Science

THE SYSTEMIC
MICROBIOME

How diverse microbial communities influence health and disease p. 932

A MULTIPLICITY OF MICROBIOMES

By **Priscilla Kelly**, **Gemma Alderton**, **Seth Thomas Scanlon**, *and* **Caroline Ash**

Humans have many distinct
and interconnecting microbial
populations that exert
systemic effects throughout
the body. Understanding the
ways these communities
interact provides insight into
how the collective microbiome
shapes health and disease.

The healthy human body is home to an abundance of bacteria, viruses, and fungi that are acquired in early postnatal life. The bulk of the microbiome is in the gut, but distinct microbial communities exist on the skin and in the mouth, nose, lungs, and genital tract. The human microbiome has important roles in maintaining homeostasis, and disruption of microbial colonization of an infant has systemic effects that may influence health later in life, potentially promoting the development of autoimmunity, allergies, metabolic diseases, and even cancer.

Studies of microbial niches in the skin have revealed that they are integral to the skin's multifaceted barrier function. At the same time, there is growing appreciation for how cross-talk between the skin microbiome and the immune system influences (and is influenced by) distant organs such as the gut, lungs, and brain. The microbiome in the oral cavity may also exacerbate diseases in distal sites, such as colorectal cancer, asthma, rheumatoid arthritis, and cardiovascular disease.

Given their abundance, interactions among diverse microbial communities will likely influence host physiology. If a microbial population is disrupted by antibiotics or a low-fiber diet, then pathogens can gain advantage, resulting in systemic infection and inflammation. Some emerging clinical evidence indicates that the "soup" of metabolites and signaling molecules produced by the gut microbiota could influence many diseases, not to mention behavioral states, but there are few examples of verified mechanisms.

Understanding the ecology as well as the mechanisms of cross-talk between a host and its microbiota should offer a more holistic view of human health that one day will contribute to preventing and treating a spectrum of human diseases. But we need to tool up; we are not there yet.

PERSPECTIVE

The effects of oral microbiota on health

Oral microbiota form complex biofilms that can affect local and systemic health

By **Timur Tuganbaev**[1], **Koji Yoshida**[1], **Kenya Honda**[1,2]

The oral microbiota is shaped by mutualistic coevolution with the host and by the distinct physiology of the mouth. In an evolutionary quid pro quo, the host provides the commensal bacteria with a stable ecological niche and in return the oral microbiota supports host health locally by forming symbiotic biofilms that balance pH levels and suppress pathogen growth, as well as potentially by systemically reinforcing the body's physiological processes, such as cardiovascular homeostasis. However, when a biofilm transforms into a dysbiotic state, which is no longer in homeostatic equilibrium with the host, the oral microbiota may contribute to the pathological processes of a wide range of diseases, including inflammatory bowel disease (IBD), arthritis, colorectal and pancreatic cancers, and Alzheimer's disease, by serving as a reservoir for opportunistic pathogens. This in turn presents opportunities to identify predictive disease biomarkers within the oral microbiota and the development of intervention strategies to promote oral and overall health.

The oral microbiota is the second most diverse microbial community colonizing the human body, after the intestinal microbiota. In contrast to the relatively simple spatial organization of the intestinal microbiota, the oral microbiota is formed by a collection of compositionally distinct microbial communities reflecting an array of diverse microenvironments. The host tightly controls the composition of oral microbiota by imposing selective pressures that require specialized metabolic machinery (e.g., metabolism of carbohydrates that are digested by salivary α-amylase and of highly glycosylated salivary proteins) and the ability to adhere to specific substrates (e.g., buccal epithelium, epithelium of tongue papillae, and teeth surfaces) despite salivary flow, drinking, and chewing. For instance, oral streptococci encode a family of serine-rich repeat–containing glycoproteins and antigen I/II family of adhesins to bind to the tooth pellicle, a thin layer formed by salivary proteins. Streptococci have also

[1]Department of Microbiology and Immunology, Keio University School of Medicine, Tokyo, Japan. [2]RIKEN Center for Integrative Medical Sciences (IMS), Kanagawa, Japan. Email: kenya@keio.jp

evolved amylase-binding proteins to capture human salivary amylase and use it to metabolize starch.

The host also controls the composition of oral microbiota through salivary secretory immunoglobulin A (SIgA) and antimicrobial peptides such as lysozyme and lactoferrin. Salivary SIgAs promote aggregation and subsequent elimination of potentially pathogenic bacteria, whereas SIgAs embedded in tooth pellicles and mucin layers covering oral epithelium provide binding sites for beneficial bacteria, although it is unclear how the salivary SIgAs discriminate beneficial versus pathogenic microbes. The need to control microbes with saliva is exemplified in patients with xerostomia (dry mouth), who often show an increased abundance of fungi and acid-producing and acid-resistant cariogenic bacteria (which cause tooth decay). The multiple control pathways of the host direct the assembly of spatially specialized bacterial cooperatives called oral biofilms. Within a biofilm, different bacterial species are connected through physical and metabolic associations that confer a fitness advantage to the entire microbial community. Biofilm structure also promotes horizontal transfer of genes, including antibiotic-resistance genes. Although their origin is not well defined, antibiotic-resistance genes can be detected even in the oral microbiome of healthy individuals not previously exposed to antibiotics (1). Consequently, whereas the intestinal community is sensitive to dietary change or antibiotic treatment, oral biofilms are especially stable and resilient to external perturbations.

The ability to self-assemble into complex and site-specific biofilms on the soft and hard tissues of the mouth plays a central role in oral microbiota–host interactions in health (see the figure). For example, structural analysis of supragingival bacterial biofilms revealed complex layered structures indicative of division of labor among species (2). *Corynebacterium* filaments form a structural underframe for the entire biofilm and anchor it to the salivary pellicle. At the surface of the biofilm, cocci such as *Streptococcus* consume sugars and oxygen and produce lactate, acetate, and CO_2, thereby supporting facultative anaerobic bacteria in the deeper layers of the biofilm. *Streptococcus* also fulfills a

defensive role by producing H_2O_2 that may protect the biofilm consortium and its host from infections (2).

The dorsal tongue biofilms present on the filiform papillae feature a distinct architecture and composition, which contain nitrate-reducing bacteria, such as *Neisseria* and *Rothia* (3). Nitrate is actively transported from blood into saliva by the sialin transporter and is further transformed to nitrite by oral microbes, raising the salivary nitrite concentration to 1000 times that of plasma (4). The oral nitrate-reducing bacteria are mostly facultative anaerobes that use nitrate as an alternative electron acceptor for respiration. It has been suggested that absorbed nitrite is eventually converted to NO, which in turn provides important benefits to the host including lowering blood pressure, improved endothelial function, reversal of metabolic syndrome, and reduction of oxidative stress, likely through mechanisms involving activation of the guanylate cyclase–cyclic guanosine monophosphate pathway (4). Intriguingly, the function of the filiform papillae remains unclear because they lack taste buds. They do, however, greatly increase the surface area for biofilm formation and might thus function as bioreactors that take advantage of the large enzymatic repertoire of the microbiota. These observations suggest that oral microbiota play integral roles in maintaining host health systemically and locally.

Although the host and oral microbiota probably exist in a mutually beneficial equilibrium in homeostasis, external factors that are not fully defined may trigger a vicious cycle of self-perpetuating dysbiosis. Several oral pathologies, including periodontitis, are not contagious per se but result from oral microbiota dysbiosis in which inflammation-mediated tissue destruction leads to generation of metabolites that fuel pathogenic proinflammatory microbes. In particular, *Porphyromonas gingivalis* is considered a "keystone pathogen" because it plays a critical role in maintaining the structure of an inflammatory biofilm by subverting host immune and inflammatory responses, and its impact on the community is greater than would be expected on the basis of its relative abundance (5). Gingipains, virulence factors that are secreted by *P. gingivalis*, are cysteine proteases that promote proteolytic activation of complement and induce complement

receptor–mediated inflammatory cytokine production by neutrophils. Concurrently, *P. gingivalis* inhibits neutrophil bacterial killing activity by disarming the host-protective Toll-like receptor pathways. *P. gingivalis* thus creates the inflammatory conditions that it needs to thrive, yet avoids the antibacterial response, thereby perpetuating dysbiosis and periodontitis (*5*). Gingipains also induce vascular permeability, which thus grants *P. gingivalis*, and other bacteria, systemic reach. Notably, gingipain genes of *P. gingivalis* are detectable in calcified tooth tartar samples from ancient humans (*6*),

proportion of oral microbes among the intestinal microbiota can be observed in several diseases. For example, oral microbes are significantly more abundant in the fecal microbiota of patients with IBD, primary sclerosing cholangitis (bile duct inflammation), multiple sclerosis, and gastroesophageal reflux disease compared with healthy controls (*8*). Although the enrichment of oral bacteria in the intestine could be a consequence of a depleted gut commensal population, several reports demonstrate the causal contribution of oral microbes to disease pathogenesis. For example, *Klebsiella* strains car-

in mice (*11*). Therefore, the oral cavity could act as a reservoir for opportunistic pathogens that can colonize the intestine and might causally affect chronic inflammation, as well as cancer development and metastasis.
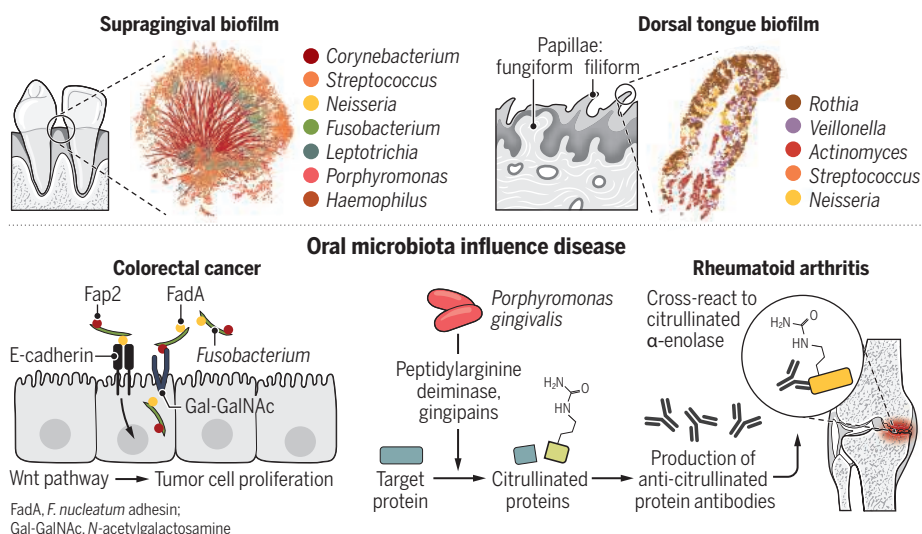
The mouth is a highly vascularized organ. As a result, oral microbes and microbial molecules might directly enter the bloodstream and contribute to the pathogenesis of systemic diseases. The development of rheumatoid arthritis is promoted through a loss of immune tolerance to citrullinated proteins. Concordantly, amounts of anti-citrullinated protein antibodies (ACPAs) are predictive for current and future disease onset. ACPA-positive individuals exhibit a higher relative abundance of *P. gingivalis* in their oral microbiota. Although it is unclear whether *P. gingivalis* plays a causal role, it may contribute to disease pathogenesis through the generation of citrullinated proteins. Gingipains proteolytically cleave bacterial and host proteins at arginine residues, and the short peptides with carboxyl-terminal arginine are immediately citrullinated by bacterial peptidylarginine deiminase (PAD) that is secreted with gingipains. It has been suggested that citrullinated proteins elicit the generation of antibodies that cross-react with bacterial and host citrullinated molecules through molecular mimicry (*12*). *Aggregatibacter actinomycetemcomitans*, associated with periodontitis, also has the capacity to produce citrullinated autoantigens through a mechanism distinct from that of *P. gingivalis*, thereby also inducing ACPAs through molecular mimicry (*13*).

Epidemiological studies have identified an association between periodontitis, cognitive function, and dementia. Recent work using brain specimens and cerebrospinal fluid from individuals diagnosed with Alzheimer's disease suggested that *P. gingivalis* could colonize the brain and induce neurodegeneration (*14*). *P. gingivalis* gingipains are found in association with neurons, tau tangles, and amyloid-β (Aβ$_{1-42}$) depositions in Alzheimer's disease specimens. Although further investigations will be required, *P. gingivalis* might translocate to the brain by hematogenous routes, leading to activation of the complement cascade, increasing amyloid-β production, augmenting proinflammatory cytokine expression, and causing neuroinflammation and neurodegeneration (*14*). These studies suggest that oral microbiota dysbiosis may have systemic ramifications that promote development of autoimmune and neurodegenerative diseases.

The oral microbiota is of particular interest in the search for taxonomic and molecular biomarkers of oral and systemic diseases owing to the ease of sampling and its compositional stability. Although the

Oral microbiota in health and disease

Oral microbiota form distinct biofilms in different regions of the mouth that protect host tissues [based on (*2,3*)]. When homeostasis between host and microbiota is lost, oral microbiota can affect distant tissues and influence, for example, colorectal cancer cell proliferation and autoimmunity that leads to rheumatoid arthritis.

Supragingival biofilm
- Corynebacterium
- Streptococcus
- Neisseria
- Fusobacterium
- Leptotrichia
- Porphyromonas
- Haemophilus

Dorsal tongue biofilm
Papillae: fungiform filiform
- Rothia
- Veillonella
- Actinomyces
- Streptococcus
- Neisseria

Oral microbiota influence disease

Colorectal cancer
Fap2 FadA
E-cadherin
Fusobacterium
Gal-GalNAc
Wnt pathway → Tumor cell proliferation

FadA, *F. nucleatum* adhesin;
Gal-GalNAc, *N*-acetylgalactosamine

Porphyromonas gingivalis
Peptidylarginine deiminase, gingipains
Target protein → Citrullinated proteins → Production of anti-citrullinated protein antibodies

Rheumatoid arthritis
Cross-react to citrullinated α-enolase

suggesting that these virulence factors have aided *P. gingivalis* to persist in the oral microbiota for 100,000 years.

In a pathological context, oral cavity–associated microbes have been detected in distant organs in humans, including the intestine, lungs, heart, and brain. Indeed, numerous studies have suggested a relationship between oral hygiene and respiratory diseases, including asthma, chronic obstructive pulmonary disease, and pneumonia (*7*). Although these findings demonstrate associations rather than causality, oral microbiota are likely microaspirated and affect the lung microbiome. Oral bacteria are also swallowed with saliva (1 ml of saliva can contain up to 10^7 to 10^8 microbes) but are exposed to gastric acid, bile acids, anaerobic environments, and numerous other colonization resistance mechanisms conferred by the gut microbiota. Therefore, intestinal colonization by oral microbiota is limited in healthy individuals. However, an increased

rying antibiotic-resistance genes isolated from saliva samples from patients with IBD can persist in the intestine of antibiotics-treated mice and drive intestinal inflammation (*8*). Consistently, intestinal enrichment of *Klebsiella* species has been associated with IBD in multiple human studies (*9*).

Opportunistic infections by oral microbes may also contribute to the development of cancer in the alimentary tract. For example, oral microbes such as *Fusobacterium*, *Parvimonas*, and *Peptostreptococcus* have been associated with colorectal cancer (CRC) development. In mice, *F. nucleatum* colonizes CRC tissues through FadA adhesin–mediated binding to host E-cadherin and accelerates proliferation of tumor cells through activation of the Wnt–β-catenin signaling pathway (*10*). Moreover, *Fusobacterium* species can migrate with metastasizing CRC cells, and intracellular *Fusobacterium*-positive CRC cells established metastatic lesions more effectively than *Fusobacterium*-negative cells

causal relationship between oral microbiota dysbiosis and systemic disease has not been fully addressed, characteristic alterations in the oral microbiota in specific disease conditions could have value as a diagnostic tool. Considering that the concurrent expansion of oral microbes in the intestine is often associated with the development of systemic inflammatory diseases, combined examination of salivary and fecal microbiota may improve diagnostic accuracy. Indeed, a combinatorial analysis of oral and stool microbiota improved discrimination of CRCs from colon polyps (15).

Conventional treatments of oral diseases have limited effects in restoring a healthy bacterial community, underscoring the need to develop approaches to manipulate the oral ecosystem. Clinical studies of fecal microbiota transplantation have raised the possibility of developing oral microbiota transplantation. Additionally, oral bacteria–derived molecules may be promising targets for intervention. For example, nitrate-containing oral hygiene products might help promote beneficial bacteria. Additional approaches to target the oral microbiota could also include genetically modified bacteria that secrete desired proteins and bacteriophages to target specific microbes. Bacteriophages have high host specificity and typically infect and kill a single bacterial species, and therefore could be used to rationally manipulate the oral microbiota structure by targeting keystone pathobionts. Employing multiple strategies simultaneously will likely maximize the benefit of future oral microbiota-targeted therapeutics for the advancement of oral and overall health. ∎

REFERENCES AND NOTES

1. J. C. Clemente et al., Sci. Adv. 1, e1500183 (2015).
2. J. L. Mark Welch, B. J. Rossetti, C. W. Rieken, F. E. Dewhirst, G. G. Borisy, Proc. Natl. Acad. Sci. U.S.A. 113, E791 (2016).
3. S. A. Wilbert, J. L. Mark Welch, G. G. Borisy, Cell Rep. 30, 4003 (2020).
4. J. O. Lundberg, M. Carlström, E. Weitzberg, Cell Metab. 28, 9 (2018).
5. T. Maekawa et al., Cell Host Microbe 15, 768 (2014).
6. J. A. Fellows Yates et al., Proc. Natl. Acad. Sci. U.S.A. 118, e2021655118 (2021).
7. L. A. Mandell, M. S. Niederman, N. Engl. J. Med. 380, 651 (2019).
8. K. Atarashi et al., Science 358, 359 (2017).
9. J. Lloyd-Price et al., Nature 569, 655 (2019).
10. M. R. Rubinstein et al., Cell Host Microbe 14, 195 (2013).
11. S. Bullman et al., Science 358, 1443 (2017).
12. N. Wegner et al., Arthritis Rheum. 62, 2662 (2010).
13. M. F. Konig et al., Sci. Transl. Med. 8, 369ra176 (2016).
14. S. S. Dominy et al., Sci. Adv. 5, eaau3333 (2019).
15. B. Flemer et al., Gut 67, 1454 (2018).

PERSPECTIVE

Modulating brain function with microbiota

Microbial metabolites identified in animal models and human neurological diseases could be therapeutic targets

By **Jane A. Foster**[1,2]

From the discovery of roles of the gut microbiota in metabolic disorders such as obesity (1) to recent discoveries of gut microbiota modulating responses to cancer immunotherapy (2), microbiome research has expanded to all areas of biomedical research over the past 15 years. A vital role for gut microbiota–brain communication in brain development, behavior, and function has emerged (3). Research using germ-free mice has played a considerable role in identifying brain systems that may be regulated by microbiota, including blood-brain barrier permeability, brain volume, neural circuitry, myelination, and alterations in microglia (3). Microbial-derived molecules, including neurotransmitters, short-chain fatty acids, bile acids, lactate, and vitamins, exert local effects in the gastrointestinal environment but can also enter the circulation to act at remote sites, including the brain (4). Recent efforts to translate preclinical findings in amyotrophic lateral sclerosis (ALS) and autism spectrum disorder (ASD) to the clinic highlight the potential for clinically important discoveries in microbiota–brain research (5–7).

Microbe–host interactions at the intestinal barrier play diverse roles in pathogen defense, immune system development, energy metabolism, and gastrointestinal physiology and homeostasis. In the context of the gut-brain axis, neural, endocrine, metabolic, and immune pathways facilitate microbiota-to-brain signaling (3). In the search to identify direct mechanisms that connect gut microbiota to brain health, the potential role of microbial metabolite pathways has become evident. Microbial genes outnumber host genes such that the genetic and metabolic potential of the metagenome substantially adds to the biochemical flexibility of the host (8). Precise mapping of the biological mechanisms that connect gut microbiota to brain function is needed to identify microbiome-associated modulators that contribute to pathophysiological processes and to assess their causative effects in human disease.

Recent studies of microbiota–brain connections have effectively translated mechanistic findings in animal models to corresponding clinical populations (5–7) (see the figure). For example, a microbial metabolite signaling pathway identified in an animal model of ALS is comparable to metabolite alterations in serum from patients with ALS (5). ALS is a neurodegenerative motor neuron disease that leads to muscle weakness and atrophy. Microbiota depletion worsened ALS-like motor symptoms in superoxide dismutase-1 mutant ($Sod1^{G93A}$) mice, a model of ALS (5). Metagenomic sequencing revealed compositional and functional differences between the gut microbiome from $Sod1^{G93A}$ mice and wild-type littermates. Notably, several bacterial genera were associated with disease features, and bacterial genes encoding enzymes involved in nicotinamide (NAM) and nicotinate, or niacin (NA), metabolism were reduced in $Sod1^{G93A}$ mice (5). Dietary NAM and NA, commonly known as vitamin B3, are precursors for nicotinamide adenine dinucleotide (NAD) (9), which plays an essential role in cellular metabolism. NAD concentrations in brain tissue, measured using magnetic resonance spectroscopy, are reduced in aging (9, 10), demonstrating the potential of this signaling cascade as a therapeutic target in aging-associated diseases.

Using a probiotic approach, colonization of antibiotic-treated $Sod1^{G93A}$ mice with ALS-associated *Akkermansia muciniphila* had a beneficial effect on disease severity, whereas colonization with ALS-associated *Ruminococcus torques* or *Parabacteroides distanonis* exacerbated ALS symptoms in antibiotic-treated $Sod1^{G93A}$ mice. Moreover, improvement in disease severity after *A. muciniphila* colonization of $Sod1^{G93A}$ mice was associated with enrichment of NAM biosynthetic intermediates in serum, and independent supplementation of $Sod1^{G93A}$ mice with NAM also improved behavioral and neurological motor tests (5).

In a cohort of 37 ALS patients, reduced serum and cerebrospinal fluid (CSF) concentrations of NAM compared with 29 healthy individuals were observed. Compositional differences in gut microbiota between ALS patients and healthy individuals were also found; however, a specific role for *Akkermansia* or other taxa in affecting NAM

causal relationship between oral microbiota dysbiosis and systemic disease has not been fully addressed, characteristic alterations in the oral microbiota in specific disease conditions could have value as a diagnostic tool. Considering that the concurrent expansion of oral microbes in the intestine is often associated with the development of systemic inflammatory diseases, combined examination of salivary and fecal microbiota may improve diagnostic accuracy. Indeed, a combinatorial analysis of oral and stool microbiota improved discrimination of CRCs from colon polyps (*15*).

Conventional treatments of oral diseases have limited effects in restoring a healthy bacterial community, underscoring the need to develop approaches to manipulate the oral ecosystem. Clinical studies of fecal microbiota transplantation have raised the possibility of developing oral microbiota transplantation. Additionally, oral bacteria–derived molecules may be promising targets for intervention. For example, nitrate-containing oral hygiene products might help promote beneficial bacteria. Additional approaches to target the oral microbiota could also include genetically modified bacteria that secrete desired proteins and bacteriophages to target specific microbes. Bacteriophages have high host specificity and typically infect and kill a single bacterial species, and therefore could be used to rationally manipulate the oral microbiota structure by targeting keystone pathobionts. Employing multiple strategies simultaneously will likely maximize the benefit of future oral microbiota-targeted therapeutics for the advancement of of oral and overall health. ∎

REFERENCES AND NOTES

1. J. C. Clemente *et al.*, *Sci. Adv.* **1**, e1500183 (2015).
2. J. L. Mark Welch, B. J. Rossetti, C. W. Rieken, F. E. Dewhirst, G. G. Borisy, *Proc. Natl. Acad. Sci. U.S.A.* **113**, E791 (2016).
3. S. A. Wilbert, J. L. Mark Welch, G. G. Borisy, *Cell Rep.* **30**, 4003 (2020).
4. J. O. Lundberg, M. Carlström, E. Weitzberg, *Cell Metab.* **28**, 9 (2018).
5. T. Maekawa *et al.*, *Cell Host Microbe* **15**, 768 (2014).
6. J. A. Fellows Yates *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2021655118 (2021).
7. L. A. Mandell, M. S. Niederman, *N. Engl. J. Med.* **380**, 651 (2019).
8. K. Atarashi *et al.*, *Science* **358**, 359 (2017).
9. J. Lloyd-Price *et al.*, *Nature* **569**, 655 (2019).
10. M. R. Rubinstein *et al.*, *Cell Host Microbe* **14**, 195 (2013).
11. S. Bullman *et al.*, *Science* **358**, 1443 (2017).
12. N. Wegner *et al.*, *Arthritis Rheum.* **62**, 2662 (2010).
13. M. F. Konig *et al.*, *Sci. Transl. Med.* **8**, 369ra176 (2016).
14. S. S. Dominy *et al.*, *Sci. Adv.* **5**, eaau3333 (2019).
15. B. Flemer *et al.*, *Gut* **67**, 1454 (2018).

PERSPECTIVE

Modulating brain function with microbiota

Microbial metabolites identified in animal models and human neurological diseases could be therapeutic targets

By **Jane A. Foster**[1,2]

From the discovery of roles of the gut microbiota in metabolic disorders such as obesity (*1*) to recent discoveries of gut microbiota modulating responses to cancer immunotherapy (*2*), microbiome research has expanded to all areas of biomedical research over the past 15 years. A vital role for gut microbiota–brain communication in brain development, behavior, and function has emerged (*3*). Research using germ-free mice has played a considerable role in identifying brain systems that may be regulated by microbiota, including blood-brain barrier permeability, brain volume, neural circuitry, myelination, and alterations in microglia (*3*). Microbial-derived molecules, including neurotransmitters, short-chain fatty acids, bile acids, lactate, and vitamins, exert local effects in the gastrointestinal environment but can also enter the circulation to act at remote sites, including the brain (*4*). Recent efforts to translate preclinical findings in amyotrophic lateral sclerosis (ALS) and autism spectrum disorder (ASD) to the clinic highlight the potential for clinically important discoveries in microbiota–brain research (*5–7*).

Microbe–host interactions at the intestinal barrier play diverse roles in pathogen defense, immune system development, energy metabolism, and gastrointestinal physiology and homeostasis. In the context of the gut-brain axis, neural, endocrine, metabolic, and immune pathways facilitate microbiota-to-brain signaling (*3*). In the search to identify direct mechanisms that connect gut microbiota to brain health, the potential role of microbial metabolite pathways has become evident. Microbial genes outnumber host genes such that the genetic and metabolic potential of the metagenome substantially adds to the biochemical flexibility of the host (*8*). Precise mapping of the biological mechanisms that connect gut microbiota to brain function is needed to identify microbiome-associated modulators that contribute to pathophysiological processes and to assess their causative effects in human disease.

Recent studies of microbiota–brain connections have effectively translated mechanistic findings in animal models to corresponding clinical populations (*5–7*) (see the figure). For example, a microbial metabolite signaling pathway identified in an animal model of ALS is comparable to metabolite alterations in serum from patients with ALS (*5*). ALS is a neurodegenerative motor neuron disease that leads to muscle weakness and atrophy. Microbiota depletion worsened ALS-like motor symptoms in superoxide dismutase-1 mutant (*Sod1*[G93A]) mice, a model of ALS (*5*). Metagenomic sequencing revealed compositional and functional differences between the gut microbiome from *Sod1*[G93A] mice and wild-type littermates. Notably, several bacterial genera were associated with disease features, and bacterial genes encoding enzymes involved in nicotinamide (NAM) and nicotinate, or niacin (NA), metabolism were reduced in *Sod1*[G93A] mice (*5*). Dietary NAM and NA, commonly known as vitamin B3, are precursors for nicotinamide adenine dinucleotide (NAD) (*9*), which plays an essential role in cellular metabolism. NAD concentrations in brain tissue, measured using magnetic resonance spectroscopy, are reduced in aging (*9, 10*), demonstrating the potential of this signaling cascade as a therapeutic target in aging-associated diseases.

Using a probiotic approach, colonization of antibiotic-treated *Sod1*[G93A] mice with ALS-associated *Akkermansia muciniphila* had a beneficial effect on disease severity, whereas colonization with ALS-associated *Ruminococcus torques* or *Parabacteroides distanonis* exacerbated ALS symptoms in antibiotic-treated *Sod1*[G93A] mice. Moreover, improvement in disease severity after *A. muciniphila* colonization of *Sod1*[G93A] mice was associated with enrichment of NAM biosynthetic intermediates in serum, and independent supplementation of *Sod1*[G93A] mice with NAM also improved behavioral and neurological motor tests (*5*).

In a cohort of 37 ALS patients, reduced serum and cerebrospinal fluid (CSF) concentrations of NAM compared with 29 healthy individuals were observed. Compositional differences in gut microbiota between ALS patients and healthy individuals were also found; however, a specific role for *Akkermansia* or other taxa in affecting NAM

concentrations in human ALS has yet to be demonstrated (*5*). Thus, this represents a potentially clinically important microbial metabolite–brain signaling pathway that might, in the future, be modulated to alleviate ALS symptoms (*5*) and potentially other diseases of aging.

Harnessing discoveries in microbiota–brain connections could improve clinical outcomes in patients with neurological disorders. Recent studies in ASD demonstrate the potential for clinically relevant findings. A substantial body of preclinical and clinical evidence demonstrates a role for the gut microbiota in ASD (*11*). Although causative microbiome-associated mechanisms in ASD in humans have not been established, case-control studies demonstrate alterations in gut microbiota composition, as well as altered plasma and fecal metabolic phenotypes, in individuals with ASD when compared with typically developing children and adolescents (*12, 13*). Moreover, several bacterial taxa and signaling pathways are associated with clinical symptoms and severity (*12*). To translate these findings, it will be necessary to move beyond correlative evidence linking microbiota to ASD to specific microbiota–brain mechanisms.

A recent study identified a microbial metabolite pathway that affects anxiety-like behavior in mice, a clinical symptom often present in ASD (*6*). Previous preclinical and clinical work implicated a pathophysiological impact of larger amounts of the microbiota-derived aromatic metabolite 4-ethylphenyl sulfate (4-EPS) on behavior in the maternal immune activation mouse model of ASD and in individuals with ASD (*12, 14*). It was demonstrated that the gut commensal bacterium *Bacteroides ovatus* can metabolize tyrosine, a common dietary amino acid, to produce *p*-coumaric acid, which in turn can be metabolized by a second commensal bacterium, *Lactobacillus plantarum*, to produce 4-EP, which is then sulfonated by host cells to produce 4-EPS. Indeed, 4-EPS was detected in the serum and brain tissue of mice colonized with bacterial strains engineered to produce 4-EP (4-EP⁺) (*6*). Brain changes in 4-EP⁺ mice included differential expression of oligodendrocyte maturation and myelination genes, lower structural connectivity, and increased anxiety-like behavior. Furthermore, treatment of 4-EP⁺ mice with clemastine fumarate, to promote oligodendrocyte maturation, improved behavior, suggesting that a gut-derived metabolite can signal to the brain to influence anxiety-like behavior through
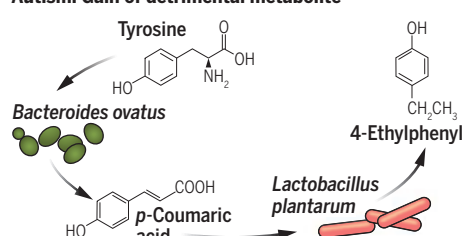
Microbial metabolite–brain signaling pathways

Mechanistic insights from animal studies show that gut bacteria–derived metabolites affect brain function. Comparable changes in gut microbial metabolites in human amyotrophic lateral sclerosis (ALS) and autism spectrum disorder suggest that exploration of therapeutic approaches, such as AB-2004 to reduce 4-ethylphenyl sulfate (4-EPS), is warranted.
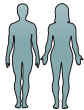

ALS: Loss of beneficial metabolite

Akkermansia muciniphila → Nicotinamide

Autism: Gain of detrimental metabolite

Tyrosine → *Bacteroides ovatus* → *p*-Coumaric acid → *Lactobacillus plantarum* → 4-Ethylphenyl

Mouse model
↓ *A. muciniphila*
↓ Serum nicotinamide
↑ Disease severity

↑ 4-EPS in serum and brain
↑ Anxiety-like behavior
Altered oligodendrocyte gene expression
↓ Structural connectivity
AB-2004 treatment normalized anxiety-like behavior

In humans
Altered metagenome
↓ Serum and cerebrospinal fluid nicotinamide

Altered metagenome
↑ Plasma and fecal 4-EPS
AB-2004 treatment reduced anxiety and irritability

effects on oligodendrocytes and myelination (*6*). The mechanism of how increased 4-EPS concentrations affect oligodendrocyte maturation, myelination, and brain connectivity has not been identified.

In a parallel study, 4-EPS was identified as an actionable therapeutic target with potential clinical benefit in ASD (*7*). A proof-of-concept experiment in mice showed that oral administration of the gastrointestinal-restricted adsorbent drug AB-2004, which binds and sequesters aromatic metabolites (*15*), reduced 4-EPS concentrations in the urine of 4-EP⁺ mice and ameliorated anxiety-like behavioral deficits (*7*). A companion phase 1 open-label clinical trial demonstrated that oral administration of AB-2004 was safe and well tolerated in adolescents with autism and reduced both anxiety and irritability, particularly in individuals with high scores on these indices at baseline (*7*). In this pilot clinical trial, treatment with AB-2004 for 8 weeks resulted in reduced urinary and plasma concentrations of 4-EPS and other gut-derived aromatic metabolites at the end of treatment, although metabolite concentrations returned to baseline 4 weeks after the end of treatment (*7*). The study also reports potential benefits on core autism behaviors, social communication, and repetitive behaviors (*7*). This preliminary observational trial suggests that targeting gut-derived metabolites in ASD may be beneficial.

Translation of mechanistic insights discovered in animal models to clinical populations can guide researchers to move beyond correlation and association toward clinically actionable microbiota–brain signaling mechanisms. Pathophysiological changes can result from depletion of a protective microbial metabolite, such as NAM, and from

elevation of a detrimental microbial metabolite, such as 4-EPS, suggesting an opportunity to provide supplements or therapeutically target the microbe and/or the metabolite, respectively. Moving from basic science discoveries to innovative clinical interventions is challenging. Bacterial taxa that play a role in animal models may not be the same in human disease. Identifying specific compositional and functional changes in microbial metabolite signaling in clinical samples is required to validate targets. The substantial clinical heterogeneity in neurological disorders is also challenging and necessitates multiple parallel approaches to develop microbial, pharmacological, dietary, and lifestyle treatment strategies to improve clinical outcomes. ∎

REFERENCES AND NOTES

1. P. J. Turnbaugh *et al.*, *Nature* **444**, 1027 (2006).
2. E. Ansaldo, Y. Belkaid, *Science* **373**, 966 (2021).
3. J. F. Cryan *et al.*, *Physiol. Rev.* **99**, 1877 (2019).
4. G. Caspani, S. Kennedy, J. A. Foster, J. Swann, *Microb. Cell* **6**, 454 (2019).
5. E. Blacher *et al.*, *Nature* **572**, 474 (2019).
6. B. D. Needham *et al.*, *Nature* **602**, 647 (2022).
7. A. Stewart Campbell *et al.*, *Nat. Med.* **28**, 528 (2022).
8. R. Sender, S. Fuchs, R. Milo, *PLOS Biol.* **14**, e1002533 (2016).
9. N. Braidy *et al.*, *Antioxid. Redox Signal.* **30**, 251 (2019).
10. X.-H. Zhu, M. Lu, B.-Y. Lee, K. Ugurbil, W. Chen, *Proc. Natl. Acad. Sci. U.S.A.* **112**, 2876 (2015).
11. H. E. Vuong, E. Y. Hsiao, *Biol. Psychiatry* **81**, 411 (2017).
12. B. D. Needham *et al.*, *Biol. Psychiatry* **89**, 451 (2021).
13. F. Liu *et al.*, *Transl. Psychiatry* **9**, 43 (2019).
14. E. Y. Hsiao *et al.*, *Cell* **155**, 1451 (2013).
15. T. Niwa *et al.*, *Nephrol. Dial. Transplant.* **6**, 105 (1991).

¹Department of Psychiatry and Behavioural Neurosciences, McMaster University, Hamilton, ON, Canada. ²Center for Depression Research and Clinical Care, Department of Psychiatry, UT Southwestern Medical Center, Dallas, TX, USA. Email: jfosterlab@gmail.com

GRAPHIC: N. CARY/SCIENCE

PERSPECTIVE

Microbiota–brain axis: Context and causality

Gut bacteria influence the brain and behavior, but causation in humans remains unclear

By **John F. Cryan**[1] *and* **Sarkis K. Mazmanian**[2]

The gut microbiota is associated with brain development and function, as well as altered emotional, motor, and cognitive behaviors in animals. However, there remains a pressing need to unravel the mechanisms and pathways of communication that underpin microbiota-brain connections. Reductionist animal models have revealed profound contributions by gut bacteria to brain activity and behavior, although the extent to which these findings translate to humans is largely unclear. Associations between fecal microbiome profiles and human behavior and neurological diseases are prevalent. Longitudinal studies that integrate genetic, environmental, and experiential factors in shaping and responding to gut microbial functions, as well as interventions that modify microbiota-brain interactions at the cellular and molecular level, will help resolve contexts in which microbes may influence the human brain and its health. Accordingly, prospects for targeting the gut, rather than the brain, to improve altered behaviors or brain pathologies appear both feasible and timely.

Exposure to microbes at birth and in early life blossoms into diverse microbial ecosystems that inhabit the skin, oral cavity, vaginal cavity, and lungs, with most human-associated bacteria harbored in the lower gastrointestinal (GI) tract. The human microbiome has been profiled in individuals from numerous geographies and with diverse conditions. Metagenomic surveys allow for reconstruction of bacterial biosynthetic pathways and provide insights into potential microbiome functions. Coupling DNA sequencing to metatranscriptomics, metabolomics, lipidomics, and proteomics offers exciting opportunities to uncover molecular mediators of the gut microbiota–brain axis.

Studies using germ-free (GF) mice, which are devoid of all microorganisms, have pro-

[1]APC Microbiome Ireland, University College Cork, Cork, Ireland. [2]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, USA. Email: j.cryan@ucc.ie; sarkis@caltech.edu

vided compelling and converging evidence that the microbiota is crucial for brain development (1). GF mice display changes in brain expression of neurotransmitters and their receptors as well as neurotrophic factors, show region-specific gene expression, and have an impaired blood-brain barrier. Even hippocampal neurogenesis and neuroplasticity in mice are affected by the absence of a gut microbiota. Oligodendrocyte maturation is restrained by specific gut microbial metabolites that alter myelination patterns within the limbic system of

Advancing human microbiota research

Most evidence linking the gut microbiota to brain phenotypes and behaviors is from animal models. These preclinical studies have informed strategies for further analyses and interventions in humans, but more rigorous human studies are needed to distinguish causality from association.

What is needed to inform causality in humans?

Define "normal" and "healthy" gut microbiota

Understand the effect of lifestyle

Perform large-scale translational and longitudinal studies

Use more accurate controls

Define directionality between cause and effect

Understand the contribution of gene-environment impacts

Document a complete inventory of human-associated microbiota

mice, which promotes anxiety-like behaviors (2). Furthermore, GF or antibiotic-treated mice have increased risk-taking, hyperactivity, and feeding behaviors and display learning and memory deficits (3). Across a variety of animal species—including flies, worms, and fish—there is growing evidence for a link between complex behaviors and gut microbiota composition (3, 4). Moreover, altered bacterial communities, antibiotic treatment, and fecal microbiota transplants (FMTs) affect symptoms and pathology in mouse models of anxiety, depression, autism spectrum dis-

order (ASD), epilepsy, Parkinson's disease (PD), amyotrophic lateral sclerosis (ALS), and Alzheimer's disease (5).
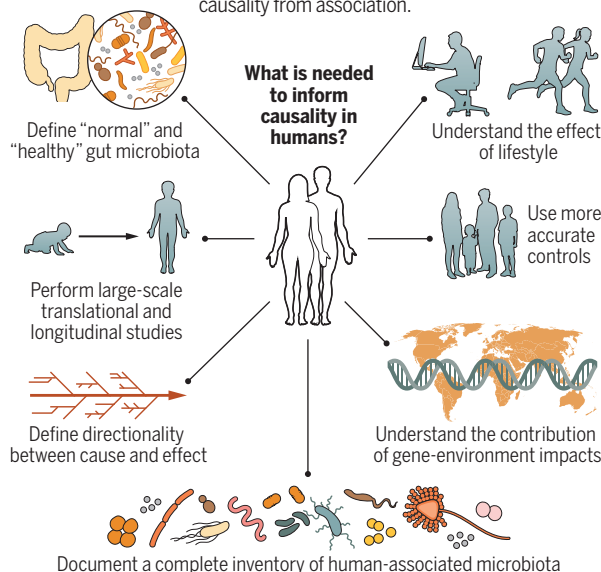
Animal studies offer valuable platforms and opportunities for hypothesis generation and testing. Preclinical models continue to evolve and improve in their translational value; for example, the immune system, liver, and microbiota of mice can be "humanized." However, it remains challenging to capture human genetic diversity and varied environmental influences in animal models, and correlating functional and structural similarities between the mouse brain and human brain has limitations. Similarly, cross-species behavioral repertoires vary considerably, hampering truly translational interpretations. Animal research has, however, convincingly demonstrated that the gut microbiota can be a consequence of the disease state, a contributor along with genetic or environmental risk, or a cause of changes to brain function, behavior, brain pathology, and disease-related symptoms. Indeed, dietary interventions or probiotic treatments that target the microbiota can ameliorate symptoms in some animal models of neurodevelopmental and neurodegenerative diseases (5–7).

Despite these breakthroughs, human studies have lagged behind animal research. A small number of brain imaging studies in infants or adults comparing gut microbiome profiles to brain activity scans and naturalistic behaviors have yielded intriguing associations (1). Cross-sectional studies have shown dramatic differences in the gut microbiomes and metabolomes of patients with major depression, ASD, and PD compared with matched controls. Although concordance between individual studies at the bacterial species level are often limited, larger meta-analyses that aggregate several datasets are revealing previously unrealized trends within specific disease indications. In some cases, microbiome profiles can define disease subtypes and correlate with use of medications. The quality of human microbiome research is improving, and sample sizes are increasing, with important con-

founding factors such as diet being taken into account (*8*). Accordingly, exploratory clinical findings have shown that ingestion of fermented dairy products alters brain activity in regions associated with anxiety (*9*). Psychobiotics can improve mood and affect (emotional states) as well as anxiety that is associated with irritable bowel syndrome. Modulation of gut-derived metabolites by an oral drug, antibiotic treatments, or FMT appear to improve ASD-related behaviors and GI symptoms in various investigational and open-label trials (*10*, *11*). Although early microbiota-based interventions appear promising, replication and well-controlled studies are needed to move from correlative research to validation of disease modification.

