inflammatory caspases (caspase-1, caspase-4, caspase-5, and caspase-11), which become auto-activated by proximity (*1–3*). The activated pro-inflammatory caspases then cleave gasdermin D (GSDMD) in the cytosol to liberate an N-terminal fragment that assembles into cell membrane pores to cause an inflammatory cell death called pyroptosis, in which the cell membrane is rapidly permeabilized to release inflammatory mediators (*4–9*). Caspase-1 also processes the pro-inflammatory interleukin-1 (IL-1) family cytokines, which lack a conventional secretion signal, and the processed inflammatory cytokines are then

released through GSDMD pores (*10*, *11*). Pyroptosis and IL-1 family cytokines recruit and activate immune cells to the site of infection to orchestrate host defense against invading pathogens. During pathogenic *Yersinia* infection, the *Yersinia* effector protein YopJ inhibits transforming growth factor–β–activated kinase 1 (TAK1) or IκB kinase (IKK) to trigger an alternate pyroptotic pathway mediated by Toll-like receptors (TLRs) or death receptors that instigate the formation of a complex involving the adapter Fas-associated death domain (FADD) and receptor-interacting serine-threonine protein kinase 1 (RIPK1) and recruitment and activation of caspase-8, which cleaves GSDMD (*12*, *13*). Blockade of TAK1 kinase activity with the specific inhibitor 5z-7-oxozeaenol (5z7) mimics the effect of YopJ and pathogenic *Yersinia* infection to cause RIPK1–caspase-8–dependent GSDMD-mediated pyroptosis in the presence of TLR ligands or tumor necrosis factor–α (TNF-α). However, not much is known about how this alternative caspase-8–dependent pyroptotic pathway is activated or regulated.

To determine the molecular mechanisms underlying activation of RIPK1–caspase-8–dependent pyroptosis during *Yersinia* infection, we performed an unbiased genome-wide CRISPR screen. We infected Cas9-expressing immortalized mouse bone marrow–derived macrophages (iBMDMs) with lentiviruses encoding a library of single-guide RNAs (sgRNAs) and treated them with extracellular lipopolysaccharide (LPS)/5z7 to trigger caspase-8–dependent pyroptosis or electroporated them with LPS to trigger caspase-11–mediated GSDMD

cleavage for comparison (Fig. 1A). Surviving cells were isolated and retreated twice more to enrich for pyroptosis-resistant cells. The genomes of resistant cells were sequenced to identify enriched sgRNAs that knocked out genes required for pyroptosis. The screen identified multiple genes in the lysosomal membrane–anchored folliculin (Flcn)–folliculin-interacting protein 2 (Fnip2)–Rag-Ragulator complex as necessary for caspase-8–mediated pyroptosis but not caspase-11–mediated pyroptosis.

Genome-wide CRISPR screen identifies the Rag-Ragulator complex as critical for pyroptosis induced by LPS/5z7 treatment

We identified several sgRNAs enriched in cells that survived three sequential exposures to stimuli that triggered caspase-8– or caspase-11–dependent pyroptosis compared with medium-treated control cells (Fig. 1, A to C, and fig. S1). Only a few sgRNA target genes were enriched in cells that survived both caspase-8– and caspase-11–mediated pyroptosis. Cells that survived LPS transfection were significantly enriched for sgRNAs designed to knock out *Gsdmd*, *Casp-11*, and *Irf2*, all known genes implicated in the non-canonical inflammasome pathway, including the gene encoding Irf2, which transcriptionally activates *Gsdmd* expression (*14*). Caspase-8–induced pyroptosis-resistant cells were highly enriched for sgRNAs targeting genes involved in tumor necrosis factor receptor (TNFR) signaling, mitochondrial electron transport chain (ETC) complex I, and the lysosome-associated Rag-Ragulator complex. The 27 most enriched sgRNAs, which were increased 164- to 2500-fold compared with mock-treated cells, included four sgRNAs targeting the gene encoding TNFR1; eight sgRNAs targeting four genes encoding subunits of ETC complex I (*Ndufa6*, *Ndufa13*, *Ndufaf4*, and *Ndufa9*), and 15 sgRNAs targeting genes encoding five components of the Rag-Ragulator supercomplex (Lamtor1, Lamtor4, RagC, Flcn, and Fnip2), which serves as a platform for recruiting Raptor-mTOR (the mechanistic target of rapamycin) to the lysosomal surface and regulating its activation (*15*). The recovery of TNFR1 sgRNAs in LPS/5z7-treated, but not LPS-transfected, surviving cells confirmed the selectivity of the screen. The identification of ETC complex I gene sgRNAs was not surprising because mitochondrial reactive oxygen species, of which ETC complex I is a major source, have been implicated in augmenting GSDM pore formation (*16*, *17*). By comparison, neither *Casp-8*– nor *Gsdmd*-targeting guides were enriched, likely because deficiency of caspase-8 or Gsdmd could lead to Ripk3-dependent or Gsdme-directed cell death (*18–20*). The Rag-Ragulator supercomplex is composed of nine subunits: Ragulator, a pentameric scaffold (Lamtor1-5) that tethers the complex to the lysosomal membrane by lipidation of Lamtor1 (*21*, *22*); two Rag GTPases (GTP-loaded RagA or

[1]The Joint Center for Infection and Immunity between Guangzhou Institute of Pediatrics, Guangzhou Women and Children's Medical Center, Guangzhou 510623, China, and Institut Pasteur of Shanghai, Chinese Academy of Sciences, Shanghai 200031, China. [2]The Center for Microbes, Development and Health, Key Laboratory of Molecular Virology and Immunology, Institut Pasteur of Shanghai, Chinese Academy of Sciences, Shanghai 200031, China. [3]University of Chinese Academy of Sciences, Shanghai 200031, China. [4]Program in Cellular and Molecular Medicine, Boston Children's Hospital, Boston, MA 02115, USA. [5]Department of Pediatrics, Harvard Medical School, Boston, MA 02115, USA. [6]Department of Biomedical Informatics, Harvard Medical School, Boston, MA 02115, USA. [7]Department of Dermatology and Harvard Skin Disease Research Center, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA. [8]Department of Medicine, Harvard Medical School, Boston, MA 02115, USA.
*Corresponding author. Email: judy.lieberman@childrens.harvard.edu (J.L.); xingliu@ips.ac.cn (X.L.)
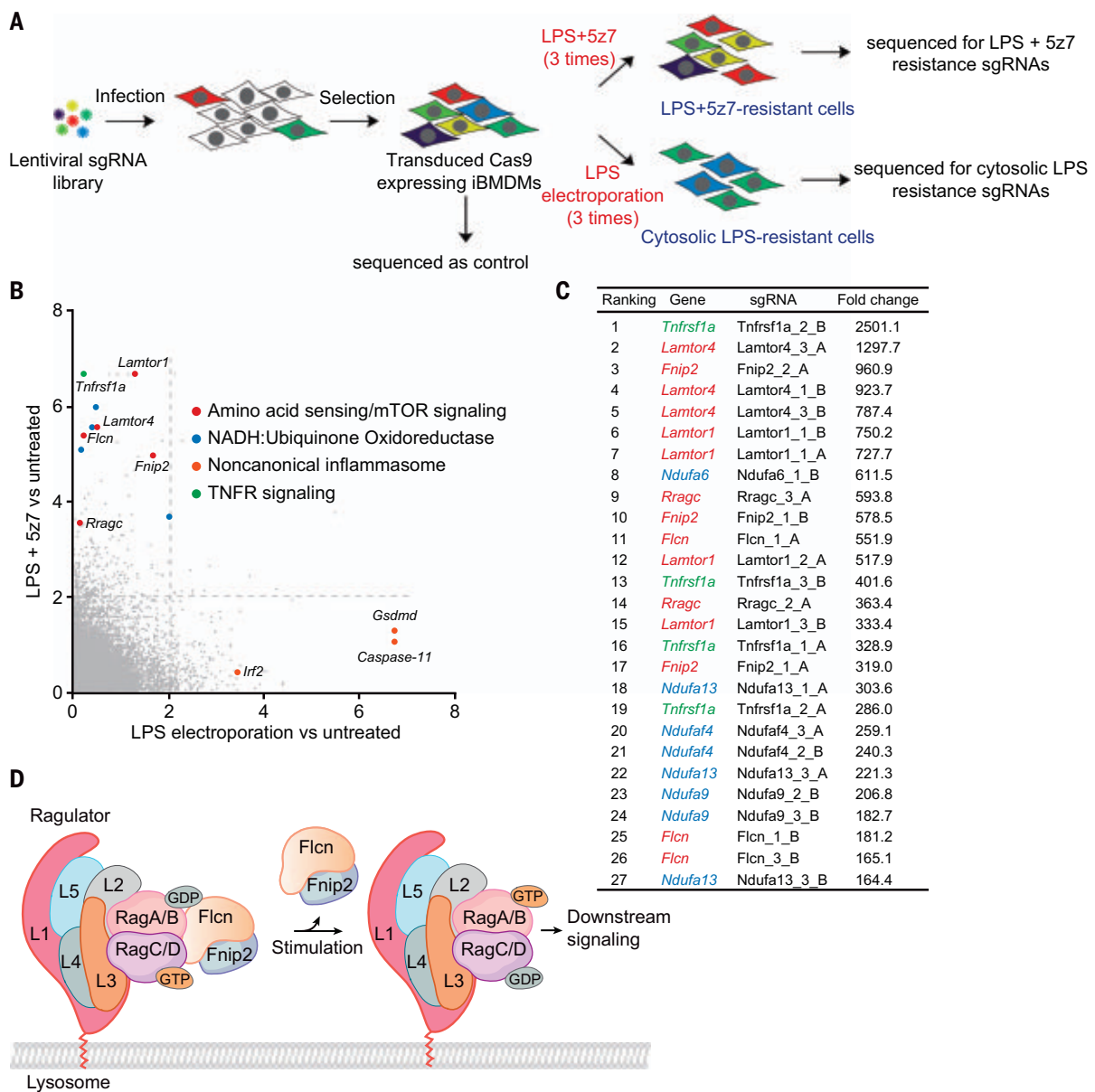†These authors contributed equally to this work.

Fig. 1. A genome-wide CRISPR screen identifies multiple components of the FLCN–FNIP2–Rag-Ragulator complex as needed for caspase-8–mediated pyroptosis. (**A**) Schematic workflow of the genome-wide CRISPR screening procedure in iBMDMs. (**B**) −lgRRA (robust rank aggregation) values calculated by the MAGeCK algorithm indicate the relative enrichment of genes after treatment with LPS electroporation (*x* axis) or LPS/5z7 (*y* axis) relative to controls in positive selection. The dotted lines represent *P* = 0.01. (**C**) List of sgRNA hits from the screening in iBMDMs treated with LPS/5z7. Corresponding genes targeted with multiple sgRNAs and fold enrichments are shown. Green indicates genes encoding for the TNFR, red is the FLCN–FNIP2–Rag-Ragulator complex, and blue is the mitochondrial electron transport complex I. (**D**) Diagram of the FLCN–FNIP2–Rag-Ragulator complex.

RagB and GDP-loaded RagC or RagD); and FLCN complexed with its binding partners, Fnip1 or Fnip2 (FLCN-interacting protein 1/2), which facilitates the activation of the Rag heterodimer by catalyzing GTP hydrolysis of RagC or RagD (Fig. 1D). Because the identification of the lysosomal Rag-Ragulator complex was unexpected and dominated the list of most enriched sgRNAs, subsequent experiments focused on examining the putative role of Rag-Ragulator in caspase-8–mediated pyroptosis.

Ragulator complex is required for caspase-8–mediated, but not inflammasome-activated, pyroptosis

To verify the contribution of the Rag-Ragulator top gene hits to LPS/5z7-induced pyroptosis, we used CRISPR-Cas9–mediated gene editing to knock out *Lamtor1*, *Lamtor4*, *Rragc*, *Flcn*, and *Fnip2* in iBMDMs. For each gene, two clones transduced with distinct gRNAs that target different regions of each gene were obtained. Gene editing was verified by DNA

sequencing and immunoblot (Fig. 2A and fig. S2). Concurrent treatment of wild-type (WT) iBMDMs with LPS and 5z7 resulted in rapid pyroptotic cell death, as assessed by lactate dehydrogenase (LDH) release, but iBMDMs deficient in each of the Rag-Ragulator genes tested were highly resistant to LPS/5z7-induced pyroptosis (Fig. 2A and fig. S3A). Consistent with the dependency of caspase-8–mediated pyroptosis on these genes, caspase-8 maturation and GSDMD cleavage to the p30 and p20
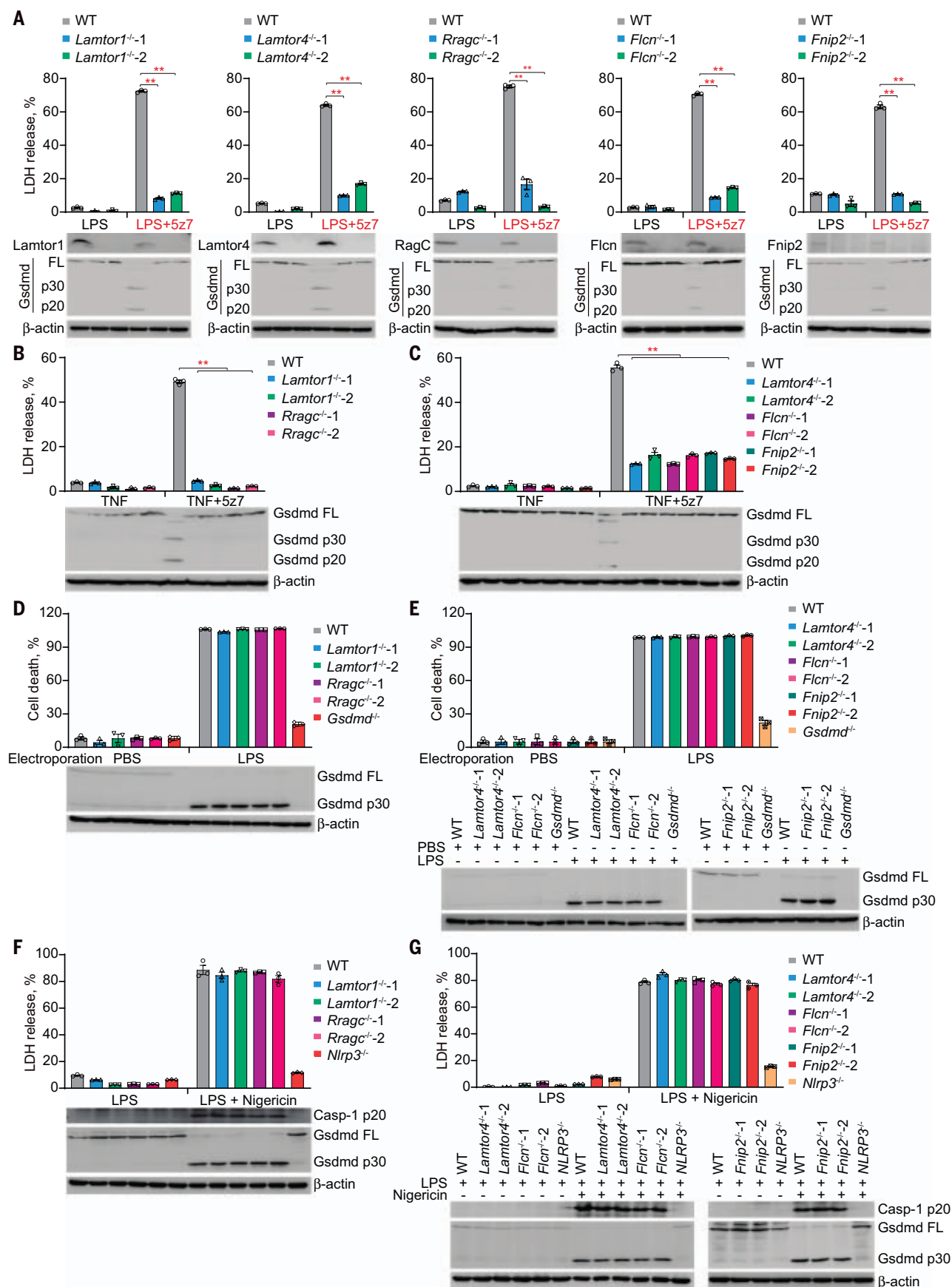
Fig. 2. FLCN–FNIP2–Rag-Regulator gene deficient cells are resistant to LPS/5z7 treatment but sensitive to inflammasome-induced pyroptosis. (**A**) The indicated iBMDMs were challenged with LPS or LPS/5z7. Cell death was measured by LDH release after 2.5 hours. Specific gene ablation was confirmed and Gsdmd cleavage was examined by immunoblot, as shown in the bottom panel. (**B**, **C**, **F**, and **G**) The indicated iBMDMs were treated with TNF-α (abbreviated to TNF in the figure), TNF-α/5z7 [(B) and (C)], or LPS/nigericin [(F) and (G)]. Cell death was measured by LDH release after 2.5 hours [(B) and (C)] or 1.5 hours [(F) and (G)]. (**D** and **E**) The indicated iBMDMs were electroporated with LPS. Cell death was assessed and calculated by measuring the ATP level after 2.5 hours. Gsdmd cleavage and caspase-1 processing (during canonical inflammasome activation) were examined by immunoblot. Graphs show mean ± SEM of triplicate wells. Data are representative of at least three independent experiments. Data were analyzed using a two-tailed Student's *t* test. **P < 0.01.

fragments, which were detected by immunoblot in LPS- and 5z7-treated WT iBMDMs, were not detected in the knockout cells (Fig. 2A and fig. S3B). iBMDMs knocked out for the same set of Rag-Regulator genes also survived treatment with an alternate stimulus of caspase-8–dependent pyroptosis (TNF-α plus 5z7) (*12*) and showed no sign of GSDMD cleavage (Fig. 2, B and C). Caspase-3/7 cleavage, occurring downstream of caspase-8 activation (*13*, *23*), was blunted by Rag-Regulator deficiency upon LPS/5z7 treatment (fig. S3B). By contrast, Rag-Regulator deficiency had no effect on caspase-3/7 cleavage or activation in etoposide-induced apoptosis (fig. S3C), indicating the specific role of Rag-Regulator in LPS/5z7-induced pyroptosis. To determine whether the Rag-Regulator complex was also involved in inflammasome and inflammatory caspase–mediated pyroptosis, WT and knockout clones of iBMDMs were treated with two stimuli of canonical inflammasomes and caspase-1–mediated pyroptosis [LPS plus nigericin to activate the NLRP3 inflammasome or anthrax-protective antigen (PA) plus lethal factor-flagellin fusion protein (LFN-Fla) to activate the NLRC4 inflammasome] or were electroporated with LPS to induce the noncanonical caspase-11 inflammasome and pyroptosis (Fig. 2, D to G, and fig. S4). WT iBMDMs and iBMDMs knocked out for the five Rag-Regulator genes were all similarly killed by activating the canonical or noncanonical inflammasomes. As positive controls and as expected, iBMDMs genetically deficient in *Gsdmd* were resistant to electroporated LPS, iBMDMs deficient in *Nlrp3* were resistant to LPS plus nigericin, and iBMDMs deficient in *Nlrc4* were resistant to PA plus LFN-Fla. In these experiments cell death was assessed either by an ATP assay measuring surviving cells or an LDH assay measuring pyroptosis. GSDMD cleavage, assessed by immunoblot, correlated with cell death, as expected. Thus, *Lamtor1*, *Lamtor4*, *Rragc*, *Flcn*, and *Fnip2* are selectively required for RIPK1–caspase-8–dependent pyroptosis but not for inflammasome-mediated pyroptosis.

The Rag-Regulator complex recruits RIPK1 and caspase-8 to the lysosomal membrane

Because the Rag-Regulator complex is localized to the lysosomal membrane and is required for caspase-8–mediated pyroptosis, we postulated that RIPK1 and caspase-8 might be recruited to and activated on the lysosomal membrane after LPS/5z7 treatment. In the absence of cFLIP$_L$,

TLR4 signaling typically causes the assembly of a complex containing a death domain–containing protein FADD and RIPK1 that recruits and activates caspase-8 (*24*). The complex also assembles in response to LPS/5z7 (*24*). Consistently, within 100 min of adding LPS/5z7, but not LPS on its own, FADD immunoprecipitated with RIPK1, procaspase-8, and the p43 subunit of activated caspase-8, indicating the formation of a FADD–RIPK1–caspase-8–containing complex that activated caspase-8 (Fig. 3A). Although cleaved caspase-8 levels increased by 150 min, the RIPK1 band was strongly diminished and the interaction of FADD with both RIPK1 and caspase-8 became undetectable, suggesting that the complex was transient and disintegrated, potentially because of RIPK1 degradation. To determine whether the complex depended on Rag-Regulator, the experiment was repeated in iBMDMs knocked out for *Rragc*. In RagC-deficient cells treated with LPS/5z7, FADD did not pull down RIPK1 or caspase-8, RIPK1 was not degraded, and the p43-activated caspase-8 band was barely detected (Fig. 3A). These data suggest that assembly of the FADD–RIPK1–caspase-8 complex depends on Rag-Regulator and might be recruited to lysosomes for subsequent function. Indeed, treatment with LPS/5z7 significantly and specifically induced colocalization of caspase-8 and RIPK1 with lysosomes, but not endoplasmic reticulum, mitochondria, or early endosomes, and was dependent on Lamtor1 (Fig. 3, B to D, and fig. S5). To confirm that RIPK1 and caspase-8 colocalize with Rag-Regulator after LPS/5z7 stimulation, hemagglutinin (HA)–tagged RIPK1 (Fig. 3E) and caspase-8 (Fig. 3F) were coexpressed with Flag-tagged Lamtor1, RagC, Flcn, or Fnip2 in human embryonic kidney (HEK) 293T cells, and coimmunoprecipitation was assessed by anti-HA immunoprecipitation followed by anti-Flag immunoblot. Coimmunoprecipitation of both RIPK1 and caspase-8 with RagC and Flcn was detected but not with Lamtor1 or Fnip2 (Fig. 3, E and F). Moreover, interaction of endogenous RIPK1 and caspase-8 with RagC could both be detected in iBMDMs treated with LPS/5z7 (Fig. 3G). Compared with GTP-loaded inactive RagC (Q119L), GDP-loaded active RagC (S74N) showed stronger binding to RIPK1/caspase-8 (Fig. 3, H and I). Mapping of interactive domains by truncation revealed that the kinase domain (KD) of RIPK1 and the caspase domain (CD) of caspase-8 were required and sufficient for RagC binding (fig. S6).

These domains are distinct from the domains implicated in RIPK1 binding to caspase-8 and FADD through a death domain (DD)–death effector domain (DED) interaction and DD-DD interaction, respectively (*25*–*27*), and in caspase-8 binding to FADD through a DED-DED interaction (*28*, *29*). Because caspase-8–mediated pyroptosis depended on Lamtor1 and Fnip2, we suspect that RIPK1 and caspase-8 associate with the large supercomplex, but their interaction with these other components could be indirect. This result suggested that RIPK1 and caspase-8 directly interacted with RagC and Flcn, which are in contact with each other in the cryo–electron microscopy structure of the human Rag-Regulator supercomplex (*30*). Thus, Rag-Regulator acts as a lysosomal platform for recruiting RIPK1 and procaspase-8 and activating caspase-8.

Neither mTOR activity nor Tfeb/Tfe3 is involved in caspase-8– or inflammasome-induced pyroptosis

Because the Rag-Regulator complex regulates mTORC1 activation, we next investigated whether caspase-8–mediated pyroptosis is regulated by mTORC1. To inhibit mTORC1, iBMDMs were treated with rapamycin, a potent allosteric mTORC1 inhibitor, or with Torin 1, a selective ATP-competitive mTORC1/2 inhibitor (*31*). Both inhibited mTORC1 kinase activity in iBMDMs at the doses used and caused dephosphorylation of the mTORC1 target p70 S6 kinase (fig. S7A). However, incubation of iBMDMs with rapamycin or Torin 1 before treatment with LPS/5z7 had no effect on GSDMD cleavage or on the extent or kinetics of pyroptotic cell death, as assessed by propidium iodide (PI) uptake (Fig. 4, A and B). Furthermore, treatment with 5-aminoimidizole-4-carboxamide riboside (AICAR), which blocks mTOR signaling by activating AMP-activated protein kinase (AMPK), did not affect LPS/5z7-triggered GSDMD cleavage or pyroptosis (Fig. 4C and fig. S7B). Thus, mTOR kinase activity does not affect caspase-8–induced pyroptosis, although the Rag-Regulator complex is required. These mTORC1 inhibitors also did not affect electroporated LPS–induced noncanonical inflammasome or LPS plus nigericin–induced canonical inflammasome activation of GSDMD cleavage or pyroptosis (Fig. 4, D and E). Thus, mTOR activity is not involved in pyroptosis triggered by TAK1 inhibition or inflammasome signaling. In addition, simultaneous knockdown
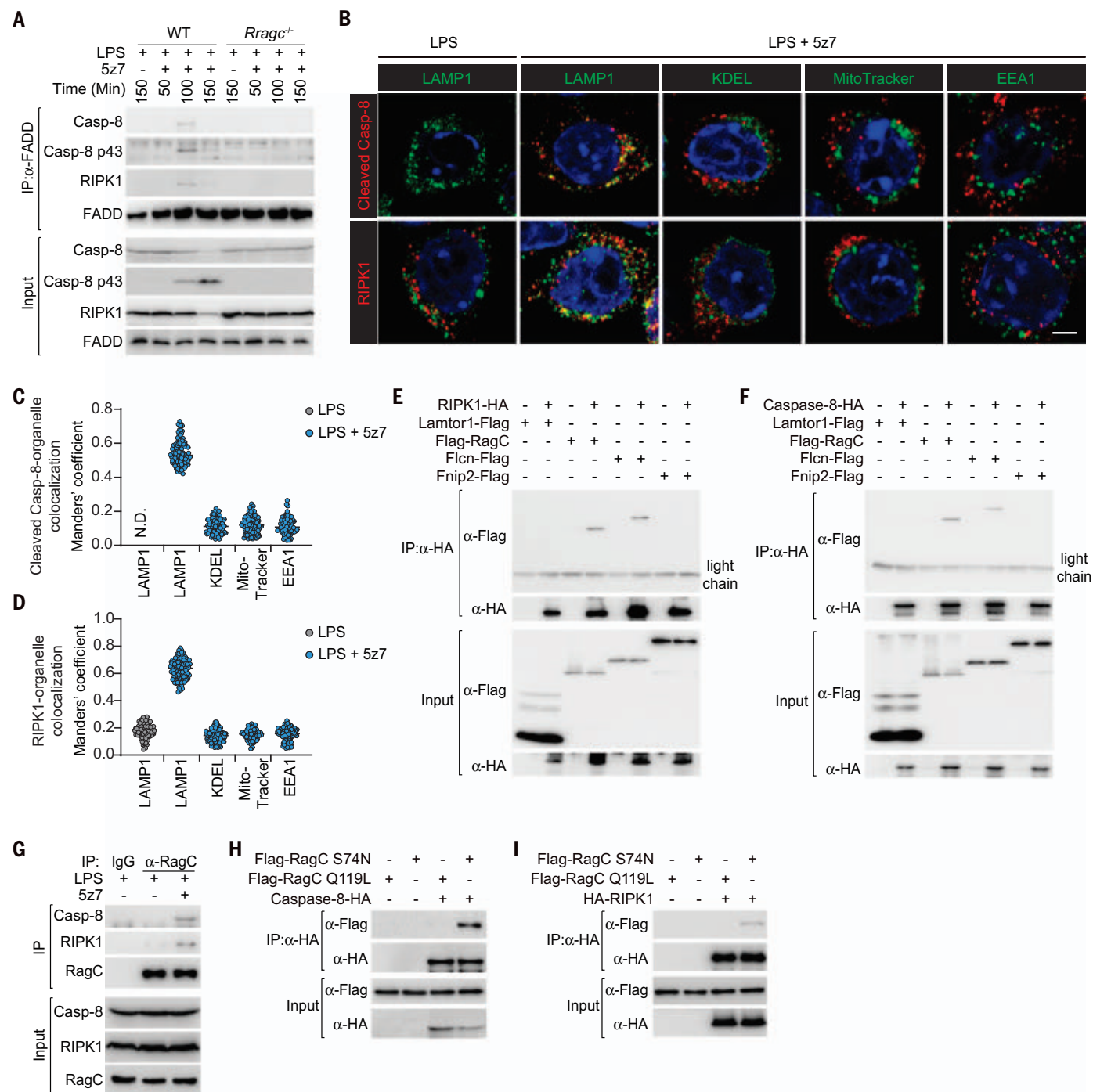
Fig. 3. Recruitment and activation of an LPS/5z7-induced FADD–RIPK1–caspase-8 complex to lysosomes depends on binding to RagC and Flcn in the Rag-Ragulator complex. (**A**) WT and *Rragc*[−/−] iBMDMs were treated with LPS/5z7 for the indicated times. Endogenous FADD complex was immunoprecipitated with anti-FADD antibody and analyzed by immunoblot with the indicated antibodies. (**B**) Representative confocal fluorescence images of LPS or LPS plus 5z7–treated iBMDMs co-stained for cleaved caspase-8 or RIPK1 with the indicated organelle markers (LAMP1, lysosomes; KDEL, endoplasmic reticulum; MitoTracker, mitochondria; EEA1, early endosomes) and DAPI. Scale bars, 2 μm. (**C** and **D**) Quantification of colocalization of cleaved caspase-8 (C) and RIPK1 (D) with organelles in multiple confocal images [as in (B)] by calculating Manders' overlap coefficient (each point represents a single cell). N.D., not detected. (**E**, **F**, **H**, and **I**) Lysates of HEK293T cells transfected with the indicated plasmids were assayed for immunoprecipitation with anti-HA antibody and analyzed by immunoblot probed with the indicated antibodies. (**G**) Lysates of iBMDMs treated with LPS or LPS/5z7 were assayed for immunoprecipitation with anti-RagC antibody and analyzed by immunoblot probed with the indicated antibodies. Data are representative of at least three independent experiments.
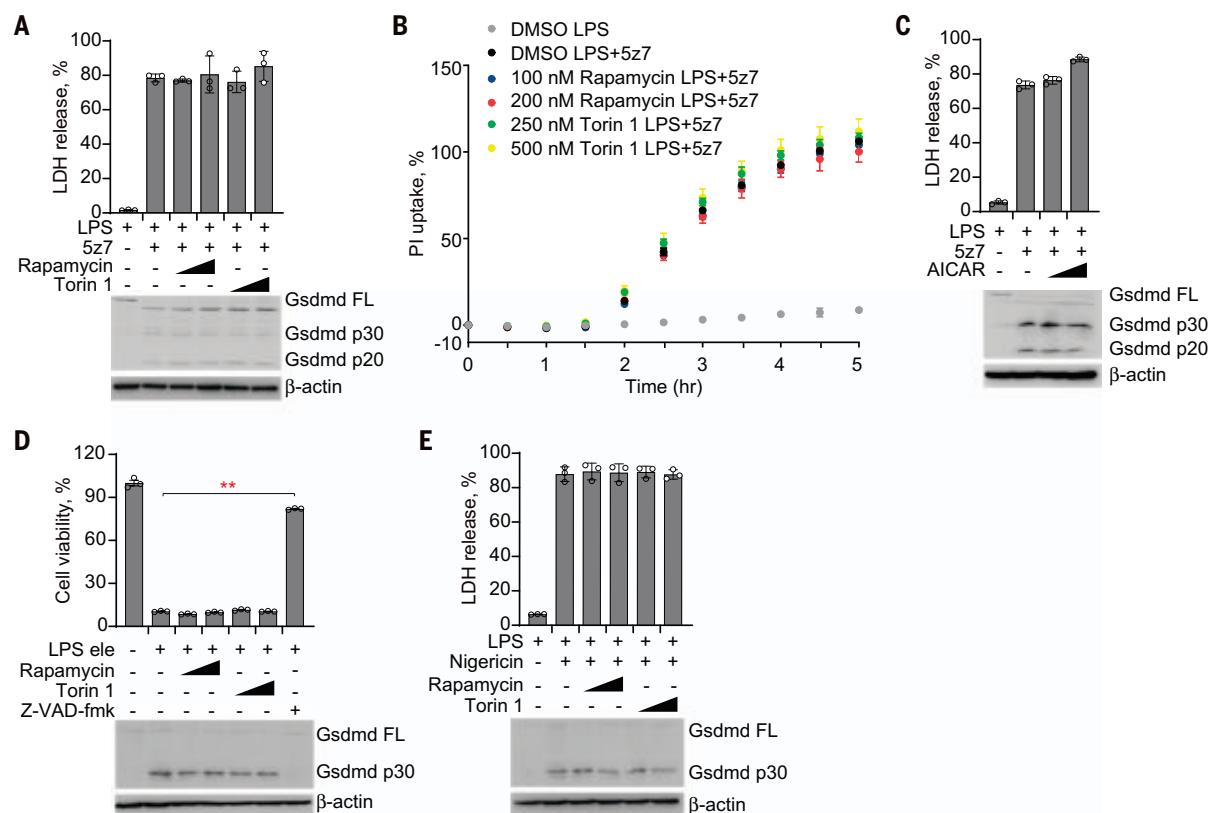
Fig. 4. mTORC1 activity is not required for LPS/5z7- or inflammasome-triggered pyroptosis. (**A** and **B**) iBMDMs preincubated with rapamycin or Torin 1 for 2 hours were treated with LPS or LPS/5z7, and cell death was measured by LDH release after 2.5 hours (A) or by entry of PI into cells (B). (**C**) iBMDMs preincubated with AICAR for 2 hours were treated with LPS or LPS/5z7, and cell death was measured by LDH release 2.5 hours later. (**D**) iBMDMs preincubated with rapamycin or Torin 1 for 2 hours were electroporated (ele) with LPS, and cell viability was assessed by measuring the ATP level 2.5 hours later. (**E**) iBMDMs preincubated with rapamycin or Torin 1 for 2 hours were treated with LPS or LPS/nigericin, and cell death was measured by LDH release 1.5 hours later. Gsdmd cleavage was examined by immunoblot. Graphs show mean ± SEM of triplicate wells. Data are representative of at least three independent experiments. Data were analyzed using a two-tailed Student's *t* test. **P < 0.01.

of the MiT/TFE family transcription factors Tfeb and Tfe3, master regulators of lysosomal biogenesis and autophagy in response to various stresses, showed no effect on LPS/5z7-triggered pyroptosis of Ragulator/Rag/Flcn-deficient cells (fig. S8, A to D). Forced nuclear localization and activation of Tfeb (S141, 210A) also did not affect LPS/5z7-induced pyroptosis (fig. S8, E and F), ruling out the possibility that the defects of Flcn-Rag-Ragulator-knockout cells in LPS/5z7-triggered pyroptosis resulted from constitutive nuclear localization of Tfeb/Tfe3 and subsequent induction of downstream autophagy-lysosomal gene expression (*32*, *33*).

Lysosomal tethering of Ragulator is required for TAK1 inhibition–induced pyroptosis

We next investigated whether lysosomal localization of Rag-Ragulator is required for TAK1 inhibition–induced pyroptosis. Rag-Ragulator is tethered to the cytosolic side of the lysosomal membrane by lipidation of Lamtor1 at its N-terminal G2 and C3C4 residues (*21*, *22*). To address this question, we ectopically expressed WT or a *Lamtor1* mutant of its three

N-terminal lipidation sites to alanine (Lamtor1 3A) in *Lamtor1*[−/−] iBMDMs. Lamtor1 3A was no longer tethered to the lysosome and redistributed to the cytosol (Fig. 5A). Rescue of *Lamtor1*[−/−] by ectopic expression of WT *Lamtor1* in iBMDMs restored LPS/5z7-induced pyroptosis, as assessed by measuring LDH release, PI uptake, and pyroptotic morphological changes (ballooning cell membrane) (Fig. 5, B to D). By contrast, Lamtor1 3A did not rescue LPS/5z7-induced pyroptosis, indicating that Ragulator needs to be tethered to the lysosome to promote LPS/5z7-induced pyroptosis. Overexpression of Lamtor1 3A disrupted the cleavage and lysosomal localization of caspase-8 after LPS/5z7 treatment (fig. S9). Next, we used immunoblotting to determine which step in RIPK1–caspase-8–GSDM–mediated pyroptosis was regulated by Ragulator (Fig. 5E). In WT iBMDMs, phosphorylation of RIPK1 at Ser166 was detected within 100 min of adding LPS/5z7, which was followed by detection of RIPK1 cleavage and cleavage of caspase-8 and GSDMD to their active fragments at 150 min. GSDME was also cleaved concurrently to produce

a p30 active N-terminal fragment. The detection of activated caspase-8, GSDMD, and GSDME fragments coincided with the kinetics of membrane disruption recorded by PI uptake in WT iBMDMs (Fig. 5C). By contrast, *Lamtor1*-deficient iBMDMs showed severely blunted RIPK1 phosphorylation and cleavage and caspase-8, GSDMD, and GSDME activation (Fig. 5E). Reexpression of Lamtor1 restored the defects in RIPK1 and caspase-8 activation as well as GSDM cleavage in Lamtor1-deficient cells. Thus, the activation of RIPK1 and the downstream caspase-8–GSDM pyroptotic pathway only occurs when the RIPK1–caspase-8–containing complex is anchored to the lysosomal membrane by Ragulator.

Next, we examined the role of lysosomal Ragulator in pyroptosis induced by *Y. pseudo-tuberculosis*. Like LPS/5z7 treatment, *Yersinia* infection in WT iBMDMs triggered massive pyroptosis, as assessed by LDH release, which, as expected, was associated with caspase-8 and GSDMD cleavage to their active forms (Fig. 5F). For the first 60 min after infection, before pyroptosis occurred, bacterial replication as
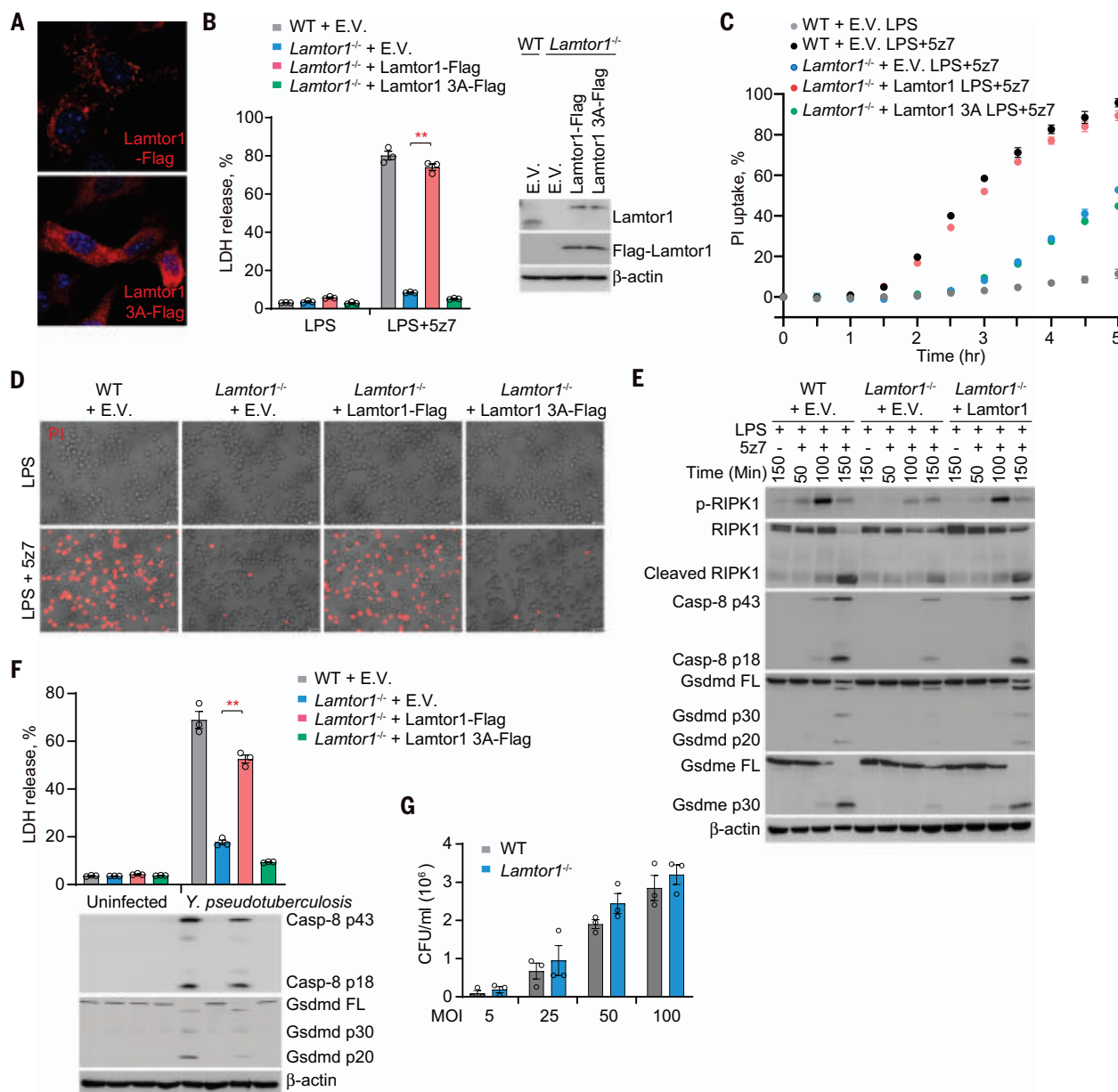
Fig. 5. Pyroptosis activated by LPS/5z7 or pathogenic *Yersinia* depends on lipidated Lamtor1. (**A**) The subcellular localization of Lamtor1 and Lamtor1 3A mutant was imaged by confocal microscopy. (**B**, **C**, **D**, and **F**) The indicated iBMDMs were treated with LPS, LPS/5z7 [(B) to (D)], or *Yersinia* (F). Cell death was measured by LDH release after 2.5 hours (B) or after 5 hours (F), by the entry of PI into cells in real time (C), or as observed by phase-contrast fluorescence microscopy (D). (**E**) Full-length and cleaved products of caspase-8, RIPK1, Gsdmd, and Gsdme from whole-cell lysates of the indicated iBMDMs treated with LPS or LPS/5z7 for the indicated times. (**G**) The amount of *Yersinia* taken up by WT or *Lamtor1*$^{−/−}$ iBMDMs was quantified by counting colony-forming units (CFUs). Graphs in (B), (C), (F), and (G) show mean ± SEM of triplicate wells. Data are representative of at least three independent experiments. Data were analyzed using a two-tailed Student's *t* test. **$P < 0.01$. E.V., empty vector.

measured by colony counts was comparable in WT and *Lamtor1*$^{−/−}$ iBMDMs (Fig. 5G). Lamtor1 deficiency also did not affect the secretion and activity of YopJ (fig. S10). However, cell death and caspase-8 and GSDMD processing were strongly reduced in *Lamtor1*$^{−/−}$ iBMDMs (Fig. 5F). Pyroptosis was rescued by ectopic expression of Flag-tagged WT Lamtor1 but not Lamtor1 3A. Thus, lysosomal targeting of Lamtor1 is also a prerequisite for *Yersinia*-triggered pyroptosis.

Lysosomal Rag GTPase drives RIPK1–caspase-8–induced pyroptosis

Our genetic screen and its validation (Figs. 1 and 2) suggested that RagC and Flcn-Fnip, which activates RagC's GTPase activity by converting RagC into the GDP-loaded form (*22*, *34*, *35*), were required for RIPK1–caspase-8–mediated pyroptosis. To determine which form of RagC promotes TAK1 inhibition–triggered pyroptosis, we reintroduced RagC

WT, the GTP-bound form of RagC (RagC Q119L), or the GDP-bound form of RagC (RagC S74N) into RagC-deficient iBMDMs. Ectopic expression of either RagC WT or the GDP-bound form of RagC restored LPS/5z7-mediated pyroptosis to *Rragc*-knockout iBMDMs, as assessed by LDH release, PI uptake, and ballooning morphology, whereas the GTP-bound form of RagC failed to do so (Fig. 6, A to C). RagC deficiency inhibited activation of RIPK1, caspase-8, GSDMD, and
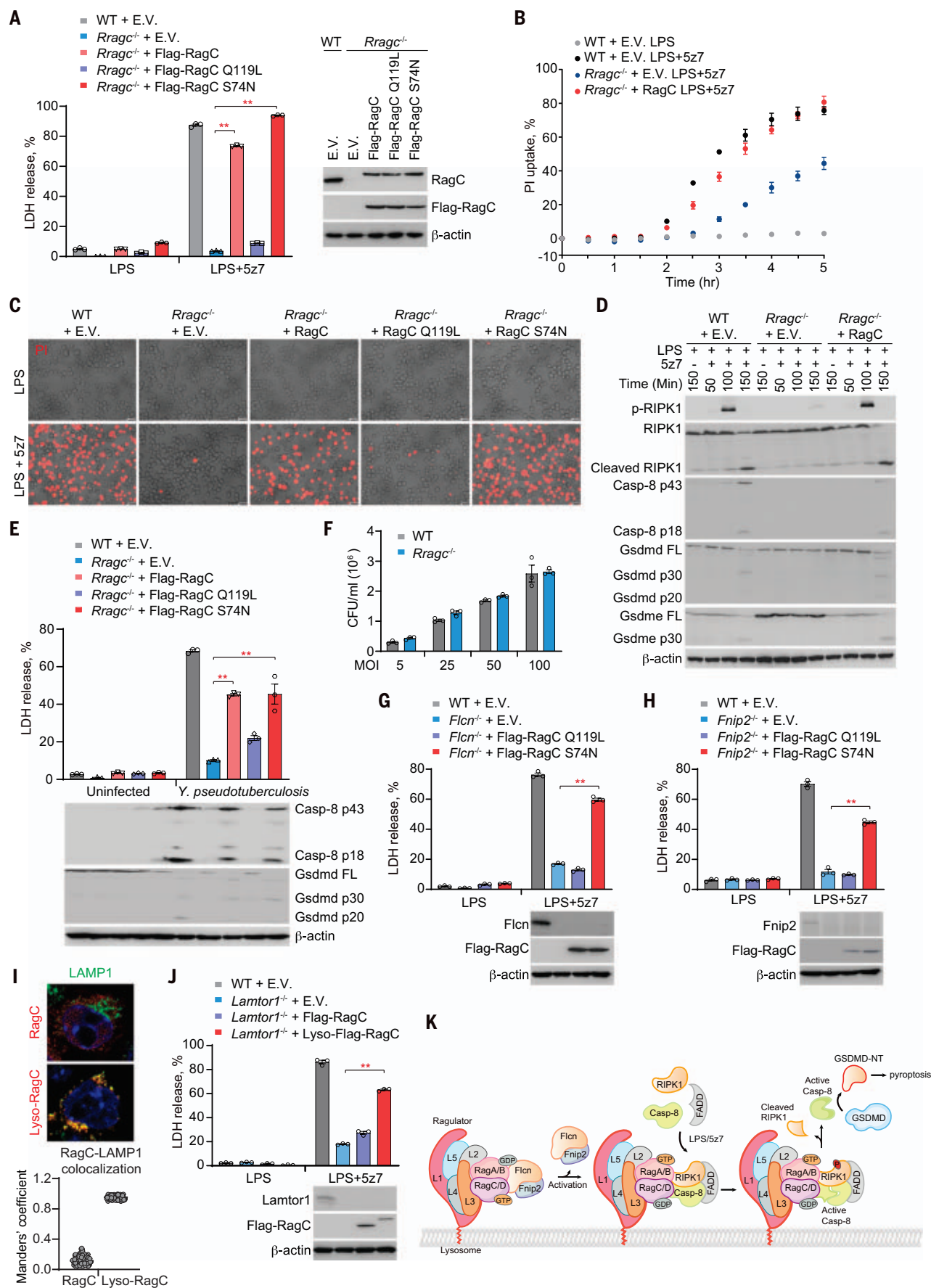
Fig. 6. GTPase activity of lysosomal membrane-anchored RagC is required for caspase-8 activation and pyroptosis triggered by LPS/5z7 or pathogenic *Yersinia*. (**A**, **B**, **C**, and **E**) The indicated iBMDMs were treated with LPS, LPS/5z7, or *Yersinia*. Cell death was measured by LDH release after 2.5 hours (A) or 5 hours (E), by the entry of PI into cells in real time (B), or as observed by phase-contrast fluorescence microscopy (C). (**D**) Full-length and cleaved products of caspase-8, RIPK1, Gsdmd, and Gsdme from whole-cell lysates of the indicated iBMDMs treated with LPS or LPS/5z7 for the indicated times. (**F**) The amount of *Yersinia* taken up by WT or *Rragc*$^{-/-}$ iBMDMs was quantified by counting CFUs. (**G**, **H**, and **J**) WT, *Flcn*$^{-/-}$, *Fnip2*$^{-/-}$, or *Lamtor1*$^{-/-}$ iBMDMs reintroduced with the indicated plasmids were treated with LPS/5z7 and assayed for LDH release 2.5 hours later. (**I**) Lysosomal localization of ectopically expressed RagC and Lyso-RagC in *Lamtor1*$^{-/-}$ iBMDMs was assessed by confocal microscopy (representative images, upper panel). Colocalization of RagC with LAMP1 was analyzed by calculating Manders' overlap coefficient (lower panel). (**K**) Model of RIPK1–caspase-8 activation mediated by Rag-Ragulator complex. Graphs in (A), (B), (E) to (H), and (J) show mean ± SEM of triplicate wells. Data are representative of at least three independent experiments. Data were analyzed using a two-tailed Student's *t* test. **$P < 0.01$.

GSDME, which could be rescued by reexpression of RagC (Fig. 6D). Similarly, ectopic expression of WT or GDP-bound RagC in RagC-deficient iBMDMs made them susceptible to pyroptosis caused by infection of pathogenic *Yersinia*, whereas the GTP-bound form of RagC did not (Fig. 6E). The presence of RagC did not affect bacterial replication at early time points or the secretion and activity of YopJ (Fig. 6F and fig. S10). By comparison, *Salmonella* infection–induced, caspase-1–dependent GSDMD cleavage and pyroptosis in iBMDMs were not affected by *Rragc* knockout or knockout of any of the other Rag-Ragulator complex genes tested (*Lamtor1*, *Lamtor4*, *Flcn*, and *Fnip2*) (fig. S11). Thus, the GDP-bound form of RagC selectively mediates caspase-8–induced pyroptosis. We hypothesized that Ragulator and Flcn-Fnip function upstream of the Rag GTPase to tether it to the lysosome and maintain it in its GDP-bound state. To test this hypothesis, we performed rescue experiments in Flcn- or Fnip-deficient iBMDMs by reintroducing RagC (Fig. 6, G and H). Ectopic expression of GDP-bound RagC restored susceptibility of Flcn- or Fnip2-deficient iBMDMs to LPS/5z7-triggered pyroptosis, whereas the GTP-bound form of RagC did not. To determine whether Ragulator's role was primarily to direct RagC to the lysosomal membrane, we generated a RagC construct in which the lysosomal targeting sequence of Lamtor1 was added to the N terminus of RagC (called Lyso-RagC). Although lysosomal localization of WT RagC was lost in Lamtor1-deficient cells, the lysosomal targeting sequence relocalized Lyso-RagC to the lysosome when Lamtor1 was lacking (Fig. 6I). Ectopic expression of Lyso-RagC made Lamtor1-deficient cells sensitive to LPS/5z7-triggered pyroptosis (Fig. 6J). In addition, LPS/5z7 treatment did not affect the subcellular localization of Flcn (fig. S12). Thus, lysosomal targeting of GDP-bound RagC is sufficient to trigger caspase-8–mediated pyroptosis.

Discussion

Recent studies revealed an alternative inflammasome- and inflammatory caspase–independent, but caspase-8– and gasdermin-dependent, pyroptotic pathway activated by TLR/TNF-α signaling during pathogenic *Yersinia* infection (*12*, *13*). Here, in an unbiased CRISPR screen, we discovered a critical and unexpected role of the lysosomal membrane–tethered Rag-Ragulator supercomplex. Rag-Ragulator served as a platform for activating a FADD–RIPK1–caspase-8 complex formed in response to TLR/TNFR ligation, and Rag GTPase activity was critical for triggering the pathway (Fig. 6K). There is still much to learn. Future kinetics, biochemistry, and live-cell imaging experiments can determine whether the FADD-RIPK1–containing complex recruits caspase-8 before or after Rag-Ragulator binding and whether it contains other components or if its formation is regulated by posttranslational modifications of its component proteins. Our data suggest that the death-receptor–triggered complex binds to RagC and Flcn, but does the Rag-Ragulator complex remain intact after binding or do some components dissociate? Are other proteins recruited? Although we found that mTORC1 activity did not affect caspase-8 activation or pyroptosis, does the FADD–RIPK1–caspase-8 complex compete for binding to Rag-Ragulator with mTORC1? Does cellular metabolism (amino acid and ATP availability and cellular stress, which regulate mTORC1), affect the execution of this cell death pathway? One study suggests that lysosomes are disrupted during inflammasome-mediated pyroptosis (*36*). Do the gasdermins form pores in lysosomal membranes to disrupt lysosomes and cause their proteases to be released into the cytosol and wreak additional damage? RIPK1 phosphorylation is Rag dependent and occurs once the complex is bound to Rag-Ragulator. Activated caspase-8 cleaved and activated both GSDMD and GSDME, as has been shown previously (*12*, *13*), but the mechanism for GSDME cleavage is uncertain. Does caspase-8 directly cleave GSDME or is cleavage indirect through caspase-8 cleavage of caspase-3, a known activator of GSDME? The role of Rag-Ragulator in pyroptosis described here confirms the key role of Rag-Ragulator as a central hub for monitoring environmental cues. This role is now extended to include infection as well as nutrient and energy availability and helping to decide not only whether a cell proliferates but also whether it survives. Future studies can leverage these insights to explore more mechanistic details of pyroptosis, as well as to manipulate this process for therapeutic benefits.

Materials and methods
Plasmids, antibodies, and reagents
psPAX2 (#12260), pMD2.G (#12259), lentiCas9-Blast (#52962), lentiCRISPR v2 (#52961), pFastBacDual with LFn + Fla (#84866), and pFastBacDual with PA+His tag (#84870) constructs were obtained from Addgene. Mouse *Ripk1*, *Casp-8*, *Lamtor1*, *Rragc*, *Flcn*, *Fnip2*, and *Tfeb* were amplified by polymerase chain reaction (PCR) from the mouse cDNA library and cloned into mammalian expression vectors (pHAGE-BSD-Flag, pHAGE-Ble, pcDNA3-NHA, or pcDNA3-CHA). All point mutations and truncations were generated using the overlap PCR method. For short hairpin RNA (shRNA) cloning, annealed shRNA oligos were ligated into pLKO.1 (Addgene #8453). The oligo sequences for targeting the genes of interest were as follows: negative control (5'-CCT AAG GTT AAG TCG CCC TCG-3'), *Tfeb* (5'-GCA GGC TGT CAT GCA TTA TAT-3'), and *Tfe3* (5'-GCC TAA CAT CAA ACG CGA GAT-3'). All plasmids were verified by sequencing.

Monoclonal anti-Flag antibody (F3165) and anti–β-actin antibody (A1978) were from Sigma-Aldrich. Antibodies against HA (#3724, #2367), Flag (#14793), TFEB (#32361), TFE3 (#14779), EEA1 (#48453), p70 S6 Kinase (#2708), Phospho-p70 S6 Kinase (#9234), p38 (#9212), Phospho-p38 (#4511), Erk (#4695), Phospho-Erk (#4370), cleaved caspase-8 (#8592), caspase-8 (#4927), caspase-3 (#9665), caspase-7 (#12827), RIPK1 (#3493), Phospho-RIPK1 (#31122), RagC (#3360, #9480), Lamtor1 (#8975), Lamtor4 (#12284), and FLCN (#3697) were from Cell Signaling Technology. Monoclonal anti-FADD antibody (sc-166516) and monoclonal anti-KDEL antibody (sc-58774) were from Santa Cruz Biotechnology. Monoclonal anti-LAMP1 antibody was from Developmental Studies Hybridoma Bank (1D4B). Monoclonal anti-GSDMD antibody (ab209845), monoclonal anti-GSDME antibody (ab215191), polyclonal anti-FNIP2 antibody (ab106611), monoclonal anti–pro caspase-8 antibody (ab108333), monoclonal anti-LAMP1 antibody (ab208943), and monoclonal anti-FADD antibody (ab124812) were from Abcam. Anti-YopJ antibody was kindly shared by Professor Zongmin Du (Beijing Institute of Microbiology and Epidemiology, China).

LPS (#L4524), etoposide (#E1383), polybrene (#H9268), rapamycin (#V900930), 5z7 (O9890),

and AICAR (#A9978) were from Sigma-Aldrich. Recombinant murine TNF-α (#315-01A) was from Peprotech. Z-VAD-fmk (#550377) was from BD Pharmingen. Ultrapure LPS (tlrl-3pelps), nigericin (tlrl-nig), and blasticidin (anti-bl) were from Invivogen. Torin 1 (#MB3467) was from Meilunbio. Puromycin (#A1113802), zeocin (#R25005), and MitoTracker (#M7510) were from Invitrogen.

Cell culture and stimulation

HEK293T, mouse embryonic fibroblasts, and C57BL/6 mouse iBMDM cells, described previously (7), were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum, 1× GlutaMAX (Thermo-Fisher, #35050061), 1× penicillin-streptomycin (ThermoFisher Scientific, #15140148), and 50 μM 2-Mercaptoethanol (β-ME), and verified to be free of mycoplasma contamination. To induce pyroptosis, cells were stimulated with 10 ng/ml LPS plus 250 nM 5z7, 50 ng/ml TNF-α plus 250 nM 5z7, 1 μg/ml LPS plus 20 μM nigericin, or 2 μg/ml PA plus 1 μg/ml LFn-Fla. For noncanonical inflammasome activation, 2 million iBMDMs were mixed with 2 μg of ultrapure LPS in 10 μl of buffer R and electroporated using the Neon Transfection System (ThermoFisher Scientific) with parameters (1400 V, 10 ms, and 3 pulses). To induce apoptosis, iBMDMs were treated with 50 μM etoposide. Transient transfection of HEK293T cells was performed with Lipofectamine 3000 (Invitrogen) according to the manufacturer's instructions. For stable transfection, lentiCas9-Blast, pHAGE-BSD-Flag, or pHAGE-Ble containing the gene of interest, together with packaging plasmids psPAX2 and pMD2.G in a ratio of 5:3:2, were transiently transfected into HEK293T cells. At 72 hours after transfection, the cell culture supernatant containing lentivirus was collected and filtered through a 0.45-μm membrane (Millipore, #SLHV033RB) to remove cell debris. iBMDMs were then infected with the lentivirus for 3 days before adding 10 μg/ml blasticidin or 400 μg/ml zeocin for stable transfection selection.

Genome-wide CRISPR-Cas9 screen

The mouse CRISPR-knockout pooled library for genome-wide CRISPR/Cas9 screen was obtained from Addgene (#1000000053) and amplified according to the protocol provided by the manufacturer. For lentivirus production, mouse GeCKO A and B library plasmids were mixed equally and transfected into HEK293T cells together with the packaging plasmids psPAX2 and pMD2.G. Seventy-two hours later, lentivirus was collected and the viral titer was measured with the QuickTiter Lentivirus Titer kit (Cell Biolabs, #VPK-107). For the genome-wide screen, Cas9 stably expressing iBMDMs were seeded in 10-cm dishes

(2×10^6 cells/dish), and 7×10^7 cells were infected with the lentivirus-containing sgRNA library at a multiplicity of infection (MOI) of 0.3. Sixty hours later, cells were treated with puromycin to remove uninfected cells. Six days after that, the transduced cells were seeded in 40×10 cm dishes (8×10^6 cells/dish) and treated with 10 ng/ml LPS plus 125 nM 5z7 for 6 hours. Three days later, the surviving cells were reseeded and treated with 10 ng/ml LPS plus 250 nM 5z7 overnight. Two days later, surviving cells were again treated with 20 ng/ml LPS plus 250 nM 5z7 overnight before surviving cells were harvested. For the noncanonical inflammasome screen, 3×10^8 transduced cells were electroporated with ultra LPS for three rounds. Surviving cells and untreated transduced cells (as the control sample) were harvested and lysed in the SNET buffer [20 mM Tris-HCl pH 8.0, 5 mM EDTA, 400 mM NaCl, 1% SDS and 400 μg/ml Proteinase K (NEB, P8107S)]. Genomic DNAs were prepared by using phenol-chloroform extraction and isopropanol precipitation and amplified by two-step PCR using the 2× Hieff Canace Gold PCR Master Mix (Yeasen, #10149ES01). The samples were quantified and sequenced on a HiSeq 2500 (Illumina) by GENEWIZ. Sequencing data were further processed and analyzed using the MAGeCK algorithm (37). MAGeCK built a linear model to estimate the variance of guide RNA (gRNA) read counts, evaluated the gRNA abundance changes between control and treatment conditions, and assigned P values for positive and negative selection.

Generation of knockout iBMDMs by CRISPR-Cas9

HEK293T cells were transfected with sgRNA-expressing lentiCRISPR v2 together with psPAX2 and pMD2.G at a ratio of 5:3:2. Then, 72 hours later, the supernatant containing lentivirus was collected and filtered through a 0.45-μm membrane (Millipore) for subsequent iBMDM infection. Three days after infection, cells were treated with 5 μg/ml puromycin before diluting to single clones cultured in 96-well plates. The candidate knockout clones were verified by sequencing of genomic DNA and immunoblot. sgRNA sequences for targeting the genes of interest were as follows: *Lamtor1* (5'-TGG ACC GGG CAA GGC AGT AC-3'/5'-GCT CTT CTT TCG CAT CCA CG-3'), *Lamtor4* (5'-AGC CAG TGC CAT CTC GGA GT-3'/5'-TAG ACT TCC GCA CTG ACC CA-3'), *Rrag*c (5'-TTT CTG TAC CAC CTT ACT GA-3'/5'-TCA TAA GAC TGC ATA TCC AC-3'), *Flcn* (5'-GGC TGC CGG TCA CTT GCC GT-3'/5'-GCC TGC TAC CGC ATG CCT TC-3'), *Fnip2* (5'-ACC GTA TGT AGT GTA TCT TC-3'/5'-ACT TTA CTA ATC ATC AGT TG-3'), *Gsdmd* (5'-AGC ATC CTG GCA TTC CGA G-3'), *Nlrp3* (5'-GAA GAT TAC CCG CCC GAG AA-3'), and *Nlrc4* (5'-TGT TTC GAA TAG TCC CCC CC-3').

Recombinant protein purification

PA and LFn-flagellin recombinant proteins were purified from Sf9 cells. Seventy-two hours after P3 baculovirus infection, cells were resuspended in lysis buffer (50 mM Tris-HCl, pH 8.0, 300 mM NaCl, 10 mM β-ME, 5 mM imidazole, and 1% Triton X-100) and then clarified by centrifugation at 42,000 rpm for 2 hours at 4°C. Proteins were purified using Ni-NTA agarose (QIAGEN), eluted with elution buffer (50 mM Tris-HCl pH 8.0, 300 mM NaCl, 10 mM β-ME, 500 mM imidazole), and further purified by Superdex 200 (10/300) gel-filtration chromatography and mono-Q ion exchange.

Bacterial strains and culture conditions

The *Y. pseudotuberculosis* YPIII strain, a gift from Dr. Shiyun Chen (Wuhan Institute of Virology, Chinese Academy of Sciences), was grown overnight in 2× YT broth at 26°C. On the day of infection, bacteria were diluted 1:50 into 2× YT plus 20 mM $MgCl_2$ and 20 mM sodium oxalate and grown for 2 hours at 26°C, followed by a shift to 37°C for 2 hours. Bacteria were then washed in phosphate-buffered saline (PBS; Invitrogen) and added to cells at a MOI of 40. Next, 100 μg/ml gentamicin was added to the cultures 2 hours after infection. To quantify the number of bacteria that had been taken up by cells, iBMDMs were infected with the *Y. pseudotuberculosis* YPIII strain at the MOI indicated in the corresponding figures. Thirty minutes later, cells were washed with PBS three times and gentamicin was added to kill extracellular bacteria. Then, intracellular bacteria were released by treating cells with 0.05% Triton X-100 before lysates were serially diluted and plated on 2× YT agar. Bacterial colonies were counted after 1 day of culture at 37°C. For the *Salmonella typhimurium* infection assay, bacteria were grown overnight in Luria-Bertani (LB) broth at 37°C. On the day of infection, bacteria were diluted 1:50 into LB broth and grown for 4 hours at 37°C. Bacteria were then washed with PBS and added to cells at a MOI of 20. Next, 100 μg/ml gentamicin was added to the cultures 0.5 hours after infection to kill extracellular bacteria.

Cell death assays

Cell death and viability of stimulated macrophages were determined by measuring LDH release using the CytoTox 96 assay kit (Promega, #G1780) and ATP level using the CellTiter-Glo Luminescent Cell Viability Assay (Promega, #G7570), respectively. Luminescence and absorbance were measured on a BioTek Synergy H1 plate reader. For kinetic cytotoxicity assay by monitoring cell permeability, the cell culture medium was changed into buffer B [25 mM HEPES, pH 7.4, 120 mM NaCl, 4 mM KCl, 1.5 mM $CaCl_2$, 1 mM $MgCl_2$, 5mM glucose, 0.1% bovine serum albumin (BSA)] containing 2 μg/ml PI (Invitrogen, #P3566), and the

fluorescence (excitation: 535 nm, emission: 617 nm) was continuously recorded after treatment for 5 hours at 30-min intervals using a BioTek Synergy plate reader.

Immunoprecipitation and immunoblotting

For the immunoprecipitation assay, cells were harvested and resuspended in lysis buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1% Triton X-100) supplemented with a complete protease inhibitor cocktail (Sigma-Aldrich). Cell extracts were then incubated with the indicated antibodies for 4 hours at 4°C before adding protein A-G agarose beads for 2 hours. The beads were extensively washed at least three times with wash buffer (50 mM Tris-HCl, pH 7.4, 300 mM NaCl, 0.5% Triton X-100) and bound proteins were eluted by boiling with SDS-loading buffer. For immunoblot analysis, immunoprecipitation samples or whole-cell lysate samples were subjected to electrophoresis through SDS–polyacrylamide gel electrophoresis gels. The separated proteins were then transferred to a polyvinylidene fluoride membrane (Millipore). Immunoblot was probed with the indicated antibodies. The protein bands were visualized using a SuperSignal West Pico chemiluminescence ECL kit (Pierce).

Immunostaining and confocal microscopy

iBMDMs grown on coverslips were stimulated with 10 ng/ml LPS plus 250 nM 5z7. Cells were fixed 100 min later with 4% paraformaldehyde for 20 min before permeabilization using 0.1% Triton X-100 (10 min) and blocking with 5% BSA (1 hour). Then, cells were immunostained with the indicated primary antibodies, followed by incubation with the corresponding fluorescent-conjugated secondary antibodies (Jackson ImmunoResearch). Nuclei were counterstained with 4,6-diamidino-2-phenylindole (DAPI; Cell Signaling Technology). Slides were mounted using Aqua-Poly/Mount (Dako). Images were captured using a Zeiss 880 laser scanning confocal microscope at 63× magnification and analyzed using Zeiss Zen software. Manders' overlap coefficient was calculated using ImageJ (where each point represents a single cell, with 100 cells per sample). All images are representative of at least three independent experiments.

Statistics

Student's t test (two-tailed) was used for the statistical analysis of all experiments. $P < 0.05$ was considered significant.

REFERENCES AND NOTES

1. P. Broz, V. M. Dixit, Inflammasomes: Mechanism of assembly, regulation and signalling. *Nat. Rev. Immunol.* **16**, 407–420 (2016). doi: 10.1038/nri.2016.58; pmid: 27291964
2. N. M. de Vasconcelos, M. Lamkanfi, Recent insights on inflammasomes, gasdermin pores, and pyroptosis. *Cold Spring Harb. Perspect. Biol.* **12**, a036392 (2020). doi: 10.1101/cshperspect.a036392; pmid: 31570336
3. D. R. Green, The coming decade of cell death research: Five riddles. *Cell* **177**, 1094–1107 (2019). doi: 10.1016/j.cell.2019.04.024; pmid: 31100266
4. N. Kayagaki *et al.*, Caspase-11 cleaves gasdermin D for non-canonical inflammasome signalling. *Nature* **526**, 666–671 (2015). doi: 10.1038/nature15541; pmid: 26375259
5. J. Shi *et al.*, Cleavage of GSDMD by inflammatory caspases determines pyroptotic cell death. *Nature* **526**, 660–665 (2015). doi: 10.1038/nature15514; pmid: 26375003
6. J. Ding *et al.*, Pore-forming activity and structural autoinhibition of the gasdermin family. *Nature* **535**, 111–116 (2016). doi: 10.1038/nature18590; pmid: 27281216
7. X. Liu *et al.*, Inflammasome-activated gasdermin D causes pyroptosis by forming membrane pores. *Nature* **535**, 153–158 (2016). doi: 10.1038/nature18629; pmid: 27383986
8. L. Sborgi *et al.*, GSDMD membrane pore formation constitutes the mechanism of pyroptotic cell death. *EMBO J.* **35**, 1766–1778 (2016). doi: 10.15252/embj.201694696; pmid: 27418190
9. R. A. Aglietti *et al.*, GsdmD p30 elicited by caspase-11 during pyroptosis forms pores in membranes. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 7858–7863 (2016). doi: 10.1073/pnas.1607769113; pmid: 27339137
10. R. Heilig *et al.*, The Gasdermin-D pore acts as a conduit for IL-1β secretion in mice. *Eur. J. Immunol.* **48**, 584–592 (2018). doi: 10.1002/eji.201747404; pmid: 29274245
11. C. L. Evavold *et al.*, The pore-forming protein Gasdermin D regulates interleukin-1 secretion from living macrophages. *Immunity* **48**, 35–44.e6 (2018). doi: 10.1016/j.immuni.2017.11.013; pmid: 29195811
12. P. Orning *et al.*, Pathogen blockade of TAK1 triggers caspase-8-dependent cleavage of gasdermin D and cell death. *Science* **362**, 1064–1069 (2018). doi: 10.1126/science.aau2818; pmid: 30361383
13. J. Sarhan *et al.*, Caspase-8 induces cleavage of gasdermin D to elicit pyroptosis during *Yersinia* infection. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E10888–E10897 (2018). doi: 10.1073/pnas.1809548115; pmid: 30381458
14. N. Kayagaki *et al.*, IRF2 transcriptionally induces *GSDMD* expression for pyroptosis. *Sci. Signal.* **12**, eaax4917 (2019). doi: 10.1126/scisignal.aax4917; pmid: 31113851
15. L. Bar-Peled, D. M. Sabatini, Regulation of mTORC1 by amino acids. *Trends Cell Biol.* **24**, 400–406 (2014). doi: 10.1016/j.tcb.2014.03.003; pmid: 24698685
16. J. Hirst, Mitochondrial complex I. *Annu. Rev. Biochem.* **82**, 551–575 (2013). doi: 10.1146/annurev-biochem-070511-103700; pmid: 23527692
17. C. Rogers *et al.*, Gasdermin pores permeabilize mitochondria to augment caspase-3 activation during apoptosis and inflammasome activation. *Nat. Commun.* **10**, 1689 (2019). doi: 10.1038/s41467-019-09397-2; pmid: 30976076
18. W. J. Kaiser *et al.*, RIP3 mediates the embryonic lethality of caspase-8-deficient mice. *Nature* **471**, 368–372 (2011). doi: 10.1038/nature09857; pmid: 21368762
19. A. Oberst *et al.*, Catalytic activity of the caspase-8-FLIP(L) complex inhibits RIPK3-dependent necrosis. *Nature* **471**, 363–367 (2011). doi: 10.1038/nature09852; pmid: 21368763
20. D. Vercammen *et al.*, Inhibition of caspases increases the sensitivity of L929 cells to necrosis mediated by tumor necrosis factor. *J. Exp. Med.* **187**, 1477–1485 (1998). doi: 10.1084/jem.187.9.1477; pmid: 9565639
21. S. Nada *et al.*, The novel lipid raft adaptor p18 controls endosome dynamics by anchoring the MEK-ERK pathway to late endosomes. *EMBO J.* **28**, 477–489 (2009). doi: 10.1038/emboj.2008.308; pmid: 19177150
22. Y. Sancak *et al.*, Ragulator-Rag complex targets mTORC1 to the lysosomal surface and is necessary for its activation by amino acids. *Cell* **141**, 290–303 (2010). doi: 10.1016/j.cell.2010.02.024; pmid: 20381137
23. H. I. Muendlein *et al.*, ZBP1 promotes LPS-induced cell death and IL-1β release via RHIM-mediated interactions with RIPK1. *Nat. Commun.* **12**, 86 (2021). doi: 10.1038/s41467-020-20357-z; pmid: 33397971
24. H. I. Muendlein *et al.*, cFLIPₗ protects macrophages from LPS-induced pyroptosis via inhibition of complex II formation. *Science* **367**, 1379–1384 (2020). doi: 10.1126/science.aay3878; pmid: 32193329
25. I. Matsuda *et al.*, The C-terminal domain of the long form of cellular FLICE-inhibitory protein (c-FLIPL) inhibits the interaction of the caspase 8 prodomain with the receptor-interacting protein 1 (RIP1) death domain and regulates caspase 8-dependent nuclear factor κB (NF-κB) activation. *J. Biol. Chem.* **289**, 3876–3887 (2014). doi: 10.1074/jbc.M113.506485; pmid: 24398693
26. Y. H. Park, M. S. Jeong, H. H. Park, S. B. Jang, Formation of the death domain complex between FADD and RIP1 proteins in vitro. *Biochim. Biophys. Acta* **1834**, 292–300 (2013). doi: 10.1016/j.bbapap.2012.08.013; pmid: 22922561
27. H. Anderton *et al.*, RIPK1 prevents TRADD-driven, but TNFR1 independent, apoptosis during development. *Cell Death Differ.* **26**, 877–889 (2019). doi: 10.1038/s41418-018-0166-8; pmid: 30185824
28. M. Muzio *et al.*, FLICE, a novel FADD-homologous ICE/CED-3-like protease, is recruited to the CD95 (Fas/APO-1) death—Inducing signaling complex. *Cell* **85**, 817–827 (1996). doi: 10.1016/S0092-8674(00)81266-0; pmid: 8681377
29. T. M. Fu *et al.*, Cryo-EM structure of caspase-8 tandem DED filament reveals assembly and regulation mechanisms of the death-inducing signaling complex. *Mol. Cell* **64**, 236–250 (2016). doi: 10.1016/j.molcel.2016.09.009; pmid: 27746017
30. K. Shen *et al.*, Cryo-EM structure of the human FLCN-FNIP2-Rag-Ragulator complex. *Cell* **179**, 1319–1329.e8 (2019). doi: 10.1016/j.cell.2019.10.036; pmid: 31704029
31. C. C. Thoreen *et al.*, An ATP-competitive mammalian target of rapamycin inhibitor reveals rapamycin-resistant functions of mTORC1. *J. Biol. Chem.* **284**, 8023–8032 (2009). doi: 10.1074/jbc.M900301200; pmid: 19150980
32. R. E. Lawrence *et al.*, Structural mechanism of a Rag GTPase activation checkpoint by the lysosomal folliculin complex. *Science* **366**, 971–977 (2019). doi: 10.1126/science.aax0364; pmid: 31672913
33. G. Napolitano *et al.*, A substrate-specific mTORC1 pathway underlies Birt-Hogg-Dubé syndrome. *Nature* **585**, 597–602 (2020). doi: 10.1038/s41586-020-2444-0; pmid: 32612235
34. L. Bar-Peled, L. D. Schweitzer, R. Zoncu, D. M. Sabatini, Ragulator is a GEF for the rag GTPases that signal amino acid levels to mTORC1. *Cell* **150**, 1196–1208 (2012). doi: 10.1016/j.cell.2012.07.032; pmid: 22980980
35. Z. Y. Tsun *et al.*, The folliculin tumor suppressor is a GAP for the RagC/D GTPases that signal amino acid levels to mTORC1. *Mol. Cell* **52**, 495–505 (2013). doi: 10.1016/j.molcel.2013.09.016; pmid: 24095279
36. N. M. de Vasconcelos, N. Van Opdenbosch, H. Van Gorp, E. Parthoens, M. Lamkanfi, Single-cell analysis of pyroptosis dynamics reveals conserved GSDMD-mediated subcellular events that precede plasma membrane rupture. *Cell Death Differ.* **26**, 146–161 (2019). doi: 10.1038/s41418-018-0106-7; pmid: 29666477
37. W. Li *et al.*, Quality control, modeling, and visualization of CRISPR screens with MAGeCK-VISPR. *Genome Biol.* **16**, 281 (2015). doi: 10.1186/s13059-015-0843-6; pmid: 26673418

ScienceSignaling.org

PUT YOUR RESEARCH OUT IN FRONT

Submit your research:
cts.ScienceMag.org

Science Signaling

AAAS

Twitter: @SciSignal
Facebook: @ScienceSignaling

REPORTS

CORONAVIRUS

mRNA vaccination boosts cross-variant neutralizing antibodies elicited by SARS-CoV-2 infection

Leonidas Stamatatos[1,2]*, Julie Czartoski[1], Yu-Hsin Wan[1], Leah J. Homad[1], Vanessa Rubin[1], Hayley Glantz[1], Moni Neradilek[1], Emilie Seydoux[1], Madeleine F. Jennewein[1], Anna J. MacCamy[1], Junli Feng[1], Gregory Mize[1], Stephen C. De Rosa[1,3], Andrés Finzi[4,5,6], Maria P. Lemos[1], Kristen W. Cohen[1], Zoe Moodie[1], M. Juliana McElrath[1,2,7]*, Andrew T. McGuire[1,2,3]*

Emerging severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) variants have raised concerns about resistance to neutralizing antibodies elicited by previous infection or vaccination. We examined whether sera from recovered and naïve donors, collected before and after immunizations with existing messenger RNA (mRNA) vaccines, could neutralize the Wuhan-Hu-1 and B.1.351 variants. Prevaccination sera from recovered donors neutralized Wuhan-Hu-1 and sporadically neutralized B.1.351, but a single immunization boosted neutralizing titers against all variants and SARS-CoV-1 by up to 1000-fold. Neutralization was a result of antibodies targeting the receptor binding domain and was not boosted by a second immunization. Immunization of naïve donors also elicited cross-neutralizing responses but at lower titers. Our study highlights the importance of vaccinating both uninfected and previously infected persons to elicit cross-variant neutralizing antibodies.

The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) betacoronavirus first emerged in the Hubei Province of China in late 2019 and has since infected more than 115 million people and caused more than 2.5 million deaths in 192 countries (1–3). Infection is mediated by the viral spike protein (S), which is composed of an S1 domain that contains an N-terminal domain (NTD), a C-terminal domain (CTD), and a receptor binding domain (RBD) that mediates attachment to the entry receptor angiotensin-converting enzyme 2 (ACE2) as well as an S2 domain that contains the fusion machinery (4–8).

Preexisting immunity to SARS-CoV-2 is associated with protection against reinfection in humans (9–11) and in nonhuman primates (12, 13). Although the correlates of protection in humans against repeat infection or after vaccination have not been firmly established, neutralizing antibodies (nAbs) are thought to be an important component of a protective immune response against SARS-CoV-2 (14, 15). In support of this, passive transfer of nAbs limits respiratory tract infection and protects against infection in animal models (16–20), and nAbs may contribute to protection against infection in humans (9). SARS-CoV-2 infection rapidly elicits nAbs (16, 21–24) that decline, but remain detectable, over several months (25–29).

Most serum nAb responses elicited during natural infection are directed at the RBD (21, 23, 30, 31). Numerous neutralizing anti-RBD monoclonal antibodies (mAbs) have been characterized, the most potent of which block the RBD-ACE2 interaction (16, 17, 22–24, 32–37). Neutralizing mAbs that bind regions of the viral spike have also been identified (24, 33, 38–42).

Two mRNA-based vaccines (Pfizer-BioNTech BNT162b2 and Moderna mRNA-1273) have received emergency use authorization in several countries. Both vaccines encode a stabilized ectodomain version of the S protein derived from the Wuhan-Hu-1 variant isolated in December 2019 (43), show >94% efficacy at preventing COVID-19 illness (44–47), and elicit nAbs (48, 49).

Because of the high global burden of SARS-CoV-2 transmission, viral evolution is occurring. Recently, viral variants of concern have emerged in the UK (B.1.1.7), South Africa (B.1.351), and Brazil (P.1) that harbor specific mutations in their S proteins that may be associated with increased transmissibility (50–55).

Of particular concern are mutations found in the B.1.351 lineage, which is defined by the D80A (amino acid substitution from aspartic acid to alanine at position 80) and D215G mutations in the NTD; the K417N, E484K, and N501Y mutations in the RBD; and the D614G mutation in S1 (52, 56). An A701V mutation in S2 is also observed at high frequencies, where-

as deletions in residues 242 to 244 as well as R246I and L18F mutations in the NTD are present at lower frequencies (52). (Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.)

The B.1.1.7, B.1.351, and P.1 lineages all harbor a N501Y mutation in the RBD, which increases the affinity for the ACE2 receptor (57, 58), and a D614G mutation, which increases virion spike density, infectivity, and transmissibility (59, 60). The B.1.351 and P.1 lineages also share the E484K mutation in the RBD, and both variants are mutated at position 417 (K417T in P.1).

Mutations found in emergent S variants decrease sensitivity to neutralization by mAbs, convalescent plasma, and sera from vaccinated individuals (27, 37, 58, 61–70). As a result, there is concern that these and other emerging variants can evade nAb responses generated during infection with variants that were circulating earlier in the pandemic and also nAb responses elicited by vaccines based on the S protein of the Wuhan-Hu-1 variant. There is concern that these mutations are responsible for the reduced efficacy observed in ongoing trials of SARS-CoV-2 vaccines in South Africa (71, 72).

Here, we evaluated the neutralization susceptibility of spike variants harboring lineage-defining and prevalent B.1.351 mutations to sera from two groups. Sera were collected from 15 donors with previously confirmed SARS-CoV-2 infection [referred to as previously infected donors (PIDs)] before and after one or two immunizations with either mRNA vaccine and from 13 uninfected donors who received two doses of the above vaccines [referred to as naïve donors (NDs); tables S1 and S2].

Antibody neutralization experiments were performed with pseudoviruses expressing either the full-length Wuhan-Hu-1 S or either of two versions of the B.1.351 lineage S—one herein referred to as B.1.351, containing the lineage-defining S mutations D80A, D215G, K417N, E484K, N501Y, and D614G and the A701V mutation that is highly prevalent in this lineage, and a second variant that also includes a Δ242-243 deletion (B.1.351-Δ242-243). The viral stocks were appropriately diluted to achieve comparable entry levels during the neutralization experiments (fig. S1).

We first evaluated the neutralizing potency of several mAbs isolated from nonvaccinated patients infected early in the pandemic. These mAbs target different epitopes: three against the RBD (CV30, CV3-1, and CV2-75) and one against the NTD (CV1) (fig. S2). CV30 is a member of the VH3-53 class of antibodies that bind to the receptor binding motif (RBM) (22, 32, 73–78). It makes direct contact with the K417 and

[1]Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA. [2]Department of Global Health, University of Washington, Seattle, WA, USA. [3]Department of Laboratory Medicine and Pathology, University of Washington, Seattle, WA, USA. [4]Centre de Recherche du CHUM, Montréal, QC, Canada. [5]Département de Microbiologie, Infectiologie et Immunologie, Université de Montréal, Montreal, QC, Canada. [6]Department of Microbiology and Immunology, McGill University, Montreal, QC, Canada. [7]Department of Medicine, University of Washington, Seattle, WA, USA.
*Corresponding author. Email: amcguire@fredhutch.org (A.T.M.); lstamata@fredhutch.org (L.S.); jmcelrat@fredhutch.org (M.J.M.)

Fig. 1. B.1.351 variants show decreased susceptibility to neutralizing mAbs. (A to E) The ability of the indicated mAbs to neutralize Wuhan-Hu-1, B.1.351, and B.1.351–Δ242-243 pseudovirus infectivity in 293T-hACE2 cells was measured as indicated. The epitope specificity of each mAb is shown in parentheses. EBV, Epstein-Barr virus. Data points represent the mean of two technical replicates. Data are representative of two independent experiments.
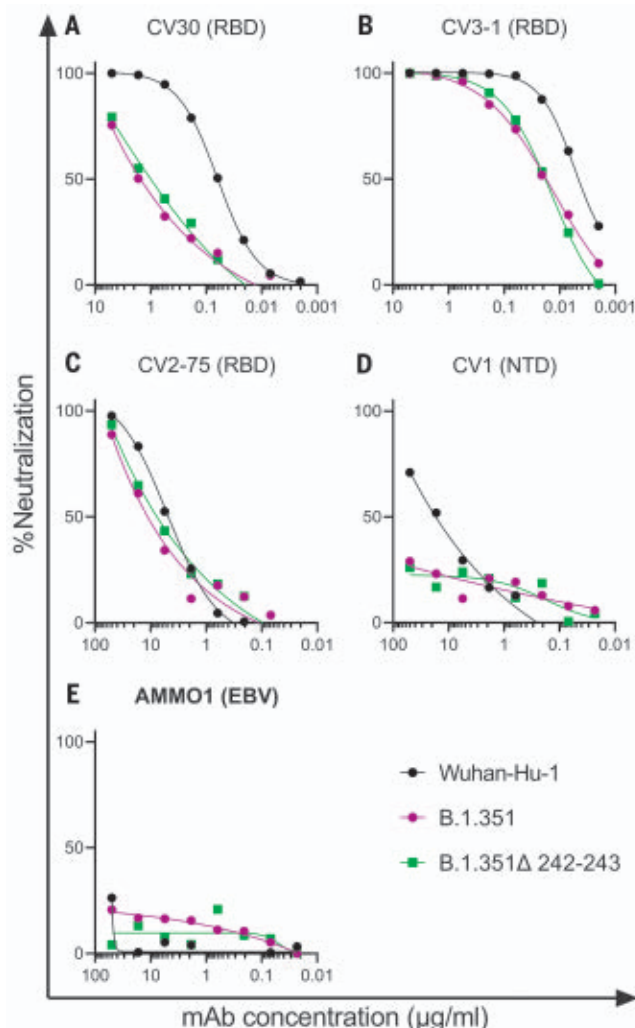
N501 residues in the RBM that are mutated in the B.1.351 and P.1 lineages; however, unlike other known VH3-53 mAbs, it does not contact E484 (78). The neutralization potency of this mAb was ~10-fold weaker toward both B.1.351 variants (Fig. 1A). Similarly, the non–VH3-53 mAb CV3-1 was three- to fourfold less potent against the B.1.351 variants (Fig. 1B), whereas CV2-75 was modestly less effective (Fig. 1C). By contrast, the anti-NTD CV1 mAb was unable to neutralize either B.1.351 variant (Fig. 1D). As expected, the control anti–Epstein-Barr virus mAb AMMO1 was nonneutralizing (79) (Fig. 1E). Collectively, these data indicate that the B.1.351 variants tested here are more resistant to neutralization by mAbs isolated from subjects infected by viral variants from early in the pandemic. We therefore examined whether the B.1.351 variants are resistant to nAb responses elicited by the Pfizer-BioNTech or Moderna mRNA vaccines in both PIDs and NDs.

The RBD-specific immunoglobulin G (IgG), IgM, and IgA binding responses to the RBD from the Wuhan-Hu-1 variant were measured before (on average, 202 days after symptom onset; table S1) and either 5 to 29 days (table S1) after the first and second immunizations in the PIDs or 6 to 28 days after the second immunization in the NDs. Three PIDs experienced asymptomatic SARS-CoV-2 infection (donors D, L, and M; table S1), two of whom, L and M, did not have detectable anti-RBD IgG antibodies before immunization, whereas the third, D, had low but detectable serum anti-RBD IgG antibody titers (Fig. 2A). In the 13 PIDs with RBD-specific IgG antibodies before vaccination, a single dose of either vaccine boosted these titers ~500-fold (Fig. 2A). Across all PIDs, there was a 200-fold increase in median RBD-specific IgA titers after vaccination (Fig. 2B). Overall, in PIDs, a single vaccine dose elicited 4.5-fold higher IgG and 7.7-fold higher IgA titers compared with two vaccinations in NDs. RBD-specific IgM titers were generally lower and were not significantly boosted in response to vaccination in PIDs (Fig. 2C). In PIDs, a concomitant increase in RBD- (Fig. 2D) and S-specific IgG+ (Fig. 2E) memory B cell fre-

quencies took place after vaccination. The two PIDs that lacked RBD-specific IgG titers before immunization (donors L and M) also lacked RBD-specific IgG+ memory B cells (Fig. 2D) and had lower frequencies of S-specific IgG+ memory B cells after vaccination. Consistent with the serology data, an increase in the frequency of IgA+ (Fig. 2F) but not IgM+ spike-specific memory B cells was observed (fig. S3). Vaccination also induced S-specific CD4+ T cell responses (Fig. 2G).

Sera from 12 of 15 PIDs sampled before vaccination neutralized the Wuhan-Hu-1 SARS-CoV-2 variant (Fig. 3A and fig. S4). The nonneutralizing sera were from the three asymptomatic PIDs who had low or undetectable anti-RBD IgG titers (Fig. 3A, dashed lines, and fig. S4). Prevaccine sera from the NDs were also nonneutralizing (fig. S5). Consistent with the observed increase in binding antibodies after a single immunization in PIDs with preexisting RBD-specific IgG titers, the median half-maximal neutralizing titers [half-maximal inhibitory dilution (ID50)] were boosted ~1000-fold after the first dose, whereas the second dose had no effect (Fig. 3A). In the two PIDs lacking RBD-specific IgG titers before vaccination, the first vaccine dose elicited lower neutralizing titers (ID50 = ~30 in donor L and ~200 in donor M; Fig. 3A). In the NDs, two doses of the vaccine elicited ID50 titers that were ~10- and 5-fold lower than those elicited by one or two doses in the PIDs, respectively (Fig. 3A and fig. S6). Collectively, these data indicate that in PIDs who generate adequate immunological memory to the RBD, a single vaccine dose elicits an anamnestic response resulting in RBD-binding and nAb responses that are superior to a two-dose regimen in uninfected donors. A similar boost in binding and/or vaccine-matched neutralizing titers has been observed in PIDs who received a single mRNA vaccine dose in two recent studies (80, 81).

We next evaluated the ability of sera collected before and after immunization in NDs and PIDs to neutralize the more resistant B.1.351 and B.1.351–Δ242-243 pseudoviruses. These variants are 0.5 and 0.7% divergent from the Wuhan-Hu-1 variant. We also included SARS-CoV-1 pseudoviruses in this analysis as a representative variant that is even more dissimilar to the vaccine. SARS-CoV-1 and SARS-CoV-2 are 24, 26, and 50% divergent in the overall S protein, RBD, and RBM, respectively (82). Consequently, several mAbs that potently neutralize SARS-CoV-2 fail to bind SARS-CoV-1 (16, 22–24).

Before vaccination, 5 of 15 sera from PIDs neutralized B.1.351, and only three had ID50 titers above 100 (Fig. 3, B and E, and fig. S4); 7 of 15 neutralized B.1.351–Δ242-243, and only one had titers above 100 (Fig. 3, C and E, and fig. S4). Only two prevaccine PID sera achieved 80% neutralization of B.1.351, and only one achieved 80% neutralization of B.1.351–Δ242-243

(fig. S7A). The median ID_{50} of the prevaccine sera against the Wuhan-Hu-1 variant was significantly higher than that against B.1.351 or B.1.351–Δ242-243 (Fig. 3E). Consistent with the high level of sequence disparity, sera from only one PID showed very weak neutralizing activity toward SARS-CoV-1 before vaccinatio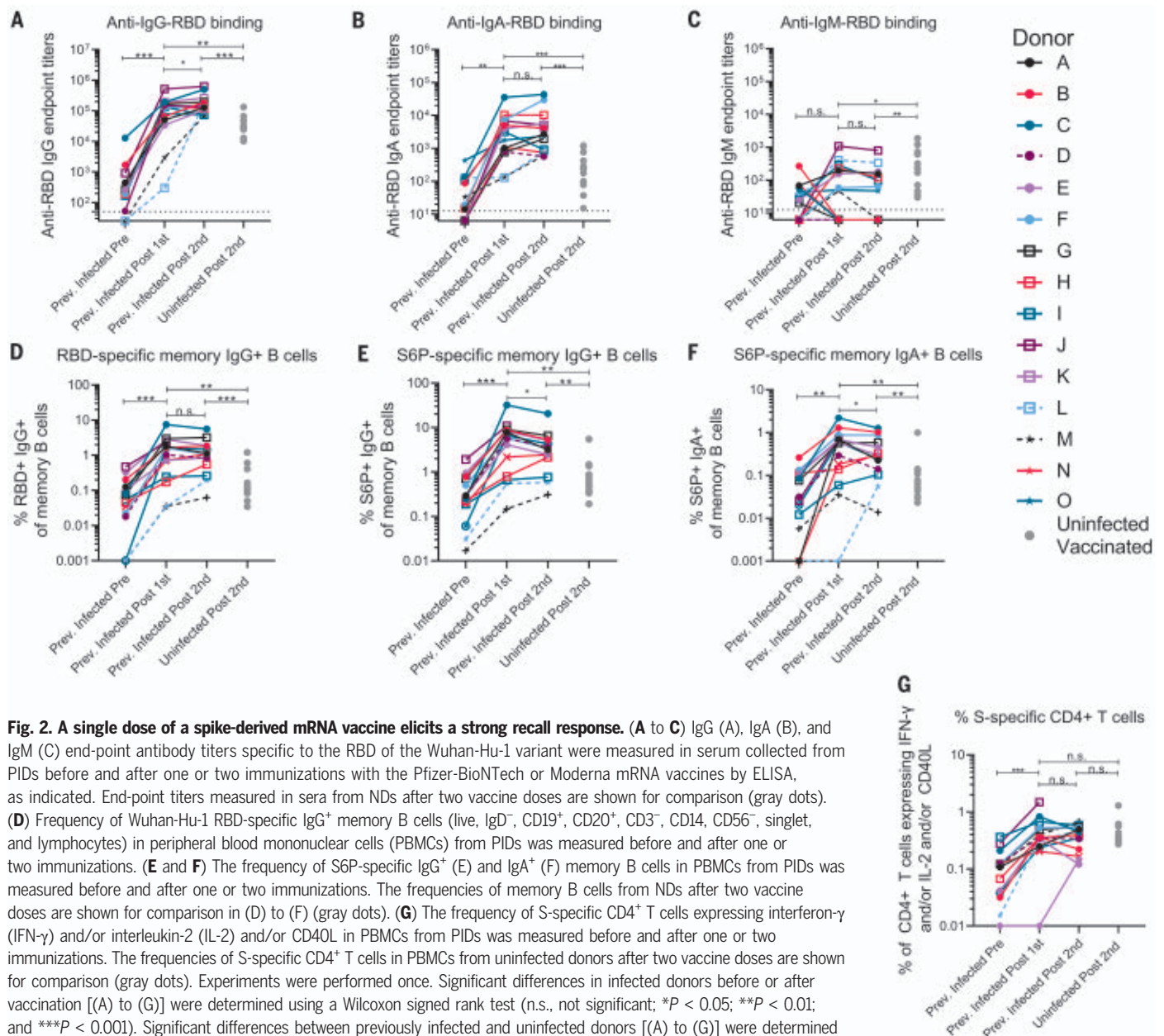n (Fig. 3, D and E, and fig. S7). A single immunization boosted the nAb titers against all three SARS-CoV-2 variants and SARS-CoV-1 in 13 of 15 PIDs (Fig. 3, A to D); however, the median ID_{50} titers were ~3-fold lower against B.1.351, ~10-fold lower against B.1.351–Δ242-243, and 100-fold lower against SARS-CoV-1 than against Wuhan-Hu-1 (Fig. 3E). A single immunization did not elicit nAbs

against the B.1.351 variants or SARS-CoV-1 in the two asymptomatic donors who lacked RBD-specific IgG memory (donor L and M; Fig. 3, A to D, and Fig. 3E, open circles). The median ID_{80} values were also lower for the B.1.351 and B.1.351–Δ242-243 variants compared with the Wuhan-Hu-1 variant (fig. S7A).

The neutralizing titers elicited by a single immunization in PIDs were significantly higher than those elicited by two immunizations in NDs against all pseudoviruses tested—10-fold higher against Wuhan-Hu-1 (Fig. 3A), 20-fold higher against B.1.351 (Fig. 3B), 30-fold higher against B.1.351–Δ242-243 (Fig. 3C), and 7-fold higher against SARS-CoV-1 (Fig. 3D). Only 8 of 13 vaccinated NDs were able to achieve 80%

neutralization of B.1.351–Δ242-243, and none could achieve 80% neutralization of SARS-CoV-1 (fig. S7B).

The B.1.351 and B.1.351–Δ242-243 variants contain three RBD mutations that affect the neutralization potency of anti-RBD mAbs (Fig. 1). Moreover, preexisting anti-RBD IgG memory appears to be important for a robust recall response to vaccination. To determine the relative contribution of anti-RBD antibodies to serum neutralization, we depleted RBD-specific antibodies from the sera of 10 PIDs after one vaccination and from nine NDs after two vaccinations. This approach efficiently removed RBD-specific (Fig. 4, A and C) but not anti-S2P–specific antibodies from sera, as measured by enzyme-linked

Fig. 2. A single dose of a spike-derived mRNA vaccine elicits a strong recall response. (**A** to **C**) IgG (A), IgA (B), and IgM (C) end-point antibody titers specific to the RBD of the Wuhan-Hu-1 variant were measured in serum collected from PIDs before and after one or two immunizations with the Pfizer-BioNTech or Moderna mRNA vaccines by ELISA, as indicated. End-point titers measured in sera from NDs after two vaccine doses are shown for comparison (gray dots). (**D**) Frequency of Wuhan-Hu-1 RBD-specific IgG$^+$ memory B cells (live, IgD$^-$, CD19$^+$, CD20$^+$, CD3$^-$, CD14, CD56$^-$, singlet, and lymphocytes) in peripheral blood mononuclear cells (PBMCs) from PIDs was measured before and after one or two immunizations. (**E** and **F**) The frequency of S6P-specific IgG$^+$ (E) and IgA$^+$ (F) memory B cells in PBMCs from PIDs was measured before and after one or two immunizations. The frequencies of memory B cells from NDs after two vaccine doses are shown for comparison in (D) to (F) (gray dots). (**G**) The frequency of S-specific CD4$^+$ T cells expressing interferon-γ (IFN-γ) and/or interleukin-2 (IL-2) and/or CD40L in PBMCs from PIDs was measured before and after one or two immunizations. The frequencies of S-specific CD4$^+$ T cells in PBMCs from uninfected donors after two vaccine doses are shown for comparison (gray dots). Experiments were performed once. Significant differences in infected donors before or after vaccination [(A) to (G)] were determined using a Wilcoxon signed rank test (n.s., not significant; *$P < 0.05$; **$P < 0.01$; and ***$P < 0.001$). Significant differences between previously infected and uninfected donors [(A) to (G)] were determined using a Wilcoxon rank sum test (*$P < 0.05$; **$P < 0.01$; and ***$P < 0.001$).

Fig. 3. Preexisting SARS-CoV-2 nAb responses are boosted by a single dose of a spike-derived mRNA vaccine. (**A** to **D**) The serum dilution resulting in 50% neutralization (ID_{50}) of Wuhan-Hu-1 (A), B.1.351 (B), B.1.351–Δ242-243 (C), and SARS-CoV-1 (D) pseudoviruses was measured in PIDs before and after one or two immunizations with the Pfizer-BioNTech or Moderna vaccines and in NDs after two vaccine doses, as indicated. Data points between PIDs who were symptomatic and asymptomatic are connected by solid and dashed lines, respectively, in (A) to (D). (**E**) Serum dilution resulting in 50% neutralization (ID_{50}) from PIDs before (squares) and after (circles) a single immunization with the Pfizer-BioNTech or Moderna vaccines against Wuhan-Hu-1, B.1.351, B.1.351–Δ242-243, and SARS-CoV-1 pseudoviruses, as indicated. PIDs who were asymptomatic and negative for anti-IgG RBD antibodies and RBD-specific IgG$^+$ memory B cells before vaccination are shown as open circles. (**F**) Neutralizing potency (ID_{50}) of serum from NDs after two immunizations with the Pfizer-BioNTech or Moderna vaccines against the indicated pseudoviruses. Each data point represents a different donor, and the horizonal bars represent the medians in (E) and (F). The dashed lines demarcate the lowest serum dilutions tested. Experiments were performed once. Significant differences in infected donors before or after vaccination, or from the same time point against different variants, were determined using a Wilcoxon signed rank test (*$P < 0.05$; **$P < 0.01$; and ***$P < 0.001$). Significant differences between previously infected and uninfected donors were determined using a Wilcoxon rank sum test (*$P < 0.05$; **$P < 0.01$; and ***$P < 0.001$).

immunosorbent assay (ELISA) (Fig. 4, B and D). This depletion abrogated serum neutralization of Wuhan-Hu-1 virus (Fig. 4, C and F), which suggests that most nAbs elicited or boosted by vaccination target this subdomain.

The above results indicate that in NDs, two doses of either the Pfizer-BioNTech or Moderna vaccines elicited nAb titers against the vaccine-matched Wuhan-Hu-1, lower titers against B.1.351, and even lower titers against B.1.351–Δ242-243. Reduced sensitivity to vaccine-elicited nAbs has been reported for other B.1.351 variants (*66*, *83*, *84*).

Similarly, sera from PIDs who experienced symptomatic SARS-CoV-2 infection and who had detectable anti-RBD IgG titers before vaccination displayed generally weak nAb titers against Wuhan-Hu-1 at 1 to 9 months after infection and lower or nonexistent titers against the B.1.351 variants, in agreement with another study (*69*). However, as long as RBD-specific IgG$^+$ memory B cell and antibody responses were generated during infection, a single immunization with either mRNA vaccine elicited a robust recall response that boosted the autologous neutralizing titers by ~1000-fold, and these antibody responses cross-neutralized the B.1.351 variants, but at lower titers. In most of the previously infected vaccinees, the anti–B.1.351–Δ242-243 neutralizing titers were comparable to those against the vaccine-matched Wuhan-Hu-1 in uninfected vaccinees. This is notable, as these titers were associated with 95% protection from COVID-19 in phase 3 trials (*44*, *46*, *48*, *49*). Moreover, vaccine-elicited antibody responses also neutralized SARS-CoV-1 but with much lower potencies. Collectively, our data suggest that the two mRNA vaccines that are based on the Wuhan-Hu-1 variant can elicit and/or boost nAb responses but that their potency is reduced against divergent variants.

Here, we show that the cross-nAb responses generated after immunization in previously infected subjects are a result of anti-RBD antibodies. Combined with the observation that the vaccines elicited nAb responses that are less potent against the B.1.351 variant with the Δ242-243 deletion in the NTD, this suggests that NTD mutations can modulate the sensitivity of emerging variants to anti-RBD nAbs. By contrast, the NTD region itself, which appears to tolerate antigenic variation in SARS-CoV-2 and other coronaviruses (*50*, *52*, *55*, *85*), does not appear to be the target of cross-nAbs elicited by infection or vaccination. We note that there are other less-frequent mutations associated with this lineage, such as L18F, Δ244, L244H, and R246I, that were not examined here, which may further increase resistance to vaccine-elicited antibodies. In this study, a pseudovirus assay was used to measure nAbs. Several studies have now shown that authentic virus and pseudovirus neutralization correlate quite well (*16*, *86*, *87*). Although the absolute

Fig. 4. Vaccine-elicited nAbs target the RBD. RBD-binding antibodies were adsorbed from sera from PIDs after receiving a single vaccine dose or from NDs after receiving two vaccine doses using Wuhan-Hu-1 RBD immobilized to magnetic beads. (**A** and **B**) Antibody binding in undepleted or RBD-depleted sera from PIDs was measured to RBD at a 1:500 dilution (A) and S2P at a 1:4500 dilution (B) by ELISA, as indicated. A_{450}, absorbance at 450 nM. (**C**) The serum dilution resulting in 50% neutralization (ID_{50}) of the Wuhan-Hu-1 pseudovirus was measured in undepleted or RBD-depleted sera from the PIDs in (A) and (B). (**D** and **E**) Antibody binding in undepleted and RBD-depleted sera from NDs was measured to RBD at a 1:500 dilution (D) and S2P at a 1:500 dilution (E) by ELISA. (**F**) The percent neutralization of a 1:120 dilution of undepleted or RBD-depleted sera from the donors in (D) and (E) was measured against the Wuhan-Hu-1 pseudovirus. Experiments were performed once.
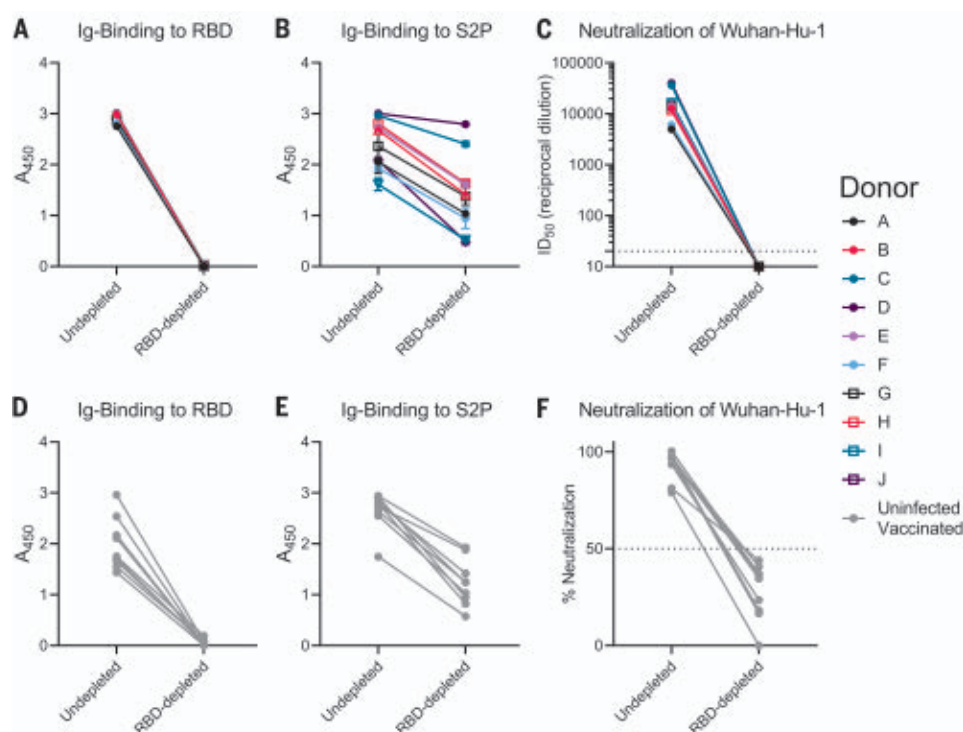
sensitivity of the authentic and pseudovirus assays may differ, we anticipate that the relative differences we report here will not vary between the two.

Although the correlates of protection for SARS-CoV-2 vaccines have not been established, studies in nonhuman primates indicate that even low titers of nAbs are sufficient to prevent experimental SARS-CoV-2 infection, particularly if CD8[+] T cell responses are mounted (*18*). Our study suggests that most previously infected subjects will benefit from a single immunization with either the Pfizer-BioNTech or Moderna vaccines, as it will lead to significant increases in serum nAb responses against vaccine-matched and emerging variants. The observation that a second dose administered 3 to 4 weeks after the first did not further boost neutralizing titers in PIDs who have clear evidence of RBD-directed immunological memory before vaccination suggests that the second dose of an mRNA vaccine could be delayed in some persons who have previously been infected with SARS-CoV-2. Longitudinal monitoring of the nAb titers before and after the first dose should be used to determine the necessity or optimal timing of the second dose in the context of previous infection.

REFERENCES AND NOTES

1. E. Dong, H. Du, L. Gardner, *Lancet Infect. Dis.* **20**, 533–534 (2020).
2. P. Zhou *et al.*, *Nature* **579**, 270–273 (2020).
3. N. Zhu *et al.*, *N. Engl. J. Med.* **382**, 727–733 (2020).
4. A. C. Walls *et al.*, *Cell* **181**, 281–292.e6 (2020).
5. M. Hoffmann *et al.*, *Cell* **181**, 271–280.e8 (2020).
6. M. Letko, A. Marzi, V. Munster, *Nat. Microbiol.* **5**, 562–569 (2020).
7. X. Ou *et al.*, *Nat. Commun.* **11**, 1620 (2020).
8. D. Wrapp *et al.*, *Science* **367**, 1260–1263 (2020).
9. A. Addetia *et al.*, *J. Clin. Microbiol.* **58**, e02107-20 (2020).
10. I. W. Pray *et al.*, *MMWR Morb. Mortal. Wkly. Rep.* **69**, 1600–1604 (2020).
11. S. F. Lumley *et al.*, *N. Engl. J. Med.* **384**, 533–540 (2021).
12. A. Chandrashekar *et al.*, *Science* **369**, 812–817 (2020).
13. W. Deng *et al.*, *Science* **369**, 818–823 (2020).
14. D. S. Stephens, M. J. McElrath, *JAMA* **324**, 1279–1281 (2020).
15. K. P. O'Callaghan, A. M. Blatz, P. A. Offit, *JAMA* **324**, 437–438 (2020).
16. T. F. Rogers *et al.*, *Science* **369**, 956–963 (2020).
17. S. J. Zost *et al.*, *Nature* **584**, 443–449 (2020).
18. K. McMahan *et al.*, *Nature* **590**, 630–634 (2021).
19. A. Baum *et al.*, *Science* **370**, 1110–1115 (2020).
20. A. Schäfer *et al.*, *J. Exp. Med.* **218**, e20201993 (2021).
21. M. S. Suthar *et al.*, *Cell Rep. Med.* **1**, 100040 (2020).
22. E. Seydoux *et al.*, *Immunity* **53**, 98–105.e5 (2020).
23. D. F. Robbiani *et al.*, *Nature* **584**, 437–442 (2020).
24. P. J. M. Brouwer *et al.*, *Science* **369**, 643–650 (2020).
25. J. M. Dan *et al.*, *Science* **371**, eabf4063 (2021).
26. L. B. Rodda *et al.*, *Cell* **184**, 169–183.e17 (2021).
27. C. Gaebler *et al.*, *Nature* **591**, 639–644 (2021).
28. J. Seow *et al.*, *Nat. Microbiol.* **5**, 1598–1607 (2020).
29. F. Muecksch *et al.*, *J. Infect. Dis.* **223**, 389–398 (2021).
30. L. Piccoli *et al.*, *Cell* **183**, 1024–1042.e21 (2020).
31. T. L. Steffen *et al.*, bioRxiv 2020.08.21.261727 [Preprint]. 22 August 2020. https://doi.org/10.1101/2020.08.21.261727.
32. C. O. Barnes *et al.*, *Cell* **182**, 828–842.e16 (2020).
33. L. Liu *et al.*, *Nature* **584**, 450–456 (2020).
34. J. Hansen *et al.*, *Science* **369**, 1010–1014 (2020).
35. B. Ju *et al.*, *Nature* **584**, 115–119 (2020).
36. Y. Cao *et al.*, *Cell* **182**, 73–84.e16 (2020).
37. A. J. Greaney *et al.*, *Cell Host Microbe* **29**, 44–57.e9 (2021).
38. X. Chi *et al.*, *Science* **369**, 650–655 (2020).
39. M. McCallum *et al.*, bioRxiv 2021.01.14.426475 [Preprint]. 14 January 2021. https://doi.org/10.1101/2021.01.14.426475.
40. G. Cerutti *et al.*, *Cell Host Microbe* **29**, 819–833.E7 (2021).
41. G. Song *et al.*, bioRxiv 2020.09.22.308965 [Preprint]. 23 September 2020. https://doi.org/10.1101/2020.09.22.308965.
42. C. Wang *et al.*, bioRxiv 2020.10.20.346916 [Preprint]. 20 October 2020. https://doi.org/10.1101/2020.10.20.346916.
43. F. Wu *et al.*, *Nature* **579**, 265–269 (2020).
44. L. R. Baden *et al.*, *N. Engl. J. Med.* **384**, 403–416 (2021).
45. K. S. Corbett *et al.*, *Nature* **586**, 567–571 (2020).
46. F. P. Polack *et al.*, *N. Engl. J. Med.* **383**, 2603–2615 (2020).
47. A. B. Vogel *et al.*, bioRxiv 2020.09.08.280818 [Preprint]. 8 September 2020. https://doi.org/10.1101/2020.09.08.280818.
48. L. A. Jackson *et al.*, *N. Engl. J. Med.* **383**, 1920–1931 (2020).
49. E. E. Walsh *et al.*, *N. Engl. J. Med.* **383**, 2439–2450 (2020).
50. A. Rambaut *et al.*, "Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations" (Virological, 2021); https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563.
51. E. Volz *et al.*, medRxiv 2020.12.30.20249034 [Preprint]. 4 January 2021. https://doi.org/10.1101/2020.12.30.20249034.
52. H. Tegally *et al.*, medRxiv 2020.12.21.20248640 [Preprint]. 22 December 2020. https://doi.org/10.1101/2020.12.21.20248640.
53. N. G. Davies *et al.*, *Science* **372**, eabg3055 (2021).
54. E. C. Sabino *et al.*, *Lancet* **397**, 452–455 (2021).
55. N. R. Faria *et al.*, "Genomic characterisation of an emergent SARS-CoV-2 lineage in Manaus: preliminary findings" (Virological, 2021); https://virological.org/t/genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-manaus-preliminary-findings/586.
56. Á. O'Toole *et al.*, "Tracking the international spread of SARS-CoV-2 lineages B.1.1.7 and B.1.351/501Y-V2" (Virological, 2021); https://virological.org/t/tracking-the-international-spread-of-sars-cov-2-lineages-b-1-1-7-and-b-1-351-501y-v2/592.
57. K. K. Chan, T. J. C. Tan, K. K. Narayanan, E. Procko, *Sci. Adv.* **7**, eabf1738 (2021).
58. T. N. Starr *et al.*, *Science* **371**, 850–854 (2021).
59. L. Zhang *et al.*, *Nat. Commun.* **11**, 6013 (2020).
60. B. Zhou *et al.*, *Nature* **592**, 122–127 (2021).
61. Z. Liu *et al.*, bioRxiv 2020.11.06.372037 [Preprint]. 11 January 2021. https://doi.org/10.1101/2020.11.06.372037.
62. A. Baum *et al.*, *Science* **369**, 1014–1018 (2020).
63. Q. Li *et al.*, *Cell* **182**, 1284–1294.e9 (2020).
64. Y. Weisblum *et al.*, *eLife* **9**, e61312 (2020).
65. E. C. Thomson *et al.*, *Cell* **184**, 1171–1187.e20 (2021).
66. P. Wang *et al.*, bioRxiv 2021.01.25.428137 [Preprint]. 12 February 2021. https://doi.org/10.1101/2021.01.25.428137.
67. C. Graham *et al.*, bioRxiv 2021.02.03.429355 [Preprint]. 3 February 2021. https://doi.org/10.1101/2021.02.03.429355.
68. X. Shen *et al.*, *Cell Host Microbe* **29**, 529–539.E3 (2021).
69. C. K. Wibmer *et al.*, *Nat. Med.* **27**, 622–625 (2021).
70. Z. Wang *et al.*, *Nature* **592**, 616–622 (2021).
71. Johnson & Johnson, "Johnson & Johnson Announces Single-Shot Janssen COVID-19 Vaccine Candidate Met Primary Endpoints in Interim Analysis of its Phase 3 ENSEMBLE Trial"

(2021); www.prnewswire.com/news-releases/johnson–johnson-announces-single-shot-janssen-covid-19-vaccine-candidate-met-primary-endpoints-in-interim-analysis-of-its-phase-3-ensemble-trial-301218035.html.
72. Novavax, Inc., "Novavax COVID-19 Vaccine Demonstrates 89.3% Efficacy in UK Phase 3 Trial" (2021); https://ir.novavax.com/news-releases/news-release-details/novavax-covid-19-vaccine-demonstrates-893-efficacy-uk-phase-3.
73. M. Yuan et al., Science 369, 1119–1123 (2020).
74. N. C. Wu et al., Cell Rep. 33, 108274 (2020).
75. Y. Wu et al., Science 368, 1274–1278 (2020).
76. R. Shi et al., Nature 584, 120–124 (2020).
77. S. Du et al., Cell 183, 1013–1023.e13 (2020).
78. N. K. Hurlburt et al., Nat. Commun. 11, 5413 (2020).
79. J. Snijder et al., Immunity 48, 799–811.e9 (2018).
80. S. Saadat et al., JAMA 325, 1467–1469 (2021).
81. F. Krammer et al., N. Engl. J. Med. 384, 1372–1374 (2021).
82. Y. Wan, J. Shang, R. Graham, R. S. Baric, F. Li, J. Virol. 94, e00127-20 (2020).
83. J. Liu et al., N. Engl. J. Med. 384, 1466–1468 (2021).
84. K. Wu et al., N. Engl. J. Med. 384, 1468–1470 (2021).
85. K. E. Kistler, T. Bedford, eLife 10, e64509 (2021).
86. D. Pinto et al., Nature 583, 290–295 (2020).
87. R. Yang et al., Biosaf. Health 2, 226–231 (2020).
88. National Research Council Canada, "HEK293SF-3F6 and HEK293-6E expression platforms (L-10894 / 11266 / 11565)" (2019); https://nrc.canada.ca/en/research-development/intellectual-property-licensing/hek293sf-3f6-hek293-6e-expression-platforms-l-10894-11266-11565.

CORONAVIRUS

Prior SARS-CoV-2 infection rescues B and T cell responses to variants after first vaccine dose

Catherine J. Reynolds[1]†, Corinna Pade[2]†, Joseph M. Gibbons[2]†, David K. Butler[1], Ashley D. Otter[3], Katia Menacho[4], Marianna Fontana[5,6], Angelique Smit[5], Jane E. Sackville-West[7], Teresa Cutino-Moguel[4], Mala K. Maini[8], Benjamin Chain[8], Mahdad Noursadeghi[8], UK COVIDsortium Immune Correlates Network‡, Tim Brooks[3], Amanda Semper[3], Charlotte Manisty[4,9], Thomas A. Treibel[4,9], James C. Moon[4,9], UK COVIDsortium Investigators‡, Ana M. Valdes[10,11], Áine McKnight[2]§, Daniel M. Altmann[12]§, Rosemary Boyton[1,13]§*

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) vaccine rollout has coincided with the spread of variants of concern. We investigated whether single-dose vaccination, with or without prior infection, confers cross-protective immunity to variants. We analyzed T and B cell responses after first-dose vaccination with the Pfizer/BioNTech messenger RNA vaccine BNT162b2 in health care workers (HCW) followed longitudinally, with or without prior Wuhan-Hu-1 SARS-CoV-2 infection. After one dose, individuals with prior infection showed enhanced T cell immunity, antibody-secreting memory B cell response to the spike protein, and neutralizing antibodies effective against variants B.1.1.7 and B.1.351. By comparison, HCW receiving one vaccine dose without prior infection showed reduced immunity against variants. B.1.1.7 and B.1.351 spike mutations resulted in increased, abrogated, or unchanged T cell responses, depending on human leukocyte antigen (HLA) polymorphisms. Single-dose vaccination with BNT162b2 in the context of prior infection with a heterologous variant substantially enhances neutralizing antibody responses against variants.

D uring worldwide rollout of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) vaccines, it is vital to understand how vaccination influences immune responses and protection among those who have had prior natural SARS-CoV-2 infection. This is a knowledge gap because a history of previous infection was an exclusion criterion in phase 3 vaccine trials (1). Countries have adopted diverse approaches—among them, the UK policy to maximize deployment of first doses to the largest possible number of people by extending the time interval to second dose. At the end of 2020, it became apparent that several virus variants had emerged (2, 3) and that these might affect vaccine rollout. The B.1.1.7 variant, possessing the spike Asn^{501}→Tyr

(N501Y) mutation, first emerged in the UK in December 2020 and spread rapidly (4). Additional variants of concern (VOC) include the B.1.351 variant, which emerged at about the same time in South Africa, and the P.1 variant, which emerged in January 2021 in Brazil. In addition to the N501Y mutation, both of these variants have the E484K mutation, which is implicated in escape from neutralizing antibodies (5, 6).

The Pfizer/BioNTech mRNA vaccine BNT162b2 encodes a prefusion-stabilized, membrane-anchored SARS-CoV-2 full-length spike protein modified by two proline substitutions (1, 7, 8). A two-dose regimen of 30 μg BNT162b2, 21 days apart, confers 95% protection against Wuhan-Hu-1 SARS-CoV-2 (1), eliciting high nAb titers as well as CD4 and CD8 cell responses (8). When given as a single 60-μg dose, BNT162b1 induced virus Ab neutralization, but T cell responses were reduced compared with the standard prime-boost regime (8). A single 30-μg dose of BNT162b1 was not reported beyond day 21. However, the cumulative incidence of COVID-19 cases among 21,676 placebo and 21,699 vaccine recipients diverged 12 days after the first dose, indicating possible early-onset first-dose protection (1). For those who were previously infected, single-dose vaccination may act as a boost after natural infection. Therefore, we aimed to test the impact of prior SARS-CoV-2 infection on T and B cell responses to first-dose vaccination.

To do this, we analyzed T and B cell immunity after the first 30-μg dose of the Pfizer/BioNTech mRNA vaccine BNT162b2 in a cohort of UK hospital health care workers (HCW)

[1]Department of Infectious Disease, Imperial College London, London, UK. [2]Blizard Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK. [3]National Infection Service, Public Health England, Porton Down, UK. [4]St Bartholomew's Hospital, Barts Health NHS Trust, London, UK. [5]Royal Free London NHS Foundation Trust, London, UK. [6]Division of Medicine, University College London, London, UK. [7]James Wigg Practice, Kentish Town, London, UK. [8]Division of Infection and Immunity, University College London, London, UK. [9]Institute of Cardiovascular Science, University College London, London, UK. [10]Academic Rheumatology, Clinical Sciences, Nottingham City Hospital, Nottingham, UK. [11]NIHR Nottingham Biomedical Research Centre, Nottingham University Hospitals NHS Trust and University of Nottingham, Nottingham, UK. [12]Department of Immunology and Inflammation, Imperial College London, London, UK. [13]Lung Division, Royal Brompton and Harefield Hospitals, London, UK.
*Corresponding author: Email: r.boyton@imperial.ac.uk
†These authors contributed equally to this work.
‡UK COVIDsortium Investigators and UK COVIDsortium Immune Correlates Network collaborators and affiliations are listed in the supplementary materials.
§These authors contributed equally to this work.

Fig. 1. Impact of prior natural infection with SARS-CoV-2 during the first wave on T and B cell responses to a single dose of the mRNA SARS-CoV-2 vaccine BNT162b2.
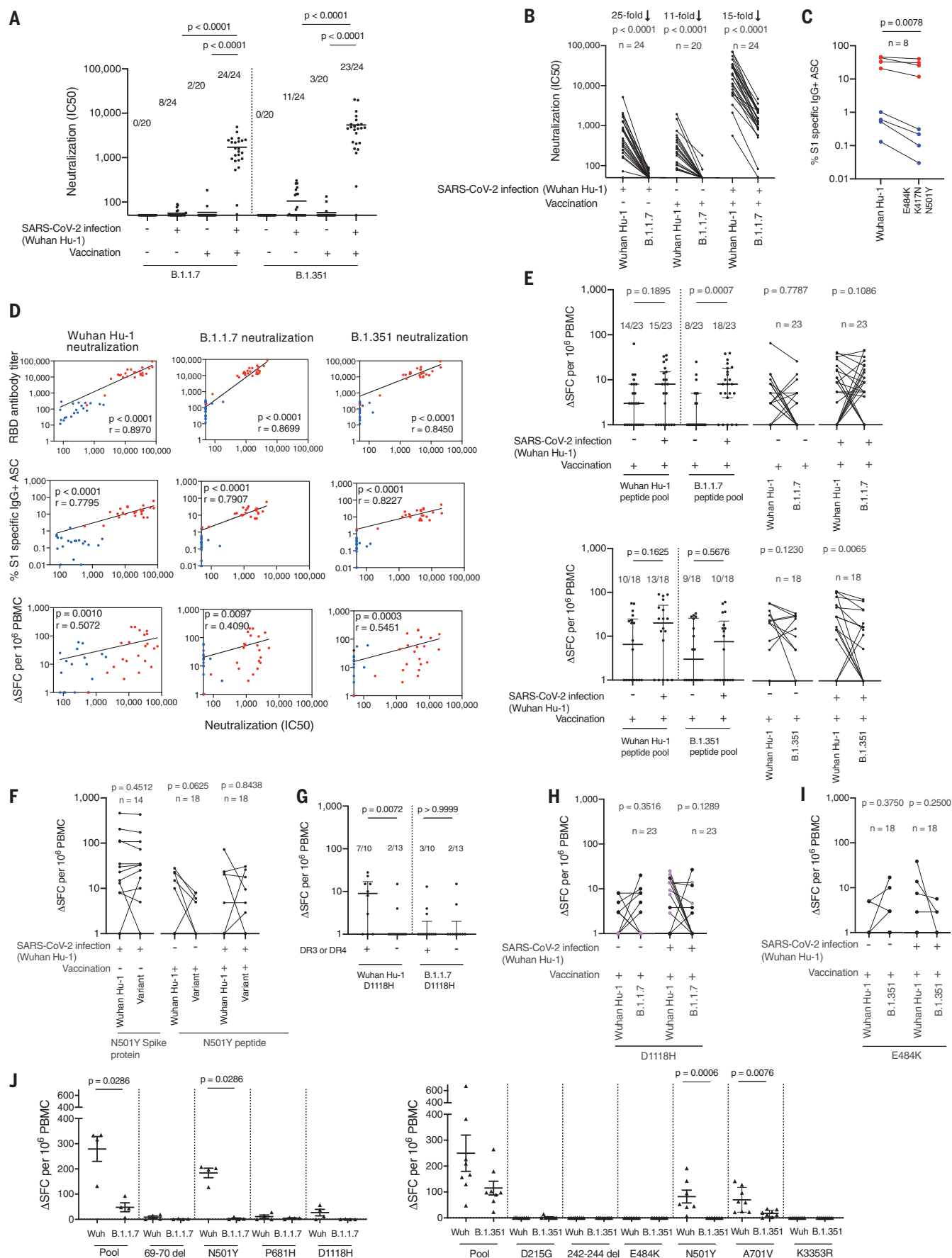(**A**) Nucleocapsid Abs measured by electrochemiluminescence immunoassay analyzer (ECLIA) in serum samples from HCW with ($n = 25$ individuals) and without ($n = 26$ individuals) laboratory-confirmed SARS-CoV-2 infection (Wuhan-Hu-1, during the first wave) 3 weeks after a single dose of the mRNA SARS-CoV-2 vaccine BNT162b2. (**B**) Magnitude of T cell response to spike protein and spike mapped epitope peptides (MEPs) in HCW with and without laboratory-confirmed SARS-CoV-2 infection ($n = 23$ per group). Data are shown prevaccination (16 to 18 weeks after infection) and 3 weeks after the first-dose vaccination (week 42) with line at geometric mean. (**C**) Proportion of HCW with ($n = 23$) and without ($n = 23$) laboratory-confirmed SARS-CoV-2 infection (during the first wave) with a T cell response to spike protein within the range of 0, 1 to 19, 20 to 79, and >80 ΔSFC/10^6 PBMC before and 3 weeks after first-dose vaccination. (**D**) Magnitude of T cell response to spike protein in HCW without a history of SARS-CoV-2 infection, plotted pairwise at 16 to 18 weeks and 42 weeks (3 weeks after first-dose vaccination). (**E**) Percentage of S1-specific IgG$^+$ antibody-secreting cells (ASCs) in vaccinated HCW with ($n = 23$) and without ($n = 22$) prior SARS-CoV-2 infection and in unvaccinated HCW with ($n = 12$) and without ($n = 5$) prior infection. Line at geometric mean. (**F**) RBD Ab titers measured by ECLIA in serum samples from HCW with ($n = 25$) and without ($n = 26$) laboratory-confirmed SARS-CoV-2 infection following first-dose vaccination. (**G**) Neutralizing antibody titer (IC$_{50}$) against Wuhan-Hu-1 authentic virus in HCW with ($n = 24$) and without ($n = 20$) laboratory-confirmed SARS-CoV-2 infection. Line at arithmetic mean. (**H**) Correlation between percentage of S1-specific ASC and magnitude of T cell response to spike protein in vaccinated HCW with ($n = 21$, red) and without ($n = 19$, blue) a history of SARS-CoV-2 infection during the first wave. (**I**) Correlation between percentage of S1-specific ASC and RBD Ab titer in HCW with ($n = 23$, red) and without ($n = 23$, blue) a history of SARS-CoV-2 infection. [(A), (B), (E), and (F)] Numbers of HCW in each group with detectable responses are shown. [(F) and (G)] Data are shown prevaccination (16 to 18 weeks after infection) and 3 weeks after the first-dose vaccination (week 42). [(A), (D), and (F)] Wilcoxon matched-pairs signed rank test. [(B), (C), (E), and (G)] Kruskal Wallis multiple comparison analysis of variance (ANOVA) with Dunn's correction. [(H) and (I)] Spearman's rank correlation. Ab, antibody; HCW, health care workers; RBD, receptor binding domain; S1, spike subunit 1; SFC, spot forming cells.