Genome-wide association studies (GWASs) have used large, collaborative team approaches to increase sample size and uncover genetic risks for complex diseases. However, the microbiome is more diverse and dynamic than the human genome, features that present distinct challenges. The complexity of human microbiotas in people living in different locations and environments, its temporal variability shaped by lifestyle, and a still incomplete inventory of human-associated bacteria pose barriers to understanding its causal effects on human outcomes. Bacterial and viral community profiling of fecal DNA has generated evidence of microbiome changes in a variety of neurological disorders, but critical information that is required to fully interpret these results is still missing. For example, several psychiatric and neurodegenerative disorders are accompanied by intestinal problems; disentangling whether the microbiota contributes to gut symptoms or was reshaped by the pathology and/or lifestyle associated with GI manifestations will affect interpretation of observed microbiome differences. The relative lack of longitudinal studies does not inform whether microbiome changes precede a clinical diagnosis.

Most cohorts compare community controls to patients and do not capture the effects of shared environments and diets on the microbiota that can be assessed through analysis of household controls and/or siblings. Gathering genetic information, such as known risk alleles associated with neuropsychiatric or neurodegenerative disorders, would allow more refined analyses on the level of contribution by the gut microbiota to impacts of gene-environment interactions. Therefore, efforts to predict causality by the microbiota to outcomes in humans (even in limited contexts) requires approaches that capture multiple variables that can influence the microbiota's plasticity over time, confounds not encountered in simplified animal models.

A key aspect of microbiome-brain research is its cross-disciplinary nature, and varied expertise is needed to unravel complex effects on human neurologic or behavioral outcomes (see the figure). Large natural history studies that follow human populations in their native settings over time are ideal platforms for exploring the contributions of the microbiota to effects of interest because accessible fecal samples can be collected with accompanying lifestyle data, supplying not just longitudinal information about microbiome profiles but also real-time tracking of lived experiences that may influence its composition and function. Notably, these experiences include environmental changes as well as emotional stressors, physical stressors, circadian rhythms, injury, infection, and neurologic disease diagnosis (e.g., potential brain-to-body feedback). Similar to natural history studies, birth cohorts are powerful for studying certain conditions that manifest early in life, such as ASD. This approach enables microbiome profiling in high-risk populations and can start during pregnancy, with disease expected to be diagnosed 3 to 5 years after birth. Microbiome assessments in birth cohorts can be coupled with genetic testing, metabolic profiling, and diet and lifestyle information. Longitudinally correlating blood or urinary biomarkers, immune profiles, and microbial products with behavioral and neurobiological end points will yield insights into cellular and molecular mechanisms that may mediate gut-brain interactions.

Replication across cohorts will be critical but poses logistical and material challenges. Defining the directionality of cause and effect in large population studies would provide unprecedented opportunities for generating hypotheses that can be tested in humans and experimental systems. Also, the relative contribution of the microbiota to normal behavioral responses in the absence of a disease state requires validation—for example, exploring whether the microbiota plays a role in homeostatic interoceptive (internal) responses, determining how microbiota effects on immunity or metabolism shape everyday behaviors, and unraveling whether interindividual variance in cognitive, social, or stress responses is mediated by microbial influences.

Notably, it is conceivable that the influences of gut bacteria on human conditions are not causal, outside of certain types of infections. By setting homeostatic thresholds for the activity and function of the immune, metabolic, and endocrine systems, an "unhealthy" microbiota may not itself cause illness but rather make its host less resilient to the effects of stressors that include genetic risks, unhealthy diets or lifestyles, or encountered emotional or physical stresses. This idea of combinatorial etiologies, wherein the microbiota represents one component of a complex disease, is supported by seminal observations that mice that harbor specific genetic or environmental risks for ALS, ASD, and PD display varying symptoms depending on their microbiome composition (*12–15*). Accordingly, interventions to establish a "healthy" microbiota through diet or FMT, or correcting subtle microbial deviations with probiotics or prebiotics, may confer biological resilience that mitigates underlying genetic or other contributions of illnesses that affect the brain. Perhaps the view that animals have intimately coevolved with their microbiotas, which fundamentally interact widely and dynamically with the entire body, explains why so many outcomes and diseases have been attributed to gut bacterial function (or dysfunction). Robustness is achieved by overlapping and complementary functions of a complex microbiota, and depletion or changes in microbial diversity may erode resilience to a stressor, thereby increasing disease risk.

The prospect of microbiota-based interventional studies for proof-of-concept or treatment objectives is exciting to contemplate. There is already evidence that FMTs in adult humans are feasible, scalable, and safe (*12*). Successful interventions that target the microbiota, not the human, and improve altered behaviors or brain pathology would help validate that gut bacteria contribute to a neurologic condition. Additionally, the potential of delivering drugs with targets in the gut, rather than current approaches that require traversing the blood-brain barrier, may offer greater therapeutic tractability and safety. ∎

REFERENCES AND NOTES

1. J. F. Cryan *et al.*, *Physiol. Rev.* **99**, 1877 (2019).
2. B. D. Needham *et al.*, *Nature* **602**, 647 (2022).
3. H. E. Vuong, J. M. Yano, T. C. Fung, E. Y. Hsiao, *Annu. Rev. Neurosci.* **40**, 21 (2017).
4. M. P. O'Donnell, B. W. Fox, P.-H. Chao, F. C. Schroeder, P. Sengupta, *Nature* **583**, 415 (2020).
5. P. Fang, S. A. Kazmi, K. G. Jameson, E. Y. Hsiao, *Cell Host Microbe* **28**, 201 (2020).
6. E. Blacher *et al.*, *Nature* **572**, 474 (2019).
7. E. Y. Hsiao *et al.*, *Cell* **155**, 1451 (2013).
8. R. Gacesa *et al.*, *Nature* **604**, 732 (2022).
9. K. Tillisch *et al.*, *Gastroenterology* **144**, 1394 (2013).
10. A. Stewart Campbell *et al.*, *Nat. Med.* **28**, 528 (2022).
11. D.-W. Kang *et al.*, *Sci. Rep.* **9**, 5821 (2019).
12. C. R. Kelly *et al.*, *Gastroenterology* **160**, 183 (2021).
13. A. Burberry *et al.*, *Nature* **582**, 89 (2020).
14. T. R. Sampson *et al.*, *Cell* **167**, 1469 (2016).
15. D. Matheoud *et al.*, *Nature* **571**, 565 (2019).

REVIEW

Microbiota and maintenance of skin barrier function

Tamia A. Harris-Tryon[1] and Elizabeth A. Grice[2]

Human skin forms a protective barrier against the external environment and is our first line of defense against toxic, solar, and pathogenic insults. Our skin also defines our outward appearance, protects our internal tissues and organs, acts as a sensory interface, and prevents dehydration. Crucial to the skin's barrier function is the colonizing microbiota, which provides protection against pathogens, tunes immune responses, and fortifies the epithelium. Here we highlight recent advances in our understanding of how the microbiota mediates multiple facets of skin barrier function. We discuss recent insights into pathological host–microbiota interactions and implications for disorders of the skin and distant organs. Finally, we examine how microbiota-based mechanisms can be targeted to prevent or manage skin disorders and impaired wound healing.

Humans live in partnership with their microbiota, complex communities of bacteria, fungi, and viruses that inhabit the body's surfaces. These relationships have been forged and challenged over millions of years of coevolution. Thus, it is unsurprising that our microbial cohabitants are major participants in shaping and maintaining essential physiological processes. As our outermost barrier against the external environment, the skin is colonized by a distinctive commensal microbiota that stimulates and educates defense and immune responses, contributes to proper differentiation and epithelialization, and even provides direct defense against pathogenic microorganisms.

In this Review, we examine our current understanding of the skin's microbial ecology and highlight recent insights into the microbiota's role in shaping and fortifying the barrier function of the skin. We also consider pathological microbe–host interactions and their role in skin disease and disruptions to other organ systems. Finally, we consider how these interactions could be leveraged to prevent or treat skin disease and impaired wound healing.

Microbial ecology of the human skin

Human skin, with its hypersaline and acidic environment and low nutrient availability, is distinct from other mucosa and epithelia. Both culture-based approaches and metagenomic profiling strategies of increasing resolution demonstrate that the human skin microbiota comprises a restricted set of bacterial, fungal, and viral inhabitants (1–3). Within the skin, bacteria predominantly belong to three phyla, Actinobacteria, Firmicutes, and Proteobacteria, with associated bacteriophages further modulating bacterial community dynamics and virulence. Eukaryotic viruses are also present,

although generally in lower numbers. The skin also houses eukaryotic organisms, but these are less abundant than bacteria. *Malassezia* species, for example, predominate among fungal communities throughout adult human skin (Fig. 1), and the eight-legged arachnid genus *Demodex* resides within the hair follicle. Initial colonization of neonatal skin is suspected to occur during delivery. In infants born vaginally, *Lactobacillus*, *Prevotella*, or *Sneathia* species are transferred to the skin during passage through the cervix and vagina. These species disappear by 6 weeks of age, when the microbiota begins to develop a more skin-like profile enriched with *Staphylococcus* and *Corynebacterium* species. In infants born by cesarian section, *Staphylococcus*, *Corynebacterium*, and *Cutibacterium* species predominate, without a preceding vaginal signature (4).

One of the major drivers of the skin's microbial ecology is the pilosebaceous unit, a skin

appendage that contains the hair follicle and its associated sebaceous gland. Sebaceous glands excrete a waxy, oily substance called sebum that emolliates the skin and selects for microbial species that metabolize the nutrients it contains. The pilosebaceous unit is also hypoxic, providing an ideal environment for the facultative anaerobe *Cutibacterium acnes*, which dominates this niche. Other signature bacteria of the skin surface include the coagulase-negative staphylococci (CoNS) species, such as *Staphylococcus epidermidis*, which are equipped to adhere and persist on human skin and tolerate the conditions. Skin sites with higher moisture and occlusion (e.g., the groin, axilla, and umbilicus) are enriched by *Corynebacterium* species that can be lipid-dependent and slow-growing in culture (Fig. 1).

The skin microbiota experiences a major shift at puberty when sex hormones drive maturation of the sebaceous gland and initiate sebum production. The introduction of lipid-rich sebum drives expansion of lipophilic *C. acnes* and *Malassezia* spp. on the skin surface, which correlates with serum sex hormone concentrations (5). Outside of puberty, the strains of bacteria and fungi colonizing the skin remain relatively stable within an individual over time (2). This stability is remarkable given the continuous disturbances imposed by lifestyle, the environment, and other host-specific factors, including constant shedding of terminally differentiated keratinocytes (squames) and secretions of sweat and sebum (Fig. 1). *C. acnes* and the pilosebaceous unit are likely major stabilizers of this effect in the human skin microbiome. The spatial architecture and hypoxic

[1]Department of Dermatology, University of Texas Southwestern Medical Center, Dallas, TX, USA. [2]Departments of Dermatology and Microbiology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA.
Email: tamia.harris-tryon@utsouthwestern.edu (T.A.H.-T.); egrice@pennmedicine.upenn.edu (E.A.G.)
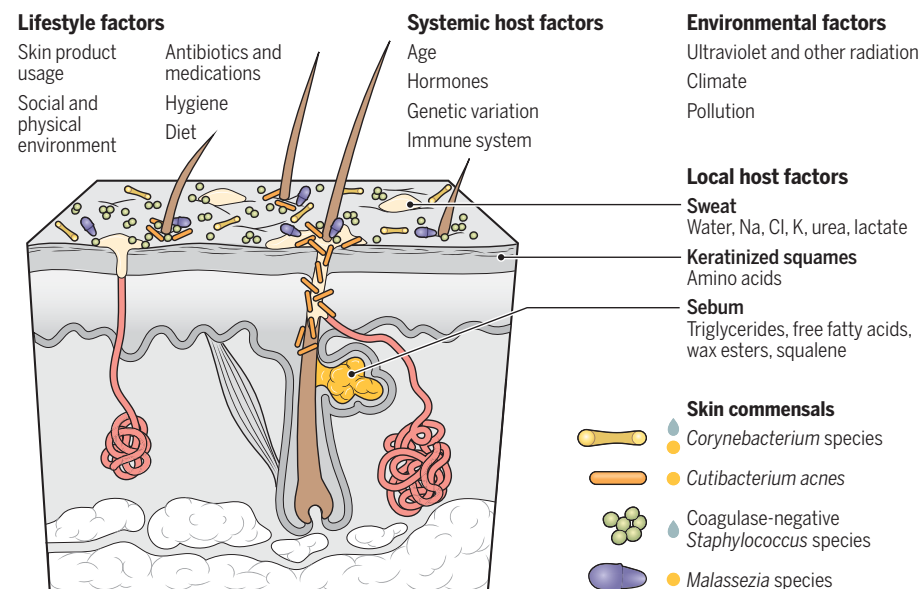
Fig. 1. Factors that influence the microbial colonization of skin. Local, systemic, environmental, and lifestyle factors together contribute to colonization and stability of the skin microbiome. Pictured are common human skin commensals. Hydrophilic and lipophilic microbes are indicated with a water droplet (blue) and oil droplet (yellow), respectively.

conditions of the pilosebaceous unit impose a bottleneck on *C. acnes*, allowing early colonizing strains of *C. acnes* to predominate a given follicle with limited competition (*6*). Although pulse disturbance experiments with topical antiseptics displace sensitive lower-abundance taxa, *C. acnes* remains during community recovery (*7*). By contrast, longer disruptions to the microbiota through the use of systemic antibiotics impose long-lasting effects at the community level and drive selection for antibiotic-resistant staphylococci. Systemic antibiotics also increase gene mobilization among the microbiome, which is indicative of a stress response (*8*).

Microbes fortify multiple facets of the skin barrier

The skin is a formidable structure composed of a stratified, cornified epithelium of keratinocytes, which undergo terminal differentiation. These physical structures are further fortified by chemical and immunological features that enhance the barrier. The skin microbiota affects all aspects of the skin barrier, while also directly interacting with commensal and pathogenic microbes encountered at the surface (Fig. 2). We next discuss how microbes interact with the skin barrier's microbial, chemical, and innate and adaptive immune components.

Microbial barrier

The skin microbiota itself is a barrier against invasion, colonization, and infection by foreign and pathogenic microbes. Living in polymicrobial communities, skin microbes vie for resources and have evolved mechanisms to directly antagonize their rivals. Multiple CoNS species, such as *Staphylococcus hominis*, produce antibiotics with unique chemistry and potent inhibitory activity against the major skin pathogen *Staphylococcus aureus* (*9*). Other species such as *Staphylococcus capitis* antagonize *S. aureus* through interference with the accessory gene regulator (agr) quorum sensing pathways, which are required for *S. aureus* virulence (*10, 11*). Notably, many of these antagonistic mechanisms synergize with host antimicrobial responses. For example, lugdunin, a peptide antibiotic produced by *Staphylococcus lugdunensis*, induces keratinocytes to produce the antimicrobial peptide LL-37 and neutrophil chemoattractant CXCL8 through the Toll-like receptor–myeloid differentiation primary response protein 88 (TLR–MyD88) pathway (*12*). Competitive mechanisms are not
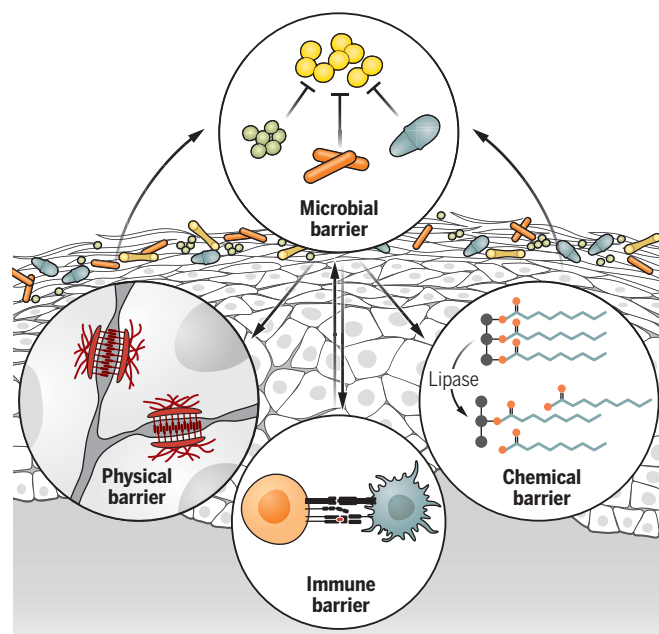
Fig. 2. The skin microbiota mediates multiple levels of barrier function. Skin microbes form the first barrier against the environment through various mechanisms of colonization resistance, including resource exclusion, direct inhibition, and/or interference. The skin microbiota also contributes to the differentiation and epithelialization of the physical skin barrier. Microbes boost the chemical barrier of the skin by producing lipases that digest sebum triglycerides to free fatty acids, which amplify the acidity of skin and restrict colonization by transient and pathogenic species. Finally, microbes stimulate innate and adaptive immune defenses such as release of antimicrobial peptides, induction of neonatal tolerance, and development of protective immunity.

limited to CoNS species in the skin microbiota. *C. acnes* competes to maintain its niche in the human pilosebaceous unit, with specific strains producing a thiopeptide antibiotic, cutimycin, that limits *S. aureus* colonization (*13*). How these individual interactions coalesce in a community setting and how this affects the community structure and function remain unclear.

Physical barrier

Keratinocytes undergo a program of tightly regulated terminal differentiation to form the stratum corneum, a process that also can be mediated by the microbiota. Self-renewing basal keratinocytes exit the cell cycle and acquire the machinery (e.g., intermediate filaments and lipid granules) that together form the "bricks and mortar" of the permeability barrier. This barrier also directly interfaces with the microbiota, resident or transient, and is subject to microbial regulation. The microbiota is required for normal skin barrier structure and function in mice and promotes differentiation and epithelial integrity through signaling of the keratinocyte aryl hydrocarbon receptor (AHR) (*14*). Skin bacteria also secrete sphingomyelinases that process lamellar lipids into ceramides (*15*), a critical component of the stratum corneum.

Chemical barrier

In addition to the physical distance from the environment provided by the corneocytes, keratinocytes, and skin lipids, the acidic skin surface creates a chemical environment that restricts bacterial colonization. Both *C. acnes* and *Corynebacterium* spp. secrete lipases that hydrolyze free fatty acids from triglycerides in sebum (*16, 17*). Free fatty acids further augment skin immunity by directly inhibiting bacteria and by stimulating the expression of human β-defensin 2 (hBD-2) (*18*). *C. acnes* also binds directly to free fatty acids, suggesting that the availability of free fatty acids facilitates the colonization of *C. acnes*.

Innate immune barrier

The microbiota is intimately associated with the skin epithelium, and the host and microbe have the capacity for cross-talk. Microbes can stimulate a range of innate immune responses that often depend on the metabolic and inflammatory contexts. For example, filamentous and yeast forms of *Candida albicans* stimulate distinct immune responses in the skin (*19*). Similarly, the T cell response to *S. epidermidis* in skin requires the expression of specific glycans on the bacterial surface that interact with C-type lectins on host innate immune cells (*20*). Oxygen availability can also affect host–microbial interactions at the skin surface. The microaerophilic bacterium *C. acnes* ferments the glycerol backbone of triglyceride and generates short-chain fatty acids (SCFAs). In turn, SCFAs inhibit histone deacetylases (HDACs), which can act as epigenetic regulators of the immune system (*21*). Unlike in the gastrointestinal tract, where SCFAs have anti-inflammatory effects on gut immunity, SCFAs have proinflammatory effects in the skin. By means of keratinocytes, *C. acnes*–derived SCFAs inhibit HDAC8 and HDAC9 and stimulate inflammation through TLR signaling (*22*). In the sebaceous gland, SCFAs derived from *C. acnes* fermentation augment inflammation through the activation of the free fatty acid receptor (*23*). Thus, a microbe's metabolic and inflammatory context can result in distinct types of immune responses.

Skin microbes further bolster skin immunity by stimulating the production of host-derived antimicrobial peptides and proteins (AMPs), which act as natural antibiotics. The expression of the AMP LL-37, a fragment of the protein cathelicidin, increases in response to activation of TLR signaling initiated by microbial signals (*24*). In addition to the cathelicidin family of AMPs, the skin also generates members of the

ILLUSTRATION: KELLIE HOLOSKI/*SCIENCE*

β-defensin family with bactericidal action against *Escherichia coli* and *S. aureus* strains. The sebaceous gland responds to Gram-negative lipopolysaccharide by generating the small proline-rich proteins SPRR1 and SPRR2, which directly disrupt negatively charged bacterial membranes (*25*). The skin also produces numerous cationic intrinsically disordered proteins with broad antimicrobial activity (*26*). These AMPs act in concert to provide the skin with a range of antimicrobial defenses against the microbes encountered in the environment.

The skin microbiota also helps coordinate innate immune responses during wound repair. Similar to observations made in the lung and gut, the commensal microbiota in the skin elicits a type I interferon (IFN) response during this process (*27*). In response to microbial stimuli, neutrophils express CXCL10, which recruits activated plasmacytoid dendritic cells (pDCs) to sites of injury. pDCs generate type I IFNs, which accelerate wound repair through stimulation of fibroblast and macrophage growth factor responses. Indeed, the recruitment of antigen-presenting cells to the skin is microbiota-dependent (*28*). Similarly, microbes enhance skin regeneration in wound repair and hair follicle neogenesis through a process that requires interleukin-1 receptor (IL-1R)–MYD88 signaling (*29*).

Adaptive immune barrier

The skin is home to a diverse repertoire of adaptive immune cells, among them vast pools of resident memory T cells poised to respond to various environmental stimuli, including pathogenic and commensal microbes. In early infancy, exposure to the skin commensal *S. epidermidis* mediates the influx of regulatory T cells (T$_{regs}$) into the skin (*30*). This wave of T$_{reg}$ migration occurs concurrently with hair follicle development and requires the production of chemokines generated by the hair follicle keratinocytes (*30, 31*). T$_{regs}$, along with many other immune cell subsets in the skin, ultimately reside adjacent to the hair follicle, with specificity to the microbial antigens detected during this developmental window.

In a parallel process, mucosal-associated invariant T (MAIT) cells are acquired in infancy during a similar time-restricted developmental window. MAIT cells are absent in germ-free mice, and their development requires vitamin B2 metabolites that are only produced by bacteria and fungi, not mammalian cells. In the thymus, exposure to 5-(2-oxopropylideneamino)-6-D-ribitylaminouracil, a bacterial metabolite of vitamin B2 trafficked to the thymus from mucosal sites, mediates MAIT cell expansion and targeting to the skin and mucosal sites (*32, 33*). Microbial cell surface molecules can also act as signals to the host. Most species of *Corynebacterium* contain mycolic acid in their cell envelope. Mycolic acid from *Corynebacterium* species can promote γδ T cell accumulation in an IL-23–dependent manner under steady state.

However, this interaction is context-dependent, as a high-fat diet instead promotes cutaneous inflammation (*17*). Thus, the inflammatory milieu present at the time of microbial exposure affects the immune response within the skin. Taken together, these findings highlight the key role that microbes play in the recruitment and stimulation of immune cells in the skin.

Pathological microbial–host interactions and skin disorders

From an ecological standpoint, microbial communities are inevitably destined to change in structure and function when their niche is disrupted. As such, an altered skin microbiome is more often the rule than the exception in skin disease. Shifts in resource availability and, in some cases, complete devastation of their habitat are factors that drive the depletion of normal skin residents in favor of opportunists. Owing to the tight interconnectivity of the microbiota with its host, it is difficult to distinguish between "the chicken and the egg" in the absence of experimental approaches that rely on cultured isolates in experimental model systems. Whether causative or a consequence, altered microbial communities can mediate tissue damage and/or inflammation across a variety of skin disorders.

S. aureus is a frequent opportunist of the skin and overwhelms the commensal microbiota in barrier disorders such as atopic dermatitis (AD) and skin wounds (*34–36*). In the setting of disease, *S. aureus* can evade the host immune response to establish chronic infection. Moreover, *S. aureus* and some CoNS species produce proteases and other factors that further damage the barrier and drive pathological inflammatory responses (*11, 37*). In addition to direct damage, *S. aureus* interferes with adaptive immune responses by producing alpha toxins that trigger IL-1R–mediated inflammation and prevent the accumulation of *S. aureus*–specific T$_{regs}$ and the development of tolerance to *S. aureus* later in life (*38*).

In AD and other dysbiotic contexts where an opportunist overtakes the ecosystem, there is a depletion of the commensal microbes and their mediators that previously supported the skin's barrier defenses. For example, tryptophan metabolites are reduced in AD skin. When these metabolites are therapeutically administered, inflammation in mouse models of AD is attenuated by AHR (*39*). Coal tar, one of the oldest therapies for AD, can activate the AHR in the skin to drive differentiation programs, AMP expression, and normalization of the microbiome (*40*). Dysregulation of tryptophan catabolism by the microbiota may also contribute to hidradenitis suppurativa (HS), a condition characterized clinically by festering wounds of the armpits and groin whose pathogenesis is poorly understood. HS lesions are also deficient in AHR activation, which coincides with a depletion of tryptophan-metabolizing micro-

biota (*41*). Thus, microbial metabolites produced by skin commensals are depleted during disease states, which may maintain and exacerbate inflammation and barrier disruption.

A dysregulated or dysfunctional immune system also has impacts on the skin microbiota, which can further exacerbate disease. *S. aureus* and *S. epidermidis*, for example, are more abundant and cause greater amounts of skin damage in Netherton syndrome patients, who have a genetic defect in the skin protease inhibitor lymphoepithelial Kazal-type–related protease inhibitor 1 (LEKTI-1) (*42*). Skin infections and shifts in skin colonization can also occur in humans with primary immunodeficiency disorders. For example, in patients with dedicator of cytokinesis protein 8 (DOCK8) deficiency, the cutaneous virome is enriched with a diversity of eukaryotic viruses, including human papilloma viruses (*43*). Genetic deletion studies in murine models further highlight the key role that the immune system plays in restricting the microbiota. Mice devoid of type 2 innate lymphoid cells (ILC2s) have enlarged sebaceous glands and generate greater amounts of antimicrobial lipids that restrict colonization of Gram-positive commensals (*44*). By contrast, mice lacking T cells and epidermal expression of the transcription factor JunB are unable to control *S. aureus* inflammation at the skin surface and recapitulate several aspects of atopic inflammation (*45*).

As shotgun metagenomics and culture-based investigations have advanced, it is becoming clear that strain-level variations of human skin commensals and pathogens also have an impact on disease pathogenesis. Specific strains of *S. aureus* correlate with disease severity and clinical outcomes in different contexts (*34, 36*). Acneic skin is colonized with *C. acnes* strains that inherently produce greater amounts of the proinflammatory metabolite porphyrin compared with *C. acnes* strains recovered from healthy skin (*46*). Moreover, porphyrin production by *C. acnes* is under the control of vitamin B12. A vitamin B12 supplementation study in humans showed increased porphyrin production leading to acne development (*47*). This is a potential molecular mechanism that may explain how the same species of bacteria can both cause disease and reside as a member of the healthy skin microbiota. A challenge going forward will be to identify specific markers of virulence within bacterial strains that can inform management strategies for problematic microbial burdens.

Another advantage of culture-independent approaches is that they have greatly facilitated the identification of fastidious anaerobic microbiota in skin disease and wounds. In chronic wounds and HS, mixed communities of Gram-positive anaerobic bacteria can inhabit deeper tissues of the skin (*48*). Although the skin commensal *C. acnes* may be found in some abundance in these mixed communities, other players,

such as *Anaerococcus* spp., *Porphyromonas* spp., *Finegoldia* spp., *Veillonella* spp., and *Peptostreptococcus* spp., are more common. In chronic wounds, the persistence of anaerobic communities after debridement is associated with poor wound outcomes (*36*, *49*). The challenges of isolation and study of anaerobes, especially in mixed communities, are limiting factors in advancing our understanding of their role in skin disorders such as HS and chronic wounds.

Psoriasis is another common inflammatory skin disease. In contrast to AD and other barrier disorders, alterations to the skin microbiota in psoriasis are more subtle, less consistent across studies, and are more weakly associated with disease (*35*, *50*, *51*). Thus, there is currently limited evidence that the skin microbiota drives psoriasis pathogenesis. However, in a psoriasis mouse model, *C. albicans* exposure augmented T helper 17 cell immunity with increased infiltration of and IL-17 production by αβ T cells sensitized by *Candida* (*52*). Microbial–host interactions at other mucosal sites are hypothesized to contribute to psoriasis as well, as discussed in the next section.

Systemic roles for the skin microbiome

There is increasing evidence that skin damage and sensitization can affect other barrier sites, such as the intestine and the lung (Fig. 3). For example, superficial skin damage causes keratinocytes to release IL-33 systemically. In synergy with IL-25, IL-33 triggers the activation of ILC2s in the intestine to generate IL-4. This, in turn, stimulates the expansion of mast cells in the intestine, where they are poised to respond to food allergens and mediate anaphylaxis (*53*). Wounding of the skin also augments intestinal inflammation in dextran sodium sulfate–induced colitis mouse models, which mimic inflammatory bowel disease. Cross-talk between the skin and gut depends on the production of hyaluronan fragments generated in the dermis during injury that stimulate intestinal fibroblasts to differentiate into proinflammatory adipocytes through a process called reactive adipogenesis. These reactive adipocytes propagate gut inflammation through the production of AMPs and other inflammatory mediators (*54*).

Skin sensitization also affects the lungs. Epidemiological evidence demonstrates that many patients progress through an "atopic march," first presenting with the skin barrier condition AD and subsequently developing allergic rhinitis, food allergies, and asthma (*55*). Epicutaneous exposure to *S. aureus* stimulates keratinocytes to produce IL-36, which amplifies serum immunoglobulin E (IgE) levels. Mice lacking the IL-36 receptor do not develop elevated IgE in response to *S. aureus* and are also protected from allergen-specific lung inflammation (*56*). These findings provide evidence for skin exposure to microbial pathogens as an initiating event in systemic inflammation. However, it is notable that the skin has the capacity to control and restrict commensal responses independently of other mucosal sites through a sophisticated network of immune strategies (*57*). More work is needed to uncover the many ways that these regulatory mechanisms, which contribute to sustained compartmentalization, malfunction in disease.

The gut microbiome can also affect skin inflammation. For example, type 3 inflammation in a mouse psoriasis model is dampened in germ-free mice, which lack a microbiome (*58*). Moreover, mice that are sensitized to allergens in the intestine through oral administration develop antigen-specific T cells in the skin after epicutaneous challenge with the same antigen (*59*). In both cases, activation of the intestinal immune networks affects the amplitude of the inflammatory signals in the skin. Thus, alterations in the gut microbiome may affect skin immunity, although clear targets for therapeutic avenues to influence skin disease through modulation of the intestinal microbes remain undefined. What has been shown is that the dietary impacts on the gut microbiome, especially dietary fiber, have meaningful effects on systemic immunity (*60*). Cutaneous innate immune responses are also linked to the gut, where adequate expression of AMPs that protect against bacterial skin infection is dependent on dietary vitamin A (*61*). Together, these findings strengthen our molecular understanding of the importance of diet in the development of host immunity.

An emerging area of investigation is the interface between skin microbiota and the neuroimmune axis. Bacteria can directly activate sensory neurons in the skin and cause pain through the production of pore-forming toxins (*62*). As in interactions with other aspects of the host, variation at the strain level drives variable responses, depending on the presence of specialized toxins and quorum sensing systems. Sensory neurons in the skin are also directly activated by the fungal pathogen *C. albicans*, and stimulation is required for γδ T cell immunity to control cutaneous candidiasis through release of neuropeptide CGRP (*63*). By contrast, the pathogen *Streptococcus pyogenes*, which causes necrotizing fasciitis, directly activates nociceptor neurons by secreting streptolysin S, which in turn promotes neuropeptide CGRP release and inhibits killing of *S. pyogenes*. In this context, CGRP antagonism prevents necrotizing infection (*64*). Although these studies have focused on skin pathogens in neuroimmune interactions, how skin commensals, at the community level, contribute to our sensory perceptions under homeostatic conditions remains under investigation.

Outlook and conclusions

The application of molecular, culture-independent techniques to survey microbial communities has reinvigorated the study of skin microbiota and its role in dermatological health
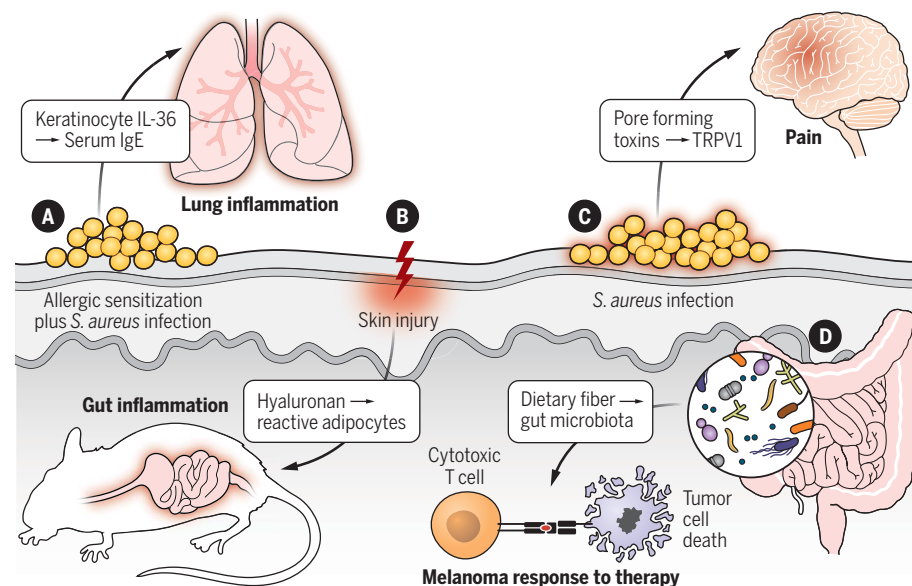
Fig. 3. Skin cross-talk with other organ systems is mediated by the microbiota. Emerging evidence highlights the role of skin cross-talk with distant organ systems, which is driven by host–microbiota interactions. Depicted are four examples of cross-talk between skin and other organ systems. (**A**) Allergic sensitization of skin, together with *S. aureus* infection, results in IL-36–dependent lung inflammation, suggesting a potential mechanism for the "atopic march." (**B**) Skin injury releases hyaluronan fragments systemically, which drives reactive adipogenesis, gut inflammation, and dysbiosis in murine models. (**C**) During infection, *S. aureus* releases pore-forming toxins that are implicated in directly activating nociceptors and causing pain. TRPV1, transient receptor potential cation channel subfamily V member 1. (**D**) The gut microbiome and dietary fiber contribute to melanoma response to immune checkpoint therapy, driving cytotoxic T cell accumulation and killing of tumor cells.
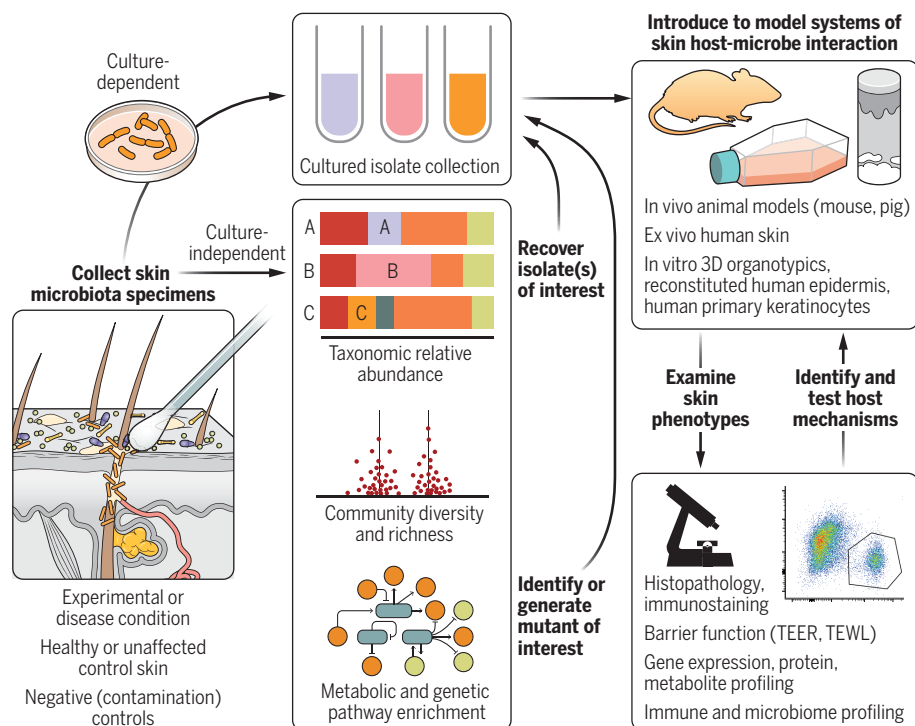
Fig. 4. Framework for examining functional and mechanistic attributes of the skin microbiota.
Illustrated is an example pipeline for using skin microbiome profiles to inform functional and mechanistic investigations. Applying a combination of culture-independent and culture-dependent approaches allows examination of specific microbes or consortia in the context of skin model systems. Furthermore, shotgun metagenomic sequencing can inform which metabolic or genetic pathways to target in the microbiota, which then can be functionally interrogated by examining mutants for these pathways. Deep phenotyping of skin or skin models colonized or associated with clinical isolates, mutants, and/or consortia facilitates hypothesis generation regarding host interactive pathways, which then can be tested in this pipeline. TEER, transepithelial electrical resistance; TEWL, transepidermal water loss.

and disease. Interpreting results from skin microbiota surveys, such as 16S rRNA gene sequencing or shotgun metagenomics, and using these data to guide a deeper understanding of functional and mechanistic attributes of the microbiota remains a challenge for the field. We posit that application of both culture-independent and culture-dependent approaches, within an ecological framework of community-wide interactions, can advance our understanding of homeostatic and pathological mechanisms at the host–microbiota interface (Fig. 4). Skin microbiome surveys can reveal the composition and diversity of the community and be combined with culture-based approaches to identify isolates or target the isolation of community members. These isolates can then be used individually or in combination to test their effects on the skin in vivo with model organisms, in vitro using human skin cells and constructs, or ex vivo using human skin explants. In this framework, shotgun metagenomics of ample read depths can be a useful approach for strain-level characterization of the community or can identify functional genetic pathways that are enriched within a sample. Genetic deletion of these pathways or genes can then be targeted in clinical isolates or obtained from collections of laboratory isolates, to test the role of these pathways in driving skin phenotypes. Deep-skin phenotyping, for example, using unbiased assays to measure gene expression can further define how the host responds to the skin microbiota.

Therapeutic advancement based on the skin microbiome will rely on such approaches to identify candidates for either enhancement or depletion in the community. Bacteriotherapy, or transplantation of live, defined bacteria, is currently under development for the treatment of AD using a strain of *S. hominis* that was isolated from healthy human skin and inhibits *S. aureus* (*9*). Phase 1 trials in *S. aureus*–positive AD patients (*n* = 54) indicate the safety of *S. hominis A9* and demonstrate a reduction of *S. aureus* colonization, although overall clinical severity of disease was not significantly affected (*65*). Moreover, a placebo-controlled trial using lysates of the Gram-negative *Vitreoscilla filiformis* has proven beneficial through the stimulation of IL-10–producing dendritic cells within the skin (*66*). Screening strategies may also be tailored to identify microbes that activate or repress host pathways of interest. The "Flowers' Flora" consortium, for example, was

developed by screening human skin commensals for AHR activation in keratinocytes (*14*). Colonization with this consortium improved barrier function in germ-free mice and reduced disease severity in murine models of AD. The Gram-negative *Roseomonas mucosa*, isolated from healthy human skin, was once explored for the treatment of AD, but clinical trials were ultimately discontinued owing to failure to meet end points (*67*). Other microbiota-based therapeutic approaches are less developed in the skin but could include phage-directed therapies to target pathogens, engineering commensals to express molecules of benefit, and/or prebiotic approaches to modify the habitat and thus the microbiota (*68*).

Although microbes have tremendous therapeutic potential, an ecological perspective of community-level interactions between host and microbes is needed to inform efforts to manipulate the microbiome. Selection of a consortium with desired functional attributes is only the first step, as major obstacles to the delivery and stable engraftment of transplanted communities remain. The host and the endogenous microbiota have powerful effects on the establishment, persistence, growth, and long-term impact of a transplanted community and are likely factors that influence the engraftment of a transplant (*69*). Spatial architecture of the pilosebaceous unit likely limits the complete removal of skin microbes, even with topical treatments meant to sterilize skin. These protected structures may serve as reservoirs to "reseed" the skin microbiome following disturbance. Additionally, it is now apparent that skin microbes are highly specialized to their niches, reflecting millions of years of adaptation to human skin, and not only interact with the local tissue microenvironment but drive signals at distant organs as well. Unsurprisingly, disrupting microbe–host relationships in the skin has consequences on organ structure and function. Introducing a new member to the community undoubtedly triggers responses from both host and microbial cells. Understanding how skin microbial communities interact with the host and each other is crucial to inform transplantation strategies and all types of microbial-based therapeutics that target this interface for the prevention and treatment of skin disorders.

REFERENCES AND NOTES
1. G. D. Hannigan *et al.*, *mBio* **6**, e01578-15 (2015).
2. J. Oh *et al.*, Temporal stability of the human skin microbiome. *Cell* **165**, 854–866 (2016).
3. S. Saheb Kashaf *et al.*, *Nat. Microbiol.* **7**, 169–179 (2022).
4. D. M. Chu *et al.*, *Nat. Med.* **23**, 314–326 (2017).
5. J. Park *et al.*, *J. Invest. Dermatol.* **142**, 212–219 (2022).
6. A. Conwill *et al.*, *Cell Host Microbe* **30**, 171–182.e7 (2022).
7. A. J. SanMiguel *et al.*, *J. Invest. Dermatol.* **138**, 2234–2243 (2018).
8. J.-H. Jo *et al.*, *Sci. Transl. Med.* **13**, eabd8077 (2021).
9. T. Nakatsuji *et al.*, *Sci. Transl. Med.* **9**, eaah4680 (2017).
10. A. E. Paharik *et al.*, *Cell Host Microbe* **22**, 746–756.e5 (2017).
11. M. R. Williams *et al.*, *Sci. Transl. Med.* **11**, eaat8329 (2019).
12. K. Bitschar *et al.*, *Nat. Commun.* **10**, 2730 (2019).
13. J. Claesen *et al.*, *Sci. Transl. Med.* **12**, eaay5445 (2020).