A

p = 0.0020
p = 0.0006

0/26 0/24 0/26 22/24 23/25 23/25

N relative antibody index

1,000
100
10
1

SARS-CoV-2 infection (Wuhan Hu-1): − − − + + +
Vaccination: − − + − − +
Weeks: 16-18 28-30 42 16-18 28-30 42

B

p = 0.0440
p = 0.0021 p = 0.0557

p = 0.0070
p = 0.0014 p = 0.0010

3/23 14/23 16/23 22/23   11/23 19/23 14/23 22/23

ΔSFC per 10^6 PBMC

1,000
100
10
1

SARS-CoV-2 infection (Wuhan Hu-1): − + − +   − + − +
Vaccination: − − + +   − − + +
Spike protein   Spike MEP

C

■ 80+ (ΔSFC)
▨ 20-79 (ΔSFC)
▧ 1-19 (ΔSFC)
□ 0 (ΔSFC)

p = 0.0440   p = 0.0557

Spike protein

SARS-CoV-2 infection (Wuhan Hu-1): − + − +
Vaccination: − − + +

D

p = 0.0089

ΔSFC per 10^6 PBMC

1000
100
10
1

SARS-CoV-2 infection (Wuhan Hu-1): − −
Vaccination: − +

E

p < 0.0001
p = 0.0208

p = 0.0381

0/5 10/12 20/22 23/23

% S1 specific IgG+ ASC

100
10
1
0.1
0.01

SARS-CoV-2 infection (Wuhan Hu-1): − + − +
Vaccination: − − + +

F

p < 0.0001   p = 0.0043
p < 0.0001

0/26 0/26 26/26 22/24 22/24 25/25

RBD antibody titer

10,000
1,000
100
10
1

SARS-CoV-2 infection (Wuhan Hu-1): − − − + + +
Vaccination: − − + − − +
Weeks: 16-18 28-30 42 16-18 28-30 42

G

p < 0.0001
p = 0.0002
p = 0.0046
p < 0.0001

2/20 22/24 20/20 24/24

Neutralization (IC50)

100,000
10,000
1,000
100

SARS-CoV-2 infection (Wuhan Hu-1): − + − +
Vaccination: − − + +

H

% S1 specific IgG+ ASC

100
10
1
0.1
0.01

p = 0.0261
r = 0.4842

p = 0.6267
r = 0.1193

1 10 100 1,000
ΔSFC per 10^6 PBMC

I

% S1 specific IgG+ ASC

100
10
1
0.1
0.01

p = 0.0008
r = 0.6502

p = 0.5479
r = 0.1321

10 100 1,000 10,000 100,000
RBD antibody titer

Fig. 2. Impact of vaccination and prior natural infection with SARS-CoV-2 during the first wave on T and B cell responses to the UK B.1.1.7 and South African B.1.351 variants. (**A**) Neutralizing antibody (nAb) titer (IC_{50}) against B.1.1.7 and B.1.351 authentic virus in HCW with ($n = 24$) and without ($n = 20$) laboratory-confirmed SARS-CoV-2 infection (Wuhan-Hu-1). Lines at arithmetic mean. Data are shown prevaccination (16 to 18 weeks after infection) and 3 weeks after the first-dose vaccination (week 42). (**B**) nAb (IC_{50}) titers against Wuhan-Hu-1 and B.1.1.7 authentic viruses plotted pairwise by individual. (**C**) Percentage of Wuhan-Hu-1 S1 and S1 containing variant mutations (E484K, K417N, and N501Y) specific IgG$^+$ antibody-secreting cells (ASCs) in vaccinated HCW with ($n = 4$) and without ($n = 4$) prior SARS-CoV-2 infection. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr. (**D**) Correlations between nAb (IC_{50}) titers of Wuhan-Hu-1, B.1.1.7, or B.1.351 authentic virus and RBD Ab titer, percentage of S1-specific ASC, and magnitude of T cell response to S1 protein in vaccinated HCW with ($n = 22$ to 24, red) and without ($n = 18$ to 20, blue) a history of SARS-CoV-2 infection. (**E**) Magnitude of T cell response to Wuhan-Hu-1, B.1.1.7, or B.1.351 peptide pools in vaccinated HCW with ($n = 23$ or 18) and without ($n = 23$ or 18) SARS-CoV-2 infection (Wuhan-Hu-1), plotted as grouped data (median plus interquartile range) and pairwise for each individual. (**F**) Magnitude of T cell response to Wuhan-Hu-1 S1 protein and N501Y variant spike RBD protein in unvaccinated HCW with laboratory-confirmed SARS-CoV-2 infection ($n = 14$) or to Wuhan-Hu-1 and N501Y mutated peptide in vaccinated HCW with ($n = 18$) and without ($n = 18$) a history of SARS-CoV-2 infection, plotted pairwise by individual. (**G**) Magnitude of T cell response to Wuhan-Hu-1 or B.1.1.7 D1118H peptide in vaccinated HCW with a history of SARS-CoV-2 infection ($n = 23$), plotted by DRB1*0301 or DRB1*0401 status. Lines at median plus interquartile range. (**H**) Magnitude of T cell response to Wuhan-Hu-1 or B.1.1.7 D1118H peptide in vaccinated HCW with ($n = 23$) and without ($n = 23$) a history of SARS-CoV-2 infection, plotted pairwise by individual and with individuals carrying DRB1*0301 or DRB1*0401 alleles marked in purple. (**I**) Magnitude of T cell response to Wuhan-Hu-1 or B.1.351 E484K mutated peptide in vaccinated HCW with ($n = 18$) and without ($n = 18$) a history of SARS-CoV-2 infection, plotted pairwise by individual. (**J**) Magnitude of T cell response to Wuhan-Hu-1 (Wuh), B.1.1.7, or B.1.351 peptide pools and individual peptides in Wuhan-Hu-1 peptide immunized HLA-DRB1*04:01 transgenic mice (left-hand panel, $n = 4$; right-hand panel, $n = 8$; lines at arithmetic mean + SEM). (A) Kruskal Wallis multiple comparison ANOVA with Dunn's correction. [(B), (C), (E) (right-hand panels), (F), (H), and (I)] Wilcoxon matched-pairs signed rank test. (D) Spearman's rank correlation. [(E) (left-hand panels), (G), and (J)] Mann-Whitney U test. ASC, antibody-secreting cells; HCW, health care workers; RBD, receptor binding domain; S1, spike subunit 1; SFC, spot forming cells.

(*9–12*). The COVIDsortium HCW cohort has been studied longitudinally since the end of March 2020, providing accurate infection and immune history in the context of genotyping, including human leukocyte antigen (HLA) imputation (*10–12*). Our aim was to compare T and B cell immunity after a first dose of vaccine in December 2020 in postinfection (after natural infection), vaccinated postinfection (vaccination in the context of prior SARS-CoV-2 infection), and vaccinated naïve (single-dose vaccination) individuals. We sought to explore whether there is evidence for altered T cell recognition of the B.1.1.7 and B.1.351 variants and, in particular, of the N501Y mutation shared by several VOC.

The UK has deployed a heterodox vaccination regimen to maximize immune protection and slow spread of the B.1.1.7 lineage, giving an initial 30-μg dose of BNT162b2 followed by boosting up to 12 weeks later (*13*). A cross-sectional substudy ($n = 51$ individuals) of the existing longitudinal HCW cohort (*9–12*) was recruited 22 (±2) days after the first dose. After the start of the study, the majority of acute infections had already occurred among this cohort (*11*). At the time of receiving their first vaccine dose in December 2020, prior to the emergence of VOC, 25 individuals were ~39 weeks removed from SARS-CoV-2 infection with the Wuhan-Hu-1 strain, and 26 were confirmed uninfected, having tested negative in longitudinal serology for spike and nucleocapsid (N) proteins (table S1 and fig. S1).

We first measured SARS-CoV-2 N antibody longitudinally up to 16 to 18 weeks, then at 28 to 30 weeks, and finally at 42 weeks after recruitment, to confirm that there was no laboratory evidence of new infection at the time of drawing blood for the vaccine study at 42 weeks; none of the previously uninfected HCW had become seropositive (Fig. 1A). T cell responses to spike protein and mapped epitope peptides (MEPs) in either postinfection, vaccinated postinfection, and vaccinated naïve individuals were compared (Fig. 1B). Ninety-six percent (22/23) of vaccinated postinfection individuals mounted a T cell response to spike protein compared with 70% (16/23) of vaccinated naïve individuals, with a fourfold increase in the magnitude of the T cell response. Furthermore, while the T cell response to spike protein in vaccinated naïve individuals increased ($P = 0.0440$), it was lower than that of vaccinated postinfection individuals ($P = 0.0557$) (Fig. 1C). As expected, there was no significant change in T cell response to N (a measure of immunity to natural infection) (fig. S2A).

Paired analysis of T cell immunity to spike protein in previously uninfected individuals, analyzed at the 16- to 18-week time point and 3 weeks after vaccination, showed a significantly increased response ($P = 0.0089$) (Fig. 1D). Three individuals who previously showed a response, despite lack of laboratory evidence for infection (therefore presumably a cross-reactive response to an endemic human coronavirus), showed an unchanged or decreased response to spike after vaccination.

The size of the SARS-CoV-2 spike subunit 1 (S1)–specific memory B cell (MBC) pool was investigated by B cell enzyme-linked immunosorbent spot (ELISpot) assay (Fig. 1E and fig. S2B). As for T cell responses, the number of S1-specific immunoglobulin G (IgG$^+$) antibody-secreting cells (ASCs) was far greater in vaccinated postinfection individuals than in vaccinated naïve individuals ($P < 0.0001$). Prior infection generated a 63-fold increase in S1-specific ASCs. There were no preexisting S1-specific ASCs in uninfected HCW before vaccination. Twenty of 22 vaccinated naïve individuals had detectable S1-specific ASCs composing 0.02 to 1.54% of the MBC pool. By comparison, all vaccinated postinfection individuals had detectable S1-specific ASCs (1.90 to 50% of the MBC pool). We previously reported (*14*) spike receptor binding domain (RBD) enhanced Ab responses in the vaccinated postinfection group. In this work, the vaccinated naïve group attained antibody titers similar to those of the postinfection group at 16 to 18 weeks and 28 to 30 weeks (Fig. 1F). Vaccinated naïve individuals demonstrated a lower nAb response to wild-type virus than was seen after natural infection at 16 to 18 weeks, although this did not achieve statistical significance. In line with the findings for MBC and RBD binding, there was a significantly enhanced nAb response in vaccinated postinfection individuals compared with the vaccinated naïve group (Fig. 1G), with a mean value of 25,273 compared with 420, that is, a 60-fold increase. To put this in context, these values are 43-fold higher than the values recorded after two vaccine doses in the phase 1 trial (*7*). There was no correlation between the magnitude of the spike protein T cell response and the percentage of S1-specific ASCs (Fig. 1H). As expected, there was a positive correlation between the percentage of S1-specific ASCs and the serum titer of RBD antibody in the vaccinated postinfection individuals [correlation coefficient (r) = 0.6502; $P = 0.0008$] (Fig. 1I). After vaccination, two previously infected individuals showed lower percentages of S1-specific memory B cells

and reduced serum RBD-specific antibody levels than the rest of the group; prior infection involving case-definition symptoms tended to be associated with a higher specific B cell frequency than milder disease (Fig. 1F and fig. S2C). These individuals who, despite infection, had also not shown a detectable T cell response (one never seroconverted, and the other rapidly became seronegative during longitudinal follow-up) had a poor or absent response to infection that was only minimally overcome by vaccination.

The data in Fig. 1 indicate that there is a strong prime-boosting effect of prior infection on single-dose vaccination. Augmentation is seen more strongly in MBC frequency, anti-RBD, and nAb responses than for T cell response frequency. Furthermore, there was no correlation between S1 ASC frequency and T cell response frequency (Fig. 1H). There is, however, a correlation between S1 ASC and RBD antibody titers, indicating that individuals with higher numbers of MBCs mount stronger antibody responses, and individuals who had experienced infection clustered at the higher end of this response (Fig. 1I).

Shortly before the vaccination program was initiated, several VOC emerged, including B.1.1.7. This variant has nine mutations in the spike protein. Several studies have reported weaker nAb responses to B.1.1.7 relative to the previously circulating Wuhan-Hu-1 strain (2–6, 15–18). The majority of SARS-CoV-2 immune naïve individuals made no nAb response to the B.1.1.7 (18/20) and B.1.351 (17/20) variants after single-dose vaccination. In contrast, almost all vaccinated postinfection individuals made a strong nAb response to the B.1.1.7 (24/24) and B.1.351 (23/24) variants after a single-dose vaccination, with a 46-fold (B.1.1.7) and 63-fold (B.1.351) increase in mean nAb half-maximal inhibitory concentration (IC_{50}) in vaccinated postinfection individuals compared with vaccinated naïve individuals. In a paired analysis, we observed in vitro significantly reduced nAb potency to authentic B.1.1.7 variant (mean: 35) with a 96% fall compared to that of Wuhan-Hu-1 (mean: 866; $P < 0.0001$) in sera from individuals with a past medical history of natural infection (Fig. 2B). Worryingly, after single-dose vaccination, 90% (18/20) of vaccinated naïve individuals showed no detectable nAbs ($IC_{50} < 50$) against B.1.1.7 (mean IC_{50}: 37; range: 0 to 184; $P = 0.2090$), but they did show demonstrable nAb responses to Wuhan-Hu-1 SARS-CoV-2 virus (mean IC_{50}: 420; range: 80 to 2004; $P = 0.0046$). In contrast, all vaccinated postinfection individuals responded to single-dose vaccination with substantially enhanced nAb responses, neutralizing not just Wuhan-Hu-1 SARS-CoV-2 (mean IC_{50}: 25,273; range: 581 to 76,369) but also the B.1.1.7 (mean IC_{50}: 1717; range: 52 to 4919) and B.1.351 (mean IC_{50}: 5451; range: 41 to 20,411) variants (Fig. 2, A and B,

and fig. S3). We show a 93% reduction in neutralization (IC_{50}) responses to the SARS-CoV-2 B.1.1.7 variant (mean: 1717) compared with the Wuhan-Hu-1 (mean: 25,273) virus in vaccinated postinfection individuals. However, despite this fall, the majority (22/24) remain within a "protective threshold." This was not the case for vaccinated naïve individuals. There was a 91% reduction in neutralization (IC_{50}) responses against the SARS-CoV-2 B.1.1.7 variant (mean: 37) compared with the Wuhan-Hu-1 virus (mean: 420), resulting in the majority of individuals (19/20) falling below the "protective threshold." This result was mirrored in the SARS-CoV-2 S1-specific MBC pool, where reduced numbers of S1-specific IgG+ ASC are seen (in vaccinated naïve individuals compared with vaccinated postinfection individuals) responding to S1 antigen containing the N501Y, K417N, and E484K mutations. Prior infection substantially enhances the specific MBC pool after single-dose vaccination (Fig. 2C). We looked at correlations between RBD binding antibodies, B cell responses, T cell responses, and IC_{50}, comparing neutralization of Wuhan-Hu-1, B.1.1.7, and B.1.351 live virus (Fig. 2D). Despite the lower neutralization of B.1.1.7 and B.1.351 variants, the pattern was retained of strong correlation between RBD antibody titer and S1-specific B cell frequency and neutralization and somewhat weaker correlation between T cell response and neutralization.

A lack of Ab-mediated protection in single-dose vaccinees could be mitigated by a broader repertoire of T cell responses (18). To investigate differences in T cell recognition, we designed peptide pools covering the affected regions of Wuhan-Hu-1, B.1.1.7, and B.1.351 variant sequence (table S2). We compared T cell responses to these peptide pools in peripheral blood mononuclear cells (PBMCs) from vaccinated postinfection and vaccinated naïve individuals (Fig. 2E). Responses in postinfection vaccinees were in general higher than in the vaccinated naïve individuals (note an enhanced response to the B.1.1.7 peptide pool). T cell responses were heterogeneous; responses to variant pools could be either higher or lower than to Wuhan-Hu-1 pools. Alterations in affinity for the T cell receptor can lead to altered peptide ligand effects and differential polarization of cytokine effector programs, as we have previously observed in Zika virus infection (19). We wondered whether this was also occurring for SARS-CoV-2; however, we found no evidence for immune deviation to interleukin (IL)–4, IL-5, IL-10, IL-13, IL-17A, or IL-23 (fig. S3).

For B.1.1.7 and B.1.351, attention has centered on the N501Y mutation, as this is implicated in altered angiotensin-converting enzyme 2 (ACE2) binding and enhanced infectivity and transmission but is also a target for B and T cell recognition. We initially looked at T cell

responses after natural infection and found that at 16 to 18 weeks postinfection, the N501Y mutation appeared to have no substantial differential impact on the T cell response (Fig. 2F), unlike nAb recognition (5).

The specific impact of any T cell epitope changes on the immune response against VOC depends on changes in peptide binding to the peptide-presenting HLA molecules. Because the HLA complex is the most polymorphic part of the human genome, any alteration to core HLA binding motifs will differentially affect people with certain HLA alleles over others. We performed in silico analysis (using NetMHCIIpan) to predict which of the B.1.1.7 and B.1.351 mutations were found in HLA core binding motifs and how this might affect binding to common HLAII alleles (DRB1*0101, DRB1*0301, DRB1*0401, DRB1*0701, DRB1*1101, DRB1*1301, and DRB1*1501) (tables S3 and S4). Some of the mutations did not fall in a region predicted to bind the HLAII alleles tested (D3L, T716I, T1001I, A1708D, and 3675-7 SGF del). Although several mutations were not predicted to significantly change affinity for the HLAII alleles, others did show predicted differential affinities depending on host HLAII type (tables S3 and S4). Analyzing altered responses to the D1118H mutation, we noted that individuals who carried DRB1*0301 and DRB1*0401 showed enhanced T cell responses to the Wuhan-Hu-1 peptide compared with those who did not ($P = 0.0072$) (Fig. 2G). T cell responses to the variant peptide appeared to be reduced in individuals carrying DRB1*0301 and DRB1*0401 (Fig. 2H). There is a basis for this in terms of differential HLAII binding as the D-to-H mutation is predicted to lose the T cell epitope for people carrying DRB1*0301 and DRB1*0401 but not, for example, in those who carry DRB1*0701 or DRB1*1501, who would be predicted to show an enhanced response (table S3). People carrying DRB1*1301 are predicted to gain a response as a consequence of this mutation. Analyzing responses to the E484K mutation seen in B.1.351 and P.1 variants, we noted that it did not fall in a region predicted to bind the HLAII alleles tested (table S4). The mutation appeared to have no substantial or differential impact on T cell responses (Fig. 2I).

When we primed transgenic mice expressing human HLA-DRB1*0401 with the Wuhan-Hu-1 peptide pool, T cell responses to the B.1.1.7 variant peptide pool were significantly reduced ($P = 0.0286$) (Fig. 2J). Furthermore, the T cell response to the spike N501Y mutation common to all three of the current VOC was ablated.

In this HCW cohort, vaccinated naïve individuals made an anti-S1 RBD Ab response with a mean titer of ~100 U/ml at 22 (±2) days after vaccination, roughly equivalent to the mean peak Ab response after natural infection (14). However, the spike T cell response after one

dose was lower than after natural infection, and for 30% of vaccinees, no response could be measured. However, T cell responses are enhanced fourfold in those vaccinated postinfection. This T cell enhancement is small relative to the 63-fold change in ASCs and the corresponding 140-fold change in Roche anti-S (RBD) Ab levels we observed after one vaccine dose in HCW vaccinated postinfection (14). While much has been written about the impact of rapidly waning serum antibodies, our findings confirm that MBCs are nevertheless primed and able to contribute a rapid, large response to repeat exposure. The rather large effect on B cell priming and restimulation, relative to T cells, in previously infected single dose–vaccinated individuals may reflect the fact that, among the nuanced differences between the licensed SARS-CoV-2 vaccines, aspects of the mRNA adjuvant effect appear to skew immunity to high nAb titers, which may underpin its high efficacy. Our evidence for enhanced vaccine responses after infection supports the case that only one vaccine dose is necessary to maximize immune protection for SARS-CoV-2–experienced individuals (14, 20).

It is notable that the high IC_{50} titers in those vaccinated after infection provide such a large protective margin that responses to authentic B.1.1.7 and B.1.351 variants are also high. In contrast, nAb responses in individuals several months on from mild infection show much lower IC_{50} titers against B.1.1.7 and B.1.351, often less than 100. Similarly, the majority of responses in naïve individuals after one dose show weak recognition of B.1.1.7 and B.1.351. This finding indicates potentially poor protection against B.1.1.7 and B.1.351 in individuals who have experienced natural infection or who have only had one vaccine dose.

It is important to map the effect of VOC mutations on any evasion of T cell immunity. The case has been made that reductions in antibody neutralization of mutant spike may be mitigated by protective T cells (8). A case has been made for the role of T cells as correlates of protection (21). Our evidence from this analysis 22 (±2) days after one dose is that T cell immunity is mostly variably low but also relatively unperturbed by the N501Y mutation. The other mutations we considered that overlay CD4 epitopes were, as might be predicted, distributed across the range of HLAII polymorphisms. Those alleles associated with loss of CD4 response to the variant pool tended to be those with a lysine in pocket 4 of the groove (HLA-DR residue 71β), whereas those with an increased response to the variant pool tended to be those with a smaller amino acid, alanine. In HLA-DRB1*0401 transgenics, we confirmed that in the context of a given HLAII heterodimer, the N501Y mutation can result in ablation of this part of the T cell response,

demonstrating that HLA polymorphisms are likely to be significant determinants of responder and nonresponder status with respect to vaccine escape.

SARS-CoV-2 immunity now encompasses postinfection plus either zero, one, or two vaccine doses and first and second dose naïve vaccinated. Single-dose vaccination after infection achieves similar levels of S1 RBD binding antibodies to two doses in naïve vaccinated individuals and second-dose vaccination in one-dose vaccinated postinfection individuals offers no additional enhancement (22). Moving forward, it will be important to resolve the quantitative and qualitative differences between these groups in terms of neutralizing antibody repertoire as well as phenotype and durability of memory B and T cell responses. Durability of immunity to natural infection and after vaccination as well as sustained vaccine efficacy and vaccine escape need to be monitored over time.

REFERENCES AND NOTES

1. F. P. Polack et al., N. Engl. J. Med. 383, 2603–2615 (2020).
2. J. R. Mascola, B. S. Graham, A. S. Fauci, JAMA 325, 1261–1262 (2021).
3. D. M. Altmann, R. J. Boyton, R. Beale, Science 371, 1103–1104 (2021).
4. A. Rambaut et al., Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations, virological.org (2020); https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563.
5. Z. Wang et al., Nature 592, 616–622 (2021).
6. D. A. Collier et al., Nature 593, 136–141 (2021).
7. E. E. Walsh et al., N. Engl. J. Med. 383, 2439–2450 (2020).
8. U. Sahin et al., Nature 586, 594–599 (2020).
9. T. A. Treibel et al., Lancet 395, 1608–1610 (2020).
10. J. B. Augusto et al., Wellcome Open Res. 5, 179 (2020).
11. C. J. Reynolds et al., Sci. Immunol. 5, eabf3698 (2020).
12. C. Manisty et al., EBioMedicine 65, 103259 (2021).
13. Nat. Med. 27, 183 (2021).
14. C. Manisty et al., Lancet 397, 1057–1058 (2021).
15. X. Shen et al., Cell Host Microbe 29, 529–539.e3 (2021).
16. A. J. Greaney et al., Cell Host Microbe 29, 463–476.e6 (2021).
17. P. Supasa et al., Cell 184, 2201–2211.e7 (2021).
18. A. Muik et al., Science 371, 1152–1153 (2021).
19. C. J. Reynolds et al., Sci. Rep. 8, 672 (2018).
20. F. Krammer et al., N. Engl. J. Med. 384, 1372–1374 (2021).
21. A. Bertoletti, A. T. Tan, N. Le Bert, Oxford Open Immunology 2, iqab006 (2021).
22. R. R. Goel et al., Sci. Immunol. 6, eabi6950 (2021).

PALEOANTHROPOLOGY

A Middle Pleistocene *Homo* from Nesher Ramla, Israel

Israel Hershkovitz[1,2]*†, Hila May[1,2]*†, Rachel Sarig[2,3]*†, Ariel Pokhojaev[1,2,3], Dominique Grimaud-Hervé[4], Emiliano Bruner[5], Cinzia Fornai[6,7], Rolf Quam[8,9,10], Juan Luis Arsuaga[9,11], Viktoria A. Krenn[6,7], Maria Martinón-Torres[5,12], José María Bermúdez de Castro[5,12], Laura Martín-Francés[5,12], Viviane Slon[1,2,13], Lou Albessard-Ball[4,14], Amélie Vialet[4], Tim Schüler[15], Giorgio Manzi[16], Antonio Profico[14,16], Fabio Di Vincenzo[16]‡, Gerhard W. Weber[7,17]†, Yossi Zaidner[18,19]†

It has long been believed that Neanderthals originated and flourished on the European continent. However, recent morphological and genetic studies have suggested that they may have received a genetic contribution from a yet unknown non-European group. Here we report on the recent discovery of archaic *Homo* fossils from the site of Nesher Ramla, Israel, which we dated to 140,000 to 120,000 years ago. Comprehensive qualitative and quantitative analyses of the parietal bones, mandible, and lower second molar revealed that this *Homo* group presents a distinctive combination of Neanderthal and archaic features. We suggest that these specimens represent the late survivors of a Levantine Middle Pleistocene paleodeme that was most likely involved in the evolution of the Middle Pleistocene *Homo* in Europe and East Asia.

Recent dental (*1*), mandibular (*2*), genetic (*3*, *4*), and demographic (*5*) studies have predicted the existence of an as yet unidentified African or West Asian Middle Pleistocene (MP) population that contributed to the evolution of the Neanderthal clade. This contrasts with the traditional view that considered the European continent as the sole place of origin of the Neanderthals and their direct ancestors.

Here we report on the discovery of several fossils from the recently excavated MP open-air site of Nesher Ramla (NR), central Israel (Fig. 1), in association with stone artifacts, and faunal remains (*6*).

A nearly complete right parietal bone and four fragments from the left parietal bone represent the NR-1 fossil (Fig. 2A and fig. S1). The NR-2 fossil is an almost complete mandible, missing only the left ramus, the right condylar process, and the mandibular angle of the right ramus (Fig. 3). The lower left second molar (NR-2 M_2) and most of the dental roots are still in place (fig. S2). Both NR-1 and NR-2 were found in situ within the lowest archaeological layer (Fig. 1D, Unit VI), together with animal bones and flint tools, and most likely represent the same individual (supplementary text A).

Unit VI is assigned an age of 140 to 120 thousand years (ka) ago, based on the electron spin resonance–uranium series (ESR-US) dates of animal teeth recovered in this unit (with a weighted mean of 125.8 ± 5.9 ka). This age was corroborated by a series of thermoluminescence (TL) dates of burnt flints from the archaeological layer immediately above the fossil (Unit V). This layer yielded a weighted mean of 127.6 ± 4.0 ka (confirmed by isochron analysis), which is in agreement with the ESR-US dates obtained for this unit, ranging between ~128 and ~120 ka (a weighted mean of 122.3 ± 3.3 ka). This chronological information is consistent with the previously published optically stimulated luminescence (OSL) dates for the entire archaeological sequence [ranging ~170 and ~78 ka; (*6*)].

The preserved anatomical elements were thoroughly described and analyzed in comparison to a large number of fossils of different periods (table S1), using a combination of traditional approaches based on linear and angular measurements, as well as three-dimensional (3D) landmark-based geometric morphometric (GM) methods (supplementary text C to E).

The overall morphology of the NR-1 parietal bones (supplementary text C and tables S2 and S3) is indicative of an archaic, low cranial vault, which is typical of MP *Homo* specimens and is substantially different from early and recent *H. sapiens*, which instead manifest a curved parietal bone with a pronounced eminence (*7*, *8*).

Further support for the rather archaic morphology of the NR *Homo* comes from the angle formed by the coronal and sagittal sutures (c/s angle), 91.1° in the NR-1 specimen. This angle increased during the evolution of Pleistocene *Homo* (fig. S3): *H. erectus* and African MP *Homo* exhibit a mean angle of 92.1° ± 2.1°; the angle opens to 94.9° ± 3.4° in European MP *Homo*/Neanderthals and reaches 99.4° ± 4.2° in early and recent *H. sapiens*. The c/s angle is significantly different between these three groups (*H* = 22.5, *p* < 0.001). The c/s angle for NR-1 is similar to that of archaic *Homo*, particularly African MP *Homo* (91.1° ± 1.1°), and falls outside the range of variation of *H. sapiens*.

The NR-1 parietal bone is considerably thick, mainly in the parietal eminence area (figs. S4 and S5). Regarding this aspect, the NR-1 parietal is similar to that of European MP *Homo* specimens (e.g., Petralona, Atapuerca SH) (fig. S4). It is generally thicker than the parietal of Neanderthals (e.g., Amud 1, Guattari, and La Chapelle-aux-Saints) and most early *H. sapiens* (except for Laetoli H18 and Omo 2), and it is much thicker than that of recent *H. sapiens*.

The 3D GM analysis, used to assess NR-1 shape variation with respect to a comparative sample of Pleistocene and recent *Homo* (supplementary text, Fig. 2C, fig. S1, and table S1), confirms the archaic morphology of NR-1. The first three principal components (PCs) explain 74.5% of the total shape variance. The first PC (34.9%) differentiates early and recent *H. sapiens* from all other groups, including Asian *H. erectus*, European and African MP *Homo*, and Neanderthals, owing to their marked curvature along both the sagittal and the coronal planes (Fig. 2C). The second PC (21.3%) is not taxonomically informative (fig. S1). The third PC (18.3%) separates Asian *H. erectus* and African MP *Homo* from Neanderthals and European MP *Homo* (Fig. 2C), based on the relative development of the parietal eminence and its relative antero-posterior position. The European MP group is characterized by an antero-posteriorly and supero-inferiorly flatter parietal bone (Fig. 2C). NR-1 is distinct from *H. sapiens*; it is at an intermediate position between the Neanderthal and MP *Homo* clusters (Fig. 2C). An unrooted phylogenetic analysis, based on the mean shape of each *Homo* group, placed NR-1 close to the origin of the branch

[1]Department of Anatomy and Anthropology, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel. [2]The Shmunis Family Anthropology Institute, the Dan David Center for Human Evolution and Biohistory Research, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel. [3]Department of Oral Biology, the Goldschleger School of Dental Medicine, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel. [4]UMR7194, HNHP, Département Homme et Environnement, Muséum national d'Histoire naturelle, CNRS, UPVD, Paris, France. [5]CENIEH (National Research Center on Human Evolution), Burgos, Spain. [6]Institute of Evolutionary Medicine, University of Zurich, Zurich, Switzerland. [7]Department of Evolutionary Anthropology, University of Vienna, Vienna, Austria. [8]Department of Anthropology, Binghamton University (SUNY), Binghamton, NY, USA. [9]Centro UCM-ISCIII de Evolución y Comportamiento Humanos, Madrid, Spain. [10]Division of Anthropology, American Museum of Natural History, New York, NY, USA. [11]Departamento de Geodinámica, Estratigrafía y Paleontología, Facultad de Ciencias Geológicas, Universidad Complutense de Madrid, Ciudad Universitaria s/n, 28040, Madrid, Spain. [12]Department of Anthropology, University College London, London, UK. [13]Department of Human Molecular Genetics and Biochemistry, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel. [14]PalaeoHub, Department of Archaeology, University of York, York, UK. [15]Thuringian State Office for the Preservation of Historical Monuments and Archaeology Weimar, Germany. [16]Department of Environmental Biology, Sapienza University of Rome, Roma, Italy. [17]Core Facility for Micro-Computed Tomography, University of Vienna, Vienna, Austria. [18]Institute of Archaeology, The Hebrew University of Jerusalem, Jerusalem, Israel. [19]Zinman Institute of Archaeology, University of Haifa, Haifa, Mount Carmel, Israel.

*Corresponding author. Email: anatom2@tauex.tau.ac.il (I.H.); mayhila@tauex.tau.ac.il (H.M.); sarigrac@tauex.tau.ac.il (R.S.) †These authors contributed equally to this work. ‡Present address: Natural History Museum, University of Florence, Florence, Italy.

leading to African MP *Homo*, close to the split from the *H. erectus* branch and to European MP *Homo* and Neanderthals (including Atapuerca SH), and far from early and recent *H. sapiens* (Fig. 2B).

With regard to the configuration of the endoparietal surface (fig. S5), NR-1 is polygonal, i.e., the surface is clearly oriented according to three distinct planes (fig. S5). Instead, Neanderthals and *H. sapiens* manifest an arched endoparietal surface. The flatness of the superior parietal lobule, seen in the NR-1 virtual endocast (fig. S6), is one of the most characteristic features of MP *Homo* (*9, 10*). Other important characteristics of the NR-1 endocast and that are also typical of MP *Homo* are the very low position of the maximum endocranial width at the superior part of the first temporal convolution (fig. S6), the very short parietal lobe (fig. S7), and the differing lengths of the maximal endocranial width and intraparietal width as well as their posterior position on the parietal bone (table S3). These features can sometimes also be seen in Neanderthals (table S3) (*9, 10*). Conversely, recent *H. sapiens* specimens exhibit subequal maximal endocranial and intraparietal widths, which are located much higher and more anteriorly than in NR-1 (*9*).

The vascular pattern of the middle meningeal vessels in NR-1 is simple. Only a few, short ramifications are visible and anastomoses are absent, as is the case in other MP *Homo* and Neanderthals (figs. S9 and S10) (*11*). The posterior branch of the middle meningeal vessel in NR-1 is as developed as the anterior one, a pattern persistent among MP *Homo*. Both Neanderthals (e.g., La Quina H5 and La Chapelle-aux-Saints) and recent *H. sapiens* show a dominance of the anterior branch; the latter also possesses complex vascular endocranial imprints (fig. S10).

The NR-2 specimen is a robust mandible (Fig. 3); the corpus is medio-laterally wide, and the cortical bone is thick (Fig. 3, fig. S11). Its most pronounced feature is the short ramus relative to the body height, with a sturdy, low, and wide coronoid process (Fig. 3). This specimen displays several archaic features (e.g., no trigonum mentale or incurvatio mandibulae, a wide incisura submentalis, a developed planum alveolare, a strongly developed planum triangulare, and a mandibular corpus that presents fairly parallel alveolar and basal margins) commonly seen in MP *Homo* (*12, 13*) (supplementary text D and table S4A).

We combined taxonomically relevant mandibular features into a hierarchical clustering analysis (fig. S12). Modern and Pleistocene humans form the two main clusters: NR-2 is placed on a side branch of the latter, together with MP *Homo* from Atapuerca SH, Tighenif 3, Arago XIII, and one Neanderthal (fig. S12). The discrete traits underscore the mosaic nature of the NR mandible, showing archaic morphology together with some Neanderthal traits.
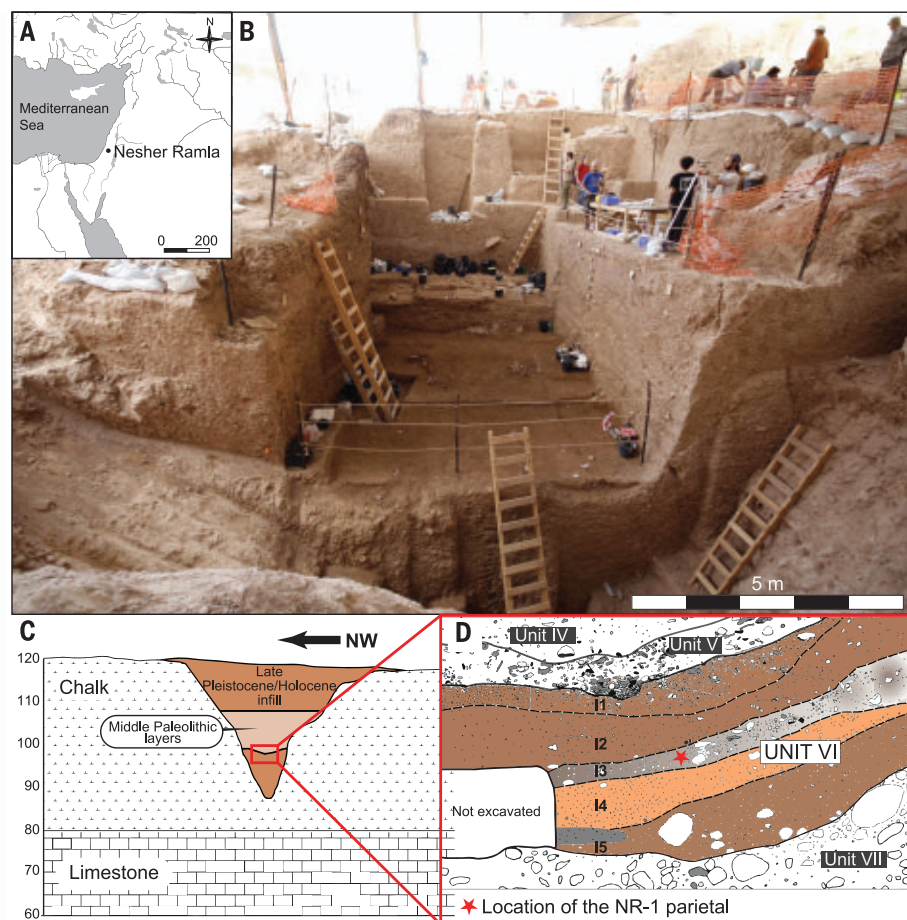
Fig. 1. The Nesher Ramla site and its stratigraphy. (**A**) Map of the eastern Mediterranean region showing the location of the site. (**B**) View of the excavation from the east. (**C**) Section through the general archaeological sequence placed within the local geology (the vertical scale on the left shows meters above sea level). (**D**) Magnification of Units IV to VII. A red star denotes the location of the NR-1 (parietal bone) 8 m below the surface in Unit VI. NR-2 (mandible) was found in the same unit, 2 m to the north.

The metric dimensions of the NR-2 mandibular body are presented in fig. S13. The symphyseal area is considerably thick (16.6 mm), close to the values of European MP *Homo* mandibles (16.9 ± 2.1 mm), and moderately tall (33.7 mm), close to the Neanderthal mean (34.0 ± 4.6 mm). The body (measured between the first and the second molar) is thick (17.7 mm), within the range of European MP *Homo* (18.1 ± 3.1 mm), yet taller (32.7 mm) than that of European MP *Homo* (30.2 ± 1.6 mm) and Neanderthals (29.9 ± 3.3 mm), close to the values of early *H. sapiens* (33.0 ± 4.0 mm).

The results of the 3D GM analysis (fig. S14 and tables S5A and S5B) for the NR-2 mandible are illustrated in Fig. 3C. The first two principal components explain 47.5% of the total variance. Variation along PC1 (37.9%) is driven by changes in the length of the mandibular body, the shape of the ramus (shorter and broader among archaic *Homo*), and the expression of the mental area. Variation along PC2 (9.6%) reflects changes in the body height

(mainly in the mental region), and the transition from a body's parallel alveolar and basal margins to ones that converge posteriorly. In the PC1-PC2 plot, early and recent *H. sapiens* separate from the other *Homo* specimens, whereas European MP *Homo* and Atapuerca SH are distinguished from Neanderthals (including the Levantine Amud 1) and Asian *H. erectus*. NR-2 falls between Neanderthals and the European MP *Homo* specimens (including Atapuerca SH), far outside the range of the variation of *H. sapiens*. The phylogenetic analysis, based on the mandibular mean shape of each hominin group, placed NR-2 on a separate branch (together with Tabun C1), close to the split between MP European fossils and Neanderthals, and far from *H. erectus*, African MP *Homo*, and *H. sapiens* (Fig. 3B). This result, based on metrics alone, largely echoes the results of the cluster analysis based on discrete traits and confirms that NR-2 belongs to an archaic group with Neanderthal affinities.

The lower second molar (NR-2 M$_2$) is complete and shows some occlusal wear causing a slight exposure of the dentine horns (Fig. 4A and supplementary text E). The occlusal surface of the NR-2 M$_2$ reveals four well-developed cusps and a hypoconulid. The presence of five main cusps is typical for most (70%) of the Atapuerca SH *Homo* (*14*) and Neanderthals (*15*). The NR-2 M$_2$ has a clear continuous mid-trigonid crest and a discontinuous distal trigonid crest on the dentine surface, corresponding to grade 3 of Bailey *et al.* (*16*) (fig. S15). A mid-trigonid crest is present in more than 90% of Neanderthals and MP *Homo* from Atapuerca SH (*14*, *15*). A grade 3 expression of the mid-trigonid crest, as in the NR-2 M$_2$, is present in nearly 60% of the Neanderthal specimens, but it is absent in *H. sapiens* (*16*). The

Qesem Cave M$_2$ specimen (QC-J15) (*17*) shows a similar pattern of a continuous mid-trigonid crest and a discontinuous distal trigonid crest (fig. S15). The Ehringsdorf G specimen presents only a mid-trigonid crest (but no distal crest) (fig. S15), whereas the Mauer specimen does not manifest a mid-trigonid crest at all.

The NR-2 M$_2$ has a single, pyramidal root bifurcating at the apical fourth of the root (Fig. 4, C and D). The large pulp cavity extends to the middle of the root and branches out into short root canals that extend into the apices, a configuration of the roots known as taurodontism (Fig. 4). This pyramidal root, with a taurodontic pulp cavity, is frequent in Neanderthals (*18*). In modern humans, the second lower molars possess separate mesial and distal roots with some variation in the canals. The root of the NR-2 M$_2$ (Fig. 4 and fig. S15) is relatively long (16.4 mm), falling toward the higher end of the range of the variation of both Upper Paleolithic *H. sapiens* (11.3 to 16.8 mm) and Neanderthals (14.3 to 16.5 mm).

The 3D GM analysis for the dentinal crown shape (landmark configuration combining the information from the enamel-dentin junction or EDJ, and that from the cemento-enamel junction or CEJ: fig. S16 for the measurement template, table S6 for the landmark definitions, and fig. S17 for the PC1-PC2 plot and the PC1-PC3 plot) showed that the NR-2 M$_2$ falls at the upper distant margin of the Neanderthal range, close to the Krapina specimens and Ehringsdorf G.

Shape variation along PC1 (30.6% of the total variance) is driven by the relative height of the crown and by the bucco-lingual expansion of the EDJ relative to the dentine outline. Like the M$_2$ of Neanderthals and *H. sapiens*, NR-2 M$_2$ exhibits a relatively high crown and a bucco-lingually expanded EDJ. Along PC2 (14.7%), the NR-2 M$_2$ plots toward the most extreme range of the distribution, opposite to the *H. sapiens*, Atapuerca SH, and African MP specimens. The associated shape is characterized by the expansion of the distal aspect of the dentine crown, a feature that NR-2 M$_2$ shares with some Neanderthal specimens (Krapina and El Sidrón) and the European MP *Homo* Ehringsdorf G (supplementary text E). Differently from the parietal and mandible, the unrooted phylogenetic trees' construction, based on the combined CEJ-EDJ data (Fig. 4B), resulted in a clear affiliation with Neanderthals, whereas Qesem QC-J15 associated with Atapuerca SH. Concerning crown size, NR2 M$_2$ is outside the modern human range (fig. S18).

The cumulative evidence from the three analyzed anatomical elements (parietal bone, mandible, and M$_2$) reveal a unique combination of archaic and Neanderthal features, supporting the existence of a local, Levantine population at the final MP. The results of the
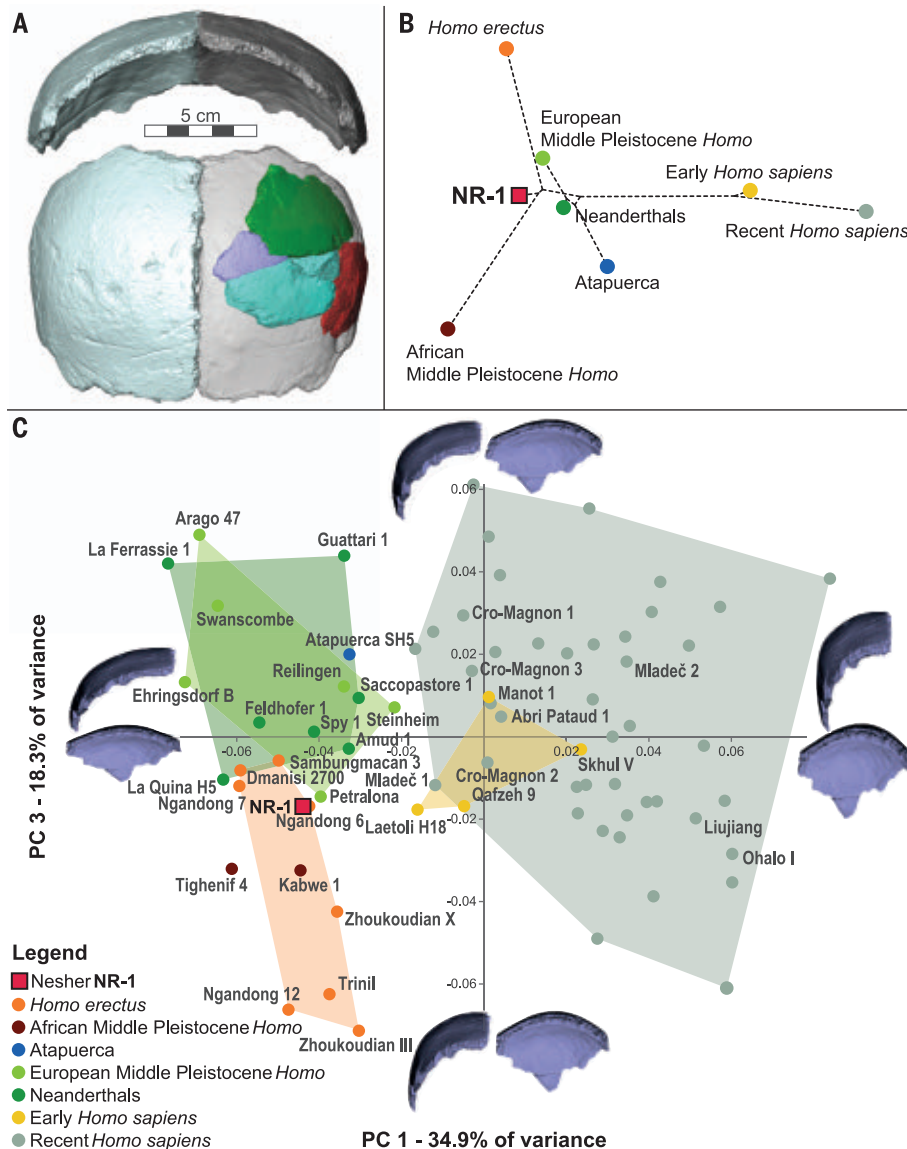
Fig. 2. The NR-1 parietal bones and their analyses. (A) Virtual reconstruction of the middle portion of the calvarium including both parietals. The almost complete right parietal was mirrored-imaged and the left parietal fragments (represented in various colors) were superimposed, showing a very good overlap with the contralateral part. **(B)** An unrooted phylogenetic tree constructed on the mean shape of each group using the neighbor-joining method. NR-1 is positioned on the line leading to African MP *Homo*, close to the splitting point from *H. erectus* and not far from the European MP *Homo* and Neanderthals. **(C)** PCA plot in shape space for the parietal bone. PC1 separates early and recent *H. sapiens* from the rest of the *Homo* sample. European MP *Homo* and Neanderthals (including the Levantine Amud 1) overlap and are distinguished from Asian *H. erectus* and African MP *Homo* along PC3; NR-1 is intermediate between these clusters and is close to Petralona and some late Asian *H. erectus*. Extreme shapes along the PCs are shown from the anterior and medial views.

quadratic discriminant analyses (QDAs) (table S7) reinforce this observation, showing that an affiliation of the NR fossils with early and recent *H. sapiens* is highly unlikely, but that it is impossible to establish whether NR fossils are more likely to be classified as MP *Homo*, Neanderthal, or *H. erectus* (the latter for the parietal only). Consequently, the discriminant function plot (fig. S1) shows that the NR-1 parietal falls between the *H. erectus*/African MP *Homo* group and the European MP *Homo*/ Neanderthals, with a similar likelihood of belonging to either cluster (*H. erectus* = 0.41, MP *Homo* = 0.34, Neanderthal = 0.25, based on the first three PCs).

The earliest that Neanderthal features in Levantine fossils have been discernible in the MP was around 400 ka ago at Qesem Cave (*19*), the earliest modern humans were present in the Levant around 180 ka ago (*20*), and unequivocal Neanderthals did not appear in the Middle East before ~70 ka ago. NR bridges a gap in this record, by displaying a highly heterogeneous, yet archaic morphology. The parietal documents a rather archaic shape of the braincase; the mandible is similar to that of MP *Homo*; the molar is quite Neanderthal-like, similar to Ehringsdorf G.

Arsuaga *et al.* (*21*) advocated an earlier evolutionary development of the masticatory apparatus, compared with the braincase in Neanderthals. Similarly, the Jebel Irhoud fossils from North Africa possess a more primitive neurocranium but a more *H. sapiens*–like face and dentition (*22*). Archaic populations carrying Neanderthal-like features were also present across much of the Eurasian continent during the MP, revealed by the Chinese findings of Maba, Xujiayao, and Xuchang (*23–27*). The existence of MP Asian populations deviating markedly from the *H. erectus* paradigm has been repeatedly proposed, for instance, for the Tongzi teeth (*28*) or the Sambungmacan 3 cranium (*29*); the latter (together with Ngandong 6 and 7) shows strong morphological affinities with the NR-1 parietals.

The NR fossils could represent late-surviving examples (140 to 120 ka) of a distinctive Southwest Asian MP *Homo* group, predating Levantine Neanderthals from Amud, Kebara, and Ein Qashish (70 to 50 ka). On the basis of their mosaic morphology showing a different degree of Neanderthal features, other MP Levantine fossils, whose taxonomic affinities have long been debated, from the sites of Qesem Cave (*19*), Zuttiyeh Cave (*30*), and probably Tabun Cave (*31*), might also be attributed to this group (supplementary text F). Adopting the cautious approach advocated by Mayr (*32*), we suggest addressing this Levantine MP paleodeme as the "Nesher Ramla *Homo*." Its presence from ~420 to 120 ka ago in a geographically restricted area may have allowed for repeated interbreeding with modern
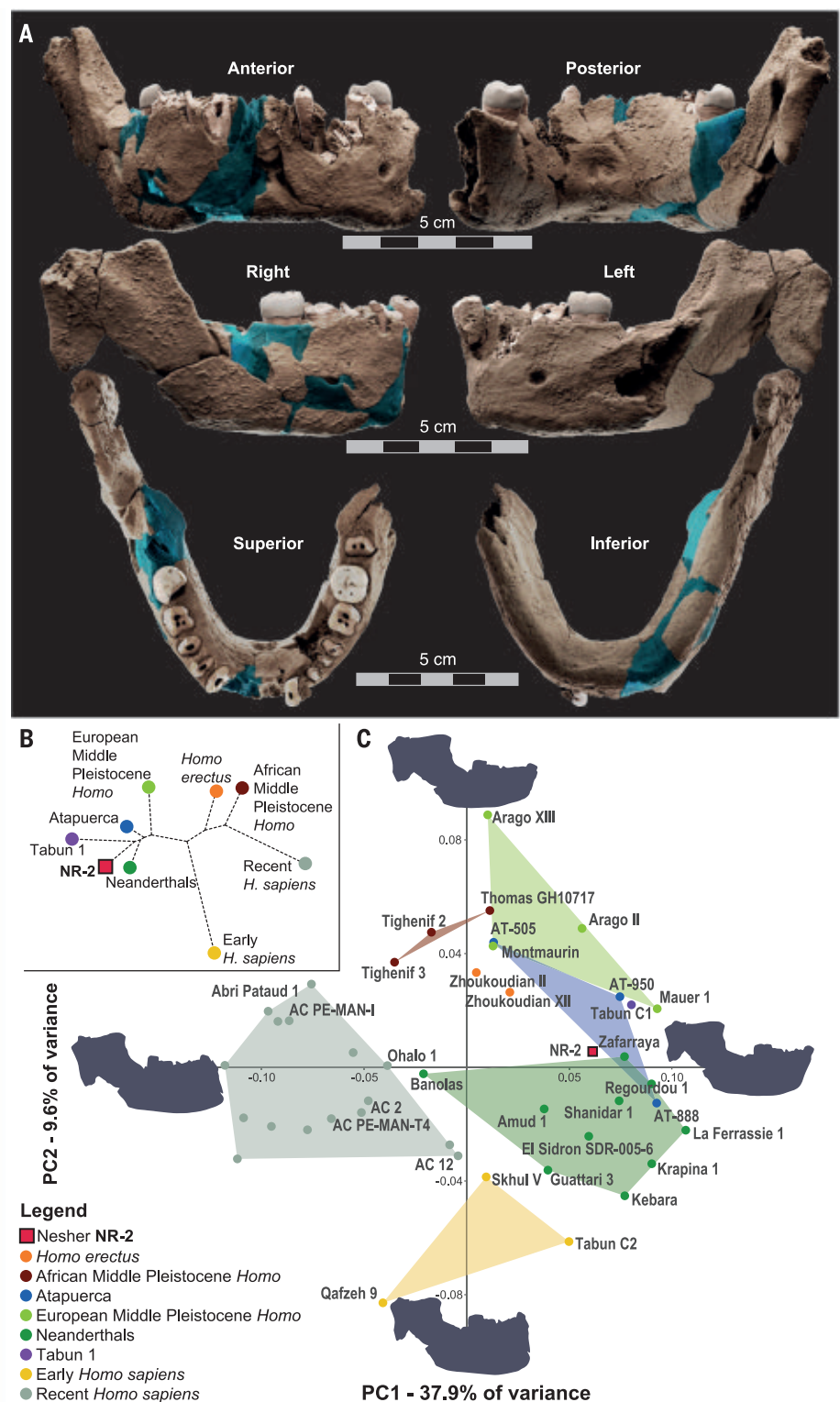
Fig. 3. The NR-2 mandible and its analysis. (**A**) Different views of the reconstructed mandible. (**B**) Unrooted phylogenetic tree constructed using the mean shape of each *Homo* group using the neighbor-joining method. NR-2 and Tabun C1 have a common ancestor that is close to the split between Neanderthals and European MP *Homo* and far from early and recent *H. s*apiens, as well as from *H. erectus* and African MP *Homo*. (**C**) PCA plot for the mandible in shape space. The combination of PC1 and PC2 separates early and recent *H. sapiens* from the other *Homo* specimens and distinguishes European MP *Homo* and Atapuerca from Neanderthals (including the Levantine Amud 1) and Asian *H. erectus*. NR-2 plots between Neanderthals and European MP *Homo*. Extreme shapes along PC1 and PC2 are shown from a lateral view.
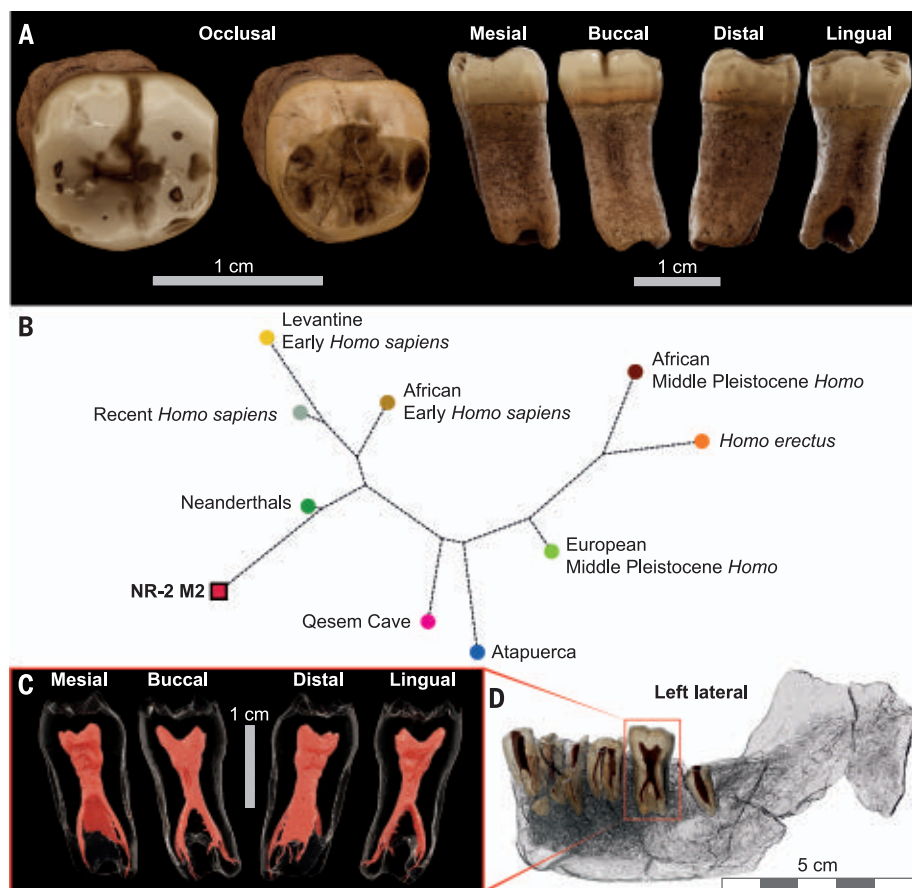
Fig. 4. The lower left second molar (NR2 M$_2$) and its analysis. (**A**) From the left, the NR2 M$_2$ in an occlusal view, with and without an enamel cap, and in mesial, buccal, distal, and lingual views. The mesial and distal interproximal wear facets are visible in the buccal and lingual views. (**B**) An unrooted phylogenetic tree construction, based on the mean shape of each hominin group using the neighbor-joining method. NR2 M$_2$ is close to Neanderthals and far from *H. sapiens*. (**C**) Root canals (pulp cavities) in the mesial, buccal, distal, and lingual views. The pyramidal roots and the taurodontic pulp cavity extend into the apex before branching out into short root canals. (**D**) The dental roots (in brown) and their canals (in red) are presented in "glass" images of the mandible from a lateral view, revealing the presence of taurodontism.

human populations such as the people from Misliya Cave (*20*), a notion also supported by their shared technological tradition [(*6*); supplementary text F]. This scenario is compatible with evidence of an early (200 to 400 ka ago) gene flow between modern humans and Neanderthals (*3, 4*) and helps explain the variable expression of the dental and skeletal features of later Levantine fossils from the Skhul and Qafzeh populations, a phenomenon noted by anthropologists since the 1930s (*31, 33*). Moreover, a recent study of the Atapuerca SH and Arago dental remains (*1*) suggested the existence of more than one *Homo* lineage in MP Europe [see also (*34*)] and hypothesized the contribution of Levantine *Homo* groups carrying Neanderthal-like traits to European *Homo* lineages. The NR *Homo*, carrying Neanderthal-like traits, could thus represent the "source" population postulated in the demographic "sources and sinks" model (*5*), according to which Western Europe was repopulated through a series of successive migrations.

REFERENCES AND NOTES

1. J. M. Bermúdez de Castro et al., *Quat. Sci. Rev.* **217**, 45–61 (2019).
2. M. Roksandic, P. Radović, J. Lindal, *Quat. Int.* **466**, 66–81 (2018).
3. C. Posth et al., *Nat. Commun.* **8**, 16046 (2017).
4. M. Meyer et al., *Nature* **531**, 504–507 (2016).
5. R. W. Dennell, M. Martinón-Torres, J. M. Bermúdez de Castro, *Quat. Sci. Rev.* **30**, 1511–1524 (2011).
6. Y. Zaidner et al., *Science* **372**, 1429–1433 (2021).
7. E. Bruner et al., *Coll. Antropol.* **28**, 99–112 (2004).
8. P. Gunz, S. Neubauer, B. Maureille, J.-J. Hublin, *Curr. Biol.* **20**, R921–R922 (2010).
9. D. Grimaud-Hervé, *L'évolution de l'encéphale chez Homo erectus et Homo sapiens: Exemples de l'Asie et de l'Europe.* Cah. paléoanthropologie (CNRS éditions, 1997).
10. E. M. Poza-Rey, A. Gómez-Robles, J. L. Arsuaga, *J. Hum. Evol.* **129**, 67–90 (2019).
11. E. Bruner, S. Mantini, A. Perna, C. Maffei, G. Manzi, *Eur. J. Morphol.* **42**, 217–224 (2005).
12. A. Vialet, M. Modesto-Mata, M. Martinón-Torres, M. Martínez de Pinillos, J.-M. Bermúdez de Castro, *PLOS ONE* **13**, e0189714 (2018).
13. A. Mounier, F. Marchal, S. Condemi, *J. Hum. Evol.* **56**, 219–246 (2009).
14. M. Martinón-Torres, J. M. Bermúdez de Castro, A. Gómez-Robles, L. Prado-Simón, J. L. Arsuaga, *J. Hum. Evol.* **62**, 7–58 (2012).
15. S. Bailey, *Period. Biol.* **108**, 253–267 (2006).
16. S. E. Bailey, M. M. Skinner, J. J. Hublin, *Am. J. Phys. Anthropol.* **145**, 505–518 (2011).
17. I. Hershkovitz et al., *Quat. Int.* **398**, 148–158 (2016).
18. K. Kupczik, J.-J. Hublin, *J. Hum. Evol.* **59**, 525–541 (2010).
19. I. Hershkovitz et al., *Am. J. Phys. Anthropol.* **144**, 575–592 (2011).
20. I. Hershkovitz et al., *Science* **359**, 456–459 (2018).
21. J. L. Arsuaga et al., *Science* **344**, 1358–1363 (2014).
22. J.-J. Hublin et al., *Nature* **546**, 289–292 (2017).
23. X. J. Wu, E. Bruner, *Am. J. Phys. Anthropol.* **160**, 633–643 (2016).
24. X.-J. Wu, I. Crevecoeur, W. Liu, S. Xing, E. Trinkaus, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 10509–10513 (2014).
25. Z.-Y. Li et al., *Science* **355**, 969–972 (2017).
26. S. Xing, M. Martinón-Torres, J. M. Bermúdez de Castro, X. Wu, W. Liu, *Am. J. Phys. Anthropol.* **156**, 224–240 (2015).
27. W. Liu et al., *J. Hum. Evol.* **64**, 337–355 (2013).
28. S. Xing, M. Martinón-Torres, J. M. Bermúdez de Castro, *J. Hum. Evol.* **130**, 96–108 (2019).
29. S. Márquez, K. Mowbray, G. J. Sawyer, T. Jacob, A. Silvers, *Anat. Rec.* **262**, 344–368 (2001).
30. S. E. Freidline, P. Gunz, I. Janković, K. Harvati, J. J. Hublin, *J. Hum. Evol.* **62**, 225–241 (2012).
31. A.-M. Tillier, *J. Isr. Prehist. Soc.* **35**, 439–450 (2005).
32. E. Mayr, in *Cold Spring Harbor Symposia on Quantitative Biology* (Cold Spring Harbor Laboratory Press, 1950), vol. 15, pp. 109–118.
33. B. Arensburg, A. Belfer-Cohen, in *Neandertals and Modern Humans in Western Asia*, T. Akazawa, K. Aoki, O. Bar-Yosef, Eds. (Plenum, 1998), pp. 311–322.
34. G. Manzi, *Quat. Int.* **411**, 254–261 (2016).