14. A. Uberoi *et al.*, *Cell Host Microbe* **29**, 1235–1248.e8 (2021).
15. Y. Zheng *et al.*, *Cell Host Microbe* **30**, 301–313.e9 (2022).
16. L. Bomar, S. D. Brugger, B. H. Yost, S. S. Davies, K. P. Lemon, *mBio* **7**, e01725–e15 (2016).
17. V. K. Ridaura *et al.*, *J. Exp. Med.* **215**, 785–799 (2018).
18. T. Nakatsuji *et al.*, *J. Invest. Dermatol.* **130**, 985–994 (2010).
19. S. W. Kashem *et al.*, *Immunity* **42**, 356–366 (2015).
20. Y. E. Chen *et al.*, bioRxiv 664656 [Preprint] (2019). https://doi.org/10.1101/664656.
21. P. V. Chang, L. Hao, S. Offermanns, R. Medzhitov, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 2247–2252 (2014).
22. J. A. Sanford *et al.*, *Sci. Immunol.* **1**, eaah4609 (2016).
23. J. A. Sanford, A. M. O'Neill, C. C. Zouboulis, R. L. Gallo, *J. Immunol.* **202**, 1767–1776 (2019).
24. Y. Lai *et al.*, *J. Invest. Dermatol.* **130**, 2211–2221 (2010).
25. C. Zhang *et al.*, *eLife* **11**, e76729 (2022).
26. T. Latendorf *et al.*, *Sci. Rep.* **9**, 3331 (2019).
27. J. Di Domizio *et al.*, *Nat. Immunol.* **21**, 1034–1045 (2020).
28. N. D. Ubags *et al.*, *J. Allergy Clin. Immunol.* **147**, 1049–1062.e7 (2021).
29. G. Wang *et al.*, *Cell Host Microbe* **29**, 777–791.e6 (2021).
30. T. C. Scharschmidt *et al.*, *Immunity* **43**, 1011–1021 (2015).
31. T. C. Scharschmidt *et al.*, *Cell Host Microbe* **21**, 467–477.e5 (2017).
32. M. G. Constantinides *et al.*, *Science* **366**, eaax6624 (2019).
33. F. Legoux *et al.*, *Science* **366**, 494–499 (2019).
34. A. L. Byrd *et al.*, *Sci. Transl. Med.* **9**, eaal4651 (2017).
35. N. Fyhrquist *et al.*, *Nat. Commun.* **10**, 4703 (2019).
36. L. R. Kalan *et al.*, *Cell Host Microbe* **25**, 641–655.e5 (2019).
37. L. Cau *et al.*, *J. Allergy Clin. Immunol.* **147**, 955–966.e16 (2021).
38. J. M. Leech *et al.*, *Cell Host Microbe* **26**, 795–809.e5 (2019).
39. J. Yu *et al.*, *J. Allergy Clin. Immunol.* **143**, 2108–2119.e12 (2019).
40. J. P. H. Smits *et al.*, *J. Invest. Dermatol.* **140**, 415–424.e10 (2020).
41. L. Guenin-Macé *et al.*, *JCI Insight* **5**, e140598 (2020).
42. M. R. Williams *et al.*, *Cell Rep.* **30**, 2923–2933.e7 (2020).
43. O. Tirosh *et al.*, *Nat. Med.* **24**, 1815–1821 (2018).
44. T. Kobayashi *et al.*, *Cell* **176**, 982–997.e16 (2019).
45. Ö. Uluçkan *et al.*, *Cell Rep.* **29**, 844–859.e3 (2019).
46. T. Johnson, D. Kang, E. Barnard, H. Li, *MSphere* **1**, e00023-15 (2016).
47. D. Kang, B. Shi, M. C. Erfe, N. Craft, H. Li, *Sci. Transl. Med.* **7**, 293ra103 (2015).
48. S. C. Williams, J. W. Frew, J. G. Krueger, *Exp. Dermatol.* **30**, 1388–1397 (2021).
49. S. Verbanic, Y. Shen, J. Lee, J. M. Deacon, I. A. Chen, *NPJ Biofilms Microbiomes* **6**, 21 (2020).
50. H.-W. Chang *et al.*, *Microbiome* **6**, 154 (2018).
51. M. A. Loesche *et al.*, *J. Invest. Dermatol.* **138**, 1973–1981 (2018).
52. C. Hurabielle *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 16465–16474 (2020).
53. J.-M. Leyva-Castillo *et al.*, *Immunity* **50**, 1262–1275.e4 (2019).
54. T. Dokoshi *et al.*, *J. Clin. Invest.* **131**, e147614 (2021).
55. A. S. Paller, J. M. Spergel, P. Mina-Osorio, A. D. Irvine, *J. Allergy Clin. Immunol.* **143**, 46–55 (2019).
56. G. J. Patrick *et al.*, *J. Clin. Invest.* **131**, e143334 (2021).
57. S. Naik *et al.*, *Science* **337**, 1115–1119 (2012).
58. Z. Zákostelská *et al.*, *PLOS ONE* **11**, e0159539 (2016).
59. M. K. Oyoshi *et al.*, *J. Clin. Invest.* **121**, 2210–2220 (2011).
60. C. N. Spencer *et al.*, *Science* **374**, 1632–1640 (2021).
61. T. A. Harris *et al.*, *Cell Host Microbe* **25**, 777–788.e8 (2019).
62. K. J. Blake *et al.*, *Nat. Commun.* **9**, 37 (2018).
63. S. W. Kashem *et al.*, *Immunity* **43**, 515–526 (2015).
64. F. A. Pinho-Ribeiro *et al.*, *Cell* **173**, 1083–1097.e22 (2018).
65. T. Nakatsuji *et al.*, *Nat. Med.* **27**, 700–709 (2021).
66. T. Volz *et al.*, *J. Invest. Dermatol.* **134**, 96–104 (2014).
67. I. A. Myles *et al.*, *Sci. Transl. Med.* **12**, eaaz8631 (2020).
68. M. T. Sorbara, E. G. Pamer, *Nat. Rev. Microbiol.* (2022).
69. J. Walter, M. X. Maldonado-Gómez, I. Martínez, *Curr. Opin. Biotechnol.* **49**, 129–139 (2018).

REVIEW

Immune-microbe interactions early in life: A determinant of health and disease long term

Petter Brodin[1,2,3]

Research on newborn immunity has revealed the importance of cell ontogeny, feto-maternal tolerance, and the transfer of maternal antibodies. Less is known about postnatal adaptation to environmental exposures. The microbiome and its importance for health have been extensively studied, but it remains unclear how mutually beneficial relationships between commensal microbes and human cells first arise and are maintained throughout life. Such immune-microbe mutualism, and perturbations thereof, is most likely a root cause of increasing incidences of immune-mediated disorders such as allergies and autoimmunity across many industrialized nations during the past century. In this Review, I discuss our current understanding of immune development and propose that mismatches among ancestral, early-life, and adult environments can explain perturbations to immune-microbe interactions, immune dysregulation, and increased risks of immune-mediated diseases.

Young children face a daunting task at birth and shortly thereafter, when substantial physiological adaptation must occur to initiate respiration, thermoregulation, and nutrition. At the same time, the child transitions from a sheltered environment in utero to the outside world, the most profound change in environmental exposures during the human lifetime. Facing this new world poses challenges for the immune system, which must protect the newborn from invasive pathogens while at the same time ensuring tolerance to beneficial microbes that are important for health. Many diseases that are increasingly prevalent in industrialized societies, such as asthma, allergies, autoimmune diseases, metabolic disorders, and neurodevelopmental alterations, have all been associated with perturbed early-life immune-microbe relationships. Immune-microbe mutualism, i.e., a mutually beneficial relationship between different species, must be established after birth because microbial colonization first happens during delivery and thereafter. To help her offspring, a mother provides passive immunity, antigen exposure in utero, selected microbes that seed and colonize the infant, and optimized, energy-dense nutrition in the form of breastmilk. Breastmilk also contains antibodies, antimicrobial peptides, and human milk oligosaccharides (HMOs) that further promote colonization by beneficial microbes and limit invasion by pathogens.

Life history theory offers a framework for explaining the long-term consequences of early-life events whereby different investments in or priorities on key traits such as reproduction, growth, and maintenance—of which immunity is a critical component—can have durable effects (1). Throughout evolution, the immune system has been shaped under strong selection pressure imposed by infectious agents, and this forms the basis for adaptive changes. Mismatches between ancestral and current environmental conditions must also be considered when trying to explain immune-mediated diseases that are influenced by the microbiota. Perturbations to evolutionary principles that govern immune system adaptation to environmental factors could provide a unifying explanation to recent changes in infant growth rates, obesity, immune-microbe dysregulation, and immune-mediated diseases such as atopy (development of allergies) and autoimmune diseases in industrialized societies. In this Review, I discuss conditions during prenatal, perinatal, and postnatal phases of development that are relevant to our understanding of the establishment of immune-microbe mutualism. I also discuss how a life history perspective and mismatches between ancestral and modern environments, as well as early-life perturbations, can have long-term consequences for health and disease.

Prenatal preparation

Balanced interactions between human cells and tissues and commensal microbes are important for metabolism, prevention of infection, maturation of lymphoid tissues, and functional development of immune cells. Several evolutionarily conserved mechanisms through which the mother prepares her offspring for postnatal life and colonization are also in operation. During pregnancy, the human immune system develops in a layered fashion (2) in which different populations of stem cells support immune cells poised toward feto-maternal tolerance early during fetal development and increasingly poised toward pathogen resistance in the later stages of pregnancy. Naïve T cells taken from mesenteric lymph nodes of fetuses in gestational weeks 18 to 22 are highly responsive to allogeneic cells and readily upregulate the transcription factor forkhead box P3 (FoxP3) and differentiate into regulatory T cells (Tregs) upon stimulation (3). Even before Tregs were described, higher frequencies of CD25-expressing memory phenotype CD4 T cells were noted in fetal versus adult spleen (4). Feto-maternal tolerance is further ensured by chimeric maternal cells that cross the placenta, further promoting the differentiation of fetal FoxP3+ Tregs (5) and seeding fetal organs where these maternal cells can persist into adulthood (6).

The exact timing of when fetal hematopoiesis shifts to adult hematopoiesis is variable among children and gradual in nature (7) but is influenced by prenatal exposures such as maternal stress and infection (Fig. 1A). In mice, a transient spike in type I interferons (IFNs) late in the prenatal period triggered the gradual switch toward adult hematopoiesis (8).

Although the trigger for this prenatal IFN-I spike is unknown [and it was observed even under germ-free conditions (8)], the fact that such cytokines are readily induced during infections and can cross the placenta suggests that prenatal infections in the mother and/or the fetus could initiate a switch from fetal to adult hematopoiesis and repurpose offspring immunity toward resistance to infections in preparation for microbial encounters (Fig. 1A). Another recent study in mice showed that mild intestinal infection of pregnant dams with the food-borne pathogen *Yersinia pseudotuberculosis* did not spread to the fetus but did trigger adaptive changes in fetal intestinal epithelial cells. This was mediated by maternal interleukin-6 (IL-6), which conferred increased resistance to intestinal infection later in life, but also predisposition to inflammatory responses in the intestine, with possible negative consequences for health (9).

Similarly, maternal stress in pregnant mice recalibrates the hypothalamic-pituitary-adrenal axis of fetuses and impairs CD8+ T cell responses in the offspring long after birth (10). Not only do pathogenic microbes induce adaptations in the fetus but so do symbiotic microbes. Symbiotic microbes transiently colonizing pregnant mice have been shown to prime fetal intestinal epithelia and recruit immune cells to ready the offspring for postnatal colonization (11). The priming of fetal immunity by maternal microbes in these mice was shown to be dependent on immunoglobulin G (IgG)–mediated transfer of microbial components across the placenta, allowing the dams to prime the fetus for exposure to the same microbes to which passive immunity, and consequently protection from invasive infection, is also provided (Fig. 1B) (11).

The presence of viable bacteria in utero has been a hotly debated topic, with conflicting results and different low-abundance strains

¹Department of Immunology and Inflammation, Imperial College London, London, UK. ²Imperial College Healthcare NHS Trust, London, UK. ³Department of Women's and Children's Health, Karolinska Institutet, Uppsala, Sweden.
*Email: p.brodin@imperial.ac.uk

reported in different studies. This is likely because of challenges in separating low-biomass bacteria in fetal intestines and amniotic fluid from environmental contaminants during tissue preparation (12, 13). Multiple studies that have controlled for environmental contaminants have failed to find evidence of fetal bacterial colonization in the absence of infection (14, 15). Conversely, in fetuses in which low-abundance bacteria have been reported, transcriptional adaptation in epithelial cells (13) and the presence of memory-phenotype T cells have been interpreted as evidence in favor of prenatal microbial colonization (16, 17). However, microbial components and metabolites transferred directly across the placenta, or in-

tion to avoid immunopathology and permit microbial colonization at birth and thereafter. Trained immunity is an additional mechanism by which fetal innate immune cells can be prepared for postnatal exposures by epigenetic adaptation (21). For example, monocytes of HIV-exposed but uninfected fetuses adapt and respond more vigorously to Toll-like receptor ligands after birth compared with HIV-unexposed controls (22). Similar adaptive changes have been reported in children born to mothers infected by malaria during pregnancy (23). Through these multiple mechanisms, fetal immune systems influenced by maternal exposures to be prepared for the challenging balancing act of permitting micro-

tial differences in microbial composition early after birth (27). Infants continue to receive microbes from their parents and other sources, but microbes seeded from the mother during vaginal labor and originating in the intestine are optimally adapted to survive in the infant gut and will predominate over time, illustrating the niche selection that occurs after birth (28). Breastfeeding promotes specific colonizing microbes in the newborn intestine, including Bifidobacteria, which are adapted to metabolize HMOs. Antibiotic exposures reduce overall microbial diversity and delay the normal development of the intestinal microbiome (29), and such early-life antibiotic exposure is associated with increased risks of developing allergy (30), asthma (31), type 1 diabetes (32), and obesity (33) later in childhood. Mechanisms underlying these altered developmental trajectories are discussed below.

A Fetal immune system and layered hematopoiesis
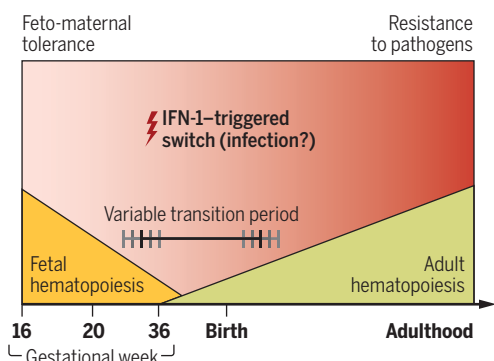
B Transplacental transfer of IgG and microbial components
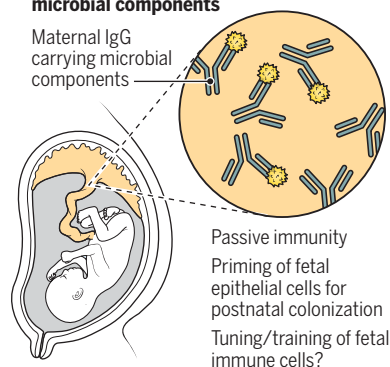
Fig. 1. Prenatal preparation. (A) Layered hematopoiesis occurs with different stem cell populations that give rise to cells poised toward feto-maternal tolerance early during pregnancy but gradually switch over toward pathogen resistance at the end of gestation in preparation for microbial colonization. The transition period is gradual in nature (7), and one trigger for this switch from fetal to adult hematopoiesis is a surge in IFN-I before birth (8). (B) Transplacental transfer of maternal IgG, and possibly microbial antigens, prepares fetal immune systems and epithelia for microbial colonization at birth and thereafter.

directly through maternal IgGs that cross the placenta, are sufficient to prime fetal intestinal epithelial cells in mice (11) and can probably also induce memory T cell differentiation. Fetal T cells are hyperresponsive to foreign and endogenous antigens (5, 18, 19). Such T cell hyperresponsiveness itself could be considered evidence for antigen scarcity in utero. This is because T cells in general require subthreshold stimulation through their T cell receptors and tune their thresholds for activation accordingly (20). An environment devoid of triggering antigens would be expected to induce lower thresholds for activation and T cell hyperresponsiveness, as seen in fetuses. Similar models have been put forward for other immune cells, such as natural killer cells (21).

An additional role of transplacental transfer of microbial components could be to provide antigens to fetal T cells such that their thresholds of activation can be tuned down in preparation for postnatal microbial coloniza-

bial colonization yet preventing dangerous infections after birth.

Birth and first encounters

Labor is initiated by hormonal factors that are triggered by physical pressure from the fetus onto the cervix, but also by inflammatory signals and the breakdown of feto-maternal tolerance. Inflammatory stimuli such as infections can trigger such loss of tolerance and initiate preterm labor (24), and specific strains of vaginal bacteria have been implicated (25). Moreover, fetal cells triggered against maternal antigens can secrete cytokines that lead to uterine contractions, illustrating the bidirectional nature of feto-maternal tolerance and its role in preterm labor (26). Once membranes rupture, a baby is exposed to microbes through the birth canal. In vaginal delivery, fecal and vaginal microbes from the mother seed the newborn skin and mouth, whereas babies born through Cesarean section are populated mainly by skin microbes, leading to substan-

Immune reactions to microbiota

In the first postnatal days, a strong systemic immune response occurs in infants, which is dominated by innate immune cells such as monocytes expanding in the blood and the elevation of circulating proinflammatory cytokines, but also by endogenous regulatory factors such as IL-1 receptor antagonist (34). Preterm babies accumulate many disease-associated environmental exposures, such as antibiotics, infections, and cesarean delivery. They also have elevated inflammatory responses in cord blood compared with term children, likely related to prenatal conditions triggering preterm delivery in the first place. After birth, preterm and term infants undergo similar changes in response to postnatal exposures that are similar to all newborns, leading to a marked convergence of immune cell and protein profiles during the first weeks of life in both preterm and term infants (35). Despite these similar adaptive changes during the first weeks of life, very preterm babies (<31 weeks of gestation) remain at much elevated risk (odds ratio ~3.6 versus term infants) of asthma (36) and other immune-mediated diseases.

After birth, cells of the immune system, particularly those at barrier surfaces, must recognize colonizing microbes and determine whether to tolerate or resist them. The establishment of host-microbe mutualism is critical for health, and several processes work to promote this. Children lacking Tregs (e.g., those with immunodysregulation polyendocrinopathy enteropathy X-linked syndrome) typically present early in life with severe inflammation involving both the intestines and skin after microbial colonization (37). In mice, a distinct population of FoxP3$^+$ Tregs are induced in the thymus during the perinatal period that are essential for maintaining self-tolerance and preventing autoimmunity later in life (38). Additional experiments in mice have shown that

the timing and location of intestinal antigen exposure is an important determinant of the responses elicited and the establishment of tolerance to commensals (*39*). During the first 10 days of postnatal life (the neonatal phase), antigens are kept separate from intestinal immune cells by a tight barrier, but from day 11 until weaning (day 21), luminal antigens are encountered mostly by immune cells in the colon. Thereafter, in the postweaning phase, antigens are encountered almost exclusively by immune cells in the small intestine (*40*). Encounters in the preweaning phase (window of opportunity) are more likely to lead to tolerance by induction of Tregs expressing the transcription factor retinoic acid–related orphan receptor γt (RORγt), influencing subsequent exposures to the same antigens later in life (*40*)(Fig. 2). The timing and location of antigen sensing is regulated by goblet cell–associated antigen passages (*41*), which are inhibited by epidermal growth factor (EGF) in breastmilk, thereby allowing for a carefully timed and localized sensing of luminal antigens (*40*) (Fig. 2). At the time of weaning in mice, there is a transient intestinal immune response toward members of the microbiota, which is modulated by dietary and microbial factors such as retinoic acid and short-chain fatty acids that further promote Tregs expressing RORγt (*42*). In the absence of such environmental factors, RORγt⁺ Tregs failed to develop, and experimental colitis induced by dextran sulfate sodium and allergic inflammation induced by oxazolone were more pronounced (*42*). Tregs induced before and during the time of weaning persist, but continuous antigen delivery is required to maintain Treg numbers and tolerance to a particular antigen.

Timing is also important for food allergy development, in which early exposure to potential allergens reduces the risk of developing allergies toward these same antigens (*43*). This effect is enhanced in breastfed infants by maternal IgG-antigen complexes, maternal EGF, and transforming growth factor β in breastmilk (*44*) (Fig. 2). Members of the microbiota also contribute and exert modulatory effects (*45*), and when specific microbes identified to be overrepresented in nonallergic children are transferred into mice, they suppress experimental allergic inflammation by inducing RORγt⁺ Tregs (*46*).

The window of opportunity for tolerance induction is less well defined in humans, but antibiotic exposure during the first 100 days of life has been associated with increased risk of atopy (*31*). This is a period when the circulating immune system in human newborns is particularly dynamic (*35*) and when CD4⁺ T cells carrying gut-homing markers are expanding in the blood, suggestive of immune reactivity to the microbiota (*34*). In contrast

to mice, this reaction is less tightly linked to weaning, but breastmilk and its components exert an important influence on the nature of early-life immune responses during microbial colonization. HMOs in the breastmilk are metabolized by microbes carrying the necessary genetic machinery, leading to the production of metabolites such as indole-3-lactic acid, which likely mediates some of the beneficial effects of breastmilk early in life (*34, 47*). IgA antibodies in breastmilk have been shown to promote colonization by selected microbes in mice by anchoring them to the mucus layer in the intestine (*48*). It is becoming increasingly clear that homeostatic B cell responses in the intestine are dominated by IgA. Conversely, more inflammatory IgG responses to commensal microbes are rare but are more abundant when tolerance is lost, as in inflammatory bowel disease (*49*). Presumably, maternal antibodies in breastmilk and those transferred across the placenta help to tip the balance toward tolerance and immune-microbe mutualism while providing passive immunity to defend from invasive bacteria. Breastfeeding and the expansion of HMO-metabolizing bacteria such as *Bifidobacteriumlongum* subspecies *infantis* (*B. infantis*) induce skewing of T cells both

in vivo and in vitro, away from T helper type 17 (T$_H$17) and T$_H$2 and toward T$_H$1 cell differentiation, and exert a dampening effect on systemic and intestinal inflammation (*34*) (Fig. 2).

Whether these differences in early-life responses to the microbiota and modulation by breastmilk will also translate into different rates of atopy development and other immune-mediated diseases remains to be seen. Clearly, there are numerous evolutionarily conserved mechanisms that promote colonization by beneficial microbes, immune tolerance to these microbes and food components, and the establishment of healthy immune-microbe mutualism early in life. So why do these increasingly fail in industrialized societies, leading to increasing rates of immune-mediated diseases?

Mismatched environments

Immune systems have evolved under strong selective pressure from microbes since the first appearance of humans in Africa 200,000 to 300,000 years ago, dispersal across the continents from 100,000 years ago, and through major lifestyle changes such as agriculturalism ~10,000 years ago (*50*). One way to explain
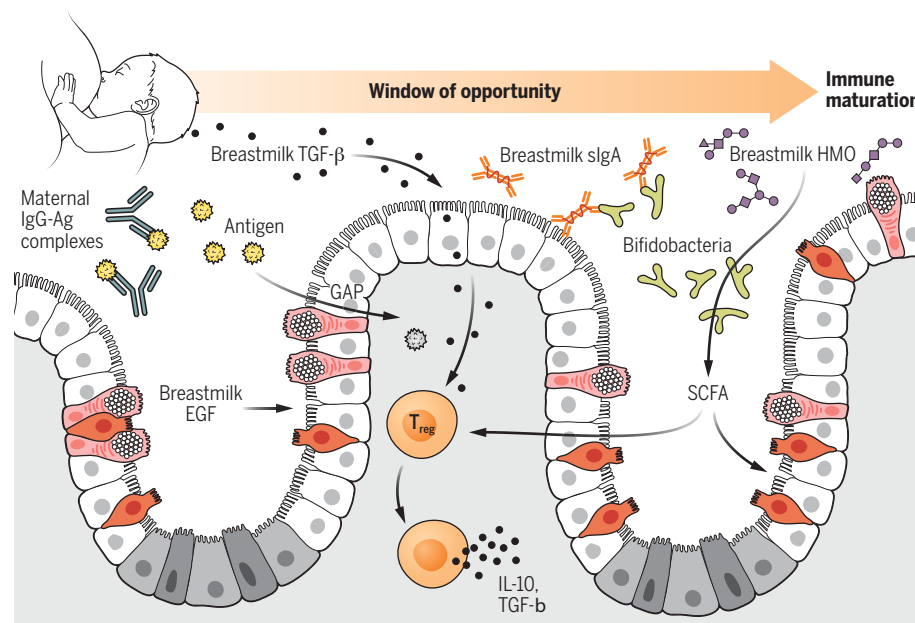
Fig. 2. Tolerance-promoting factors during a window of opportunity early in life. The induction of tolerance toward commensal microbes and harmless food antigens in the gut is promoted during a window of opportunity whereby regulated antigen sampling through goblet cell–associated antigen passages (GAPs) is regulated in time and space by the EGF provided by breastmilk (*40*). Transforming growth factor β (TGFβ) in breastmilk also promotes the differentiation of Tregs expressing the transcription factor RORγt, and HMOs provide a growth advantage to beneficial microbes such as *B. infantis*, which express all of the genes necessary for HMO metabolism. This process of HMO metabolism and microbial expansion leads to the production of specific metabolites such as indole-3-lactic acid, which skew CD4⁺ T cells away from inflammatory T$_H$17 and T$_H$2 and toward T$_H$1 and the prevention of inflammatory responses to the microbiota (*34*). Maternal IgA in breastmilk can selectively promote colonization by specific microbes by anchoring them to the mucus layer, and maternal IgG prevents invasive infections.

the increasing rates of immune-mediated diseases is the hygiene hypothesis, which was formulated based on the observation that decreasing household sizes (fewer siblings) are associated with increased rates of hay fever (*51*), leading to the proposal that reduced infectious disease exposures explains the increasing rates of immune-mediated diseases. An alternative to the hygiene hypothesis is the "old friends" hypothesis, which posits that a loss of specific microbes with which humans coevolved led to increased rates of immune-mediated disease by altering immune-microbe interactions and immune system regulation (*52*). Both of these hypotheses exemplify mismatches between ancestral environments in which humans evolved and modern environmental conditions (Fig. 3A). Such mismatches between ancestral and modern environments could involve losses of important microbes that serve valuable functions as commensals, loss of specific nutrients with important nonredundant immunological functions, or new pollutants and chemical substances that have never been encountered before and might affect immune system processes.

The reasons for priming the phenotype and locking certain conditions in place during a critical period of adaptation early in life must serve some benefit over carrying a fully flexible phenotype throughout life. One likely benefit in the case of establishing immune-microbe mutualism is that colonization and tolerance induction can occur during a period when the infant is rather immobile, physically sheltered, fed breastmilk, and immunologically protected by passive immunity and antimicrobial peptides. Locking certain aspects of the phenotype during these conditions is likely more beneficial than allowing a fully amenable phenotype throughout life and risking substantial changes upon infections and other insults. However, the benefit of early-life priming only holds if the environmental conditions experienced early in life are representative of those encountered later in life. Interpreting this from the perspective of mismatches between ancestral and modern environments suggests that a mismatch involving key factors that shape immune-microbe interactions early in life, such as colonization with key microbes, formula feeding, cesarean delivery, and antibiotic exposure, can have disproportionally large impacts on long-term health compared with, for example, antibiotic exposure later in life (Fig. 3A). This framework is meant to bring together the foundation offered by evolution with the idea of mismatched ancestral and modern environments. It also separates specific mismatched environmental exposures into those that are influential during the early-life establishment of immune-microbe interactions and hardwiring of selected elements of the phenotype versus those environmental exposures that more directly affect humans later in life, such as smoking, alcohol consumption, exposure to sun radiation, and solid food components (Fig. 3A).

Trade-offs between growth and immune function

Another evolutionary mechanism that controls overall immune function is energy allocation and trade-offs between investments in immunity (maintenance), growth, and reproduction, as described by life history theory (*1*). Humans are born highly dependent on their parents for food and protection, and there are strong incentives for newborns to grow rapidly. This is reflected by the impressive growth rates of infants, who often triple their body weight during the first year of life. Until recently, food supply has been a limiting factor for nursing mothers, and the quantity and quality of breastmilk would lead to trade-offs between investments in physical growth and immunity. For example, children who are growth restricted in utero show stronger anti-helminth IgE responses compared with nonrestricted children, even when tested in adolescence (*53*). Given that parasitic infections are a common cause of malnutrition in mothers and growth restriction in fetuses, prioritization of Th2 immunity in newborns as preparation for likely parasitic exposure after birth is a rational evolutionary outcome.

In industrialized societies, food supply is rarely a limiting factor during pregnancy or lactation. Both breastfed and formula-fed
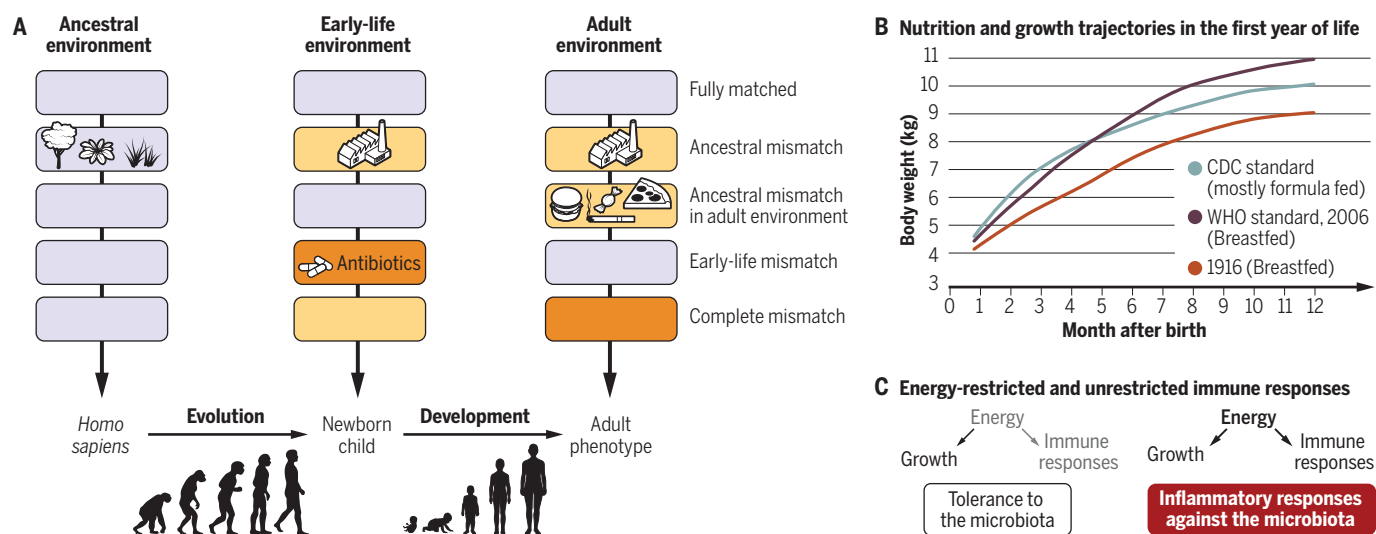
Fig. 3. Mismatch between ancestral and modern environments and altered trade-offs between growth and immunity. (**A**) During human evolution, the human immune system evolved under strong selective pressures from microbes. Given the substantial recent changes to human environmental conditions, a mismatch between ancestral and modern environments has occurred. Because of the particular importance of the critical period early in life, mismatches of key factors involved during this time (antibiotics, breastfeeding, loss of key commensals) may affect long-term health and immune-mediated diseases. Using this proposed framework of mismatched ancestral, early-life, and adult environments, the consequences of specific environmental changes can be investigated. (**B**) Trade-offs between investments in physical growth and immunity are described by life history theory (*1*). Changing growth rates over the past century (*54*–*58*) and differences between breastfed and formula-fed infants today reflect such underlying changes. (**C**) One hypothesis from these observations is that changes to environmental conditions, such as unlimited access to nutrients and energy, can derail evolutionarily conserved mechanisms for resource allocation, which also affects the immune system by removing the constraints imposed by resource allocation early in life, leading to more exuberant inflammatory responses to commensals, dysbiosis, and intestinal inflammation. This could result in increased rates of both immune-mediated diseases and obesity and metabolic disorders in the long term.

ILLUSTRATION: KELLIE HOLOSKI/*SCIENCE*

infants follow substantially different growth trajectories today compared with newborns only ~100 years ago (*54*–*56*) (Fig. 3B). In 1916, breastfed infants had a more constant growth rate during their first 8 months of life, which thereafter leveled off at a decreasing growth rate in the following months (*57*) (Fig. 3B). Historic growth rates further back during human evolution are not known but were likely even more different from the current ones. In 2006, growth charts from the US Centers for Disease Control and Prevention (CDC) showed that mostly formula-fed infants had initial growth rates that were slower than exclusively breastfed infants in the World Health Organization (WHO) standard curve, but superseded breastfed infants from 6 months onward, leading to higher average weights compared with breastfed infants at 1 year of age (*58*) (Fig. 3B). Breastfeeding is associated with reduced risks of immune-mediated diseases later in life (*59*). One hypothesis from these observations is that an additional cause of the increasing rates of immune-mediated diseases linked to exuberant inflammatory reactions to commensals and harmless allergens could be a decoupling of normal physiological mechanisms of immune regulation and inhibition based on resource allocation and trade-offs between growth and immune function (Fig. 3C). Such decoupling could contribute further to explaining unrestricted inflammatory responses early in life, impaired immune-microbial mutualism, and increasing rates of childhood obesity as seen in many industrialized societies today. To further explore this possible explanation, we need to take growth rates into account when assessing immune-mediated diseases and further investigate the mechanisms of resource allocation for growth versus immunity to better identify at-risk children and test new interventions to optimize individual trajectories for growth and immunological health.

A global view

Immune-mediated diseases differ worldwide, with high and increasing incidences of asthma, allergies, and autoimmune diseases in industrialized societies, often in urban areas far away from animals and where repeated antibiotic exposures and cleanliness have depleted many previously common microbes. One important example is *B. infantis*, which dominates the gut microbiomes of young, breastfed children in Bangladesh (*60*) and Malawi (*61*), but is nearly absent in children in Europe (*62*) and North America (*63*). This exemplifies a mismatch between ancestral and modern environments with particular importance during early-life immune system development and consequently long-term effects on diseases such as type 1 diabetes (*64*), atopy, and asthma (*31*). Efforts to restore beneficial microbes and prevent immune-mediated diseases after the

initial window of opportunity have proven challenging, and one reason for this could be that immune-microbe interactions and tissue-specific niches are less amenable to change after this initial period. The proposed potential importance of life history theory and the trade-off between investments in physical growth and immune function is an additional possible mechanism, established during human evolution, of restraining exuberant responses to colonizing microbes early in life. If true, then this additional layer of regulation, which was lost because of virtually unrestricted access to calories in modern industrialized societies, could further explain parallel increases in childhood obesity and immune-mediated diseases in many countries. Antibiotic stewardship, promotion of breastfeeding whenever possible, and compensatory nutritional strategies when breastfeeding is not possible will likely be important in reducing the escalating burden of immune-mediated diseases in the future. By learning the mechanisms of immune system regulation and trade-offs between investments in growth and immunity, we might be better able to restrain immune responses to commensals and harmless food components and thus restore tolerance and reduce risks of disease, perhaps even later in life.

Although unlimited access to energy might derail evolutionary mechanisms for restrained immunity, energy malnutrition and scarcity also have detrimental consequences for both immunity and growth. Some of these mechanisms might be mediated by altered immune-microbe interactions, as shown by transfers of microbes from malnourished children to germ-free mice, which led to growth restriction (*65*). In such situations, normal trade-offs are also derailed and immune function impaired, leading to repeated infections, chronic inflammation, and worsened malnutrition (*66*). In a recent clinical trial of 2- to 6-month-old children in Bangladesh with acute malnutrition, supplementation with *B. infantis* was able to break this vicious cycle by reducing markers of intestinal inflammation and improving weight gain (*67*).

In the context of healthy immunometabolic regulation, increased investigation into the evolutionary mechanisms of resource allocation for growth and immunity are needed in the hope of achieving a healthier balance. This might favor mutualistic immune-microbe relationships that are more consistent with ancestral environments while benefiting from the advances made with respect to hygiene, vaccinations, and an increased human life span.

REFERENCES AND NOTES

1. S. C. Stearns, *The Evolution of Life Histories* (Oxford Univ. Press, 1992); https://global.oup.com/ushe/product/the-evolution-of-life-histories-9780198577416?cc=se&lang=en&).
2. L. A. Herzenberg, L. A. Herzenberg, *Cell* **59**, 953–954 (1989).
3. J. E. Mold *et al.*, *Science* **330**, 1695–1699 (2010).
4. J. A. Byrne, A. K. Stankovic, M. D. Cooper, *J. Immunol.* **152**, 3098–3106 (1994).
5. J. E. Mold *et al.*, *Science* **322**, 1562–1565 (2008).
6. J. M. Kinder, I. A. Stelzer, P. C. Arck, S. S. Way, *Nat. Rev. Immunol.* **17**, 483–494 (2017).
7. D. G. Bunis *et al.*, *Cell Rep.* **34**, 108573 (2021).
8. Y. Li *et al.*, *Cell Stem Cell* **27**, 732–747.e7 (2020).
9. A. I. Lim *et al.*, *Science* **373**, eabf3002 (2021).
10. J. Y. Hong *et al.*, *Cell* **180**, 847–861.e15 (2020).
11. M. Gomez de Agüero *et al.*, *Science* **351**, 1296–1302 (2016).
12. M. C. Collado, N. Segata, *Nat. Med.* **26**, 469–470 (2020).
13. E. Rackaityte *et al.*, *Nat. Med.* **26**, 599–607 (2021).
14. K. M. Kennedy *et al.*, *Nat. Microbiol.* **6**, 865–873 (2021).
15. M. C. de Goffau *et al.*, *Nature* **572**, 329–334 (2019).
16. N. Li *et al.*, *Nat. Immunol.* **20**, 301–312 (2019).
17. A. Mishra *et al.*, *Cell* **184**, 3394–3409.e20 (2021).
18. J. Michaëlsson, J. E. Mold, J. M. McCune, D. F. Nixon, *J. Immunol.* **176**, 5741–5748 (2006).
19. H. E. Kay, J. Doe, A. Hockley, *Immunology* **18**, 393–396 (1970).
20. Z. Grossman, W. E. Paul, *Proc. Natl. Acad. Sci. U.S.A.* **89**, 10365–10369 (1992).
21. M. G. Netea *et al.*, *Science* **352**, aaf1098 (2016).
22. B. A. Reikie *et al.*, *J. Acquir. Immune Defic. Syndr.* **66**, 245–255 (2014).
23. K. Gbédandé *et al.*, *Infect. Immun.* **81**, 2686–2696 (2013).
24. N. Gomez-Lopez, D. StLouis, M. A. Lehr, E. N. Sanchez-Rodriguez, M. Arenas-Hernandez, *Cell. Mol. Immunol.* **11**, 571–581 (2014).
25. J. M. Fettweis *et al.*, *Nat. Med.* **25**, 1012–1021 (2019).
26. M. Frascoli *et al.*, *Sci. Transl. Med.* **10**, eaan2263 (2018).
27. M. G. Dominguez-Bello *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 11971–11975 (2010).
28. P. Ferretti *et al.*, *Cell Host Microbe* **24**, 133–145.e5 (2018).
29. N. A. Bokulich *et al.*, *Sci. Transl. Med.* **8**, 343ra82 (2016).
30. J. Metsälä *et al.*, *Epidemiology* **24**, 303–309 (2013).
31. M.-C. Arrieta *et al.*, *Sci. Transl. Med.* **7**, 307ra152 (2015).
32. M.-L. Wernroth *et al.*, *Diabetes Care* **43**, 991–999 (2020).
33. I. M. Aris *et al.*, *JAMA Netw. Open* **4**, e2116581 (2021).
34. B. M. Henrick *et al.*, *Cell* **184**, 3884–3898.e11 (2021).
35. A. Olin *et al.*, *Cell* **174**, 1277–1292.e14 (2018).
36. H. Trønnes, A. J. Wilcox, R. T. Lie, T. Markestad, D. Moster, *Pediatr. Allergy Immunol.* **24**, 782–787 (2013).
37. R. S. Wildin *et al.*, *Nat. Genet.* **27**, 18–20 (2001).
38. S. Yang, N. Fujikado, D. Kolodin, C. Benoist, D. Mathis, *Science* **348**, 589–594 (2015).
39. M. O. Dhariwala, T. C. Scharschmidt, *Trends Immunol.* **42**, 1088–1099 (2021).
40. K. A. Knoop *et al.*, *Sci. Immunol.* **2**, eaao1314 (2017).
41. J. R. McDole *et al.*, *Nature* **483**, 345–349 (2012).
42. Z. Al Nabhani *et al.*, *Immunity* **50**, 1276–1288.e5 (2019).
43. G. Du Toit *et al.*, *J. Allergy Clin. Immunol.* **122**, 984–991 (2008).
44. L. T. Yokanovich, R. D. Newberry, K. A. Knoop, *Clin. Exp. Allergy* **51**, 518–526 (2021).
45. T. Feehley *et al.*, *Nat. Med.* **25**, 448–453 (2019).
46. A. Abdel-Gadir *et al.*, *Nat. Med.* **25**, 1164–1174 (2019).
47. M. F. Laursen *et al.*, *Nat. Microbiol.* **6**, 1367–1382 (2021).
48. G. P. Donaldson *et al.*, *Science* **360**, 795–800 (2018).
49. A. Fleming, T. Castro-Dopico, M. R. Clatworthy, *Scand. J. Immunol.* **95**, e13139 (2022).
50. L. Quintana-Murci, *Cell* **177**, 184–199 (2019).
51. D. P. Strachan, *BMJ* **299**, 1259–1260 (1989).
52. G. A. W. Rook, R. Martinelli, L. R. Brunet, *Curr. Opin. Allergy Clin. Immunol.* **3**, 337–342 (2003).
53. T. W. McDade, C. W. Kuzawa, L. S. Adair, M. A. Beck, *Clin. Exp. Allergy* **34**, 44–50 (2004).
54. L. T. Weaver, *Eur. J. Clin. Nutr.* **65**, 3–9 (2011).
55. K. G. Dewey, *Biol. Neonate* **74**, 94–105 (1998).
56. K. G. Dewey *et al.*, *Pediatrics* **96**, 495–503 (1995).
57. T. B. Robertson, *Am. J. Physiol.* **41**, 547–554 (1916).
58. L. M. Grummer-Strawn, C. Reinold, N. F. Krebs; Centers for Disease Control and Prevention (CDC), *MMWR Recomm. Rep.* **59** (RR-9), 1–15 (2010).
59. B. Dawod, J. S. Marshall, M. B. Azad, *Curr. Opin. Gastroenterol.* **37**, 547–556 (2021).
60. M. N. Huda *et al.*, *Pediatrics* **143**, e20181489 (2019).
61. Ł. Grześkowiak *et al.*, *J. Pediatr. Gastroenterol. Nutr.* **54**, 812–816 (2012).
62. T. R. Abrahamsson *et al.*, *Clin. Exp. Allergy* **44**, 842–850 (2014).
63. M. B. Azad *et al.*, *CMAJ* **185**, 385–394 (2013).
64. T. Vatanen *et al.*, *Cell* **165**, 842–853 (2016).
65. L. V. Blanton *et al.*, *Science* **351**, aad3311 (2016).
66. C. D. Bourke, J. A. Berkley, A. J. Prendergast, *Trends Immunol.* **37**, 386–398 (2016).
67. M. J. Barratt *et al.*, *Sci. Transl. Med.* **14**, eabk1107 (2022).