PALEOANTHROPOLOGY

Middle Pleistocene *Homo* behavior and culture at 140,000 to 120,000 years ago and interactions with *Homo sapiens*

Yossi Zaidner[1]*, Laura Centi[1]†, Marion Prévost[1]†, Norbert Mercier[2], Christophe Falguères[3], Gilles Guérin[4], Hélène Valladas[4], Maïlys Richard[2,3,5,6], Asmodée Galy[2,7], Christophe Pécheyran[7], Olivier Tombret[3,8], Edwige Pons-Branchu[4], Naomi Porat[9], Ruth Shahack-Gross[10], David E. Friesem[10], Reuven Yeshurun[11], Zohar Turgeman-Yaffe[11], Amos Frumkin[12], Gadi Herzlinger[1,13], Ravid Ekshtain[1], Maayan Shemer[14,15], Oz Varoner[16], Rachel Sarig[17,18]‡, Hila May[18,19]‡, Israel Hershkovitz[18,19]‡

Fossils of a Middle Pleistocene (MP) *Homo* within a well-defined archaeological context at the open-air site of Nesher Ramla, Israel, shed light on MP *Homo* culture and behavior. Radiometric ages, along with cultural and stratigraphic considerations, suggest that the fossils are 140,000 to 120,000 years old, chronologically overlapping with *H. sapiens* in western Asia. Lithic analysis reveals that MP *Homo* mastered stone-tool production technologies, previously known only among *H. sapiens* and Neanderthals. The Levallois knapping methods they used are indistinguishable from that of concurrent *H. sapiens* in western Asia. The most parsimonious explanation for such a close similarity is the cultural interactions between these two populations. These findings constitute evidence of contacts and interactions between *H. sapiens* and MP *Homo*.

The emergence and expansion of *Homo sapiens* during the late Middle Pleistocene (MP) in Africa is associated with complex behaviors and technologies that typify the Middle Stone Age (MSA) (*1–5*). One of the major technological innovations of the MSA is the Levallois technology that emerged and spread across most of the African continent ~300,000 years (300 ka) ago (*1, 6*). During the late MP, the centripetal Levallois method was used as the main mode for the production of flakes and blades in many sites in Africa and western Asia (Fig. 1 and table S1). The centripetal Levallois method is a well-structured technical process executed using a set of distinct and repetitive actions (*7*). The earliest evidence of the centripetal Levallois technology was reported in the early MSA sites at the Kapthurin and Gademotta Formations (dated to 250 to 200 ka and ~276 ka, respectively) (*2, 8*). The MSA in general, and centripetal Levallois technology in particular, were found to be associated with the *H. sapiens* remains at Omo Kibish, Herto, and Aduma (*9–11*) (tables S1 and S2).

The earliest occurrences of *H. sapiens* in southwest Asia (~180 ka) are associated with Middle Paleolithic industries and Levallois technology (*12*). During Marine Isotope Stage (MIS) 5 (130 to 71 ka), all *H. sapiens* fossils in western Asia were found to be associated with the centripetal Levallois method (*13–17*). Given the prominent presence of the centripetal Levallois method in association with *H. sapiens*, it was often used as a marker of *H. sapiens* dispersal into western Asia during MIS 5 (*17–20*). In western Europe, the centripetal Levallois technology occurs sporadically from MIS 8. A systematic use of the centripetal Levallois method

in Europe is recorded at the end of MIS 5 and during MIS 4 (text S6).

Two recently discovered human fossils at Nesher Ramla (*21*) provide evidence of the presence of archaic MP *Homo* in the Levant in a Middle Paleolithic context, during a period when the area was presumably inhabited by only *H. sapiens*. This suggests a long overlap-

ping period between these two *Homo* groups (*21*). Our study of the lithic assemblage from stratigraphic Unit VI of the site, associated with the Nesher Ramla fossils (Fig. 2, B to D), indicates that late MP *Homo* fully mastered the Levallois technology. Here, we report on the cultural context, chronometric ages, and stone-tool assemblages associated with this newly discovered *Homo*.

The Nesher Ramla karst sinkhole (Fig. 2A) is located in central Israel within a chalk bedrock of Senonian age. Middle Paleolithic cultural remains were uncovered in an 8-m-thick sequence, 107.5 to 99.5 m above sea level, ~12 m below the rim of the depression (*22*). The cultural sequence consists of six archaeological units (Units I to VI; Fig. 2, B and D, and text S1). The lowermost Unit VI, the focus of this study, is ~1 m thick and is subdivided into five layers (VI I1 to I5). A right parietal human bone and an almost complete human mandible (*21*) were found in layer VI I3, which is located in the middle of a sedimentological sequence of Unit VI (Fig. 2, B and D, and fig. S1, A and B).

A combination of electron spin resonance/ uranium series (ESR/U-series), thermoluminescence (TL), and optically stimulated luminescence (OSL) dating methods was applied to date the site and the human fossils (tables S3 to S5 and text S2). Three herbivorous teeth unearthed from Unit VI (I2 and I3) were analyzed using a combined ESR/U-series approach to overcome the changes in the uranium content of the
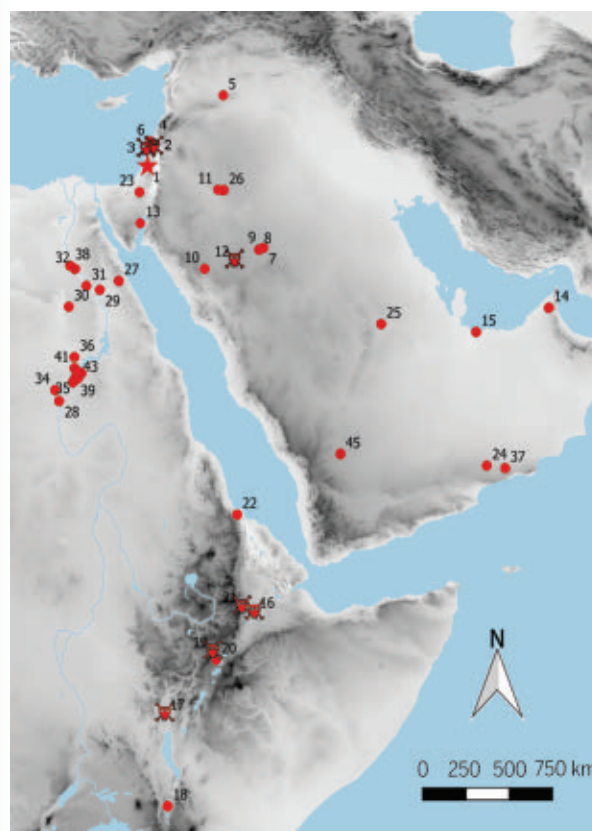

Fig. 1. Lithic assemblages with a marked component of Centripetal Levallois dated to MIS 5 and 6 in the Near East and eastern and northeastern Africa. Sites with human remains are denoted by a skull. The Nesher Ramla site is denoted by a star. The lithic assemblages marked on the map belong to different technocomplexes (see table S1 for a description of the sites and their technocomplexes). Digital elevation data are courtesy of GTOPO 30, USGS.

dental tissues that may have occurred since the burial time. The obtained ages range from 114 ± 12 ka to 140 ± 9 ka, leading to a weighted mean age of 126 ± 6 ka. The same approach yielded ages between 120 ± 9 ka and 128 ± 8 ka for animal teeth recovered in the overlying layer (Unit V) with a weighted mean of 122 ± 3 ka. Figure 3A and table S4 show equivalent doses, dose rates, and uranium uptake parameters for the enamel and dentine tissues (p- or n-values for the U-series and the Accelerating Uptake model, respectively; see text S2), and all the ESR/U-series ages. In addition, the TL dating method was applied to nine burnt flint samples collected from Unit V, ~50 cm above the fossils. The TL ages (Fig. 3A, table S3, and fig. S2) range from 191 ± 13 ka to 104 ± 11 ka; however, because these samples belong to a well-defined 20- to 40-cm-thick archaeological layer, they are likely to be coeval. Thus, the 191 ± 13 ka age appears to be an anomaly, confirmed by simple statistical tests (Chauvenet's criterion or the χ^2 test). When this result is ruled out as an outlier, the individual ages of the eight remaining flints are compatible within a 2σ error interval, and their weighted mean is 128 ± 4 ka. Moreover, because these samples have been subjected to similar external dose rates (table S3B shows small intersample external dose variations), an isochron analysis was performed (fig. S2A), resulting in an isochron age of 135 ± 13 ka, in agreement with the weighted mean TL age. This indicates that the external dose rates used for calculating the TL individual ages are most likely correct. The TL dates are indistinguishable from the ESR/U-series ages obtained for the same unit (Unit V), and they are in agreement with the weighted mean ESR/U-series date of 126 ± 6 ka for Unit VI (table S4). According to these chronometric results, Unit VI can then be confidently assigned an age of at least 120 ka, in agreement with the previously published OSL ages obtained for the whole sequence (22), indicating that the human occupation of the site occurred at the transition between MIS 6 and MIS 5.

The studies of the site-formation processes and the lithic assemblages support the results obtained by radiometric dating. Both Units V and VI exhibit similar sedimentological and micromorphological characteristics (fig. S3 and text S1). No hiatus or unconformities between the two units were observed. Micromorphological and sedimentological analyses suggest a continuous deposition by similar depositional mechanisms (23). Thus, we conclude that no gap in deposition or changes in the depositional environment occurred between the accumulation of Units VI and V. Fast and continuous deposition is also supported by the stone tools, which show similar characteristics in both units and therefore may have been
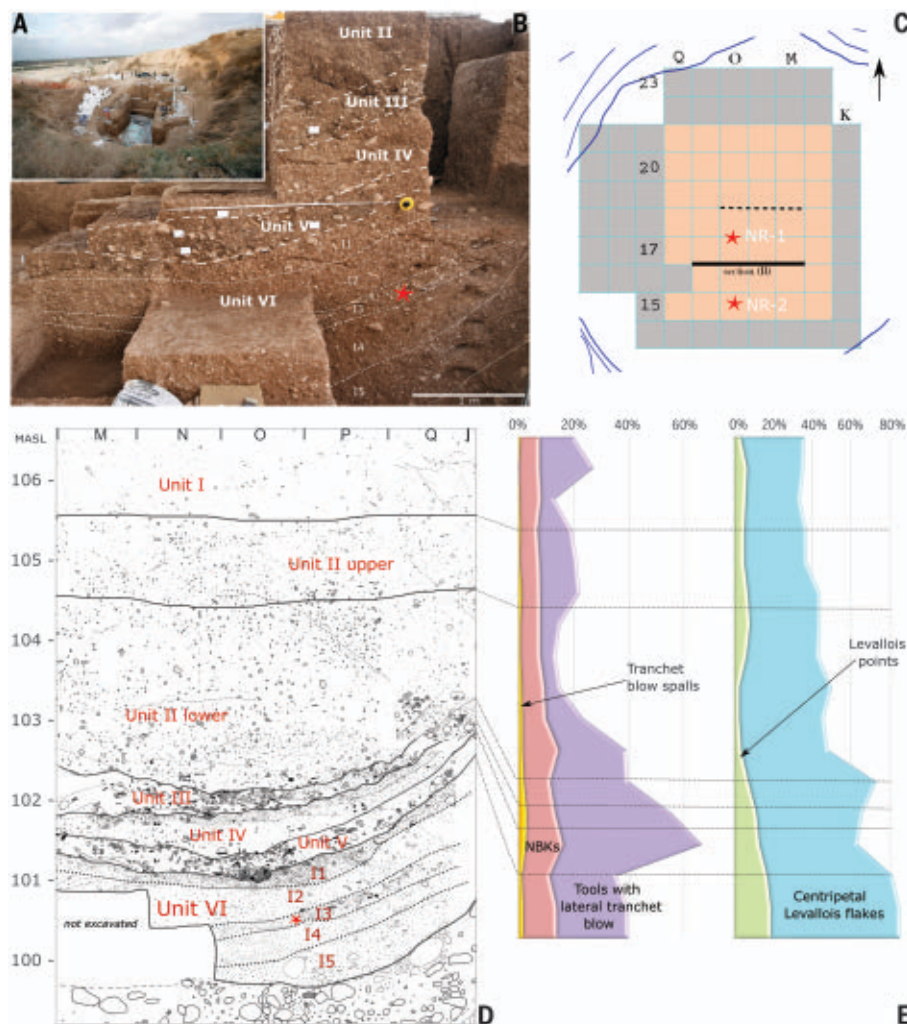

Fig. 2. The Nesher Ramla site. (**A**) General view of the site from east to west. (**B**) Photo of the stratigraphic section showing Units II lower to VI. The red star denotes the location of the MP *Homo* parietal bone. (**C**) The plan of the site. Excavated squares are denoted in gray. The studied lithic assemblage derives from the squares denoted in light brown. Red stars denote the location of the MP *Homo* fossils. The solid line denotes the location of the section shown in (B) and (D). The dashed line denotes the location of the section shown in fig. S1A. (**D**) Stratigraphic section of Nesher Ramla with its various units. The red star denotes the location of the parietal bone. (**E**) Fluctuations in the frequencies of the main lithic categories (lateral *tranchet* blow spalls from the total assemblage; tools with a lateral *tranchet* blow from the total retouched tool assemblage; Levallois points and centripetal Levallois flakes from the total Levallois assemblage). NBK, naturally backed knives.

[1]Institute of Archaeology, The Hebrew University of Jerusalem, Jerusalem, Israel. [2]Institut de Recherche sur les Archéomatériaux, UMR 5060 CNRS–Université Bordeaux Montaigne, Centre de Recherche en Physique Appliquée à l'Archéologie (CRP2A), Maison de l'Archéologie, 33607 PESSAC Cedex, France. [3]UMR7194, Départment "Homme et Environnement, Muséum national d'histoire naturelle, CNRS, UPVD, Sorbonne Université, Institut de Paléontologie Humaine, 75103 Paris, France. [4]Laboratoire des Sciences du Climat et de l'Environnement LSCE/IPSL, CEA-CNRS-UVSQ, Université Paris-Saclay, F-91198 Gif-sur-Yvette, France. [5]Department of Early Prehistory and Quaternary Ecology, University of Tübingen, 72070 Tübingen, Germany. [6]Centro Nacional de Investigación sobre la Evolución Humana, Paseo Sierra de Atapuerca 3, 09002 Burgos, Spain. [7]Université de Pau et des Pays de l'Adour, E2S UPPA, CNRS, IPREM, BP 576 64012 PAU Cedex, France. [8]UMR7209, Départment Homme et Environnement, Muséum national d'histoire naturelle, 75005 Paris, France. [9]Geological Survey of Israel, Jerusalem, Israel. [10]Department of Maritime Civilizations, Recanati Institute for Maritime Studies, Leon H. Charney School of Marine Sciences, University of Haifa, Haifa, Israel. [11]Zinman Institute of Archaeology, University of Haifa, Mount Carmel, Haifa, Israel. [12]Cave Research Center, Institute of Earth Sciences, The Hebrew University of Jerusalem, Edmond J. Safra Campus–Givat Ram, Jerusalem 9190401, Israel. [13]Department of Physics of Complex Systems, Weizmann Institute of Science, Rehovot, Israel. [14]Prehistoric Branch, Archaeological Research Department, Israel Antiquity Authority, Jerusalem, Israel. [15]Department of Bible, Archaeology and the Ancient Near East, Ben-Gurion University of the Negev, Beer-Sheva, Israel. [16]Israel Antiquities Authority, Jerusalem, Israel. [17]Department of Oral Biology, Maurice and Gabriela Goldschleger School of Dental Medicine, Tel Aviv University, Tel Aviv, Israel. [18]Shmunis Family Anthropology Institute, Dan David Center for Human Evolution and Biohistory Research, Tel Aviv University, Tel Aviv, Israel. [19]Department of Anatomy and Anthropology, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel.
*Corresponding author. Email: yzaidner@mail.huji.ac.il †These authors contributed equally to this work. ‡These authors contributed equally to this work.
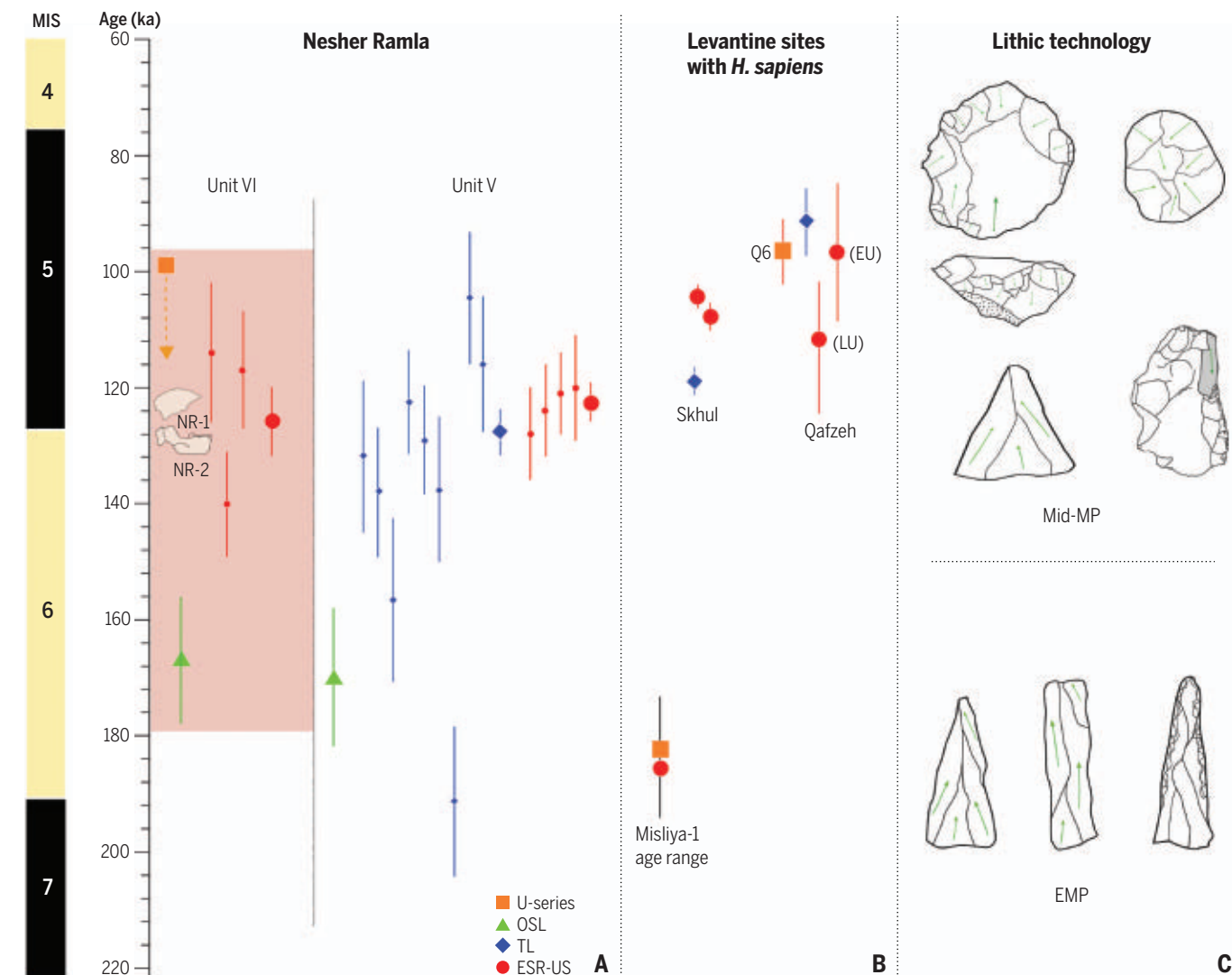
Nesher Ramla

Levantine sites with *H. sapiens*

Lithic technology

MIS

Age (ka)

Unit VI

Unit V

NR-1

NR-2

Q6 (EU)

(LU)

Skhul

Qafzeh

Mid-MP

Misliya-1 age range

EMP

■ U-series
▲ OSL
◆ TL
● ESR-US

A

B

C

Fig. 3. Chronology and position of the Nesher Ramla site in the Levantine Middle Paleolithic chrono-cultural framework. (**A**) Dating results for Units VI and V of the Nesher Ramla site. Large symbols represent weighted mean ages for ESR/U-series (circles), TL (diamonds), and OSL (triangles). Error bars represent possible ranges of the obtained ages. The orange square represents the minimum age obtained by the U-series method. NR-1 and NR-2 denote the MP *Homo* remains. (**B**) Chronology of the Middle Paleolithic sites with *H. sapiens* remains in the Levant. On the basis of mean ages (table S1), Q6 is the direct age of the Qafzeh 6 human fossil and Misliya 1 is the direct age of the Misliya *H. sapiens* maxilla. (**C**) Major lithic characteristics of Early Middle Paleolithic (EMP) and the middle Middle Paleolithic (mid-MP) in the Levant (*15*, *24*, *25*) (text S6).

manufactured by hominins with a shared cultural tradition (Fig. 2 and texts S3 and S4). Finally, Nesher Ramla's industry shows a clear similarity with the MIS 5 industries of the region (*24*) (text S6); it clearly differs from the regional Early Middle Paleolithic (EMP) industries (*25*), dating from 270 to 140 ka (*26*) (Fig. 3, B and C, and text S6). Taking into account the radiometric ages obtained and the depositional and cultural considerations, the most likely age for the Nesher Ramla *Homo* is ~140 to 120 ka.

The presence of some butchered faunal remains in anatomical articulation (fig. S1, D and E), the lithic refitting, and the in situ features (such as hearths and ash piles) indicate in situ human activities during the accumula-

tion of the Nesher Ramla archaeological sequence (*23*) (text S5). The faunal assemblage of Unit VI is dominated by tortoises and ungulates. The faunal assemblage was modified by human hunting and processing, as manifested by numerous cutmarks and hammer stone percussion marks. All skeletal parts are represented even for the largest ungulates, which testify to the hunting activities that took place at or very near the sinkhole. The presence of anatomical articulations, unbiased skeletal part profiles, and the virtual absence of abrasion indicate in situ accumulation within the sinkhole. The assemblage is generally composed of open-landscape taxa such as gazelle, equids, and ostrich, as well as animals inhabiting a wider variety of habitats, such as aurochs and boar,

and relatively small proportions of woodland-adapted fallow deer. This suggests a generally open landscape near Nesher Ramla during the accumulation of Unit VI. The ungulate composition and the dominance of tortoises are similar to other Lower Sequence units (text S5).

About 6000 artifacts (>2 cm) were excavated from Unit VI of Nesher Ramla. The lithic assemblage from Unit VI I2 to I5, directly associated with the Nesher Ramla fossils, consists of 2184 artifacts larger than 2 cm (Fig. 2C, tables S6 and S7, and text S3). All studied sublayers in which the fossils have been found, including sublayer I3, exhibit similar technological characteristics (table S6). The knapped lithic assemblage is made of flint. The hominins used local high-quality flint from the

Mishash Formation (fig. S4 and text S3). As indicated by the presence of cores, primary elements, flakes, and core maintenance products, Mishash flint was knapped at the site. The cores are completely exhausted, which suggests that hominins knapped the cores to their maximum potential. The flint, brought from distances that extend 10 km from the site, appears in minor frequencies in all technological categories; however, it is more frequent among retouched pieces, which implies that

hominins carried the retouched tools as a part of their mobile toolkit (table S8).

The Nesher Ramla *Homo* mainly used the centripetal Levallois method (figs. S5 to S7 and tables S6, S7, and S9). The lithic assemblage of Unit VI is dominated by round or rectangular wide Levallois flakes with centripetal and orthogonal scar patterns and well-prepared striking platforms (Fig. 4, fig. S5, tables S9 and S10, and text S3). The convexity of the debitage surfaces of the Levallois cores was achieved and

maintained through centripetal preparation. The assemblage exhibits a high frequency of Levallois *débordant* flakes with centripetal and orthogonal scar patterns as well as pseudo-Levallois points and flakes (Fig. 4, fig. S7, and table S9). These classic predetermining by-products of the centripetal Levallois reduction system were used to maintain the convexities of the dorsal surfaces of the Levallois cores. After the preparation of convexities, the predetermined flakes were produced by the
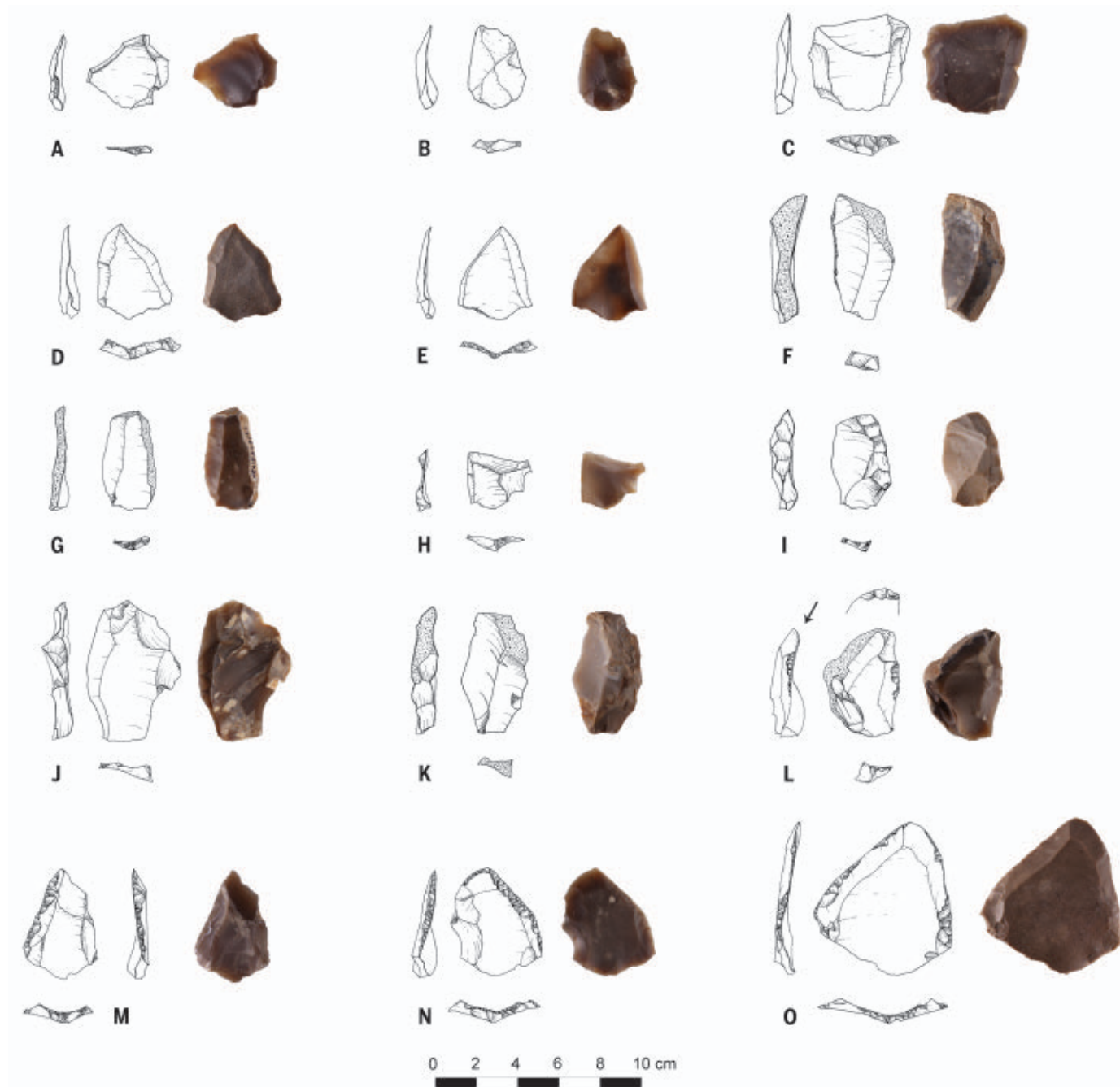
Fig. 4. Lithic assemblage of Unit VI I2 to I5. (**A** to **C**) Centripetal Levallois flakes; (**D** and **E**) Levallois points; (**F** and **G**) naturally backed knives; (**H**) pseudo-Levallois point; (**I** to **K**) *débordant* flakes; (**L**) tool with a lateral *tranchet* blow; (**M** and **N**) sidescrapers; (**O**) retouched Levallois point.

preferential method and the recurrent centripetal Levallois method (fig. S5). The centripetal Levallois method exhibits similar technological characteristics throughout the archaeological sequence of Nesher Ramla (tables S11 and S12 and text S4).

The production of Levallois points is an auxiliary reduction system recorded in Unit VI (tables S6 and S7). Levallois points occur in various frequencies throughout the site's stratigraphic sequence but are most frequent in Unit VI (table S11). Levallois points were produced by a preferential unidirectional convergent Levallois method and are mostly classical Y-shaped, with a rare use of bidirectional removals aimed at correcting the distal convexity (Fig. 4, figs. S5 and S8, and text S3). The points are symmetrical, flat, and broad-based, derived from a preferential mode of production.

Some additional distinct features of the Nesher Ramla assemblage are the systematic production of naturally backed knives and the extensive use of the lateral *tranchet* blow technique for producing tools with partly retouched and partly sharped edges (Fig. 4, figs. S7, S9, and S10, and table S13). These distinctive characteristics of the Nesher Ramla industry occur in varying frequencies throughout the entire archaeological sequence of the site, along with the centripetal Levallois technology and Levallois point production (Fig. 2E and text S4). The use of the same technologies and the production of the same set of artifacts suggest a cultural continuity in the area during the accumulation of the 8-m-thick archaeological sequence (Fig. 2, D and E, and tables S11 and S12).

MP *Homo* fossils often lack a cultural context, and their behavior and technology remain poorly known. Nonetheless, it is commonly suggested that MP *Homo* produced Lower Paleolithic industries [Acheulian, or core-on-flakes (*27–32*)]. The evidence from Nesher Ramla demonstrates that late MP *Homo* fully mastered advanced Levallois technology that until only recently was linked to either *H. sapiens* or Neanderthals. The use of the centripetal Levallois method by the Nesher Ramla *Homo* suggests caution in using lithic technology as a marker for the presence and dispersals of *H. sapiens* out of Africa in MIS 5. This is consistent with recent views that MP *Homo* could be one of the makers of MSA industries in Africa (*33*). The origin of the centripetal Levallois technology in the Levant is debatable (*13–15, 17–20*). The EMP (270 to 140 ka) industries in the Levant exhibit an entirely different technological set featuring the dominance of blades produced by non-Levallois laminar methods and the production of Levallois blanks using convergent unidirectional and bidirectional Levallois methods. The EMP industries show no evidence of the systematic use of centripetal Levallois methods (text

S6). Therefore, we suggest that the centripetal technology of the Levantine MIS 5 sites is not of local origin and that its most probable source is East and North Africa, where it represents a central component in the MIS 6 and MIS 5 assemblages (Fig. 1 and tables S1 and S2).

The centripetal Levallois technology used by the Nesher Ramla *Homo* clearly resembles the technology used in the *H. sapiens* sites of Qafzeh and Skhul, the MIS 5 sites in Arabia, and the MSA sites in North and East Africa, including sites where the remains of *H. sapiens* were found (Fig. 1, tables S1 and S2, and text S6). The resemblance lies in the fine details such as a similar mode of preparing the convexities and a similar way of processing the cores so as to produce predetermined Levallois flakes with a round or rectangular shape (tables S11 and S12 and text S6). This resulted in similar types of products, such as centripetal Levallois cores, centripetal Levallois flakes, *débordant* flakes with a centripetal scar pattern, and pseudo-Levallois flakes and points. Furthermore, Nesher Ramla hominins shared the preferential convergent unidirectional method for producing Levallois points with other Levantine MIS 5 sites (text S6) (*15*). In both Qafzeh Cave and Nesher Ramla, the preferential method is the most common one used in the production of flat and short Levallois points, as opposed to the elongated Levallois points with abrupt edges known in the Levantine EMP (*25*). The production of Levallois points by a unidirectional convergent Levallois method is rare in Africa during MIS 5, which suggests that it is mostly a Levantine feature shared by both MP *Homo* at Nesher Ramla and *H. sapiens* at the Qafzeh and Skhul caves.

Our results clearly show that *H. sapiens* and MP *Homo* shared core reduction technology during MIS 5 in a small region such as the Levant. We contend that cultural diffusion and interaction across *Homo* populations is the most likely reason for such a close cultural similarity between MP *Homo* and *H. sapiens*. These results are consistent with a growing body of genetic studies suggesting that a gene flow existed between divergent archaic *Homo* populations and *H. sapiens* during the late Middle and early Late Pleistocene (*34, 35*). Our findings provide archaeological support for close cultural interactions between different human lineages during the Middle Paleolithic period and suggest that contacts between MP *Homo* and *H. sapiens* had already occurred prior to 120 ka.

REFERENCES AND NOTES

1. D. Richter *et al.*, *Nature* **546**, 293–296 (2017).
2. K. Douze, A. Delagnes, *J. Hum. Evol.* **91**, 93–121 (2016).
3. A. S. Brooks *et al.*, *Science* **360**, 90–94 (2018).
4. C. R. Johnson, S. McBrearty, *J. Hum. Evol.* **58**, 193–200 (2010).
5. N. Porat *et al.*, *J. Archaeol. Sci.* **37**, 269–283 (2010).
6. S. P. McPherron, Z. Rezek, A. Ben-Ncer, J.-J. Hublin, *Bull. Archeol. Marocaine* **24**, 1–22 (2019).
7. E. Boëda, *Bull. Soc. Prehist. Fr.* **90**, 392–404 (1993).
8. C. A. Tryon, S. McBrearty, P.-J. Texier, *Afr. Archaeol. Rev.* **22**, 199–229 (2005).
9. J. D. Clark *et al.*, *Nature* **423**, 747–752 (2003).
10. J. Yellen *et al.*, *Paleoanthropology* **10**, 25–100 (2005).
11. J. J. Shea, *J. Hum. Evol.* **55**, 448–485 (2008).
12. I. Hershkovitz *et al.*, *Science* **359**, 456–459 (2018).
13. M. D. Petraglia, M. Haslam, D. Q. Fuller, N. Boivin, C. Clarkson, *Ann. Hum. Biol.* **37**, 288–311 (2010).
14. H. S. Groucutt *et al.*, *Nat. Ecol. Evol.* **2**, 800–809 (2018).
15. E. Hovers, *The Lithic Assemblages of Qafzeh Cave* (Oxford Univ. Press, 2009).
16. D. A. Garrod, D. M. A. Bate, *The Stone Age of Mount Carmel. Volume 1: Excavations at the Wadi El-Mughara* (Clarendon, 1937).
17. O. Bar-Yosef, in *Neandertals and Modern Humans in Western Asia*, T. Akazawa, K. Aoki, O. Bar-Yosef, Eds. (Plenum, 1998), pp. 39–56.
18. H. S. Groucutt *et al.*, *Evol. Anthropol.* **24**, 149–164 (2015).
19. J. J. Shea, *Quat. Int.* **350**, 169–179 (2014).
20. C. J. Bae, K. Douka, M. D. Petraglia, *Science* **358**, eaai9067 (2017).
21. I. Hershkovitz *et al.*, *Science* **372**, 1424–1428 (2021).
22. Y. Zaidner *et al.*, *J. Hum. Evol.* **66**, 1–17 (2014).
23. Y. Zaidner, A. Frumkin, D. Friesem, A. Tsatskin, R. Shahack-Gross, *Quat. Sci. Rev.* **138**, 76–86 (2016).
24. M. Prévost, Y. Zaidner, *PLOS ONE* **15**, e0231109 (2020).
25. Y. Zaidner, M. Weinstein-Evron, *J. Hum. Evol.* **144**, 102785 (2020).
26. I. Hershkovitz *et al.*, *Science* **362**, eaat8964 (2018).
27. E. Carbonell *et al.*, *J. Hum. Evol.* **37**, 653–693 (1999).
28. A. Ollé *et al.*, *Quat. Int.* **411**, 316–328 (2016).
29. D. Barsky, *C. R. Palevol* **12**, 305–325 (2013).
30. D. Barsky, A. M. Moigne, V. Pois, *J. Hum. Evol.* **135**, 102650 (2019).
31. P. G. Rightmire, *J. Hum. Evol.* **31**, 21–39 (1996).
32. A. de Lombera-Hermida *et al.*, *J. Hum. Evol.* **145**, 102812 (2020).
33. R. Grün *et al.*, *Nature* **580**, 372–375 (2020).
34. C. Posth *et al.*, *Nat. Commun.* **8**, 16046 (2017).
35. M. Hajdinjak *et al.*, *Nature* **555**, 652–656 (2018).

EPIGENETICS

MBD5 and MBD6 couple DNA methylation to gene silencing through the J-domain protein SILENZIO

Lucia Ichino[1,2], Brandon A. Boone[1,2], Luke Strauskulage[3,4], C. Jake Harris[2]†, Gundeep Kaur[5], Matthew A. Gladstone[2], Maverick Tan[2], Suhua Feng[2,6], Yasaman Jami-Alahmadi[7], Sascha H. Duttke[8], James A. Wohlschlegel[7], Xiaodong Cheng[5], Sy Redding[3], Steven E. Jacobsen[1,2,6,9]*

DNA methylation is associated with transcriptional repression of eukaryotic genes and transposons, but the downstream mechanism of gene silencing is largely unknown. Here, we describe two *Arabidopsis thaliana* methyl-CpG–binding domain proteins, MBD5 and MBD6, that are recruited to chromatin by recognition of CG methylation, and redundantly repress a subset of genes and transposons without affecting DNA methylation levels. These methyl readers recruit a J-domain protein, SILENZIO, that acts as a transcriptional repressor in loss-of-function and gain-of-function experiments. J-domain proteins often serve as co-chaperones with HSP70s. Indeed, we found that SILENZIO's conserved J-domain motif was required for its interaction with HSP70s and for its silencing function. These results uncover an unprecedented role of a molecular chaperone J-domain protein in gene silencing downstream of DNA methylation.

Cytosine DNA methylation (mC) in eukaryotes is typically associated with transcriptional silencing of genes and transposable elements (TEs); however, relatively little is known of the mechanism (*1, 2*). Mammalian genomes encode for several methyl-CpG–binding domain (MBD) proteins that are recruited to chromatin in part by recognition of methylated CG dinucleotides, but they also play methylation-independent roles in gene regulation (*3–7*). One prevailing model is that MBDs recruit histone deacetylase complexes to methylated DNA, causing chromatin compaction and gene silencing (*5–7*). In plants, loss of DNA methylation causes derepression of many transposons and genes (*8*), but no evidence has been found for a role of methyl readers in this process, leaving unresolved the question of what acts downstream of the methyl mark.

We recently identified two proteins named MBD5 and MBD6 from a mass spectrometry screen for methyl readers in *Arabidopsis thaliana* (*9*). MBD5 and MBD6 belong to a family of 13 members that have been identified by sequence similarity with human MBD domains (*10–12*). Outside of this domain, there is no sequence conservation between plants and animals. MBD5 and MBD6 are close relatives (*10–12*), they can interact with each other in vivo (*13, 14*), and were shown to bind methylated probes in electrophoretic mobility shift assays (*10, 15, 16*). Although a function has not been assigned to MBD5, MBD6 was shown to be required for ribosomal RNA gene regulation in allotetraploid genetic hybrids (*17*).

In plants, 5-methylcytosines are common in CG, CHG, and CHH sequence contexts (*18*). The MBD typically recognizes symmetrically methylated CG dinucleotides (*19*) but exceptions have been reported, for example, MeCP2, which can also bind mCA sites (*20*). We tested the ability of MBD5 and MBD6 to bind CG, CHG, or CHH methylation by performing fluorescence polarization (FP) assays with oligonucleotides methylated in different contexts. Both MBD5 and MBD6 showed a strong preference for CG-methylated oligonucleotides compared with unmethylated controls, but little preference was observed for CHG or CHH methylation (Fig. 1A and fig. S1). We also used DNA curtains, a single-molecule fluorescence microscopy assay, to visualize the interaction between MBD6 and flow-stretched bacteriophage λ DNA, which was methylated in vitro with the CG-specific bacterial *M.SssI* methyltransferase. MBD6 bound methylated, but not unmethylated, DNA curtains, and its enrichment profile correlated strongly with the local density of methylated CG sites (Fig. 1, B to D). To test the ability of MBD5 and MBD6 to bind methylation in natural *Arabidopsis* genomic sequences, we performed DNA affinity purification sequencing (DAP-seq) (*21*) by incubating Halo-tagged recombinant proteins with DNA extracted from wild-type plants or from

met1-3 mutant plants. The *met1-3* mutant is almost completely lacking in CG methylation because of a mutation in the maintenance *DNA METHYLTRANSFERASE 1* (*MET1*) gene but retains substantial levels of CHG and CHH methylation (*22*). We observed a strong genome-wide correlation between MBD5/6 DAP-seq enrichment and CG methylation density with DNA from a wild-type background and an almost complete loss of binding to DNA in the *met1-3* background (Fig. 1E). Only a few small peaks were retained in regions that did not completely lose CG methylation (fig. S2). Overall, these results strongly support the specificity of MBD5 and MBD6 for CG methylation in vitro.

We generated homology models of *Arabidopsis* MBD domains based on known mammalian MBD structures. High-confidence models could be determined except for the most divergent protein, MBD9, which is known not to bind methylated DNA in vivo (*23*) (fig. S3). The MBD5 and MBD6 structural models highlighted two arginine residues (R1 and R2) that are predicted to directly interact with methylated CGs by forming the previously described "methyl-Arg-G triad" (*19*) (Fig. 1F and fig. S3). We tested the importance of these residues by mutating them to alanine (MBD5[R1R2], MBD6[R1R2]) and indeed we observed a loss of specificity for binding to CG methylation in FP assays (fig. S1B).

We next investigated the genomic localization of MBD5 and MBD6 in vivo by chromatin immunoprecipitation sequencing (ChIP-seq) using FLAG-tagged transgenic lines. MBD5 and MBD6 bound methylated chromatin with a clear preference for mCG density as opposed to mCHG and mCHH density (Fig. 1, G and H). No correlation was found with the density of unmethylated CG sites (fig. S4). The MBD5[R1R2] and MBD6[R1R2] mutants showed a strong reduction of ChIP-seq enrichment (Fig. 1G and fig. S5), demonstrating that recognition of mCGs is required for recruitment of MBD5 and MBD6 to chromatin.

Methylated DNA is associated with three different chromatin states in *Arabidopsis*: euchromatic patches of RNA-directed DNA methylation (RdDM), which contain CG and non-CG methylation; pericentromeric heterochromatin, which is enriched in H3K9me2 as well as CG and non-CG methylation; and expressed genes containing gene body methylation (GbM), which are exclusively marked by CG methylation (*18*). We observed MBD5 and MBD6 ChIP-seq enrichment at a large fraction of sites in all three chromatin states, but the extent of enrichment was higher at RdDM sites compared with heterochromatin or GbM sites (fig. S6). However, the preference for RdDM sites was not observed by DAP-seq, which tests the ability of proteins to bind naked genomic DNA (fig. S6, C and D).

[1]Molecular Biology Institute, University of California Los Angeles, Los Angeles, CA 90095, USA. [2]Department of Molecular, Cell and Developmental Biology, University of California Los Angeles, Los Angeles, CA 90095, USA. [3]Department of Biochemistry and Biophysics, University of California San Francisco, San Francisco, CA 94143, USA. [4]Tetrad Graduate Program, University of California San Francisco, San Francisco, CA 94143, USA. [5]Department of Epigenetics and Molecular Carcinogenesis, University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA. [6]Eli and Edyth Broad Center of Regenerative Medicine and Stem Cell Research, University of California Los Angeles, Los Angeles, CA 90095, USA. [7]Department of Biological Chemistry, University of California Los Angeles, Los Angeles, CA 90095, USA. [8]Department of Medicine, University of California San Diego, La Jolla, CA 92093, USA. [9]Howard Hughes Medical Institute (HHMI), University of California Los Angeles, Los Angeles, CA 90095, USA.
*Corresponding author. Email: jacobsen@ucla.edu
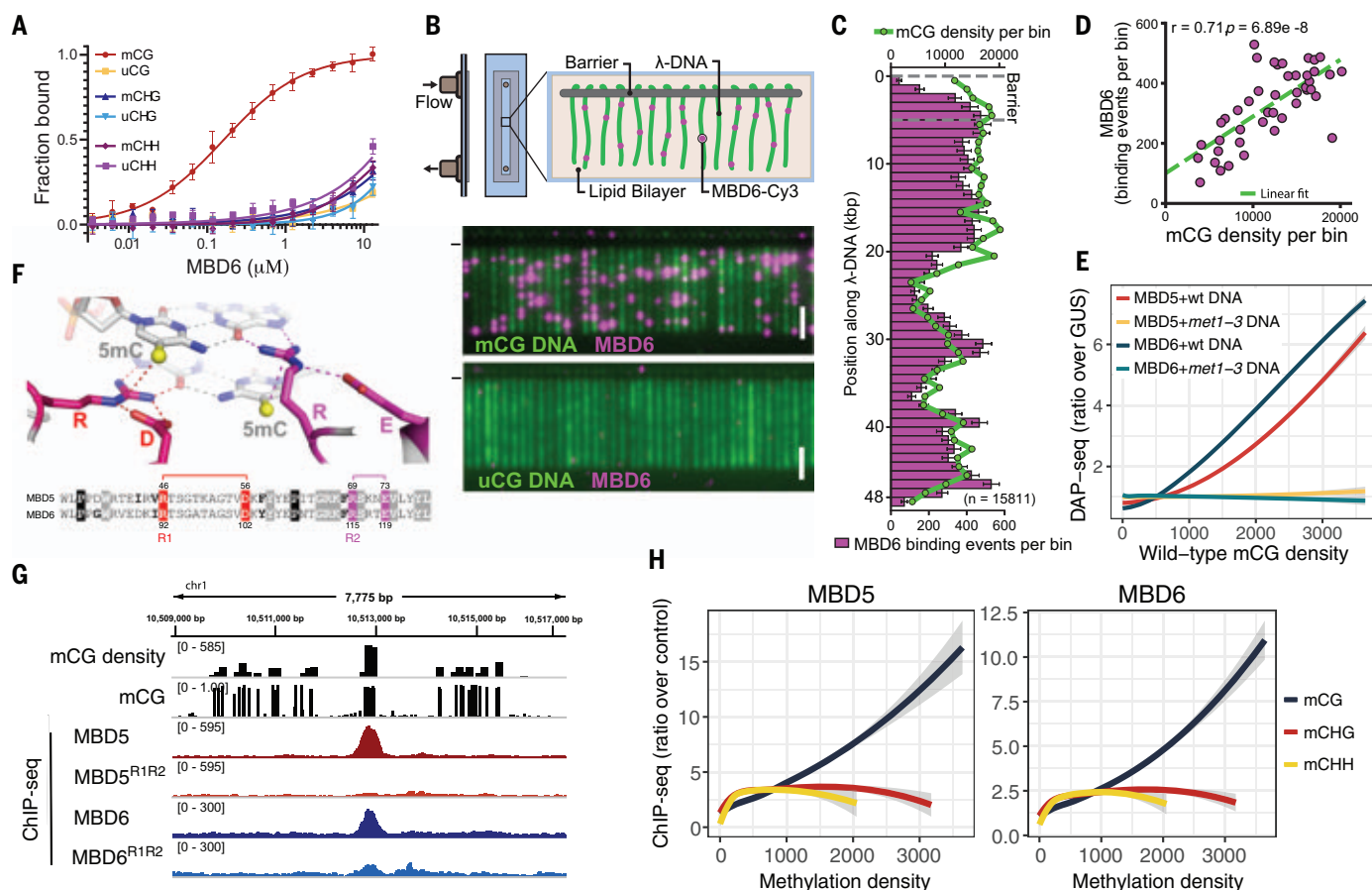†Present address: Department of Plant Sciences, University of Cambridge, Cambridge CB2 3EA, UK.

Fig. 1. MBD5 and MBD6 are CG-specific methyl readers in vitro and in vivo.
(**A**) Binding curves of MBD6 with DNA oligos methylated (m) or unmethylated (u) in the indicated contexts, measured by fluorescence polarization (N = 3, SEM). (**B**) Diagram of DNA curtain assay and representative image of YOYO-1–stained methylated (mCG) and unmethylated (uCG) DNA (green) bound by Cy3-labeled MBD6 (magenta). Horizontal lines on the lower panels indicate the chrome diffusion barriers. Scale bars, 5 μm. (**C**) Distribution of MBD6-binding events along mCG DNA overlaid with the distribution of mCG density (green line). Error bars indicate the 95% confidence intervals (CI) by bootstrap.

(**D**) Correlation scatterplot of MBD6 binding to methylated curtains and mCG density (1-kb bins). r is the Pearson correlation coefficient. (**E**) Genome-wide correlation between DAP-seq and mCG density (400-bp bins). Trend lines were calculated by locally weighted polynomial regression (loess curves). (**F**) Homology models of MBD5 and MBD6. The two arginine residues of the 5mC–Arg–G triads (R1 and R2) are shown in the sequence alignment. (**G**) Example ChIP-seq peaks at regions of dense CG methylation. (**H**) Loess curves of ChIP-seq enrichment and methylation density (400-bp bins overlapping Pol V ChIP-seq peaks). For (E) and (H), shaded area indicates 95% CI.

These observations suggest that recruitment of MBD5 and MBD6 to chromatin in vivo may be influenced by histones or other chromatin components.

To determine whether MBD5 and MBD6 regulate transcription at their targets, we performed RNA sequencing (RNA-seq) of *mbd5* and *mbd6* T-DNA mutants and of a double mutant generated by crossing them (*mbd5 mbd6*). A number of transposons and protein-coding genes were derepressed only in the double mutant, indicating genetic redundancy of *MBD5* and *MBD6* (Fig. 2A and fig. S7). We confirmed this with an independent *mbd5 mbd6* double mutant generated by CRISPR/Cas9 (figs. S7 and S8). Global run-on sequencing (GRO-seq) showed a similar pattern of changes, indicating that the derepression in *mbd5 mbd6* occurs at the transcriptional level (Fig. 2B). Most up-regulated genes and trans-

posons were not expressed in wild-type plants and showed high levels of promoter CG methylation, suggesting that they are direct targets (Fig. 2C). DNA methylation levels were not altered in *mbd5 mbd6* (Fig. 2C and fig. S9), indicating that the methyl readers act strictly downstream of DNA methylation. One of the derepressed genes was *FWA*, a well-characterized imprinted gene that is silenced by promoter methylation (24) (Fig. 2, D and E). Reintroduction into *mbd5 mbd6* mutant plants of FLAG-tagged versions of wild-type MBD5 or MBD6, but not their R1R2 mutant counterparts, was sufficient to largely rescue the derepression of *FWA* and of other genes and transposons (fig. S10). Overall, these results suggest that MBD5 and MBD6 are recruited to DNA by methylation and translate the methyl mark into gene repression at a subset of methylated sites.

We compared *mbd5 mbd6* gene expression data with those of mutants affecting different methylation pathways: *drm1 drm2* and *cmt2 cmt3* lose non-CG methylation at euchromatic RdDM sites and heterochromatic regions, respectively, whereas *met1-3* loses CG methylation genome-wide (22, 25). Most of the loci up-regulated in *mbd5 mbd6* were also derepressed in *met1-3*, indicating that they are silenced by CG methylation (fig. S11A). TEs derepressed in *mbd5 mbd6* were also longer than average and more enriched in H3K9me2, indicating that they are mostly heterochromatic TEs (fig. S11, B and C). A subset of these loci were also derepressed in *cmt2 cmt3*, but there were none derepressed in *drm1 drm2* (fig. S11A). Thus, whereas MBD5 and MBD6 are enriched at a wide range of CG-methylated sites, their repressive role is strongest at a subset of MET1-dependent heterochromatic loci. Furthermore,
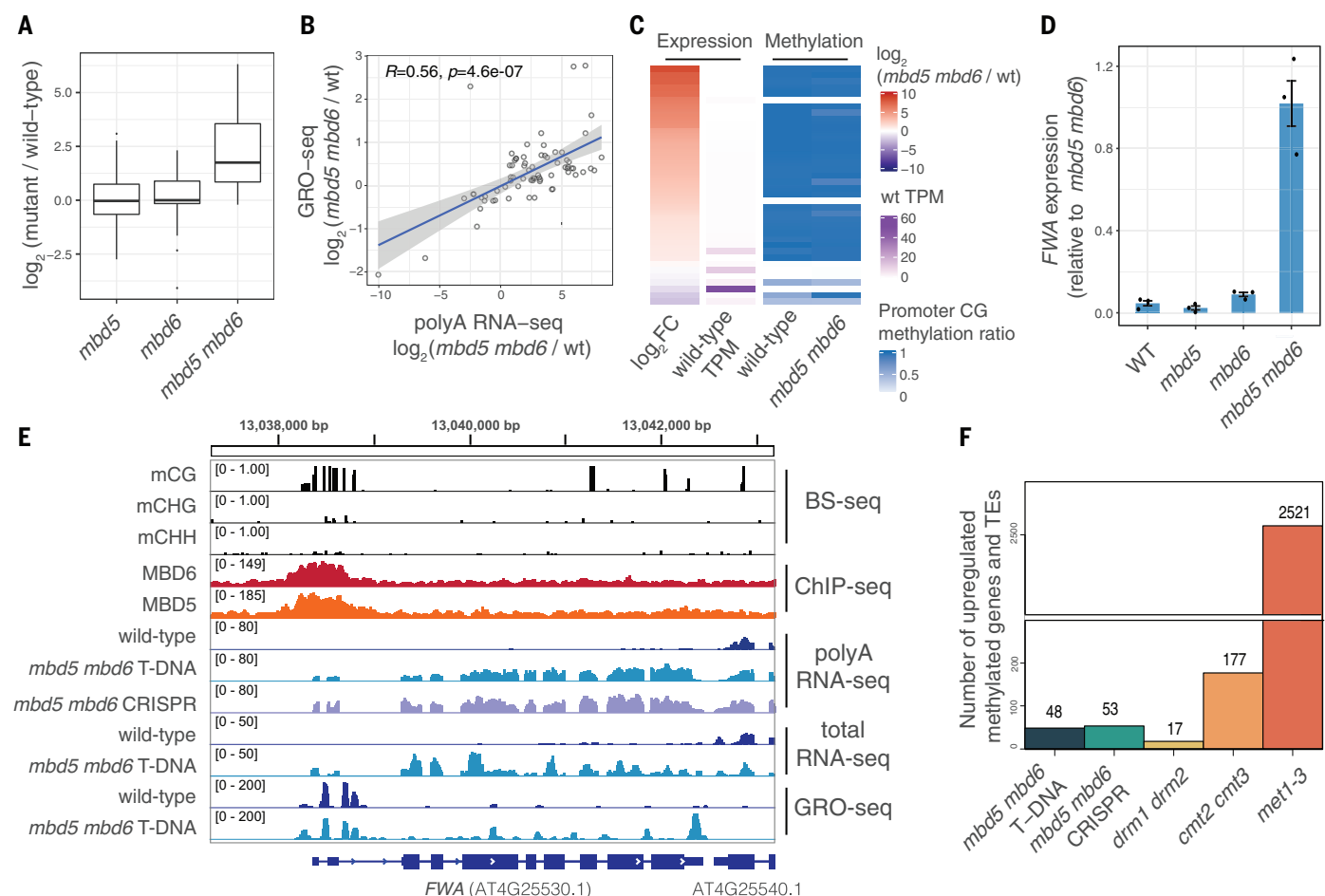
Fig. 2. MBD5 and MBD6 redundantly repress a subset of genes and transposons downstream of DNA methylation. (**A**) Boxplot of poly(A) RNA-seq for different mutants. Shown are the transcripts (genes and transposons) up-regulated in *mbd5 mbd6*. (**B**) Scatterplot comparing poly(A) RNA-seq with GRO-seq data at *mbd5 mbd6* T-DNA differential transcripts. *R* is the Spearman's rank correlation coefficient and *p* is the *P* value of the test that *R* = 0. Shaded area indicates 95% CI. (**C**) Heatmap of *mbd5 mbd6* T-DNA differential transcripts showing poly(A) RNA-seq and BS-seq data (average methylation ratio at 400-bp windows around the TSS). (**D**) Reverse transcription quantitative polymerase chain reaction (RT-qPCR) analysis of *FWA* expression normalized to *IPP2*. Dots indicate individual plants. Error bars indicate SEM. (**E**) Genome browser tracks at *FWA*. The GRO-seq enrichment at the *FWA* promoter likely corresponds to Pol V transcription. (**F**) Number of promoter methylated genes and TEs that are up-regulated in different mutants.

the number of derepressed transposons and the amplitude of derepression in *mbd5 mbd6* was much smaller than in *met1-3* (Fig. 2F and fig. S11), suggesting that MBD5 and MBD6 are not the only factors mediating repression downstream of DNA methylation.

To investigate the mechanism of action of MBD5 and MBD6, we performed immuno-precipitation–mass spectrometry (IP-MS) using the FLAG-tagged transgenic lines. Both proteins pulled down each other and three small heat shock proteins (ACD15.5, ACD21.4, and IDM3/LIL) that were previously found to interact with MBD5 and MBD7 (*13*). In addition, we detected an uncharacterized class C J-domain protein (AT5G37380) (*26, 27*) that we have named SILENZIO (SLN) (Fig. 3A and table S1). MBD5 and MBD6 also pulled down a smaller number of peptides of SUVH1, SUVH3, DNAJ1, and DNAJ2, which are components of a methyl-reader complex known to bind at RdDM sites

and up-regulate nearby protein-coding genes (*9, 28*).

We focused our further investigation on SILENZIO because of the recently described role of the J-domain proteins DNAJ1 and DNAJ2 in gene activation downstream of DNA methylation (*9, 28*). SILENZIO homologs were found to be present widely throughout the plant kingdom, but only the J-domain was conserved in animals (fig. S12). To determine whether SILENZIO was involved in gene silencing, we performed RNA-seq on an *sln* T-DNA mutant line. We found a strong correlation between the *sln* and the *mbd5 mbd6* RNA-seq data, with a similar extent of derepression of TEs and genes, including *FWA* (Fig. 3, B and C). We performed ChIP-seq with a complementing FLAG-tagged SLN line (fig. S13) and observed localization to the same sites as MBD5 and MBD6, but this localization was abolished in *mbd5 mbd6* mutants, suggest-

ing that SLN is recruited to chromatin by the methyl readers (Fig. 3, D and E, and fig. S14). Conversely, the MBD5 and MBD6 ChIP-seq signals were unaffected in *sln*, indicating that their recruitment to chromatin does not require SLN (Fig. 3, D and E, and fig. S14). Overall, these results suggest that SLN acts as a gene repressor downstream of MBD5 and MBD6.

To further test the role of SLN as a repressor, we created a fusion of SLN with ZF108, an artificial zinc finger that allows ectopic targeting of proteins to the *FWA* promoter (Fig. 3F) (*29, 30*). We transformed this fusion construct driven by the constitutive *UBIQUITIN10* promoter (*pUBQ10::ZF108-SLN*) into *fwa* epi-allele mutant plants (*24*) in which the *FWA* gene has heritably lost DNA methylation, leading to *FWA* over-expression and a late-flowering phenotype. Transgenic (T1) plants that expressed high levels of the fusion protein displayed down-regulation of *FWA*, thus supporting a role of
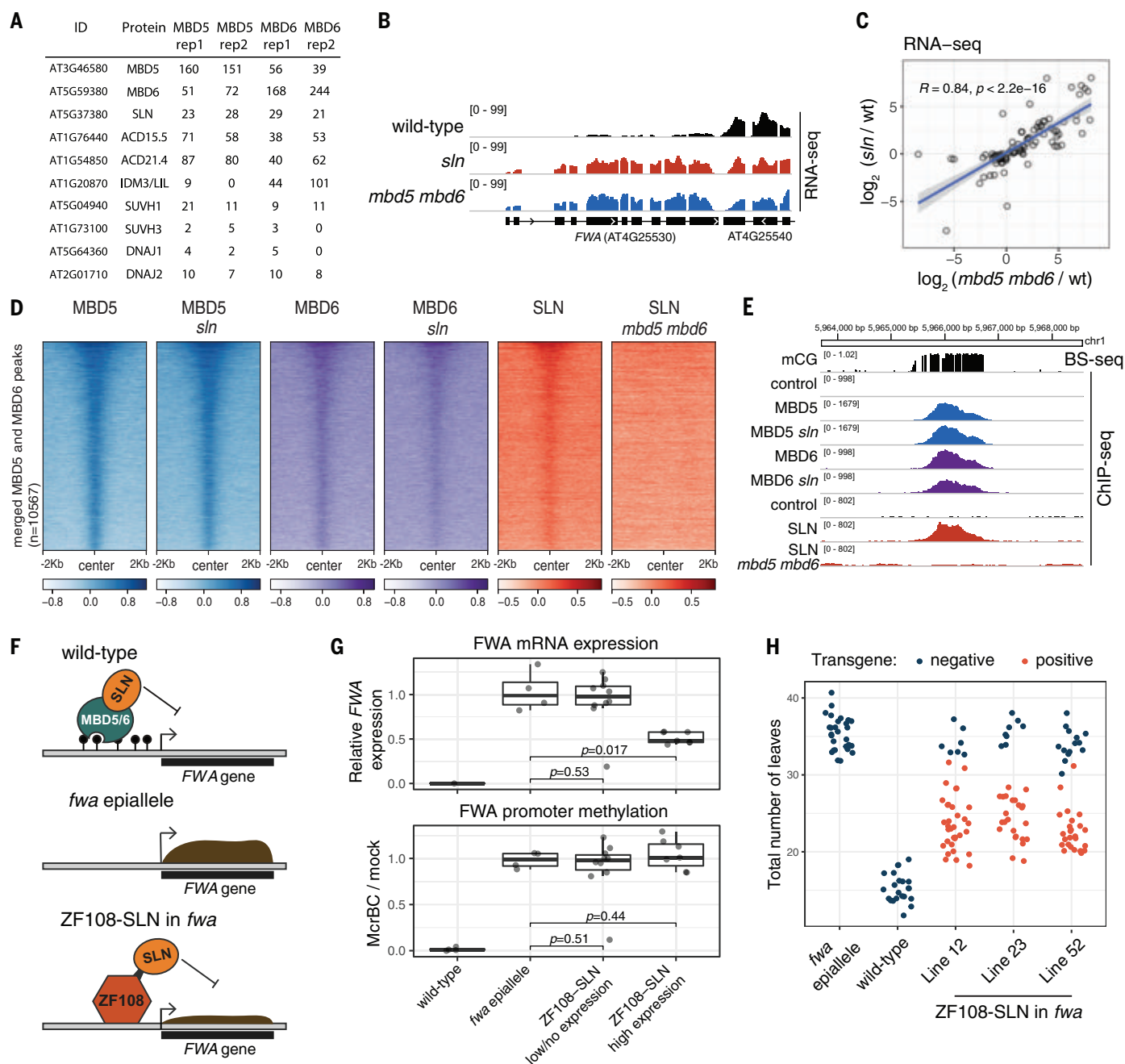
Fig. 3. SLN represses transcription downstream of MBD5 and MBD6.
(**A**) Spectral counts of proteins detected by IP-MS of FLAG-tagged MBD5 and MBD6. All proteins displayed were not detected in the no-FLAG negative control (see table S1). (**B**) RNA-seq data at *FWA*. (**C**) Scatterplot of the union of *mbd5 mbd6* CRISPR and *sln* differential transcripts. *R* is the Spearman's rank correlation coefficient and *p* is the *P* value of the test that *R* = 0. Blue line indicates the linear model fit. Shaded area indicates 95% CI. (**D**) Heatmap of ChIP-seq data (log$_2$-fold change over the no-FLAG control). (**E**) Example methylated site bound by MBD5, MBD6, and SLN in the indicated backgrounds.

(**F**) Diagram showing SLN's ectopic recruitment to unmethylated *FWA* through fusion to ZF108. (**G**) RT-qPCR analysis of *FWA* expression and McrBC-qPCR analysis of *FWA* promoter methylation in T1 lines expressing low or high levels of ZF108-SLN (Western blot in fig. S15A). Dots indicate individual plants. *p* is the *P* value of the Student's *t* test for each pair of groups. RT-qPCR data (normalized to *IPP2*) are relative to *fwa* epi-allele plants. (**H**) Flowering time (number of leaves produced before flowering) of segregating T2 populations from three transgenic lines expressing high levels of ZF108-SLN, comparing transgene-positive with null segregant (negative) plants.