REVIEW

Local barriers configure systemic communications between the host and microbiota

Qiuhe Lu and Thaddeus S. Stappenbeck*

Associations between the dynamic community of microbes (the microbiota) and the host they colonize appear to be vital for ensuring host health. Microbe-host communication is actively maintained across physiological barriers of various body sites and is mediated by a range of bidirectional secreted proteins and small molecules. So far, a range of "omics" methods have succeeded in revealing the multiplicity of associations between members of a microbiota and a wide range of host processes and diseases. Although these advances point to possibilities for treating disease, there has not been much translational success thus far. We know little about which organisms are key contributors to host health, the importance of strain differences, and the activities of much of the chemical "soup" that is produced by the microbiota. Adding to this complexity are emerging hints of the role of interkingdom interactions between bacteria, phages, protozoa, and/or fungi in regulating the microbiota-host interactions. Functional approaches, although experimentally challenging, could be the next step to unlocking the power of the microbiota.

Humans live in a microbe-dominated world. Indeed, life on Earth shows substantial evolutionary evidence of interdependency with microorganisms. Major drivers of organismal interactions occur at tissue barriers of host organs that interface with the environment and the microbiota. The composition and function of the microbiota at tissue barriers are influenced by many factors, including early-life microbial exposure, host genetics, and dietary and lifestyle habits. In turn, the microbiota in these locations affect a wide array of systemic processes, including metabolism, immunity, circadian rhythm, and behaviors. For any organism in a natural environment, a large array of factors are at work, but urbanized humans experience another level of inputs that range from ultra-processed foods, irregular eating habits, and high amounts of medication to shift work and poor sleep hygiene, depleting the microbiota and causing circadian rhythm disruption that is linked with many chronic diseases (1).

Microbiome-related research has exploded in the past two decades. The field is now emerging from an initial phase of cataloging microbial nucleic acid sequences and their metabolic products. We have created vast libraries of "multi-omics" data to mine for natural products, patterns of microbial diversity, and functional dynamics. An "omics approach" has been applied to nearly every organ and disease occurring in humans, as well as many other organisms and ecosystems. However, despite the huge research efforts and investment so far, our understanding is still too incomplete to exploit the therapeutic potential that the microbiota might offer.

How can we fill in the gaps in our understanding about host-microbiota relations? We need to move from simply making associations of microbial taxa with host phenotypes toward elucidating the mechanisms by which specific microbes contribute to healthy host physiology as well as disease. The challenge is that a microbiota may contain representatives from every branch of the tree of life, including bacteria, viruses, yeast, archaea, protozoa, and helminths. Thus far, our understanding of how these diverse microorganisms operate at tissue barriers is sparse and selective. Our hope is that emerging research will define the rules that govern these complex host-microbe interactions and establish new paradigms that accelerate the development of evidence-based therapeutics in this area.

Defining host barriers

Mucosal surfaces interface with the environment and include several distinct habitats where microbes interact with the host: oral cavity, ocular structures, genitourinary tract, upper and lower respiratory tract, skin, and gut. By far, the largest population of these microbes is in the distal gastrointestinal tract, where their amounts are similar to the abundance of microbes found in topsoil during summer months (up to 10% of wet weight biomass). The microbial population ranges widely over all other mucosal surfaces, spanning up to 13 orders of magnitude, but the community composition varies between each surface and can be altered during disease.

Mucosal barriers that prevent symbionts from overwhelming their hosts include physical, chemical, and microbial components (Fig. 1). The physical barrier is composed of epithelial cells and closely associated immune and stromal cells that are tailored for each location. Epithelial tight junction proteins are crucial for maintaining epithelial barrier integrity, and their function can be modulated by microbes.

We do know that the gut microbial product—the short-chain fatty acid (SCFA) butyrate—promotes barrier function by up-regulating the expression of tight junction proteins (2). Overlying the epithelium is a chemical barrier that is composed of combinations of mucus layers, secretory immunoglobulin A (SIgA) produced by local plasma cells, reactive oxygen species, reactive nitrogen species, and a variety of broad spectrum antimicrobial proteins (α-defensins), all of which are designed to help keep both symbionts and pathogens in check (3, 4). Mucins are highly glycosylated polymeric proteins that form a protective chemical barrier (mucus) that can help segregate microorganisms and noxious substances from the underlying epithelium (5). The intestines of germ-free mice have an underdeveloped mucus layer; here, symbionts not only stimulate the production and secretion of mucus but also promote its turnover in the lumen. The metabolites produced by the mucin-degrading bacteria *Akkermansia muciniphila* include SCFAs and sulfate that help maintain the gut lining and promote barrier function (6). Another common bacterial symbiont, *Enterococcus faecium*, secretes a peptidoglycan hydrolase (SagA) that enhances intestinal epithelial barrier function and pathogen tolerance. SagA activity generates small muropeptides that activate the host nucleotide-binding oligomerization domain–containing protein 2 (NOD2) signaling pathway to modulate expression of barrier components such as mucin 2, cryptdin 2, and the C-type lectin Reg3 gamma (RegIIIγ) (7, 8). SIgA induced by symbionts at mucosal surfaces binds to microbes and microbial products to prevent their translocation across the epithelia, which plays a key role in maintaining mucosal homeostasis and enforcing barrier function (4).

Perturbations of mucosal barriers associated with compositional changes to microbial communities are loosely described as dysbiosis, a term that characterizes microbiomes with low microbial diversity, depletion of potentially beneficial microbes, and expansion of potentially harmful microorganisms (9). Dysbiotic microbial assemblages can exacerbate disruption of the mucosal barrier. Both infectious diseases and chronic inflammatory disorders feature dysbiosis, and altered microbiota and mucosal disruption are observed in lung diseases, such as asthma (10).

Germ-free mouse models have been adopted experimentally to model humans with specific disease conditions. In a notable study, a consortium of 11 IgA-targeted bacterial strains cultured from an undernourished Malawian child was sufficient to induce a diet-dependent enteropathy in germ-free mice. The phenotype was characterized by rapid disruption of the intestinal epithelial barrier, weight loss, and sepsis (11); this methodology has since become foundational to the development of successful

Department of Inflammation and Immunity, Lerner Research Institute, Cleveland Clinic, Cleveland, OH, USA.
*Corresponding author. Email: stappet@ccf.org

food-based therapies for this form of malnutrition (12).

A "healthy" microbiota assemblage of symbiotic organisms acts in concert to exclude harmful invasive species such as *Salmonella*, which is a phenomenon known as colonization resistance. Hence, an important consequence of dysbiosis is the loss of colonization resistance to pathogens. For example, antibiotic treatment disrupts indigenous microbial communities and produces an environment permissive for opportunistic pathogens, such as *Clostridioides difficile*, to colonize and expand in the colon by promoting the germination of its spores to form vegetative cells. Two exotoxins secreted by vegetative cells—toxins A and B—damage the intestine mucosal barrier, induce a strong inflammatory response, and cause diarrhea (13).

Fecal microbiota transplantation (FMT) can successfully restore a functional microbiome and mucosal homeostasis in *C. difficile* infection. However, treatment of other intestinal diseases by FMT, such as inflammatory bowel disease (IBD), has shown only modest efficacy in clinical trials. Poor efficacy of FMT in IBD could be caused by the complexity and heterogeneity of the disease or the interpersonal diversity of donors' fecal microbiota; it could also occur because the underlying dysbiosis in IBD patients is an effect of the disease process. Other challenges of FMT include the risks of introducing poorly characterized microbial material into patients, which may include pathobionts, and unreliable long-term establishment of the transplanted microbiome (14). Nevertheless, the promising results from FMT studies in patients with *C. difficile* infection have inspired valuable mechanism-based microbiome studies on health and disease. Recently, a phase 3, double-blind, randomized, placebo-controlled clinical trial was conducted to evaluate the efficacy of oral microbiome (SER-109) therapy for treatment of recurrent *C. difficile* infection. Compared with the placebo group, the group given SER-109 showed a superior capacity to prevent recurrent infection; meanwhile, the observed safety profile across the groups remained similar (15). These studies indicate that a rational design for treatment with defined microbial communities or microbial metabolites can be therapeutically beneficial.

Monomicrobial- and polymicrobial-driven phenotypes

Microbiome studies commonly assume that the association of differentially abundant microbial taxa with a given phenotype may be causal. A common goal of such studies has been the identification of microbial taxa or products responsible for a phenotype. To do this, next-generation sequencing of specific microbes is routinely performed for bacterial 16S ribosomal RNA (rRNA) genes, eukaryotic 18S rRNA genes, or preparations that enrich for viruses (including reverse transcription for RNA viruses). An "omics-initiated approach" (e.g., Fig. 2) has successfully identified candidate microbes that are associated with a given phenotype, which are then tested by colonization experiments in either gnotobiotic or antibiotic-treated mice.

This approach has been successful in revealing several key microbial species and functions. For example, an organism called segmented filamentous bacterium has been found to be abundant in mice supplied by Taconic Biosciences but not in mice obtained from The Jackson Laboratory. These filamentous bacteria enhance T helper 17 (T_H17) cells and immunity to intestinal bacterial pathogens (16). Another example that was identified based on differences in intestinal intraepithelial lymphocyte (IEL) abundance within a mouse facility is *Lactobacillus reuteri*, which was identified through sequencing as a biomarker for mice with increased IELs. *L. reuteri* was shown to drive the IEL phenotype by monoassociation in germ-free mice (17). Similarly, the skin bacterial symbiont *Staphylococcus epidermidis* induces IL-17A$^+$CD8$^+$ T cells, which improves the skin barrier function in a model of epicutaneous infection with the fungal pathogen *Candida albicans* (18). Bacterial sequencing combined with mathematical modeling has been used to elucidate the effect of antibiotic treatment of *C. difficile* (different antibiotic treatments showed variation in intestinal damage). This approach also showed that the occurrence of a secondary bile acid–producing bacterial symbiont, *Clostridium scindens*, prevented *C. difficile* colonization (19).

Many phenotypes are not the product of a single organism, and reliance on abundance alone can be a challenging approach to identify functional communities of microbes. Here, functional screening (the screen-initiated approach shown in Fig. 2) has been successful. One approach is to make a simplified microbiota before engaging in functional low-throughput screens and tests. This rationale has been successfully applied to the isolation of bacterial spores from fecal samples from healthy human donors. The spores can then be tested iteratively in germ-free mouse colonization experiments. By this method, a consortium of 11 bacterial strains was identified that induced IFNγ$^+$CD8$^+$ T cells in mice and protected them against *Listeria monocytogenes* infection. Excitingly, this consortium also enhanced antitumor immunity in the MC38 murine colon carcinoma model (20). Using a similar approach, a consortium of 17 bacterial strains was found to induce FOXP3$^+$ regulatory T cells (T_{regs}) that attenuate 2,4,6-trinitrobenzenesulfonic acid (TNBS)–induced colitis (21). An interesting study on antibiotic resistance has identified four bacterial symbionts—*Bacteroides sartorii*, *Parabacteroides distasonis*, *Clostridium bolteae*, and *Blautia producta*—that prevented and eliminated vancomycin-resistant enterococci (VRE) colonization in ampicillin-treated mice (22). What is most interesting is that all four organisms interact in vivo: *B. producta* suppresses VRE growth, *C. bolteae* is required for the colonization of *B. producta*, and *B. sartorii*
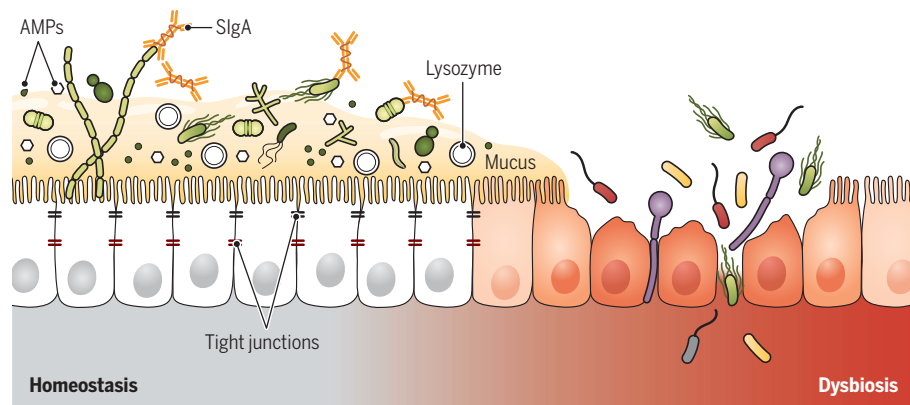
Fig. 1. A model of a generic mucosal surface showing microbiota-host interactions at the interface. In homeostasis, the host provides both physical (epithelial cells) and chemical barriers [mucus, antimicrobial peptides (AMPs), lysozyme, SIgA, etc.] to segregate symbiotic microorganisms, which usually live in harmony with the host and form an additional barrier. The tight junctions of epithelial cells are crucial for maintaining epithelial barrier integrity. Symbiotic bacteria such as *E. faecium* and *A. muciniphila* can improve host barrier function through different mechanisms. Microbiota-derived products such as SCFAs and secondary bile acids are also barrier enhancers. Interruptions in the balance of microbiota can lead to dysbiosis, which compromises barrier integrity, favors bacterial translocation, and causes expansion of harmful microbes (e.g., *C. albicans* and *C. difficile*). Thus, dysbiosis has been associated with a wide range of chronic illnesses, including IBD and asthma.
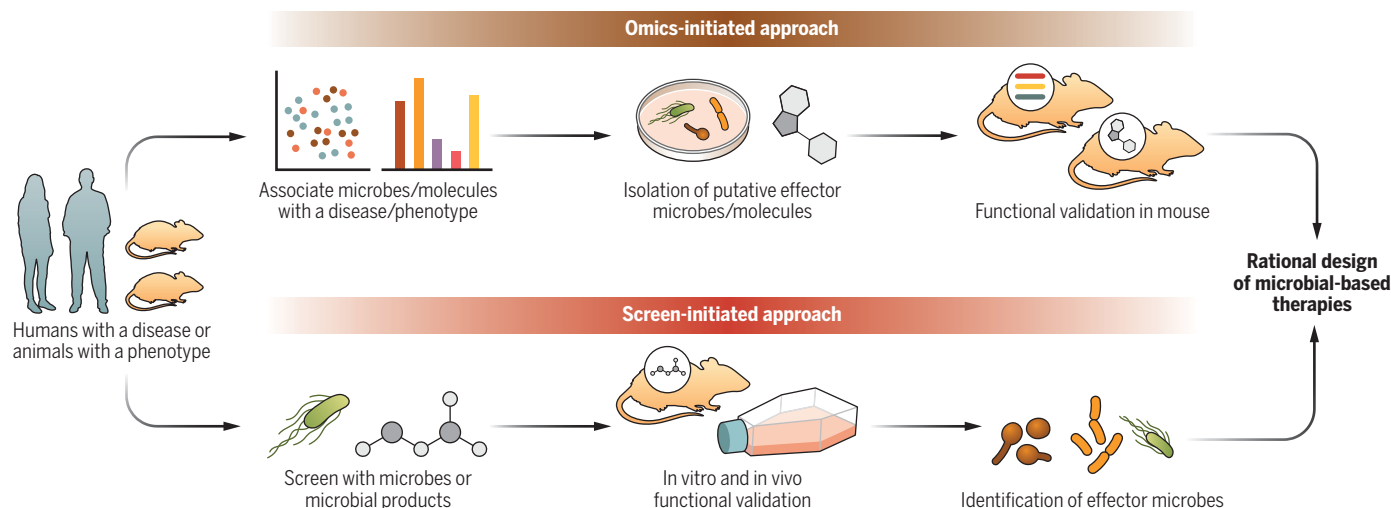
Fig. 2. Omics- and screen-initiated approaches in microbiome research.
The omics-initiated approach (top) begins with multi-omics data generated from humans with a disease or animals with a phenotype that has features of the disease state. By computational analysis of the data, individual bacteria, minimal consortia, or altered microbial products that are enriched or diminished are further tested in animal models to identify potentially causal microorganisms or their products. The screen-initiated approach (bottom) begins with a functional screen of microbial products (metabolites and enzymes) or microbes themselves on disease-relevant host pathways or cells. By using functional tests in both in vitro and in vivo studies, the causal molecules and microbes can be identified. Integration of these two approaches in microbiome studies is a very promising method to find microbial targets for therapies in human diseases.

and *P. distasonis* produce high amounts of β-lactamase, which allows *C. bolteae* and *B. producta* to colonize ampicillin-treated mice. This approach can provide mechanistic insights into interactions between the microbiota and host in different contexts.

Establishing experimental Koch's postulates

An ecological version of Koch's postulates, which are the four criteria designed to assess whether a microorganism causes a disease, offers a useful framework for functional studies of the microbiota (*23, 24*) that can be modified for experimentation with in vivo models. The goal is to show that a microbe (or microbes) isolated from an animal with a particular phenotype can be isolated in pure culture and then reintroduced into a susceptible host to reproduce the phenotype. This procedure can also be applied to microbes isolated from humans with disease and transferred into animal models to test for phenotypes that occur with the disease. For example, *Bacteroides* spp. symbionts isolated from healthy littermate mice induced intestinal inflammation in a genetically susceptible mouse: a double-knockout mouse that is deficient in both transforming growth factor β receptor II (TGFβRII) and interleukin-10R2 (IL-10R2). Although a single *Bacteroides* species was not sufficient to induce disease, these findings fulfilled Koch's postulates in a host genotype–specific fashion. More importantly, this study established a conceptual framework for assessing a microbe's disease-inducing potential (*25*). A recent study found that dual-specificity phosphatase 6 (Dusp6)–deficient mice (*Dusp6*$^{-/-}$ mice)

can counteract colonic mucosal damage in an experimental model because of enhanced colonic barrier integrity. In this case, a single bacterial strain belonging to the genus *Duncaniella* was isolated from *Dusp6*$^{-/-}$ mice, and if germ-free mice were colonized with the isolated *Duncaniella* strain, they showed less chemically induced injury. Although it is likely that other microbial strains in *Dusp6*$^{-/-}$ mice could have similar and redundant effects, this study showed the potential of using a specific bacterial isolate to prevent intestinal epithelium damage during colitis (*26*). A similar discovery showed that colonization of murine norovirus in mice that contain a mutation in the Crohn's disease susceptibility gene [Atg16L1 hypomorphic (Atg16L1^{HM}) mice] recapitulates the abnormalities observed in the Paneth cells of conventionally raised norovirus-positive Atg16L1^{HM} mice (*27*). When colonizing conventional or gnotobiotic mice, wound healing triggered by the increased expression of chemokine CCL5 was impaired by the presence of strains of the fungus *Debaryomyces hansenii* isolated from inflamed tissues of Crohn's disease patients as well as by strains of this fungus extracted from antibiotic-treated conventional mice (*28*). Recovery of *D. hansenii* from affected mice verified that both the host and causative microbes can be potential targets for therapeutics in Crohn's disease.

The next step from ecological tests is to develop molecular tests of Koch's postulates (*29*) by determining the underlying mechanisms that might support further therapeutic studies. The human opportunistic fungal pathogen *C. albicans* colonizes many mucosal sites,

particularly in the gastrointestinal and genitourinary tracts. Alterations in the indigenous microbiota and host immunity lead to the overgrowth of *C. albicans* and the production of invasive hyphae that secrete candidalysin, a cytolytic peptide toxin. This toxin damages epithelial cells, resulting in phenotypes such as leaky gut syndrome, and causes life-threatening disseminated infections. *C. albicans* strains that lack this toxin are avirulent and have not been seen to damage epithelial cells in a mucosal-infection model (*30*). These are predominately preclinical studies and need translation to human disease to be useful for developing targets for therapies.

Interkingdom interactions within microbiota

Mucosal surfaces are densely colonized with a wide variety of microbes, some of which have potential pathogenic capabilities (i.e., bacterial and fungal pathobionts). Major perturbations of microbiota composition are commonly observed in patients with chronic inflammatory disorders, such as IBD and asthma. Dysbiosis triggered by antibiotic treatment can allow the expansion of pathobionts because of the loss of commensal microbes, especially those at mucosal surfaces. Innate inflammatory responses may cause damage to host mucosal barriers and allow invasion and systemic dissemination of pathobionts (*31*). As a result of colonization resistance by a mature microbiota, *C. albicans* cannot colonize the gastrointestinal track in adult mice. Two commensal bacteria, *B. producta* and *Bacteroides thetaiothaomicron*, antagonize *C. albicans* in the mouse intestine. These bacteria induce expression of hypoxia

inducible factor 1-α (HIF1-α) and production of the antimicrobial peptide CRAMP (LL-37), which together ensure transkingdom resistance to *C. albicans* (*32*). Overgrowth of the yeast *Pichia kudriavzevii* in infant mice and humans because of dysbiosis in early life is an associated risk for allergic airway disease later in life. Neonatal mice exposed to *P. kudriavzevii* exhibit increased inflammatory responses as they mature. Mice with the highest loads of *P. kudriavzevii* have the lowest fecal amounts of bacterially derived SCFAs. Studies using in vitro assays have reported that these SCFAs inhibit the growth and morphology of asthma-associated *P. kudriavzevii*. In addition, mice that have been administered a cocktail of SCFAs (acetate, butyrate, and propionate) show reduced colonization by *P. kudriavzevii* (*33*). Thus, although the mechanisms have yet to be defined, these studies indicate that overgrowth of one fungal pathobiont stems from a release of control by associated gut bacteria.

Likewise, *Wallemia mellicola*, a commensal fungus of both humans and mice, expands in the intestines of antibiotic-treated mice but without any accompanying overgrowth of the total fungal community. In antibiotic-treated C57BL/6 mice or altered Schaedler flora mice with a defined fungus-free bacterial community, colonization by *W. mellicola* aggravates asthma-like lung inflammation in house dust mite antigen–induced allergic airway disease (*34*). Antibiotic therapy could be an important risk factor for inflammatory disease triggered by fungal overgrowth because of unrecognized, but potentially important, mechanisms of interkingdom interactions between bacteria and fungi.

Macroparasites, including helminths and protozoa, can also have complex, reciprocal relationships with bacteria and viruses. The protozoan *Trichomonas vaginalis* is the most prevalent nonviral, sexually transmitted infection worldwide. *T. vaginalis* colonization alters the resident microbiota composition and leads to urogenital dysbiosis. However, *T. vaginalis* can itself be parasitized by *T. vaginalis* virus, which is found in ~50% of clinical isolates; when the virus is present, it associates with a greater severity of trichomoniasis in humans because the virus itself induces a Toll-like receptor 3–driven proinflammatory response (*35*). Another protozoan, *Tritrichomonas musculis*, induces IL-18–driven T_H1 and T_H17 responses in mice, which modulates the outcome of *Salmonella typhimurium*–induced enterocolitis (*36*). Further work on coinfection of mice with the helminth *Heligmosomoides polygyrus bakeri* and West Nile virus showed increased mortality, partially associated with altered gut morphology and transit, as well as translocation of gut bacteria into the spleen via blood (*37*). However, infection of mice with *H. p. bakeri* alone altered the intestinal microbiota and

increased SCFA production such that transfer of the *H. p. bakeri*–modified microbiota into antibiotic-treated or germ-free mice protected them against allergic asthma (*38*). A recent study identified a previously unknown bactericidal protein called small proline-rich protein 2A (SPRR2A) that was induced in intestinal secretory cells (goblet and Paneth cells in mouse and goblet cells in human) upon infection with *H. polygyrus*. IL-4 and IL-13 elicited by the helminth infection strongly induced SPRR2A expression. SPRR2A selectively kills Gram-positive bacteria by disrupting their membranes, thus limiting bacterial invasion of intestinal tissue during helminth infection (*39*).

Bacteriophages are predominant in the virome of the healthy gut. Phages are implicated in

"Dysbiosis triggered by antibiotic treatment can allow the expansion of pathobionts because of the loss of commensal microbes…"

horizontal gene transfer between bacterial species, helping to shape the composition of microbial assemblages and influencing mammalian immunity and physiology. Gut bacteriophages were shown to influence the dynamics of bacterial populations in germ-free mice that were colonized by a defined microbial community composed of 15 human symbionts (*40*). Bacteriophages not only target specific bacteria but also affect nonhost bacterial species through indirect effects on interbacterial interactions. Fecal metabolomics has shown that shifts in gut microbial communities caused by phage predation can have a direct impact on gut metabolites (*41*). Bacteriophages can also interact directly with the mammalian immune system because phage DNA activates interferon γ (IFN-γ) responses through the Toll-like receptor 9–dependent pathway. This interaction has been shown to exacerbate colitis. Mucosal IFN-γ responses in ulcerative colitis (UC) patients positively correlate with bacteriophage abundance (*42*).

One study has shown some early promise in the use of bacteriophages to ameliorate alcoholic liver disease. Cytolysin, an exotoxin secreted by *Enterococcus faecalis*, is strongly correlated with severity of liver disease and mortality in patients with alcoholic hepatitis. Ethanol-induced gut barrier damage likely enables translocation of cytolytic *E. faecalis* from the intestine to the liver. In gnotobiotic mice colonized with feces from patients with alcoholic hepatitis, gavage of bacteriophages targeting cytolytic *E. faecalis* effectively decreases cytolysin concentrations in the liver and reduces

the severity of ethanol-induced liver injury and steatosis (*43*).

Host-microbiota interaction through metabolites and enzymes

In addition to direct and local communication at host mucosal barriers, the microbiota produces metabolites that can communicate across host mucosae to influence host physiology systemically. Metabolites associated with various organs and diseases have been cataloged and used in functional screening (Fig. 2) (*44*). For example, gut microbe–produced metabolites that are derived from dietary tryptophan target the mammalian aryl hydrocarbon receptor (AHR) to inhibit activation of the actin regulatory proteins MyoIIA and ezrin and thereby locally strengthen intestinal barrier integrity that is mediated by tight junctions and adherens junctions (*45*). Tryptophan metabolites also remotely function as AHR agonists in astrocytes to mitigate central nervous system inflammation by potentiating the type I interferon response (*46*). Certain bacterial species in the gut ferment nondigestible dietary fibers into SCFAs—such as acetate, propionate, and butyrate—to mediate host energy balance, immune modulation, and mucosal barrier function (*47*). Butyrate locally induces the differentiation of T_{regs} in the colonic mucosa to maintain gut immune homeostasis (*48*) while modulating asthma phenotypes in the lung (*49*). These discoveries indicate that SCFAs function in a cell type–dependent manner and that local concentrations of SCFAs are important.

Another clear example of the influence that the microbiota can have on host metabolism has emerged from an untargeted mass spectrometry–based metabolomics analysis of plasma from healthy patients and patients with cardiovascular disease (Fig. 2). The disease group had increased concentrations of trimethylamine *N*-oxide (TMAO), which is produced in the liver from trimethylamine (TMA), a microbiota-derived metabolite generated from dietary L-carnitine and choline. TMAO alters host cholesterol and steroid metabolism and increases the risk for cardiovascular diseases when its concentrations are increased in plasma (*50*). A proof-of-concept study showed that inhibiting a major microbial TMA-generating enzyme (CutC/D) by a small molecule significantly reduced plasma concentrations of TMAO and effectively reduced the potential for thrombosis (*51*).

Similar to TMAO, secondary bile acids are gut microbiota–liver axis metabolites. Bile acids are a class of cholesterol derivatives synthesized in the liver (as primary bile acids) that can be modified by gut microbes into secondary bile acids, including deoxycholic acid (DCA). Bile acids have been implicated in several host-microbiota effects on host physiology. For example, at higher concentrations, DCA binds one of its receptors,

Human studies

Multiomics of the human microbiome

New hypothesis

New hypothesis

Hypothesis generation

In vitro studies

In vivo studies
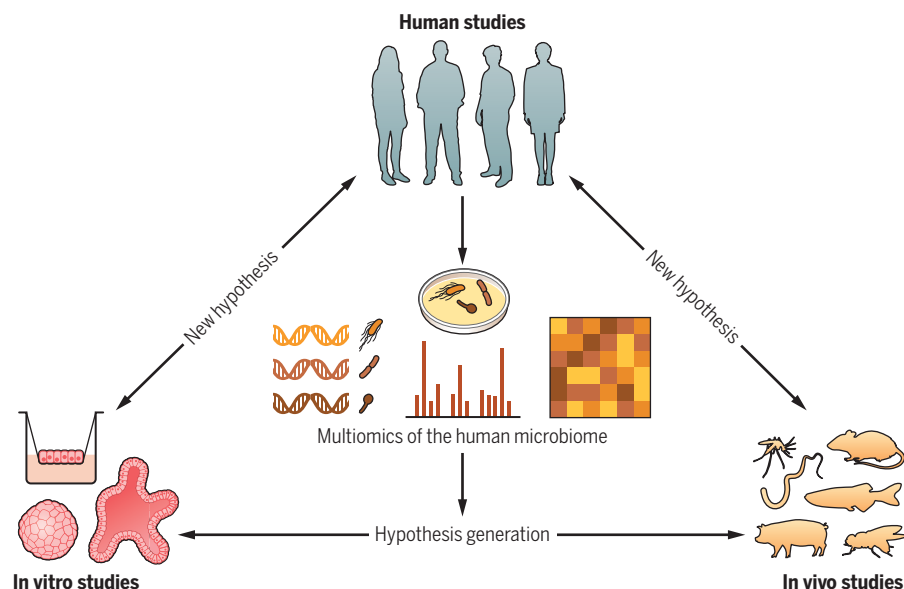
Fig. 3. Mechanistic microbiome studies with integrated approaches to find disease targets.
Approaches of microbiome studies have been established by integrating experimental and computational methodologies. The hypotheses generated from multi-omics (culturomics, metagenomics, proteomics, metabolomics, etc.) analyses of the human microbiome are tested in both in vitro and in vivo models, such as organoids, air-liquid interface cultures, and multiple animal models. New hypotheses can be generated again from in vitro or in vivo studies. In the next phase of microbiome research, we propose that there should be a major focus on mechanism-based studies to predict microbe-based therapeutic interventions.

Farnesoid X receptor (FXR), to inhibit prostaglandin E2 synthesis and promote colonic crypt regeneration after intestinal mucosal wounding (52). Secondary bile acids can mediate responses to alphavirus infection by promoting systemic type I interferon responses (53) and induce proliferation and DNA damage in intestinal cancer stem cells through FXR (54). A recent study showed that microbial enzymes abundant in the feces of patients with UC are involved in driving colitis severity upon transplant into germ-free mice (55). Together these studies indicate that microbial products can extensively modulate host physiology locally and systemically, making them attractive targets for therapeutic interventions in metabolic disorders.

Technical limitations and new avenues

Genomics approaches that infer differential abundances of microbes by analyzing bacterial 16S rRNA genes or internal transcribed spacers (ITS) of fungi can achieve genus-level sensitivity, but these approaches do not allow for differentiation between closely related species or strains. 16S rRNA gene and ITS sequencing analyses are useful for detecting high-abundance microbes, but differentially abundant microbes may prove to be only passengers, not the drivers, of a given host phenotype. Functional, low-abundance microbes that might drive a phenotype could be missed. It is therefore essential that sequencing studies are supported by experimental functional analyses in animal models and microbial isolates

from hosts with phenotypes of interest. One study showed that various *Bacteroides ovatus* isolates from human stool induced the production of strain-specific SIgA in monocolonized mice, highlighting that variable mucosal immune responses might be caused by variable microbial strains (56). A new study has also revealed genetic diversity among opportunistic *C. albicans* strains isolated from patients with UC. Strains that secrete the cell-damaging toxin candidalysin trigger intestinal inflammation when colonized in mice, reflecting disease features of individual UC patients and underscoring microbial strain–specific effects on immune-mediated inflammatory diseases (57).

One way to avoid selective vision in understanding the proximal causes of microbiota-mediated disease is to use shotgun metagenomics of entire microbiomes. This technique is culture independent, retrieves all the genomes from a microbial community, and can be used not only for taxonomic purposes but also for discovering previously uncharacterized genes, enzymes, and metabolites. Catalogs of both mouse and human gut metagenomes are now available for mining (58, 59). However, despite technical improvements, many challenges remain, especially for the assignment of genes to function or to specific microbes.

Pure culture of microorganisms remains the standard method for testing functional roles in experimental models and for testing therapeutic possibilities. In recent years, the number of microbes that can now be cultured from

both mouse and human sources has greatly expanded (60). Integration of data from culture collections with genomes, transcriptomes, and metabolomes has now made it possible to characterize the behavior and natural products of cultured and uncultured microbiota to a high level of resolution (Fig. 3) (61).

A long-standing problem in understanding the full contribution of the microbiota to host physiology is the current focus on bacteria, although increasingly more attention is being focused on fungi. Viruses are challenging to identify if they are not persistent, and most viral studies have focused on bacteriophages. Macroparasites have been largely ignored, partially owing to a lack of tools, although several studies show that host physiology and immunity can be profoundly affected by protozoan and helminth parasites in several mammalian body sites.

To date, most experiments in functional microbiome research tend to be gain-of-function studies. Loss-of-function studies at the strain or gene level can precisely define the role of a microbe or a gene in a disease model. The activity of bacteriophages shows how reducing the load of a specific pathogenic bacterium in the gut can be effective in attenuating disease (43), but the possibilities for therapeutic exploitation of phages is complicated by evolutionary and ecological considerations. Studies have shown that CRISPR-Cas9–based genome editing can be used on a complex microbiome for site-specific gene insertions and deletions of targeted members in a microbial community (62, 63). These studies show exiting potential for experimental manipulation of microbial function within a community without perturbing the whole population and without any need to culture organisms.

The laboratory mouse is the most popular animal model for microbiome research, and much of what we understand at present about the mammalian microbiota has emerged from these models. Nevertheless, microbiome studies in mice can fail in translation to humans (64). Alternative models are being developed; for example, the worm *Caenorhabditis elegans*, fruit fly *Drosophila melanogaster*, zebrafish *Danio rerio*, hydra *Hydra vulgaris*, mosquito *Anopheles gambiae* (65), and wild pig *Sus scrofa* (66) have all been used for addressing specific questions (Fig. 3). Despite these advances, a substantial challenge in microbiome research remains the heterogeneity of human microbiota, both in terms of density and of composition among individuals. Interpersonal heterogeneity may underpin variation in a range of attributes, from responses to antibiotics and metabolic profile to susceptibility and resistance to infection (67), as well as outcomes of cancer immunotherapy (68). In the future, the integration of multiple animal models together with in vitro approaches

on human cells and tissues, such as organoids (*69*) and air-liquid interface culture systems (*70*), will aid investigation of causalities between and the microbiome and human diseases (Fig. 3).

Conclusions

Microbiome research at host barriers has changed the dominant concept that microbes are generally harmful into a more nuanced understanding of the importance of resident microbes to overall host health. Mounting evidence for the role that perturbations in the microbiota play in systemic human disease has generated great interest in finding microbiota-based therapies. Despite the far greater understanding that we now have of the pivotal role that the microbiota of an organism may have in maintaining health, new approaches are still required to fully understand the complexities of temporal and spatial dynamics of the microbiota, the interkingdom interactions within microbial communities, and the constant exchange of signals between the microbiota and its host.

REFERENCES AND NOTES

1. F. Bishehsari, R. M. Voigt, A. Keshavarzian, *Nat. Rev. Endocrinol.* **16**, 731–739 (2020).
2. R. X. Wang, J. S. Lee, E. L. Campbell, S. P. Colgan, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 11648–11657 (2020).
3. A. Abot, S. Fried, P. D. Cani, C. Knauf, *Antioxid. Redox Signal.* ars.2021.0100 (2021).
4. L. V. Hooper, D. R. Littman, A. J. Macpherson, *Science* **336**, 1268–1273 (2012).
5. G. C. Hansson, *Annu. Rev. Biochem.* **89**, 769–793 (2020).
6. A. N. Skelly, Y. Sato, S. Kearney, K. Honda, *Nat. Rev. Immunol.* **19**, 305–323 (2019).
7. B. Kim *et al.*, *eLife* **8**, e45343 (2019).
8. V. A. Pedicord *et al.*, *Sci. Immunol.* **1**, eaai7732 (2016).
9. C. Petersen, J. L. Round, *Cell. Microbiol.* **16**, 1024–1033 (2014).
10. E. Mathieu *et al.*, *Front. Physiol.* **9**, 1168 (2018).
11. A. L. Kau *et al.*, *Sci. Transl. Med.* **7**, 276ra24 (2015).
12. R. Y. Chen *et al.*, *N. Engl. J. Med.* **383**, 321–333 (2020).
13. D. Paredes-Sabja, A. Shen, J. A. Sorg, *Trends Microbiol.* **22**, 406–416 (2014).
14. D. R. Plichta, D. B. Graham, S. Subramanian, R. J. Xavier, *Cell* **178**, 1041–1056 (2019).
15. P. Feuerstadt *et al.*, *N. Engl. J. Med.* **386**, 220–229 (2022).
16. I. I. Ivanov *et al.*, *Cell* **139**, 485–498 (2009).
17. L. Cervantes-Barragan *et al.*, *Science* **357**, 806–810 (2017).
18. S. Naik *et al.*, *Nature* **520**, 104–108 (2015).
19. C. G. Buffie *et al.*, *Nature* **517**, 205–208 (2015).
20. T. Tanoue *et al.*, *Nature* **565**, 600–605 (2019).
21. K. Atarashi *et al.*, *Nature* **500**, 232–236 (2013).
22. S. Caballero *et al.*, *Cell Host Microbe* **21**, 592–602.e4 (2017).
23. P. Vonaesch, M. Anderson, P. J. Sansonetti, *FEMS Microbiol. Rev.* **42**, 273–292 (2018).
24. J. A. Segre, *J. Invest. Dermatol.* **133**, 2141–2142 (2013).
25. S. M. Bloom *et al.*, *Cell Host Microbe* **9**, 390–403 (2011).
26. C. S. Chang *et al.*, *Cell Rep.* **37**, 110016 (2021).
27. K. Cadwell *et al.*, *Cell* **141**, 1135–1145 (2010).
28. U. Jain *et al.*, *Science* **371**, 1154–1159 (2021).
29. S. Falkow, *Rev. Infect. Dis.* **10**, S274–S276 (1988).
30. D. L. Moyes *et al.*, *Nature* **532**, 64–68 (2016).
31. J. Chow, H. Tang, S. K. Mazmanian, *Curr. Opin. Immunol.* **23**, 473–480 (2011).
32. D. Fan *et al.*, *Nat. Med.* **21**, 808–814 (2015).
33. R. C. Boutin *et al.*, *eLife* **10**, e67740 (2021).
34. J. H. Skalski *et al.*, *PLOS Pathog.* **14**, e1007260 (2018).
35. R. N. Fichorova *et al.*, *PLOS ONE* **7**, e48418 (2012).
36. A. Chudnovskiy *et al.*, *Cell* **167**, 444–456.e14 (2016).
37. P. Desai *et al.*, *Cell* **184**, 1214–1231.e16 (2021).
38. M. M. Zaiss *et al.*, *Immunity* **43**, 998–1010 (2015).
39. Z. Hu *et al.*, *Science* **374**, eabe6723 (2021).
40. A. Reyes, M. Wu, N. P. McNulty, F. L. Rohwer, J. I. Gordon, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 20236–20241 (2013).
41. B. B. Hsu *et al.*, *Cell Host Microbe* **25**, 803–814.e5 (2019).
42. L. Gogokhia *et al.*, *Cell Host Microbe* **25**, 285–299.e8 (2019).
43. Y. Duan *et al.*, *Nature* **575**, 505–511 (2019).
44. E. A. Franzosa *et al.*, *Nat. Microbiol.* **4**, 293–305 (2019).
45. S. A. Scott, J. Fu, P. V. Chang, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 19376–19387 (2020).
46. V. Rothhammer *et al.*, *Nat. Med.* **22**, 586–597 (2016).
47. B. van der Hee, J. M. Wells, *Trends Microbiol.* **29**, 700–712 (2021).
48. Y. Furusawa *et al.*, *Nature* **504**, 446–450 (2013).
49. A. N. Thorburn *et al.*, *Nat. Commun.* **6**, 7320 (2015).
50. W. H. Tang *et al.*, *N. Engl. J. Med.* **368**, 1575–1584 (2013).
51. A. B. Roberts *et al.*, *Nat. Med.* **24**, 1407–1417 (2018).
52. U. Jain *et al.*, *Cell Host Microbe* **24**, 353–363.e5 (2018).
53. E. S. Winkler *et al.*, *Cell* **182**, 901–918.e18 (2020).
54. T. Fu *et al.*, *Cell* **176**, 1098–1112.e18 (2019).
55. R. H. Mills *et al.*, *Nat. Microbiol.* **7**, 262–276 (2022).
56. C. Yang *et al.*, *Cell Host Microbe* **27**, 467–475.e6 (2020).
57. X. V. Li *et al.*, *Nature* **603**, 672–678 (2022).
58. A. Almeida *et al.*, *Nat. Biotechnol.* **39**, 105–114 (2021).
59. L. Xiao *et al.*, *Nat. Biotechnol.* **33**, 1103–1108 (2015).
60. J. C. Lagier *et al.*, *Nat. Microbiol.* **1**, 16203 (2016).
61. J. Lloyd-Price *et al.*, *Nature* **569**, 655–662 (2019).
62. K. N. Lam *et al.*, *Cell Rep.* **37**, 109930 (2021).
63. B. E. Rubin *et al.*, *Nat. Microbiol.* **7**, 34–47 (2022).
64. T. L. Nguyen, S. Vieira-Silva, A. Liston, J. Raes, *Dis. Model. Mech.* **8**, 1–16 (2015).
65. A. E. Douglas, *Nat. Rev. Microbiol.* **17**, 764–775 (2019).
66. H. W. Chang *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2024446118 (2021).
67. S. Alavi *et al.*, *Cell* **181**, 1533–1546.e13 (2020).
68. N. L. Ting, H. C. Lau, J. Yu, *Gut* gutjnl-2021-326264 (2022).
69. V. Bozzetti, S. Senger, *Trends Mol. Med.* **28**, 290–303 (2022).
70. G. Wilke *et al.*, *Cell Host Microbe* **26**, 123–134.e8 (2019).

Mental health care 2.0

While the COVID-19 pandemic has been a burden for our mental health, it also led to a surge in mental health care innovation. Appointments by telephone or video, as well as web- and app-based tools, have become part of a digital mental health revolution. Last year, US venture capitalists invested $5.1 billion in this area, a fivefold increase from 2019. But is this surge in activity actually leading to improvements for those with the greatest needs?

The shift to remote care addressed a pain point of the brick-and-mortar mental health system: the gap between the limited supply of providers and the increasing demand for care. Digital mental health companies made it easier to access a mental health professional, and the advent of self-help apps for mindfulness and sleep became popular as consumer products. Better access to therapy and consumer apps are undeniable improvements that could be helpful to millions.

But are these new tools serving those who need care the most? Most of these innovations do little for people suffering from a serious mental illness, such as schizophrenia and bipolar disorder. Many with serious mental illness are unemployed, too poor to pay for online therapy, and, often, on the wrong side of the digital divide. They may not be seeking online therapy and, to be clear, there is no app for the range of social and medical problems they face. They endure homelessness, incarceration, and loneliness. Can there be innovations to engage them with health care and foster recovery?

To recover, people with serious mental illness need people, place, and purpose—the "3 P's"—usually provided through supports for social engagement, housing, and employment. During the pandemic, many nonprofit providers began delivering these psychosocial supports through tablets and phones. In contrast to the traditional delivery of these services, tech-enabled services proved surprisingly efficient and scalable, demonstrating the opportunity for innovation to promote recovery. But we are just beginning to see the impact of high tech linked to high touch for people with serious mental illness.

Why have investors and entrepreneurs neglected this need? In the United States, there are 14.2 million people with serious mental illness, each representing, on average, a $1.3 million lifetime economic burden on society.

In contrast to medical and surgical costs, more mental health care is covered by federal public insurance ($110 billion) than by commercial or private insurance ($62 billion). Entrepreneurs and venture capital investors ought to be lining up at the gates of the public insurance system. Except there are no such gates.

The problem is that the health care safety net does not easily accommodate companies that want to try new ways to serve these patients, either in-person or digitally. In many states, this kind of care is carved out from other government programs and relegated to county agencies. In states where commercial insurers are involved, they tend to pay for crisis care, brief hospitalization, and medication management, but not for engagement—encouraging people to enter the system—and recovery. Although nonprofit organizations and government agencies have issued grants and launched projects to stoke innovative approaches, this temporary funding is not sufficient to attract startups that depend on a sustainable and scalable source of revenue.

> "Most of these innovations do little for people suffering from a serious mental illness..."

Fortunately, there are some recent changes that promise to better align the incentives of innovators with the greatest needs in mental public health. Payers and providers are focusing on social determinants of health such as homelessness and poverty. There is increasing awareness of the need for "whole person care" and recovery services. Most important, public and private payers are experimenting with reimbursement based on a commitment to care for an entire population, with payment dependent on outcomes. This "value-based" approach encourages entrepreneurs to take risks in developing new ways to help the safety-net population, knowing that they can profit if they can achieve better outcomes at lower costs.

What is ultimately needed is a new kind of public-private partnership to expand the scope of care and access to these resources. With Operation Warp Speed, we saw the potential of a partnership between government and the pharmaceutical industry to accelerate the development of a COVID-19 vaccine. For mental health, a similar partnership is needed between government and the digital health industry. This partnership could focus entrepreneurs on the "deep end of the pool" while helping government agencies to serve those with the greatest needs.

–Thomas Insel*

Thomas Insel is an adjunct professor in the Department of Psychiatry and Behavioral Sciences at Stanford University School of Medicine, Stanford, CA, USA, and is the former director of the US National Institute of Mental Health, Bethesda, MD, USA. tinsel0@gmail.com

> **" I'm horrified. … It's a complete and total distortion in so many ways. "**
> Behavioral economist Daniel Benjamin of the University of California, Los Angeles, to *STAT*, on the Buffalo, New York shooting suspect's citation of Benjamin's genetics research in his racist manifesto.

ANIMAL WELFARE

Judge suspends dog facility

A federal judge has issued a temporary restraining order against a major supplier of beagles for research. The action—directed at a Cumberland, Virginia, breeding facility run by Envigo—is in response to an animal welfare complaint filed last week by the U.S. Department of Justice (DOJ), which has now removed 145 puppies from the operation. Last year, U.S. Department of Agriculture inspectors found dozens of wounded, sick, and suffering animals at the facility, and records of hundreds of uninvestigated puppy deaths, prompting a crackdown by Virginia lawmakers. "[T]he Government has provided sufficient evidence that Envigo is engaged in serious and ongoing violations of the Animal Welfare Act," the judge's 21 May order states. The operation must "immediately cease breeding, selling, or otherwise dealing in beagles." Envigo denies the allegations and says it is "fully cooperating with DOJ and other involved authorities."

Will young kids get mRNA shots?

COVID-19 | More data suggest the COVID-19 vaccines made with messenger RNA (mRNA) can safely protect young children, Pfizer and its partner BioNTech said in a press release this week. Three low doses of their vaccine given to 1678 children who were at least 6 months old but younger than 5 years old led to a "strong immune response" that mirrors what has been seen in 16- to 25-year-olds who got two shots with a dose 10 times higher. The ongoing study only observed 10 illnesses, leading to a preliminary estimate of 80% efficacy against disease. Moderna earlier reported positive immune responses but a much lower efficacy against disease in a group of young children who received two shots of its mRNA vaccine given at one-fourth the dose used in adults. Both companies have asked for emergency use authorization from the U.S. Food and Drug Administration, which will convene its vaccine advisory committee to discuss the data on 15 June.

'Good faith' hackers protected

CYBERSECURITY | Researchers probing a website's security flaws in "good faith" will no longer face the risk of being prosecuted for hacking under a long-standing U.S. law, according to a new policy released last week by the Department of Justice. Research into the vulnerabilities of corporate or government sites or devices has long carried the possibility of violating the 1986 Computer Fraud and Abuse Act. But the 19 May policy revision says violating the access restrictions in a website's terms of service, for example, does not in itself warrant criminal charges, as long as federal prosecutors conclude the research was "designed to avoid any harm to individuals or the public" and not to extort the owners of a site or device.

Australia pledges climate action

CLIMATE POLICY | The Australian Labor Party's ambitious plan to cut 2005 carbon emissions 43% by 2030 helped propel it to victory in the country's 21 May elections. That target is significantly more ambitious than the 26% to 28% cut envisioned by the Liberal-National Coalition that had held power since 2013. Critics had derided the Liberal-National target as woefully insufficient. Labor's energy plan calls for increasing reliance on renewables, cutting taxes on electric vehicles, modernizing the electricity grid, and requiring the public sector to reach net zero emissions by 2030. Support for the party grew the most in parliamentary districts that had suffered recent bushfires, floods, and droughts, according to an analysis by the Climate Council, an environmental organization. "Climate action is the winner of this election," the council said.

Paleontologist statue unveiled

HISTORY OF SCIENCE | The paleontology pioneer Mary Anning finally has a statue honoring her contributions to the understanding of Earth's past. On 21 May, organizers officially unveiled the statue in Lyme Regis on southern England's Jurassic Coast. Anning collected fossils

A new life-size bronze statue of Mary Anning drew a crowd in Lyme Regis, England.

from the area's rich deposits in the early 1800s, identifying the first known ichthyosaur and England's first pterosaur. Her contributions were long overlooked, but in 2018, local resident Evie Swire, then 10 years old, and her mother Anya Pearson launched the "Mary Anning Rocks" project to raise funds for a statue to commemorate Anning. Recognition for Anning grew after the 2020 release of the film *Ammonite*, starring Kate Winslet. Swire's project eventually raised more than £100,000 for the statue, which depicts Anning with hammer and fossil in hand, overlooking the cliffs that still attract fossil hunters today.

Accuser loses defamation suit

#METOO | A judge in Peru this week ruled against archeologist Marcela Poirier in a defamation lawsuit brought by Luis Jaime Castillo Butters, a prominent Peruvian archaeologist she accused of sexual harassment. Poirier was sentenced to a $48,400 fine and a suspended jail sentence of 1 year and 8 months as a result of allegations she made on Facebook. A 2020 investigation by Castillo Butters's institution, the Pontifical Catholic University of Peru, found evidence of sexual harassment but didn't sanction him, saying the harassment predated formal rules prohibiting it. The National Academy of Sciences ousted him in 2021 after Poirier filed a complaint with the organization. She plans to appeal the judge's decision, which her supporters say is a blow to the #MeToo movement in Latin America.

NIH sets up antiviral centers

DRUG DEVELOPMENT | The National Institute of Allergy and Infectious Diseases (NIAID) last week awarded $577 million to establish nine U.S. antiviral drug discovery centers. The centers, all led by academic institutions, will focus on creating antivirals targeting SARS-CoV-2 and other viruses with pandemic potential. Targets could include: filoviruses such as Ebola and Marburg; flaviviruses, which cause yellow fever and Zika; and mosquito-born togaviruses. The new effort aims to identify novel molecular targets essential to viruses' ability to attack the body and then find small-molecule oral medications that directly block these targets, NIAID's director, Anthony Fauci, said in an 18 May press release. Recipients include the Scripps Research Institute; the University of North Carolina, Chapel Hill; Emory University; and the University of California, San Francisco.

ECOLOGY
Is a tree in Chile the world's oldest?

A famous conifer in Chile could be the world's oldest tree, according to a scientist's claim that has drawn skepticism from others. Alerce trees are long-lived, with one documented stump holding 3622 annual growth rings. That ranks second to bristlecone pines, whose oldest known living specimen resides in California and has 4853 rings. But a towering alerce in Alerce Costero National Park could be about 5500 years old, estimates environmental scientist Jonathan Barichivich of the Climate and Environmental Sciences Laboratory in Paris. In unpublished work, he paired a partial count of the tree's growth rings with statistical modeling to calculate that the alerce has an 80% probability of being more than 5000 years old. Some tree-age experts say the method seems promising, but others insist the only way to accurately date a tree is to count every ring.
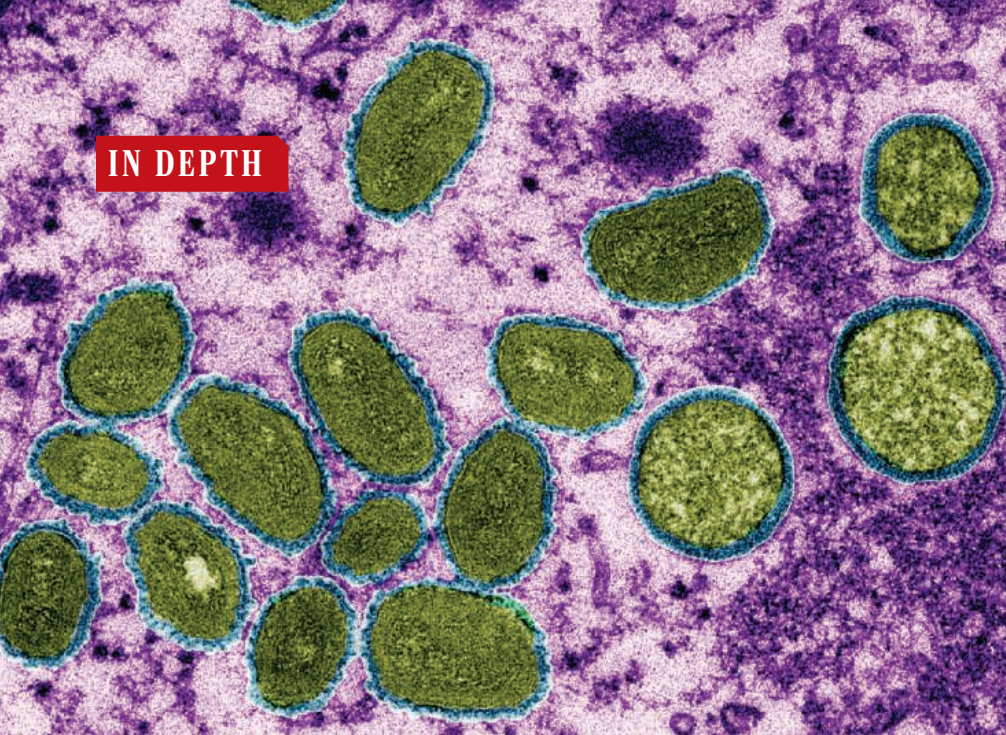
The monkeypox virus, shown here in a colored electron micrograph, is related to the smallpox virus.

GLOBAL HEALTH

Monkeypox outbreak questions intensify as cases soar

Rapid emergence of hundreds of cases around the world alarms public health officials and scientists

By **Jon Cohen**

With the COVID-19 pandemic still raging, a second public health threat has the world on high alert: a global outbreak of monkeypox. As *Science* went to press, two dozen countries in the Americas, Europe, North Africa, the Middle East, and Australia had reported nearly 300 suspected and confirmed cases of the disease, a much milder cousin of smallpox, the deadly scourge that the World Health Organization (WHO) deemed eradicated in 1980. A disproportionate number of cases are in men who have sex with men (MSM), an unusual twist given that researchers have never convincingly shown sexual transmission occurs in monkeypox.

Monkeypox is endemic in West and Central Africa, and the virus occasionally causes outbreaks elsewhere in the world, but most are quickly contained or peter out by themselves. This outbreak looks very different, says Anne Rimoin, an epidemiologist at the University of California, Los Angeles, who has long studied monkeypox in the Democratic Republic of the Congo (DRC). "We get concerned when we see a virus doing things that we don't normally see it doing," Rimoin

says. "We now have to really reevaluate what we know about monkeypox—which has all been gleaned, essentially, from low-resource settings in Africa—in very, very different kinds of populations."

"We should definitely be concerned about this new situation," agrees epidemiologist Rosamund Lewis, head of poxvirus diseases at WHO. Still, Lewis stresses that monkeypox, which causes skin lesions resembling those of smallpox, is not easily transmitted, and its spread can typically be limited by isolating cases. "It's not a concern for most

Scabs cover the skin lesions developed by two people who had monkeypox in the United States in 2003.

people," she says. What's more, several drugs and vaccines developed for smallpox can also treat and prevent monkeypox.

A WHO technical advisory group that specializes in infectious hazards, chaired by epidemiologist and monkeypox veteran David Heymann of the London School of Hygiene & Tropical Medicine, met last week to develop recommendations covering everything from the need for more aggressive surveillance to the use of vaccines. In a 23 May rapid risk assessment, the European Centre for Disease Prevention and Control warned there's a "potential risk" for human-to-animal transmission, which could enable the virus to become established in European wildlife, as it is in Africa.

U.K. officials on 7 May were the first to report a monkeypox case, a patient who had arrived from Nigeria 3 days earlier. But that person apparently did not transmit the virus to anyone else and is not the source of the wider outbreak; public health officials in Montreal say a man there had symptoms as early as 29 April. At the time, he wasn't tested for monkeypox. Most doctors would not think to do so because the disease is vanishingly rare outside of Africa and its lesions can resemble those of chickenpox and syphilis.

Given the torrent of cases reported over the past 2 weeks, it's possible the virus has been spreading under the radar for months, Heymann says. Rimoin says a much larger outbreak could be underway in Nigeria and nearby countries, from which the disease might have been exported repeatedly to the rest of the world. "We're not looking for these kinds of things, and if you don't look for them, you don't find them," she says.

"MONKEYPOX" IS A MISNOMER; the virus was discovered in 1958 in research monkeys, but its natural hosts are likely rodents and other small mammals. The virus first surfaced in humans in 1970 in what is now the DRC, causing fever, headaches, and lymph node swelling followed by an eruption of pus-filled blisters. Outbreaks start when someone comes in contact with an infected wild animal. The biggest known outbreak in the United States occurred in 2003, when 47 people were infected by pet prairie dogs that had picked up the virus from rodents imported from Ghana.

Most people recover within a few weeks. The Congo Basin strain kills up to 10% of those infected, but the current outbreak appears to only involve the West African strain, which in past outbreaks had a fatality rate of about 1%. Outbreaks "generally fizzle out on their own," Lewis notes, be-

Corrected 26 and 31 May 2022. See full text.

cause many infected people simply never infect anyone else.

Monkeypox usually spreads through close contact and respiratory droplets, but in 2017 Nigerian researchers suggested sexual transmission might have occurred in several patients with genital ulcers. Those are being seen now as well, says Fernando Simón, who directs the Spanish Ministry of Health's coordination center for health alerts and emergencies. Among the first confirmed reported cases, "most have lesions exclusively perigenital, perianal, and around the mouth," Simón says. Some of the first cases detected in Spain are MSM or transgender people who had attended a sauna in Madrid or a gay pride festival in the Canary Islands. In Belgium, several cases were linked to a gay festival in Antwerp. No studies have ever found that semen can transmit the virus. "So far, the most acceptable hypothesis is that it is transmitted after contact with lesions," Simón says. MSM and transgender communities have a high prevalence of HIV, but there's no evidence that compromised immune systems have played any role in this outbreak.

Public health officials face a tricky balance as they seek to clearly communicate the risks to the MSM community while avoiding stigmatization. "Stigma and blame undermine trust and capacity to respond effectively during outbreaks like this one," Matthew Kavanagh, deputy executive director of the Joint United Nations Programme on HIV/AIDS, said in a statement this week.

Researchers have begun sequencing viral samples, which can help trace how the virus has spread and reveal mutations that might make it more pathogenic or more transmissible. It's a laborious job: With more than 200,000 basepairs, the monkeypox genome is about seven times the size of SARS-CoV-2's and more than 20 times larger than HIV's. The first full genome, posted on 19 May by a team led by João Paulo Gomes at Portugal's National Institute of Health, showed that the strain most closely resembles viruses carried by travelers from Nigeria to Singapore, Israel, and the United Kingdom in 2018 and 2019. Sequences posted by the U.S. Centers for Disease Control and Prevention (CDC) and scientists in Belgium and Germany support that conclusion. "We're not seeing any evidence that this virus has changed to become more efficiently transmitted person to person," Jennifer McQuiston, the head of CDC's pox virus and rabies branch, said at a 23 May press conference.

Because it's a DNA virus, monkeypox has far better genetic repair mechanisms than RNA viruses such as HIV and SARS-CoV-2, which means it changes more slowly. Still, studies by Gustavo Palacios, a virologist at the Icahn School of Medicine at Mount Sinai and co-workers suggest mutations are frequent enough for investigators to create virus family trees that can help clarify how cases connect to each other.

STOPPING THE OUTBREAK will require educating people and finding cases early to reduce the chance they transmit the virus. Vaccination could play a role as well. The United Kingdom is already offering vaccines to health care workers and other contacts of known cases—a strategy called ring vaccination, which was key to the success of the smallpox eradication campaign. The U.S. has begun offering vaccination to some contacts but is undecided about health care workers caring for monkeypox patients. Rimoin says ring vaccination doesn't seem warranted yet but that might change "if in the next couple of days … we see that there are hundreds and hundreds of cases."

One of the two available vaccines, manufactured by Emergent BioSolutions, was designed as a smallpox vaccine and can protect even after exposure to monkeypox virus, when given within 4 days. But it contains a live virus called vaccinia that replicates inside the body; the similar vaccine used during smallpox eradication sometimes caused severe side effects, killing one in 1 million recipients.

The other vaccine, from Bavarian Nordic—which the United Kingdom is using—contains a nonreplicating form of vaccinia designed to cause fewer side effects. But it requires two doses, given 4 weeks apart. Animal studies suggest the first dose works faster than Emergent BioSolutions' vaccine; the second dose increases the durability of protection. Both vaccines are in short supply and typically only available via national stockpiles.

Drugs exist as well. One, tecovirimat, in 2018 became the first ever approved by the U.S. Food and Drug Administration (FDA) to treat smallpox after it proved safe in human trials and effective in animals given closely related viruses. FDA approved a second drug for smallpox, brincidofovir, last year. Although neither has been approved for monkeypox, both showed promising results against the virus in animal studies.

But a report published on 24 May in *The Lancet Infectious Diseases* reported no evidence that brincidofovir had "any convincing clinical benefit" in three patients treated in the United Kingdom over the past 3 years and said the drug had serious toxicities. Tecovirimat, in contrast, did not cause side effects and appeared to lead to a decrease in viral levels and a speedier recovery.

How far and fast the virus spreads in the coming days will determine how aggressively clinicians put these tools to the test. "These are very early days," Rimoin says. ∎

Japan tries—again—to revitalize its research

Latest effort would spend billions on a few universities, but skeptics give it long odds

By **Dennis Normile**

Alarmed by the declining stature of its universities, Japan is planning to shower up to $2.3 billion a year on a handful of schools in hopes of boosting their prominence.

The scheme was approved by the Japanese legislature on 18 May, although many details, including how to pick the favored universities, are still up in the air. But the move, under study for more than a year, has rekindled a debate among academics over how to reverse Japan's sinking research fortunes. Several previous schemes have yielded mixed results.

The new plan "aims to provide young promising scholars with the research environment that the world's top universities are supposed to offer, to dramatically enhance international collaborations, and to promote the brain circulation both domestically and internationally," says Takahiro Ueyama, a science policy specialist on the Council for Science, Technology and Innovation (CSTI), Japan's highest science advisory body, which was heavily involved in crafting the scheme.

But Guojun Sheng, a Chinese developmental biologist at Kumamoto University in Japan, is skeptical. "I am not very optimistic that this [plan] will do much to curb the slide in the ranking of Japanese research activities or international competitiveness," he says. Sheng, who previously studied and worked in China, the United States, and the United Kingdom, says the new plan does not address fundamental problems at Japanese research institutes: too few women and foreign scientists, a fear of change, and lack of support for young scientists. To get better results, "Japan has to change its research culture," he says.

Corrected 26 and 31 May 2022. See full text.

cause many infected people simply never infect anyone else.

Monkeypox usually spreads through close contact and respiratory droplets, but in 2017 Nigerian researchers suggested sexual transmission might have occurred in several patients with genital ulcers. Those are being seen now as well, says Fernando Simón, who directs the Spanish Ministry of Health's coordination center for health alerts and emergencies. Among the first confirmed reported cases, "most have lesions exclusively perigenital, perianal, and around the mouth," Simón says. Some of the first cases detected in Spain are MSM or transgender people who had attended a sauna in Madrid or a gay pride festival in the Canary Islands. In Belgium, several cases were linked to a gay festival in Antwerp. No studies have ever found that semen can transmit the virus. "So far, the most acceptable hypothesis is that it is transmitted after contact with lesions," Simón says. MSM and transgender communities have a high prevalence of HIV, but there's no evidence that compromised immune systems have played any role in this outbreak.

Public health officials face a tricky balance as they seek to clearly communicate the risks to the MSM community while avoiding stigmatization. "Stigma and blame undermine trust and capacity to respond effectively during outbreaks like this one," Matthew Kavanagh, deputy executive director of the Joint United Nations Programme on HIV/AIDS, said in a statement this week.

Researchers have begun sequencing viral samples, which can help trace how the virus has spread and reveal mutations that might make it more pathogenic or more transmissible. It's a laborious job: With more than 200,000 basepairs, the monkeypox genome is about seven times the size of SARS-CoV-2's and more than 20 times larger than HIV's. The first full genome, posted on 19 May by a team led by João Paulo Gomes at Portugal's National Institute of Health, showed that the strain most closely resembles viruses carried by travelers from Nigeria to Singapore, Israel, and the United States in 2018 and 2019. Sequences posted by the U.S. Centers for Disease Control and Prevention (CDC) and scientists in Belgium and Germany support that conclusion. "We're not seeing any evidence that this virus has changed to become more efficiently transmitted person to person," Jennifer McQuiston, the head of CDC's pox virus and rabies branch, said at a 23 May press conference.

Because it's a DNA virus, monkeypox has far better genetic repair mechanisms than RNA viruses such as HIV and SARS-CoV-2, which means it changes more slowly. Still, studies by Gustavo Palacios, a virologist at the Icahn School of Medicine at Mount Sinai and co-workers suggest mutations are frequent enough for investigators to create virus family trees that can help clarify how cases connect to each other.

STOPPING THE OUTBREAK will require educating people and finding cases early to reduce the chance they transmit the virus. Vaccination could play a role as well. The United Kingdom is already offering vaccines to health care workers and other contacts of known cases—a strategy called ring vaccination, which was key to the success of the smallpox eradication campaign. The U.S. has begun offering vaccination to some contacts but is undecided about health care workers caring for monkeypox patients. Rimoin says ring vaccination doesn't seem warranted yet but that might change "if in the next couple of days … we see that there are hundreds and hundreds of cases."

One of the two available vaccines, manufactured by Emergent BioSolutions, was designed as a smallpox vaccine and can protect even after exposure to monkeypox virus, when given within 4 days. But it contains a live virus called vaccinia that replicates inside the body; the similar vaccine used during smallpox eradication sometimes caused severe side effects, killing one in 1 million recipients.

The other vaccine, from Bavarian Nordic—which the United Kingdom is using—contains a nonreplicating form of vaccinia designed to cause fewer side effects. But it requires two doses, given 4 weeks apart—a disadvantage for people who are already exposed and need fast protection. Both vaccines are in short supply and typically only available via national stockpiles.

Drugs exist as well. One, tecovirimat, in 2018 became the first ever approved by the U.S. Food and Drug Administration (FDA) to treat smallpox after it proved safe in human trials and effective in animals given closely related viruses. FDA approved a second drug for smallpox, brincidofovir, last year. Although neither has been approved for monkeypox, both showed promising results against the virus in animal studies.

But a report published on 24 May in *The Lancet Infectious Diseases* reported no evidence that brincidofovir had "any convincing clinical benefit" in three patients treated in the United Kingdom over the past 3 years and said the drug had serious toxicities. Tecovirimat, in contrast, did not cause side effects and appeared to lead to a decrease in viral levels and a speedier recovery.

How far and fast the virus spreads in the coming days will determine how aggressively clinicians put these tools to the test. "These are very early days," Rimoin says. ∎

Japan tries—again—to revitalize its research

Latest effort would spend billions on a few universities, but skeptics give it long odds

By **Dennis Normile**

Alarmed by the declining stature of its universities, Japan is planning to shower up to $2.3 billion a year on a handful of schools in hopes of boosting their prominence.

The scheme was approved by the Japanese legislature on 18 May, although many details, including how to pick the favored universities, are still up in the air. But the move, under study for more than a year, has rekindled a debate among academics over how to reverse Japan's sinking research fortunes. Several previous schemes have yielded mixed results.

The new plan "aims to provide young promising scholars with the research environment that the world's top universities are supposed to offer, to dramatically enhance international collaborations, and to promote the brain circulation both domestically and internationally," says Takahiro Ueyama, a science policy specialist on the Council for Science, Technology and Innovation (CSTI), Japan's highest science advisory body, which was heavily involved in crafting the scheme.

But Guojun Sheng, a Chinese developmental biologist at Kumamoto University in Japan, is skeptical. "I am not very optimistic that this [plan] will do much to curb the slide in the ranking of Japanese research activities or international competitiveness," he says. Sheng, who previously studied and worked in China, the United States, and the United Kingdom, says the new plan does not address fundamental problems at Japanese research institutes: too few women and foreign scientists, a fear of change, and lack of support for young scientists. To get better results, "Japan has to change its research culture," he says.

Concerns over Japan's fading scientific clout have been growing for years. The nation's $167 billion in spending on R&D in 2020 was topped only by the United States and China, according to the Organisation for Economic Co-operation and Development (OECD). But research productivity "is markedly below [Group of 20 countries] average and citation impact is low," Clarivate's Institute for Scientific Information concluded in its 2021 annual report on G-20 research activities. An August 2021 analysis by Japan's National Institute of Science and Technology Policy (NISTEP) showed that Japan ranked fourth in its share of papers in the top 10% by number of citations from 1997 through 1999, then dropped to fifth between 2007 and 2009 and to 10th in 2017 to 2019 (see graphic, left). The drop is partly the result of the spectacular rise of China, which was not even in the top 10 in the 1990s and is now at first place. But Canada, France, Italy Australia, and India surpassed Japan as well.

What really got the attention of politicians, however, is Japan's lackluster performance in university rankings, says Yuko Harayama, a science policy expert who advises Tohoku University. The University of Tokyo is the only Japanese school in the *Times Higher Education* world university rankings, for example, and it dropped from 23rd place in 2015 to a tie for 35th place this year.

Now the government is pursuing a costly fix: a University of International Excellence program, funded by an endowment of up to 10 trillion yen ($78 billion). The fund could generate $2.3 billion annually to be shared among five to seven elite schools. Starting at the end of this year, universities will vie for inclusion in the program by presenting plans for institutional reforms and stronger research efforts. Money could start flowing in 2024. (Some of the funds will cover living and research expenses of doctoral students, not just at the selected schools but at all universities.)

The program is the latest of several attempts by the government to rejuvenate Japan's research efforts. In 2007, the Ministry of Education, Culture, Sports, Science and Technology (MEXT) launched a World Premier International Research Center Initiative (WPI) "to attract outstanding scientists to Japan" and spur other reforms. The idea was to create "something like an is-

land within a university with a completely different way of managing research activities," says Harayama, a member of the forerunner of CSTI when the WPI was planned. The 14 WPI institutes that resulted have higher proportions of internationally recruited scientists than the universities they are attached to, and two WPI directors are non-Japanese. But they have not had the reforming influence on their host universities that MEXT hoped for, Harayama says.

In 2015, Japan formed the Agency for Medical Research and Development to jump-start biomedical research, with an annual budget of $980 million. And a Moonshot Research and Development Program, launched in 2019, is disbursing $780 million over 5 years to support "high-risk, high-impact R&D" focused on seven

Declining influence

In the late 1990s, Japan ranked fourth in terms of its share of the 10% most-cited research papers. It dropped to 10th place by the end of the 2010s. The United States's share dropped as well, while China's scientific clout soared. (Credit for papers by multinational teams was split between countries.)

broad goals, including "ultra-early disease prediction and intervention," and "a sustainable global food supply."

The size and sprawling agendas of these top-down programs blur lines of responsibility and make performances difficult to evaluate, says Toshio Suda, a Japanese stem cell scientist at the National University of Singapore. They also emphasize applications more than basic research, he says.

Meanwhile, money available through MEXT's competitively reviewed Grants-in-Aid for Scientific Research, which Suda says is particularly important for younger researchers, has stagnated, hovering just below $2 billion a year for the past decade.

What's worse, Japanese universities, working with fixed block funding, "stopped giving [permanent] positions to young scientists," says Hitoshi Murayama, a theoretical physicist at the University of California, Berkeley. And those lucky enough to find appointments rarely get start-up funding, leaving them "at the mercy of senior professors in terms of resources," he adds. "The lack of independence makes it difficult for them to really kick-start their own research," says Murayama, who was the founding director of the University of Tokyo's Kavli Institute for the Physics and Mathematics of the Universe, one of the WPI institutes.

Such dismal career prospects are driving people away from academia. The number of students going for doctorates immediately after earning a master's degree has dropped 25% in 20 years, according to MEXT data. And some who do earn Ph.D.s are looking abroad for careers. Developmental biologist Kinya Ota, for example, found a position at Academia Sinica in Taiwan when nearing the end of a fixed-term appointment at a lab affiliated with RIKEN, Japan's network of national laboratories. In Taiwan, Ota got support to set up his lab from the start and, most important, "I could decide my own research direction." Ten years on, he has a permanent position and leads a small team. Tellingly, Ota says he is getting increasing numbers of queries about working overseas from younger Japanese scientists.

Rather than setting up top-down megaprograms, Sheng thinks Japan should encourage bottom-up initiatives from individual universities and institutes that might make better use of the resources. He also says greater diversity in labs, in nationality and gender, would help generate new research ideas. Women make up only 17% of Japan's research workforce, far below the OECD average of 40%.

Indeed, MEXT is studying proposals to scale up support for regional universities, provide richer stipends for graduate students, and expand opportunities for women, says Takuya Saito, director of human resources policy for the ministry. The government is aware the new plan won't fix all problems in Japanese research, Saito says: "The improvement of research capabilities in Japan as a whole is based on the recognition that support for several universities alone will not be enough." ■

An artist's conception of NASA's InSight lander, which is scheduled to be shut down by the end of 2022.

PLANETARY SCIENCE

End draws near for lander that mapped the interior of Mars

Mars InSight listened to quakes to give us our first look inside another planet

By **Jonathan O'Callaghan**

It's never easy saying goodbye. But more than 3 years after NASA's InSight lander touched down on Mars, planetary scientists are preparing for the demise of the mission, which has helped reveal how the planet is put together. Last week, NASA officials said dust accumulating on the probe's solar panels has reduced power flows to a trickle, forcing them to shut it down by the end of the year.

Launched in 2018, InSight (Interior Exploration using Seismic Investigations, Geodesy and Heat Transport) has recorded a host of accomplishments, including measuring the magnetic field of Mars and detecting more than 1300 quakes, which have enabled researchers to map its structure. It has provided "an important first step in studying the interior of Mars," says Paula Koelemeijer, a seismologist at the University of Oxford.

InSight landed in November 2018 on the Elysium Planitia, a broad, flat plain on the martian equator. Planners picked the spot because they hoped its relatively calm, quiet weather would allow one of the lander's primary instruments, an ultrasensitive seismometer, to detect seismic waves from distant marsquakes as they bounced around the interior of the planet.

After researchers switched on the instrument in February 2019, however, they heard nothing for 2 months. "We were rather nervous," recalls John Clinton, a seismologist from ETH Zürich and a co-investigator on the mission. One fear was that any marsquakes were too faint to be detected. But those worries were allayed in April 2019 when InSight recorded its first quake. Since then, more than 1300 have been documented by Clinton's Marsquake Service, an online database that catalogs InSight's daily seismic data.

Many of the quakes appear to originate in the nearby Cerberus Fossae region, but their causes remain largely unknown. Possibilities include remnant volcanic activity on the largely quiescent planet, geologic stress, and meteorite impacts. Still, by tracking the arrival of seismic waves and working out how fast they traveled through interior layers, researchers have been able to assemble a more detailed picture of the planet's core, mantle, and crust. "We now have a map of the inside of Mars," says Bruce Banerdt of NASA's Jet Propulsion Laboratory, the lead on the mission.

There have been surprises. Some researchers expected to find a relatively meager and solid core, for example. Instead, "It's a really big core," proportionally as big as Earth's, says planetary scientist Julia Semprich of the Open University, and like Earth's outer core, it is molten. But it is not clear yet whether Mars has a solid inner core like Earth's. And researchers are still analyzing data from a radio instrument designed to track the wobble of Mars's spin axis, which can offer other clues to the size and consistency of the core.

The planet's mantle, the layer sandwiched between the planet's core and crust, appears to consist of just a single rocky layer, rather than two like Earth. And the thin martian crust has either two or three layers, with possible evidence for water buried inside.

Most of the quakes detected by InSight have been relatively modest, up to a magnitude 4. But on 4 May, in the lander's waning days, it registered a rare and welcome event: a magnitude 5 quake. Throughout the mission, scientists had been hoping to capture such a tremor, which was "as large as all of the other quakes seen until then combined," says Brigitte Knapmeyer-Endrun, a planetary seismologist from the University of Cologne.

Researchers are still analyzing the waves from the powerful quake, but they hope it will deepen their understanding of the planet's structure. "There's going to be lots of information as we get our teeth into it," Banerdt says.

The $425 million mission has also documented a planetary magnetic field that is 10 times stronger than expected. That suggests magnetized rocks sit deep under the surface, storing a record of its lost magnetic field.

It hasn't been all plain sailing. A heat probe designed to measure subsurface temperatures couldn't penetrate the tough martian skin. "That was the biggest disappointment of the mission," Banerdt says.

A plan to have a second seismometer running in parallel with InSight's also failed to materialize. Europe's ExoMars mission was supposed to deliver the twin sensor, which would have enabled researchers to "see events in different regions," Knapmeyer-Endrun says. But ExoMars has faced delays.

InSight is now generating just one-tenth as much power as at the mission's start. NASA says it will cease science operations by July, then switch off the lander at the end of this year unless a passing dust devil brushes off its solar panels. That could enable the lander to "boot itself up and come back to life," Banerdt says.

Even as InSight falters, researchers are eyeing how they might follow its example by listening for quakes on other bodies. Some are planning a seismometer network for the Moon, for example, and NASA's Dragonfly mission is scheduled to take a seismometer to Saturn's moon Titan in 2026. Venus or Jupiter's icy moon Europa are also attractive targets. The InSight mission, Banerdt says, "has shown seismology for the incredibly capable technique it is." ∎

Jonathan O'Callaghan is a journalist in London.

Liviah Widders, 4, of Cincinnati, Ohio, awaited a liver transplant in December 2021.

INFECTIOUS DISEASE

Coronavirus, adenovirus, or both? Hepatitis poses mystery

Researchers scramble for data on the cause—and therefore the treatment—of liver disease in children

By **Meredith Wadman**

Physicians at pediatric liver centers in the United Kingdom have been meeting regularly to work out how best to care for children with a mysterious hepatitis that has afflicted at least 197 youngsters in the United Kingdom and more than 600 worldwide.

But they can't agree.

"It's proving immensely difficult to get everyone to agree how on Earth to manage these children," says Will Irving, a virologist at the University of Nottingham.

At the center of the debate are conflicting theories about what is causing healthy young children to suddenly become jaundiced and fall seriously ill with acute liver inflammation. One hypothesis suggests the damage is being done by adenovirus, a common childhood infection that normally causes coldlike symptoms and could be treated with an antiviral drug. Another suggests the cause is a rogue immune response to previous infection by SARS-CoV-2—which could be treated with immune-suppressing drugs such as steroids. A third hypothesis proposed this month brings them together,

suggesting adenovirus infection forms a destructive partnership with SARS-CoV-2 that sets the immune system loose on the liver.

"The clinicians have got a really difficult dilemma," Irving says. "With a sick child, do you give steroids? Do you give [an antiviral drug]? Do you give both?"

Cases remain rare but some children become seriously ill: About 9% of 180 cases in the United States have required liver transplants, the U.S. Centers for Disease Control and Prevention (CDC) said on 18 May; six children have died. In the United Kingdom, 11 children have needed transplants, with no deaths as of 16 May.

Official bodies including CDC and the UK Health Security Agency (UKHSA) have elevated the adenovirus hypothesis. Infection with adenovirus can cause hepatitis in immune-suppressed children but is not known to do so in healthy kids. But CDC says adenovirus had been found in nearly half of the U.S. cases as of 18 May "and continues to be a strong lead." The agency's most recent physician alert urges testing of suspected cases for adenovirus. Adenovirus was also found in 68% of U.K. children with hepatitis who were tested for it through 16 May. And

UKHSA, in an announcement on 20 May, said pointedly, "The investigation continues to suggest an association with adenovirus."

"The fact you have it in [nearly] 70% of cases does suggest that it must have a role," says Deirdre Kelly, a pediatric hepatologist at Birmingham Children's Hospital, one of a group of experts advising UKHSA.

But other scientists say adenovirus could be an innocent bystander. The key question, says Isabella Eckerle, a virologist at the University of Geneva, is "how specific is this finding of low amounts of adenovirus? Would we also find it in healthy kids?"

The skeptics note that liver biopsies from the children have failed to find cells stuffed with adenovirus, a classic sign of adenoviral hepatitis. They say, sometimes heatedly, that the agencies are overlooking a more likely culprit, SARS-CoV-2.

"It is profoundly embarrassing that major scientific bodies in U.S. and U.K. are using such weak circumstantial evidence," tweeted Farid Jalali, an adult hepatologist in Laguna Hills, California. He accused them of downplaying the "likely possibility that recent SARS-CoV-2 infection may be driving the increase in cases."

Jalali and others suggest SARS-CoV-2 may trigger an immune-mediated attack on the liver several weeks after infection, just as other organs can be attacked weeks later in the condition called multisystem inflammatory syndrome in children (MIS-C).

Only some of the children were known to be infected with SARS-CoV-2 around the time their hepatitis was diagnosed—15% in the United Kingdom. But a recent CDC study estimated that 75% of all U.S. children under 12 have been infected, 31% of them between December 2021 and February. The European Centre for Disease Prevention and Control on 20 May reported evidence of prior SARS-CoV-2 infection in 19 of 26 children with hepatitis. It also showed that most cases in Europe occurred during a large Omicron wave. Only a tiny minority of the affected children were vaccinated against SARS-CoV-2; vaccines are not available for children under 5 years old, the age group in which most cases occur.

The debate is not academic. It "has everything to do with whether [a] patient stays alive or not," Jalali said in an interview. If adenovirus is damaging the liver, the powerful antiviral drug cidofovir could be deployed. But if the liver damage results from a sustained immune reaction, immune-suppressing drugs could be lifesaving. You had better not be wrong, Jalali says. "If you mistakenly assume some infectious process

is actively causing liver failure, you can't go near that patient with immune-suppressive drugs," which can hobble the body's ability to fight active infection.

Petter Brodin, a pediatric immunologist and pediatrician at Imperial College London and Moshe Arditi, a pediatric infectious disease physician at Cedars-Sinai Medical Center, earlier this month published a hypothesis knitting the two viruses together. They noted that many of the cases tested in the United Kingdom harbored adenovirus-41, a strain that infects the gut, and that SARS-CoV-2 has been found to establish gut reservoirs that persist after acute infection.

Brodin and Arditi proposed that after adenovirus infects the gut, coronavirus might act as a coconspirator. A small section of the SARS-CoV-2 spike protein that has been shown to prompt a broad activation of T cells might supercharge the immune response to adenovirus, and the rogue immune response might then attack the liver. Such a mechanism, in which a piece of the coronavirus spike protein triggers an immune overreaction, has been implicated in the serious inflammation found in MIS-C.

Brodin urges clinicians to test stool and other samples for intestinal SARS-CoV-2 and a hyperactivated immune system. If the hypothesis is confirmed, he says immunosuppressive therapy would be appropriate. "[If] it's an out-of-control activation of the immune system, then you need to be very aggressive in shutting down that immune response."

Jalali, for his part, is worried by a preprint, not yet peer reviewed, posted on 14 May by scientists at Case Western Reserve University. The paper suggests the hepatitis cases are the tip of an iceberg of liver damage in children post–COVID-19. The researchers compared the electronic records of 246,000 children who got COVID-19 between March 2020 and March 2022 with those of 551,000 children who contracted other respiratory infections in that period. In the months after infection, COVID-19 infected children were 2.5 times more likely to have elevated levels of enzymes that indicate liver damage, and 3.3 times more likely to have elevated bilirubin, which can also indicate impaired liver function.

Clare Wenham, a global health policy expert at the London School of Economics whose 4-year-old son was hospitalized with hepatitis this month and treated with supportive care, has been tracking the battling theories of causation. "There just aren't enough data to really make any firm conclusions yet," she says. "You're with your clinicians and they're like: 'We just don't really know … what the trajectory is going to be.' That's what's scary as a parent." ∎

Engineered tomatoes get a healthy dose of vitamin D

Knocking out a gene boosts precursors of essential nutrient

By **Erik Stokstad**

Tomatoes get riper and tastier in the summer Sun. Two studies now suggest that with a little help from gene editing, Sun-ripened tomatoes can also stockpile a precursor molecule to vitamin D, a vital nutrient normally found mainly in animal products.