SLN as transcriptional repressor (Fig. 3G and fig. S15A). *FWA* repression was not accompanied by promoter methylation (Fig. 3G and fig. S15B), demonstrating that SLN's ability to repress transcription can be uncoupled from DNA methylation. Indeed, in the T2 segregant population, the null segregants recovered *FWA* overexpression and the corresponding late flowering time (Fig. 3H and fig. S15C). ZF108 was designed to bind *FWA*, but it also binds to thousands of off-target sites in the genome (*30*), allowing us to examine gene expression changes at these sites by performing RNA-seq in the *pUBQ10::ZF108-SLN* lines. We observed that genes with a ZF108 peak near their
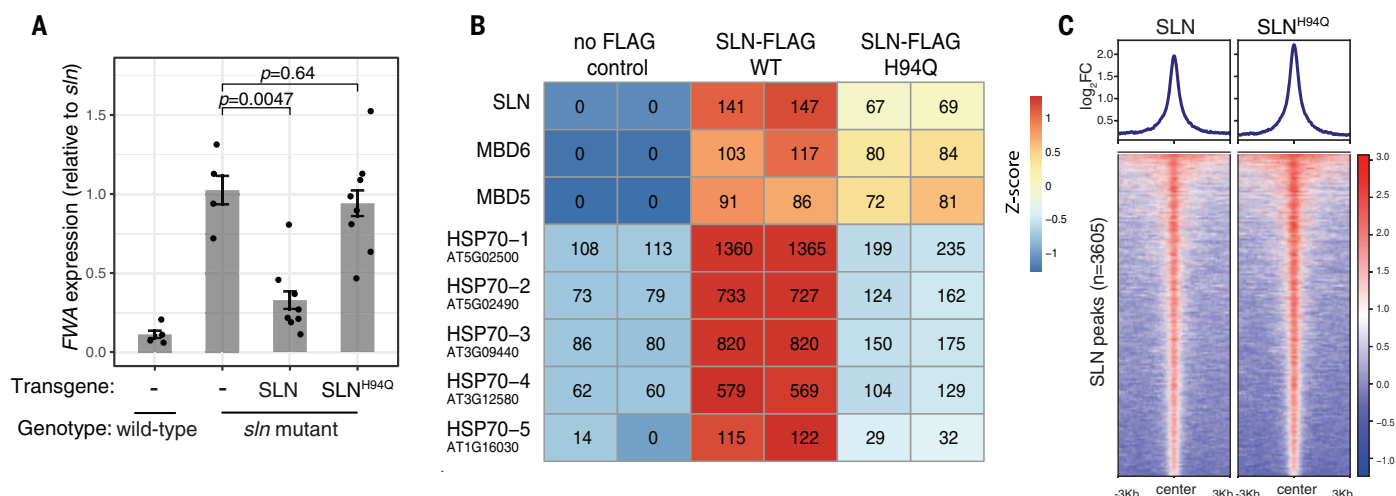
Fig. 4. SLN-silencing function requires the conserved HPD tripeptide. (**A**) RT-qPCR analysis of *FWA* expression (normalized to *IPP2*) in T1 lines expressing SLN or SLN^H94Q in the *sln* mutant background. *p* is the P value of the Student's *t* test for each pair of groups. Error bars indicate SEM. Dots indicate individual plants. (**B**) Spectral counts of proteins detected by IP-MS of wild-type and H94Q mutant SLN-FLAG (representative of two independent experiments; see table S1). (**C**) ChIP-seq of FLAG-tagged SLN and SLN^H94Q (log_2-fold change over the no-FLAG control).

promoter showed a tendency to be down-regulated (fig. S16), demonstrating that ectopic recruitment of SLN can repress many genes in addition to *FWA*.

IP-MS analysis of SLN-FLAG identified peptides corresponding to MBD5 and MBD6, as expected, but also showed a strong enrichment of five HEAT SHOCK PROTEIN 70 (HSP70) chaperones known to be constitutively expressed and localized in the nucleus (*31*) (Fig. 4B and table S1). Enrichment for HSP70s was also detected in the MBD5 and MBD6 IP-MS datasets and was lost in *sln* mutant plants (fig. S17 and table S1). This suggests that SLN mediates the interaction between the methyl readers and the HSP70s.

The canonical function of J-domain proteins is to bind clients, recruit HSP70 chaperones using a conserved HPD tripeptide, and stimulate the ATPase activity of HSP70s to increase their affinity for substrates. The HSP70-substrate interaction can induce folding, disaggregation, assembly, or disassembly of complexes involving client proteins (*32*). Mutating the histidine of the HPD tripeptide to glutamine can abrogate the J-domain–HSP70 interaction (*32*). To test whether SLN's binding to HSP70s was associated with its gene-silencing function, we generated an HPD mutant version of SLN by mutating the histidine to glutamine (SLN^H94Q) and transformed this into *sln* mutant plants. The SLN^H94Q mutant failed to rescue the de-repression of *FWA* and of the other genes and transposons, suggesting that the gene-silencing function of SLN requires the J-domain and HSP70 interaction (Fig. 4A and fig. S18, A to F). Indeed, IP-MS of SLN^H94Q showed greatly reduced enrichment of HSP70s, whereas the interaction with MBD5 and MBD6 was retained (Fig. 4B and table S1). Furthermore,

ChIP-seq enrichment of SLN on chromatin was not affected by the H94Q mutation (Fig. 4C and fig. S18, G and H). These results suggest that recruitment of SLN by the methyl readers may serve as a tether to bring the chaperone activity of SLN-HSP70s to CG-dense methylated chromatin to enforce gene silencing. The interaction between chaperones and their clients is often transient and difficult to detect by IP-MS (*32*), meaning that SLN might exert its repressive activity through recruitment, stabilization, or assembly of currently unknown repressive complexes or by targeted inhibition or disassembly of activators.

In conclusion, this work identifies a pathway that links DNA methylation to silencing of sites marked by CG methylation. The characterization of the methyl-binding proteins MBD5 and MBD6 shows that they likely act through a mechanism distinct from that of known MBD proteins in animals. The identification of the J-domain protein SILENZIO as a silencing effector further suggests that gene repression downstream of methylation is linked to chaperone activity, and this new pathway is likely to be conserved among divergent plant lineages.

REFERENCES AND NOTES

1. M. G. Goll, T. H. Bestor, *Annu. Rev. Biochem.* **74**, 481–514 (2005).
2. A. Zemach, I. E. McDaniel, P. Silva, D. Zilberman, *Science* **328**, 916–919 (2010).
3. T. Baubec, R. Ivánek, F. Lienert, D. Schübeler, *Cell* **153**, 480–492 (2013).
4. T. Shimbo, P. A. Wade, *Adv. Exp. Med. Biol.* **945**, 303–320 (2013).
5. L. D. Boxer et al., *Mol. Cell* **77**, 294–309.e9 (2020).
6. M. J. Lyst et al., *Nat. Neurosci.* **16**, 898–902 (2013).
7. X. Nan et al., *Nature* **393**, 386–389 (1998).
8. X. Zhang et al., *Cell* **126**, 1189–1201 (2006).
9. C. J. Harris et al., *Science* **362**, 1182–1186 (2018).
10. A. Zemach, G. Grafi, *Plant J.* **34**, 565–572 (2003).
11. A. Berg et al., *Nucleic Acids Res.* **31**, 5291–5304 (2003).
12. N. M. Springer, S. M. Kaeppler, *Plant Physiol.* **138**, 92–104 (2005).
13. D. Li et al., *eLife* **6**, e19893 (2017).
14. A. Zemach et al., *Plant Cell* **17**, 1549–1558 (2005).
15. F. Scebba et al., *Plant Mol. Biol.* **53**, 715–731 (2003).
16. M. Ito, A. Koike, N. Koizumi, H. Sano, *Plant Physiol.* **133**, 1747–1754 (2003).
17. S. B. Preuss et al., *Mol. Cell* **32**, 673–684 (2008).
18. J. A. Law, S. E. Jacobsen, *Nat. Rev. Genet.* **11**, 204–220 (2010).
19. Y. Liu, X. Zhang, R. M. Blumenthal, X. Cheng, *Trends Biochem. Sci.* **38**, 177–183 (2013).
20. H. W. Gabel et al., *Nature* **522**, 89–93 (2015).
21. A. Bartlett et al., *Nat. Protoc.* **12**, 1659–1672 (2017).
22. H. Stroud, M. V. C. Greenberg, S. Feng, Y. V. Bernatavichute, S. E. Jacobsen, *Cell* **152**, 352–364 (2013).
23. M. E. Potok et al., *Nat. Commun.* **10**, 3352 (2019).
24. W. J. J. Soppe et al., *Mol. Cell* **6**, 791–802 (2000).
25. H. Stroud et al., *Nat. Struct. Mol. Biol.* **21**, 64–72 (2014).
26. A. Finka, R. U. H. Mattoo, P. Goloubinoff, *Cell Stress Chaperones* **16**, 15–31 (2011).
27. V. B. V. Rajan, P. D'Silva, *Funct. Integr. Genomics* **9**, 433–446 (2009).
28. Q. Q. Zhao, R. N. Lin, L. Li, S. Chen, X. J. He, *J. Integr. Plant Biol.* **61**, 120–139 (2019).
29. L. M. Johnson et al., *Nature* **507**, 124–128 (2014).
30. J. Gallego-Bartolomé et al., *Cell* **176**, 1068–1082.e19 (2019).
31. L. Leng et al., *J. Plant Res.* **130**, 349–363 (2017).
32. R. Rosenzweig, N. B. Nillegoda, M. P. Mayer, B. Bukau, *Nat. Rev. Mol. Cell Biol.* **20**, 665–680 (2019).

CORONAVIRUS

Face masks effectively limit the probability of SARS-CoV-2 transmission

Yafang Cheng[1]*†, Nan Ma[2]†, Christian Witt[3], Steffen Rapp[4], Philipp S. Wild[4], Meinrat O. Andreae[1,5,6], Ulrich Pöschl[1], Hang Su[7,1]*

Airborne transmission by droplets and aerosols is important for the spread of viruses. Face masks are a well-established preventive measure, but their effectiveness for mitigating severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) transmission is still under debate. We show that variations in mask efficacy can be explained by different regimes of virus abundance and are related to population-average infection probability and reproduction number. For SARS-CoV-2, the viral load of infectious individuals can vary by orders of magnitude. We find that most environments and contacts are under conditions of low virus abundance (virus-limited), where surgical masks are effective at preventing virus spread. More-advanced masks and other protective equipment are required in potentially virus-rich indoor environments, including medical centers and hospitals. Masks are particularly effective in combination with other preventive measures like ventilation and distancing.

Airborne transmission is one of the main pathways for the transmission of respiratory viruses, including the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (1). Wearing face masks has been widely advocated to mitigate transmission. Masks are thought to protect people in two ways: (i) source control, reducing the emission and spread of respiratory viruses through airborne droplets and aerosols, and (ii) wearer protection, reducing the inhalation of airborne respiratory viruses.

The effectiveness of masks, however, is still under debate. Compared with N95 or FFP2 respirators, which have very low particle penetration rates (~5%), surgical and similar masks exhibit higher and more variable penetration rates (~30 to 70%) (2, 3). Given the

large number of particles emitted upon respiration and especially upon sneezing or coughing (4), the number of respiratory particles that may penetrate masks is substantial, which is one of the main reasons for doubts about their efficacy in preventing infections. Moreover, randomized clinical trials have shown inconsistent or inconclusive results, with some studies reporting only a marginal benefit or no effect of mask use (5, 6). Thus, surgical and similar masks are often considered to be ineffective. On the other hand, observational data show that regions or facilities with a higher percentage of the population wearing masks have better control of COVID-19 (7–9). So how are we to explain these contrasting results and apparent inconsistencies?

In this work, we develop a quantitative model of airborne virus exposure that can explain these contrasting results and provide a basis for quantifying the efficacy of face masks. We show that mask efficacy strongly depends on airborne virus abundance. On the basis of direct measurements of SARS-CoV-2 in air samples and population-level infection probabilities, we find that the virus abundance in most environments is sufficiently low for masks to be effective in reducing airborne transmission.

When evaluating the effectiveness of masks, we want to understand and quantify their effect on the infection probability, P_{inf}. Assuming that every inhaled single virus (virion) has

the same chance to infect a person, P_{inf} can be calculated by a single-hit model of infection

$$P_{inf} = 1 - (1 - P_{single})^{N_v} \qquad (1)$$

where P_{single} represents the infection probability for a single virus and N_v represents the total number of viruses to which the person is exposed (10). For airborne transmission, the infection probability P_{inf} for a given time period can be plotted as a function of inhaled virus number, N_v.

Figure 1 illustrates the dependence of P_{inf} on N_v based on the single-hit model (Eq. 1) and scaled by the median infectious dose $ID_{v,50}$ at which the probability of infection is 50% (10). It shows a highly nonlinear sensitivity of P_{inf} to changes in N_v. Accordingly, the same percentage of change of N_v may lead to different changes in P_{inf} depending on the absolute level of N_v. In a virus-rich regime, where N_v is much higher than $ID_{v,50}$ (Fig. 1, A and B), P_{inf} is close to unity and is not sensitive to changes in N_v. In this case, wearing a mask may not suffice to prevent infection. In a virus-limited regime, where N_v is close to or lower than $ID_{v,50}$, however, P_{inf} strongly varies with N_v, and reducing N_v by wearing a mask will lead to a substantial reduction in the infection probability (Fig. 1, C and D). Thus, we need to determine the regime of airborne virus abundance to understand mask efficacy.

Respiratory particles, including aerosol particles and larger droplets, can carry viruses and are often used to visualize the transmission of airborne viruses (4). Taking a representative average of respiratory activity (11), we find that a person typically emits a total number of ~3 × 10^6 particles during a 30-min period (supplementary text, section S1.1). This very large number implies that indoor environments are usually in a respiratory particle–rich regime. Surgical masks with particle collection efficiencies of ~50% cannot prevent the release of millions of particles per person and their inhalation by others (see green dots in Fig. 1, B and D). In other words, the human-emitted respiratory particle number is so high that we cannot avoid inhaling particles generated by another person, even when wearing a surgical mask. If every respiratory particle were to contain one or more viruses, indoor environments would often be in a virus-rich regime

[1]Max Planck Institute for Chemistry, 55128 Mainz, Germany. [2]Institute for Environmental and Climate Research, Jinan University, Guangzhou 511443, China. [3]Department of Outpatient Pneumology and Institute of Physiology, Charité Universitätsmedizin Berlin, Campus Charité Mitte, 10117 Berlin, Germany. [4]University Medical Center of the Johannes Gutenberg-University Mainz, 55131 Mainz, Germany. [5]Scripps Institution of Oceanography, University of California San Diego, La Jolla, CA 92093, USA. [6]Department of Geology and Geophysics, King Saud University, 11451 Riyadh, Saudi Arabia. [7]State Environmental Protection Key Laboratory of Formation and Prevention of Urban Air Pollution Complex, Shanghai Academy of Environmental Sciences, Shanghai 200233, China.
*Corresponding author. Email: h.su@mpic.de (H.S.); yafang.cheng@mpic.de (Y.C.)
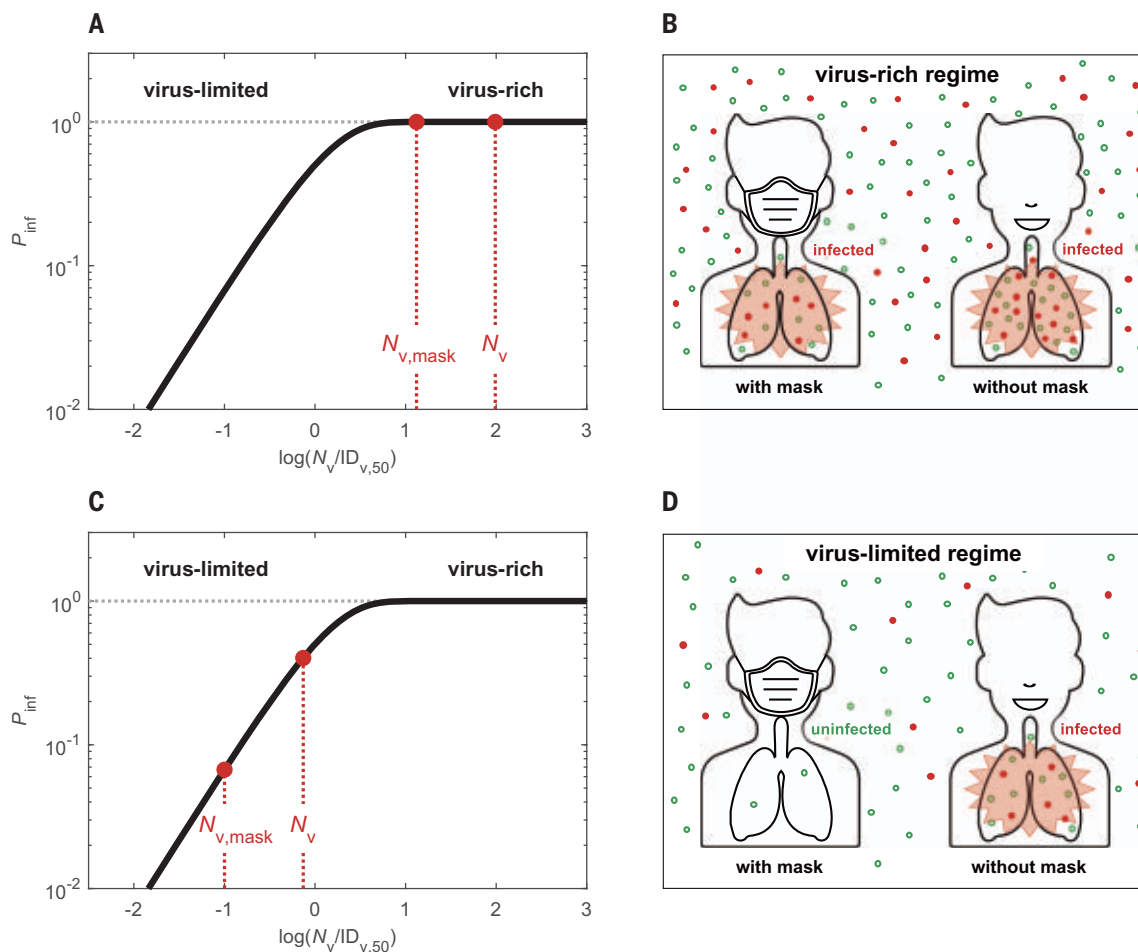†These authors contributed equally to this work.

Fig. 1. Schematic illustration of different regimes of abundance of respiratory particles and viruses. (**A** to **D**) The solid curves represent the infection probability (P_{inf}) as a function of inhaled virus number (N_v) scaled by median infectious dose $ID_{v,50}$ at which P_{inf} = 50%. In the virus-rich regime [(A) and (B)], the concentration of airborne viruses is so high that both the numbers of viruses inhaled with and without masks ($N_{v,mask}$, N_v) are much higher than $ID_{v,50}$, and P_{inf} remains close to ~1 even if masks are used. In the virus-limited regime [(C) and (D)], N_v and $N_{v,mask}$ are close to or lower than $ID_{v,50}$, and P_{inf} decreases substantially when masks are used, even if the masks cannot prevent the inhalation of all respiratory particles. In (B) and (D), the red dots represent respiratory particles containing viruses, and the open green circles represent respiratory particles without viruses. Man icon used in (B) and (D) was made by Tinu CA from www.freeicons.io, distributed under CC-BY 3.0.

because the median infectious dose $ID_{v,50}$ for respiratory diseases is typically on the order of a few tens to thousands of viruses (*12–14*).

But, does a respiratory particle–rich regime actually imply a respiratory virus–rich regime? To answer this question, we investigated characteristic virus distributions in both exhaled air samples and indoor air samples including coronaviruses (HCoV-NL63, -OC43, -229E, and -HKU1), influenza viruses (A and B), rhinoviruses, and SARS-CoV-2 (supplementary text, section S1). We find that usually just a minor fraction of exhaled respiratory particles contains viruses. In contrast to the high number of emitted respiratory particles, the number of viruses in 30-min samples of exhaled air ($N_{v,30,ex}$) are typically low, with mean values of ~53 for coronaviruses (HCoV-NL63, -OC43, -229E, and -HKU1), ~38 for influenza viruses (A and B), and ~96 for rhinoviruses (*11*) (supplementary text, section S1.2, and Fig. 2). Figure 2, A and B,

shows the infection probabilities obtained by inserting the number of exhaled viruses ($N_{v,30,ex}$) for the number of potentially inhaled viruses ($N_{v,30}$), assuming a characteristic infectious dose of $ID_{v,50}$ = 100 or 1000 viruses, respectively (*12–14*). For SARS-CoV-2 in various medical centers, we obtained mean values of $N_{v,30}$ in the range of ~1 to ~600 (*15–18*) (supplementary text, section S1.3), which correspond to P_{inf} values in the range of ~0.1% to 10% for $ID_{v,50}$ = 1000 and ~1% to 100% for $ID_{v,50}$ = 100. The wide range of $N_{v,30}/ID_{v,50}$ and P_{inf} values demonstrate that both virus-limited and virus-rich conditions can occur in indoor environments.

The high variabilities of $N_{v,30}$ and P_{inf} shown in Fig. 2, A and B, are consistent with the wide distribution of viral load observed in respiratory tract fluids (*19*) and need to be considered for estimating population-average infection probabilities, $P_{inf,pop}$ (supplementary text, section S4). For this purpose, we modeled N_v for

SARS-CoV-2 as lognormally distributed with standard deviations (σ) in the range of ~1 to 2 on the basis of recently reported distributions of the viral load of SARS-CoV-2 in respiratory fluids (*19*) (supplementary text, section S4). As shown in Fig. 2C, the population-average infection probabilities with σ > 0 are higher than in the case of uniform exposure (σ = 0) in the virus-limited regime at $P_{inf,pop}$ < ~50%. In other words, when the population-average infection probability is in the virus-limited regime with $P_{inf,pop,0}$ < 0.5 (Fig. 2C), a broader distribution (larger σ) implies an increase in the fraction of transmission events under virus-rich conditions (e.g., superspreader events), which leads to a reduction of overall mask efficacy.

The basic reproduction number for COVID-19 (R_0 ≈ 2 to 4) (*20*) can be related to a basic population-average infection probability, $P_{inf,pop,0}$, through $R_0 = P_{inf,pop,0} \times c \times d$ (*21*). With the average duration of infectiousness

Fig. 2. Infection probabilities and abundance regimes of SARS-CoV-2 and other respiratory viruses. (**A** and **B**) Individual infection probabilities (P_{inf}) plotted against inhaled virus number (N_v) scaled by characteristic median infectious doses of $ID_{v,50} = 100$ or 1000 viruses, respectively. The colored data points represent the mean numbers of viruses inhaled during a 30-min period in different medical centers in China, Singapore, and the US, according to measurement data of exhaled coronavirus, influenza virus, and rhinovirus numbers (blue circles) (*11*) and of airborne SARS-CoV-2 number concentrations (red symbols) (*15*–*18*), respectively. The error bars represent one geometric standard deviation. (**C**) Population-average infection probability ($P_{inf,pop}$) curves assuming lognormal distributions of N_v with different standard deviations of $\sigma = 0$, 1, and 2, respectively. The x axis represents the mean value of $\log(N_v/ID_{v,50})$. The shaded area indicates the level of basic population-average infection probability, $P_{inf,pop,0}$, for SARS-CoV-2, as calculated from the basic reproduction number for COVID-19 and estimated values of average duration of infectiousness and daily number of contacts.
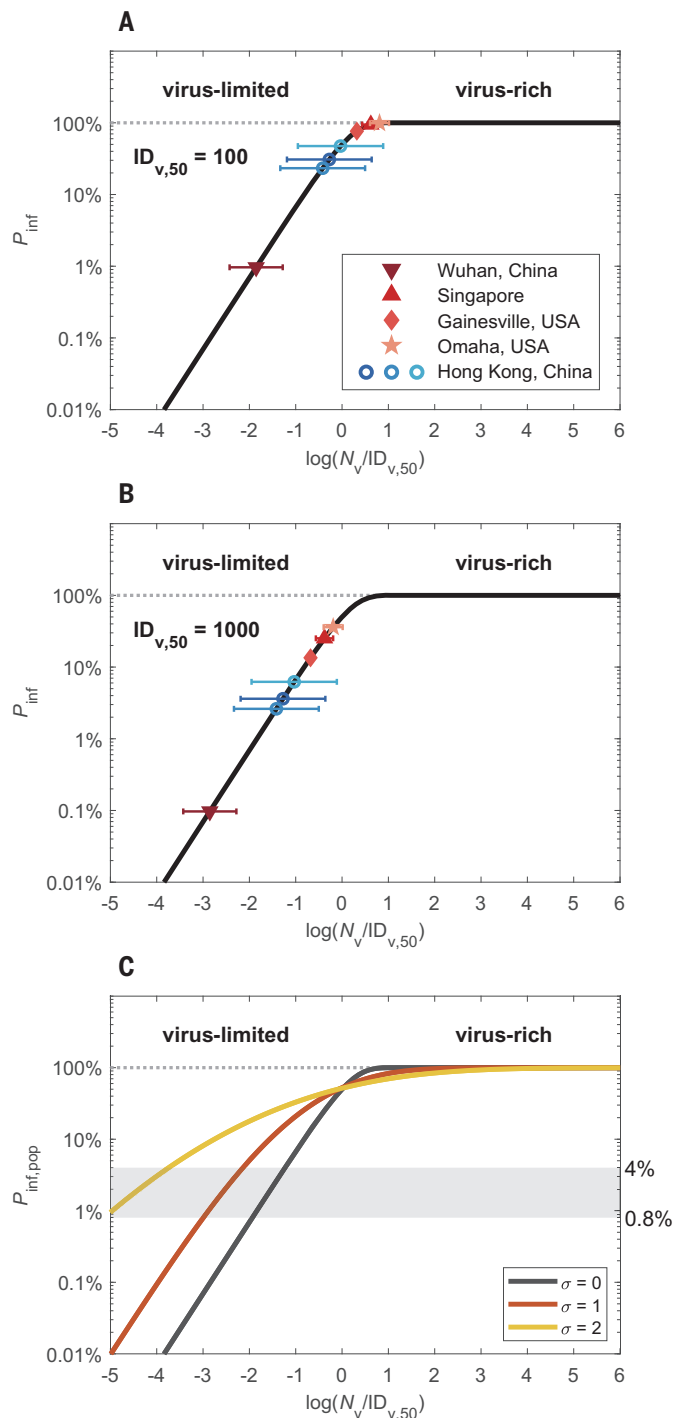
($d \approx 10$ days) and average daily numbers of human contacts ($c \approx 10$ to 25 contacts per day) (*22, 23*), we obtain estimates in the range of ~0.8% to ~4% for $P_{inf,pop,0}$, as indicated by the shaded area in Fig. 2C. The low levels of $P_{inf,pop,0}$ indicate a widespread prevalence of virus-limited conditions.

Different regimes of abundance are relevant not only for the distinction of respiratory particles and viruses, but also for different types of viruses. For example, viruses with higher transmissibility—i.e., those with higher loads and rates of emission and exhalation, greater environmental persistence, or lower $ID_{v,50}$—may result in a virus-rich regime and lead to higher basic reproduction numbers, as observed for measles and other highly infectious diseases. Our analysis shows that the levels of P_{inf} and R_0

can vary widely for different viruses. This means that aerosol transmission does not necessarily lead to a measles-like high R_0 and that relatively low values of P_{inf} and R_0 do not rule out airborne transmission. On the basis of the scaling with $ID_{v,50}$, the curves shown in Figs. 1 to 3 can easily be applied to assess the efficacy of masks and other preventive measures against new and more-infectious mutants of SARS-CoV-2, such as B.1.1.7.

Figure 3 illustrates how the efficacies of surgical masks and N95 or FFP2 masks vary between virus-limited and virus-rich conditions when masks are worn only by infectious persons (source control), only by susceptible persons (wearer protection), or by all persons (universal masking). In Fig. 3A, the population-average infection probability in the case of surgical mask use ($P_{inf,pop,mask}$) is plotted against the infection probability without masks ($P_{inf,pop}$). It shows that surgical masking achieves large reductions in infection probability when the maskless infection probability is low but increasingly smaller reductions when the maskless infection probability is high. Figure 3B shows the corresponding mask efficacy, i.e., the percentage reduction of infection probability [$\Delta P_{inf,pop}/P_{inf,pop} = (P_{inf,pop} - P_{inf,pop,mask})/P_{inf,pop}$]. It decreases slowly with increasing $P_{inf,pop}$ in the virus-limited regime, exhibits a steep decrease upon transition into the virus-rich regime as $P_{inf,pop}$ approaches unity, and goes to zero at $P_{inf,pop} = 1$. Figure 3, C and D, shows equivalent plots for N95 or FFP2 masks.

Figure 3 illustrates that source control alone is more effective than wearer protection alone but that universal masking is the most effective. This is because masks are more effective in removing larger particles (Fig. 4), and freshly generated respiratory particles are usually largest at the source, shrinking upon evaporation in indoor air (*20*). Figure 3 accounts only for airborne transmission of viruses. When considering other forms of transmission, the relative importance of source control can be even higher (supplementary text, section S5) (*20*).

The nonlinear dependence of mask efficacy on infection risk differs from the assumption that the percentage change of infection probability as a result of mask use would be proportional to the percentage change of inhaled particle number (*20*). Under this assumption, wearing a mask would have the same effect on the transmission of a virus disease at any level of infection probability. Our analysis, however, shows that the efficacy of face masks depends strongly on the level of infection probability and virus abundance: Masks reduce the infection probability by as much as their filter efficiency for respiratory particles in the virus-limited regime but much less in the virus-rich regime (Fig. 3). Accordingly, experimental investigations may find low mask efficacies when they are performed under virus-rich conditions.

Fig. 3. Reduction of airborne transmission by face masks worn by infectious persons only (source control), by susceptible persons only (wearer protection), or by all persons (universal masking). (**A** and **B**) Population-average infection probability in case of mask use ($P_{inf,pop,mask}$) plotted against infection probability without face masks ($P_{inf,pop}$) (A) and corresponding mask efficacy—i.e., relative reduction of infection probability, $\Delta P_{inf,pop}/P_{inf,pop}$—plotted against $P_{inf,pop}$ for surgical masks (B). (**C** and **D**) Same as (A) and (B) but for N95 or FFP2 masks; plots with linear scaling are shown in fig. S8. The lines represent the results obtained for source control (red line), wearer protection (yellow line), and the combination of both measures, i.e., universal masking, (blue line) in a population where the virus exposure is lognormally distributed with a standard deviation of $\sigma = 1$ (supplementary text, section S5). The shaded areas indicate the level of basic population-average infection probability, $P_{inf,pop,0}$, corresponding to the basic reproduction number for COVID-19.
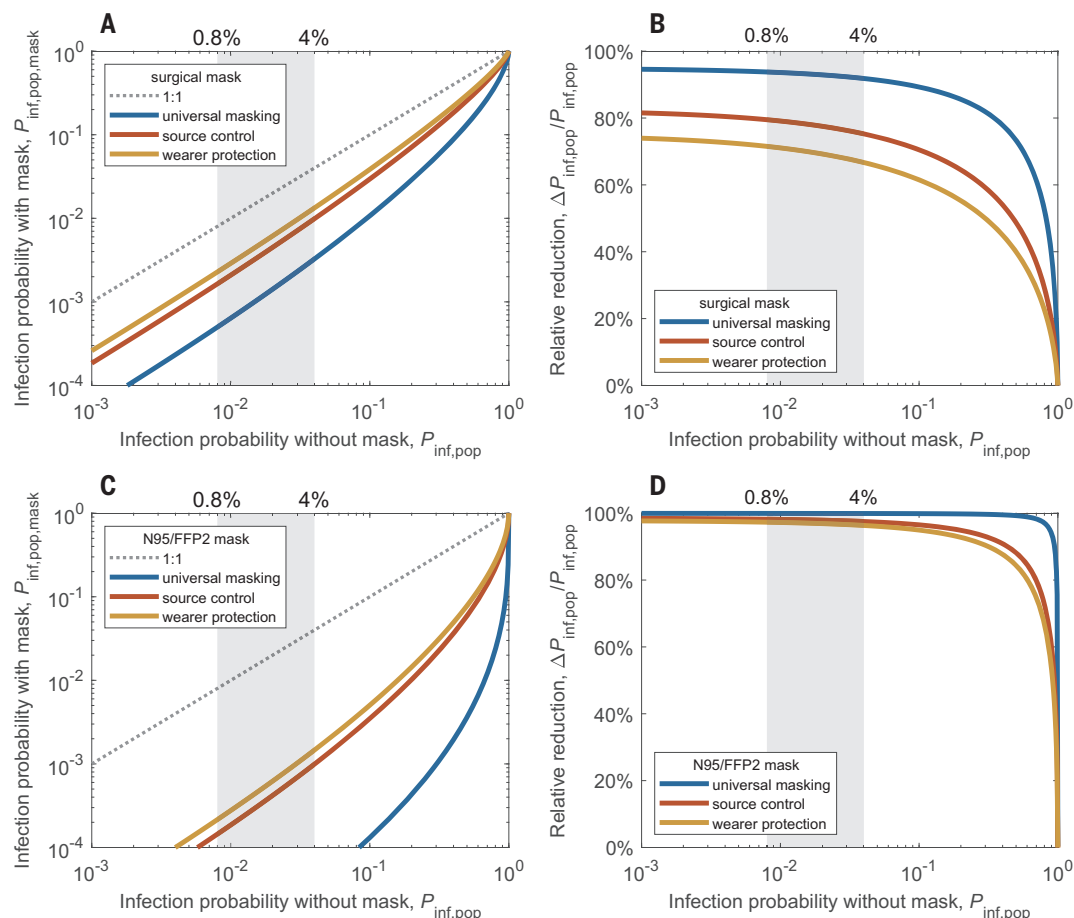
Together with other influencing factors, like consistent and correct mask use (supplementary text, section S7.3), changes between virus-rich and virus-limited conditions may contribute to divergent results reported from laboratory studies and randomized controlled trials in different environments (20) (supplementary text, section S8). Notably, the increasing effectiveness of mask use at low virus abundance implies synergistic effects of combining masks with other preventive measures that reduce the airborne-virus concentration, such as ventilation and social distancing. For example, ventilation can change an environment from virus-rich to virus-limited conditions, which may be particularly important for medical centers with relatively high SARS-CoV-2 abundances (Fig. 2 and supplementary text, section S6). On the other hand, not only the efficacy of face masks but also the efficacy of distancing may be reduced in virus-rich environments (supplementary text, section S6). The more measures that are used, the more effective each measure will be in containing the virus transmission. If the inhaled dose may also affect the severity of infections (14), as is currently being debated (24), masks may still be useful even if the reduced dose still leads to an infection.

Figure 4 shows the size distribution of respiratory particles emitted by different human activities (25–27). Aerosols are physically defined as airborne solid or liquid particles with diameters smaller than 100 µm, which can remain suspended over extended periods of time. In medical studies, however, a threshold diameter of 5 µm has often been used to distinguish between a so-called aerosol mode and a so-called droplet mode. Our analysis of measurement data from exhaled and ambient air samples indicates that the so-called aerosol mode (<5 µm) contains more viruses than the so-called droplet mode (>5 µm) (11), although the latter comprises a larger volume of liquid emitted from the respiratory tract (tables S1 and S2). This may be explained by the following mechanisms: a higher viral load occurring in the lower respiratory tract where the smaller aerosol particles are generated (28) or an enrichment of organic surfactants and viruses upon the generation of smaller aerosol particles (29). Enrichment of viruses in the aerosol mode can enhance their transmission because smaller particles remain suspended for a longer time, which leads to stronger accumulation and dispersion in the air. This may cause higher airborne virus concentrations,

inhaled virus numbers, and infection risks, especially in densely occupied rooms with poor ventilation and long periods of exposure. Moreover, small aerosol particles have a higher penetration rate and higher probability of reaching the lower respiratory tract (figs. S5 and S6).

Our analysis was focused on respiratory particles and droplets with diameters smaller than 100 µm [traditional physical definition of aerosols (30)]. Because of rapid gravitational settling, respiratory droplets larger than 100 µm are removed from the air in seconds, but they may still reach the upper respiratory tract of persons in close contact and cause infections by carrying large numbers of viruses in their very large liquid volume. For example, a single 1-mm droplet may carry as many as ~50,000 viruses in the case of a viral load of 10^8 per milliliter of respiratory fluid, which is realistic and higher than the estimated infectious dose for SARS-CoV-2 (14). Such large droplets, however, are very efficiently (~100%) removed even by simple masks (Fig. 4 and supplementary text, section S3), which further emphasizes the importance and efficacy of face masks for preventing infections. Because of the strong size dependence, and to avoid ambiguities,
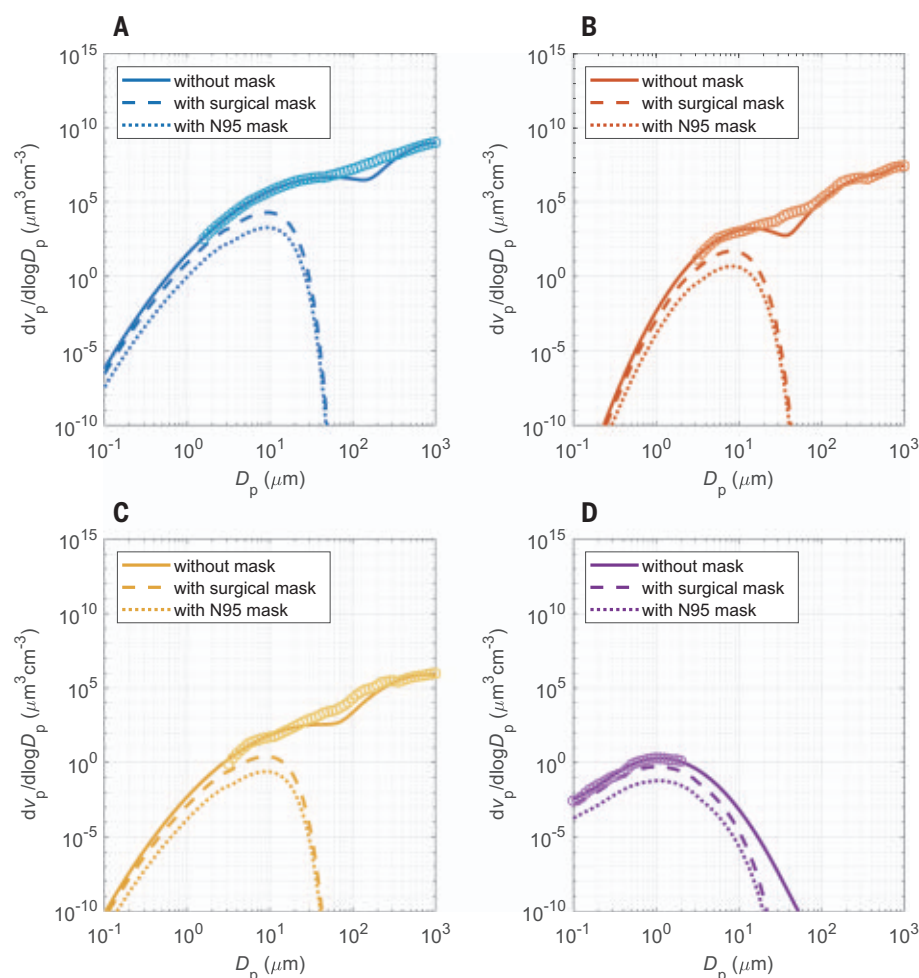
Fig. 4. Volume size distributions of respiratory particles emitted during different respiratory activities with and without masks. (**A** to **D**) Distributions for sneezing (A), coughing (B), speaking (C), and breathing (D). The open circles are measurement data obtained without masks, and the solid lines are bi- or trimodal fits to the measurement data (*25–27*) (supplementary text, section S1.1). The dashed and dotted lines were obtained by scaling with the filter efficiency curves of surgical masks and of N95 or FFP2 masks, respectively (supplementary text, section S3). The symbols v_p and D_p represent the volume concentration and diameter of respiratory particles, respectively, and $dv_p/d\log D_p$ represents the volume distribution function (supplementary text, section S1.1).

we suggest that diameter range should be explicitly specified when discussing airborne transmission by fine respiratory aerosol particles or larger droplets.

Our results have important implications for understanding and communicating preventive measures against the transmission of airborne viruses, including SARS-CoV-2. When people see images or videos of millions of respiratory particles exhaled by talking or coughing, they may be afraid that simple masks with limited filtration efficiency (e.g., 30 to 70%) cannot really protect them from inhaling these particles. However, as only few respiratory particles contain viruses and most environments are in a virus-limited regime, wearing masks can keep the number of inhaled viruses in a low-P_{inf} regime and can explain the observed efficacy of face masks in preventing the spread of COVID-19. However, unfavorable conditions and the large variability of viral loads may lead to a virus-rich regime in certain indoor environments, such as medical centers treating COVID-19 patients. In such environments, high-efficiency masks and additional protective measures like efficient ventilation should be used to keep the infection risk low. The nonlinear dependence of mask efficacy on airborne virus concentration—i.e., the higher mask efficacy at lower virus abundance—also highlights the importance of combining masks with other preventive measures. Effective ventilation and social distancing will reduce ambient virus concentrations and increase the effectiveness of face masks in containing the virus transmission. Moreover, high compliance and correct use of masks is important to ensure the effectiveness of universal mask-ing in reducing the reproduction number for COVID-19 (supplementary text, section S7.3, and fig. S11) (*20*).

REFERENCES AND NOTES

1. L. Morawska, D. K. Milton, *Clin. Infect. Dis.* **71**, 2311–2313 (2020).
2. S. A. Grinshpun *et al.*, *J. Occup. Environ. Hyg.* **6**, 593–603 (2009).
3. T. Oberg, L. M. Brosseau, *Am. J. Infect. Control* **36**, 276–282 (2008).
4. L. Bourouiba, *JAMA* **323**, 1837–1838 (2020).
5. A. E. Aiello *et al.*, *J. Infect. Dis.* **201**, 491–498 (2010).
6. H. Bundgaard *et al.*, *Ann. Intern. Med.* **174**, 335–343 (2021).
7. D. K. Chu *et al.*, *Lancet* **395**, 1973–1987 (2020).
8. R. Zhang, Y. Li, A. L. Zhang, Y. Wang, M. J. Molina, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 14857–14863 (2020).
9. J. T. Brooks, J. C. Butler, R. R. Redfield, *JAMA* **324**, 635–637 (2020).
10. C. N. Haas, J. B. Rose, C. P. Gerba, *Quantitative Microbial Risk Assessment* (Wiley, 2014).
11. N. H. L. Leung *et al.*, *Nat. Med.* **26**, 676–680 (2020).
12. T. Watanabe, T. A. Bartrand, M. H. Weir, T. Omura, C. N. Haas, *Risk Anal.* **30**, 1129–1138 (2010).
13. A. Roberts *et al.*, *J. Virol.* **79**, 503–511 (2005).
14. M. Imai *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 16587–16595 (2020).
15. P. Y. Chia *et al.*, *Nat. Commun.* **11**, 2800 (2020).
16. J. L. Santarpia *et al.*, *Sci. Rep.* **10**, 12732 (2020).
17. J. A. Lednicky *et al.*, *Aerosol Air Qual. Res.* **20**, 1167–1171 (2020).
18. Y. Liu *et al.*, *Nature* **582**, 557–560 (2020).
19. D. Jacot, G. Greub, K. Jaton, O. Opota, *Microbes Infect.* **22**, 617–621 (2020).
20. J. Howard *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2014564118 (2021).
21. P. van den Driessche, J. Watmough, in *Mathematical Epidemiology*, F. Brauer, P. van den Driessche, J. Wu, Eds. (Springer, 2008), pp. 159–178.
22. M. A. Johansson *et al.*, *JAMA Netw. Open* **4**, e2035057 (2021).
23. J. Zhang *et al.*, *Science* **368**, 1481–1486 (2020).
24. A. L. Rasmussen, K. Escandón, S. V. Popescu, *N. Engl. J. Med.* **383**, 2092–2094 (2020).
25. C. Y. H. Chao *et al.*, *J. Aerosol Sci.* **40**, 122–133 (2009).
26. J. P. Duguid, *Epidemiol. Infect.* **44**, 471–479 (1946).
27. H. Holmgren, E. Ljungström, A.-C. Almstrand, B. Bake, A.-C. Olin, *J. Aerosol Sci.* **41**, 439–446 (2010).
28. B. Bake, P. Larsson, G. Ljungkvist, E. Ljungström, A. C. Olin, *Respir. Res.* **20**, 8 (2019).
29. S. M. Burrows *et al.*, *Atmos. Chem. Phys.* **14**, 13601–13629 (2014).
30. K. A. Prather *et al.*, *Science* **370**, 303–304 (2020).
31. H. Su, Face masks effectively limit the probability of SARS-CoV-2 transmission [data set], version 1.0, Open Research Data Repository of the Max Planck Society (2021); https://dx.doi.org/10.17617/3.5d.

CATALYSIS

First-principles design of a single-atom–alloy propane dehydrogenation catalyst

Ryan T. Hannagan[1]†, Georgios Giannakakis[2]†, Romain Réocreux[3]†, Julia Schumann[3,4,5], Jordan Finzel[6], Yicheng Wang[1], Angelos Michaelides[4,5], Prashant Deshlahra[2], Phillip Christopher[6], Maria Flytzani-Stephanopoulos[2], Michail Stamatakis[3]*, E. Charles H. Sykes[1]*

The complexity of heterogeneous catalysts means that a priori design of new catalytic materials is difficult, but the well-defined nature of single-atom–alloy catalysts has made it feasible to perform unambiguous theoretical modeling and precise surface science experiments. Herein we report the theory-led discovery of a rhodium-copper (RhCu) single-atom–alloy catalyst for propane dehydrogenation to propene. Although Rh is not generally considered for alkane dehydrogenation, first-principles calculations revealed that Rh atoms disperse in Cu and exhibit low carbon-hydrogen bond activation barriers. Surface science experiments confirmed these predictions, and together these results informed the design of a highly active, selective, and coke-resistant RhCu nanoparticle catalyst that enables low-temperature nonoxidative propane dehydrogenation.

D espite their widespread prevalence in industry, the improvement of heterogeneous catalysts is usually an empirical process, and examples of rational design remain extremely rare (1, 2). Rational design is impeded in part by the enormous complexity of most heterogeneous catalysts, which often consist of metal nanoparticles (NPs) that expose a wide range of active sites with which the reactants can interact and react (3). Metal NPs can also restructure under reaction conditions (4). One approach to reduce the complexity of metal NPs is to reduce the active site to a single atom (5, 6). However, the heterogeneity of the oxide support and the adaptive nature of the local environment make rational design of even these materials challenging (7, 8).

Single-atom alloys (SAAs), a subclass of single-atom catalysts, consist of a reactive dopant metal atomically dispersed in a second metal host (9). Many SAAs are active, selective, and robust for a wide variety of reactions because the thermodynamics of mixing favor the atomic dispersion of the dopant in the more inert but more selective host (10). Furthermore, the well-defined nature of the active sites in SAAs, which have been atomically resolved in both model and NP systems, makes it possible to unambiguously model their behavior with theory (11–13).

[1]Department of Chemistry, Tufts University, Medford, MA 02155, USA. [2]Department of Chemical and Biological Engineering, Tufts University, Medford, MA 02155, USA. [3]Thomas Young Centre and Department of Chemical Engineering, University College London, Roberts Building, Torrington Place, London WC1E 7JE, UK. [4]Yusuf Hamied Department of Chemistry, University of Cambridge, Lensfield Road, CB2 1EW Cambridge, UK. [5]Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK. [6]Department of Chemical Engineering, University of California, Santa Barbara, CA 93106, USA.
*Corresponding author. Email: m.stamatakis@ucl.ac.uk (M.S.); charles.sykes@tufts.edu (E.C.H.S.)
†These authors contributed equally to this work.

Guided by theoretical predictions, we present the rational design of a RhCu SAA catalyst for nonoxidative propane dehydrogenation to two valuable compounds: propene and hydrogen (14). Our theoretical screening of SAA combinations for this reaction began by examining C–H activation of methane, which is the simplest probe molecule for establishing trends in C–H activation. We used density functional theory (DFT) to determine (i) the activation energy of the first C–H scission and (ii) the segregation energy, which reflects the propensity of the dopant atom to migrate from the surface layer to the bulk of the host metal (see computational details in the supplementary materials). Of all SAA combinations screened, the Rh-doped alloys have the lowest C–H activation energies (Fig. 1A and figs. S1 to S3). In fact, these C–H activation energies are similar to those of pure transition metals (Pt and Rh) and the industrially relevant Pt₃Sn(111) (2 × 2) intermetallic alloy (Fig. 1 and fig. S3) (15). However, of the three Rh-doped SAAs, RhCu (111) has the most favorable mixing enthalpy and segregation energy when reaction intermediates are present (Fig. 1A and figs. S1 and S2) (16). Specific to the mixing enthalpy, the energy change associated with isolated Rh atoms in a Cu(111) surface forming a dimer or a trimer is positive (>0.16 eV), indicating that the thermodynamic stability of single Rh atoms is greater than that of Rh clusters in a Cu host, which makes RhCu a promising SAA catalyst (16).

Comparison of the geometric and electronic structures of adsorbed methane on the SAA provided insight into the magnitude of the C–H activation barriers (Fig. 1B and fig. S4). On the RhCu(111) SAA surface, methane adsorbed 0.35 Å closer to the dopant atom than on pure Cu(111) (tables S1 and S2). This shorter distance increased the charge density difference Δρ($\mathbf{r}$) upon adsorption (Fig. 1B and fig. S4) by moving part of the electronic density

from the C–H internuclear region to the H···Rh internuclear region. This electronic redistribution is prominent on the RhCu(111) SAA and is observed on pure Pt(111) but is negligible on Cu(111). For the former two materials, this electronic redistribution is indicative of the preactivation of methane upon adsorption (17).

After identifying the RhCu(111) SAA as the most promising candidate for C–H activation, we calculated the energy profile for the full dehydrogenation of methane to atomic carbon on the RhCu(111) SAA, the bare Cu(111) surface, and Pt(111), one of the most effective metals for C–H activation reactions (Fig. 1C and fig. S3). Notably, the single Rh atom in the RhCu(111) SAA had an almost identical activation energy to that of pure Pt(111). However, unlike pure Pt(111), which stabilized CH₂(ads), CH(ads), and C(ads) intermediates that lead to coking, the RhCu(111) SAA inherited much of the coke resistance of pure Cu(111), with coke formation
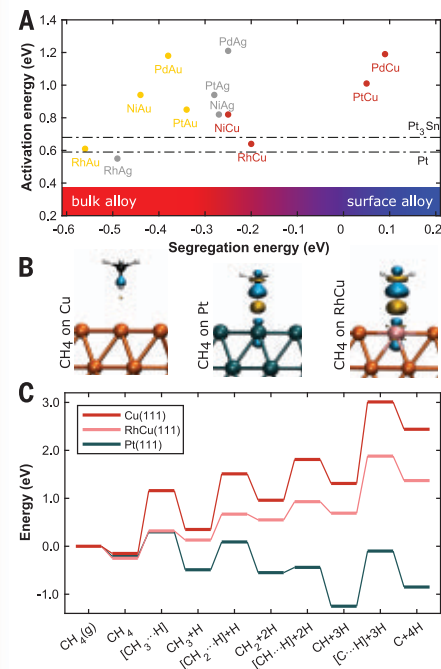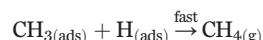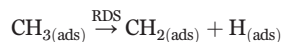

Fig. 1. Theoretical screening of SAA catalysts for C–H activation. (**A**) Activation energy of the first C–H cleavage in CH₄ plotted against segregation energy of each SAA. For comparison, the activation energies on Pt(111) and Pt₃Sn(111) (2 × 2) are shown as dash-dotted lines. (**B**) Charge density difference Δρ($\mathbf{r}$) for adsorbed CH₄ on Cu(111), Pt(111), and RhCu(111) SAA. Orange, Cu atoms; teal, Pt atoms; pink, Rh atoms. Cyan and yellow contours represent regions of electron depletion and accumulation, respectively (isovalue of ±0.01 electron Å⁻³). (**C**) Energy profile of the sequential C–H activation steps of CH₄ on various metal surfaces (more details in fig. S3). The energy is referenced with respect to the clean slab and CH₄(g). H atoms are adsorbed on the Cu host metal in the SAA case.

being exothermic on Pt(111) versus endothermic on pure Cu(111) and the RhCu(111) SAA (Fig. 1C). This is a notable result, given that Rh is generally not considered for dehydrogenation reactions because Rh NPs suffer from coking and hydrogenolysis is favored over selective dehydrogenation (18).

To test the theoretical predictions outlined above and investigate the individual reaction steps, we performed surface science and high-resolution imaging studies of model RhCu(111) SAA catalysts. Physical vapor deposition of Rh on a Cu(111) surface held at an elevated temperature yielded atomically dispersed Rh atoms that substituted Cu atoms in the surface layer (19). A typical image of the SAA surface is shown in fig. S5. Given the low desorption barrier of methane relative to typical C–H activation barriers, we used methyl iodide (CH_3I) as a probe molecule to study C–H activation in adsorbed methyl groups on the model RhCu(111) SAA surfaces. Specifically, C–I bond cleavage was facile on Cu(111), yielding methyl groups [$CH_{3(ads)}$] at low temperature that remained on the surface as it was heated. Once the rate-determining C–H activation step (RDS) occurred, yielding $H_{(ads)}$ and $CH_{2(ads)}$, gaseous methane evolved rapidly from facile coupling of $H_{(ads)}$ and $CH_{3(ads)}$, as described below (20)

$$CH_3I_{(ads)} \xrightarrow{fast} CH_{3(ads)} + I_{(ads)}$$

$$CH_{3(ads)} \xrightarrow{RDS} CH_{2(ads)} + H_{(ads)}$$

$$CH_{3(ads)} + H_{(ads)} \xrightarrow{fast} CH_{4(g)}$$

In this sense, methane evolution acted as a reporter of C–H activation, with lower-temperature methane evolution being associated with more facile C–H bond activation. Furthermore, the C–H activation barrier in CH_3 groups can be related to the C–H activation barrier in CH_4 via scaling relationships (fig. S6).

Temperature-programmed desorption (TPD) experiments on Cu(111) showed that C–H activation occurred at ~430 K (Fig. 2A), consistent with previous results (20). This finding was also consistent with our DFT calculations, which predicted that $CH_{3(ads)}$ conversion to $CH_{2(ads)}$ on Cu(111) had a barrier of 1.16 eV (Fig. 1C and fig. S3). Increasing the step density of the Cu(111) surface through argon ion sputtering resulted in a slightly lower C–H activation barrier and methane evolution at ~380 K. When individual, isolated Rh atoms were present in the Cu(111) surface, as seen in the low-temperature scanning tunneling microscopy (LT-STM) image and corresponding DFT-based image simulation in Fig. 2, B and C, several low-temperature methane evolution peaks were observed in the TPD experiments.

The lowest-temperature feature at ~160 K is caused by hydrogenation of methyl groups by hydrogen from the chamber background that

dissociates at Rh sites. This result was confirmed by experiments in which we precovered the RhCu(111) SAA surface with hydrogen before CH_3I was introduced. This low-temperature methane peak arose due to the aforementioned facile $CH_{3(ads)} + H_{(ads)} \rightarrow CH_{4(g)}$ reaction (Fig. 2A, bottom trace). The next TPD feature, at ~300 K, was caused by C–H activation of the methyl groups to form $CH_{2(ads)}$ and $H_{(ads)}$, followed by hydrogenation of the remaining methyl fragments. The temperatures at which these reaction steps occurred were consistent with our DFT calculations, which yielded a barrier for C–H activation in CH_3 on the RhCu(111) SAA of 0.54 eV, in contrast to the facile hydrogenation of CH_3 to CH_4, which has a barrier of 0.19 eV (Fig. 1C and fig. S3). Finally, the smaller CH_4 desorption feature at ~350 K is consistent with reaction at Cu step sites that were created when Rh was alloyed with Cu(111) (19). Full TPD traces for each experiment are provided in the supplementary materials (fig. S7).

To track the progression of the reaction intermediates not observable by TPD, we conducted LT-STM experiments after exposing a RhCu(111) SAA to CH_3I and annealing to differ-

ent temperatures before 5 K STM imaging (Fig. 2, D to G). After annealing to ~160 K, we observed the formation of small clusters of intact CH_3I (Fig. 2D). Heating to ~240 K resulted in dissociation of the C–I bond and $\sqrt{3} \times \sqrt{3}$ structural ordering of the $CH_{3(ads)}$ groups and $I_{(ads)}$ atoms seen in the STM images as protrusions [$I_{(ads)}$] and depressions [$CH_{3(ads)}$], confirming the facile C–I cleavage. At ~320 K, a substantial fraction of the methyl groups react and desorb as CH_4, and after a ~440 K anneal, only two-dimensional islands of I atoms (bright protrusions) in a $\sqrt{3} \times \sqrt{3}$ structure remain (21, 22). Notably, the image in Fig. 2G reveals that, aside from I atoms, which remained adsorbed on Cu(111) until ~840 K, there was no evidence of substantial coke formation, consistent with our theoretical predictions.

Guided by our theoretical calculations and surface science studies, we synthesized RhCu/ SiO_2 NP catalyst analogs with a Rh:Cu ratio of 1:100, demonstrated that they are indeed SAAs, and then tested them for the nonoxidative dehydrogenation of propane. RhCu/SiO_2 NPs of ~3-nm diameter were synthesized through galvanic replacement of silica-supported copper NPs (figs. S8 and S9) by small amounts of
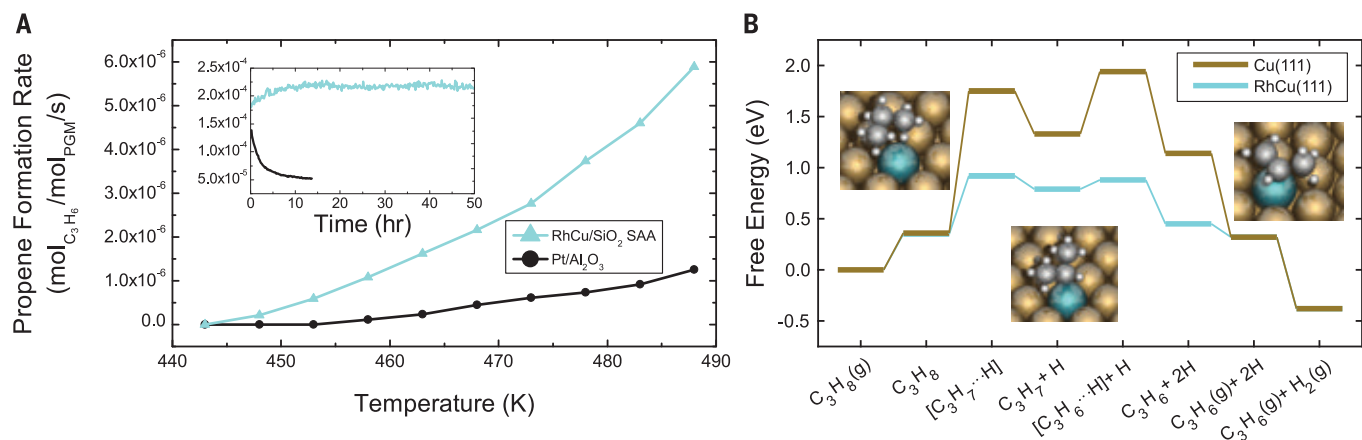
Fig. 2. Model system studies demonstrating that RhCu(111) SAAs promote low-temperature C–H activation. (**A**) TPD experiments on Cu(111), sputtered Cu(111), 0.2% RhCu(111) SAA, and 0.2% RhCu(111) SAA pre-dosed with 200 Langmuir [1 Langmuir (L) = 10^{-6} torr · s] of hydrogen. Each surface was exposed to 1.5 L of CH_3I. (**B**) STM image of an isolated Rh atom in Cu(111). (**C**) DFT-simulated STM image of an isolated Rh atom in Cu(111). (**D** to **G**) STM images from annealing experiments in which the RhCu(111) SAA was exposed to CH_3I, annealed to the indicated temperatures, and imaged at 5 K to track reaction progression.

Fig. 3. Propane dehydrogenation on RhCu SAAs. (**A**) Propene formation rate versus temperature for the nonoxidative propane dehydrogenation reaction on RhCu/SiO$_2$ SAA (Rh:Cu 1:100) and Pt/Al$_2$O$_3$ catalysts. (Inset) Stability test (50 hours) of RhCu/SiO$_2$ SAA versus Pt/Al$_2$O$_3$ at 623 K. Gas composition: 1.4 kPa of propane and 0.7 kPa of hydrogen; 90-ml-per-min flow rate. Formation rate was normalized per mole of platinum-group metal (PGM). (**B**) Free-energy diagram of RhCu(111) SAA and Cu(111) for the nonoxidative dehydrogenation of propane. Free energies were calculated at 500 K for a C$_3$H$_8$ and H$_2$ partial pressure of 1 kPa and a C$_3$H$_6$ partial pressure of 0.01 Pa. The energies are referenced with respect to the clean slabs and C$_3$H$_8$(g). H refers to a H atom adsorbed on the Cu metal surface.

Rh (*10*). CO diffuse reflectance infrared Fourier transform spectroscopy (DRIFTS) and in situ extended x-ray absorption fine structure (EXAFS) studies were performed to verify the successful alloying, as well as the absence of surface Rh aggregates. DRIFTS results (figs. S10 and S11) revealed a narrow (~15 cm^{-1} full width at half maximum) peak corresponding to linear CO adsorption on Rh atoms at 2089 cm^{-1}. Notably, no infrared signals for Rh directly on SiO$_2$ or Rh aggregates at the Cu surface were observed (*23*, *24*), thereby demonstrating that a SAA was formed, in line with our EXAFS analysis (figs. S12 and S13 and table S3).

We tested the RhCu/SiO$_2$ SAA catalyst for propane dehydrogenation in flow-reactor studies away from equilibrium in the kinetic regime. We performed a direct comparison of the RhCu/SiO$_2$ SAA catalyst with a standard consisting of Pt NPs supported on Al$_2$O$_3$ (*25*). Experiments were conducted both in the absence (fig. S14) and presence (Fig. 3A) of hydrogen in the feed (the latter experiment was performed to better simulate industrially relevant conditions in which a co-flow of H$_2$ is required to prevent Pt from coking). The data in Fig. 3A revealed that the RhCu/SiO$_2$ SAA catalyst exhibited higher activity per mole of active metal, as well as a considerably lower reaction light-off temperature, than Pt/Al$_2$O$_3$.

The high reactivity of the RhCu/SiO$_2$ SAA catalyst was coupled with Cu-like propene selectivity (100% on RhCu/SiO$_2$ SAA versus ~80% on Pt/Al$_2$O$_3$) and resistance to coking (fig. S15 and table S4). Although pure Rh NPs are poor dehydrogenation catalysts that are susceptible to coking because of the strong binding of carbon to extended Rh sites (*18*, *26*), the atomic dispersion of Rh in our SAA catalysts enabled facile C–H activation while avoiding over-dehydrogenation that leads to coke for-

mation by virtue of the endothermic nature of the coking process on single Rh atoms in a Cu host (Fig. 1C and fig. S3). This coking resistance resulted in the active, selective, and stable conversion of propane to propene and hydrogen for more than 50 hours on stream at 623 K (Fig. 3A, inset).

To further understand these experimental observations, we used DFT to calculate and compare the propane dehydrogenation reaction energetics of the RhCu(111) SAA to those of pure Cu(111), Rh(111), and Pt(111). Our DFT calculations (Fig. 3B, figs. S16 and S17, and table S5) show that the RhCu(111) SAA surface provided lower thermodynamic barriers for the formation of the isopropyl intermediate and propene than the pure Cu(111) surface. In addition, relative to Cu(111), the SAA exhibited a considerably lower kinetic barrier for the first dehydrogenation step [0.57 eV on the RhCu(111) SAA versus 1.39 eV on Cu(111)] and a very small subsequent kinetic barrier for the C$_3$H$_7$-to-C$_3$H$_6$ step (0.10 eV), indicative of a highly active catalyst. Notably, the activation barrier of propane on the RhCu(111) SAA is lower than that on pristine Pt(111) and very similar to that on Rh(111) [0.68 and 0.63 eV, respectively (fig. S16 and table S5)]. These results are consistent with experimental observations in which we measured the rate of propane dehydrogenation on the RhCu/SiO$_2$ SAA catalyst away from equilibrium over a small temperature window and found an apparent activation energy of ~0.7 eV (fig. S18). The measured apparent activation energy was identical with and without H$_2$, which suggests that there was not a substantial coverage of H on the catalyst surface at these temperatures.

One potential limitation in the widespread application of RhCu SAA catalysts is the well-known sintering of Cu NPs at elevated tem-

peratures. However, despite the expected loss of reactivity because of the sintering of Cu NPs (*27*), measurements of the initial propane dehydrogenation rates performed at 773 K showed that the RhCu/SiO$_2$ SAA catalyst performed comparably to other Pt-based catalysts previously reported (table S6) (*28*). Furthermore, we observed that addition of 1% Rh to the Cu NPs provided them with enhanced sintering resistance relative to pure Cu NPs (fig. S8). Methods known to stabilize metal NPs against sintering at high temperatures, such as atomic layer deposition of thin oxide layers, could further address this issue (*29*).