"This could be a game changer" in nations where vitamin D deficiency is a problem, says Esther van der Knaap, a plant geneticist at the University of Georgia, Athens. Biofortified plants could also help vegans get enough of the nutrient. The finding "opens up a very exciting new era for vitamin D," says nutritional scientist Susan Lanham-New of the University of Surrey.

Vitamin D helps regulate how the body uses calcium, leading to stronger bones. And there's some evidence that low levels can raise the risk of heart disease and other health problems.

Sunlight helps your body synthesize the vitamin, as ultraviolet B (UVB) radiation converts a precursor in skin into a form that the liver and kidney transform into usable vitamin D. But people living at high latitudes often aren't exposed to enough UVB, particularly in winter, to avoid deficiency. And age or darker skin can also slow synthesis.

Eating animal products—fish, eggs, and liver—that contain precursors can help make up the deficit. In addition, milk sold in the United States and some other countries is supplemented. For vegans, mushrooms and yeast are a less effective source. Another option is to take supplement pills, which are often made from lanolin, a waxy substance secreted by sheep. (Sheep can get vitamin D by licking the lanolin off their wool.)

Because tomatoes naturally make a key vitamin D precursor, two groups thought some genetic tweaking could turn them into an animal-free source of the vitamin.

This week in *Nature Plants*, a team led by Cathie Martin, a plant metabolic engineer at the John Innes Centre, reported that disabling a single gene created tomatoes that could each provide 20% of the recommended daily allowance of vitamin D in the United Kingdom. And in a March preprint posted

> ## "A very exciting new era for vitamin D."
> **Susan Lanham-New**,
> University of Surrey

at Research Square, a group led by plant geneticist Sunghwa Choe of Seoul National University reported that disabling a related gene produced tomatoes with even higher levels of vitamin D precursors.

Normally, tomatoes make a precursor called provitamin D3 but convert it into other compounds using enzymes coded for by two genes, *7-DR1* and *7-DR2*. These compounds help the plant grow and protect it from stress. The researchers suspected that knocking out either of these genes would cause the plant to accumulate provitamin D3, which when exposed to sunlight transforms into a second precursor—previtamin D3—that people can use. Leaving the other gene intact would help the plant function normally. "This seemed like a real opportunity," Martin says.

It worked. Martin's team knocked out *7-DR2*, producing tomatoes that, after ripening and exposure to sunlight, offered as much previtamin D3 as two medium eggs. The content can be increased by slicing the tomato first, and likely even more by drying them in the Sun.

Choe's group disabled the other gene, *7-DR1*. Although the researchers did not expose the biofortified fruit to UVB, they estimate that each tomato can contain up to 100 micrograms of provitamin D3. Levels remained high after a month of freeze-dried storage. "We think that the molecule is pretty stable in the fruit," Choe says.

So far, the modified tomatoes have been grown only in laboratory greenhouses. Field tests will be crucial to seeing whether the plants can thrive under real-world stress, and Martin will begin a field trial next month. Choe hopes to start one this summer. Researchers will also need to show that the body can absorb the previtamin D3 in the tomatoes and convert it to vitamin D.

Despite the potential benefits, some consumers might not accept gene-edited tomatoes, notes Kevin Cashman, a nutritional scientist at University College Cork. But if biofortified tomatoes do make it to market, plant physiologist Dominique Van Der Straeten of Ghent University writes in a commentary in *Nature Plants*, they could mark "a leap forward in decreasing our dependence on animal-based foods." ∎

is actively causing liver failure, you can't go near that patient with immune-suppressive drugs," which can hobble the body's ability to fight active infection.

Petter Brodin, a pediatric immunologist and pediatrician at Imperial College London and Moshe Arditi, a pediatric infectious disease physician at Cedars-Sinai Medical Center, earlier this month published a hypothesis knitting the two viruses together. They noted that many of the cases tested in the United Kingdom harbored adenovirus-41, a strain that infects the gut, and that SARS-CoV-2 has been found to establish gut reservoirs that persist after acute infection.

Brodin and Arditi proposed that after adenovirus infects the gut, coronavirus might act as a coconspirator. A small section of the SARS-CoV-2 spike protein that has been shown to prompt a broad activation of T cells might supercharge the immune response to adenovirus, and the rogue immune response might then attack the liver. Such a mechanism, in which a piece of the coronavirus spike protein triggers an immune overreaction, has been implicated in the serious inflammation found in MIS-C.

Brodin urges clinicians to test stool and other samples for intestinal SARS-CoV-2 and a hyperactivated immune system. If the hypothesis is confirmed, he says immunosuppressive therapy would be appropriate. "[If] it's an out-of-control activation of the immune system, then you need to be very aggressive in shutting down that immune response."

Jalali, for his part, is worried by a preprint, not yet peer reviewed, posted on 14 May by scientists at Case Western Reserve University. The paper suggests the hepatitis cases are the tip of an iceberg of liver damage in children post–COVID-19. The researchers compared the electronic records of 246,000 children who got COVID-19 between March 2020 and March 2022 with those of 551,00 children who contracted other respiratory infections in that period. In the months after infection, COVID-19 infected children were 2.5 times more likely to have elevated levels of enzymes that indicate liver damage, and 3.3 times more likely to have elevated bilirubin, which can also indicate impaired liver function.

Clare Wenham, a global health policy expert at the London School of Economics whose 4-year-old son was hospitalized with hepatitis this month and treated with supportive care, has been tracking the battling theories of causation. "There just aren't enough data to really make any firm conclusions yet," she says. "You're with your clinicians and they're like: 'We just don't really know … what the trajectory is going to be.' That's what's scary as a parent." ■

Engineered tomatoes get a healthy dose of vitamin D

Knocking out a gene boosts precursors of essential nutrient

By **Erik Stokstad**

Tomatoes get riper and tastier in the summer Sun. Two studies now suggest that with a little help from gene editing, Sun-ripened tomatoes can also stockpile a precursor molecule to vitamin D, a vital nutrient normally found mainly in animal products.

"This could be a game changer" in nations where vitamin D deficiency is a problem, says Esther van der Knaap, a plant geneticist at the University of Georgia, Athens. Biofortified plants could also help vegans get enough of the nutrient. The finding "opens up a very exciting new era for vitamin D," says nutritional scientist Susan Lanham-New of the University of Surrey.

Vitamin D helps regulate how the body uses calcium, leading to stronger bones. And there's some evidence that low levels can raise the risk of heart disease and other health problems.

Sunlight helps your body synthesize the vitamin, as ultraviolet B (UVB) radiation converts a precursor in skin into a form that the liver and kidney transform into usable vitamin D. But people living at high latitudes often aren't exposed to enough UVB, particularly in winter, to avoid deficiency. And age or darker skin can also slow synthesis.

Eating animal products—fish, eggs, and liver—that contain precursors can help make up the deficit. In addition, milk sold in the United States and some other countries is supplemented. For vegans, mushrooms and yeast are a less effective source. Another option is to take supplement pills, which are often made from lanolin, a waxy substance secreted by sheep. (Sheep can get vitamin D by licking the lanolin off their wool.)

Because tomatoes naturally make a key vitamin D precursor, two groups thought some genetic tweaking could turn them into an animal-free source of the vitamin.

This week in *Nature Plants*, a team led by Cathie Martin, a plant metabolic engineer at the John Innes Centre, reported that disabling a single gene created tomatoes that could each provide 20% of the recommended daily allowance of vitamin D in the United Kingdom. And in a March preprint posted

> ## "A very exciting new era for vitamin D."
> **Susan Lanham-New**,
> University of Surrey

at Research Square, a group led by plant geneticist Sunghwa Choe of Seoul National University reported that disabling a related gene produced tomatoes with even higher levels of vitamin D precursors.

Normally, tomatoes make a precursor called provitamin D3 but convert it into other compounds using enzymes coded for by two genes, *7-DR1* and *7-DR2*. These compounds help the plant grow and protect it from stress. The researchers suspected that knocking out either of these genes would cause the plant to accumulate provitamin D3, which when exposed to sunlight transforms into a second precursor—previtamin D3—that people can use. Leaving the other gene intact would help the plant function normally. "This seemed like a real opportunity," Martin says.

It worked. Martin's team knocked out *7-DR2*, producing tomatoes that, after ripening and exposure to sunlight, offered as much previtamin D3 as two medium eggs. The content can be increased by slicing the tomato first, and likely even more by drying them in the Sun.

Choe's group disabled the other gene, *7-DR1*. Although the researchers did not expose the biofortified fruit to UVB, they estimate that each tomato can contain up to 100 micrograms of provitamin D3. Levels remained high after a month of freeze-dried storage. "We think that the molecule is pretty stable in the fruit," Choe says.

So far, the modified tomatoes have been grown only in laboratory greenhouses. Field tests will be crucial to seeing whether the plants can thrive under real-world stress, and Martin will begin a field trial next month. Choe hopes to start one this summer. Researchers will also need to show that the body can absorb the previtamin D3 in the tomatoes and convert it to vitamin D.

Despite the potential benefits, some consumers might not accept gene-edited tomatoes, notes Kevin Cashman, a nutritional scientist at University College Cork. But if biofortified tomatoes do make it to market, plant physiologist Dominique Van Der Straeten of Ghent University writes in a commentary in *Nature Plants*, they could mark "a leap forward in decreasing our dependence on animal-based foods." ■

SCIENTIFIC WORKFORCE

Hughes bets 'happy labs' will boost diversity

150 early-career scientists will get lavish funding based on research excellence

By **Jeffrey Mervis**

The Howard Hughes Medical Institute (HHMI) is taking a novel approach to the chronic problem of under-representation in U.S. academic science: lavishly fund a relatively small number of young scientists and then sit back and let their "happy labs" produce a more diverse workforce.

On 26 May, HHMI announced it will award $8.6 million over the next decade to each of 150 life scientists in tenure-track positions. The unprecedented amount of funding for early-career researchers is expected to draw thousands of applicants. Winners, who will also become HHMI employees, do not themselves need to be from a group under-represented in science. (In fact, federal law prohibits the use of race or gender as a criterion for hiring.) But a commitment to a diverse, equitable, and inclusive (DEI) work environment is essential.

"Excellence comes first," says Leslie Vosshall, HHMI's chief scientific officer, who notes that the winners will be required to spend at least 80% of their time at the bench. "But we're also looking for people to whom students representing all metrics of diversity will flock, making their labs hotbeds of diversity within the department."

The $1.5 billion initiative, called the Freeman Hrabowski Scholars Program, is named for the retiring president of the University of Maryland (UMD), Baltimore County, who has spent 30 years perfecting and then replicating the much-admired Meyerhoff Scholars Program to boost the number of minority scientists. Hrabowski, a mathematician who is Black, thinks the new program has great potential for "moving the needle," as he puts it.

"It takes a scientist to produce another scientist," he says, endorsing HHMI's decision to focus on research excellence. "The Hughes brand will be a magnet for attracting talent." With universities eager to do more to foster diversity, Hrabowski says, the program is also well-timed. "I'm more

encouraged by this than by anything I've seen since we began Meyerhoff," he says.

The magnitude of the HHMI program makes it "potentially transformative," agrees Stephen Thomas, a professor of public health at UMD who also directs the mentoring component of a diversity initiative funded by the National Institutes of Health (NIH). But Thomas, who is Black, says a well-funded lab doesn't always translate into an inclusive environment. Excellent scientists of any demographic can fail as mentors, he says. "The current system of training researchers is calcified and creates far too many tormentors instead of mentors," he asserts.

Mentoring will be key, Thomas says. He notes NIH's career development grants re-

> "We're … looking for people to whom students representing all metrics of diversity will flock."
>
> **Leslie Vosshall**, Howard Hughes Medical Institute

quire applicants to designate a senior scientist who will tutor them. "Who is going to mentor the HHMI scholars?" he asks.

Vosshall says current HHMI investigators—all senior scientists—will fill that role, serving as "buddies" who will advise scholars "on all aspects of developing their research program, coping with setbacks, and mentoring their people." But Thomas would prefer a more formal arrangement. "It takes a lot of training to become a good mentor," he says.

The program's success will also depend on finding scientists who are truly committed to increasing diversity. "We've spent thousands of hours over the last 6 months figuring out how to do that," Vosshall says.

She notes that HHMI rejected one hiring tool used by a growing number of universities trying to diversity their faculty.

"We explicitly did not go the route of 'diversity statements' because they more or less look the same," she says. "In fact, people often cut and paste the statements of others, and you don't really learn anything."

Instead, applicants will be asked to discuss their own experiences with discrimination and explain how they would do things differently. For example, she says, "if you have been in a lab where African American or Latinx scientists have not had a happy experience and have been pushed out of science, how are you going to build a space where that doesn't happen?"

The initial winnowing of what Vosshall predicts may be a pool of 2000 or more applicants will be "based solely on a page and a half response to those questions," she says, with the identities of the applicants hidden to avoid reviewer bias. The 300 or so who survive that first cut will then be judged entirely on their scientific potential, she adds. The first cohort of 30 winners will be selected in March 2023, and HMMI plans to hold competitions every other year for four subsequent cohorts of similar size.

The success of the program won't be measured by the diversity of those cohorts, however. Instead, Vosshall plans to use two yardsticks. The first is the relative demographics of the students who train in the winning labs compared with those in other labs. The second is the ability of those students to foster diversity after they get their own academic positions.

"Our expectation is that happy and inclusive labs will attract a more diverse pool of students," she says about the first goal. As for the second: "I think having 3000 people who have been given the space and the grace to be able to do their science without being constantly under assault will make an enormous difference" when they set up their own labs. ∎

FALLING STAR

It was a source of pride for decades. But Brazil's National Institute for
Space Research is rapidly losing funding, brains—and hope.

By **Sofia Moutinho**
and **Herton Escobar**

Carefully stowed away in a locker, Evlyn Novo keeps a collection of framed plaques honoring her time as a researcher at the Brazilian National Institute for Space Research (INPE). Novo joined the institute as a young remote sensing specialist in 1975 to work on a pioneering effort to use satellite data to monitor deforestation in the Amazon. Over her career, she helped INPE develop into one of the flagships of Brazilian science—a global leader in watching tropical forests from space. For every 5 years Novo spent at INPE, she received a commemorative plaque honoring her service. She was looking forward to getting the 10th one.

But with only 2 years to go until that milestone, Novo, 69, has come to a heart-breaking decision: She has lost faith in the institution's future and will retire from INPE by the end of this year. "I don't want to be the one to turn off the lights," she says.

INPE is in decline, and Novo sees it everywhere. A few years ago, the office lights often stayed on until late at night at INPE's main campus in São José dos Campos, near São Paulo, where staff and students analyzed remote sensing data, built satellites, and modeled weather and climate. Today, the institute struggles to pay its electricity bills. Potholes pepper the campus streets and sidewalks are broken. They are the physical symptoms of a much larger institutional crisis marked by sharp budget cuts, a shrunken staff, and relentless attacks by Brazilian President Jair Bolsonaro and his supporters. "It is a climate of total dismay," Novo says.

Other sectors in Brazilian science are backsliding, too. A faltering economy and shifting political priorities have led to steep cuts in funding for science at many universities and federal research institutions. But INPE's downfall has been particularly painful to watch, many say, because of its international prestige, its role in protecting the Amazon, and the way Bolsonaro has added insult to injury by trying to discredit the institute's work.

Today, Brazil's monitoring programs for deforestation and wildfires are imperiled, the supercomputer that runs climate models is aging and unreliable, and INPE's satellite development program is on hold, lacking funds to advance planned missions and launches. A beacon of Brazil's scientific prowess has become a symbol of science's struggle for survival there, in an underfunded and

INPE's satellite assembly hall in 2020. Today, little is happening in the facility as funds have dried up.

politically menacing environment.

INPE's director, Clezio De Nardin, acknowledges that his institute is in trouble. In an interview with *Science*, he says INPE needs at least twice its current operating budget of 92 million reais ($18 million) a year to fulfill its missions. But he blames the cuts on Brazil's economic problems, not on politics. "I don't believe any ruler in good conscience would act to destroy an institution that produces essential infrastructure for its own country," De Nardin says. "Especially because defunding the space sector will have consequences in society for decades."

In an email to *Science*, Brazil's Ministry of Science, Technology and Innovation (MCTI) pointed to a few positive signs: This year, INPE's discretionary funding went up by 33%, bringing it back to 2019 levels, and grants from the National Fund for Scientific and Technological Development (FNDCT) will add an additional boost.

But researchers say the amount of grant support is uncertain, and the extra funding won't be nearly enough to reverse years of decline. By now, many staffers have become demoralized, says Gino Genaro, a senior satellite technologist at INPE: "People are distressed without knowing what to do and what the future will hold."

INPE WAS FOUNDED in 1961, when the space race was in full swing and the Cold War at its peak. Inspired by a visit from Soviet cosmonaut Yuri Gagarin, then-President Jânio Quadros embraced an idea from the Brazilian Interplanetary Society to launch a national space institute. In its early years, INPE created and consolidated research and postgraduate programs in meteorology, astrophysics, and remote sensing.

In the 1970s, it became the first space agency in any country to monitor forests using satellite data, obtained by the U.S. Landsat Program. In 1989, the agency launched the Brazilian Amazon Deforestation Satellite Monitoring Program (PRODES), which provides yearly and historical deforestation trends for the world's largest rainforest. In 2004, INPE added a system named

Real-Time Deforestation Detection (DETER), which uses real-time images from a variety of satellites, some of them developed partly in Brazil, to send daily and monthly updates of fires and other causes of forest loss to enforcement officials.

Data from PRODES and DETER helped Brazil create and enforce policies that were key to reducing annual deforestation in the Amazon by 82% between 2004 and 2014. Both programs are "fundamental for understanding the agricultural expansion in Brazil

Institute in free fall

Once a beacon of Brazil's scientific prowess, the National Institute for Space Research (INPE) has suffered major losses in its overall budget (top) and funding for two world-renowned rainforest monitoring programs (bottom) over the past decade. Its staff has shrunk by almost one-quarter (middle).

Total budget (millions of reais)

After a record low in 2021, INPE saw its budget increase to 92.3 million reais ($18 million) this year.

A one-time increase in 2017 allowed INPE to contract the launch of the Amazonia-1 satellite.

● Researchers ● Technicians and administrative staff

Staff

Amazon monitoring budget (millions of reais)

and carbon emissions from deforestation in the Amazon," says Douglas Morton, a remote sensing specialist at NASA's Goddard Space Flight Center. Morton says researchers worldwide prize INPE's data, which it made publicly available on the internet in 2003, a pioneering step that the United States and other countries would follow. The institute is also an important training center for remote sensing scientists, Morton adds.

Another source of pride at INPE is the Integration and Testing Laboratory (LIT), which assembles satellites. The only lab of its kind in the Southern Hemisphere, LIT has tested and developed many satellites in the past 35 years, most in partnership with other countries. The most successful and long-standing project was a collaboration with China to launch six imaging satellites, the first one in 1999, that gave Brazil its own satellite views for the first time. Two are still used today to monitor fires, deforestation, and land use.

"INPE is certainly one of the most important research institutions in Latin America," says computer scientist Gilberto Câmara, a career researcher who led the institute from 2006 to 2013. During Câmara's tenure, in 2010, Brazil invested $13 million in a new Cray XT6 supercomputer, for use by INPE teams and researchers elsewhere in Brazil. Nicknamed Tupã, after an Indigenous South American god of thunder, it was one of the 30 most powerful computers in the world at the time—and it was hailed as another milestone for Brazilian science.

But the string of successes wouldn't last.

INPE'S FORTUNES TURNED about a decade ago, during the first term of leftist President Dilma Rousseff. Federal support for science, including at INPE, suffered in a maelstrom of economic and political troubles. The turmoil culminated in Rousseff's impeachment and removal in 2016, and Bolsonaro's election in 2018.

The decline accelerated after Bolsonaro took office, despite his campaign promises to prioritize science. His government cut MCTI's overall budget by 35% in the first 3 years of his administration, to 8.3 billion reais ($1.7 billion). Despite a partial recovery this year, after Congress banned the

An image from the Amazonia-1 satellite shows a dark island of rainforest in the Parakanã Indigenous Territory in Para state, surrounded by fields and pastures. Monitoring deforestation has helped Brazil protect the Amazon.

nal scientists who have stopped using Tupã.

There is no money for a replacement. As a makeshift solution, INPE bought a second, smaller and less powerful machine in 2018 that took over weather forecasting and other everyday operational processes, while the old supercomputer is mostly used for research. Last year, with funds from the United Nations Development Programme, it bought a set of additional processors for Tupã. But if the old machine dies, Brazil's climate modeling program might shut down. The agency has submitted a request to the Brazilian Innovation Agency for 200 million reais ($40 million) to upgrade its entire computer system. Even if the project is approved, it will take at least a year until the new machinery is ready to work.

LIT, the satellite development facility, has also suffered heavily. Its last major project was the Amazonia-1, the first satellite entirely developed in Brazil, launched from India in February 2021. It added an extra eye in the sky to watch over the Amazon, increasing the frequency at which images of the region are generated and allowing faster deforestation alerts.

Amazonia-1's launch was a cause for celebration, but the success masked many problems. A lack of funds had delayed the mission by 3 years. Most researchers working on it had lost their grant by that time; the grants had to be reinstated temporarily to make the launch possible. And funds are lacking to develop two long-planned companion satellites. "The discontinuity of this project is a shot at the heart of the Brazilian space program," says former INPE Director Ricardo Galvão. As a continental country with huge areas of native vegetation and crops, Brazil needs a whole constellation of new satellites to meet its remote sensing demands, Galvão says.

To save on LIT's formidable electricity bill, INPE has reduced its operations, shutting down cleanrooms from time to time. The room where spacecraft were assembled now houses the skeleton of a satellite, built from spare pieces of the Amazonia-1, "to show off to visiting politicians and journalists," an employee tells *Science*. LIT has long tested products for industry on the side, such as cars, phones, and even bathtubs; today that is practically all it does.

De Nardin says the institute is looking for funds to develop new Amazonia satellites. But meanwhile, Genaro, who has worked at the institute's space systems division for 20 years, says, "All we have are loose ideas on paper, without any budget or a team to execute them."

MONEY IS NOT THE ONLY resource in short supply. There's also a dearth of brains, especially young ones. INPE has seen a strik-

government from freezing FNDCT funds, MCTI's overall budget for this year is still 34% smaller than 5 years ago, not counting inflation.

INPE's annual budget, meanwhile, plunged by 63% from 2010, to a record low of just 76 million reais ($15 million) in 2021. Even after the recent boost, it is still "obviously behind what is needed," De Nardin tells *Science*.

The consequences have been felt at every corner of the institute. The budget dedicated to the two rainforest monitoring programs declined by 70% in the past decade, to 2.7 million reais ($540,000) last year—about 0.1% of what New York City spends on its Department of Parks & Recreation annually. "If there are more budget constraints, we will have to stop doing something; maybe reduce the area we monitor," says INPE's senior technologist Cláudio Almeida, who coordinates the two programs.

Keeping an eye on other corners of Brazil's rich biodiversity has also been challenging. One program tracks deforestation and wildfires in the Cerrado, 2 million square kilometers of shrubland bordering the Amazon forest that covers most of central Brazil.

Among the world's most biodiverse savannas, the Cerrado is under heavy pressure; more than half has already been cleared for crops and cattle. Last year, Almeida's team avoided shutting down the program only by transferring money from another project at the last minute. They hope a FNDCT grant of 15 million reais ($3 million), released in late April, will keep the program going another 3 years.

The funding crisis is "very worrisome," says Julia Shimbo, an ecologist at the nongovernmental Amazon Environmental Research Institute. INPE's data are not only used for public policies on deforestation, she notes, but are also a reference for international agreements on greenhouse gas emissions to which Brazil has committed.

The cash crunch creates other problems as well. After 12 years, Tupã is on its last legs. Processors frequently get overheated, circuit breaks are not uncommon, and the institute can barely afford the electricity to run the computer. "This obsolescence affects the country's climate research, weather forecasts, and INPE's mission itself," says environmental physicist Paulo Artaxo of the University of São Paulo's main campus, one of many exter-

ing reduction in personnel over the past 2 decades as researchers retired with no one to replace them. The last time INPE was allowed to add permanent staff was in 2012. The cash-strapped government is reluctant to allow hiring, in part because Brazil's generous pension system has become a huge financial burden. Other federal research institutes and universities face the same problem.

As a result, the number of full-time employees at INPE has dropped from some 2000 in 1990 to about 753 today. (Of those, 146 are researchers, 460 work in technical positions, and the remainder in management and administration jobs.) Like most other groups at INPE, Almeida's Amazon monitoring team relies mostly on temporary researchers with external grants who often don't stay long because of payment delays and the instability of living without a job contract.

"Back in the day, it was superhard just to find room to put up a desk at INPE," Almeida recalls. Nowadays, he says, "any fellow [grantee] has their own office." It's not just empty, he says. "The institute is growing old."

Younger scientists see little future there. Digital ecologist Thiago Silva did his postdoctoral research at INPE between 2010 and 2013, using satellite data to study wetland dynamics in the Amazon. Working under Novo, he was one of two contenders selected to be hired by the institute in 2012. He turned down the offer, fearing he might end up isolated and without enough money—or even colleagues—to continue his work in the long term. Silva moved to Scotland in 2019 to become a professor at the University of Stirling. "We end up having to leave Brazil to keep our research when it would be much more beneficial to stay in the country," he says.

Novo's last four graduate students have left the country as well. Most of their research lines were discontinued. "The situation is nerve-racking," Novo says. "You spend years building a lab, buying equipment, and forming human resources. And all that can disappear overnight because those in charge of policies lack long-term vision."

The government says INPE can't complain. Marcos Pontes, who left as Brazil's minister of science, technology, and innovation in April to run for Congress, told a local radio station in February that the institute was "one of the most privileged" units within the science ministry. Pontes, best known as the country's first and only astronaut—he spent 9 days on the International Space Station in 2006—said every institute supported by MCTI had suffered restrictions and no one could accuse him of "making deliberate cuts" to INPE. De Nardin agrees that the government isn't singling INPE out for cuts.

Many critics see it differently. They say Bolsonaro's government has intentionally targeted INPE, angered by its leadership and transparency in monitoring deforestation and environmental crimes, which critics say the government's pro-development policies have encouraged. Since Bolsonaro came to power in 2019, deforestation and fires in the Amazon have risen to their highest levels in more than a decade. "INPE is paying the price of the dismantling of the environmental sector in Brazil," says biologist Izabella Teixeira, who served as environment minister from 2010 to 2016. To her, INPE's crisis reflects the Bolsonaro government's "ideological contempt for the environment and science."

The unruly president has not hidden his dissatisfaction with INPE. "What happens with a lot of INPE reports ... is that they just copy the previous year's reports," he said in 2019. Bolsonaro accused INPE of "lying" about an uptick in forest destruction and had Galvão—a renowned physicist—fired from the director's post for confronting him in public about it. (Galvão was replaced by Darcton Damião, a retired air force colonel with a master's degree in remote sensing from INPE, who was succeeded by De Nardin in October 2020.)

Many INPE employees and researchers who spoke with *Science* believe the institute is a victim of politics, but did not want to say so on the record. "You see that people have a scream stuck in their throats because they don't dare to express their ideas, fearing retaliation," Novo says.

"Little by little, they are making INPE disappear," says Thelma Krug, who worked at the institute for 37 years, where she helped

create the Amazon monitoring programs and had a leading role as an environmental data analyst. Krug, one of the three vice presidents of the Intergovernmental Panel on Climate Change, retired in 2019. "I left because I could not agree to not speaking about what was going on at INPE," she says.

IF INPE STOPS monitoring deforestation and fires, companies and nongovernmental institutions based in Brazil and other countries can fill the gap. But the weakening of the agency is also a weakening of Brazil's

Funding cuts have imperiled the National Institute for Space Research's leading role in weather forecasting and climate modeling in Brazil.

sovereignty, Câmara says: "It is fundamental for the state to have the competence to produce its own data." Many scientists say INPE can only recover if Bolsonaro is defeated during the presidential election in October and a new government takes over. The campaign has yet to begin, and Bolsonaro's main competitor in the polls, former President Luis Inácio Lula da Silva, has not announced specific plans for science, although he has alluded to prioritizing investments in science and education. INPE flourished and Amazon deforestation went down sharply during Lula da Silva's previous presidency, between 2003 and 2010.

Novo is less optimistic. "Even if there is a change in government, it will take a long time for INPE to recover," she says. Novo could have retired 18 years ago but stayed because she loves her job. Now, she is just waiting for her last students to finish their theses. "I want to cry when I think about INPE's destiny," she says, her voice cracking. ∎

Sofia Moutinho is a science journalist in Rio de Janeiro. Herton Escobar is a science journalist in São Paulo.

ENVIRONMENTAL SCIENCE

No model is perfect, but some are useful

Agent-based model should inform the action plan to curb algal blooms in Lake Erie

By **Irina D. Ofițeru**[1] and **Cristian Picioreanu**[2]

Cyanobacterial algal blooms, particularly the ones that produce powerful toxins, are a grave threat to lake ecosystems around the world. One such toxin is microcystin, which was the cause of a drinking water crisis in 2014 in Toledo, Ohio. In an effort to curb algal blooms in Lake Erie, American and Canadian government agencies agreed on a plan to control excessive algal growth in the lake (1). The plan emphasizes the need to reduce phosphorus, but there is a growing body of literature that suggests the need for both phosphorus and nitrogen reduction (2). On page 1001 of this issue, Hellweger et al. (3) report mathematical modeling results that support the latter management strategy. The results suggest that the action plan developed by the US Environmental Protection Agency for only phosphorus reduction may have the opposite effect, increasing the microcystin concentration in Lake Erie.

Hellweger et al. propose an agent-based model of *Microcystis* (a producer of microcystin) that includes a representation of growth and toxin release by the cyanobacterium, which mainly obtains its energy

up has been an issue for agent-based models. Because these models are inherently intricate, they are laborious to develop and require extensive computational power. Moreover, most agent-based models are built in the developer's favorite programming language (Hellweger *et al.* used Fortran). The stochastic features of agent-based models, such as the randomness of the agents' movements, make software errors more difficult to identify, especially when their experimental validation is limited. But, just as one would not expect simplicity from a model of the cosmos, one should also not expect it from a model that involves even simplified metabolic pathways, with different time and space scales. After all, there are 1 billion times as many bacteria on Earth as there are stars in the observable Universe (*6*).

There is no easy solution for scaling up, and compromises must always be made. Hellweger *et al.* simplified the spatial structure of the lake by considering it "well mixed" and therefore homogeneous. This is a limitation of their current model, because spatial structure (e.g., three-dimensional variations in concentrations, temperature, etc.) plays an important role and may appear at different length scales (*7*). The lake forms a complex ecosystem, with multiple microbial species interacting in different ways with the nutrients, and considering more microbial diversity could change the model outcome. Nevertheless, the current assumptions may have been the necessary compromise, dictated also by the multitude of sampling points used in the available experimental databases.

Attempts to integrate "everything" in an agent-based model across multiple scales have been reported before (*8, 9*) but without clear and targeted questions, and because they lack the experimental data for validation, they can be regarded more as proof-of-principle exercises. By comparison, Hellweger *et al.* had a clear question and a plethora of experimental data to test their models. Arguably, their code, although available to be downloaded on Zenodo, will not attract many users for several reasons. It is usually difficult to delve into someone else's code, even when reasonably well documented. Furthermore, even if Fortran is still heavily used by physicists or in large-scale climate models, there exist other more user-friendly or open-access simulators like NetLogo, which uses Scala and Java (*10*); IDynoMiCS, which uses

through photosynthesis. Agent-based models first appeared several decades ago to represent a population of autonomous agents that interact with one another and with their environment, thus generating the emergent properties of the systems (*4*). This type of model has been used in sociology and economics. However, in environmental microbiology, these models were rarely applied in large-scale analyses because of the need for millions of agents to produce statistically relevant results (*5*). This require-

ment made this approach computationally prohibitive. Therefore, agent-based models have not been commonly used to support environmental management policies of high impact.

Hellweger *et al.* took the plunge and integrated a sophisticated agent-based model comprising toxigenic and nontoxigenic *Microcystis* cells with some of their specific biology at the Lake Erie scale. In this process, they made a gargantuan effort to compile, reconcile, and integrate various experimental data, both from the lab and the field. As the authors rightly point out, their model's main advantage is to reproduce the observed patterns reflected in the hundreds of experiments gathered from the literature. The inability to scale

[1]School of Engineering, Newcastle University, Merz Court, Newcastle upon Tyne NE1 7RU, UK. [2]Water Desalination and Reuse Center, Biological and Environmental Sciences and Engineering Division, King Abdullah University of Science and Technology, Thuwal 23955-6900, Saudi Arabia. Email: dana.ofiteru@ncl.ac.uk

Java (*11*); CellModeller, which uses Python (*12*); NUFEB, which uses C++ (*13*); and BioDynaMo, which also uses C++ and is developed with CERN openlab (*14*). However, each of these simulators is serving fairly limited audiences and very likely requires good coding skills for additional development. This is a real limitation in academia, where very few are using software professionals to implement models, though attempts exist (*15*). One can only hope that the model developed by Hellweger *et al.* is an impetus to the modeling community to build more user-friendly and verifiable models that can be used to inform environmental policies.

How sophisticated should the model be? Hellweger *et al.* also propose a simpler empirical model that predicts similar outcomes and suggests that both nitrogen and phosphorus need to be reduced in Lake Erie to control the algal blooms. Critics will rightly refer to the principle of Occam's razor and doubt that the more complicated model is absolutely necessary. The answer to this could be that without access to the ground truth, one cannot know a priori if the simpler model would give a "good enough" representation of the system. For such an important environmental management decision that will affect millions of people's health and their drinking water source, we ought to try both, especially considering that the United States and Canada are due to revise and adjust their domestic action plans for Lake Erie in 2023. Based on the results reported by Hellweger *et al.*, and on the increasing body of literature, the need to also consider nitrogen reduction should be assessed in the control of algal blooms. ∎

REFERENCES AND NOTES

1. US Environmental Protection Agency (EPA), "U.S. action plan for Lake Erie" (EPA, 2018).
2. H. W. Paerl *et al.*, *Hydrobiologia* **847**, 4359 (2020).
3. F. L. Hellweger *et al.*, *Science* **376**, 1001 (2022).
4. M. Griesemer, S. S. Sindi, in *Microbial Systems Biology*, vol. 2349 of *Methods in Molecular Biology*, A. Navid, Ed. (Humana Press, 2022), pp. 367–380.
5. C. Picioreanu, J.-U. Kreft, M. C. M. Van Loosdrecht, *Appl. Environ. Microbiol.* **70**, 3024 (2004).
6. T. P. Curtis, W. T. Sloan, *Curr. Opin. Microbiol.* **7**, 221 (2004).
7. V. Gordon, L. Bakhtiari, K. Kovach, *Phys. Biol.* **16**, 041001 (2019).
8. V. Gogulancea *et al.*, *Front. Microbiol.* **10**, 1871 (2019).
9. I. D. Ofiţeru, M. Bellucci, C. Picioreanu, V. Lavric, T. P. Curtis, *Water Res.* **50**, 382 (2014).
10. S. Tisue, U. Wilensky, paper presented at the Fifth International Conference on Complex Systems, Boston, MA, 16 to 21 May 2004.
11. L. A. Lardon *et al.*, *Environ. Microbiol.* **13**, 2416 (2011).
12. T. J. Rudge, P. J. Steiner, A. Phillips, J. Haseloff, *ACS Synth. Biol.* **1**, 345 (2012).
13. B. Li *et al.*, *PLOS Comput. Biol.* **15**, e1007125 (2019).
14. L. Breitwieser *et al.*, *Bioinformatics* **38**, 453 (2021).
15. D. Vedder, M. Ankenbrand, J. Sarmento Cabral, *Methods Ecol. Evol.* **12**, 2324 (2021).

Delta-scale solutions for human-scale needs

Global studies inform river management needed for landscape sustainability

By **Paola Passalacqua** and **Andrew J. Moodie**

I n 1855, the Yellow River in China suddenly and rapidly changed its path, generating catastrophic flooding in a process known as avulsion. The natural disaster lasted for years and damaged cities and villages along the river for hundreds of kilometers (*1*). A more recent avulsion happened at the Kosi River fan in India in 2008, which generated flooding that killed more than 400 people and displaced 3 million. Are these disasters predictable? Could human lives and livelihoods be saved if information about future avulsions was at hand? The availability of remotely sensed data and analysis techniques has made possible new observations of avulsions. On page 987 of this issue, Brooke *et al.* (*2*) present a global analysis of river avulsions, highlighting controls on avulsions and helping predict where the next avulsion might happen.

Avulsions occur along rivers from mountains to the coast. In confined environments, they are typically located at the mouth of canyons (*3*). In both coastal and inland deltas, which form where a river flows into a body of standing water such as a sea or a lake, avulsions are expected to be located in the backwater region (*4*), which is defined as the portion of a river that responds to changes in the downstream receiving body of water. Brooke *et al.* identified an additional behavior of avulsions in deltas in their global analysis. They analyzed 80 avulsion events on deltas. Of these, 38% were located upstream of where they are normally expected (see the figure). This behavior, primarily found in steep sediment-rich deltaic systems in tropical islands and deserts, where flood-driven erosion can extend upstream of the backwater region, will have wide-reaching effects in the future (*5*). According to the authors, their analysis implies that an increase in sediment down rivers, because of land use and climate change, could cause rivers to shift their avulsion events from the backwater region to more upstream locations. As a result, upstream communities could be

When a river changes its course

Avulsion-driven channel relocation operates at larger spatiotemporal scales than engineering interventions. Sustainability cannot be achieved through superposition of local short-lived interventions and requires a system-scale approach that accounts for human perspective.

1 Avulsions in deltas are expected to happen within the backwater region, which is the portion of the river that responds to changes in the downstream body of water.

2 A newly categorized type of avulsion may occur further upstream than what is commonly expected. This type of avulsion can affect communities that have not been exposed to avulsions. before.

GRAPHIC: N. DESAI/*SCIENCE*

Java (*11*); CellModeller, which uses Python (*12*); NUFEB, which uses C++ (*13*); and BioDynaMo, which also uses C++ and is developed with CERN openlab (*14*). However, each of these simulators is serving fairly limited audiences and very likely requires good coding skills for additional development. This is a real limitation in academia, where very few are using software professionals to implement models, though attempts exist (*15*). One can only hope that the model developed by Hellweger *et al.* is an impetus to the modeling community to build more user-friendly and verifiable models that can be used to inform environmental policies.

How sophisticated should the model be? Hellweger *et al.* also propose a simpler empirical model that predicts similar outcomes and suggests that both nitrogen and phosphorus need to be reduced in Lake Erie to control the algal blooms. Critics will rightly refer to the principle of Occam's razor and doubt that the more complicated model is absolutely necessary. The answer to this could be that without access to the ground truth, one cannot know a priori if the simpler model would give a "good enough" representation of the system. For such an important environmental management decision that will affect millions of people's health and their drinking water source, we ought to try both, especially considering that the United States and Canada are due to revise and adjust their domestic action plans for Lake Erie in 2023. Based on the results reported by Hellweger *et al.*, and on the increasing body of literature, the need to also consider nitrogen reduction should be assessed in the control of algal blooms. ∎

REFERENCES AND NOTES

1. US Environmental Protection Agency (EPA), "U.S. action plan for Lake Erie" (EPA, 2018).
2. H. W. Paerl *et al.*, *Hydrobiologia* **847**, 4359 (2020).
3. F. L. Hellweger *et al.*, *Science* **376**, 1001 (2022).
4. M. Griesemer, S. S. Sindi, in *Microbial Systems Biology*, vol. 2349 of *Methods in Molecular Biology*, A. Navid, Ed. (Humana Press, 2022), pp. 367–380.
5. C. Picioreanu, J.-U. Kreft, M. C. M. Van Loosdrecht, *Appl. Environ. Microbiol.* **70**, 3024 (2004).
6. T. P. Curtis, W. T. Sloan, *Curr. Opin. Microbiol.* **7**, 221 (2004).
7. V. Gordon, L. Bakhtiari, K. Kovach, *Phys. Biol.* **16**, 041001 (2019).
8. V. Gogulancea *et al.*, *Front. Microbiol.* **10**, 1871 (2019).
9. I. D. Ofiţeru, M. Bellucci, C. Picioreanu, V. Lavric, T. P. Curtis, *Water Res.* **50**, 382 (2014).
10. S. Tisue, U. Wilensky, paper presented at the Fifth International Conference on Complex Systems, Boston, MA, 16 to 21 May 2004.
11. L. A. Lardon *et al.*, *Environ. Microbiol.* **13**, 2416 (2011).
12. T. J. Rudge, P. J. Steiner, A. Phillips, J. Haseloff, *ACS Synth. Biol.* **1**, 345 (2012).
13. B. Li *et al.*, *PLOS Comput. Biol.* **15**, e1007125 (2019).
14. L. Breitwieser *et al.*, *Bioinformatics* **38**, 453 (2021).
15. D. Vedder, M. Ankenbrand, J. Sarmento Cabral, *Methods Ecol. Evol.* **12**, 2324 (2021).

Delta-scale solutions for human-scale needs

Global studies inform river management needed for landscape sustainability

By **Paola Passalacqua** *and* **Andrew J. Moodie**

In 1855, the Yellow River in China suddenly and rapidly changed its path, generating catastrophic flooding in a process known as avulsion. The natural disaster lasted for years and damaged cities and villages along the river for hundreds of kilometers (*1*). A more recent avulsion happened at the Kosi River fan in India in 2008, which generated flooding that killed more than 400 people and displaced 3 million. Are these disasters predictable? Could human lives and livelihoods be saved if information about future avulsions was at hand? The availability of remotely sensed data and analysis techniques has made possible new observations of avulsions. On page 987 of this issue, Brooke *et al.* (*2*) present a global analysis of river avulsions, highlighting controls on avulsions and helping predict where the next avulsion might happen.

Avulsions occur along rivers from mountains to the coast. In confined environments, they are typically located at the mouth of canyons (*3*). In both coastal and inland deltas, which form where a river flows into a body of standing water such as a sea or a lake, avulsions are expected to be located in the backwater region (*4*), which is defined as the portion of a river that responds to changes in the downstream receiving body of water. Brooke *et al.* identified an additional behavior of avulsions in deltas in their global analysis. They analyzed 80 avulsion events on deltas. Of these, 38% were located upstream of where they are normally expected (see the figure). This behavior, primarily found in steep sediment-rich deltaic systems in tropical islands and deserts, where flood-driven erosion can extend upstream of the backwater region, will have wide-reaching effects in the future (*5*). According to the authors, their analysis implies that an increase in sediment down rivers, because of land use and climate change, could cause rivers to shift their avulsion events from the backwater region to more upstream locations. As a result, upstream communities could be

When a river changes its course

Avulsion-driven channel relocation operates at larger spatiotemporal scales than engineering interventions. Sustainability cannot be achieved through superposition of local short-lived interventions and requires a system-scale approach that accounts for human perspective.

1 Avulsions in deltas are expected to happen within the backwater region, which is the portion of the river that responds to changes in the downstream body of water.

2 A newly categorized type of avulsion may occur further upstream than what is commonly expected. This type of avulsion can affect communities that have not been exposed to avulsions before.

GRAPHIC: N. DESAI/*SCIENCE*

exposed to avulsion hazards that they have never experienced before.

Avulsions are the ultimate delta process. They operate at the largest time and space scales of a delta system—happening over centuries to millennia and covering the entire delta landscape up to thousands of square kilometers. River avulsions have extremely positive and negative effects on the landscape and its communities. They distribute sediment widely across the entire deltaic region, building and maintaining delta plains. For example, coastal Louisiana was shaped and built through multiple switches in the course of the Mississippi River over millennia (6). However, avulsions also present an obvious hazard to communities living near rivers. Hundreds of millions of people living in coastal areas around the world are at risk of powerful flooding resulting from sudden and catastrophic avulsions, in addition to the constant threat of sea-level rise and other climate crises. The infrequent nature of avulsions, compared to more frequent extreme weather events and the continuous effect of sea-level rise, makes avulsions a less-discussed topic despite their catastrophic effects.

The dichotomy between the dynamics required by river systems to be sustainable and the stability desired by communities living on these systems is the grand challenge for delta management (7). Human civilizations have established themselves in proximity to water, which in turn exposes communities to a combination of acute (e.g., flooding events, avulsions) and chronic (e.g., river migration, sea-level rise) stresses (8). For centuries, societies have implemented various engineering interventions to protect communities and their livelihoods (9), such as building 2500 km of levees along the Mississippi River and extensive embankments that create the low-lying lands on the Ganges-Brahmaputra-Meghna delta. However, as impressive as these engineering feats are, human-scale interventions provide only local and temporary stability relative to the temporal and spatial scale of avulsions. This can contribute to a false sense of protection (10, 11) and even amplify degradation at the system scale by limiting natural sediment dispersal (12–14).

Can engineering interventions be de-signed to address sustainability at the delta system scale while simultaneously providing protection to the communities living on the landscape? Here lies the conflict between the temporal and spatial scales needed for landscape sustainability versus those that are important for human-scale stability. Designing for immediate stability provides short-lived protection over local spatial scales, but these local interventions tend to not achieve landscape sustainability. Interventions to achieve long-term system-scale sustainability should include river diversions that approximate the sediment dispersal achieved by large spatial- and temporal-scale avulsions.

Effective placing of river diversions requires a comprehensive understanding of natural avulsion timing and location. The ability to predict avulsions is improving and moving toward providing robust predictions that governments and decision-makers can use to implement interventions at the system scale. Researchers must focus on integrating process models and remotely sensed global patterns to provide actionable predictions of landscape change over the coming decades. Moreover, river diversion placement must account for the socioeconomic aspects of these interventions and consider how community needs can be integrated with the large spatial and temporal scales of avulsions (15). Only through such an approach will we be able to find a compromise between the human desire for stability and the dynamics that are needed for a landscape to be sustainable over time. ∎

> "Interventions... should include river diversions that approximate the sediment dispersal achieved by... avulsions."

REFERENCES AND NOTES

1. C. Xue, *Mar. Geol.* **113**, 321 (1993).
2. S. Brooke *et al.*, *Science* **376**, 987 (2022).
3. T. C. Blair, J. G. McPherson, *J. Sediment. Res.* **64**, 450 (1994).
4. D. J. Jerolmack, *Quat. Sci. Rev.* **28**, 1786 (2009).
5. J. Best, *Nat. Geosci.* **12**, 7 (2018).
6. D. E. Frazier, *Trans. Gulf Coast Assoc. Geol. Soc.* **27**, 287 (1967).
7. P. Passalacqua *et al.*, *Earth's Future* **9**, e2021EF002121 (2021).
8. M. G. Macklin, J. Lewin, *Quat. Sci. Rev.* **114**, 228 (2015).
9. R. Dodgen, *Controlling the Dragon: Confucian Engineers and the Yellow River in Late Imperial China* (Univ. of Hawaii Press, 2001).
10. G. Di Baldassarre *et al.*, *Hydrol. Sci. J.* **54**, 1007 (2009).
11. S. N. Lane, C. Landström, S. J. Whatmore, *Philos. Trans.- Royal Soc., Math. Phys. Eng. Sci.* **369**, 1784 (2011).
12. J. Pethick, J. D. Orford, *Global Planet. Change* **111**, 237 (2013).
13. S. E. Munoz *et al.*, *Nature* **556**, 95 (2018).
14. L. W. Auerbach *et al.*, *Nat. Clim. Chang.* **5**, 153 (2015).
15. A. J. Moodie, J. A. Nittrouer, *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2101649118 (2021).

Cockrell School of Engineering, The University of Texas at Austin, Austin, TX USA. Email: paola@austin.utexas.edu; amoodie@utexas.edu

MEDICINE

Remote control of the heart and beyond

A resorbable closed-loop sensor-actuator implant can temporarily control heart rate

By **Wolfram-Hubertus Zimmermann**[1,2,3,4,5,6]

External and internal electrical pacing of the heart are fundamental interventions in patients with cardiovascular disease (1). Recently, wearables, such as the Apple Watch and Fitbit devices, have been introduced to the consumer market to monitor key bodily functions such as heart rate and rhythm, blood oxygenation, blood pressure, and body temperature (2). On page 1006 of this issue, Choi *et al.* (3) go beyond sensing by reporting a resorbable closed-loop sensor-actuator (see the figure), with the eventual aim of controlling heart function in patients with a postsurgical risk of bradycardia (slow heart rate). This technology is wireless, circumventing common shortcomings of implanted devices, such as drive-line infections or the need for surgical procedures to remove or replace, for example, pacemaker leads or batteries. The demonstrated cardiac application of this technology in rats, dogs, and ex vivo human heart preparations could improve outpatient surveillance, allowing for earlier release from the hospital and remote monitoring of patients living in medically underserved areas.

Choi *et al.* made use of a previously introduced design strategy, comprising water soluble metals (molybdenum and silicon) and degradable polymers [polyurethane and poly(lactic-co-glycolic acid)], to fabricate resorbable devices (4). In addition to providing preclinical proof of concept

[1]Institute of Pharmacology and Toxicology, University Medical Center Göttingen, Göttingen, Germany. [2]German Center for Cardiovascular Research (DZHK), Partner Site Göttingen, Göttingen, Germany. [3]Cluster of Excellence "Multiscale Bioimaging: from Molecular Machines to Networks of Excitable Cells" (MBExC), University of Göttingen, Göttingen, Germany. [4]German Center for Neurodegenerative Diseases (DZNE), Göttingen, Germany. [5]Fraunhofer Institute for Translational Medicine and Pharmacology (ITMP), Göttingen, Germany. [6]Campus-Institute Data Science (CIDAS), University of Göttingen, Göttingen, Germany. Email: w.zimmermann@med.uni-goettingen.de

exposed to avulsion hazards that they have never experienced before.

Avulsions are the ultimate delta process. They operate at the largest time and space scales of a delta system—happening over centuries to millennia and covering the entire delta landscape up to thousands of square kilometers. River avulsions have extremely positive and negative effects on the landscape and its communities. They distribute sediment widely across the entire deltaic region, building and maintaining delta plains. For example, coastal Louisiana was shaped and built through multiple switches in the course of the Mississippi River over millennia (6). However, avulsions also present an obvious hazard to communities living near rivers. Hundreds of millions of people living in coastal areas around the world are at risk of powerful flooding resulting from sudden and catastrophic avulsions, in addition to the constant threat of sea-level rise and other climate crises. The infrequent nature of avulsions, compared to more frequent extreme weather events and the continuous effect of sea-level rise, makes avulsions a less-discussed topic despite their catastrophic effects.

The dichotomy between the dynamics required by river systems to be sustainable and the stability desired by communities living on these systems is the grand challenge for delta management (7). Human civilizations have established themselves in proximity to water, which in turn exposes communities to a combination of acute (e.g., flooding events, avulsions) and chronic (e.g., river migration, sea-level rise) stresses (8). For centuries, societies have implemented various engineering interventions to protect communities and their livelihoods (9), such as building 2500 km of levees along the Mississippi River and extensive embankments that create the low-lying lands on the Ganges-Brahmaputra-Meghna delta. However, as impressive as these engineering feats are, human-scale interventions provide only local and temporary stability relative to the temporal and spatial scale of avulsions. This can contribute to a false sense of protection (10, 11) and even amplify degradation at the system scale by limiting natural sediment dispersal (12–14).

Can engineering interventions be de-

signed to address sustainability at the delta system scale while simultaneously providing protection to the communities living on the landscape? Here lies the conflict between the temporal and spatial scales needed for landscape sustainability versus those that are important for human-scale stability. Designing for immediate stability provides short-lived protection over local spatial scales, but these local interventions tend to not achieve landscape sustainability. Interventions to achieve long-term system-scale sustainability should include river diversions that approximate the sediment dispersal achieved by large spatial- and temporal-scale avulsions.

Effective placing of river diversions requires a comprehensive understanding of natural avulsion timing and location. The ability to predict avulsions is improving and moving toward providing robust predictions that governments and decision-makers can use to implement interventions at the system scale. Researchers must focus on integrating process models and remotely sensed global patterns to provide actionable predictions of landscape change over the coming decades. Moreover, river diversion placement must account for the socioeconomic aspects of these interventions and consider how community needs can be integrated with the large spatial and temporal scales of avulsions (15). Only through such an approach will we be able to find a compromise between the human desire for stability and the dynamics that are needed for a landscape to be sustainable over time. ∎

> "Interventions... should include river diversions that approximate the sediment dispersal achieved by... avulsions."

REFERENCES AND NOTES

1. C. Xue, *Mar. Geol.* **113**, 321 (1993).
2. S. Brooke *et al.*, *Science* **376**, 987 (2022).
3. T. C. Blair, J. G. McPherson, *J. Sediment. Res.* **64**, 450 (1994).
4. D. J. Jerolmack, *Quat. Sci. Rev.* **28**, 1786 (2009).
5. J. Best, *Nat. Geosci.* **12**, 7 (2018).
6. D. E. Frazier, *Trans. Gulf Coast Assoc. Geol. Soc.* **27**, 287 (1967).
7. P. Passalacqua *et al.*, *Earth's Future* **9**, e2021EF002121 (2021).
8. M. G. Macklin, J. Lewin, *Quat. Sci. Rev.* **114**, 228 (2015).
9. R. Dodgen, *Controlling the Dragon: Confucian Engineers and the Yellow River in Late Imperial China* (Univ. of Hawaii Press, 2001).
10. G. Di Baldassarre *et al.*, *Hydrol. Sci. J.* **54**, 1007 (2009).
11. S. N. Lane, C. Landström, S. J. Whatmore, *Philos. Trans. Royal Soc., Math. Phys. Eng. Sci.* **369**, 1784 (2011).
12. J. Pethick, J. D. Orford, *Global Planet. Change* **111**, 237 (2013).
13. S. E. Munoz *et al.*, *Nature* **556**, 95 (2018).
14. L. W. Auerbach *et al.*, *Nat. Clim. Chang.* **5**, 153 (2015).
15. A. J. Moodie, J. A. Nittrouer, *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2101649118 (2021).

Cockrell School of Engineering, The University of Texas at Austin, Austin, TX USA. Email: paola@austin.utexas.edu; amoodie@utexas.edu

Remote control of the heart and beyond

A resorbable closed-loop sensor-actuator implant can temporarily control heart rate

By **Wolfram-Hubertus Zimmermann**[1,2,3,4,5,6]

External and internal electrical pacing of the heart are fundamental interventions in patients with cardiovascular disease (1). Recently, wearables, such as the Apple Watch and Fitbit devices, have been introduced to the consumer market to monitor key bodily functions such as heart rate and rhythm, blood oxygenation, blood pressure, and body temperature (2). On page 1006 of this issue, Choi *et al.* (3) go beyond sensing by reporting a resorbable closed-loop sensor-actuator (see the figure), with the eventual aim of controlling heart function in patients with a postsurgical risk of bradycardia (slow heart rate). This technology is wireless, circumventing common shortcomings of implanted devices, such as drive-line infections or the need for surgical procedures to remove or replace, for example, pacemaker leads or batteries. The demonstrated cardiac application of this technology in rats, dogs, and ex vivo human heart preparations could improve outpatient surveillance, allowing for earlier release from the hospital and remote monitoring of patients living in medically underserved areas.

Choi *et al.* made use of a previously introduced design strategy, comprising water soluble metals (molybdenum and silicon) and degradable polymers [polyurethane and poly(lactic-co-glycolic acid)], to fabricate resorbable devices (4). In addition to providing preclinical proof of concept

¹Institute of Pharmacology and Toxicology, University Medical Center Göttingen, Göttingen, Germany. ²German Center for Cardiovascular Research (DZHK), Partner Site Göttingen, Göttingen, Germany. ³Cluster of Excellence "Multiscale Bioimaging: from Molecular Machines to Networks of Excitable Cells" (MBExC), University of Göttingen, Göttingen, Germany. ⁴German Center for Neurodegenerative Diseases (DZNE), Göttingen, Germany. ⁵Fraunhofer Institute for Translational Medicine and Pharmacology (ITMP), Göttingen, Germany. ⁶Campus-Institute Data Science (CIDAS), University of Göttingen, Göttingen, Germany. Email: w.zimmermann@med.uni-goettingen.de

for their use in monitoring and pacing the heart, the authors demonstrated important additional features of their device such as magnetic resonance imaging (MRI) compatibility, mechanical robustness, rechargeability, rate-adapted pacing capabilities, and secure data processing.

Triggering electrical pulsing in response to sensing defined biosignals adds a new layer of complexity to human-machine interfacing. But is such a disruptive technology ready to be deployed in clinical practice? There are a number of key issues that need to be addressed: the degree to which the obtained data are reliable, how safety and effectiveness can be ensured, and how misuse can be prevented.

The spectrum of wearable devices is growing rapidly, with the promise that close monitoring of fundamental bodily functions can support lifestyle changes. Whether such devices can be used beyond the consumer market to sustain a healthy lifestyle and, ultimately, to inform life-saving medical interventions remains to be determined by well-controlled clinical investigations with confirmation of safety and effectiveness. Implantable pacemakers, defibrillators, and electrocardiogram (ECG) monitors have been in broad clinical use for decades. More recent developments allow telemonitoring of heart rate and rhythm as well as pulmonary artery pressure by permanently implanted microelectromechanical systems (MEMS) with documented clinical utility in preventing sudden cardiac death and reducing hospitalization for worsening of heart failure (5, 6).

Whether transient devices, made from degradable materials, will obtain similarly reliable data for exploitation as clinical biomarkers while also allowing autonomous organ control will need to be documented in patients. The required information goes beyond what is necessary for the clearance of consumer market wearables as class I and II medical devices by the US Food and Drug Administration (FDA) or CE marking by the European Medicines Agency (EMA). Most notably, clinically obtained evidence for safe and efficacious use in a designated indication, such as bradycardia, will be required for premarket approval of implantable closed-loop recorders, which will be considered class III medical devices.

In the case of using closed-loop sensor-actuator systems with pacemaker functions in patients, risks of inaccurate pacing will have to be determined and minimized (7). This is not a trivial task because such devices must be able to make highly accurate ECG recordings for a clear dissection of signal from noise (a common problem in contemporary pacers and defibrillators) and

Day 1 Bioresorbable pacemaker

10 mm

Day 14

10 mm

Day 33

10 mm

Day 62 No longer visible

10 mm
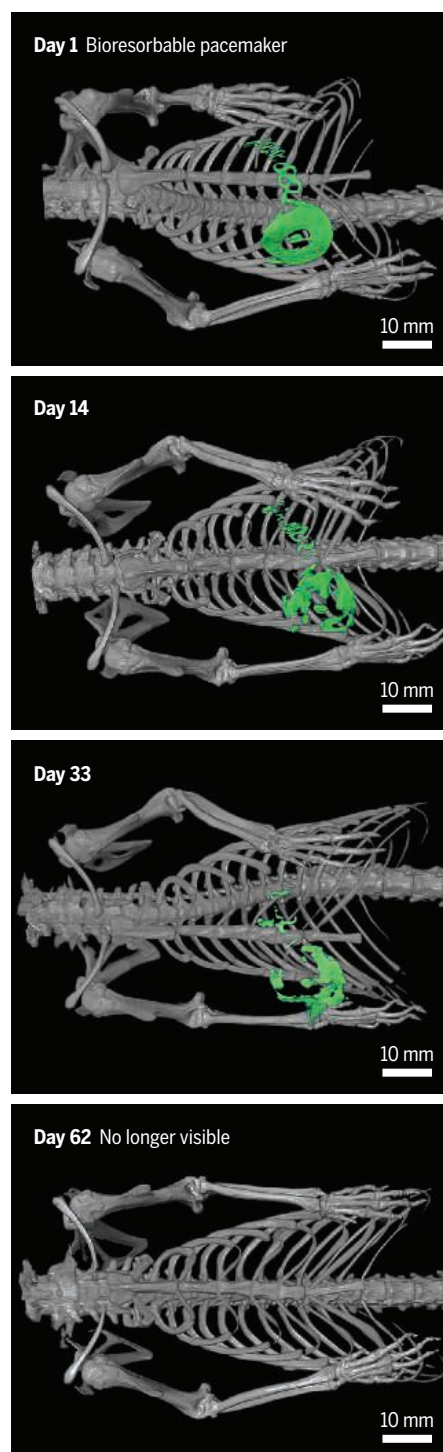
Computed tomography images show resorption of the pacemaker device in rats over the course of 9 weeks.

appropriately manage pacing. It remains to be determined in clinical trials whether the resorption of the implant, foreign body reactions, patient-specific anatomical features of the heart (such as hypertrophy, dilation, scarring, fibrosis, fat deposition), or the presence of other implants will have an impact on real-life ECG signal recordings. Local steroid administration, introduction

of improved materials, and the implementation of explainable artificial intelligence algorithms may help to further improve relevant signal recovery.

Implanted sensors will collect highly personalized data, which could be misused or even manipulated. Patients need to be made aware of the related risks and countermeasures in case of, for example, an unwanted loss of control of the actuator component of the autonomically acting closed-loop device. Moreover, designated data protection measures need to be established that restrict access to recorded information to dedicated and trained medical staff.

The technology described by Choi et al. could have broad applications in sensing and controlling the function not only of the heart but also of other excitable organs or tissue. In some cases, transient support will be desirable and sufficient, such as in patients with transient paralysis. There may be other scenarios where chronic or even permanent use may be of interest, such as in patients with a dysfunctional pacing function of the sinus node (sick sinus syndrome) or permanent paralysis after, for example, traumatic denervation.

Although the use of closed-loop sensor actuators can be readily envisioned as a bridge to recovery, as in the proposed case of postsurgical recovery from bradycardia, additional biological repair may be required to overcome the need for extended electrical support. Support of endogenous and exogenous tissue regeneration and repair of the heart has been an intense topic of research, and a first clinical trial on sustainable remuscularization of the failing heart is underway (i.e., the BioVAT-HF-DZHK20 trial: NCT04396899). By combining electrical and tissue engineering, it seems plausible that electromechanical activity of engrafted cells or tissue may, in the future, be controllable to facilitate electromechanical integration. Similar applications may be envisioned in other diseases caused by the dysfunction of electrically excitable cells. ∎

REFERENCES AND NOTES

1. M. Glikson et al., Eur. Heart J. **42**, 3427 (2021).
2. K. Bayoumy et al., Nat. Rev. Cardiol. **18**, 581 (2021).
3. Y. S. Choi et al., Science **376**, 1006 (2022).
4. Y. S. Choi et al., Nat. Commun. **11**, 5990 (2020).
5. H. Pürerfellner et al., Heart Rhythm **12**, 1113 (2015).
6. M. R. Cowie et al., ESC Heart Fail. **9**, 48 (2022).
7. J. Sperzel, C. W. Hamm, A. Hain, Herzschrittmacherther. Elektrophysiol. **31**, 273 (2020).

CREDIT: BASED ON (3)

Antibiotics made to order

New lipopeptide antibiotics provide hope in the fight against multidrug-resistant bacteria

By **Ryan F. Seipke**

Antimicrobial drug resistance is a global threat to human health. There is an urgent need to discover new antibiotics whose modes of action circumvent prevalent clinical resistance mechanisms. Most antibiotics in clinical use are microbial natural products or their derivatives, whose production is encoded by a biosynthetic gene cluster (BGC) (*1*). Traditional antibiotic discovery strategies involve screening large microbial strain collections for antibiotic activity, followed by a resource-intensive pursuit of pure material for further characterization. This pipeline is hampered by challenges isolating unexplored microbial taxa and because most BGCs are not expressed during laboratory studies (*2, 3*). On page 991 of this issue, Wang *et al.* (*4*) use in silico discovery of BGCs and chemical synthesis of their predicted products to identify a new lipopeptide that is active against multidrug-resistant (MDR) clinical isolates. Another report by this group also used this approach to identify a promising new antibiotic (*5*), highlighting its utility.

Wang *et al.* (*4*) analyzed ~10,000 bacterial genomes, hunting for BGCs encoding lipopeptides, a clinically deployed antibiotic class with diverse modes of action (*6*). The authors prioritized BGCs phylogenetically unrelated to those previously characterized in the hope that they would produce new antibiotics. They identified a distinct lipopeptide BGC harbored by *Paenibacillus mucilaginosus*. Rather than pursue a time-consuming, culture-dependent approach to produce and purify the compound, they capitalized on the power of bioinformatic algorithms to predict possible compounds produced by the enzymatic machinery encoded by the BGC and then chemically synthesized them. They used this so-called "synthetic-bioinformatic natural product (synBNP)" approach to synthesize eight possible compounds predicted from the *P. mucilaginosus* BGC. One compound, which the authors named cilagicin, possessed bactericidal activity against several MDR Gram-positive bacteria. Cilagicin was active against difficult-to-treat *Clostridioides difficile* and vancomycin-resistant enterococci in vitro, which are considered urgent and serious threats by the US Centers for Disease Control and Prevention (CDC) (*7*).

During their experiments, Wang *et al.* (*4*) discovered that a cell wall precursor accumulated in cilagicin-treated cultures. This observation suggested that cilagicin inhibits cell wall biosynthesis, the same target of important classes of antibiotics, such as β-lactams (e.g., penicillins and carbapenems) and glycopeptides (e.g., vancomycin) (*8, 9*). The authors established that cilagicin inhibits cell wall biosynthesis through sequestration of the lipid carrier molecule undecaprenyl phosphate (C55-P) and its inactive form, undecaprenyl pyrophosphate (C55-PP). C55-PP is produced de novo and is dephosphorylated to C55-P during transport of cell wall precursors across the cytoplasmic membrane. Upon delivery of its cargo, C55-P is rephosphorylated and returns to the inner leaflet of the membrane to replenish the dwindling supply of C55-PP. Thus, by sequestering both C55-PP and C55-P, cilagicin blocks bacterial transport of essential cell wall building blocks, which arrests production of the cell wall and ultimately causes cell death.

A handful of other antibiotics can bind to either C55-PP or C55-P, but overall, this mode of action is underexploited, and resistance to antibiotics that target only one of these occurs readily. Notably, Wang *et al.* (*4*) did not observe evolution of resistance to cilagicin over the course of a 25-day experiment in which *Staphylococcus aureus* was serially passaged in culture medium containing a subinhibitory concentration of cilagicin, whereas resistance readily developed to bacitracin or amphomycin, agents that bind only to C55-PP or C55-P, respectively (*10, 11*). The

Discovery of antibiotic candidates

Bioinformatics was used to identify biosynthetic gene clusters (BGCs) within bacterial genomes that produce lipopeptide antibiotics. Chemical synthesis of the predicted compounds from the cilagicin and macolacin BGCs and further chemical optimization resulted in two new promising antibiotics.

Bioinformatics

Cilagicin biosynthetic gene cluster

Macolacin biosynthetic gene cluster

Chemical synthesis of predicted natural products

Cilagicin

Macolacin

Optimization of the lipid moiety

Cilagicin-BP

Biphenyl-macolacin

Promising antibiotics active against multidrug-resistant bacteria

Faculty of Biological Sciences, Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds, UK. Email: r.seipke@leeds.ac.uk

lack of detectable resistance to cilagicin is likely linked to its ability to bind both C55-PP and C55-P, because changes to two distinct molecular targets must evolve for resistance to develop. The binding to multiple targets may be an important consideration when developing future antibiotics.

The same research team also recently used their synBNP approach to overcome colistin resistance (5). Resistance to colistin, another lipopeptide antibiotic, raised substantial concern when a resistance determinant encoded by a gene called *mcr-1* (mobilized colistin resistance 1) spread rapidly in pathogenic enteric bacteria around the globe. Widespread dissemination of the *mcr-1* gene jeopardized the utility of colistin as the last line of defense against infections caused by MDR Gram-negative bacteria (12). Gram-negative bacteria have a cell wall that is encased by an additional lipid outer membrane, which is a permeability barrier to many small molecules. This limits the number of antibiotics in our anti-Gram-negative arsenal that target the cell wall or other targets within. Colistin has potent antibiotic activity against Gram-negative bacteria because it binds to lipopolysaccharides and phospholipids in the outer membrane, displacing divalent cations, which disrupts membrane integrity and ultimately leads to cell death (13, 14).

Wang *et al.* (5) set out to identify BGCs that encoded the production of colistin analogs, with the clever rationale that nature may have figured out how to diversify the antibiotic to overcome resistance. Like their work with cilagicin, the authors focused their attention on a single BGC and synthesized its predicted product, which they named macolacin, which possessed antibacterial activity against colistin-resistant bacteria. The authors were able to further improve macolacin activity by optimizing its lipid moiety (see the figure), which facilitates interaction with the membrane. One improved derivative, biphenyl-macolacin, outperformed the parent molecule and possessed potent in vitro activity against intrinsically colistin-resistant *Neisseria gonorrhoeae,* and carbapenem-resistant and extensively drug-resistant *Acinetobacter baumannii,* which are recognized as urgent threats by the CDC (7).

Although cilagicin and macolacin showed promising in vitro activity against problematic MDR bacterial pathogens, the real question was how well do these agents perform in an infection model? The answer will dictate the future of these agents as therapeutics. In both studies, Wang *et al.* (4, 5) assessed each of these compounds in a mouse infection model. For cilagicin, there was an initial setback. High levels of serum binding to cilagicin blocked its antibacterial activity. The authors overcame this hurdle by altering the lipid component of cilagicin, ultimately utilizing the same biphenyl moiety used to improve macolacin, as a strategy to reduce serum binding. Such hit-to-lead optimization is a key feature in antibiotic development (15). The new structure, cilagicin-BP (see the figure), was as efficacious as vancomycin when used to treat mice infected with MDR *S. aureus* and even more so when used to treat *Streptococcus pyogenes* infection (which is not MDR). The efficacy of biphenyl-macolacin was also evaluated in mice infected with either carbapenem-resistant *A. baumannii* engineered to express the *mcr-1* colistin resistance gene, or an *mcr-1*–expressing clinical isolate of *A. baumannii* that is resistant to all antibiotics tested. Treatment with colistin did not reduce the bacterial load beyond that used to establish the infection, whereas treatment with biphenyl-macolacin reduced the bacterial load by five orders of magnitude.

Many promising antibiotic compounds fall by the wayside because of low production titer during microbial fermentation. Aside from a rare handful of compounds, chemical synthesis is ultimately used to produce the quantity, and notably, the chemical diversity of analogs necessary to define the clinical potential of a lead pharmacophore. In two studies, Wang *et al.* not only produced two new biologically inspired antibiotics but established a route for their synthesis and generation of analogs. The next major steps for their development are absorption, distribution, metabolism, excretion, and toxicity studies, which may reveal the need for further structural optimization before entry into clinical trials. Although clinical deployment of cilagicin and macolacin may take time, Wang *et al.* (4, 5) have established an inspirational interdisciplinary roadmap for future antibiotic discovery that may tip the scales in our fight against antimicrobial resistance. ∎

REFERENCES AND NOTES
1. D. J. Newman, G. M. Cragg, *J. Nat. Prod.* **83**, 770 (2020).
2. P. A. Hoskisson, R. F. Seipke, *mBio* **11**, e02642 (2020).
3. L. L. Ling *et al.*, *Nature* **517**, 455 (2015).
4. Z. Wang *et al.*, *Science* **376**, 991 (2022).
5. Z. Wang *et al.*, *Nature* **601**, 606 (2022).
6. I. W. Hamley, *Chem. Commun.* **51**, 8574 (2015).
7. Centers for Disease Control and Prevention, "Antibiotic Resistance Threats in the United States, 2019" (2019); www.cdc.gov/drugresistance/biggest-threats.html.
8. K. Kitano A. Tomasz, *Antimicrob. Agents Chemother.* **16**, 838 (1979).
9. M. A. T. Blaskovich *et al.*, *ACS Infect. Dis.* **4**, 715 (2018).
10. K. J. Stone, J. L. Strominger, *Proc. Natl. Acad. Sci. U.S.A.* **68**, 3223 (1971).
11. T. Schneider *et al.*, *Antimicrob. Agents Chemother.* **53**, 1610 (2009).
12. Y.-Y. Liu *et al.*, *Lancet Infect. Dis.* **16**, 161 (2016).
13. M. Schindler, M. J. Osborn, *Biochemistry* **18**, 4425 (1979).
14. A. Sabnis *et al.*, *eLife* **10**, e65836 (2021).
15. M. Miethke *et al.*, *Nat. Rev. Chem.* **5**, 726 (2021).

How full is the evolutionary fuel tank?

A meta-analysis quantifies the heritable genetic variance in fitness—the fuel of evolution

By **Bruce Walsh**

In 1898, Hermon Bumpus gathered 136 house sparrows immobilized by an ice storm, noting that the averages of several morphological traits differed between survivors and nonsurvivors. This was one of the first attempts to measure the phenotypic selection component of Charles Darwin's thesis, that adaptation is driven by heritable traits that affect fitness. Since then, a vast literature on quantifying associations between trait values and fitness has emerged (1). The quantification of Darwin's second evolution component—that such traits are heritable—required the development of quantitative genetics by Ronald Fisher in 1918 (2). Although the selection and genetics components can be combined to determine the expected change in any trait, of greater interest is the general adaptive potential of a population. On page 1012 of this issue, Bonnet *et al.* (3) present a meta-analysis of 19 studies showing the abundance of heritable variations in fitness and the potential for adaptation.

Fisher famously stated that "natural selection is not evolution," meaning that if a trait is not heritable, no amount of selection will result in a change in the offspring of surviving parents. Fisher's key to deciphering heritability was noting that parents pass along specific variants of a gene (alleles), rather than entire genotypes, to their offspring. The sum of all the single-allele effects for a given trait carried by an individual is defined as their breeding value (BV) for that trait. BVs are best understood in terms of deviations from the mean, so that a random individual has an expected BV of zero, which implies that its offspring will, on average, be average. The expected deviation of an offspring from the population

Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ, USA. Email: jbwalsh@u.arizona.edu

lack of detectable resistance to cilagicin is likely linked to its ability to bind both C55-PP and C55-P, because changes to two distinct molecular targets must evolve for resistance to develop. The binding to multiple targets may be an important consideration when developing future antibiotics.

The same research team also recently used their synBNP approach to overcome colistin resistance (5). Resistance to colistin, another lipopeptide antibiotic, raised substantial concern when a resistance determinant encoded by a gene called *mcr-1* (mobilized colistin resistance 1) spread rapidly in pathogenic enteric bacteria around the globe. Widespread dissemination of the *mcr-1* gene jeopardized the utility of colistin as the last line of defense against infections caused by MDR Gram-negative bacteria (12). Gram-negative bacteria have a cell wall that is encased by an additional lipid outer membrane, which is a permeability barrier to many small molecules. This limits the number of antibiotics in our anti-Gram-negative arsenal that target the cell wall or other targets within. Colistin has potent antibiotic activity against Gram-negative bacteria because it binds to lipopolysaccharides and phospholipids in the outer membrane, displacing divalent cations, which disrupts membrane integrity and ultimately leads to cell death (13, 14).

Wang *et al.* (5) set out to identify BGCs that encoded the production of colistin analogs, with the clever rationale that nature may have figured out how to diversify the antibiotic to overcome resistance. Like their work with cilagicin, the authors focused their attention on a single BGC and synthesized its predicted product, which they named macolacin, which possessed antibacterial activity against colistin-resistant bacteria. The authors were able to further improve macolacin activity by optimizing its lipid moiety (see the figure), which facilitates interaction with the membrane. One improved derivative, biphenyl-macolacin, outperformed the parent molecule and possessed potent in vitro activity against intrinsically colistin-resistant *Neisseria gonorrhoeae,* and carbapenem-resistant and extensively drug-resistant *Acinetobacter baumannii,* which are recognized as urgent threats by the CDC (7).

Although cilagicin and macolacin showed promising in vitro activity against problematic MDR bacterial pathogens, the real question was how well do these agents perform in an infection model? The answer will dictate the future of these agents as therapeutics. In both studies, Wang *et al.* (4, 5) assessed each of these compounds in a mouse infection model. For cilagicin, there was an initial setback. High levels of serum binding to cilagicin blocked its antibacterial activity. The authors overcame this hurdle by altering the lipid component of cilagicin, ultimately utilizing the same biphenyl moiety used to improve macolacin, as a strategy to reduce serum binding. Such hit-to-lead optimization is a key feature in antibiotic development (15). The new structure, cilagicin-BP (see the figure), was as efficacious as vancomycin when used to treat mice infected with MDR *S. aureus* and even more so when used to treat *Streptococcus pyogenes* infection (which is not MDR). The efficacy of biphenyl-macolacin was also evaluated in mice infected with either carbapenem-resistant *A. baumannii* engineered to express the *mcr-1* colistin resistance gene, or an *mcr-1*–expressing clinical isolate of *A. baumannii* that is resistant to all antibiotics tested. Treatment with colistin did not reduce the bacterial load beyond that used to establish the infection, whereas treatment with biphenyl-macolacin reduced the bacterial load by five orders of magnitude.

Many promising antibiotic compounds fall by the wayside because of low production titer during microbial fermentation. Aside from a rare handful of compounds, chemical synthesis is ultimately used to produce the quantity, and notably, the chemical diversity of analogs necessary to define the clinical potential of a lead pharmacophore. In two studies, Wang *et al.* not only produced two new biologically inspired antibiotics but established a route for their synthesis and generation of analogs. The next major steps for their development are absorption, distribution, metabolism, excretion, and toxicity studies, which may reveal the need for further structural optimization before entry into clinical trials. Although clinical deployment of cilagicin and macolacin may take time, Wang *et al.* (4, 5) have established an inspirational interdisciplinary roadmap for future antibiotic discovery that may tip the scales in our fight against antimicrobial resistance. ∎

REFERENCES AND NOTES

1. D. J. Newman, G. M. Cragg, *J. Nat. Prod.* **83**, 770 (2020).
2. P. A. Hoskisson, R. F. Seipke, *mBio* **11**, e02642 (2020).
3. L. L. Ling *et al.*, *Nature* **517**, 455 (2015).
4. Z. Wang *et al.*, *Science* **376**, 991 (2022).
5. Z. Wang *et al.*, *Nature* **601**, 606 (2022).
6. I. W. Hamley, *Chem. Commun.* **51**, 8574 (2015).
7. Centers for Disease Control and Prevention, "Antibiotic Resistance Threats in the United States, 2019" (2019); www.cdc.gov/drugresistance/biggest-threats.html.
8. K. Kitano A. Tomasz, *Antimicrob. Agents Chemother.* **16**, 838 (1979).
9. M. A. T. Blaskovich *et al.*, *ACS Infect. Dis.* **4**, 715 (2018).
10. K. J. Stone, J. L. Strominger, *Proc. Natl. Acad. Sci. U.S.A.* **68**, 3223 (1971).
11. T. Schneider *et al.*, *Antimicrob. Agents Chemother.* **53**, 1610 (2009).
12. Y.-Y. Liu *et al.*, *Lancet Infect. Dis.* **16**, 161 (2016).
13. M. Schindler, M. J. Osborn, *Biochemistry* **18**, 4425 (1979).
14. A. Sabnis *et al.*, *eLife* **10**, e65836 (2021).
15. M. Miethke *et al.*, *Nat. Rev. Chem.* **5**, 726 (2021).

How full is the evolutionary fuel tank?

A meta-analysis quantifies the heritable genetic variance in fitness—the fuel of evolution

By **Bruce Walsh**

I n 1898, Hermon Bumpus gathered 136 house sparrows immobilized by an ice storm, noting that the averages of several morphological traits differed between survivors and nonsurvivors. This was one of the first attempts to measure the phenotypic selection component of Charles Darwin's thesis, that adaptation is driven by heritable traits that affect fitness. Since then, a vast literature on quantifying associations between trait values and fitness has emerged (1). The quantification of Darwin's second evolution component—that such traits are heritable—required the development of quantitative genetics by Ronald Fisher in 1918 (2). Although the selection and genetics components can be combined to determine the expected change in any trait, of greater interest is the general adaptive potential of a population. On page 1012 of this issue, Bonnet *et al.* (3) present a meta-analysis of 19 studies showing the abundance of heritable variations in fitness and the potential for adaptation.

Fisher famously stated that "natural selection is not evolution," meaning that if a trait is not heritable, no amount of selection will result in a change in the offspring of surviving parents. Fisher's key to deciphering heritability was noting that parents pass along specific variants of a gene (alleles), rather than entire genotypes, to their offspring. The sum of all the single-allele effects for a given trait carried by an individual is defined as their breeding value (BV) for that trait. BVs are best understood in terms of deviations from the mean, so that a random individual has an expected BV of zero, which implies that its offspring will, on average, be average. The expected deviation of an offspring from the population

Ecology and Evolutionary Biology, University of Arizona, Tuscon, AZ, USA. Email: jbwalsh@u.arizona.edu

mean is simply the average of the BVs of its parents. As a result, parents with exceptional BVs have offspring that, on average, deviate substantially from the population mean. Conversely, offspring from parents with modest BVs fall close to the mean.

The spread of individual BVs is a measure of the evolutionary potential of a trait. This is the basis of the additive variance of a trait, defined as the variance among BVs for that trait in a given population. If this variance is small, offspring have very little resemblance to their parents, whereas if it is large, exceptional parents tend to have exceptional offspring. If there is no additive variance for a trait, it will not evolve. More generally, if there is no additive variance in fitness, no trait will genetically respond to selection.

Bonnet *et al*. used total number of offspring, also known as lifetime reproductive success (LRS), as the measure for the fitness of an individual. The LRS parameter is converted to relative fitness simply by dividing LRS for an individual by the average LRS of the population, which allows for quantifiable comparisons across studies. In statistical terms, if a population shows a significant additive variance in fitness among its individuals, this implies that parents with higher LRSs than the population average also have a high BV for LRS, and thus their children also tend to have high LRSs.

Estimating BVs, and thus additive variance, is a common problem in modern animal breeding, built around using pedigree information. A BV exists even when the trait is not displayed, as it is a measure of

domesticated population to a small wild population has been somewhat rocky (*4*). Domesticated pedigrees tend to be much deeper and denser than those for natural populations, resulting in greater precision in BV estimates. Furthermore, fitness is a problematic trait for standard pedigree methods, which assume trait values are continuous and follow a Gaussian distribution, whereas fitness data are highly discrete—a parent can only have an integer number of offspring, with a large point mass at zero, that is, individuals with zero offspring. Although there have been a few attempts to estimate the additive variance in fitness in wild populations using standard pedigree methods, the failure of the Gaussian assumption suggests that these are likely rather biased.

Bonnet *et al*. extended these pedigree methods by using a discrete Poisson distribution with an inflated zero value instead of a Gaussian and provided a much better fit for the fitness data. Using the improved fitting, their resulting average estimate of the additive variance in relative fitness, $V_A(w)$, was two- to fourfold larger than previous values. To put it in a more tangible context, this means that if the fitness of a population drops by a third, it would take roughly 10 generations to recover back to normal fitness levels. Hence, populations with shorter generation times might have a better chance to somewhat mitigate anthropogenic changes.

In nature, the target of selection is almost certainly a constantly shifting, high-dimensional (i.e., multi-trait) phenotype that may poorly project onto individual traits or even a set of traits. Most studies of adaptation are structured around some assumed edifice of traits that affects fitness. A poor choice of traits can give a misleading impression of population adaptation. Fortunately, an estimate of $V_A(w)$ provides an upper bound, and therefore a maximal possible change in any trait independent of selection. For example, a typical trait heritability of 0.3 will mean that 30% of the trait variation is due to variance in BVs, and the maximal possible change in the average value of a trait in the population is about one standard deviation every four generations. A more reliable way to estimate $V_A(w)$ can help to better quantify the nature of selection and the robustness of a population to major environmental changes. ■

South African meerkats (*Suricata suricatta*) were among the populations examined by Bonnet *et al*., who used the total number of pups in a mother's lifetime as the measure of her reproductive fitness.

Thus, one of the holy grails in evolutionary genetics is to estimate the additive variation in fitness itself, which gives a general measure of the evolutionary potential of a population and places limits on the maximal response for any trait. This challenging estimation problem was tackled by Bonnet *et al*. using a collection of 19 long-term vertebrate population studies (covering a total of 561 cohorts and ~250,000 individuals) from North America, Europe, Africa, and Oceania. The meta-analysis showcases an immense, but doable, effort in estimating this fundamental evolutionary parameter.

how exceptional an offspring from that parent would be, if produced. In the case of milk production, information on the BV of a bull is provided by the observed yields of his mother, sisters, and daughters. The same pedigree machinery used by breeders can, in theory, be applied in natural populations to estimate the additive variance of any measured trait. Pedigrees for natural populations can be constructed using molecular markers, and closed populations of vertebrates are well suited for such analyses. Even with perfect pedigrees, the transition of pedigree methods from a large and well-structured

REFERENCES AND NOTES

1. B. Walsh, M. Lynch, *Evolution and Selection of Quantitative Traits* (Oxford Univ. Press, 2018).
2. M. Lynch, B. Walsh, *Genetics and Analysis of Quantitative Traits* (Sinauer, 1998).
3. T. Bonnet *et al*., *Science* **376**, 1012 (2022).
4. J. D. Hadfield, A. J. Wilson, D. Garant, B. C. Sheldon, L. E. B. Kruuk, *Am. Nat.* **175**, 116 (2010).