REFERENCES AND NOTES

1. F. Besenbacher et al., Science **279**, 1913–1915 (1998).
2. S. Alayoglu, A. U. Nilekar, M. Mavrikakis, B. Eichhorn, Nat. Mater. **7**, 333–338 (2008).
3. R. van Lent et al., Science **363**, 155–157 (2019).
4. F. Tao et al., Science **322**, 932–934 (2008).
5. B. Qiao et al., Nat. Chem. **3**, 634–641 (2011).
6. B. C. Gates, M. Flytzani-Stephanopoulos, D. A. Dixon, A. Katz, Catal. Sci. Technol. **7**, 4259–4275 (2017).
7. J. Hulva et al., Science **371**, 375–379 (2021).
8. L. DeRita et al., Nat. Mater. **18**, 746–751 (2019).
9. G. Kyriakou et al., Science **335**, 1209–1212 (2012).
10. R. T. Hannagan, G. Giannakakis, M. Flytzani-Stephanopoulos, E. C. H. Sykes, Chem. Rev. **120**, 12044–12088 (2020).
11. M. T. Darby, M. Stamatakis, A. Michaelides, E. C. H. Sykes, J. Phys. Chem. Lett. **9**, 5636–5646 (2018).
12. G. Giannakakis, M. Flytzani-Stephanopoulos, E. C. H. Sykes, Acc. Chem. Res. **52**, 237–247 (2019).
13. M. T. Darby, R. Réocreux, E. C. H. Sykes, A. Michaelides, M. Stamatakis, ACS Catal. **8**, 5038–5050 (2018).
14. National Academies of Sciences, Engineering, and Medicine, The Changing Landscape of Hydrocarbon Feedstocks for Chemical Production: Implications for Catalysis: Proceedings of a Workshop (The National Academies Press, 2016).
15. D. Gerceker et al., ACS Catal. **7**, 2088–2100 (2017).
16. M. T. Darby, E. C. H. Sykes, A. Michaelides, M. Stamatakis, Top. Catal. **61**, 428–438 (2018).
17. V. Fung, G. Hu, B. Sumpter, J. Mater. Chem. A **8**, 6057–6066 (2020).
18. J. H. Sinfelt, D. J. C. Yates, J. Catal. **8**, 82–90 (1967).
19. R. T. Hannagan et al., ChemCatChem **12**, 488–493 (2019).
20. C. Chao-Ming, B. E. Bent, Surf. Sci. **279**, 79–88 (1992).
21. M. Pascal et al., Surf. Sci. **512**, 173–184 (2002).
22. M. D. Marcinkowski et al., Nat. Chem. **10**, 325–332 (2018).

23. S. Gonzalez, C. Sousa, F. Illas, *Int. J. Mod. Phys. B* **24**, 5128–5138 (2010).
24. J. A. Anderson, C. H. Rochester, Z. Wang, *J. Mol. Catal. Chem.* **139**, 285–303 (1999).
25. J. J. H. B. Sattler, J. Ruiz-Martinez, E. Santillan-Jimenez, B. M. Weckhuysen, *Chem. Rev.* **114**, 10613–10653 (2014).
26. N. Raman *et al.*, *ACS Catal.* **9**, 9499–9507 (2019).
27. M. V. Twigg, M. S. Spencer, *Top. Catal.* **22**, 191–203 (2003).
28. G. Sun *et al.*, *Nat. Commun.* **9**, 4454 (2018).
29. B. J. O'Neill *et al.*, *Angew. Chem. Int. Ed.* **52**, 13808–13812 (2013).

SUPERCONDUCTIVITY

Discovery of a Cooper-pair density wave state in a transition-metal dichalcogenide

Xiaolong Liu[1]†, Yi Xue Chong[1]†, Rahul Sharma[1,2], J. C. Séamus Davis[1,3,4,5]*

Pair density wave (PDW) states are defined by a spatially modulating superconductive order parameter. To search for such states in transition-metal dichalcogenides (TMDs), we used high-speed atomic-resolution scanned Josephson-tunneling microscopy. We detected a PDW state whose electron-pair density and energy gap modulate spatially at the wave vectors of the preexisting charge density wave (CDW) state. The PDW couples linearly to both the *s*-wave superconductor and the CDW and exhibits commensurate domains with discommensuration phase slips at the boundaries, conforming those of the lattice-locked commensurate CDW. Nevertheless, we found a global $\delta\Phi \cong \pm 2\pi/3$ phase difference between the PDW and CDW states, possibly owing to the Cooper-pair wave function orbital content. Our findings presage pervasive PDW physics in the many other TMDs that sustain both CDW and superconducting states.

Transition-metal dichalcogenides (TMDs) are a rich platform for the exploration of quantum matter (*1–6*). In this context, a fundamental but elusive state is the electron-pair density wave (PDW) (*7*). Theoretically, the PDW state of TMDs was predicted to be generated by magnetic field (*8*), spin-valley locking (*9*), a CDW lock-in transition (*10*), and doping TMD bilayers (*11*); PDW melting into a charge-6*e* superfluid was also predicted (*12*). However, detecting the PDW state in a TMD experimentally has been challenging.

A familiar TMD state is the charge density wave (CDW): a charge density field $\rho_C(\boldsymbol{r})$ that modulates spatially at wave vectors $\boldsymbol{Q}_C^i (i =$ 1, 2, 3), which are not crystal reciprocal lattice vectors. Such a charge density modulation

$$\rho_C^i(\boldsymbol{r}) = \rho_i(\boldsymbol{r})e^{i\boldsymbol{Q}_C^i \cdot \boldsymbol{r}} + \rho_i^*(\boldsymbol{r})e^{-i\boldsymbol{Q}_C^i \cdot \boldsymbol{r}} \quad (1)$$

has a mean-field order parameter $\left\langle C_{\boldsymbol{k}}^\dagger C_{\boldsymbol{k}+\boldsymbol{Q}_C^i} \right\rangle$, where $C_{\boldsymbol{k}}^\dagger$ is the creation operator and $C_{\boldsymbol{k}}$ is the annihilation operator for single-electron $\boldsymbol{k}$-space eigenstates. The simplest TMD superconductor state is spatially homogeneous but breaks gauge symmetry

$$\Delta_S(\boldsymbol{r}) = \Delta_0 e^{i\phi} \quad (2)$$

with a mean-field order parameter $\left\langle C_{\boldsymbol{k}}^\dagger C_{-\boldsymbol{k}}^\dagger \right\rangle$. By contrast, a PDW state is described by an electron-pair field $\Delta_P(\boldsymbol{r})$ that modulates spatially at one or more wave vectors $\boldsymbol{Q}_P^i$

$$\Delta_P^i(\boldsymbol{r}) = \left[\Delta_i(\boldsymbol{r})e^{i\boldsymbol{Q}_P^i \cdot \boldsymbol{r}} + \Delta_i^*(\boldsymbol{r})e^{-i\boldsymbol{Q}_P^i \cdot \boldsymbol{r}}\right]e^{i\theta} \quad (3)$$

This state also breaks gauge symmetry, and its mean-field order parameter is $\left\langle C_{\boldsymbol{k}}^\dagger C_{-\boldsymbol{k}+\boldsymbol{Q}_P^i}^\dagger \right\rangle$. Sophisticated atomic-scale visualization of

TMD states by using single-electron tunneling (*13–16*) has revealed CDW quantum phase transitions (*13*), a CDW Bragg glass (*14*), interfacial band alignment (*15*), and strain control of the CDW state (*16*). But to detect and image a PDW state in TMDs remained an experimental challenge.

Experimentally, the total electron-pair density $\rho_{CP}(\boldsymbol{r})$ might be visualized by measuring Josephson critical-current $I_J(\boldsymbol{r})$ to a superconducting scanning tunneling microscopy (STM) tip (*17*) because $\rho_{CP}(\boldsymbol{r}) \propto I_J^2(\boldsymbol{r})R_N^2(\boldsymbol{r})$, where R_N is the normal-state junction resistance (*18*, *19*). But this has proven impractical because the thermal fluctuation energy k_BT typically exceeds the Josephson energy $E_J = \Phi_0 I_J/2\pi$, where

$$I_J = \frac{\pi\Delta(T)}{2eR_N}\tanh\left[\frac{\Delta(T)}{2k_BT}\right] \quad (4)$$

(k_B, Boltzmann constant; T, temperature; $2e$, electron-pair charge; and Φ_0, magnetic flux quantum). Instead, when $E_J < k_BT$, the tip-sample Josephson junction exhibits a phase-diffusive (*20–22*) steady state at voltage V and electron-pair current

$$I_{CP}(V) = \frac{1}{2}I_J^2 ZV/\left(V^2 + V_c^2\right) \quad (5)$$

Here, $V_c = 2eZk_BT/\hbar$, where Z is the high-frequency impedance in series with the voltage source and $\hbar$ is Planck's constant h divided by 2π. From Eq. 5

$$dI_{CP}/dV \equiv g(V) = \frac{1}{2}I_J^2 Z\left(V_c^2 - V^2\right)/\left(V_c^2 + V^2\right)^2 \quad (6)$$

yields $g(0) \propto I_J^2$ (fig. S1) [(*23*), section 1]. Thus, spatially resolved measurements of $g(0, \boldsymbol{r})$ can provide a practical means (*24–28*) to image $I_J(\boldsymbol{r})$, so that the electron-pair density can then be visualized as $N_{CP}(\boldsymbol{r}) \equiv g(\boldsymbol{r}, 0)R_N^2(\boldsymbol{r}) \propto \rho_{CP}(\boldsymbol{r})$ [(*23*), section 1].

[1]Laboratory of Atomic and Solid State Physics, Department of Physics, Cornell University, Ithaca, NY 14850, USA. [2]Department of Physics, University of Maryland, College Park, MD 20740, USA. [3]Department of Physics, University College Cork, Cork T12 R5C, Ireland. [4]Max Planck Institute for Chemical Physics of Solids, D-01187 Dresden, Germany. [5]Clarendon Laboratory, University of Oxford, Oxford OX1 3PU, UK.
*Corresponding author. Email: jcseamusdavis@gmail.com
†These authors contributed equally to this work.

We studied bulk crystals of 2H-NbSe$_2$, a quasi–two-dimensional TMD with a robust CDW state (29). It has a hexagonal layered structure with Se-Se separation d and a Fermi surface with pockets surrounding the Γ and K points (fig. S2). The CDW phase transition at $T \approx 33.5$ K generates crystal and charge density modulations at three in-plane wave vectors $\boldsymbol{Q}_C^i \approx \{(1,0); (1/2, \sqrt{3}/2); (-1/2, \sqrt{3}/2)\} 2\pi/3a_0$ ($a_0 = \sqrt{3}d/2$ is the unit cell dimension), and the s-wave superconductivity (SSC) transition at $T_C \approx 7.2$ K completely gaps the Fermi

Fig. 1. Simultaneous single-electron and electron-pair tunneling spectroscopy.
(**A**) Topographic image $T(\boldsymbol{r})$ of Se-termination surface of NbSe$_2$ measured at $T = 290$ mK. (Inset) The Fourier transform $T(\boldsymbol{q})$ with Bragg peaks $\boldsymbol{Q}_B^i \approx \{(1,0); (1/2, \sqrt{3}/2); (-1/2, \sqrt{3}/2)\} 2\pi/a_0$ indicated with gray circles and the CDW peaks $\boldsymbol{Q}_C^i \approx \{(1,0); (1/2, \sqrt{3}/2); (-1/2, \sqrt{3}/2)\} 2\pi/3a_0$ indicated with black circles. (**B**) Typical differential tunneling conductance spectrum $g(V) \equiv dI/dV(V)$ from a Nb scan-tip to NbSe$_2$ surface at $T = 290$ mK. The range of energies at which CDW modulations are intense in $g(V)$ is indicated approximately with red arrows. (**C**) Energy range in (B) is zoomed to show typical $g(V)$ characteristic owing to the combination of the superconducting energy gaps Δ_T of the Nb tip and Δ of the NbSe$_2$. The range of energies at which superconducting coherence peaks are intense in $g(V)$ is indicated with green arrows. (**D**) Measured electron-pair tunnel current $I_{CP}(V)$ in the phase-diffusive Josephson effect energy range $|E| \lesssim 100$ μeV, with the range of energies at which electron-pair current is maximum ($\pm I_m$) indicated with blue arrows. (**E**) Energy in (C) is zoomed to show phase-diffusive Josephson effect energy range, and the measured $g(V)$ whose $g(0) \propto I_J^2$ from Eq. 6 is indicated with a blue arrow.
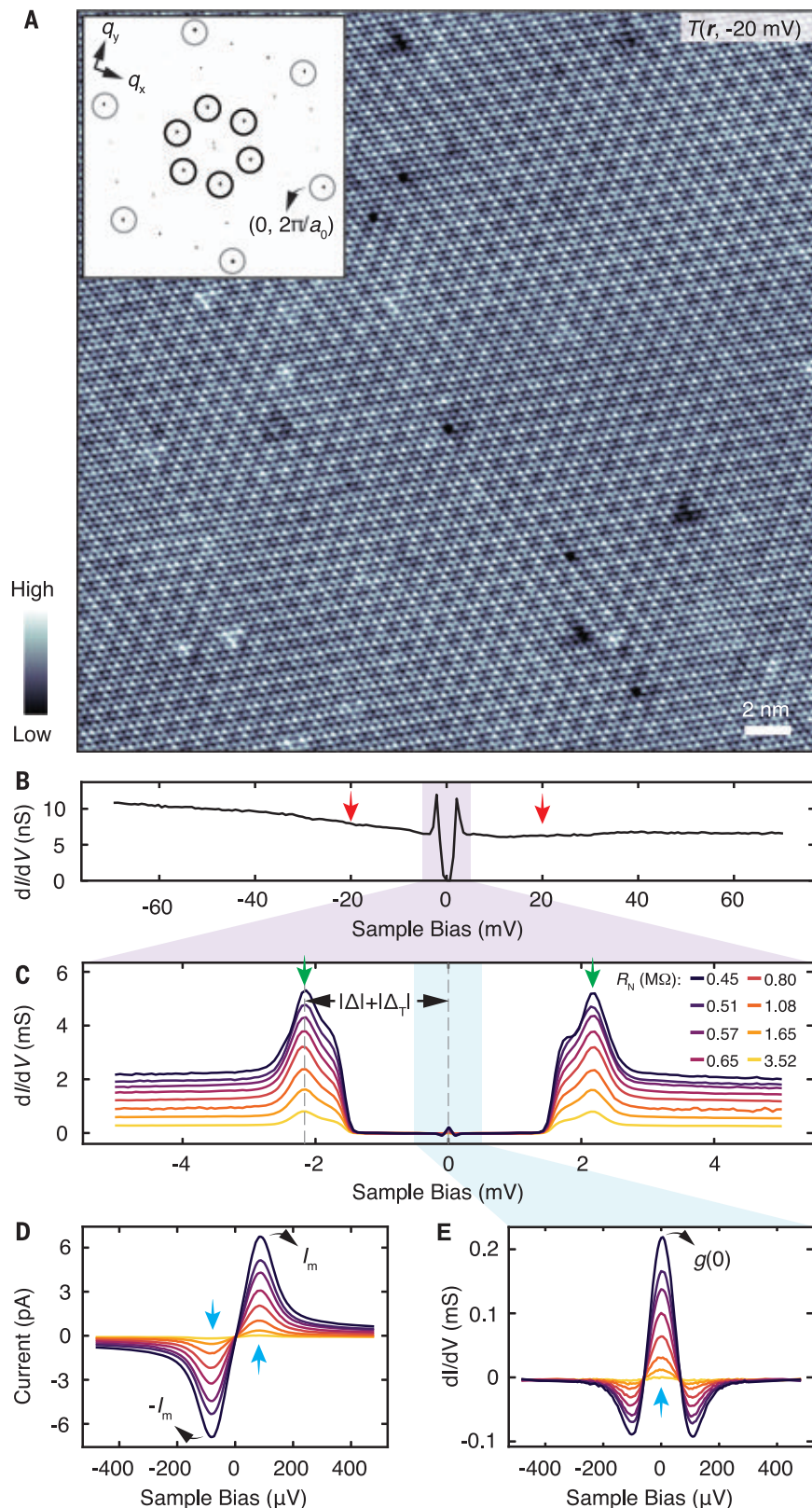
surface. We used atomic-resolution super-conducting scan tips made of Nb (17) with a standard tip-energy-gap $|\Delta_T| \approx 0.9$ meV (fig. S3).

A typical topographic image $T(\boldsymbol{r}, V)$ of the Se-termination layer of NbSe$_2$ when using such tips is shown in Fig. 1A, with the CDW modulations appearing as $3a_0$ periodic intensity amplifications [Fig. 1A, inset, $T(\boldsymbol{q}, V)$] (13, 14). A typical differential tunneling conductance spectrum $dI/dV|_V \equiv g(V)$ is shown in Fig. 1B. To simultaneously visualize the CDW, SSC, and any putative PDW states, a dynamic range exceeding 10^4 is required in the tip-sample voltage, spanning the CDW range from above ~50 mV (Fig. 1B), to the SSC energy gap range ~1 mV (Fig. 1C), to the Josephson pair-current range approaching ~10 µV (Fig. 1, D and E). Visualizing the quasiparticle densities $N_Q(\boldsymbol{r})$ of both CDW and SSC uses single-electron tunneling at energies indicated with the red and green arrows in Fig. 1, B and C, respectively. Visualizing electron-pair density $N_{CP}(\boldsymbol{r})$ of the condensate uses the phase-diffusive Josephson tunneling current $I_{CP}(V)$ or $g(0)$, indicated with the blue arrows in Fig. 1, D and E.

At $T = 290$ mK, we first imaged $N_Q(\boldsymbol{r}) \equiv g(\boldsymbol{r}, -20$ mV) at $V = -20$ mV, where CDW intensity is strong (13), with the results shown in Fig. 2A. Next, we imaged the normal-state resistance (fig. S4) of the tip-sample Josephson junction $R_N(\boldsymbol{r}) \equiv I^{-1}(\boldsymbol{r}, -45$ mV) (Fig. 2B). Third, we studied the electron-pair current by measuring $g(\boldsymbol{r}, 0)$ (Eq. 6) (Fig. 2C). All four independent images $T(\boldsymbol{r}, V)$: $N_Q(\boldsymbol{r})$: $R_N(\boldsymbol{r})$: $g(\boldsymbol{r}, 0)$ are registered to each other with precision of $\delta x \approx \delta y \lesssim 15$ pm (fig. S5) [(23), section 2]. This constitutes a typical data set for visualizing the crystal, CDW, SSC, and PDW states simultaneously; its acquisition required developing high-speed scanned Josephson-tunneling microscopy (SJTM) imaging protocols (fig. S6) [(23), section 3]. Eventually, to visualize the electron-pair density, we used the data in Fig. 2, B and C, to derive $N_{CP}(\boldsymbol{r}) \equiv g(\boldsymbol{r}, 0)R_N^2(\boldsymbol{r})$ (Fig. 2D). Here, we see electron-pair density modulations at three in-plane wave vectors $\boldsymbol{Q}_P^i \approx \{(1, 0); (1/2, \sqrt{3}/2); (-1/2, \sqrt{3}/2)\}2\pi/3a_0$, indicated with the blue circles in Fig. 2D, $N_{CP}(\boldsymbol{q})$ inset. Ultimately, in Fig. 2E, $N_C(\boldsymbol{r})$ of the CDW contains only the charge density modulations at $\boldsymbol{Q}_C^i$ (Fig. 2A), whereas in Fig. 2F, the simultaneous $N_P(\boldsymbol{r})$ of the PDW contains only the electron-pair density modulations at $\boldsymbol{Q}_P^i$ (Fig. 2D). The data in Fig. 2F represent the observation of a PDW state in a TMD material, NbSe$_2$.

In a PDW state, the energy gap $\Delta_P(\boldsymbol{r})$ should also modulate at $\boldsymbol{Q}_P$ (Eq. 3). We define the total gap $|\Delta(\boldsymbol{r})| = |\Delta_P(\boldsymbol{r}) + \Delta_S|$ to be half the energy separation between two coherence peaks minus $|\Delta_T|$ (Fig. 1C). Our measured $|\Delta(\boldsymbol{r})|$ then exhibits modulations at three

wave vectors $\boldsymbol{Q}_P^i \approx \{(1, 0); (1/2, \sqrt{3}/2); (-1/2, \sqrt{3}/2)\}2\pi/3a_0$ (fig. S7) [(23), section 4]. This confirms independently, by use of single-electron tunneling, the existence of a PDW state in NbSe$_2$. Its gap modulation amplitude $|\Delta_P| < 0.01|\Delta_0|$ [(23), section 4]. A plot of the measured Fourier amplitudes of simultaneous $N_Q(|\boldsymbol{q}|)$ and $N_{CP}(|\boldsymbol{q}|)$ in the directions of $\boldsymbol{Q}_P^i \approx \boldsymbol{Q}_C^i$ is shown in Fig. 3A. The key maxima occur near $|q| = 2\pi/3a_0$, establishing quantitatively that $|\boldsymbol{Q}_P^i| = |\boldsymbol{Q}_C^i| \pm 1\%$. But, although

imaged in precisely the same field of view (FOV), the charge density modulations (Fig. 2E) and electron-pair density modulations (Fig. 2F) appear distinctly different, with normalized cross correlation coefficient $\eta \approx -0.4$.

Possible microscopic mechanisms for a PDW state include Zeeman splitting (30, 31) of a Fermi surface (not relevant here) and strongly correlated electron-electron interactions generating intertwined CDW and PDW states (32, 33). But whatever the microscopic PDW
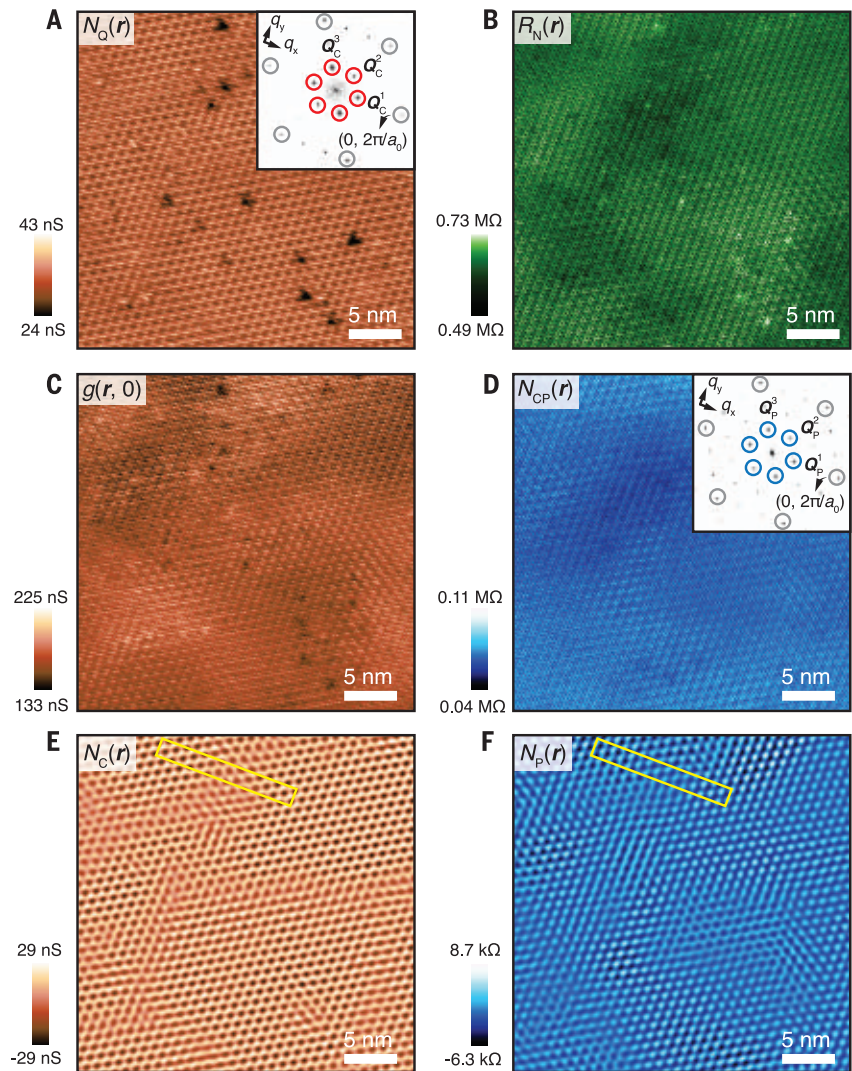

Fig. 2. Atomic-scale electron-pair tunneling visualization of a PDW state. (**A**) Measured $N_Q(\boldsymbol{r}) \equiv g(\boldsymbol{r}, -20$ mV) in the same FOV as Fig. 1A with pixel size ~30 pm at $T = 290$ mK. (Inset) $N_Q(\boldsymbol{q})$, with CDW peaks indicated with red circles. (**B**) Simultaneously measured $R_N(\boldsymbol{r}) = I^{-1}(\boldsymbol{r}, -4.5$ mV) as in (A). The purpose of this measurement is to establish the normal-state tip-sample junction resistance. (**C**) Simultaneously measured $g(r, 0) \propto I_J^2(\boldsymbol{r})$ as in (A). (**D**) Measured electron-pair density $N_{CP}(\boldsymbol{r}) \equiv g(\boldsymbol{r}, 0)R_N^2(\boldsymbol{r})$ from (B) and (C). (Inset) The PDW peaks in $N_{CP}(\boldsymbol{q})$ are indicated with blue circles. (**E**) Pure CDW charge density modulations $N_C(\boldsymbol{r})$ from (A). These are visualized at wave vectors $\boldsymbol{Q}_C^i \cong \{(1, 0); (1/2, \sqrt{3}/2); (-1/2, \sqrt{3}/2)\}2\pi/3a_0$ by means of Fourier filtering $N_Q(\boldsymbol{r})$ at the CDW peaks indicated with red circles. (**F**) Pure electron-pair density modulations $N_P(\boldsymbol{r})$ from (D). These are visualized at wave vectors $\boldsymbol{Q}_P^i \cong \{(1, 0); (1/2, \sqrt{3}/2); (-1/2, \sqrt{3}/2)\}2\pi/3a_0$ by means of Fourier filtering $N_{CP}(\boldsymbol{r})$ at the PDW peaks indicated with blue circles. There is a virtual absence of influence by impurity atoms or atomic-scale defects on the PDW state, as also seen in fig. S13.
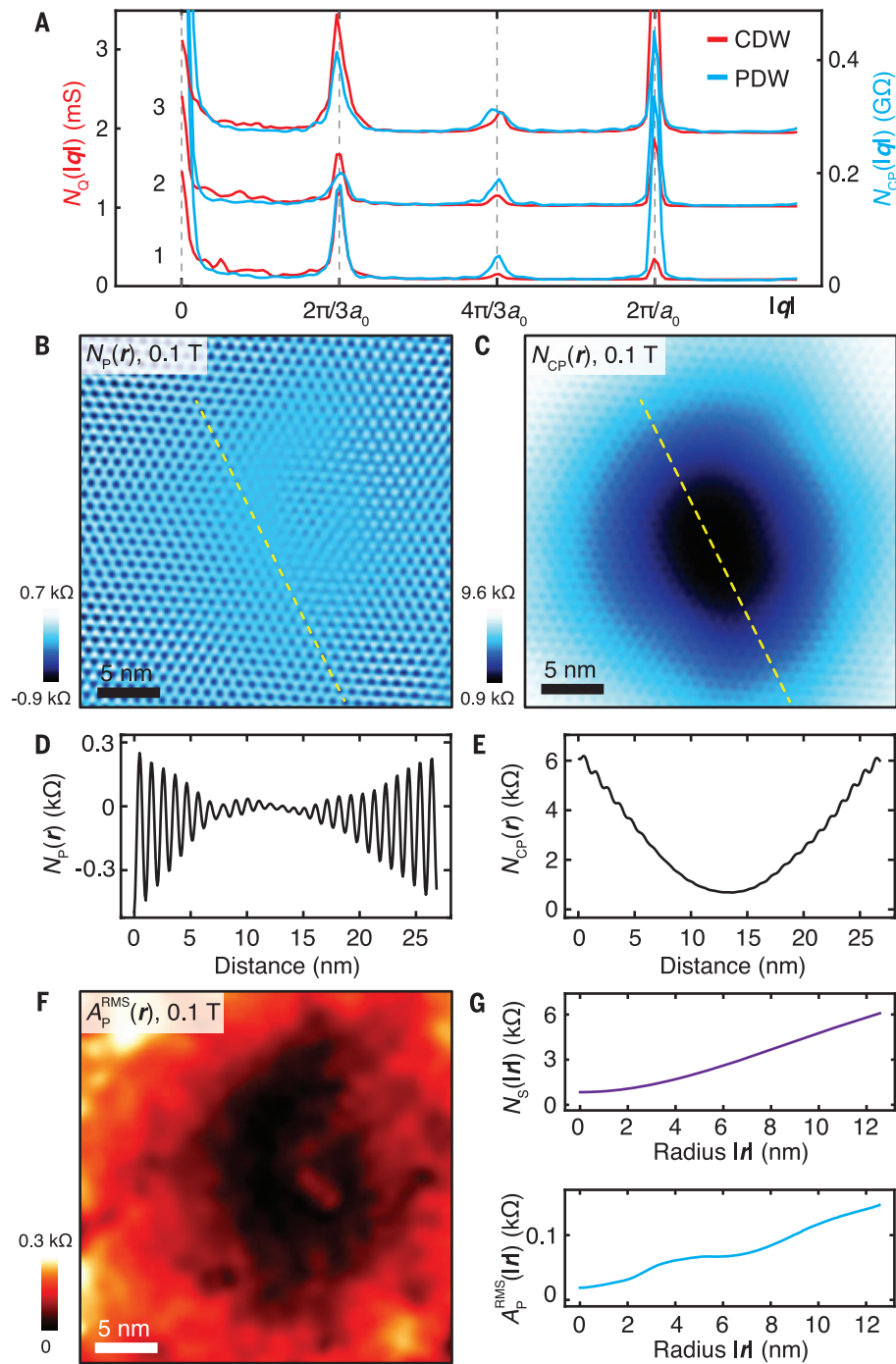
Fig. 3. Mutual decay of superconductivity and PDW into quantized vortex core. (**A**) Simultaneously measured amplitudes of charge density modulations $N_Q(|q|)$ (red) and electron-pair density modulations $N_{CP}(|q|)$ (blue) at $T = 290$ mK, where $Q_P^i \approx Q_C^i$ is evident. (**B**) Measured PDW electron-pair density modulations at Q_P^i, $N_P(r)$, centered on the core of a quantized vortex at $B = 0.1$ T and $T = 290$ mK. (**C**) Measured electron-pair density $N_{CP}(r)$ centered on vortex in (B). (**D**) Line profile of $N_P(r)$ along the yellow dashed line in (B). (**E**) Line profile of $N_{CP}(r)$ along the yellow dashed line in (C). (**F**) Measured PDW amplitude $A_P^{RMS}(r)$ centered on the vortex core. (**G**) The azimuthally averaged $N_S(|r|)$ centered on the vortex core symmetry point, and similarly the azimuthally averaged RMS amplitude of all three PDW modulations $A_P^{RMS}(|r|)$.

mechanism for NbSe$_2$, Ginzburg-Landau (GL) theory allows a general analysis of interactions between SSC and CDW states. Consider a didactic GL free-energy density

$$\mathcal{F} = \mathcal{F}_S + \mathcal{F}_C + \mathcal{F}_P + \sum_i \left(\lambda_i \rho_i \Delta_S^* \Delta_i + c.c. \right) \quad (7)$$

Here, $\mathcal{F}_S$, $\mathcal{F}_C$, and $\mathcal{F}_P$ are the free energy densities of a SSC state (Eq. 2), a CDW state (Eq. 1), and a PDW state (Eq. 3), respectively. The term $\lambda_i \rho_i \Delta_S^* \Delta_i$ represents lowest-order coupling of the SSC and CDW states with a PDW and induces $\Delta_P^i(r)$ at wave vectors $Q_P^i = Q_C^i$ owing to interactions of $\rho_C^i(r)$ and Δ_S. But the relative spatial arrangements of $\rho_C^i(r)$ and $\rho_P^i(r) \propto \left| \Delta_P^i(r) \right|^2$ are ambiguous because if $\rho_C^i(r) \propto \cos\left[Q_C^i \cdot r + \Phi_C^i(r) \right]$ and $\rho_P^i(r) \propto \cos\left[Q_P^i \cdot r + \Phi_P^i(r) \right]$ and $Q_P^i = Q_C^i$, the phase difference $\delta\Phi^i(r) \equiv \Phi_P^i(r) - \Phi_C^i(r)$ cannot be predicted from Eq. 7.

To explore the GL predictions, we next visualized quasiparticle density $N_Q(r)$ and electron-pair density $N_{CP}(r)$ centered on a quantized vortex core (figs. S8 and S9) [(23), section 5]. $N_P(r)$ of the PDW is shown in Fig. 3B, and the total $N_{CP}(r)$ is shown in Fig. 3C. The background superfluid density $N_S(r) = N_{CP}(r) - N_P(r)$ is azimuthally symmetric about the core, which is consistent with previous experiments (34) and GL theory (35). The mutual decay of the PDW and SSC into the vortex along the yellow dashed lines in Fig. 3, B and C, is visualized in Fig. 3, D and E. More quantitatively, if $N_C^i(r) = A_C^i(r) \cos\left[Q_C^i \cdot r + \Phi_C^i(r) \right]$ and $N_P^i(r) = A_P^i(r) \cos\left[Q_P^i \cdot r + \Phi_P^i(r) \right]$, the combined PDW amplitude is represented by $A_P^{RMS}(r) \sqrt{\sum_{i=1}^{3} \left[A_P^i(r) \right]^2 / 3}$ in Fig. 3F [(23), section 6], demonstrating its mutual decay with $N_S(r)$ of SSC in Fig. 3G. This is as expected within GL theory (Eq. 7) for a CDW-induced PDW state.

Even though the PDW state is strongly linked to the parent SSC state (Fig. 3) and to the modulation wave vectors of the CDW state (Figs. 2D and 3A), the two modulating states appear spatially disparate at the atomic scale (Fig. 2, E and F). To explore this unexpected phenomenon, we visualized the amplitude and phase of the CDW and PDW for all three wave vectors $Q_C^i \approx Q_P^i$ (fig. S10). Shown in Fig. 4A is measured $A_C^1(r)$ from Fig. 2A, and shown in Fig. 4B is the simultaneously measured $A_P^1(r)$ from Fig. 2D. Both show nanoscale variations in the magnitude of their order parameters that are spatially alike, which is consistent with Eq. 7. Shown in Fig. 4, C and D, are the $\Phi_C^1(r)$ and $\Phi_P^1(r) - 2\pi/3$ simultaneously obtained with Fig. 4, A and B, which are very similar but only when a phase shift of $2\pi/3$ is subtracted everywhere from the measured $\Phi_P^1(r)$. In Fig. 4E, we show the histogram of $\delta\Phi^1(r) = \Phi_P^1(r) - \Phi_C^1(r)$ from Fig. 4, C and D, and in Fig. 4F, we show a combined histogram of all $\left| \delta\Phi^1(r) \right|$ ($i = 1, 2, 3$). Hence, the relative spatial phase of the PDW and CDW states is globally $|\delta\Phi| \approx \pm 2\pi/3$. Experimentally measured $N_C^1(r)$ and $N_P^1(r)$ are shown in Fig. 4G merging with simultaneously measured topography $T(r)$ from the same FOV (Fig. 2, E and F, yellow boxes), revealing that an a_0 displacement between $N_C^1(r)$ and $N_P^1(r)$ generates this universal $\pm 2\pi/3$ phase shift.
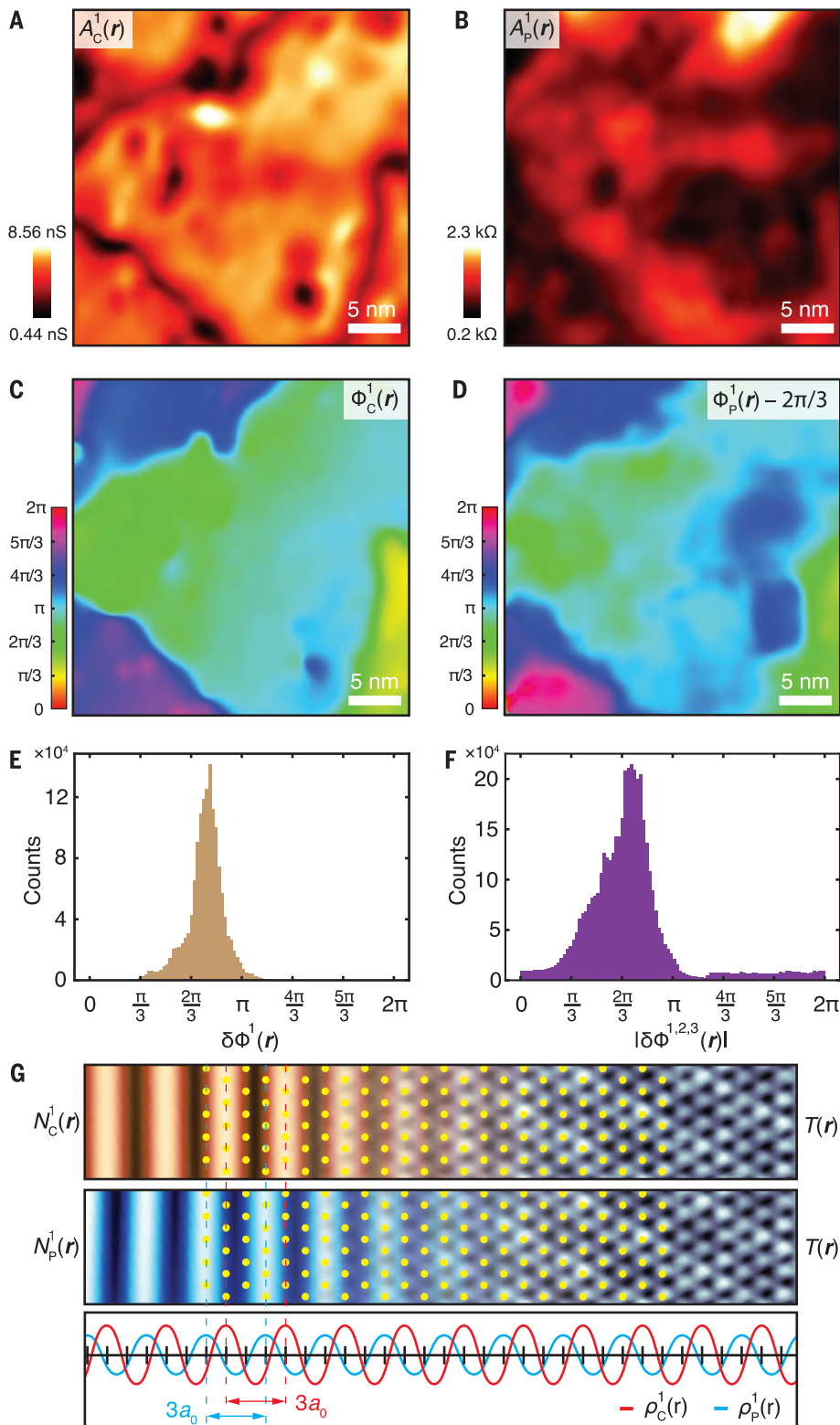
Fig. 4. Inter-state discommensuration of the CDW and PDW states. (**A**) Measured CDW amplitude of $A_C^1(\boldsymbol{r})$ for modulations at $\boldsymbol{Q}_C^1$ from Fig. 2A. All data shown in (A) to (D) were measured in the FOV of Fig. 1A, with the resulting electronic structure images shown in Fig. 2. (**B**) Simultaneously measured PDW amplitude of $A_P^1(\boldsymbol{r})$ for modulations at $\boldsymbol{Q}_P^1$ from Fig. 2D. (**C**) Measured CDW spatial phase $\Phi_C^1(\boldsymbol{r})$ for modulations at $\boldsymbol{Q}_C^1$ [simultaneous with (A)]. Our measurements found that $\Phi_C^1(\boldsymbol{r})$ is virtually independent of the bias voltage V for -70 mV $\leq V \leq -10$ mV (fig. S14) and impervious to atomic defects (fig. S15). (**D**) Measured PDW spatial phase $\Phi_P^1(\boldsymbol{r}) - \frac{2\pi}{3}$ for modulations at $\boldsymbol{Q}_P^1$ [simultaneous with (B)]. (**E**) Histogram of $\delta\Phi^1(\boldsymbol{r}) \equiv \Phi_P^1(\boldsymbol{r}) - \Phi_C^1(\boldsymbol{r})$ from (C) and (D). (**F**) Histograms of $\left|\delta\Phi^i(\boldsymbol{r})\right| \equiv \left|\Phi_P^i(\boldsymbol{r}) - \Phi_C^i(\boldsymbol{r})\right|$ for $i = 1, 2, 3$ from Fig. 2, A, and D. This result is also independently supported by the cross-correlation coefficient ($\eta_E = -0.44$) between $N_C(\boldsymbol{r})$ and $N_P(\boldsymbol{r})$ (Fig. 2, E and F) closely matching that of simulated images ($\eta_S = -0.5$) with $2\pi/3$ inter-state phase difference (fig. S16). (**G**) (Top and middle) Experimentally measured $N_C^1(\boldsymbol{r})$ and $N_P^1(\boldsymbol{r})$ merging with simultaneously measured topography $T(\boldsymbol{r})$ from the same FOV (Fig. 2, E and F, yellow boxes) with Se atoms indicated by yellow dots. (Bottom) Schematic of the PDW:CDW inter-state discommensurations with $\delta\Phi^1(\boldsymbol{r}) = \frac{2\pi}{3}$.

So, what generates and controls this complex new PDW state at atomic scale? First, Bloch-state modulations at crystal-lattice periodicity will lead inevitably to lattice-periodic modulations of $N_Q(\boldsymbol{r})$, $N_{CP}(\boldsymbol{r})$, and $\Delta(\boldsymbol{r})$ [(*23*), section 7]. However, at a more sophisticated and specific level, a multiband plus anisotropic energy-gap theory of NbSe$_2$ has been developed to describe superconductive electronic structure modulations at the atomic scale (*36*). Beyond this, lattice strain is important in CDW physics of TMD materials (*13*, *16*). Lattice-locked 3×3 commensurate CDW domains occur in NbSe$_2$, separated by discommensurations at which the CDW phase jumps by $\delta\Phi_C = \pm 2\pi/3$ (*28*). We detected these $\delta\Phi_C = \pm 2\pi/3$ discommensurations, for example, in $\Phi_C^1(\boldsymbol{r})$ (Fig. 4C and fig. S11) and found $\delta\Phi_P =$

±2π/3 phase slips for the PDW state at virtually identical locations, along its domain boundaries in Fig. 4D. This might be expected if the PDW is induced by the CDW coupling to the superconductivity because the PDW domains would replicate those of the preexisting CDW. Moreover, the interstate phase-difference $|\delta\Phi| = |\Phi_P(r) - \Phi_C(r)| \approx \pm 2\pi/3$ occurs universally (Fig. 4, C and D), not just at the commensurate domain boundaries. Hence, the simplest overall explanation is that the global phase shift $|\delta\Phi|$ does not originate from an independent lattice-lock-in of the PDW [(23), section 8].

As to atomic-scale interactions between the CDW and the SSC states, one must consider Cooper pairing in the presence of the CDW periodic potential $V(r)$. Solving the linearized superconducting gap equation does generate a nonzero $\Delta_P(r)$ with $Q_P = Q_C$ (37). More intuitively, electron pairing occurs not only at momenta $(k, -k)$ but also $(k + G, -k)$ and $(k, -k + G)$, where G is a reciprocal-lattice vector of the CDW state: $G = mQ_C$; $m = 0$, ±1, ±2, ... (fig. S12). The consequent electron-pair density at lowest order in G [(23), section 9] is

$$\rho_P(r) \propto \cos\left(Q_C \cdot r + \Phi_C^{Q_C} + \delta\Phi\right) \quad (8)$$

Thus, the electron-pair density modulates spatially at the wave vectors $\pm Q_C$ owing to the finite center-of-mass electron-pair momentum $(\pm\hbar Q_C)$ imposed by the CDW. Moreover, this same approach shows that a phase difference $|\delta\Phi| = \left|\Phi_P^{Q_C} - \Phi_C^{Q_C}\right|$ is determined by the k-space structure factor of the electron-pair wave function [(23), section 9]. Last, at the single-atom scale, we found that impurity atoms leave the PDW state virtually unperturbed (fig. S13), implying that Anderson's theorem also pertains to an s-wave PDW.

The techniques and observations reported here herald abundant and exciting PDW physics in the many TMDs that, like NbSe₂, sustain both CDW and superconducting states.

REFERENCES AND NOTES

1. S. Manzeli, D. Ovchinnikov, D. Pasquier, O. V. Yazyev, A. Kis, *Nat. Rev. Mater.* **2**, 17033 (2017).
2. A. Kogar et al., *Science* **358**, 1314–1317 (2017).
3. P. Li et al., *Nat. Commun.* **8**, 2150 (2017).
4. J. Gooth et al., *Nat. Commun.* **9**, 4093 (2018).
5. S. Ran et al., *Science* **365**, 684–687 (2019).
6. L. Jiao et al., *Nature* **579**, 523–527 (2020).
7. D. F. Agterberg et al., *Annu. Rev. Condens. Matter Phys.* **11**, 231–270 (2020).
8. D. Möckli, Y. Yanase, M. Sigrist, *Phys. Rev. B* **97**, 144508 (2018).
9. J. Venderley, E.-A. Kim, *Sci. Adv.* **5**, eaat4698 (2019).
10. C. Chen, L. Su, A. H. Castro Neto, V. M. Pereira, *Phys. Rev. B* **99**, 121108 (2019).
11. K. Slagle, L. Fu, *Phys. Rev. B* **102**, 235423 (2020).
12. D. F. Agterberg, M. Geracie, H. Tsunetsugu, *Phys. Rev. B Condens. Matter Mater. Phys.* **84**, 014513 (2011).
13. A. Soumyanarayanan et al., *Proc. Natl. Acad. Sci. U.S.A.* **110**, 1623–1627 (2013).
14. J. Okamoto, C. J. Arguello, E. P. Rosenthal, A. N. Pasupathy, A. J. Millis, *Phys. Rev. Lett.* **114**, 026802 (2015).
15. H. M. Hill, A. F. Rigosi, K. T. Rim, G. W. Flynn, T. F. Heinz, *Nano Lett.* **16**, 4831–4837 (2016).
16. S. Gao et al., *Proc. Natl. Acad. Sci. U.S.A.* **115**, 6986–6990 (2018).
17. S. H. Pan, E. W. Hudson, J. C. Davis, *Appl. Phys. Lett.* **73**, 2992–2994 (1998).
18. J. Šmakov, I. Martin, A. V. Balatsky, *Phys. Rev. B Condens. Matter Mater. Phys.* **64**, 212506 (2001).
19. M. Graham, D. K. Morr, *Phys. Rev. Lett.* **123**, 017001 (2019).
20. O. Naaman, W. Teizer, R. C. Dynes, *Phys. Rev. Lett.* **87**, 097004 (2001).
21. Y. M. I. Anchenko, *Sov. Phys. JETP* **55**, 2395–2402 (1969).
22. G.-L. Ingold, H. Grabert, U. Eberhardt, *Phys. Rev. B Condens. Matter* **50**, 395–402 (1994).
23. Materials and methods and supplementary text are available as supplementary materials.
24. H. Kimura, R. P. Barber Jr., S. Ono, Y. Ando, R. C. Dynes, *Phys. Rev. Lett.* **101**, 037002 (2008).
25. V. Crespo et al., *Physica C* **479**, 19–23 (2012).
26. M. H. Hamidian et al., *Nature* **532**, 343–347 (2016).
27. M. T. Randeria, B. E. Feldman, I. K. Drozdov, A. Yazdani, *Phys. Rev. B* **93**, 161115 (2016).
28. D. Cho, K. M. Bastiaans, D. Chatzopoulos, G. D. Gu, M. P. Allan, *Nature* **571**, 541–545 (2019).
29. W. L. McMillan, *Phys. Rev. B Condens. Matter* **14**, 1496–1502 (1976).
30. P. Fulde, R. A. Ferrell, *Phys. Rev.* **135** (3A), A550–A563 (1964).
31. A. I. Larkin, Y. N. Ovchinnikov, *Sov. Phys. JETP* **20**, 762–762 (1965).
32. E. Fradkin, S. A. Kivelson, J. M. Tranquada, *Rev. Mod. Phys.* **87**, 457–482 (2015).
33. P. Choubey et al., *Proc. Natl. Acad. Sci. U.S.A.* **117**, 14805–14811 (2020).
34. H. Suderow, I. Guillamón, J. G. Rodrigo, S. Vieira, *Supercond. Sci. Technol.* **27**, 063001 (2014).
35. A. A. Abrikosov, *J. Exp. Theor. Phys.* **32**, 1147–1182 (1957).
36. I. Guillamon, H. Suderow, F. Guinea, S. Vieira, *Phys. Rev. B Condens. Matter Mater. Phys.* **77**, 134505 (2008).
37. K. Machida, T. Koyama, T. Matsubara, *Phys. Rev. B Condens. Matter* **23**, 99–105 (1981).
38. X. Liu, Y. X. Chong, R. Sharma, J. C. S. Davis, Data associated with "Discovery of a Cooper-pair density wave state in a transition-metal dichalcogenide". Zenodo (2021).

ORGANIC CHEMISTRY

A tautomeric ligand enables directed C–H hydroxylation with molecular oxygen

Zhen Li[1]†, Zhen Wang[1]†, Nikita Chekshin[1], Shaoqun Qian[1], Jennifer X. Qiao[2], Peter T. Cheng[2], Kap-Sun Yeung[3], William R. Ewing[2], Jin-Quan Yu[1]*

Hydroxylation of aryl carbon–hydrogen bonds with transition metal catalysts has proven challenging when oxygen is used as the oxidant. Here, we report a palladium complex bearing a bidentate pyridine/pyridone ligand that efficiently catalyzes this reaction at ring positions adjacent to carboxylic acids. Infrared, x-ray, and computational analysis support a possible role of ligand tautomerization from monoanionic (L,X) to neutral (L,L) coordination in the catalytic cycle of aerobic carbon–hydrogen hydroxylation reaction. The conventional site selectivity dictated by heterocycles is overturned by this catalyst, thus allowing late-stage modification of compounds of pharmaceutical interest at previously inaccessible sites.

Hydroxylation of aromatic compounds is an important chemical transformation in various metabolic pathways and in organic synthesis because of the ubiquity of the phenol motif. In living systems, monooxygenases catalyze hydroxylation of aryl C–H bonds with molecular oxygen (1, 2). Although aerobic oxidations of alcohols and olefins (3–7) and catalytic C–H hydroxylation using organometallic complexes with PhIO and peroxides as oxidants have also been extensively investigated (8–19), the development of transition metal catalysts for C–H hydroxylation with molecular oxygen has met with limited success (Fig. 1A) (20). To date, studies have been directed toward understanding the reactivity of L₂Pt(II)-Me₂ with O₂ (21). In one particularly illuminating example, a tripod ligand–supported Pt(II)-alkyl complex was shown to react with O₂ to give a Pt-OOH species that could potentially lead to C–O bond formation (22). Although analogous Pd(II)-alkyl complexes could react with O₂ through a radical chain mechanism (23), Mirica and

¹The Scripps Research Institute, La Jolla, CA 92037, USA. ²Discovery Chemistry, Bristol-Myers Squibb, Princeton, NJ 08543, USA. ³Bristol-Myers Squibb Research and Development, Cambridge, MA 02142, USA.
*Corresponding author. Email: yu200@scripps.edu
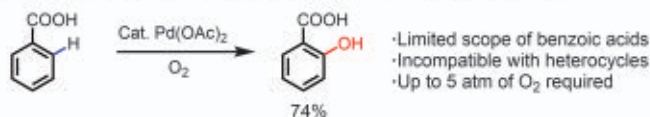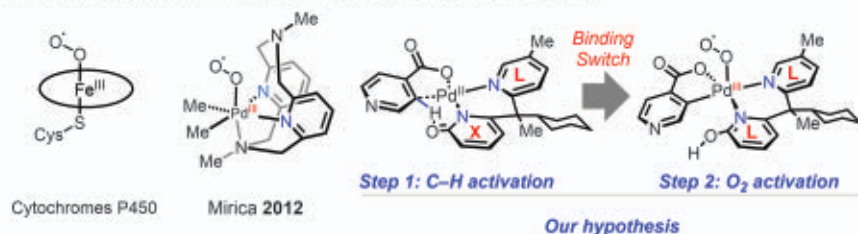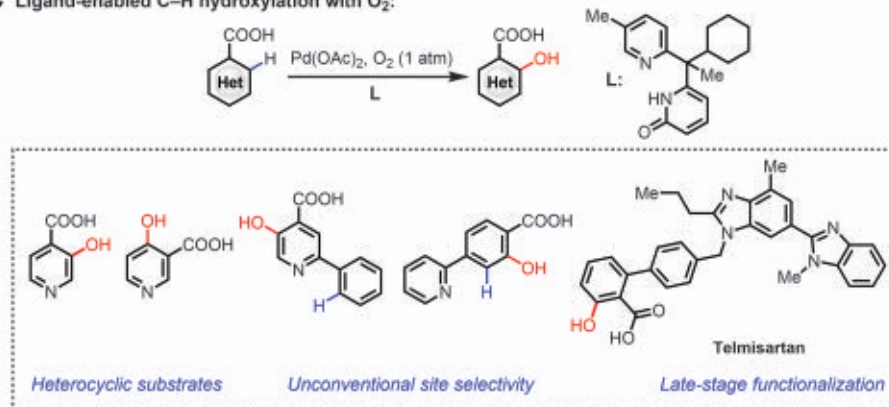†These authors contributed equally to this work.

Fig. 1. C–H hydroxylation with molecular oxygen. (**A**) Early observation of directed C–H hydroxylation using molecular oxygen. (**B**) Inspiration from monooxygenases and organometallic complexes. (**C**) Ligand-enabled C–H hydroxylation of heterocyclic carboxylic acids with O_2.

colleagues have recently disclosed an organometallic model implicating O_2 activation by a tripod ligand–supported Pd(II)alkyl species (Fig. 1B) (*24*). Electron paramagnetic resonance (EPR) and mass spectrometry supported a two-electron oxidation of synthetic Pd(II)alkyl complexes to form Pd(III)alkyl(OO·) and subsequently Pd(IV)alkyl(OOH) species. This species then oxidizes another Pd(II) complex to give two molecules of Pd(IV)(OH), which undergo C–O reductive elimination (*25*). Whereas these studies indicate the feasibility of O_2 activation by Pd(II) complexes, merging this elementary step with C–H activation to close the catalytic cycle remains a tremendous challenge. A viable catalytic system requires a ligand that is compatible with all crucial elementary steps: the activation of C–H bonds of synthetically relevant substrates, oxidation with O_2, and subsequent C–OH reductive elimination.

Vedernikov and colleagues observed an intriguing Pd(II)-catalyzed acetoxylation reaction of benzylic C–H bonds using O_2 as the oxidant (*26*). Although this reactivity was limited to quinoline-directed cyclopalladation pro-

cesses, this acetoxylation reaction provides an alternative system to investigate the activation of O_2 by Pd(II) species. A single example of aerobic aryl C–H hydroxylation catalyzed by Pd(OAc)$_2$ salts has been observed (Fig. 1A) (*20*). However, the scope of this protocol was limited to benzoic acids and, for some cases, even required the use of 5 atm of O_2 to achieve reasonable reactivity. Efforts to elucidate the mechanism of the reaction for further improvement have been futile because of the lack of supporting ligands and well-defined catalysts. In the past decades, we have developed several classes of ligands that can accelerate Pd-catalyzed C–H activation reactions. Unfortunately, none of these ligands is capable of facilitating the oxidation–hydroxylation step. On the basis of the aforementioned mechanistic studies on O_2 activation and our previous understanding of ligand cooperation in the C–H activation, we believe that a conventional ligand with a single coordination mode would be unlikely to accommodate both types of reactivity. Because the pyridone motif was recently identified as an efficient ligand for the cleavage

of C–H bonds (*27*), we therefore aimed to design a bifunctional tautomeric ligand scaffold that could switch between pyridone and pyridine coordination modes, potentially leveraging one ligand motif (pyridone) to facilitate C–H activation and another one (pyridine) to promote O_2 activation (Fig. 1B).

Here, we report the design of a Pd(II) catalyst bearing a bidentate pyridine-pyridone ligand that enables C–H hydroxylation of both aromatic and heteroaromatic carboxylic acids (Fig. 1C). By altering our ligand design approach from the conventional five-membered-ring metallacycle with Pd to a six-membered-ring metallacycle, we aimed to facilitate tautomerization of the L,L-ligand mode (pyridine-pyridine) into the L,X-ligand mode (pyridine-pyridone) required for C–H activation (Fig. 2A). This enhanced propensity for tautomerization is supported by both computational analysis and infrared (IR) spectroscopy. The ability of the ligand to switch between the two coordination modes could be important for both C–H activation and oxidation with O_2 to proceed, thus completing the catalytic cycle.

Our initial experimental effort was guided by our previous finding that Pd(OAc)$_2$ catalyzes the *ortho*-C–H hydroxylation of a few benzoic acids with molecular oxygen (*20*). The lack of ligands and the well-defined catalyst structure has substantially hampered our efforts to advance mechanistic understanding and further design of superior catalysts. The recent development of several classes of ligands for Pd-catalyzed C–H activation reactions led us to focus on ligand design for C–H hydroxylation of previously unreactive heteroaryl acids. Although none of the previously developed ligands showed any reactivity with O_2, we were intrigued by our previous observation that pyridone ligands as the internal base can accelerate C–H activation (*27*). We thus began to examine the ligand effect of a wide range of pyridone ligands using an isonicotinic acid substrate **1**, which previously showed no reactivity under the ligandless conditions. Unfortunately, no desired product was observed with a wide range of ligands (**L1** to **L19**; see table S1a). We reasoned that an effective ligand must perform a dual function: countering inhibition from heterocycle coordination and accelerating the C–H activation step while simultaneously facilitating the oxidation of Pd(II) to Pd(IV) by O_2, as demonstrated by Mirica and colleagues in an organometallic model containing bidentate pyridine ligands (Fig. 1B). We thus began to explore bidentate ligands containing both pyridone and pyridine that could potentially meet these criteria, with the L,X-type coordination incorporating the pyridone as an internal concerted metalation deprotonation (CMD)–active base for C–H cleavage and the L,L-type coordination incorporating the hydroxypyridine mirroring

A Catalyst design based on ligand tautomerization:
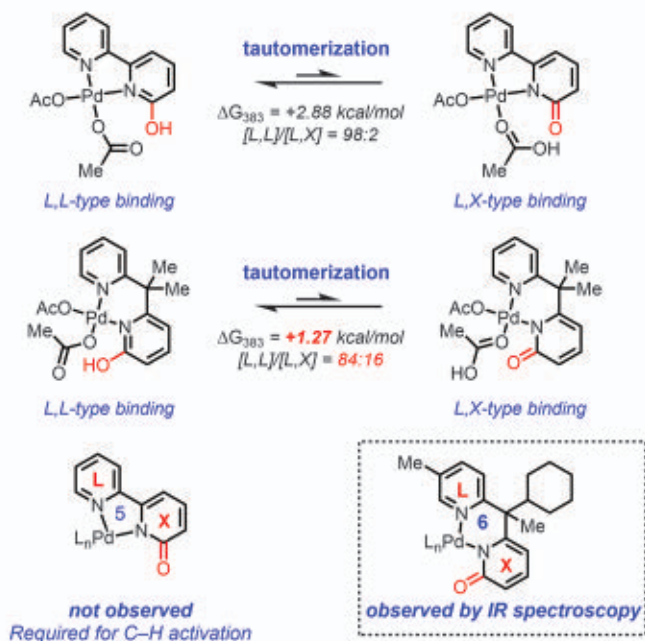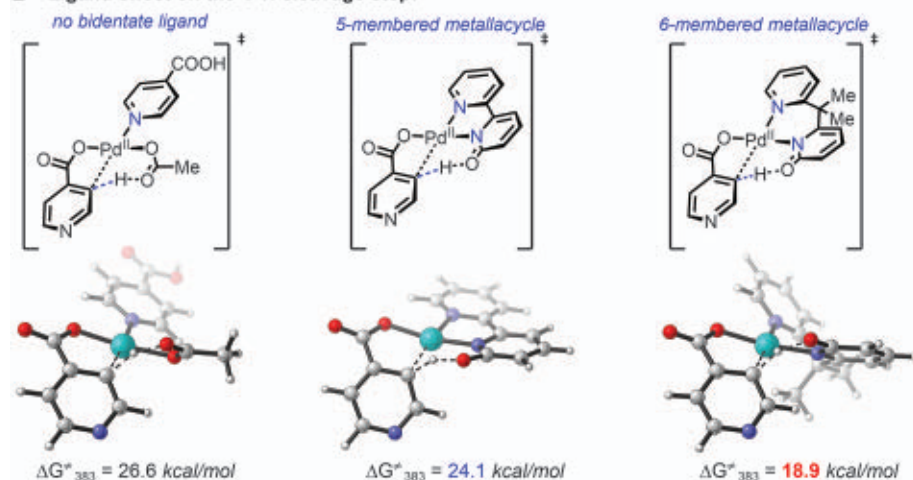

B Ligand effect on the C-H cleavage step:

Fig. 2. Ligand design and computational rationalization. (**A**) Catalyst design based on ligand tautomerization. (**B**) Ligand effect on the C–H cleavage step. See the supplementary materials for computational details.

Mirica's approach for oxygen activation. A variety of pyridine-pyridone ligands that could form five-membered metallacycles with Pd(II) were synthesized and tested under the standard conditions (**L20** to **L25**; see table S1b), yet no C–H hydroxylation was observed.

Although L,L- or L,X-type coordination motifs of these ligands are both possible, as shown by the synthesis of corresponding complexes (*28*), comparative density functional theory (DFT) analysis of Pd-ligand chelates showed that the L,X coordination mode of the six-membered metallacycle is more accessible compared with that of the

five-membered one by 1.6 kcal/mol at the reaction temperature (Fig. 2A). This reasoning led us to prepare bidentate pyridine-pyridone ligands that would form six-membered metallacycles so that the energy barrier to switch between L,L- and L,X-type coordination modes might be lowered (**L26** to **L45**; see table S1b). Fortunately, ligand **L33** afforded the desired *ortho*-hydroxylated product in 48% yield, demonstrating the feasibility of this ligand design principle. A wide range of analogous bidentate pyridone ligands bearing different substitutions were then synthesized and tested (**L34** to **L45**). Although electronic tuning of the

pyridine ring affected the reactivity (**L28** to **L32**; see table S1b), substitution at the benzylic position improved the yield more substantially. Ligand **L40**, which bears a phenyl substituent, provided the desired product in 55% yield. Ligand **L42**, featuring a cyclohexyl group, further improved the yield to 72%. The presence of bulky substituents could enhance binding through the Thorpe-Ingold effect, as shown in **L42** (Fig. 3). This reaction can also proceed using air, albeit in lower yield (**2a**, 20%). Dimethylformamide (DMF) was identified as the best solvent for this transformation and 1,4-benzoquinone (BQ) was used as an additive to further boost the yield (see table S5).

The use of $^{18}O_2$ (97% purity) gave the desired product with 95% ^{18}O incorporation, supporting the direct incorporation of O_2 into the hydroxylation pathway. Although no direct evidence has been obtained for the involvement of the putative Pd-OOH intermediate shown in Fig. 1B, no radical intermediates were detected in EPR experiments using 5,5-dimethyl-1-pyrroline *N*-oxide as a trap, which lowers the likelihood of radical chain processes (see the supplementary materials). At this stage, whether our reaction involves the 4e oxidation pathway proposed by Bercaw and Mirica (*22, 24*) remains to be ascertained. To obtain further evidence in support of the role of tautomerization, we used IR spectroscopy to monitor the coordination mode of both **L20** and **L42** by titrating these ligands into a solution of Pd(OAc)$_2$ in dimethoxyethane (DME). The optimal solvent, DMF, interferes with IR detection in the spectroscopic region of interest, although the reaction also proceeds in DME with diminished product yields. The observed peak at 1640 cm^{-1} with **L42** is consistent with the presence of the pyridone-binding motif (*28*), whereas the absence of the corresponding carbonyl signal suggests a predominant L,L-type pyridine-pyridine binding for five-membered coordinating counterpart **L20**. In addition, x-ray crystallographic characterization of a Pd(II)–**L42** complex and comparison with a known pyridone–Pd(II) complex (*28*) also revealed a carbonyl motif in the coordination structure based on bond length (see the supplementary materials for experimental details). Finally, the lack of deuterium incorporation in the presence of AcOH-*d*$_4$ with ligand **L20** also suggests that this ligand-Pd complex is not reactive in C–H activation even if a small amount of the pyridone form can be accessed. DFT analysis of the proposed C–H cleavage transition states (Fig. 2B) showed the free energy barrier with the six-membered ligand metallacycle to be 18.9 kcal/mol at the reaction temperature, compared with 24.1 kcal/mol with the five-membered ligand metallacycle and 26.6 kcal/mol with substrate as a ligand (acetate as the
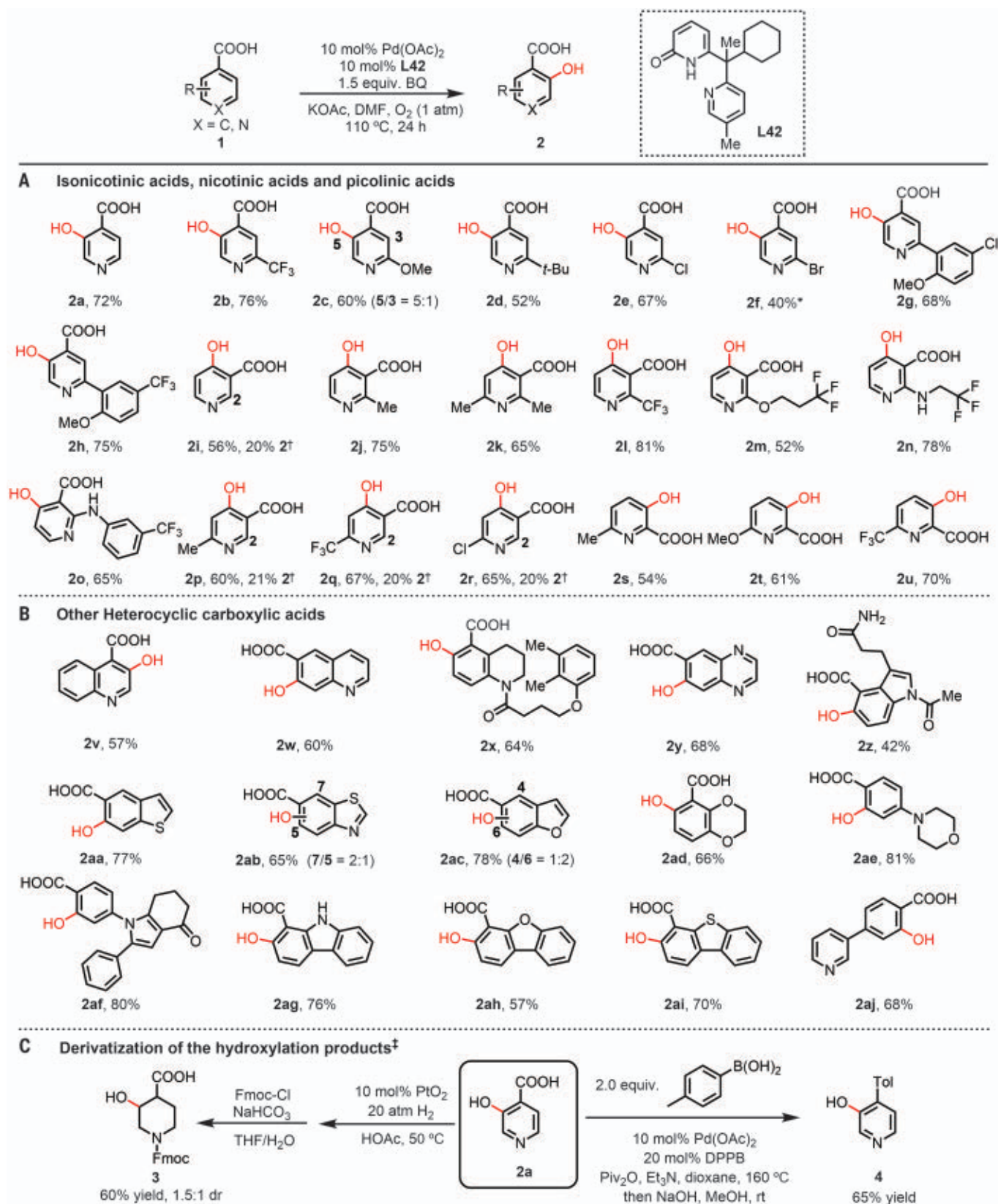
Fig. 3. Bidentate pyridone ligand enabled C–H hydroxylation with molecular oxygen and derivatization of the product. The reaction was performed with carboxylic acid (0.1 mmol), Pd(OAc)$_2$ (0.01 mmol), **L42** (0.01 mmol), BQ (0.15 mmol), 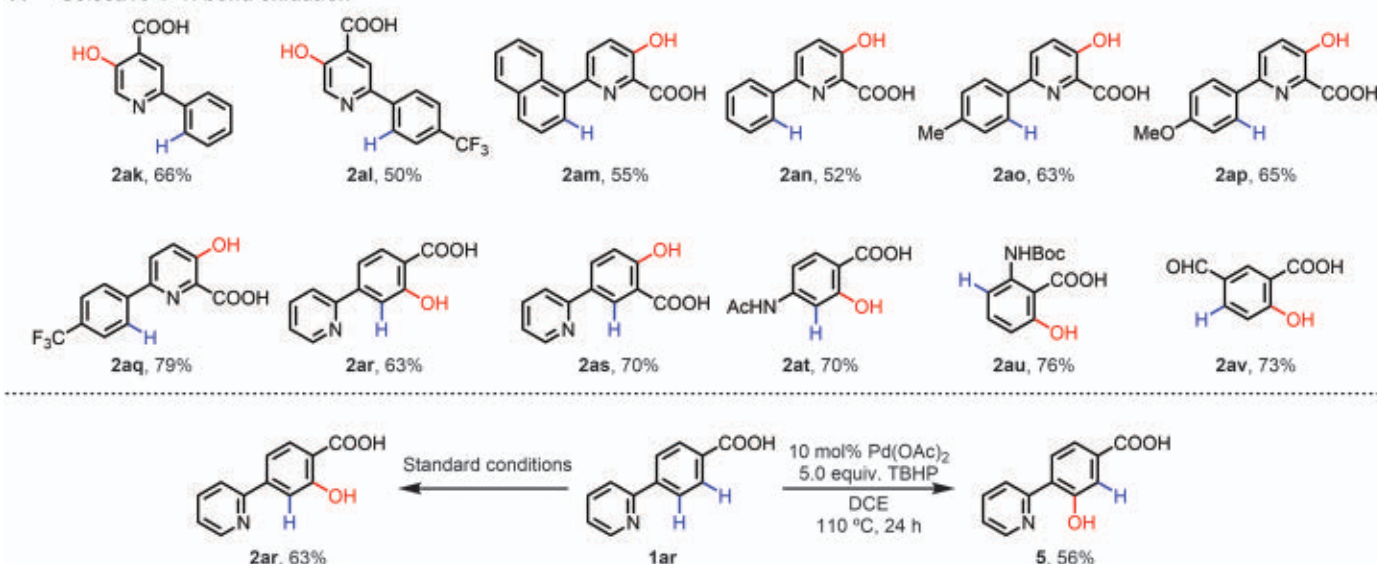and KOAc (0.2 mmol) in DMF (0.8 ml) under O$_2$ (1 atm) at 110°C for 24 hours. Isolated yield. *48 hours †Determined by ^{1}H-NMR using 1,3,5-trimethoxybenzene as the internal standard. ‡See the supplementary materials for experimental details. Fmoc, fluorenylmethyloxycarbonyl; DPPB, 1,4-bis(diphenylphosphino)butane; Tol, p-tolyl.

CMD base). Compared with the planar five-membered binding ligand, the nonplanarity of the six-membered binding ligand allows it to adopt a puckered (boat) metallacycle conformation, resulting in a better positioning of the pyridone motif for C–H cleavage and reducing the ligand and substrate distortion in the corresponding transition state. Whereas both DFT studies and IR data support the presence of pyridone form when coordinated with Pd(II), as well as pyridone-assisted C–H cleavage accompanied by tautomerization from the pyridone to pyridine, which might facilitate the subsequent O_2 activation step, the exact coordination structure of the catalyst during each catalytic step remains to be elucidated.

With the optimized conditions and the best ligand in hand, we subjected a wide range of pharmaceutically important heterocyclic benzoic acids to the hydroxylation reaction (Fig. 3). Isonicotinic acids containing electron-withdrawing (**1b**), electron-donating (**1c** and **1d**), and halogen (**1e**, **1f**) substituents were all reactive, affording the hydroxylated product at the less hindered position with exclusive monoselectivity. Heterocyclic biaryl substrates, common scaffolds in drug molecules, were also smoothly hydroxylated in high yields (**2g** and **2h**). Unsubstituted nicotinic acid gave a mixture of 4-hydroxynicotinic acid (**2i**, major) and 2-hydroxynicotinic acid (minor). Substrates **1j** and **1k** provided *ortho*-hydroxylation products selectively without any trace of benzylic C–H

oxidation, rendering a radical pathway unlikely. Various 2-substituted nicotinic acids afforded 4-hydroxylated products in high yield regardless of their electronic properties (**2l** to **2o**). The 2-alkylamino and 2-arylamino groups, which are often incompatible with C–H activation reactions because of their strongly coordinating nature, were both well tolerated (**2n** and **2o**). For 6-substituted nicotinic acids (**2p** to **2r**), both 2- and 4-hydroxylated products were obtained. Picolinic acids were also compatible substrates for this reaction, affording corresponding 3-hydroxypicolinic acids in high yields (**2s** to **2u**). Picolinic acids are usually unreactive substrates in C–H activation because of their bidentate chelation. Considering the Kolbe-Schmitt reaction (which is used in
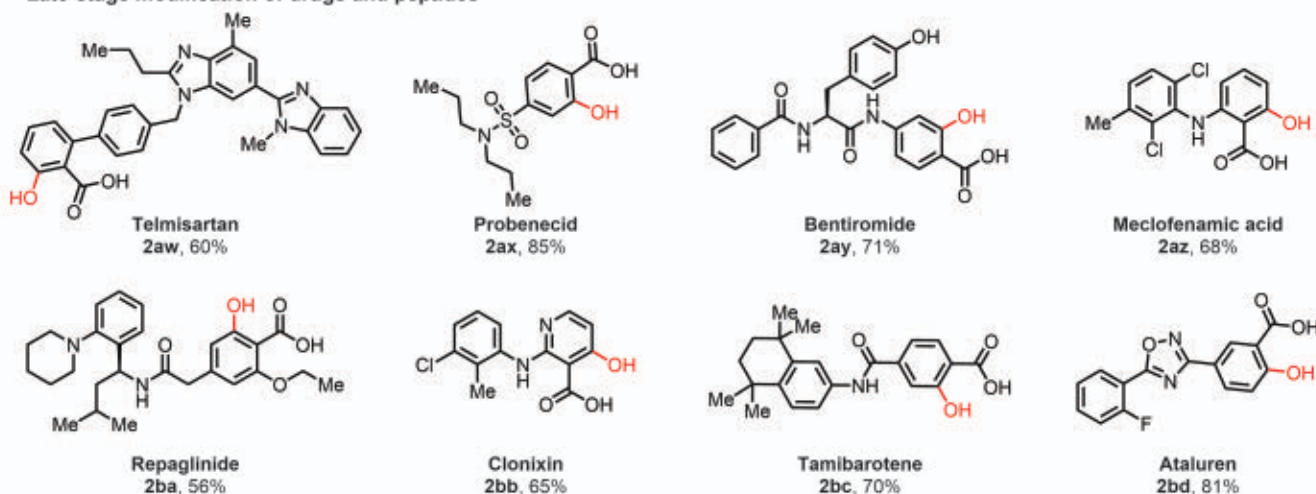

Fig. 4. Selective C H hydroxylation and late-stage modification. Standard conditions entailed carboxylic acid (0.1 mmol), Pd(OAc)$_2$ (0.01 mmol), **L42** (0.01 mmol), BQ (0.15 mmol), and KOAc (0.2 mmol) in DMF (0.8 ml) under O_2 (1 atm) at 110°C for 24 hours. Isolated yields are reported. TBHP, tert-butyl hydroperoxide; DCE, 1,2-dichloroethane.

the industrial production of salicylic acids) is incompatible with electron-deficient arenes (*29*), this protocol provides a valuable method with which to synthesize 2-hydroxy heteroaryl acids from readily available precursors.

Hydroxylated product **2a** could further undergo hydrogenation under 20 atm of H_2 using PtO_2 as a catalyst, providing facile access to substituted piperidine **3**. In addition, carboxylic acids could also serve as useful synthetic handles for subsequent transformations into a wide range of other functional groups (*30*, *31*). For example, we successfully performed a decarboxylative coupling reaction between **2a** with *p*-tolylboronic acid, forming bi-aryl product **4**. Considering the impact of various benzoquinone analogs on Pd(II)–catalyzed C–H activation and C–C bond–forming reactions (*32*, *33*), we rescreened other benzoquinone analogs under the optimized conditions (see table S6) and found that the use of bulkier 2-(*tert*-butyl)benzoquinone reduced the Pd catalyst loading to 2 mol%.

To further explore the scope of this hydroxylation reaction, we next evaluated other heteroaromatic carboxylic acids in addition to pyridine. Quinolines, tetrahydroquinoline, and quinoxaline were all compatible, affording hydroxylated products in good yields (**2v** to **2y**). Indoles, which are often unstable under oxidizing conditions, were also tolerated under our reaction conditions (**2z**). Carboxylic acids featuring benzothiophene (**2aa**), benzothiazole (**2ab**), benzofuran (**2ac**), benzodioxane (**2ad**), morpholine (**2ae**), pyrrole (**2af**), carbazole (**2ag**), dibenzofuran (**2ah**), and dibenzothiophene (**2ai**) structures were all successfully hydroxylated. Hydroxylation of pharmaceutically important bi-aryl heterocyclic carboxylic acid (**2aj**) also gave the desired product in 68% yield.

The extraordinary compatibility of this catalyst with heterocycles without using a strong directing group prompted us to explore site selectivity in the presence of multiple directing groups (Fig. 4). Overcoming the heterocycle-directed site selectivity typically requires the installation of an external directing group onto the native functional groups (*34*). Reversing the heterocycle-directed to simple carboxylic acid–directed C–H activation remains an unsolved challenge for Pd(II) catalysts. Because 2-pyridyl is considered one of the strongest directing groups for $C(sp^2)$–H activation, we examined biaryl substrates in which the pyridyl nitrogen and carboxyl group are either on the same or different aryl rings. Unexpectedly, the weakly coordinating carboxyl group completely overpowered the 2-pyridyl–directing effect in our hydroxylation reaction. For example, 2-aryl isonicotinic acids and 6-aryl picolinic acids consistently afforded carboxyl-directed hydroxylation product in 50 to 79% yields (**2ak** to **2aq**). Different scaffolds with 2-pyridyl and the carboxyl group on different rings also

afforded the site selectivity governed by the carboxyl group (**2ar** and **2as**). Exclusive deuterium incorporation into the position *ortho* to the carboxyl group rather than the pyridyl group in the presence of AcOH-d_4 suggests that carboxyl-directed C–H palladation is predominantly favored by this new ligand (see figs. S1 to S3). Switching to the ligandless conditions (*19*) with TBHP as the oxidant, substrate **1ar** underwent pyridine-directed hydroxylation to give **5**. The unconventional site selectivity was also obtained in the presence of a range of other commonly used native directing groups including -NHAc, -NHBoc, and aldehyde (**2at** to **2av**). The aldehyde remained intact under this oxidative condition (**2av**). The ability to tune the site selectivity by choosing different conditions affords a great potential for rapidly diversifying molecules by sequential site-selective C–H activation.

In principle, late-stage modification of complex natural products and drug molecules by site-selective C–H activation is a powerful approach to rapidly optimize bioactivity of drug candidates. Among the wide range of Pd-catalyzed C–H activation reactions, the lack of compatibility with heterocycles and site selectivity is a major obstacle toward reducing this approach to practice. In light of the importance of installing hydroxyl groups into drug candidates for late-stage functionalization, we subjected a number of commercial drugs to our standard reaction conditions. The antihypertension drug telmisartan was *ortho*-hydroxylated in 60% yield (**2aw**). Various pharmaceuticals and peptides, including probenecid, bentiromide, meclofenamic acid, repaglinide, clonixin, tamibarotene, and ataluren, were all successfully hydroxylated at the *ortho* positions to their carboxylic acids with high efficiency (**2ax** to **2bd**), especially notable because these substrates are unreactive under ligandless conditions. These derivatives are not only valuable for potentially repurposing medicine for different diseases but are also useful for studying drug metabolism and pharmacokinetics vital in the discovery process.

REFERENCES AND NOTES

1. E. O'Reilly, V. Köhler, S. L. Flitsch, N. J. Turner, *Chem. Commun.* **47**, 2490–2501 (2011).
2. S. Chakrabarty, Y. Wang, J. C. Perkins, A. R. H. Narayan, *Chem. Soc. Rev.* **49**, 8137–8155 (2020).
3. J. Piera, J. E. Bäckvall, *Angew. Chem. Int. Ed.* **47**, 3506–3523 (2008).
4. G. J. ten Brink, I. W. Arends, R. A. Sheldon, *Science* **287**, 1636–1639 (2000).
5. E. M. Ferreira, B. M. Stoltz, *J. Am. Chem. Soc.* **123**, 7725–7726 (2001).
6. D. R. Jensen, J. S. Pugsley, M. S. Sigman, *J. Am. Chem. Soc.* **123**, 7475–7476 (2001).
7. S. S. Stahl, *Angew. Chem. Int. Ed.* **43**, 3400–3420 (2004).
8. T. C. Lau, C. M. Che, W. O. Lee, C. K. Poon, *J. Chem. Soc. Chem. Commun.* **21**, 1406–1407 (1988).
9. J. T. Groves, P. Viski, *J. Am. Chem. Soc.* **111**, 8537–8538 (1989).
10. M. W. Grinstaff, M. G. Hill, J. A. Labinger, H. B. Gray, *Science* **264**, 1311–1313 (1994).
11. S. Taktak, M. Flook, B. M. Foxman, L. Que Jr., E. V. Rybak-Akimova, *Chem. Commun.* 2005, 5301–5303 (2005).
12. S. Das, C. D. Incarvito, R. H. Crabtree, G. W. Brudvig, *Science* **312**, 1941–1943 (2006).
13. M. S. Chen, M. C. White, *Science* **318**, 783–787 (2007).
14. J. B. C. Mack *et al.*, *J. Am. Chem. Soc.* **141**, 972–980 (2019).
15. S. Hansson, A. Heumann, T. Rein, B. Åkermark, *J. Org. Chem.* **55**, 975–984 (1990).
16. P. L. Alsters, H. T. Teunissen, J. Boersma, G. van Koten, *Recl. Trav. Chim. Pays Bas* **109**, 487–489 (1990).
17. T. Yoneyama, R. H. Crabtree, *J. Mol. Catal. Chem.* **108**, 35–40 (1996).
18. A. R. Dick, K. L. Hull, M. S. Sanford, *J. Am. Chem. Soc.* **126**, 2300–2301 (2004).
19. R. Giri *et al.*, *Angew. Chem. Int. Ed.* **44**, 7420–7424 (2005).
20. Y. H. Zhang, J.-Q. Yu, *J. Am. Chem. Soc.* **131**, 14654–14655 (2009).
21. V. V. Rostovtsev, J. A. Labinger, J. E. Bercaw, T. L. Lasseter, K. I. Goldberg, *Organometallics* **17**, 4530–4531 (1998).
22. V. V. Rostovtsev, L. M. Henling, J. A. Labinger, J. E. Bercaw, *Inorg. Chem.* **41**, 3608–3619 (2002).
23. L. Boisvert, M. C. Denney, S. K. Hanson, K. I. Goldberg, *J. Am. Chem. Soc.* **131**, 15802–15814 (2009).
24. F. Tang, Y. Zhang, N. P. Rath, L. M. Mirica, *Organometallics* **31**, 6690–6696 (2012).
25. F. Qu, J. R. Khusnutdinova, N. P. Rath, L. M. Mirica, *Chem. Commun.* **50**, 3036–3039 (2014).
26. D. Wang, P. Y. Zavalij, A. N. Vedernikov, *Organometallics* **32**, 4882–4891 (2013).
27. P. Wang *et al.*, *Nature* **551**, 489–493 (2017).
28. V. Salamanca, A. Toledo, A. C. Albéniz, *J. Am. Chem. Soc.* **140**, 17851–17856 (2018).
29. A. S. Lindsey, H. Jeskey, *Chem. Rev.* **57**, 583–620 (1957).
30. L. Guo, M. Rueping, *Acc. Chem. Res.* **51**, 1185–1195 (2018).
31. C. Liu, C. L. Ji, X. Hong, M. Szostak, *Angew. Chem. Int. Ed.* **57**, 16721–16726 (2018).
32. D.-H. Wang, K. M. Engle, B.-F. Shi, J.-Q. Yu, *Science* **327**, 315–319 (2010).
33. C. A. Salazar *et al.*, *Science* **370**, 1454–1460 (2020).
34. Y.-J. Liu *et al.*, *Nature* **515**, 389–393 (2014).

FERROELECTRICS

Stacking-engineered ferroelectricity in bilayer boron nitride

Kenji Yasuda[1]*, Xirui Wang[1], Kenji Watanabe[2], Takashi Taniguchi[2], Pablo Jarillo-Herrero[1]*

Two-dimensional (2D) ferroelectrics with robust polarization down to atomic thicknesses provide building blocks for functional heterostructures. Experimental realization remains challenging because of the requirement of a layered polar crystal. Here, we demonstrate a rational design approach to engineering 2D ferroelectrics from a nonferroelectric parent compound by using van der Waals assembly. Parallel-stacked bilayer boron nitride exhibits out-of-plane electric polarization that reverses depending on the stacking order. The polarization switching is probed through the resistance of an adjacently stacked graphene sheet. Twisting the boron nitride sheets by a small angle changes the dynamics of switching because of the formation of moiré ferroelectricity with staggered polarization. The ferroelectricity persists to room temperature while keeping the high mobility of graphene, paving the way for potential ultrathin nonvolatile memory applications.

F erroelectric materials with an electric-field switchable polarization offer a wide range of technological applications, such as nonvolatile memories, high-permittivity dielectrics, electromechanical actuators, and pyroelectric sensors (1). Thinning down vertical ferroelectrics is one of the essential steps for the implementation of ferroelectric nonvolatile memory as part of the quest for

¹Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. ²National Institute for Materials Science, Namiki 1-1, Tsukuba, Ibaraki 305-0044, Japan.
*Corresponding author. Email: yasuda@mit.edu (K.Y.); pjarillo@mit.edu (P.J.-H.)

denser storage and lower power consumption (1). Room-temperature ferroelectricity down to atomic thicknesses was, however, difficult to access because of the depolarization effect until the recent development of three series of materials: epitaxial perovskites (2, 3), HfO_2-based ferroelectrics (4), and low-dimensional van der Waals (vdW) ferroelectrics (5–13). Among them, 2D vdW ferroelectrics present opportunities to integrate high-mobility materials such as graphene into ferroelectric field-effect transistors while keeping their properties intact, attributable to the absence of dangling bonds (14). Their uniform atomic thickness

also makes them ideal as ferroelectric tunnel barriers for use in ferroelectric tunnel junctions (15). Despite the potential importance for application as a ferroelectric memory, only a few examples of 2D vertical ferroelectrics—$CuInP_2S_6$, In_2Se_3, $MoTe_2$, and WTe_2—have been discovered so far (9–13); the candidate materials have been largely limited by the requirement of the polar space group in the original layered bulk crystal.

The development of vdW assembly enabled the engineering of heterostructures with physical properties beyond the sum of those of the individual layers (16). For example, the Dirac band structure of graphene is dramatically transformed when it is aligned with hexagonal boron nitride (hBN) or stacked with another slightly rotated graphene sheet. The modified band structures have led to the discovery of a variety of emergent phenomena related to electron correlations and topology beyond expectations from the original band structure (17–24). Here, we demonstrate that the vdW stacking modifies not only the electronic band structure but also the crystal symmetry, thereby enabling the design of ferroelectric materials out of nonferroelectric parent compounds. We use BN as an example, but the same procedure can be applied to other bipartite honeycomb 2D materials, such as 2H-type transition metal dichalcogenides (TMDs) (25). Bulk hBN crystals realize AA′ stacking, as shown in Fig. 1A. This 180°-rotated natural stacking order restores
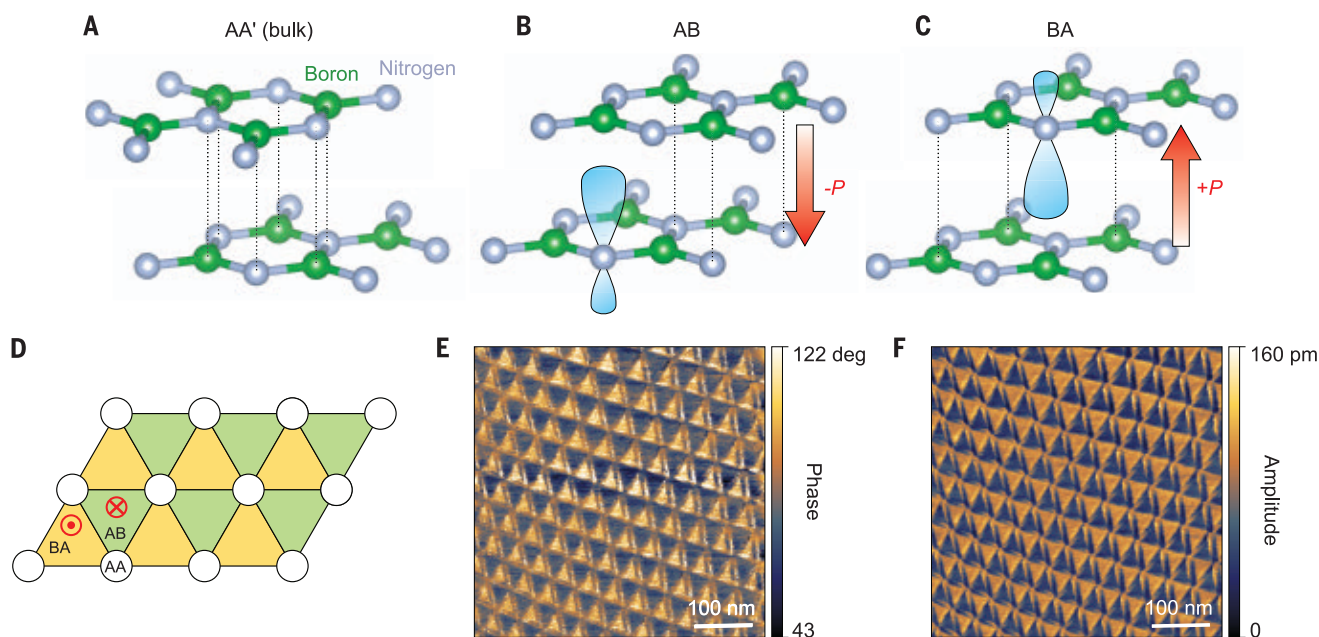
Fig. 1. Polarization in AB-stacked bilayer BN. (**A**) Illustration of the atomic arrangement for AA′ stacking, the bulk form of hBN. Nitrogen and boron atoms are shown in silver and green, respectively. (**B** and **C**) Illustration of the atomic arrangement for AB and BA stacking. The vertical alignment of nitrogen and boron atoms distorts the $2p_z$ orbital of nitrogen (light blue), creating an out-of-plane electric dipole. (**D**) Illustration of a small-angle twisted bilayer BN after the atomic reconstruction. The reconstruction creates relatively large AB (green) and BA (yellow) domains, with small AA regions (white) and domain walls in between (black). The red circled dot and red circled X represent up and down polarization, respectively. (**E** and **F**) Vertical PFM phase and amplitude images of twisted bilayer BN. Scale bars are 100 nm. The contrast at the domain wall, different from the AB and BA domain regions, likely originates from the flexoelectric effect (36, 42).
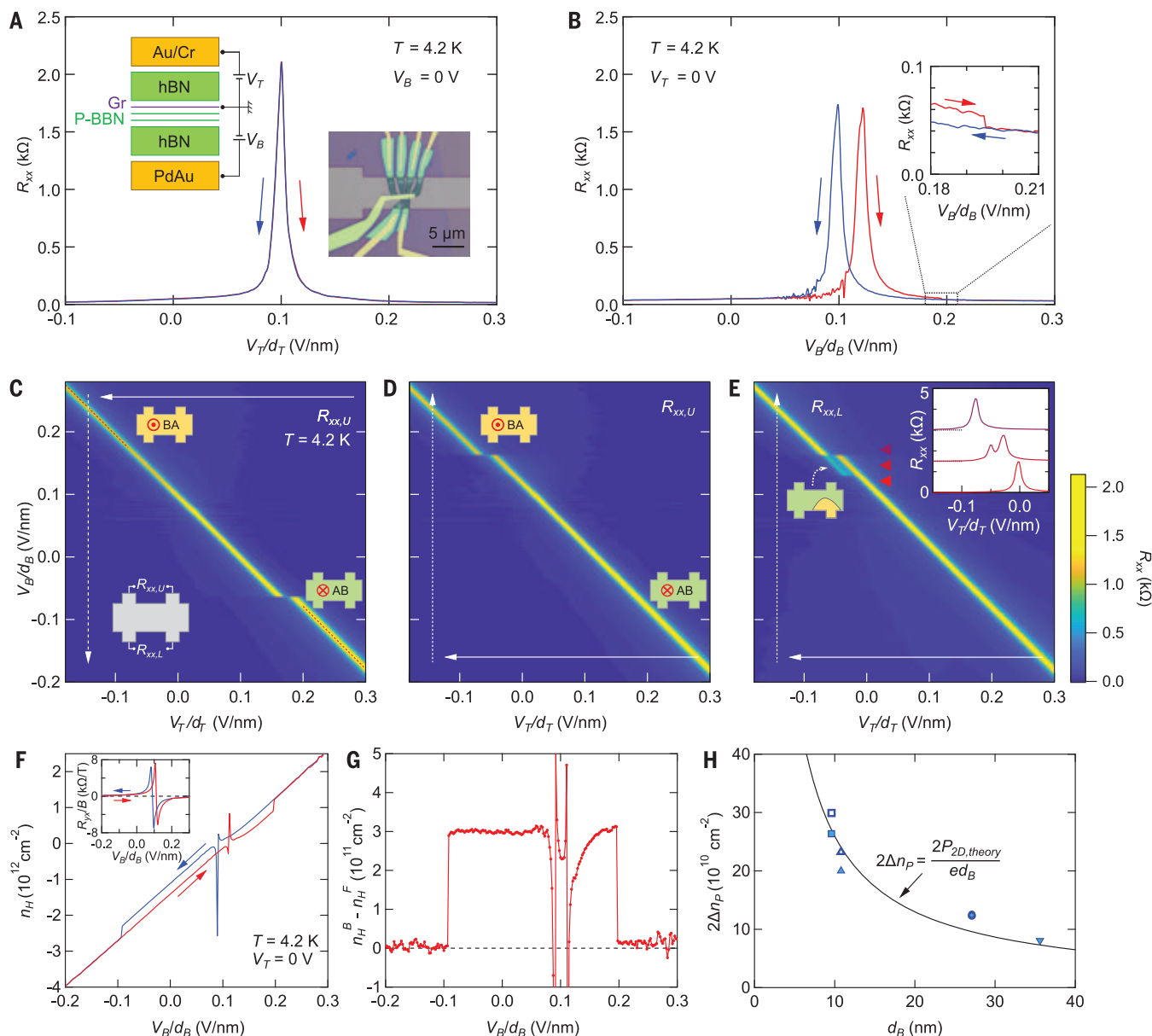
Fig. 2. Ferroelectric switching in parallel-stacked bilayer BN. (**A**) Resistance R_{xx} of graphene for device P1 as a function of V_T/d_T, the top gate voltage V_T divided by the thickness of top hBN d_T. V_T/d_T is scanned in the backward (forward) direction starting from +0.36 V/nm (−0.36 V/nm) in the blue (red) curve. Note that we only show the relevant scan range around the resistance peak in the figure. The inset on the left shows the schematic device structure. The inset on the right shows an optical micrograph of the device. Gr, graphene. (**B**) Resistance R_{xx} as a function of V_B/d_B, the bottom gate voltage V_B divided by the distance between graphene and bottom gate electrode d_B. V_B/d_B is scanned in the backward (forward) direction starting from +0.42 V/nm (−0.42 V/nm) in the blue (red) curve. The inset shows the enlarged plot around 0.20 V/nm. (**C**) Resistance $R_{xx,U}$ measured with the upper voltage contacts of device P1 (as displayed in the inset on the lower left) as a function of V_B/d_B and V_T/d_T. We repeatedly scanned V_T/d_T (fast scan, solid arrow) in the backward direction while gradually changing V_B/d_B (slow scan, dashed arrow). V_B/d_B is changed in the backward direction starting from +0.42 V/nm. Note that we only show the relevant scan range in the figure. The insets on the upper left and lower right show the schematic domain configuration and the polarization direction (red circled dot and red circled X).

The red dashed lines represent the charge neutrality points. (**D**) The same as (C) with V_B/d_B changed in the forward direction starting from −0.42 V/nm. (**E**) The same as (D) for the resistance $R_{xx,L}$ measured with the lower voltage contacts. The inset on the right shows the line cuts at the fixed V_B/d_B locations indicated by the red triangles. Each curve is offset by 1.5 kilohms for clarity. The inset on the left shows the schematic domain configuration during the ferroelectric switching (fig. S15). (**F**) Hall carrier density n_H measured as a function of V_B/d_B. V_B/d_B is scanned in the backward (forward) direction starting from +0.42 V/nm (−0.42 V/nm) in the blue (red) curve. The inset shows the Hall resistance R_{yx} as a function of V_B/d_B under magnetic field (B) = 0.5 T. (**G**) The difference of Hall carrier density in the backward and forward scan, $n_H^B - n_H^F$. (**H**) Twice the induced carrier density by the polarization of P-BBN, $2\Delta n_P$, plotted against d_B for four devices studied in this work. The $2\Delta n_P$ of each device is shown with a different shape; square (P1), triangle (P2), inverse-triangle (T1), and circle (T2). The filled and hollow symbols represent $2\Delta n_P$ estimated from the horizontal shift of the resistance peak and the Hall resistance, respectively. Note that two markers of device T2 almost overlap with each other. The black curve is the theoretical curve calculated from the polarization obtained from the Berry phase calculation, $P_{2D,theory}$ (25).

Fig. 3. Ferroelectric switching in twisted bilayer BN. (**A**) Resistance R_{xx} of device T1 as a function of V_B/d_B and V_T/d_T. The insets show schematic illustrations of the domain configurations. We repeatedly scanned V_T/d_T (fast scan, solid arrow) while gradually changing V_B/d_B (slow scan, dashed arrow). V_B/d_B is changed in the backward direction starting from +0.42 V/nm. The size of the domain is not to scale. (**B**) Spatial average of polarization of bilayer BN, $\langle P \rangle$, estimated from the two-peak fitting as a function of the applied electric field V_B/d_B for a twisted device, T1 (solid lines), and a nontwisted device, P1 (dashed lines). The blue and red curves are backward and forward scans, respectively. $\langle P \rangle$ of device P1 is estimated by taking

the average of the polarization measured with the upper voltage contacts and the lower voltage contacts. We expect that small, but finite, AB (BA) regions remain, even at $\langle P \rangle = 1(-1)$, in a twisted device as depicted in the insets of (A), although they are too small to be clearly detected with our resistance measurement scheme.
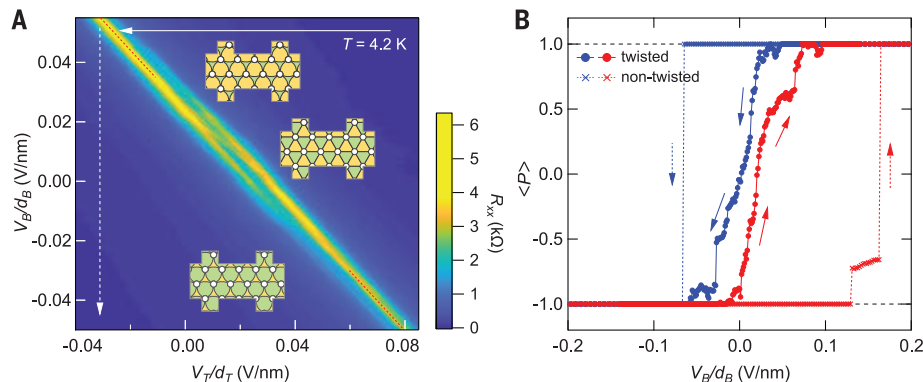
the inversion symmetry broken in the monolayer. However, if two BN monolayer sheets are stacked without rotation (parallel stacking, P), it has been theoretically (26, 27) and experimentally (28–31) shown that polar AB or BA stacking orders (Fig. 1, B and C, respectively) are formed. These configurations are local energy minima in a parallel-stacked form and are realized as metastable crystal structures (26, 27). In AB (BA) stacking, the B (N) atoms in the upper layer sit above the N (B) atoms in the lower layer while the N (B) atoms in the upper layer lay above the empty site at the center of the hexagon in the lower layer. The vertical alignment of the $2p_z$ orbitals of N and B distorts the orbital of N, creating an electric dipole moment (fig. S3). As a result, AB and BA stacking will exhibit out-of-plane polarization in the opposite directions (25).

We demonstrate the polarization of AB-stacked bilayer BN by vertical piezoelectric force microscopy (PFM). We fabricated nearly 0° bilayer BN devices by using the "tear and stack" method, where half of a monolayer BN flake is picked up and stacked on top of the remaining half (32, 33). In twisted bilayer BN, lattice relaxation leads to the formation of a moiré pattern consisting of AB and BA lattice networks with topological defects (AA regions), as in the case of twisted bilayer graphene and TMDs (Fig. 1D) (34–36). However, unlike twisted bilayer graphene, the low crystalline symmetry of BN creates a distinctive moiré pattern with staggered polarization in the AB and BA domains (25). PFM measurements on a small-angle twisted bilayer BN show a triangular pattern with finite contrast between the adjacent triangles (Fig. 1, E and F), whereas no moiré pattern is observed in the topographic image (fig. S6). The different piezoresponse in the AB and BA domains evidences the opposite out-of-plane polarizations in these domains. In a larger area scan (figs. S7 and S8), the periodicity of the triangular pattern varies at wrinkles and bubbles as the rotational angle changes. The triangular contrast

does not show up in the monolayer BN region, confirming the interlayer interaction origin of the polarization (figs. S7 and S8). The stacking order–dependent out-of-plane polarization presents the interesting possibility that the polarization can be switched by an in-plane interlayer shear motion of one-third of the unit cell (25), which is distinct from the switching mechanism of conventional ferroelectrics (1).

To study the change of the polarization under the electric field, we fabricated dual-gated vdW heterostructure devices composed of metal top gate (Au/Cr)/hBN/graphene/0° parallel stacked bilayer BN (P-BBN)/hBN/metal bottom gate (PdAu) (e.g., device P1), as schematically shown in the inset of Fig. 2A. Zero-degree stacking of P-BBN allows the entire region to be a single domain of AB or BA stacking without forming the moiré pattern. Here, the graphene sensitively detects the extra charge carriers induced by the polarization of P-BBN. Figure 2A shows the resistance of the graphene sensor as a function of the top gate voltage, V_T (for both forward and backward gate sweep directions), which exhibits a typical maximum without hysteresis. By contrast, the forward and backward scans of the resistance versus the bottom gate voltage, V_B (Fig. 2B) shows hysteresis, exhibiting maxima at about 0.10 and 0.12 V/nm for the backward and forward scans, respectively. In addition, we observe a resistance step at around 0.20 V/nm in the forward scan, as displayed in the inset. As discussed later, this bistability is attributed to the polarization switching of P-BBN by the applied electric field.

Dual-gate scanning allows independent control of the carrier density of graphene and the electric field across the P-BBN, because the top gate primarily changes the former (figs. S10 and S11), whereas the bottom gate changes both. In a standard dual-gated graphene device, a measurement of the resistance versus the top and bottom gate voltages results in a single diagonal feature, a maximum resistance ridge, corresponding to the charge neutrality condi-

tion. The diagonal feature stems from the fact that the induced carrier density follows the equation $n = \varepsilon_{hBN}(V_B/d_B + V_T/d_T)/e$, where ε_{hBN} is the dielectric constant of hBN, d_B (d_T) is the distance between graphene and the bottom (top) gate electrode, and e is the elemental charge. By contrast, two parallel-shifted diagonal lines are observed in a dual-gate scan for our P-BBN device (Fig. 2C). The shift reflects an abrupt change in the induced carrier density, Δn_P, caused by the switching of the electric polarization of P-BBN: As the polarization switches from up (BA stacking) to down (AB stacking) at $V_B/d_B = -0.06$ V/nm, the total induced carrier density changes from $\varepsilon_{hBN}(V_B/d_B + V_T/d_T)/e + \Delta n_P$ to $\varepsilon_{hBN}(V_B/d_B + V_T/d_T)/e - \Delta n_P$, leading to the shift of the charge neutrality resistance peak. Similarly, the forward scan of the bottom gate shows the polarization switching from down to up at $V_B/d_B = 0.16$ V/nm (Fig. 2D). Notably, the resistance measured using the lower voltage contacts exhibits an intermediate, two-peak behavior during the switching (Fig. 2E). This indicates the coexistence of micrometer-scale AB and BA domains and provides a hint to the dynamics of polarization switching. Namely, the domain wall is pinned in the middle of the Hall bar at around $V_B/d_B = 0.13$ V/nm, followed by the depinning at around $V_B/d_B = 0.16$ V/nm.

To further investigate the ferroelectric properties of P-BBN, we measured the carrier density n_H of graphene extracted from Hall resistance measurements (Fig. 2F). Hysteretic behavior with an abrupt jump in n_H is observed when sweeping V_B, which is attributed to the ferroelectric switching. The subtraction of the forward and backward sweeps gives a magnitude of $2\Delta n_P$, which equals 3.0×10^{11} cm^{-2} (Fig. 2G). This value is consistent with $2\Delta n_P = 2.6 \times 10^{11}$ cm^{-2} estimated from the horizontal shift of the charge neutrality resistance peak in the dual-gate scan (Fig. 2, C to E). Δn_P allows us to calculate the magnitude of the polarization of AB-stacked bilayer BN. According to a simple model calculation (see fig. S14 for details), the
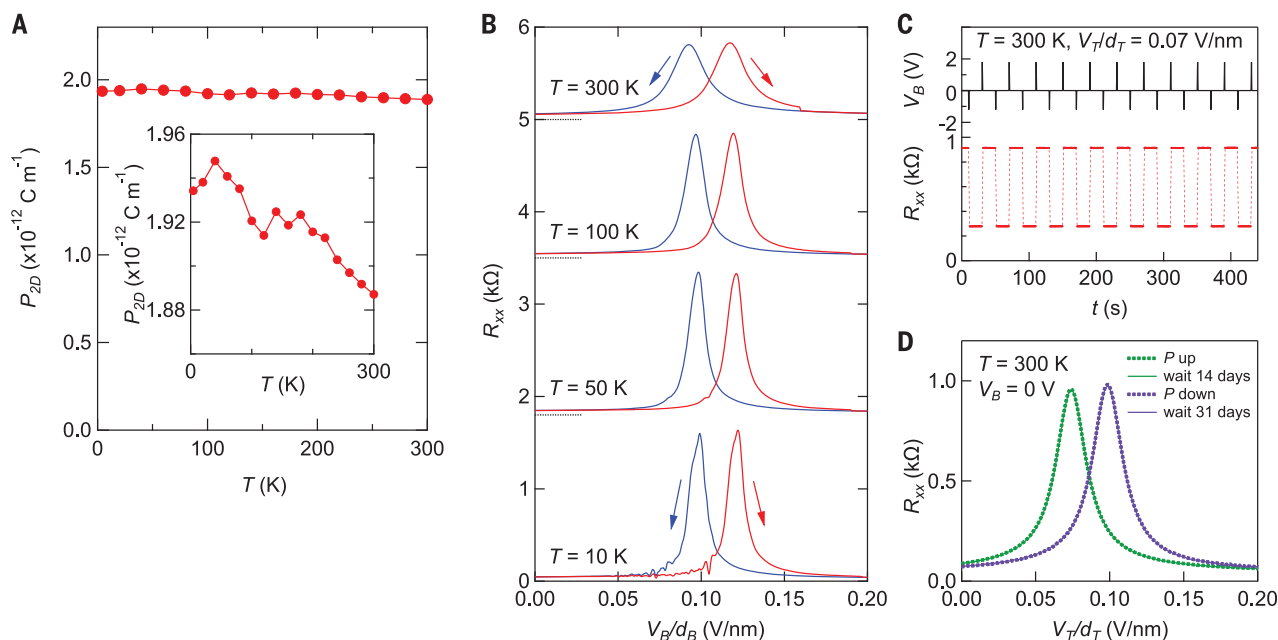
Fig. 4. Room-temperature operation of a ferroelectric field-effect transistor.
(**A**) Temperature dependence of the magnitude of the polarization P_{2D} for device P1. The inset shows a zoom-in of the vertical axis. (**B**) Hysteresis of resistance at various temperatures. V_B/d_B is scanned in the backward (forward) direction starting from +0.42 V/nm (−0.42 V/nm) in the blue (red) curve. Each curve is offset for clarity. The offset values are shown as dotted lines. (**C**) Resistance (red curve) after the repeated application of a voltage pulse of V_B = +1.8 V and V_B = −1.2 V (black curve), which corresponds to V_B/d_B = +0.19 and −0.13 V/nm, respectively. The measurement is performed at T = 300 K and V_T/d_T = 0.07 V/nm. (**D**) Stability of polarization at room temperature. V_T/d_T is scanned in the forward direction. The dotted green (purple) curve is measured at V_B = 0 V right after applying V_B/d_B = +0.31 V/nm (−0.26 V/nm) to induce polarization up (down). The solid green curve is measured after applying V_B/d_B = +0.31 V/nm to induce polarization up and then leaving the device at V_B = 0 V and T = 300 K for 14 days. The solid purple curve is measured after applying V_B/d_B = −0.26 V/nm to induce polarization down and then leaving the device at V_B = 0 V and T = 300 K for 31 days. Each of the two curves almost exactly overlap, showing the robustness of polarization direction for a long period.

2D polarization follows $P_{2D} = e\Delta n_P d_B$; namely, the electric dipole moment between the bottom gate and graphene is equal to the magnitude of the polarization of bilayer BN. Figure 2H shows our measurement of $2\Delta n_P$ for four different devices studied in this work, which indeed exhibit an inverse proportional behavior with respect to d_B. The magnitude of the polarization estimated from these data points is P_{2D} = 2.25 (0.37) × 10^{-12} C m^{-1} (corresponding to P_{3D} = 0.68 μC cm^{-2}). This agrees well with the theoretically calculated magnitude of the polarization of AB-stacked bilayer BN from a Berry phase calculation, $P_{2D,theory}$ = 2.08 × 10^{-12} C m^{-1} (25, 37).

Having established the ferroelectric nature of P-BBN, we next studied how the moiré superlattice affects the ferroelectric switching in a small-angle twisted bilayer BN. Here, owing to the opposite polarization of AB and BA stacking regions (Fig. 1, D to F), each domain with staggered polarization is expected to expand or shrink, through domain wall motion, when a vertical electric field is applied. Figure 3A shows the dual-gate scan of the resistance of graphene for a 0.6°–rotation angle twisted bilayer BN (device T1). It exhibits two parallel diagonal peaks, each corresponding to the AB or BA domains, similar to Fig. 2C. However, rather than an abrupt

transition between the two lines, a gradual shift in weight from one to the other takes place along the diagonal. Thus, the magnitude of each peak gives the relative proportion of AB and BA domain sizes, or the average polarization, as a function of the applied electric field (fig. S17). The electric field dependence of the polarization (Fig. 3B) highlights the difference between the twisted and nontwisted devices. First, the coercive field is much smaller for the twisted bilayer BN than for the nontwisted P-BBN. Secondly, the polarization switching occurs gradually, in contrast to the sharp switching of the nontwisted device. In a nontwisted device, a domain wall moves over the device scale during the switching, as shown in Fig. 2, C to E, and is likely to be pinned by strong pinning centers. By contrast, each domain wall in a twisted device moves only by a moiré length scale and will experience weaker pinning centers, leading to the small coercive field. In addition, the different pinning strength of each domain wall leads to the gradual switching. Thus, the global rotation of the two layers modifies the dynamics of the ferroelectric switching behavior.

Finally, we studied the temperature dependence of the ferroelectricity in P-BBN. Notably, the polarization measured from Δn_P is almost independent of temperature (Fig. 4A

and fig. S22) up to room temperature. The nearly temperature-independent ferroelectric polarization presumably reflects the distinctive coupling between the out-of-plane polarization and the in-plane shear motion in P-BBN. The strong intralayer covalent bonding inhibits the in-plane thermal vibration of atoms, making the polarization insensitive to temperature (38). Correspondingly, the ferroelectric hysteresis is observable up to room temperature despite the temperature-induced broadening of the resistance peak (Fig. 4B). Such hysteretic behavior allows us to deterministically write the polarization by a voltage pulse of only a few volts and read it in a nonvolatile way, as shown in Fig. 4C. We also checked the stability of the ferroelectric polarization by keeping the sample at 0 V at room temperature for an extended period after setting the polarization to up or down (Fig. 4D). The resistance remains almost the same after at least a month (the longest period measured); namely, P-BBN retains its polarization over a technologically relevant time scale. Hence, the present result points to the potential use of P-BBN/graphene as a ferroelectric field-effect transistor with an ultrahigh mobility of graphene of around 5 × 10^4 cm^2 V^{-1} s^{-1} at room temperature (figs. S19 to S21).

The designer approach for engineering vdW ferroelectrics and moiré ferroelectrics

demonstrated in this study can be extended to other bipartite honeycomb 2D materials, such as semiconducting 2H-type TMDs like MoS_2 and WSe_2, metallic and superconducting ones like NbS_2 and $NbSe_2$, and group III chalcogenides like GaS, GaSe, and InSe (25). The inversion symmetry breaking of these synthetic ferroelectrics will be coupled to the electronic band structures in a tunable manner through polarization switching. In addition to interesting physics resulting from the modification of the intrinsic properties of each material, such engineered ferroelectrics and moiré systems may substantially expand the capabilities of 2D materials for electronic, spintronic, and optical applications (15, 39).

We note that (40) and a paper in the same issue (41) report related findings.

REFERENCES AND NOTES

1. K. Uchino, *Ferroelectric Devices* (CRC Press, 2009).
2. D. D. Fong et al., *Science* **304**, 1650–1653 (2004).
3. H. Wang et al., *Nat. Commun.* **9**, 3319 (2018).
4. U. Schröeder, C. S. Hwang, H. Funakubo, *Ferroelectricity in Doped Hafnium Oxide: Materials, Properties and Devices* (Woodhead Publishing, 2019).
5. C. Cui, F. Xue, W.-J. Hu, L.-J. Li, *npj 2D Mater. Appl.* **2**, 18 (2018).
6. M. Wu, P. Jena, *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **8**, e1365 (2018).
7. A. V. Bune et al., *Nature* **391**, 874–877 (1998).
8. K. Chang et al., *Science* **353**, 274–278 (2016).
9. F. Liu et al., *Nat. Commun.* **7**, 12357 (2016).
10. Y. Zhou et al., *Nano Lett.* **17**, 5508–5513 (2017).
11. C. Cui et al., *Nano Lett.* **18**, 1253–1258 (2018).
12. S. Yuan et al., *Nat. Commun.* **10**, 1775 (2019).
13. Z. Fei et al., *Nature* **560**, 336–339 (2018).
14. C. R. Dean et al., *Nat. Nanotechnol.* **5**, 722–726 (2010).
15. E. Y. Tsymbal, H. Kohlstedt, *Science* **313**, 181–183 (2006).
16. A. K. Geim, I. V. Grigorieva, *Nature* **499**, 419–425 (2013).
17. B. Hunt et al., *Science* **340**, 1427–1430 (2013).
18. C. R. Dean et al., *Nature* **497**, 598–602 (2013).
19. L. A. Ponomarenko et al., *Nature* **497**, 594–597 (2013).
20. R. V. Gorbachev et al., *Science* **346**, 448–451 (2014).
21. Y. Cao et al., *Nature* **556**, 43–50 (2018).
22. Y. Cao et al., *Nature* **556**, 80–84 (2018).
23. A. L. Sharpe et al., *Science* **365**, 605–608 (2019).
24. M. Serlin et al., *Science* **367**, 900–903 (2020).
25. L. Li, M. Wu, *ACS Nano* **11**, 6382–6388 (2017).
26. G. Constantinescu, A. Kuc, T. Heine, *Phys. Rev. Lett.* **111**, 036104 (2013).
27. S. Zhou, J. Han, S. Dai, J. Sun, D. J. Srolovitz, *Phys. Rev. B* **92**, 155438 (2015).
28. J. H. Warner, M. H. Rümmeli, A. Bachmatiuk, B. Büchner, *ACS Nano* **4**, 1299–1304 (2010).
29. C.-J. Kim et al., *Nano Lett.* **13**, 5660–5665 (2013).
30. S. M. Gilbert et al., *2D Mater.* **6**, 021006 (2019).
31. H. J. Park et al., *Sci. Adv.* **6**, eaay4958 (2020).
32. K. Kim et al., *Nano Lett.* **16**, 1989–1995 (2016).
33. Y. Cao et al., *Phys. Rev. Lett.* **117**, 116804 (2016).
34. H. Yoo et al., *Nat. Mater.* **18**, 448–453 (2019).
35. A. Weston et al., *Nat. Nanotechnol.* **15**, 592–597 (2020).
36. L. J. McGilly et al., *Nat. Nanotechnol.* **15**, 580–584 (2020).
37. R. D. King-Smith, D. Vanderbilt, *Phys. Rev. B Condens. Matter* **47**, 1651–1654 (1993).
38. Q. Yang, M. Wu, J. Li, *J. Phys. Chem. Lett.* **9**, 7160–7164 (2018).
39. J. Sung et al., *Nat. Nanotechnol.* **15**, 750–754 (2020).
40. C. R. Woods et al., *Nat. Commun.* **12**, 347 (2021).
41. M. Vizner Stern et al., *Science* **372**, 1462–1466 (2021).
42. See supplementary materials.
43. K. Yasuda, X. Wang, K. Watanabe, T. Taniguchi, P. Jarillo-Herrero, Replication Data for: Stacking-engineered ferroelectricity in bilayer boron nitride. Harvard Dataverse (2021); https://doi.org/10.7910/DVN/JNXOIM.

FERROELECTRICS

Interfacial ferroelectricity by van der Waals sliding

M. Vizner Stern[1], Y. Waschitz[1], W. Cao[2], I. Nevo[1], K. Watanabe[3], T. Taniguchi[3], E. Sela[1], M. Urbakh[2], O. Hod[2], M. Ben Shalom[1]*

Despite their partial ionic nature, many-layered diatomic crystals avoid internal electric polarization by forming a centrosymmetric lattice at their optimal van der Waals stacking. Here, we report a stable ferroelectric order emerging at the interface between two naturally grown flakes of hexagonal boron nitride, which are stacked together in a metastable non-centrosymmetric parallel orientation. We observe alternating domains of inverted normal polarization, caused by a lateral shift of one lattice site between the domains. Reversible polarization switching coupled to lateral sliding is achieved by scanning a biased tip above the surface. Our calculations trace the origin of the phenomenon to a subtle interplay between charge redistribution and ionic displacement and provide intuitive insights to explore the interfacial polarization and its distinctive "slidetronics" switching mechanism.

T he ability to locally switch a confined electrical polarization is a key functionality in modern technologies, where storing and retrieving a large volume of information is vital (1). The need to reduce the dimensions of individually polarized domains, from the ~100-nm² scale (2, 3) toward the atomic scale, is rising (4). The main challenges involve long-range dipole interactions, which tend to couple the individual domain polarization orientations (5). Likewise, surface effects and external strains that are difficult to control become dominant once the surface-to-volume ratio increases (6). To overcome these challenges, one can consider layered materials, such as hexagonal boron nitride (h-BN) and transition-metal dichalcogenides (TMDs), where the bulk volume can be reduced to the ultimate atomic-thickness limit, whereas the crystalline surface remains intact (7). However, it is rare to find a spontaneous net electric polarization in two dimensions (2D) that is sufficiently large to read and write under ambient conditions (8–10). Furthermore, in naturally grown h-BN and TMD crystals, polarization is eliminated by the formation of a centrosymmetric van der Waals (vdW) structure that is lower in energy than other metastable stacking configurations. Here, we break this symmetry by controlling the twist angle between two h-BN flakes and find an array of permanent and switchable polarization domains at their interface. The polarization is oriented normal to the plane, and its amplitude is in good agreement with previous first-principle predictions for a two-layered system (11) and with our detailed multilayer calculations.

To identify which stacking modes can carry polarization, we present in Fig. 1A six different high-symmetry configurations of bilayer h-BN. The stacking configurations are divided into two groups termed "parallel" and "antiparallel" twist orientations (12); within each group, a relative lateral shift by one interatomic distance switches the stacking configuration in a cyclic manner. Typically, the crystal grows in the optimal antiparallel (AA') configuration

¹School of Physics and Astronomy, Tel Aviv University, Israel. ²Department of Physical Chemistry, School of Chemistry, The Raymond and Beverly Sackler Faculty of Exact Sciences and The Sackler Center for Computational Molecular and Materials Science, Tel Aviv University, Tel Aviv 6997801, Israel. ³National Institute for Material Science, Tsukuba, Japan.
*Corresponding author. Email: moshebs@tauex.tau.ac.il

Fig. 1. High-symmetry interlayer stacking configurations. (**A**) Top view illustration of two layers. For clarity, atoms of the top layer are represented by smaller circles. Within each group of parallel or antiparallel twist orientations, a relative lateral shift by one lattice spacing results in a cyclic switching between three high-symmetry stacking configurations. (**B**) Calculated local-registry index (LRI) map of the atomic overlaps (*17*) in a rigid structure made of two *h*BN layers stacked with a twist angle of 0.5°. Blue regions correspond to AA stacking, whereas AB/BA stacking appears in orange (LRI = 0.86). (**C**) Calculated LRI map after geometry relaxation of the structure presented in (B). Large domains of uniform untwisted AB/BA stacking appear, at the expense of the preoptimized AA regions. The twist is accumulated in smaller AA-like regions and in the ~10-nm-wide incommensurate domain walls (bright lines); see fig. S3 and (*17*) for further discussion. (**D**) Cross-sectional illustration of two few-layered flakes (blue and light blue regions mark the top and bottom flakes, respectively) of naturally grown *h*-BN (AA'), which are stacked with no twist. Plus (minus) signs mark boron (nitrogen) sites. A topographical step of a single-layer switches between parallel and antiparallel stacking orientations at the interface between the two flakes. Vertical charge displacements in eclipsed/hollow pairs (vertical/diagonal ellipses) and the resulting net polarization P$_z$ are marked by arrows.
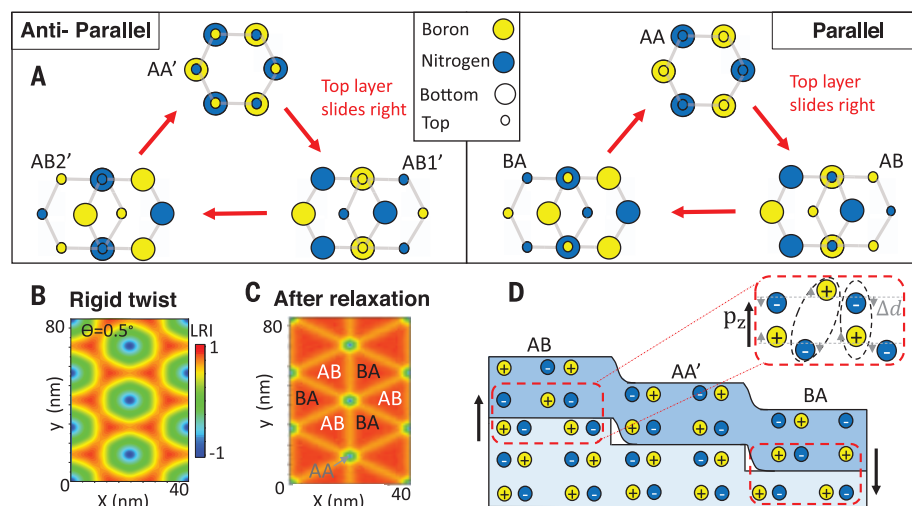
with full overlap between nitrogen (boron) atoms of one layer and boron (nitrogen) atoms of the adjacent layer (*13*). In the parallel twist orientation, however, the fully eclipsed configuration (AA) is unstable because it forces pairs of bulky nitrogen atoms atop each other, resulting in increased steric repulsion (*14*). Instead, a lateral interlayer shift occurs to a metastable AB stacking with only half of the atoms overlapping, whereas the other half are aligned with the empty centers of the hexagons in the adjacent layer (*15, 16*). The AB and BA stacking form equivalent lattice structures (only flipped), and all depicted antiparallel configurations (AA', AB1', AB2') are symmetric under spatial inversion.

To explore these different configurations, we artificially stamped two exfoliated *h*-BN flakes on top of each other, each consisting of a few AA' stacked layers, with a minute twist angle between the otherwise parallel interfacial layers (*17*). The small twist imposes interlayer translations that evolve continuously and form a moiré pattern owing to the underlying crystal periodicity (Fig. 1B). In this rigid lattice picture, the three nearly commensurate stacking configurations (AB, BA, AA) appear at adjacent positions in space. Notably, this picture breaks for a sufficiently small twist angle as a result of structural relaxation processes, as shown by our molecular dynamics calculation based on dedicated interlayer potentials (Fig. 1C) (*17, 18*). Instead, the system divides into large domains of reconstructed commensurate AB and BA stackings separated by sharp incommensurate domain walls that accommodate the global twist (fig. S3, A and B) (*19–22*). Notably, near the center of the extended commensurate domains, perfectly aligned configurations

are obtained with no interlayer twist. In the experiments, we also introduce a topographic step at the interface between the flakes. A step thickness of an odd number of layers guarantees antiparallel stacking (AA', AB1', or AB2') on one side and parallel stacking (AA, AB, or BA) on its other side (Fig. 1D). Thus, one can compare all possible configurations at adjacent locations in space.

To measure variations in the electrical potential, V_{KP}, at surface regions of different stacking configurations, we place the *h*-BN sandwich on a conducting substrate (Si/SiO$_2$, graphite, or gold) and scan an atomic force microscope operated in a Kelvin probe mode (KPFM) (Fig. 2A) (*17*). The potential map above the various stacking configurations is shown in Fig. 2B. We find clear domains (black and white) of constant V_{KP}, extending over areas of several square micrometers, which are separated by narrow domain walls. Dark gray areas of average potential are observed above (i) positions where only one *h*-BN flake exists (outside the dashed yellow line); (ii) above two flakes but beyond the topographic step marked by dashed red lines in Fig. 2B (and topography map fig. S2), as expected; and (iii) beyond topographic folds that can further modify the interlayer twist angle (dashed green lines). These findings confirm that white and black domains correspond to AB and BA interfacial stacking that host a permanent out-of-plane electric polarization. Such polarization is not observed at the other side of the step, where centrosymmetric AA', AB1', AB2' configurations are obtained, or at the AA configuration expected at domain-wall crossings (blue dots in Fig. 1C). Sufficiently far from the domain walls (toward the center of each domain), a

constant potential is observed with a clear difference ΔV_{KP} between the AB and BA domains, as shown in Fig. 2C. Whereas KPFM measurement nullifying the tip response at the electric bias frequency gives an underestimated potential difference of $\Delta V_{KP} \sim 100$ mV because of averaging contributions from the tip's cantilever (*17*), more quantitative measurements obtained through sideband tip excitations yield ΔV_{KP} values ranging between 210 and 230 mV for both closed-loop scans and local open-loop measurements (fig. S1). Similar values are measured for several samples with different substrate identities and various thicknesses of the top *h*-BN flake (for flakes thicker than 1 nm), and when using different AFM tips. These findings confirm that ΔV_{KP} is an independent measure of the intrinsic polarization of the system that, in turn, is confined within a few interfacial layers.

Although our force field calculations for slightly twisted bilayer *h*-BN show a uniform triangular lattice of alternating AB and BA stacked domains (Fig. 1C), in the experiment we observe large variations in their lateral dimensions and shape. This indicates minute deviations in the local twist, which are unavoidable in the case of small twist angles (*19–22*). Specifically, the ~1-μm^2 domains in the left part of Fig. 2B correspond to a global twist of less than 0.01° (*23*). In addition, any external perturbation to the structure, caused either by transferring it to a polymer, heating, or directly pressing it with the AFM tip, usually resulted in a further increase in domain size. In a few cases, high-temperature annealing resulted in a global reorientation to a single domain flake, many micrometers in dimensions. This behavior confirms the metastable
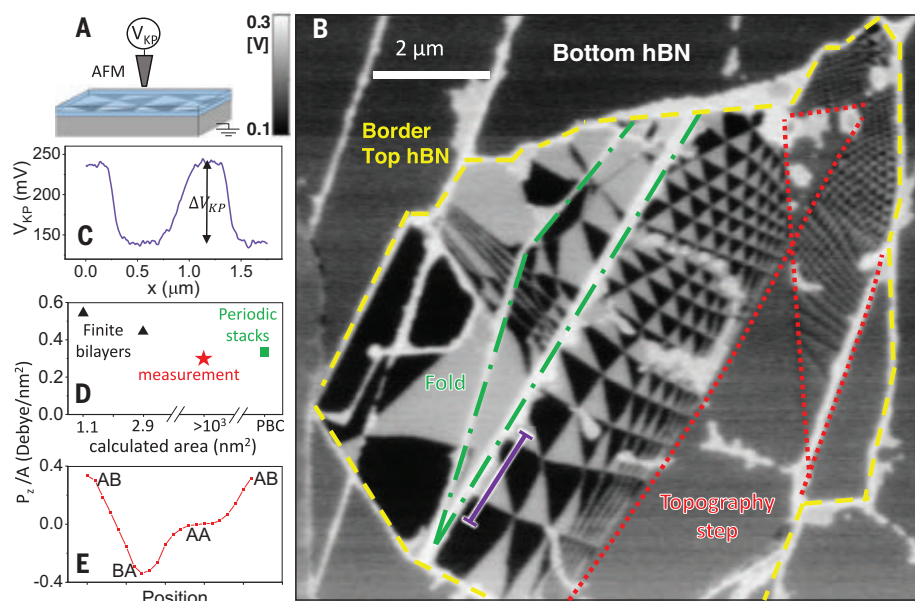
Fig. 2. Direct measurement of interfacial polarization. (**A**) Illustration of the experimental setup. An atomic force microscope is operated in Kelvin-probe mode to measure the local potential modulation, V_{KP}, at the surface of two 3-nm-thick h-BN flakes, which are stacked with a very small twist angle. (**B**) V_{KP} map showing oppositely polarized domains of AB/BA stacking (black and white), ranging in area between ~0.01 and 1 μm^2 and separated by sharp domain walls. (**C**) Surface potential measured along the purple line marked in (B) by first-harmonic KPFM. (**D**) DFT calculations of the polarization, P_z, per unit area obtained for finite h-BN bilayer flakes of different lateral dimensions (1.1 and 2.9 nm^2, black triangles) and for laterally periodic stacks made of 2 to 10 layers (marked as PBC, green square). The red star marks the polarization value evaluated from the measured ΔV_{KP} data. (**E**) Calculated polarization for different interlayer shifts.

nature of the AB/BA stacking mode, as well as the possible superlubric nature of the interface (*24, 25*). At the other extreme, much smaller domains are observed in the top right-hand section of Fig. 2B. The smallest triangle edge that we could identify over many similar flakes was 60 nm in length, which corresponds to a twist angle of 0.24°. We therefore conclude that, below this angle, atomic reconstruction to create untwisted domains is energetically favored. Naturally, this constitutes a lower bound on the maximal angle for domain formation as smaller domains below our experimental resolution may form at larger twist angles.

To trace the microscopic origin of the measured polarization, we performed a set of density functional theory (DFT) calculations on finite bilayer and quad-layer h-BN flakes. For the finite bilayer calculations, we constructed two model systems, where hydrogen passivated h-BN flakes of either 1 or 3 nm^2 surface area are stacked in the AB stacking mode (fig. S4). The calculated polarizations per unit area, P_z/A, of these systems were 0.55 and 0.45 Debye/nm^2, respectively (black triangles in Fig. 2D), pointing perpendicular to the interface only (table S1). Because edge effects may influence the results of such finite system calculations (*17*) (fig. S4), we performed com-

plementary laterally periodic system calculations at various thicknesses. The detailed methods used for these calculations are discussed in (*17*). For the AB stacked periodic bilayer, we find a polarization of $P_z/A =$ 0.33 Debye/nm^2, changing by only 10% when including additional two and eight AA' stacked layers above and below the AB stacked interface (see Fig. 2D and fig. S7). Adding more AB-stacked layers, however, reveals a linear increase of the calculated polarization with the number of added interfaces (fig. S9). The latter is a highly appealing control mechanism to engineer the magnitude of the polarization in future 2D systems, independent of external effects such as surface chemistry and local strains. Lateral shifts of the periodic bilayer system show a gradual evolution of the polarization when shifting between the AA, AB, and BA stacking modes, from zero to +0.33 and −0.33 Debye/nm^2, respectively (Fig. 2E). This is crucial when considering the complex response at domain walls, where lattice deformations induce additional flexo (*26*) and piezo (*27*) responses. In fig. S8, we present also the charge density redistribution in the periodic bilayer system owing to interlayer coupling. The corresponding interlayer potential difference at the experimental configuration is calculated to be $\frac{1}{2}\Delta V_{KP} = 120$mV (fig. S7),

in excellent agreement with the side-band measurements (red star in Fig. 2D). Similar results were reported in recent computational studies (*11*). However, to obtain quantitative agreement with the experimental measurements, the results should be carefully converged with respect to the various calculation parameters (figs. S5 and S7). By further applying an electric field of 0.1 V/nm normal to the plane, we find, from our calculations, a minute piezo-electric deformation of 0.1 pm and a 5% variation in the polarization magnitude (fig. S10).

It is instructive to further translate the measured potential difference into intralayer displacements in a simplistic point-like charges model (Δd in Fig. 1D), where each atom is allowed to displace from its layer's basal plane in the vertical direction. With the lattice site density of $n = 37$ nm^{-2} and the on-site charge value, $q \sim e/2$, for single-layered h-BN (*28, 29*), our measured ΔV_{KP} gives out-of-plane atomic displacement of the order of $\Delta d = \Delta V_{KP}\varepsilon_0/ 4nq \sim 1 \times 10^{-3}$Å ($\varepsilon_0$ is vacuum permittivity), which is much smaller than the intralayer (1.44 Å) and interlayer (3.30 Å) spacings. This implies that the polarization is determined by a delicate competition between the various interlayer interaction components and charge redistribution. Intuitively, we expect the vdW attraction to vertically compress the nonoverlapping interfacial sites (diagonal dashed ellipse in Fig. 1D) closer together than the overlapping sites (vertical dashed ellipse in Fig. 1D), which are more prone to Pauli repulsion. This direction of motion, for example, reduces the average interlayer separation and favors Bernal (AB like) stacking in graphite over the AA configuration (*30*). In h-BN, however, the partial ionic nature of the two lattice sites (*12, 31, 32*) makes the fully eclipsed AA' stacking more stable (*13*). Hence, imposing a polar AB interface, as in our case, may favor overlapping sites of opposite charges to come closer together than the nonoverlapping pairs and the polarization to point in the opposite direction.

To quantify these arguments, we present a reduced classical bilayer model that captures the intricate balance between Pauli, vdW, and Coulomb interatomic interactions at different stacking modes. In our model (*17*), the interfacial energy $E = \frac{1}{2}\sum_{i,j}\left[4\varepsilon\left(\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^6\right) + \frac{q_iq_j}{r_{ij}}\right]$ includes a Lennard-Jones (LJ) potential characterized by the cohesive energy, ε, and the interlayer spacing scale, σ, and Coulomb interactions between the dimensionless partial atomic charges on the boron and nitrogen sites $q = \pm q_i/e$. Although neglecting any charge transfer processes between the layers that are explicitly taken into account in our DFT calculations, this model captures both the magnitude and orientation of the polarization by

Fig. 3. Dynamic flipping of polarization orientation by domain-wall sliding. Kelvin-probe maps measured consecutively from left to right above a particular flake location showing domains of up (white) and down (black) polarizations. The middle image was taken after

biasing the tip by a fixed DC voltage of −20 V and scanning it above the blue square region shown on the left-hand image. Then the tip was biased by 10 V and scanned again over the same region before taking the right-hand image. Consecutive domain-wall positions are marked by dashed red, green, and yellow lines. Larger white (black) domains appear after positive (negative) bias scans as a result of domain-wall motion beyond the scan area. Note that the number of domain walls is apparently not altered.
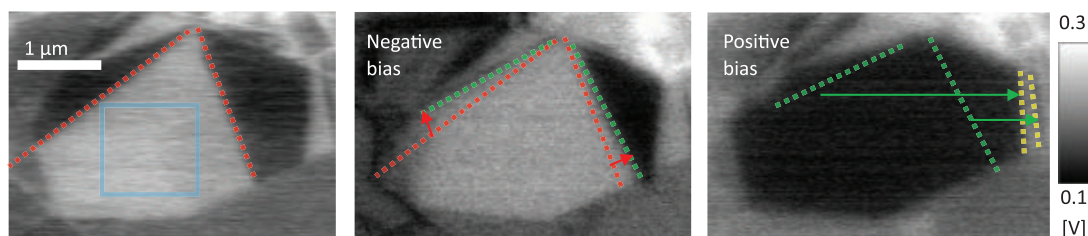
adjusting the ratio between ε and Coulomb scales $\propto \varepsilon / q^2$ (fig. S11). Our detailed DFT calculations indicate that in bilayer h-BN, the net polarization is oriented as marked by the arrows in Fig. 1D.

The permanent polarization observed in separated domains, whose dimensions can be tuned by the twist angle, each exhibiting a distinct and stable potential, may be useful in applications. To that end, however, one should identify additional ways to control the local orientation beyond the twisting mechanism. Reversible switching between AB and BA configurations, accompanied by polarization inversion, can be achieved through relative lateral translation by one atomic spacing (1.44 Å), as illustrated in Fig. 1A. Similar sliding between different stacking configurations was recently demonstrated in multilayered graphene. It was shown that both mechanical (33) and electric perturbations (34) can push domain walls, practically modifying the local stacking. In the present h-BN interface, however, the polar switching calls for a preferred up or down orientation that can be predetermined by the user. To obtain such a spatially resolved control, we scanned a biased tip above an individual domain to induce a local electric field normal to the interface. The polarization images before and after the biased scans are presented in Fig. 3. We observe a redistribution of domain walls to orient the local polarization with the electric field under the biased tip. For example, after scanning a negatively biased tip above the region marked by the blue square, we observed a large white domain due to the motion of the walls marked by dashed red (green) line before (after) the scan. A successive scan by a positively biased tip resulted in practically complete domain polarization flipping. Hence, by applying negative or positive bias to the tip, it is possible to determine the polarization orientation of the underlying domain. Similar switching behavior was attained above different domains within the same interface and for several measured structures (fig. S12). We note that domain-wall motion is observed for electric field values exceeding

~0.3 V/nm and when operating the biased scan in a pin-point mode (17).

Our results therefore demonstrate that the broken symmetry at the interface of parallel-stacked h-BN flakes gives rise to an out-of-plane two-dimensional polarization confined within a few interfacial layers that can be locally detected and controlled. Although the h-BN system, with only two different light atoms, offers a convenient experimental and computational testbed and allows for intuitive interpretations, first-principle analysis (11) predicts similar phenomena to occur in other, more complex biatomic vdW crystals, such as various TMDs (35, 36). Notably, the origin of the polarization and the sliding inversion mechanism presented herein are fundamentally different from the minute deformations of tightly bonded atoms in common non-centrosymmetric 3D bulk crystals. The "slidetronics" switching involves lateral motion by a full lattice spacing in a weakly coupled interface under ambient conditions. The associated sliding order parameter reveals vortices patterns around the AA points (Figs. 1C and 2B) with topological aspects resembling the hexagonal manganite system (37). Unlike the 3D manganites, the present 2D structure allow relaxation processes through the delamination and formation of bubbles, or the annealing of walls at the open edges (fig. S12). In the present study, however, we focus on the physics away for the domain wall and toward the domain center, where no twist, moiré pattern, or strain are considered. The sensitivity of the system to the delicate interplay between vdW attraction, Pauli repulsion, Coulomb interactions, and charge redistribution implies that external stimuli such as pressure, temperature, and/or electric fields may be used to control the polarization, thus offering many opportunities for future research.

We note that a paper in the same issue (38) reports similar findings. During the review process of this manuscript, similar experimental findings were also reported in (39), and a ferroelectric response in aligned bilayer graphene was reported in (40).

REFERENCES AND NOTES

1. M. Lines, A. Glass, *Principles and Applications of Ferroelectrics and Related Materials* (Oxford Univ. Press, 2001).
2. N. Setter et al., *J. Appl. Phys.* **100**, 051606 (2006).
3. J. F. Scott, *Ferroelectrics* **314**, 207–222 (2005).
4. J. F. Scott, *ISRN Mater. Sci.* **2013**, 1–24 (2013).
5. M. Dawber, K. M. Rabe, J. F. Scott, *Rev. Mod. Phys.* **77**, 1083–1130 (2005).
6. J. Müller, P. Polakowski, S. Mueller, T. Mikolajick, *ECS J. Solid State Sci. Technol.* **4**, N30 (2015).
7. Y. Cao et al., *Nano Lett.* **15**, 4914–4921 (2015).
8. D. D. Fong et al., *Science* **304**, 1650–1653 (2004).
9. S. S. Cheema et al., *Nature* **580**, 478–482 (2020).
10. F. Liu et al., *Nat. Commun.* **7**, 1–6 (2016).
11. L. Li, M. Wu, *ACS Nano* **11**, 6382–6388 (2017).
12. G. Constantinescu, A. Kuc, T. Heine, *Phys. Rev. Lett.* **111**, 036104 (2013).
13. R. S. Pease, *Nature* **165**, 722–723 (1950).
14. S. Zhou, J. Han, S. Dai, J. Sun, D. J. Srolovitz, *Phys. Rev. B Condens. Matter Mater. Phys.* **92**, 155438 (2015).
15. S. M. Gilbert et al., *2D Mater.* **6**, 021006 (2019).
16. J. H. Warner, M. H. Rümmeli, A. Bachmatiuk, B. Büchner, *ACS Nano* **4**, 1299–1304 (2010).
17. Materials, methods, and additional information are available as supplementary materials.
18. T. Maaravi, I. Leven, I. Azuri, L. Kronik, O. Hod, *J. Phys. Chem. C* **121**, 22826–22835 (2017).
19. J. S. Alden et al., *Proc. Natl. Acad. Sci. U.S.A.* **110**, 11256–11260 (2013).
20. M. R. Rosenberger et al., *ACS Nano* **14**, 4550–4558 (2020).
21. A. Weston et al., *Nat. Nanotechnol.* **15**, 592–597 (2020).
22. H. Yoo et al., *Nat. Mater.* **18**, 448–453 (2019).
23. T. A. Green, J. Weigle, *Helv. Phys. Acta* **21**, 217 (1948).
24. C. R. Woods et al., *Nat. Phys.* **10**, 451–456 (2014).
25. O. Hod, E. Meyer, Q. Zheng, M. Urbakh, *Nature* **563**, 485–492 (2018).
26. L. J. McGilly et al., *Nat. Nanotechnol.* **15**, 580–584 (2020).
27. P. Ares et al., *Adv. Mater.* **32**, e1905504 (2020).
28. W. M. Lomer, K. W. Morton, *Proc. Phys. Soc. A* **66**, 772–773 (1953).
29. Y.-N. Xu, W. Y. Ching, *Phys. Rev. B Condens. Matter* **44**, 7787–7798 (1991).
30. J.-C. Charlier, X. Gonze, J.-P. Michenaud, *Europhys. Lett.* **28**, 403–408 (1994).
31. N. Marom et al., *Phys. Rev. Lett.* **105**, 046801 (2010).
32. O. Hod, *J. Chem. Theory Comput.* **8**, 1360–1369 (2012).
33. L. Jiang et al., *Nat. Nanotechnol.* **13**, 204–208 (2018).
34. M. Yankowitz et al., *Nat. Mater.* **13**, 786–789 (2014).
35. J. Sung et al., *Nat. Nanotechnol.* **15**, 750–754 (2020).
36. T. I. Andersen et al., https://arxiv.org/abs/1912.06955 arXiv [cond-mat.mes-hall] (2019).
37. S. Artyukhin, K. T. Delaney, N. A. Spaldin, M. Mostovoy, *Nat. Mater.* **13**, 42–49 (2014).
38. K. Yasuda, X. Wang, K. Watanabe, T. Taniguchi, P. Jarillo-Herrero, *Science* **372**, 1458–1462 (2021).
39. C. R. Woods et al., *Nat. Commun.* **12**, 347 (2021).
40. Z. Zheng et al., *Nature* **588**, 71–76 (2020).
41. M. Vizner Stern et al., Replication Data for: Interfacial Ferroelectricity by van der Waals Sliding, Zenodo (2021).

MATERIALS SCIENCE

Pressure-driven fusion of amorphous particles into integrated monoliths

Zhao Mu[1]†, Kangren Kong[1]†, Kai Jiang[2], Hongliang Dong[3,4], Xurong Xu[5], Zhaoming Liu[1]*, Ruikang Tang[1,6]*

Biological organisms can use amorphous precursors to produce inorganic skeletons with continuous structures through complete particle fusion. Synthesizing monoliths is much more difficult because sintering techniques can destroy continuity and limit mechanical strength. We manufactured inorganic monoliths of amorphous calcium carbonate by the fusion of particles while regulating structurally bound water and external pressure. Our monoliths are transparent, owing to their structural continuity, with a mechanical strength approaching that of single-crystal calcite. Dynamic water channels within the amorphous bulk are synergistically controlled by water content and applied pressure and promote mass transportation for particle fusion. Our strategy provides an alternative to traditional sintering methods that should be attractive for constructing monoliths of temperature-sensitive biominerals and biomaterials.

norganic materials, especially minerals and ceramics, play important roles in modern society (1–4), but the manufacture of their monoliths is a great challenge (5). In practice, many inorganic materials are produced in powder forms and then consolidated by pressing and sintering (6–10). However, mass transportation among particles is often insufficient through sinter treatment, complete particle-particle fusion cannot be achieved within bulk materials, and particle boundaries remain. Owing to internal discontinuities, the properties of sintered inorganic bulks are not ideal (11, 12), especially with respect to mechanical strength (13). In nature, biological organisms [e.g., sea urchins (14) and coccoliths (15)] can produce inorganic

skeletons that have continuous structures with flexible morphology (16). Accordingly, these inorganic skeletons are superior to artificial skeletons because of their structural integration. Increasing evidence has demonstrated that these biological organisms use amorphous particles as precursors to produce skeletons via particle-particle fusion (17–19). Inspired by this biological phenomenon, we suggest that the manufacture of monolithic inorganic materials can be achieved by fusion of their amorphous precursors under pressure (P). However, many previous attempts using amorphous mineral phases such as amorphous calcium carbonate (ACC) and amorphous calcium phosphate have revealed that external P frequently induces crystallization of amorphous particles rather than particle-particle fusion (20–23). These amorphous mineral phases contain abundant water, which induces dissolution and recrystallization of the intermediate amorphous phases under P.

By controlling the water content (n) within amorphous precursors and the external P, we can achieve biomineralization-like particle fusion, which follows the discovery of pressure-driven mass transportation through dynamic water channels in amorphous phases. By using ACC as a typical amorphous precursor, we experimentally demonstrated this control of

complete fusion of ACC particles to result in a transparent calcium carbonate monolith. The monolith exhibits optimal mechanical strength with a hardness (H) of 2.739 GPa and a Young's modulus (E) of 49.672 GPa; these values are superior to those of common cement materials and almost reach those of single-crystal calcite (24).

The chemical formula of ACC is $CaCO_3 \cdot nH_2O$, with n frequently being ≥ 1 (20, 25, 26). In our experiments, we synthesized ACC particles with an average size of 450 ± 100 nm (fig. S1) while controlling n between 0 and 1.6 (fig. S2A). We confirmed that ACC particles with $n \geq 1$ cannot coalesce under external P, because phase transformation–based crystallization happens instead (fig. S3), but keeping $n < 1$ enables us to avoid this issue. We used ACC with $n = 0.7$ (fig. S2B) for our typical samples in a subsequent compaction treatment under P. In all of our experiments, the treatment time was 5 min at room temperature. At $P \leq 0.5$ GPa, we did not observe fusion because we could observe the particles and their boundaries under scanning electron microscopy (SEM) (Fig. 1A). When P increased to 0.8 GPa, some ACC particles began to coalesce, but we could still identify their original spherical morphology. The particles further coalesced at a large scale, and the spherical shape disappeared as P increased to 1.5 GPa. At P = 2.0 GPa, all ACC particles coalesced completely to form a continuous and uniform bulk. To clearly express the coalescence behavior of ACC particles, we additionally processed their SEM images with binarization (Fig. 1A; blue areas represent non-coalesced particles). The structural discontinuities decreased as P increased from 0.2 to 2.0 GPa, and finally, all individual particles were fused into an integral whole. We used Au nanoparticles to confirm complete coalescence at the nanoscale (Fig. 1B and fig. S4). In this examination, the ACC particle surfaces were labeled with Au nanoparticles (Fig. 1C). After fusion treatment, these Au particles "moved" from the surface into the interior and were completely surrounded by the uniform ACC phase (Fig. 1D and fig. S5). We verified the surrounding

[1]Department of Chemistry, Zhejiang University, Hangzhou 310027, China. [2]Engineering Research Center of Nanophotonics & Advanced Instrument (Ministry of Education), Department of Materials, East China Normal University, Shanghai 200241, China. [3]Center for High Pressure Science and Technology Advanced Research, Shanghai 201203, China. [4]State Key Laboratory of High Performance Ceramics and Superfine Microstructure, Shanghai Institute of Ceramics, Chinese Academy of Sciences, Shanghai 200050, China. [5]Qiushi Academy for Advanced Studies, Zhejiang University, Hangzhou 310027, China. [6]State Key Laboratory of Silicon Materials, Zhejiang University, Hangzhou 310027, China.
*Corresponding author. Email: oldliu@zju.edu.cn (Z.L.); rtang@zju.edu.cn (R.T.) †These authors contributed equally to this work.

ACC by element mapping (fig. S6) and electron irradiation–induced crystallization (fig. S7). Because we did not observe a gap or pore structure in the Au-labeled monolith, we again confirmed complete particle-particle coalescence.

We performed in situ high-pressure synchrotron radiation x-ray diffraction (XRD) analysis (Fig. 1E) and Fourier transform infrared spectroscopy (fig. S8) to show that the resulting monolith was still ACC and did not undergo crystallization during treatment. This observation is not consistent with previous studies (20, 21). We argue that it follows from our observations that the phase stability of ACC improved as n decreased. During the fusion process, the apparent density (which includes the void volumes) increased, along with the bulk transparency (Fig. 1F). The apparent density of the compacted ACC bulk was ~1.30 g/cm^3 at 0.2 GPa, and it increased to 2.45 g/cm^3 at 2.0 GPa. Because the documented densities of ACC are 1.62 to 2.18 g/cm^3 (27–29), the value of 2.45 g/cm^3 follows from the formation of a more densified ACC bulk. A comparison of samples before and after the pressure treatment showed that the water content in ACC remained almost unchanged (fig. S2C), which excludes the possibility of water removal from ACC (30). Thus, we attribute the apparent density increase to the elimination of defects and pores by particle-particle fusion. In addition, under the same pressure treatment, the ACC with a lower water content (n = 0.3) has a higher apparent density, 2.60 g/cm^3. The theoretical density of vaterite, a single crystal of calcium carbonate, is 2.71 g/cm^3, and the theoretical density of ACC is in the range of 2.18 to 2.65 g/cm^3 (29, 31) at different water contents. This close-packed characteristic follows the continuous internal structure of the resulting ACC monolith.

Microstructural defects in materials commonly cause light scattering, which leads to opacity. The constructed bulk compacted at P = 0.2 GPa is white and nontransparent (inset of Fig. 2A), which is due to internal discontinuities (Fig. 1A). When P was increased to 2.0 GPa, the degree of transparency increased substantially, and the constructed bulk became almost transparent (Fig. 1F and inset of Fig. 2A), which indicates that light scattering decreased as a result of improved internal continuity. We used nanoindentation tests to examine the mechanical strengths of the bulks constructed under different P values (Fig. 2A). We calculated E and H using the load-displacement curve. The bulk material constructed at 0.2 GPa performed poorly, with E = 0.100 ± 0.020 GPa and H = 0.003 ± 0.001 GPa. At 2.0 GPa, E and H of the resulting material increased to 32.783 ± 3.537 GPa and 1.962 ± 0.327 GPa, respectively, which are substantially higher than those of the calcite bulk (fig. S9) constructed with calcite powder at the same P, which had E and H of 0.100 to 2.700 GPa and 0.010 to 0.040 GPa, respectively (Fig. 2A). In comparison with the original ACC particles, the monolith exhibited improved stability with increased kinetic persistence against crystallization (fig. S10), which is attributed to its reduced specific surface area by particle fusion.

To provide a comprehensive understanding of the P-induced fusion of ACC with different n values, we generated a phase diagram of ACC from our collected data. The phase diagram shows three major zones of crystallization, fusion, and no fusion (Fig. 2C). The phase diagram reveals that both P and n have important effects on the behaviors of the ACC. At low P, the driving force is insufficient, so all primary ACCs with different n values are neither fused nor crystallized. As P increases, the driving force becomes sufficient, and an effect of n exists. ACC takes different paths, depending on the value of n. ACC with abundant water (n > 0.7 to 1.2, which is related to P) is more likely to crystallize (path 1) to form calcite or vaterite (Fig. 2G). With decreasing n, P-induced coalescence (path 2) becomes increasingly favorable, which leads to particle-particle coalescence rather than crystallization (n < 0.7 to 1.2, which is related
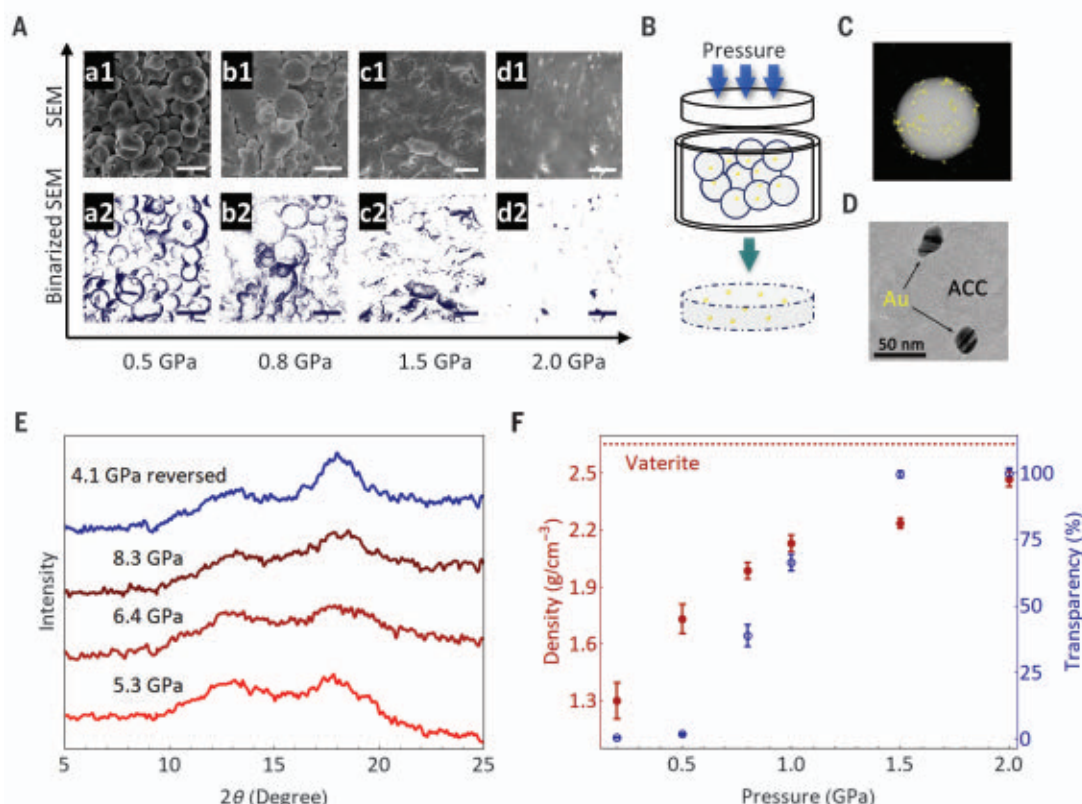

Fig. 1. Coalescence of ACC particles under pressure.
(**A**) SEM images of ACC particles compacted under 0.5, 0.8, 1.5, and 2 GPa, respectively (a1 to d1). Images a2 to d2 correspond to binary processing of images a1 to d1. Blue, noncoalescent area; white, coalescent area. Scale bars, 1 μm. (**B**) Scheme of coalescence of ACC particles labeled with Au nanoparticles under pressure. (**C**) Characterization of ACC labeled with Au nanoparticles. (**D**) Transmission electron microscopy cross-sectional view of compacted ACC labeled with Au. (**E**) In situ high-pressure synchrotron radiation XRD analysis of ACC particles. 2θ, diffraction angle. (**F**) Changes in apparent density and transparency of ACC bulk as pressure increases. Error bars indicate the SD of the average for five measurements.

to P) (Fig. 2, E and F). However, sufficient water is important to maintain the shape stability of ACC. If n is less than ~0.1, ACC neither coalesces nor crystallizes, but cracks under P (path 3 in Fig. 2D) and forms numerous tiny particles. The diagram shows that the synergistic effect of n and P is key to ensuring particle-particle fusion. Generally, n (the water in ACC) includes mobile and structural water (32). The thermogravimetric analysis curves indicate the domination of structural water when n < 0.6 (fig. S2D). Figure 2C shows that increasing the amount of either structural water (n = 0.1 to 0.6) or mobile water (n = 0.6 to 1.1) promotes the fusion process, because both are involved in the pressure-driven treatment. Furthermore, our Raman study implied a degenerate state of water under high pressure for fusion (fig. S11) (30). According to the phase diagram, both n and P can be further optimized, and the "best" ACC monolith was prepared under conditions of

$n \approx 0.3$ and $P = 3.0$ GPa (the maximum P in our experiment), with E and H values of 49.673 ± 3.490 GPa and 2.739 ± 0.249 GPa, respectively (fig. S12). This fused ACC monolith is superior to many cements and calcium carbonate-based composites and is even similar to single-crystal calcite (Fig. 2B) (24).

We used in situ Raman spectroscopy of ACC at a high P and molecular simulation to understand P-driven coalescence and crystallization. The peak at ~712 cm^{-1} in the Raman spectrum (fig. S13) corresponds to the in-plane bending (ν_4) of carbonate molecules. The peak had a blueshift and redshift during compression and pressure release, respectively (Fig. 3A). The same changes also appeared in two other peaks at 490 and 1082 cm^{-1}. In ACC, H bonds are considered the primary interaction between water and carbonate and result in a redshift in the C=O vibration frequency (29). The apparent blueshift with increasing P implies the weakening of these H bonds, which reflects

dissociation between water and carbonates. This effect contributed to the loss of absorbed water (33) after the separation of H_2O and CO_3^{2-} in the ACC phase. With separation, water and ions aggregate to form water clusters, which provide migration channels to enhance mass transportation within ACC (34).

We employed a bulk phase of ACC to understand its internal ion transport with different n and P values. The coordination environment, presented by pair distribution functions (PDFs), at 2 GPa is clearly different from those at lower P values (Fig. 3, B and C), indicating that diffusivity varies with pressure change. The structural changes reflected in the PDFs (Fig. 3, B and C, and fig. S14) verify the dissociation between water and carbonate and the aggregation of desorbed water from the carbonate, showing consistency between the Raman and simulation results. We can use percolation theory (35) for clusters to understand the relation between the structural change and diffusivity in a semiquantitative way. With aggregation, the changes in the scale and morphology of water clusters are substantial and act on mass transfer. We used 3.8 Å as the cutoff distance to analyze water clusters (34). The calculation results intuitively reflect the aggregation of water, forming larger clusters (Fig. 3, D and E) under P. However, clusters exist under dynamic equilibrium (Fig. 3F and fig. S15) instead of as a localized framework so that clusters will disperse throughout the whole system with sufficient time. Moreover, the ability of mass transfer D for units in water clusters is related to the scale length λ_m (35), which is affected by the scale and the fractal dimension

$$D \propto D_0 \frac{\lambda_m^{2-H+1/\nu}}{\lambda_0^{1/\nu+H-2}} \quad (1)$$

where D_0 is the diffusivity of a single, unclustered water molecule; λ_0 is the unit size of water; and $H \in (-1/\nu, 0)$ and $\nu \approx 0.9$ are coefficients related to the fractal dimension of water clusters (35). The diffusivity of all water $\bar{D}$ is the arithmetic mean of D

$$\bar{D} = \frac{\sum i N_i D_0 \lambda_i^{2-H_i+1/\nu_i} \lambda_0^{2-1/\nu_i-H_i}}{\sum i N_i} \quad (2)$$

where i is the number of molecules contained in a water cluster and N_i is the number of water clusters that contain i molecules. For all water clusters, λ_0 can be regarded as the same. D_0 as an intensive quantity can be affected by the coordination environment. Considering the maintenance of the ACC frame during pressing (fig. S13), D_0 is regarded as a constant, and the dominant factors that determine $\bar{D}$ are scale and morphology. We employed two methods (34) to measure the fractal dimension of clusters, both of which show the invariance of the fractal dimension for large clusters (figs. S16 and S17). Because large clusters contribute much more than small clusters to $\bar{D}$ (fig. S18),
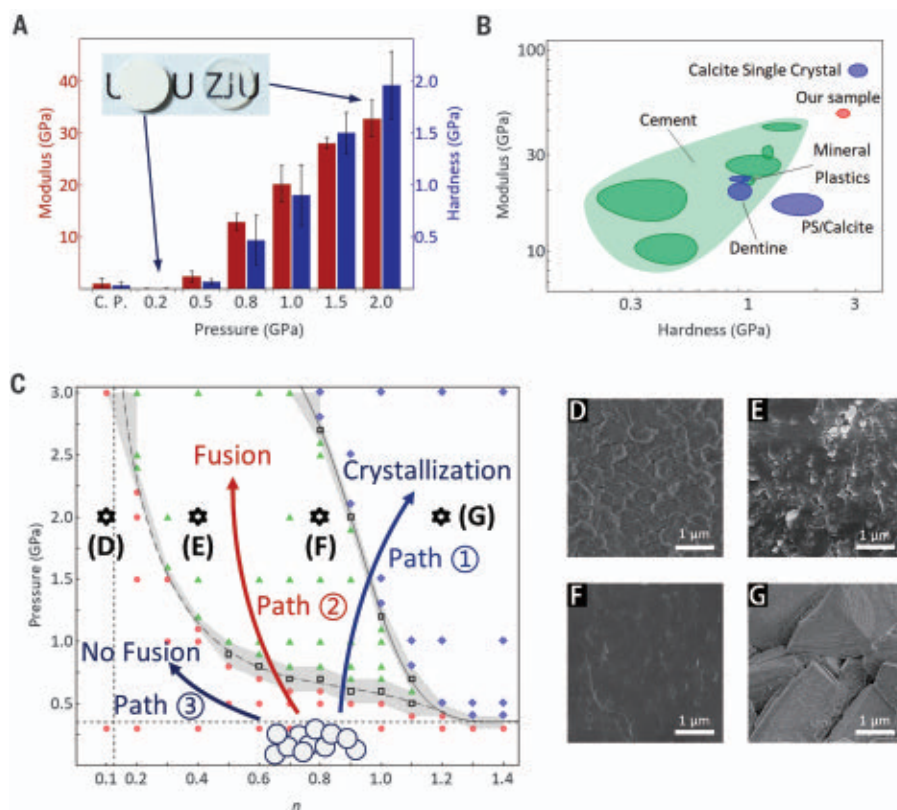

Fig. 2. Behaviors of particle fusion to bulks. (**A**) Nanoindentation tests of ACC bulks, which are obtained at different pressures from 0.2 to 2.0 GPa. C. P. represents pure calcite bulk, which is obtained from calcite particles at 2 GPa. Error bars indicate the SD of the average for more than five measurements. (Inset) Optical photos of the ACC bulks obtained under 0.2 and 2.0 GPa, respectively. The diameter of each ACC bulk is 4 mm. (**B**) Comparison of mechanical properties between prepared ACC bulks through coalescence of ACC particles (our sample) with other materials. PS/Calcite, calcite-filled polystyrene. (**C**) Phase diagram of ACC particles with different n under different pressures. Red circles, nonfusion conditions; green triangles, fusion conditions; blue rhombuses, crystallization conditions; black squares, conditions in which more than one phenomenon is detected. Gray regions represent the boundaries of fusion-nonfusion, fusion-crystallization, and nonfusion-crystallization areas. (**D** to **G**) SEM images of representative coalescence behaviors, corresponding to points (D) to (G) in panel (C).
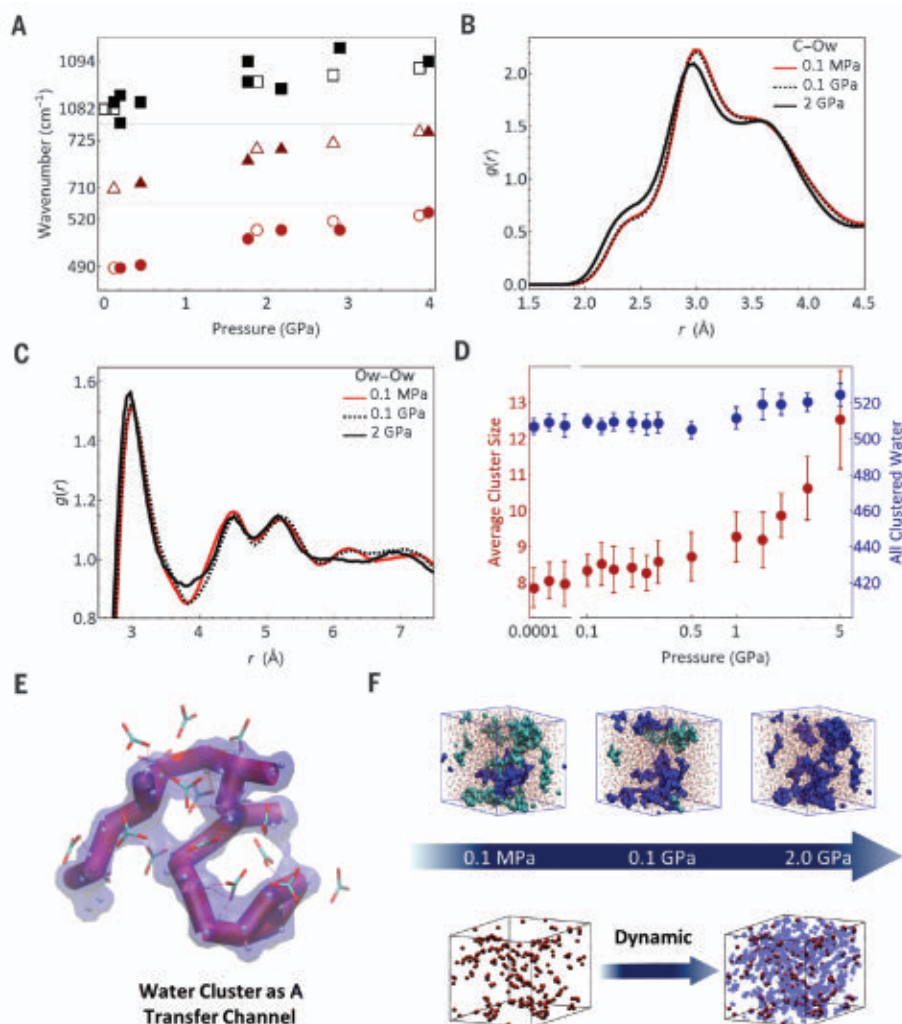
Fig. 3. Evolution of water clusters during pressure treatment. (**A**) Peak displacement of the in situ Raman spectrum in a compression-decompression process. Solid points and hollow points respectively represent compression and decompression processes. (**B** and **C**) PDFs [denoted by $g(r)$] of the C-Ow (carbon atom of carbonate–oxygen atom of water) and Ow-Ow pairs calculated according to simulations under different pressures. r, radius. (**D**) Changes in average water cluster size and total number of water clusters. The average water cluster size represents the average number of water molecules that are contained in clusters. "All clustered water" represents the total number of water molecules that form clusters. Error bars indicate the SD of the average for successive 5-ns data windows. (**E**) Illustration of a typical water cluster. (**F**) (Top) Illustration of the aggregation between clusters with increasing pressure. A main water cluster (blue) will gradually aggregate with other small clusters (cyan) to form a lager cluster. (Bottom) Illustration of the dynamic equilibration between clusters under pressure. Water in a cluster (red) will form new clusters (light blue) while disconnecting and reaggregating.

approximately $P = 1.0$ to 2.0 GPa is close to the boundary of the fusion and nonfusion areas in the phase diagram. Although the discussion above concerns water, it should be noted that the diffusivity of ions is correlated with water (Fig. 4C), owing to the interactions between ions and water. Strong water diffusion must be accompanied by corresponding strong ion diffusion in ACC (37, 38). From this analysis, we suggest that gradual aggregation leads to a qualitative change in the capacity of mass transportation for the whole system, finally leading to complete fusion.

Our results above could also be used to understand the role of water content in the experimental phase diagram. According to Eq. 3, the cluster distribution N_i apparently controls the diffusivity. N_i is directly determined by water content. For two types of ACCs with different n values of 0.5 and 0.7, the largest water cluster sizes are approximately 18 and 77 under ambient conditions. Considering the high order of i in Eq. 3, the gap in diffusivity is particularly evident and will widen under high P. This gap suggests why fusion is difficult to observe in areas with low water content. However, after n increases to approximately 0.8 to 0.9, a type of water cluster that expands the boundary of the simulation cell appears (Fig. 4D)—it is termed the infinite cluster (39). Typically (34), the critical condition of n for infinite clusters is independent of the simulation scale. The value of i of these infinite clusters is considered infinite; then, $\bar{D}$ approaches infinity as well, which contradicts common sense. Actually, the infinite clusters exceed the domain of the definition applied for finite clusters in Eq. 1, so Eq. 3 becomes invalid. Any point of the solid can reach from the outside to within this kind of infinite cluster (40) to enable long-range mass transfer, which results in particle fusion as well as phase separation (crystallization). We note that $n = 0.8$ to 0.9 is close to the critical value of the phase diagram for different tendencies previously reported (34).

In our experiments, the fusion process of the pure ACC particles is independent of their sizes, synthesis methods, or pressing device (opened or closed) for appropriate values of n (figs. S20, S2C, and S21) (30). However, humidity and temperature are the surrounding conditions that may affect the formation of monoliths as well as their structural uniformity and continuity, because the induced crystallizations are competitive with fusion (figs. S2A and S10) (30). Moreover, high temperature (>50°C) can result in decreased n for the ACC samples (fig. S2D). Natural and many synthetic ACC samples contain additives (41, 42). Although fusion could proceed in the presence of additives such as Mg^{2+}, citrate, and polyacrylic acid, higher pressures were always required (fig. S22). The interactions between additives and water or ions (43–48) decrease their initial diffusion
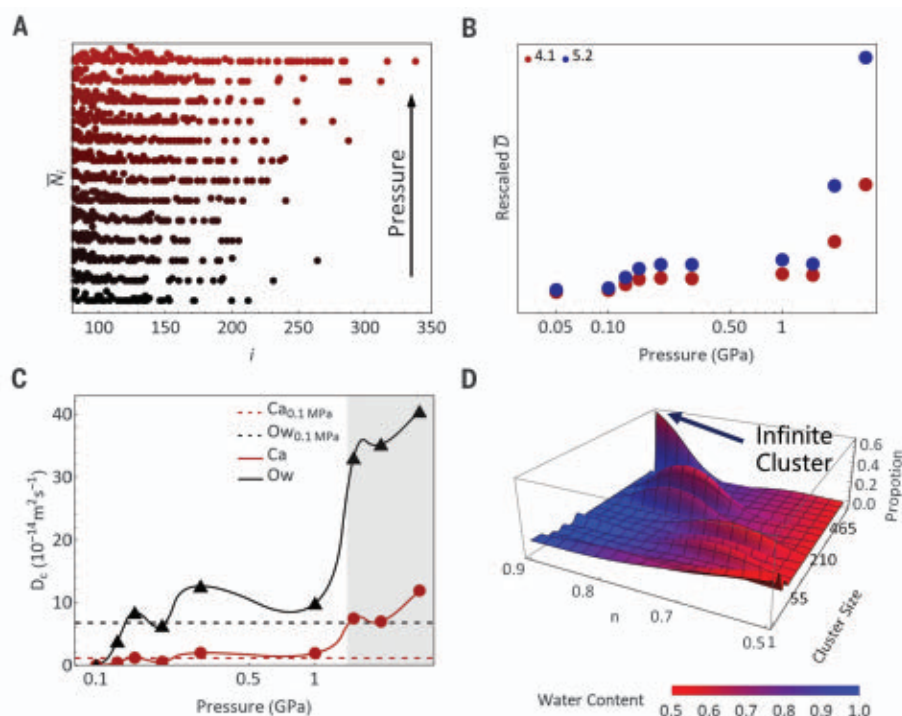
the fractal dimension is approximately considered to be the same for each cluster. Thus, λ_i can be substituted by i, and $\bar{D}$ can be written as follows

$$\bar{D} \propto \frac{\sum i^d N_i}{\sum i N_i}, d = 3 - H_i + 1/\nu \quad (3)$$

where d is a coefficient with a value of 4.1 to 5.2 (considered constant during pressing). Because clusters exist in dynamic equilibrium, N_i is replaced by the time-averaged $\bar{N}_i$ in

statistical calculations. According to Eq. 3, molecules in a cluster that contains more molecules or has simpler morphology have stronger diffusivity. Specifically, these water molecules are less restrained by carbonates (36). As expected, N_i becomes larger along with P (Fig. 4A), and $\bar{D}$ increases slowly until $P = 2.0$ GPa, then rises sharply with increasing P (Fig. 4B), whether near the lower bound of $d = 4.1$ or the upper bound of $d = 5.2$, consistent with another result calculated from the free path (Fig. 4C and fig. S19). This critical P at

Fig. 4. Understandings of diffusivity in ACC. (A) $\bar{N}_i$ becomes larger with increasing pressure. **(B)** Calculated diffusivity with increasing pressure, for d = 4.1 and 5.2. **(C)** Diffusion coefficient D_c of water and Ca (calcium ions). The diffusion coefficient of carbonate is approximately equal to that of Ca. The gray area is where ACC particles fuse. **(D)** Distribution of water clusters directly affected by water content. The largest cluster is an infinite cluster throughout the whole system.

coefficients (*30*). Therefore, the control of additives is an approach to regulate pressure-induced fusion. The similar pressure-driven fusion of amorphous particles could be extended to other inorganic ionic compounds such as calcium phosphate and magnesium carbonate, demonstrating its general applicability (fig. S23).

This pressure-driven fusion through dynamic water channels provides a feasible method to achieve inorganic monoliths with optimized characteristics from their amorphous particles. Furthermore, the resulting monoliths can be used as conformable precursors for moldable preparation of crystalline materials by a shape- and mechanical strength–preserving solid-state crystallization (fig. S24) (*49*). Our preparation of a calcite bulk from the ACC particles is analogous to biomineralization of complete skeletons by using numerous amorphous particles.

Understanding the role of water channels promotes the mass transportation of solid inorganic particles, favoring the fusion of particle precursors for monolithic material construction. This approach is distinct from the classical sintering approach because high temperature is necessary in sintering to improve mass transportation, which is incompatible with thermally sensitive materials. Our findings focus on the internal structure

of solid inorganic particles to improve mass transportation, which is extremely suitable for the construction of most thermally sensitive materials, such as biominerals and biomaterials. This understanding establishes an alternative strategy for the construction of continuously structured materials, enabling large-scale and efficient manufacturing of inorganic monoliths.

REFERENCES AND NOTES

1. F. L. Riley, *J. Am. Ceram. Soc.* **83**, 245–265 (2000).
2. M. A. Meyers, P.-Y. Chen, A. Y.-M. Lin, Y. Seki, *Prog. Mater. Sci.* **53**, 1–206 (2008).
3. H. H. K. Xu et al., *Bone Res.* **5**, 17056 (2017).
4. C. Hu, Z. Li, *Constr. Build. Mater.* **90**, 80–90 (2015).
5. N. T. K. Thanh, N. Maclean, S. Mahiddine, *Chem. Rev.* **114**, 7610–7630 (2014).
6. J. Guo et al., *Angew. Chem. Int. Ed.* **55**, 11457–11461 (2016).
7. C. Wang et al., *Science* **368**, 521–526 (2020).
8. H. Zhang, B.-N. Kim, K. Morita, H. Y. Keijiro Hiraga, Y. Sakka, *Sci. Technol. Adv. Mater.* **12**, 055003 (2011).
9. J. W. Evans, L. De Jonghe, Eds., in *The Production and Processing of Inorganic Materials* (Springer, 2016), pp. 383–439.
10. Z. Mu, R. Tang, Z. Liu, *Nanomaterials* **11**, 241 (2021).
11. J. Nowotny, T. Bak, T. Burg, M. K. Nowotny, L. R. Sheppard, *J. Phys. Chem. C* **111**, 9769–9778 (2007).
12. J. Salzman, C. Uzan-Saguy, B. Meyler, R. Kalish, *Phys. Stat. Sol. A* **176**, 683–687 (1999).
13. N. Ding, C.-M. L. Wu, H. Li, *Phys. Chem. Chem. Phys.* **16**, 23716–23722 (2014).
14. M. Albéric et al., *Cryst. Growth Des.* **18**, 2189–2201 (2018).
15. J. R. Young et al., *Rev. Mineral. Geochem.* **54**, 189–215 (2003).
16. L. Addadi, S. Weiner, *Phys. Scr.* **89**, 098003 (2014).
17. Y. Politi, T. Arad, E. Klein, S. Weiner, L. Addadi, *Science* **306**, 1161–1164 (2004).
18. J. Mahamid et al., *Proc. Natl. Acad. Sci. U.S.A.* **107**, 6316–6321 (2010).
19. L. Yang, C. E. Killian, M. Kunz, N. Tamura, P. U. P. A. Gilbert, *Nanoscale* **3**, 603–609 (2011).
20. H. Du et al., *Chem. Mater.* **32**, 4282–4291 (2020).
21. T. Yoshino, K. Maruyama, H. Kagi, M. Nara, J. C. Kim, *Cryst. Growth Des.* **12**, 3357–3361 (2012).
22. T. Mass et al., *Proc. Natl. Acad. Sci. U.S.A.* **114**, E7670–E7678 (2017).
23. C. Y. Sun et al., *Proc. Natl. Acad. Sci. U.S.A.* **117**, 30159–30170 (2020).
24. C. Merkel et al., *J. Struct. Biol.* **168**, 396–408 (2009).
25. H. Du et al., *J. Am. Chem. Soc.* **140**, 14289–14299 (2018).
26. A. V. Radha, T. Z. Forbes, C. E. Killian, P. U. P. A. Gilbert, A. Navrotsky, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 16438–16443 (2010).
27. J. Bolze et al., *Langmuir* **18**, 8364–8369 (2002).
28. J. Bolze, D. Pontoni, M. Ballauff, T. Narayanan, H. Cölfen, *J. Colloid Interface Sci.* **277**, 84–94 (2004).
29. A. Fernandez-Martinez, B. Kalkan, S. M. Clark, G. A. Waychunas, *Angew. Chem. Int. Ed.* **52**, 8354–8357 (2013).
30. Materials and methods are available as supplementary materials.
31. M. Saharay, A. O. Yazaydin, R. J. Kirkpatrick, *J. Phys. Chem. B* **117**, 3328–3336 (2013).
32. C. Günther, A. Becker, G. Wolf, M. Epple, *Z. Anorg. Allg. Chem.* **631**, 2830–2835 (2005).
33. S. A. Parry, A. R. Pawley, R. L. Jones, S. M. Clark, *Am. Mineral.* **92**, 525–531 (2007).
34. Y. G. Bushuev, A. R. Finney, P. M. Rodger, *Cryst. Growth Des.* **15**, 5269–5279 (2015).
35. M. B. Isichenko, *Rev. Mod. Phys.* **64**, 961–1043 (1992).
36. R. J. Reeder et al., *Cryst. Growth Des.* **13**, 1905–1914 (2013).
37. A. C. S. Jensen et al., *J. Phys. Chem. C* **122**, 3591–3598 (2018).
38. A. Koishi et al., *J. Phys. Chem. C* **122**, 16983–16991 (2018).
39. A. Geiger, H. E. Stanley, *Phys. Rev. Lett.* **49**, 1895–1898 (1982).
40. M. Rosso, J. F. Gouyet, B. Sapoval, *Phys. Rev. Lett.* **57**, 3195–3198 (1986).
41. F. C. Meldrum, H. Cölfen, *Chem. Rev.* **108**, 4332–4432 (2008).
42. S. Weiner, Y. Levi-Kalisman, S. Raz, L. Addadi, *Connect. Tissue Res.* **44** (suppl. 1), 214–218 (2003).
43. E. Loste, R. M. Wilson, R. Seshadri, F. C. Meldrum, *J. Cryst. Growth* **254**, 206–218 (2003).
44. S. Raz, S. Weiner, L. Addadi, *Adv. Mater.* **12**, 38–42 (2000).
45. F. Lippmann, in *Sedimentary Carbonate Minerals* (Springer, 1973), p. 79.
46. D. Gebauer, H. Cölfen, A. Verch, M. Antonietti, *Adv. Mater.* **21**, 435–439 (2009).
47. D. J. Tobler et al., *Adv. Funct. Mater.* **25**, 3081–3090 (2015).
48. X.-R. Xu et al., *J. Cryst. Growth* **310**, 3779–3787 (2008).
49. Z. Liu et al., *Nature* **574**, 394–398 (2019).

INDIANA UNIVERSITY
SCHOOL OF MEDICINE

Indiana University School of Medicine (IUSM) is seeking applications for **Tenure-Track Faculty positions** at the **Assistant/Associate/Full Professor** levels in
- Cancer Drug Discovery
- Area of Protein Engineering or Synthetic biology

This search is associated with the Department of Biochemistry & Molecular Biology (BIOM) and the NCI designated IU Simon Comprehensive Cancer Center (IUSCCC). Applications will be considered from all areas of basic and applied cancer research with an emphasis on NMR- and/or cryoEM-based small molecule discovery approaches, development and implementation of novel therapeutics and strategies for cancer treatment. Science directed toward the identification or redesign of biological molecules to create new therapeutics, especially cancer immunotherapeutics.

Successful applicants are expected to have/develop a strong, leadership in multi-investigator and/or independent extramurally funded research program, participate in training students and fellows, and engage in research interactions with the clinical programs complementary to their scientific interests. A competitive salary, startup funds, and space will be provided.

Interested individuals should submit a statement outlining their approach to integrating diversity in their pursuit of academic excellence in their cover letter, curriculum vitae, cover letter, summary of past accomplishments and future research plans, and the names and addresses of 3-5 references in electronic format. The search committee will begin considering applications immediately and on an on-going basis until positions are filled. Candidates can submit applications at **https://indiana.peopleadmin.com/postings/10957**. Additionally, candidates can submit applications by email at **biomfac@iupui.edu**.

The Indiana University School of Medicine (https://medicine.iu.edu/) is the largest medical school in the US and is annually ranked among the top medical schools in the US by US News & World Report. More information about the sponsoring units can be found at the following websites: BIOM (http://medicine.iu.edu/departments/biochemistry-molecular-biology/) and IUSCC (http://www.cancer.iu.edu/).

As the nation's largest medical school, IUSM is committed to being an institution that not only reflects the diversity of the learners we teach and the patient populations we serve, but also pursues the values of diversity, equity and inclusion that inform academic excellence. We desire candidates who enhance our representational diversity, as well as those whose work contributes to equitable and inclusive learning and working environments for our students, staff, and faculty. IUSM strives to take an anti-racist stance, regularly evaluating and updating its policies, procedures, and practices to confer equitable opportunities for contribution and advancement for all members of our community. We invite individuals who will join us in our mission to advance gender and racial equity to transform health and wellbeing for all throughout the state of Indiana. More information about diversity and inclusion efforts at IUSM can be found at https://faculty.medicine.iu.edu/diversity.

Indianapolis is the capital and most populous city in the State of Indiana. It is growing economically thanks to a strong corporate base anchored by the life sciences. Indiana is home to one of the largest concentrations of health sciences companies in the nation. Indianapolis has a sophisticated blend of charm and culture with a wonderful balance of business and leisure. The growing residential base is supported by rich amenities and quality of life - the city possesses a variety of professional sports, arts venues and outdoor recreation areas. Residents of this dynamic city, and surrounding suburbs, enjoy leading educational systems and top-ranked universities, paired with a diverse population. Indianapolis International Airport is a top-ranked international airport, being named "Best Airport in North America" by Airports Council International for many years. For additional information on life in Indy: https://faculty.medicine.iu.edu/relocation.

THE UNIVERSITY OF ARIZONA — COLLEGE OF MEDICINE

ASSOCIATE DEPARTMENT HEAD & ASSOCIATE OR FULL PROFESSOR
Department of Immunobiology
College of Medicine, University of Arizona
Health Sciences

The University of Arizona College of Medicine-Tucson (COM-T) is seeking an outstanding scientist who works in the areas of immunology and/or microbiology to serve as Associate Department Head (ADH) of the Department of Immunobiology. The Department hosts a group of well-funded investigators with research interests in basic immunology, microbiology, and host-pathogen interactions. The ADH will serve a key role in the administration of the Department including functioning as a liaison between Department faculty, students, staff, and COM-T administration. A successful ADH candidate will engage in leadership and innovation in support of the research, teaching, and service missions of the Department, including, but not limited to, faculty promotion, tenure and evaluation reviews; trainee progress assessment; curricular and educational issues; and hiring. In addition, the ADH must be prepared and willing to step in for the Department Head in any facet of DH function.

Diversity, equity, and inclusion are the core values of the Department of Immunobiology, COM-T and UArizona, and candidates of all backgrounds are welcome and invited to apply. The successful candidate will join the UA Department of Immunobiology (primary) at the rank of Associate or Full Professor. If appropriate, the successful candidate will also join one or more of the UA Centers or institutes. Tenure can be awarded if applicable.

Competitive candidates are expected to have a track record of national or international prominence, scholarly productivity and demonstrated ability to secure independent, peer-reviewed funding such as an NIH R01. Competitive candidates will also have a history of successfully working with a diverse group of students, trainees, and colleagues.

Modern, functional shared resources in clinical trials, biostatistics, bioinformatics, mouse models, and tissue acquisition and analysis, flow cytometry, mass spectrometry, structural biology and functional genomics are available to all University scientists.

Applicants must have a PhD and/or MD degree at a current rank of Associate (minimum 3 years in the rank) or Full Professor. Salary and start-up funds are attractive and commensurate with qualifications and experience.

The University of Arizona has been recognized on Forbes list of America's Best Employers in the United States and has been awarded the Work-Life Seal of Distinction by World@Work! For more information about working at the University of Arizona, please visit https://talent.arizona.edu.

For full details and qualifications, and to complete an on-line application, see req3686 at https://talent.arizona.edu/. Outstanding UA Benefits! The University of Arizona is an Equal Opportunity Employer Minorities/Women/Vets/Disabled.

myIDP:
A career plan customized for you, by you.

For your career in science, there's only one

Science

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.

myIDP

Visit the website and start planning today!
myIDP.sciencecareers.org

Science Careers In partnership with:

AAAS

FASEB
Federation of American Societies
for Experimental Biology

UMASS University of Massachusetts Medical School

UCSF
University of California
San Francisco

UIC

BURROUGHS WELLCOME FUND

Hear all voices

n May, as our institution began to open back up and whispers about in-person lab meetings started to float around, I panicked. In the months preceding the pandemic, I had cochlear implants surgically placed to address steadily worsening hearing loss. As I began to retrain my brain to hear—a difficult and tiring process that required me to match the new sounds I was hearing with my previous knowledge of the auditory world—we went into lockdown. Amid the isolation and uncertainty, I discovered a bright spot: Online meetings worked well for me. With everyone speaking into computer microphones, I could hear and focus. What would happen as we returned to "normal"?

My hearing loss began when I was in graduate school; no one has been able to determine the cause. At first, I naïvely thought I could overcome it with hard work and dedication. I began to record every one-on-one meeting because even with the most intense focus I just couldn't catch everything, and taking notes is not an option when you rely on lip reading. In seminars, I sat in the front row, exhausting myself as I strained to both hear the speaker and process the science. So, in the fourth year of my postdoc, I decided cochlear implants were the right next step. There was no guarantee of success, and even a best case scenario would not mean regaining "normal" hearing—hearing aids and cochlear implants can't re-create all of our ears' amazing tricks—but I had nothing to lose.

The week after I began to hear with both implants, I attended a conference. I scrutinized every listening environment to plan my equipment and where I needed to be located during each event. I attached my minimicrophone to the podium and had my poster moved to a quiet corner. It worked OK—but it was still exhausting, and I felt excluded from so much. I opted out of the valuable (and noisy) networking sessions to take the first shuttle back to my hotel, where I fell asleep instantly.

When meetings went virtual during the pandemic, my experience was much better. The sound quality on Zoom was crisp and clean, without the overwhelming ambient noises of a physical office space or meeting room. I could sit back and let the sound come to me instead of straining to hear. It was easier to focus, participate, and think deeply about the science. I began to leave seminars feeling invigorated rather than needing a nap.

With reopening on the horizon, I feared I would lose this progress. I can hear better than I could before the pandemic; I've gotten used to my implants and can now techni-

"I began to leave seminars feeling invigorated rather than needing a nap."

cally pass a hearing test. But I will always need quiet environments and microphones because I am, after all, hearing electronically.

I didn't want to be forced back into hustling to comprehend. So, with my stomach clenched with anxiety, I set up a meeting with my adviser to make a hearing plan for our lab meetings. I did not want to distract or burden my 30 lab mates, so the simplest solution seemed to be to ask the presenter to use a microphone and repeat audience questions before answering them.

Initially, it went OK. But as the presentation gave way to lengthy technical questions and in-depth discussion, the system broke down. I found myself leaning forward, my neck tensed as I turned back and forth from speaker to audience, straining to hear.

I was tempted to accept that I wasn't going to catch everything, but I reminded myself of the relaxed experience I had on Zoom. I took a deep breath, turned to my adviser, and said, "This isn't working." We grabbed a microphone and asked the audience to pass it around, but it still wasn't enough. By the end of the 3-hour meeting, five microphones were bouncing around the socially distanced group, and I could sit back in my chair, take in the slides, and let the audio come to me.

Now, a few weeks later, my anxiety is gradually giving way to scientific curiosity and questions enabled by feeling included and worthy of belonging. A few of my lab mates have even thanked me because they, too, were unable to hear people without microphones. Maybe instead of rushing back to "normal," we can all take this opportunity to create environments that are more welcoming—for everyone. ∎

Carolyn J. Adamski is an HHMI postdoctoral fellow at Baylor College of Medicine and the Jan and Dan Duncan Neurological Research Institute.

ILLUSTRATION: ROBERT NEUBECKER

Science Webinars help you keep pace with emerging scientific fields!

Stay informed about scientific breakthroughs and discoveries.

Gain insights into current research from top scientists.

Take the opportunity to ask questions during live broadcasts.

Get alerts about upcoming free webinars.

Sign up at: webinar.sciencemag.org/stayinformed

EMP☉WER
WITH EVIDENCE

AAAS | ANNUAL MEETING

SUBMIT YOUR STUDENT E-POSTER ABSTRACT BY SEPTEMBER 14!

E-posters will be presented virtually and include **a single-page PDF with static text**, submitted by student presenters, accompanied with **oral video presentation**.

The 2022 AAAS Annual Meeting will be held in a **hybrid format**, taking place online and in Philadelphia February 17-20.

aaas.org/meetings | #AAASmtg