10.1126/science.abo4624

CLIMATE POLICY

Actions for reducing US emissions at least 50% by 2030

Policies must help decarbonize power and transport sectors

By **John Bistline[1], Nikit Abhyankar[2], Geoffrey Blanford[1], Leon Clarke[3,4], Rachel Fakhry[5], Haewon McJeon[3,6], John Reilly[7], Christopher Roney[1], Tom Wilson[1], Mei Yuan[7], Alicia Zhao[3]**

Pathways for limiting global warming to 1.5° and 2°C generally involve net-zero greenhouse gas (GHG) emissions economy-wide near mid-century and halving emissions over the next decade (*1*). Updated pledges by countries and companies around the 2021 United Nations climate conference reflect this sense of urgency. The updated US pledge to reduce net emissions 50 to 52% by 2030 (*2*) would represent a tripling of the pace of historical reductions (see materials and methods S3). We report on a six-model intercomparison of potential actions to reach the US target of at least 50% GHG reductions by 2030. This analysis helps identify which findings are more robust or uncertain given different model structures and input assumptions [see supplementary materials (SM)]. Models highlight the central roles of clean electricity and electrification, the large scale of deployment needed relative to historical levels and scenarios with only current policies, and a range of benefits from near-term action.

The US pledge reflects a much more stringent 2030 target than has been previously analyzed. The existing literature has primarily focused on the actions needed to achieve longer-term deep decarbonization goals in the 2050 time frame (*3*) and has provided less detailed analysis of pathways in the next decade, a gap addressed by several recent studies (*4–9*). The six models in those studies (the basis of this comparison) are among the most widely applied and detailed models of the US energy system, which make them well-suited to provide information for policy-makers and other stakeholders on concrete actions to support nearer-term tar-

[1]Electric Power Research Institute, Palo Alto, CA, USA.
[2]Lawrence Berkeley National Laboratory, Berkeley, CA, USA.
[3]University of Maryland, College Park, MD, USA. [4]Bezos Earth Fund, Washington, DC, USA. [5]Natural Resources Defense Council, New York, NY, USA. [6]Pacific Northwest National Laboratory, College Park, MD, USA. [7]MIT Joint Program on the Science and Policy of Global Change, Cambridge, MA, USA.
Email: jbistline@epri.com

gets. Some models use a top-down approach to identify least-cost emission reduction actions, whereas others use a bottom-up, sector-specific suite of measures and incentives to reflect policy proposals. Comparing the scenarios modeled here with ones that represent current policies and technological trends indicates the magnitude of implementation gaps that need to be closed through strengthened policies and incentives in the years until 2030. Insights from this modeling may be relevant for other countries, suggesting initial steps they can take toward more sustainable, affordable, reliable, and equitable clean energy transitions.

ROBUST ACTIONS, ON TARGET
The bulk of reductions come from the electric sector and transport

Although specific sectoral contributions vary, all models studied indicate that most GHG reductions by 2030 come from the power and transportation sectors, which account for 69 to 89% of reductions (see the first figure). A highly consistent finding is the large role of the electric power sector in accelerating change through direct emissions reductions, primarily through fuel switching, and through end-use electrification to reduce fossil fuel use and emissions in transport, industry, and buildings. Direct CO_2 reductions from the power sector account for 48 to 66% of total 2030 reductions across models.

All sectors are involved in reaching the 2030 target (see the first figure), including enhancing the land sink and reducing non-CO_2 GHG emissions (see SM). Some sectors have lower-cost reductions available and are assumed to be able to move more quickly (e.g., electrification of light-duty vehicles, whereas others need to be set up early to make deeper reductions feasible and affordable in later decades. The relative roles of economic sectors depend on the time scale used for assessing impacts of short-lived climate pollutants. Although this analysis uses standard 100-year Global Warming Potentials to align with Paris Agreement accounting requirements, mitigation opportunities for methane-dominated sectors such as agriculture and fuels production would be increased, and mitiga-

tion for CO_2-dominated sectors such as electricity and transport would be lowered (*10*), if considering impacts over shorter horizons.

Energy efficiency (broadly conceived as reductions in energy use per unit of service demand or economic activity), cleaner electricity, and rapid electrification are key pillars for both near- and long-term emissions reductions, though models differ on how fast change can occur in each sector and the relative roles of these factors in reducing 2030 emissions (figs. S3 and S4).

The scale and pace of transformations to reach the 2030 target require immediate and sustained efforts

In the power sector, the average annual additions of wind and solar capacity increase by two to seven times their historical levels in the last decade to meet the 2030 target (29 to 91 GW/year across models, as shown in the second figure). Another model-consistent finding is that coal capacity retirements meet or exceed historical levels, which lead to ~90 to 100% reductions in coal generation by 2030 (fig. S12). Half of the participating models also deploy gas with carbon capture and sequestration (CCS) by 2030, increasing from 0 GW today to 0 to 70 GW by 2030. These transformations exceed the pace and degree of change projected to occur with current federal and state policies (fig. S5).

Despite broad agreement on the need for substantial power sector decarbonization, models differ in the degree of investment in various low-carbon technologies. Key differences include the level of electrification (fig. S8), share of electricity generated by renewables (fig. S12), ratio of wind to solar builds (see the second figure), extent of new gas capacity builds (to replace retiring coal capacity and balance renewables), role of emerging technologies (e.g., CCS), and extent of infrastructure buildout (e.g., transmission). Variations in investment mixes across models are due to a combination of differences in input assumptions and model structure such as temporal resolution, capital costs of generation options, electricity demand, constraints on technological choice sets, and various policy measures applied to achieve the target reduction level (see SM).

Technologies for decarbonizing end uses exhibit similarly rapid deployment. Electric vehicle (EV) shares as a fraction of new light-duty sales increase from around 4% in 2021 to 34 to 100% by 2030 with an average of 67% (fig. S10). The Biden administration set a 50% EV sales target by 2030 (*11*), but model results suggest that EV deployment may have to exceed that level to reach the 2030 target. These shares are higher than EV shares in 2030 with current policies and incentives (16 to 77% with a 38%

average) (fig. S11). Despite rapid growth in new sales for passenger vehicles and other end-use technologies, turnover dynamics and inertia mean that the stock share and emissions impact lag new sales shares (fig. S11). In addition to transport, many of the roughly 150 million US households and businesses would make investments related to space heating and cooling, water heating, efficiency improvements, and appliance purchases over the next decade, leading to electricity's share of final energy across the economy to increase from 21% today to 25 to 35% by 2030 (fig. S9).

POLICY AND IMPLEMENTATION
Strong policies are needed, and many different combinations of policies and incentives can be used

With current policies and technological trends (i.e., model "reference" scenarios), continued electric sector emissions reductions, energy efficiency improvements, and vehicle electrification lead to 6 to 28% reductions in energy-related CO_2 emissions by 2030, falling far short of the 50% below 2005 emissions target (fig. S2).

Modeling efforts include a range of policy levers to close this gap (table S2), including clean energy tax credits, electric sector standards, end-use equipment rebates, efficiency standards, and carbon pricing. Some models rely primarily on an economy-wide GHG cap or carbon price to reach the 2030 goal, whereas others use a broader suite of options. Estimates of the marginal cost of CO_2 reductions, which reflect the cost of the last, or most expensive, ton of CO_2 reduced to meet the emission goal, provide an indicator of the level of policy stringency required to reach the target. For those studies using carbon pricing instruments, the marginal abatement cost ranges from $36 to $155 per ton of CO_2 in 2030 across models with an average of $84 per ton of CO_2 (table S3).

The large range of carbon prices reflects differences in model assumptions about technological costs and policies (see materials and methods S5 and table S2). To the extent that other measures such as vehicle standards, state and local policies, and tax credits are also included, they can reduce the CO_2 price, making it an incomplete measure of the policy support needed (12).

That many combinations of policies and incentives can be used to reach the 2030 target indicates that there are many pathways to halving emissions, though questions remain about which options are effective and politically durable, as well as about the roles of different actors (e.g., federal, state, and other subnational policies and their interactions).

Ensuring grid dependability will be key, and post-2030 reductions will be needed from other sectors

All scenarios indicate high shares of solar and wind technologies (fig. S5), plant closures (fig. S7), and an increasing reliance on electricity as a result of electrification (fig. S8). These drivers suggest that the dependability of power systems (including resource adequacy, stability, and resiliency) will be an ongoing focus for planners, system operators, and policy-makers. The substantial decline in coal use (fig. S12) also requires attention to a just transition for individuals and communities affected by these changes and their complex political economy (13). Additionally, given the large and rapid buildouts of resources, beyond historic levels, required to meet the 2030 target (see the second figure), institutional innovation will likely be needed to deploy commercial technologies at a faster pace, including expediting siting and permitting.

Actions in electricity and other sectors to meet the 2030 target are largely focused on deploying existing technologies and taking advantage of technological progress from previous decades (e.g., cost reductions and performance improvements in renewables, EVs, heat pumps), which were facilitated by policies and incentives to encourage innovation and early adoption. Continued encouragement of innovation is critical for making nascent technologies ready to scale to meet post-2030 targets. A largely decarbonized electric sector can help reduce emissions in other sectors directly through electrification and indirectly through electricity-derived fuels (14).

IMPLICATIONS

Models agree that near-term actions to reduce GHG emissions to meet the 2030 target can produce a range of benefits, and studies highlight several categories.

Magnitudes of GHG reductions are similar across models (see the first figure), though few studies explicitly monetize these benefits (or, conversely, the costs from inaction) by multiplying these changes by the social cost of GHGs or similar approaches, which can estimate damages from rising temperatures, extreme weather events, and other climate impacts. However, one analysis reaching a 53% reduction in net GHGs by 2030 indicates annual climate benefits of ~$140 billion per year (7). Near-term action, especially reductions of high–warming potential GHGs such as methane, can reduce the rate of worsening climate impacts and probabilities of triggering positive climate feedbacks (6).

Meeting the 2030 GHG target has side benefits of substantial reductions in non-CO_2 air pollutants such as sulfur dioxide (SO_2), nitrogen oxides (NO_x), and particulate matter, reducing deaths and illnesses from air pollution (5, 9). For instance, studies estimate the health benefits from reducing SO_2 and NO_x from the power sector alone to equal tens of billion of dollars annually by 2030 (7, 15). Accelerated electrification can amplify the air quality benefits of electric sector decarbonization. Many 2030 target studies quantify reductions in these pollutant emissions (5, 7, 9), and some have included estimates of the value of reductions in death and illness from pollutant exposure (7, 9). Individual studies discuss additional benefits that a 2030 target and associated policies could bring, including increasing jobs (4, 6, 7), encouraging technological progress and innovation (5, 6), boosting international competitiveness (6), and improving distributional outcomes for lower-income households (9).

Emissions reductions by sector and model

Historical emissions and 100-year Global Warming Potential values are based on the US Environmental Protection Agency's "Inventory of US Greenhouse Gas Emissions and Sinks." "Other CO_2" refers to non-energy CO_2 emissions where specified.

Legend: Land · Non-CO_2 Greenhouse gases (GHGs) · Other CO_2 · Transport · Buildings · Industry · Electric

GHG reductions from 2005 by sector (Mt-CO_2e)

(chart with y-axis categories: 2019, and for 2030: EDF-NEMS, GCAM-USA-AP, LBNL, PATHWAYS, REGEN, REGEN E+, USREP-ReEDS; top x-axis −4000 to 500; bottom x-axis Net GHG reductions from 2005 (% change) −60 to 7.5)

Net GHG reductions from 2005 (% change)
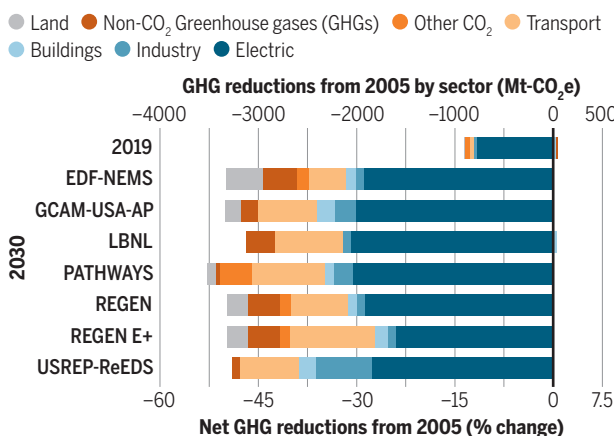
Mt-CO_2e, metric tons of CO_2 equivalent; EDF-NEMS, Environmental Defense Fund–National Energy Modeling System (6); GCAM-USA-AP, Global Change Analysis Model for the US (8); LBNL, Lawrence Berkeley National Laboratory models (4); PATHWAYS, Regional Investment and Operations Model supply-side model and EnergyPATHWAYS demand-side model (7); REGEN, US Regional Economy, Greenhouse Gas, and Energy model (5); REGEN E+, US-REGEN model with accelerated electrification (5); USREP-ReEDS, US Regional Energy Policy-Regional Energy Deployment System model (9).

Electric sector transitions by technology

Historical and projected electric sector capacity additions are shown. Projections show the average annual build rates for utility-scale capacity through 2030. Historical values come from Form EIA-860 data. Additional information on participating models and study assumptions can be found in materials and methods S1 and S2.

Legend:
- Other
- Hydro
- Nuclear
- Gas CCS
- Gas
- Coal
- Storage
- Solar
- Wind

Average build rate to 2030
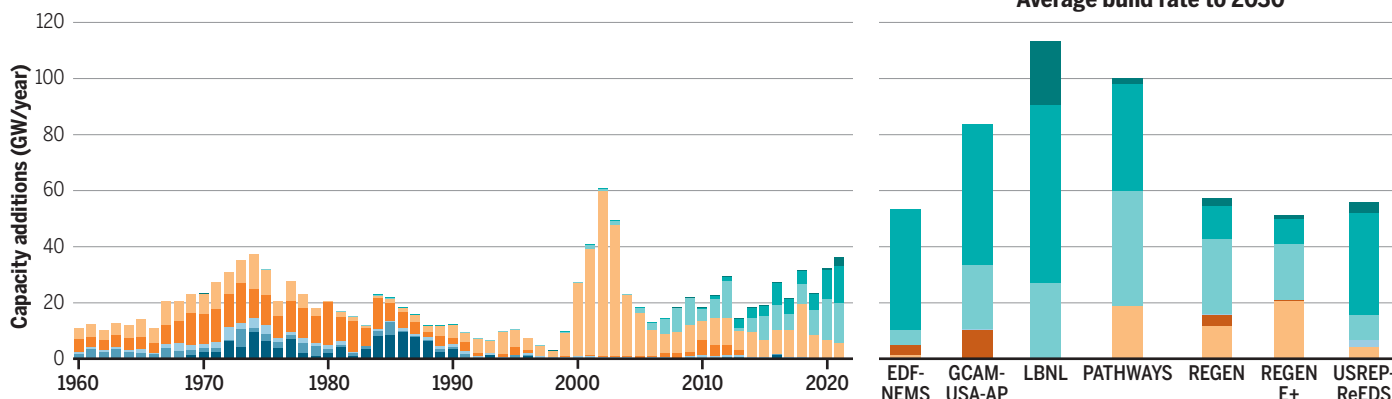
A complementary question to the benefits assessment is to quantify costs associated with reaching the 2030 target and their distribution. Analyses point in broadly comparable directions, but there are differences in magnitudes and viewpoints about which cost metrics are appropriate and how these compare with public health and environmental benefits (12). The two models in this study that calculate levels of expenditures on energy services required to undertake this transformation (including investment costs for supply- and demand-side resources, along with fuels and maintenance) relative to a reference scenario suggest that incremental costs would be $90 billion to $100 billion (PATHWAYS) and $110 billion to $280 billion (REGEN) per year by 2030. Economic impacts and equity outcomes ultimately depend on policy implementation details, which vary across scenarios (table S2). For instance, one model examined how a revenue-generating economy-wide emissions cap can lead to net benefits for lower-income households when permit revenues are distributed to households in equal lump-sum payments (9).

CONCLUSIONS

This study identifies robust findings across models about actions to reach the US 2030 GHG emissions target, using different modeling and policy approaches to provide more detailed projections, actionable information, and confidence about low- or no-regrets strategies. It highlights the considerable pace and scale of change needed—removing barriers to expedite supply- and demand-side buildouts (e.g., CCS technologies, EV charging infrastructure), maintaining reliability with higher renewables shares, and designing policies that encourage affordable and equitable decarbonization. These comparisons

also quantify differences across models in the extent of renewables deployment, degree of electrification, role of CCS, and mitigation outside of the power and transport sectors.

These independent models can provide greater confidence that the next decarbonization steps are clearer and more affordable. The substantial cost reductions for key technologies such as EVs and renewables make their roles larger than previously projected. Although such changes have costs, they also have substantial benefits—including immediate and localized benefits of air quality improvements—and present opportunities for a more equitable distribution of both, which warrant further investigation.

Realizing these benefits requires strong supporting policies. Although this analysis shows that there are many policy pathways to reach 2030 targets, models agree that additional federal policies and incentives are necessary catalysts. The US Senate is currently debating whether to move forward with tax credits, which could provide over $500 billion to accelerate the deployment of renewables, EVs, and other low-emitting technologies. Public policy, private investment, and innovation today can help to reach the 2030 target and scale the technologies of tomorrow to reduce emissions not only from electricity and transport but from across the economy. ∎

REFERENCES AND NOTES

1. J. Rockström et al., Science **355**, 1269 (2017).
2. US Government, "Reducing Greenhouse Gases in the United States: A 2030 Emissions Target" (2021); https://www4.unfccc.int/sites/ndcstaging/PublishedDocuments/United%20States%20of%20America%20First/United%20States%20NDC%20April%2021%202021%20Final.pdf.
3. US Government, "The Long-Term Strategy of the United States: Pathways to Net-Zero Greenhouse Gas Emissions by 2050" (2021); https://www.whitehouse.gov/wp-content/uploads/2021/10/US-Long-Term-Strategy.pdf.
4. N. Abhyankar et al., "Illustrative Strategies for the United States to Achieve 50% Emissions Reduction by 2030" (Lawrence Berkeley National Laboratory, 2021).
5. J. Bistline et al., "Strategies and Actions for Achieving a 50% Reduction in U.S. Greenhouse Gas Emissions by 2030" (3002023165, Electric Power Research Institute, 2021).
6. "Recapturing U.S. Leadership on Climate: Setting an Ambitious and Credible Nationally Determined Contribution" (Environmental Defense Fund, 2021).
7. R. Fakhry, S. Yeh, "The Biden Administration Must Swiftly Commit to Cutting Climate Pollution at Least 50% by 2030" (IB: 21-03-A, National Resources Defense Council, 2021).
8. N. Hultman et al., "An All-In Climate Strategy Can Cut U.S. Emissions by 50% by 2030" (America Is All In, 2021).
9. M. Yuan et al., "Meeting Potential New U.S. Climate Goals" (Report 351, MIT Joint Program on the Science and Policy of Global Change, 2021).
10. I. B. Ocko et al., Science **356**, 492 (2017).
11. The White House, "Executive Order on Strengthening American Leadership in Clean Cars and Trucks" (2021); https://www.whitehouse.gov/briefing-room/presidential-actions/2021/08/05/executive-order-on-strengthening-american-leadership-in-clean-cars-and-trucks/.
12. S. Paltsev, P. Capros, Clim. Change Econ. (Singap.) **4** (suppl. 1), 1340003 (2013).
13. E. Grubert, Science **370**, 1171 (2020).
14. J. Bistline, Joule **5**, 2551 (2021).
15. C. Driscoll et al., "An 80x30 Clean Electricity Standard: Carbon, Costs, and Health Benefits" (Clean Energy Futures, 2021).
16. J. Bistline et al., Zenodo (2022); https://doi.org/10.5281/zenodo.5874740.

GRAPHIC: K. FRANKLIN/SCIENCE

SCIENCE AND SOCIETY

The myth of the lost white tribe

A historian confronts the stories used to sell the American manifest destiny project

By **Laura Stark**

The phrase "origin story" has a whiff of fabulation. For avowed storytellers, untruth is the heart of the joke shared with a winking audience. But *In Whose Ruins*, by Alicia Puglionesi, traces 19th-century origin stories passed off and then passed on as empirical fact—with the effect of placing white people as rightful inheritors of American lands stewarded by Native communities. There are no individual evildoers in the book, which follows people, both white and minoritized, whose actions range from the harebrained to the heroic. The culprit of the book is capitalism itself.

The action unfolds in the industrial age of the American 19th century and extends to the 21st. The "power" of the subtitle refers to both political and material power—energy sources given market value under industrial capitalism, including hydropower, nuclear power, and oil. The book visits four modest places with marvelous landforms, the gravitational centers of the book around which human desires orbit—and there is much to covet. It is rare for a work of political passion to be such a well-told story. The book is inquisitive and sympathetic, artful not dour, worried about political overreach but going for it nonetheless.

One archetypal white American origin story is the myth of the lost white tribe. In the 19th century, a vaguely documented theory took hold among white elites and everyday folks that a tribe of white people inhabited North America prior to the people we rightly regard as Indigenous to the land. At stake was the project of manifest destiny, the idea that white settlers were the (Christian) God-given inheritors of the land, a notion that still has legs today.

Along the Ohio River, an ancient burial mound bulges above farm fields, a site that white settlers called the Grave Creek Mound. Native people recognized this type of earthwork as a sacred internment site for their dead. But for settlers, the enormous size of the complex indicated it was built by people with ingenuity and engineering skills—traits they refused to believe Native people possessed. White onlookers were so invested in their disbelief that they pressed it into fact, attributing the earthwork to a lost white tribe.

The myth of this lost tribe wandered northward into white people's explanations for oil wells dug into the Appalachian foothills, apparent when settlers arrived. Native communities had dug holes to draw petroleum for purposes of ceremony, fuel, and trade. For very different reasons, white speculators pocked the same ground when oil was given commodity value in the late 19th century. They made shallow use of techniques appropriated from Native life—as mere guides to the oil's location—without the corollary commitment to long-term care.

Yet Native scholars and communities retained the empirical and cultural histories of each of the book's four sites. A pile of rocks marked with ancient Native petroglyphs sat in the brambly back corner of a Baltimore park until Native organizers and archaeologists reinstalled them to the site from which they were extracted. In the early 20th century, well-intended white preservationists blasted the rock images from a riverbed to "rescue" them from the encroaching capitalist state. When the Conowingo Dam was built, it flooded the lands and submerged the inscribed rocks, whose brilliance some white scientists and citizens had attributed to—guess who?—America's lost white tribe.

By the early 20th century, the myth of the lost white race began to recede, but the anxieties that nourished it sprouted another theory that justified white dominance over Native communities. The theory of the "vanishing Indian" taught that Native people would soon be extinct, but not for obvious reasons of dispossession and extermination.

Puglionesi travels to the Sonoran Desert, where the theory of the vanishing Indian drove a white romance with Native lifeways and an authenticity industry. This commercial market mushroomed alongside the Manhattan Project and postwar nuclear bomb tests on Native lands and people.

Despite their whiff of falseness, the theories of the lost white tribe and vanishing Indian work a strange magic. They reproduce white dominance, while the landscapes to which the stories refer appear as evidence of dominance's end. For Puglionesi, the end of the world as we know it is not a tragedy but a hope: "So many of our stories dance around the horror of a treasure that is cursed, and the fear of letting it go. The point is to see that not as ruin, but as life itself." ∎

In Whose Ruins
Alicia Puglionesi
Scribner, 2022. 368 pp.

The reviewer is at the Department of History and the Department of Medicine, Health, and Society, Vanderbilt University, Nashville, TN 37235, USA. Email: laura.stark@vanderbilt.edu

SCIENCE & FOOD

Two farmers work a field supported by the farming cooperative Mitoonini in Kirinyaga, Kenya, in 2017.

PODCAST

Food for All in Africa: Sustainable Intensification for African Farmers
Gordon Conway, Ousmane Badiane, Katrin Glatzel
Comstock Publishing Associates, 2019. 342 pp.

Africa's food producers face distinctive challenges, ranging from the geopolitical to the environmental. Meeting the nutritional needs of the continent requires efforts that recognize these challenges and take advantage of the region's assets. This week on the *Science* podcast, Ousmane Badiane discusses strategies for combating hunger and outlines how genetically enhanced crops and homegrown technologies could help increase the continent's food production sustainably. https://bit.ly/3wftLDo

PHOTO: MICHAEL GOTTSCHALK/PHOTOTHEK/GETTY IMAGES

SCIENCE AND SOCIETY

The myth of the lost white tribe

A historian confronts the stories used to sell the American manifest destiny project

By **Laura Stark**

The phrase "origin story" has a whiff of fabulation. For avowed storytellers, untruth is the heart of the joke shared with a winking audience. But *In Whose Ruins*, by Alicia Puglionesi, traces 19th-century origin stories passed off and then passed on as empirical fact—with the effect of placing white people as rightful inheritors of American lands stewarded by Native communities. There are no individual evildoers in the book, which follows people, both white and minoritized, whose actions range from the harebrained to the heroic. The culprit of the book is capitalism itself.

The action unfolds in the industrial age of the American 19th century and extends to the 21st. The "power" of the subtitle refers to both political and material power—energy sources given market value under industrial capitalism, including hydropower, nuclear power, and oil. The book visits four modest places with marvelous landforms, the gravitational centers of the book around which human desires orbit—and there is much to covet. It is rare for a work of political passion to be such a well-told story. The book is inquisitive and sympathetic, artful not dour, worried about political overreach but going for it nonetheless.

One archetypal white American origin story is the myth of the lost white tribe. In the 19th century, a vaguely documented theory took hold among white elites and everyday folks that a tribe of white people inhabited North America prior to the people we rightly regard as Indigenous to the land. At stake was the project of manifest destiny, the idea that white settlers were the (Christian) God-given inheritors of the land, a notion that still has legs today.

Along the Ohio River, an ancient burial mound bulges above farm fields, a site that white settlers called the Grave Creek Mound. Native people recognized this type of earthwork as a sacred internment site for their dead. But for settlers, the enormous size of the complex indicated it was built by people with ingenuity and engineering skills—traits they refused to believe Native people possessed. White onlookers were so invested in their disbelief that they pressed it into fact, attributing the earthwork to a lost white tribe.

The myth of this lost tribe wandered northward into white people's explanations for oil wells dug into the Appalachian foothills, apparent when settlers arrived. Native communities had dug holes to draw petroleum for purposes of ceremony, fuel, and trade. For very different reasons, white speculators pocked the same ground when oil was given commodity value in the late 19th century. They made shallow use of techniques appropriated from Native life—as mere guides to the oil's location—without the corollary commitment to long-term care.

Yet Native scholars and communities retained the empirical and cultural histories of each of the book's four sites. A pile of rocks marked with ancient Native petroglyphs sat in the brambly back corner of a Baltimore park until Native organizers and archaeologists reinstalled them to the site from which they were extracted. In the early 20th century, well-intended white preservationists blasted the rock images from a riverbed to "rescue" them from the encroaching capitalist state. When the Conowingo Dam was built, it flooded the lands and submerged the inscribed rocks, whose brilliance some white scientists and citizens had attributed to—guess who?—America's lost white tribe.

By the early 20th century, the myth of the lost white race began to recede, but the anxieties that nourished it sprouted another theory that justified white dominance over Native communities. The theory of the "vanishing Indian" taught that Native people would soon be extinct, but not for obvious reasons of dispossession and extermination.

Puglionesi travels to the Sonoran Desert, where the theory of the vanishing Indian drove a white romance with Native lifeways and an authenticity industry. This commercial market mushroomed alongside the Manhattan Project and postwar nuclear bomb tests on Native lands and people.

Despite their whiff of falseness, the theories of the lost white tribe and vanishing Indian work a strange magic. They reproduce white dominance, while the landscapes to which the stories refer appear as evidence of dominance's end. For Puglionesi, the end of the world as we know it is not a tragedy but a hope: "So many of our stories dance around the horror of a treasure that is cursed, and the fear of letting it go. The point is to see that not as ruin, but as life itself." ∎

In Whose Ruins
Alicia Puglionesi
Scribner, 2022. 368 pp.

The reviewer is at the Department of History and the Department of Medicine, Health, and Society, Vanderbilt University, Nashville, TN 37235, USA. Email: laura.stark@vanderbilt.edu

SCIENCE & FOOD

Two farmers work a field supported by the farming cooperative Mitoonini in Kirinyaga, Kenya, in 2017.

PODCAST

Food for All in Africa: Sustainable Intensification for African Farmers
Gordon Conway, Ousmane Badiane, Katrin Glatzel
Comstock Publishing Associates, 2019. 342 pp.

Africa's food producers face distinctive challenges, ranging from the geopolitical to the environmental. Meeting the nutritional needs of the continent requires efforts that recognize these challenges and take advantage of the region's assets. This week on the *Science* podcast, Ousmane Badiane discusses strategies for combating hunger and outlines how genetically enhanced crops and homegrown technologies could help increase the continent's food production sustainably. https://bit.ly/3wftLDo

ASTRONOMY

Implied but unseen

A journalist probes dark matter and those who study it

By **Catherine Heymans**

One hundred years ago in May 1922, the Dutch astronomer Jacobus Kapteyn published a study of the motion of stars in the Milky Way that led to a visionary conclusion: The existence of any invisible "dark matter" in the Universe can be detected through its gravitational effect (*1*). A new era of research was born—one that is celebrated by Dutch science journalist Govert Schilling in his new book, *The Elephant in the Universe.*

It takes a village to raise a child, and scientific theories likewise require the efforts of many to reach maturity. Schilling takes readers on a journey through the stories of a veritable army of physicists from around the globe as they have sought to understand the dark matter and dark energy that appear to make up 95% of our Universe.

Many readers will be familiar with the story of pioneering astronomer Vera Rubin. Harnessing the power of Kent Ford's revolutionary image-intensifying technology, Rubin put Kapteyn's theory to the test, measuring the gravity-induced motion of hot gas clouds in distant galaxies.

In Schilling's retelling, Ford happily reminisces about his colleague, who died in 2016. A photo displayed in his living room shows Rubin, clad in a thick winter coat and gloves, painstakingly collecting light on a frozen mountaintop during each 3-hour-long exposure. (Rubin and Ford's work started at a time when telescopes were still hand-driven.)

Throughout the 1970s, the pair studied 21 galaxies, all exhibiting unexpected behavior that could only be explained by the existence of dark matter. This work is widely recognized as the first conclusive evidence for dark matter, and with the construction of the Vera C. Rubin Observatory in Chile, which Schilling visits, we will soon be able

The reviewer is Astronomer Royal for Scotland and is at the Institute for Astronomy, University of Edinburgh, Edinburgh, UK, and the German Centre for Cosmological Lensing, Ruhr-University Bochum, Bochum, Germany. Email: heymans@roe.ac.uk

to confront theories that seek to explain the Universe's dark side.

Schilling delves behind this well-known narrative, introducing readers to the lesser-known "Westerbork cowboys," the Dutch radio astronomers who, in fact, beat Rubin and Ford to their conclusive measurement by several years in work that Rubin herself cited. Has science forgotten about their work, published in Albert Bosma's doctoral thesis, because radio astronomy was not fully established at the time? Or was it because of the cowboys' "tacky" behavior, which prompted the director of the competing Mullard Radio Astronomy Observatory to write a formal letter of complaint to their boss? We will never know for sure, but Schilling is keen to correct the history books on this matter.

Schilling does not shy away from complex experiments and theories that have led to an as-yet-undetected cold dark matter particle becoming the corner-stone of modern physics. He dives headfirst into such topics as CP violation and Bose-Einstein condensates as well as other graduate-level theoretical particle physics that underpins the alternative axion theory for dark matter.

There is a joy to finding these challenging yet clearly explained concepts interspersed between charming stories of the scientists responsible for discovering them. Schilling relates, for example, how Samuel Ting once received a speeding ticket on the way to the launch of his own particle detector on the last space shuttle *Endeavour* mission and how Martin Rees and Joel Primack, killing time together while colleagues skied at a con-

The Elephant in the Universe: Our Hundred-Year Search for Dark Matter
Govert Schilling
Belknap Press, 2022.
376 pp.

ference, discovered that a cold dark matter particle could explain the observed pattern of large-scale structures in the Universe. He writes of the chance Scandinavian ferry meeting between Marc Davis and George Efstathiou that led to the birth of computational simulations of the whole Universe and even joins the story himself, describing how he has followed his subjects over the lifetimes of both his and their careers.

As the decades pass, the nature of research shifts from a moment when academics had the freedom and time to pursue interdisciplinary research to today's era of fast-paced competition. Mistakes are made, and individuals who should know better intimidate and bully. Robert Kirshner, a key player in the discovery of dark energy, laments, "What's the strongest force in the Universe? It's not gravity, it's jealousy." And while the decade-long search for the weakly interacting massive particle (WIMP) dark matter candidate continues to come up empty-handed, fear is in the air as we move from the era of the "WIMP miracle" into a fast-approaching "dark crisis."

Will dark matter and dark energy become the ether of the 21st century? Schilling introduces readers to an array of alternative theories, including those that question the likes of Einstein and Newton. One thing is for certain, as Schilling's book makes clear: There are thousands of researchers eager to find out. ∎

REFERENCES AND NOTES
1. J. C. Kapteyn, *Astrophys. J.* **55**, 302 (1922).

10.1126/science.abq1661

The Alpha Magnetic Spectrometer on the International Space Station measures antimatter.

PHOTO: NASA

Brazil is using the disruption of Russian fertilizer imports to justify allowing mining in Indigenous lands.

Edited by **Jennifer Sills**

War serves as excuse for Amazon destruction

The Russian invasion of Ukraine is providing a new excuse for Amazon destruction. On 2 March (7 days after the invasion began), Brazilian President Jair Bolsonaro called on the National Congress to give urgency to approving bill PL 191/2020 to supply Brazil's agribusiness with the fertilizers the country normally imports from Russia (*1*). On 9 March, Bolsonaro's supporters in Congress passed a motion granting "urgent" status to the bill, allowing it to bypass the normal committee process. The motion was approved but has not been finalized out of fear of international boycotts of Brazilian minerals.

PL 191/2020 would open Indigenous lands to mining, hydroelectric dams, and logging, cattle ranching, and industrial agriculture by non-Indigenous entrepreneurs and companies (*2*). It was submitted to the National Congress in February 2020 by President Bolsonaro and gained priority in February 2021 when both houses of Congress were captured by the coalition of parties that supports Bolsonaro's anti-environmental agenda (*3*). However, it stalled until the invasion of Ukraine.

Indigenous lands are essential to maintaining the environmental services of Brazil's Amazon rainforest, given that these lands protect more forest than do federal "conservation units" (*4*). The supposed justification for approving the bill is the alleged need to extract potassium, a key component of fertilizers, from the Autazes mine in Brazil's state of Amazonas. However, Autazes is not located on officially recognized Indigenous land, making the passage of the bill irrelevant to this goal (*5*). Furthermore, Russia is not the world's only source of potassium: Other countries, led by Canada, are major international suppliers (*6*). Even within Brazil, mining Amazonia is not required to obtain potassium. The largest potassium deposits are in the state of Minas Gerais, and Brazil's deposits outside of Amazonia could supply the country with potassium until at least 2100 (*7*).

Passing PL 191/2020 would not achieve the goals articulated by its supporters. Even if it affected the status of the planned mine in Autazes, the project would not be functional in time to solve Brazil's predicted fertilizer shortage for the coming year. Imports will have to supplement the supply. Major mining companies in Brazil are now opposing PL191/2020 (*8*). International actors, including the Canadian companies invested in Autazes (*9*), should not agree to initiate this project given the false pretenses under which the bill is being advanced. Moreover, if the bill goes into effect without changes, mineral importers should follow through with potential boycotts to show that the irresponsible actions by Brazil have consequences.

Paulo Arthur A. Trindade[1,2]*, **Juliana S. Araújo**[2,3], **Valter M. Azevedo-Santos**[1,2,4], **Friedrich W. Keppeler**[1,2,5], **Tommaso Giarrizzo**[1,2,6], **Philip M. Fearnside**[7]

[1]Federal University of Pará, Belém, PA, Brazil. [2]Aquatic Ecology Group, Belém, PA, Brazil. [3]Paulista State University, Sorocaba, SP, Brazil. [4]Federal University of Tocantins, Porto Nacional, TO, Brazil. [5]Center for Limnology, University of Wisconsin–Madison, Madison, WI 53706, USA. [6]Marine Sciences Laboratory, Federal University of Ceara, Fortaleza, CE, Brazil. [7]National Institute for Research in the Amazon, Manaus, AM, Brazil.
*Corresponding author.
Email: pauloaatrindade@gmail.com

REFERENCES AND NOTES

1. R. D. Coletta, "Bolsonaro usa possível falta de fertilizantes da Rússia para defender mineração em terras indígenas," *Folha de São Paulo* (2022); https://www1.folha.uol.com.br/mercado/2022/03/bolsonaro-usa-possivel-falta-de-fertilizantes-para-defender-mineracao-em-terras-indigenas.shtml?origin=folha [in Portuguese].
2. Câmara dos Deputados, 2020 PL 191/2020 (2020); www.camara.leg.br/proposicoesWeb/fichadetramitacao?idProposicao=2236765 [in Portuguese].
3. L. Ferrante, P. M. Fearnside, *Science* **371**, 898 (2021).
4. E. M. Nogueira, A. M. Yanai, S. S. Vasconcelos, P. M. L. A. Graça, P. M. Fearnside, *Reg. Environ. Change* **18**, 573 (2018).
5. M. Santilli, "Bolsonaro mente sobre potássio em terras indígenas," Instituto Socioambiental (2022); https://neomondo.org.br/2022/03/09/bolsonaro-mente-sobre-potassio-em-terras-indigenas/ [in Portuguese].
6. Government of Canada, "Potash facts" (2022); www.nrcan.gc.ca/our-natural-resources/minerals-mining/minerals-metals-facts/potash-facts/20521.

7. T. Pimentel, "Estudo da UFMG contradiz Bolsonaro sobre exploração de potássio em terras indígenas para fertilizantes," *g1 Minas* (2022); https://g1.globo.com/mg/minas-gerais/noticia/2022/03/07/estudo-da-ufmg-contradiz-bolsonaro-sobre-exploracao-de-potassio-em-terras-indigenas-para-fertilizantes.ghtml [in Portuguese].
8. M. Pajolla, "Após repercussão negativa, mineradoras abandonam defesa de PL que atinge terras indígenas," *Brasil de Fato* (2022); www.brasildefato.com.br/2022/03/15/apos-repercussao-negativa-mineradoras-abandonam-defesa-de-pl-que-atinge-terras-indigenas [in Portuguese].
9. A. Boadle, "Bolsonaro pushes for mining on Brazil's indigenous lands, citing fertilizer pinch," *NASDAQ* (2022); www.nasdaq.com/articles/bolsonaro-pushes-for-mining-on-brazils-indigenous-lands-citing-fertilizer-pinch.

10.1126/science.abq3611

Drug building blocks and libraries at risk in Ukraine

Ukrainian chemical vendors such as Life Chemicals, Otava, and Enamine supply more than 50% of the 400,000 chemical building blocks used by scientists internationally to synthesize new drugs (*1*). These building blocks also serve as the substrates for libraries of more than 40 billion readily accessible or "tangible" molecules, which can be constructed by stitching the building blocks into larger molecules and are synthesized as needed. Drawing from these molecules has revealed novel leads for multiple drug targets (*2–6*). More than 90% of this tangible chemical space also comes from Ukraine. The Russian invasion of Ukraine endangers both the unique building block library and the 40 billion–molecule tangible library that derives from it.

Ordinarily, more than 2000 new building blocks are synthesized per month by Enamine alone (*7*)—the collection is the fruit of decades of synthesis by thousands of chemists. For the first month of the war, production stopped entirely. Work has tentatively restarted, but the reprieve is fragile.

These building blocks and tangible libraries have been constructed over the past 30 years, as has the expertise to assemble them into the kaleidoscope of compounds that, until 24 February, were openly available to the world's drug hunters and chemical biologists. International aid should include protection for Ukrainian scientists, their families, and their labs and reagents. Improved safety for supply lines into and out of Ukraine will preserve access to and the integrity of the building-block and tangible chemical libraries. Once the war has ended, it will be vital to help the Ukrainian scientific community rebuild, allowing them to continue their work to the benefit of all.

Ivan S. Kondratov[1,2], Yurii S. Moroz[3,4], John J. Irwin[5], Brian K. Shoichet[5]*
[1]Enamine Ltd., Kyïv 02094, Ukraine. [2]V.P. Kukhar

Institute of Bioorganic Chemistry & Petrochemistry, National Academy of Sciences of Ukraine, Kyïv 02660, Ukraine. [3]Chemspace, Kyïv 02094, Ukraine. [4]Taras Shevchenko National University of Kyïv, Kyïv 01601, Ukraine. [5]Department of Pharmaceutical Chemistry and Quantitative Biology Institute, University of California San Francisco, San Francisco, CA 94143, USA.
*Corresponding author.
Email: shoichet@cgl.ucsf.edu

REFERENCES AND NOTES
1. Y. Zabolotna *et al.*, *J. Chem. Inf. Model* **62**, 2171 (2021).
2. J. Lyu *et al.*, *Nature* **566**, 224 (2019).
3. R. M. Stein *et al.*, *eLife* **9**, e53779 (2020).
4. C. Gorgulla *et al.*, *Nature* **580**, 663 (2020).
5. A. A. Sadybekov *et al.*, *Nature* **601**, 452 (2022).
6. A. Alon *et al.*, *Nature* **600**, 759 (2021).
7. Enamine, Building Blocks Catalog; https://enamine.net/building-blocks/building-blocks-catalog.

10.1126/science.abq7841

Protecting global marine animal forests

Despite years of awareness raised through international initiatives such as the UN Climate Change Conference in Glasgow (COP26) (*1*), the Decade of Ocean Science for Sustainable Development (2021–2030) (*2*), and the Ecosystem Restoration decade (*3*), human activities continue to deeply transform marine ecosystems (*4*). Sustainable Development Goal 14—Life Below Water—is the least-funded and most-underrepresented objective in the European Union (*2*), with decades of budget allocation delay for ocean research and conservation. Benthic ecosystems suffer from the effects of bottom trawling, urban and agricultural pollution, bioinvasions, climate change, and other anthropogenic pressures (*5*). Among these ecosystems, marine animal forests, which are dominated by benthic suspension feeders such as sponges, hard corals, and gorgonians, form three-dimensional habitats (*6*) that are particularly vulnerable to disturbances (*7*).

Marine animal forests include habitats ranging from coastal to deep sea, representing one of the largest biomes on Earth (*8*). The forests are ecologically relevant as biodiversity hotspots and nursery grounds, and evidence suggests that they have the potential to provide ecosystem services (*9*), especially by ameliorating the effects of climate change by immobilizing carbon (*10*). However, information on distribution, population dynamics, connectivity, and ecosystem functioning of key marine animal forest species is still lacking. Although technology is available to gather the needed data (*11*), these targets are not currently a political priority.

Anthropogenic disturbances, past and present (*12*), jeopardize the ecological processes of marine animal forests and threaten the services they provide to human societies worldwide. These communities urgently need conservation, monitoring, and restoration far beyond the efforts made up to now. Protecting marine animal forests requires scientific, social, and political investment in increasing our knowledge. With a more extensive understanding, we will be able to properly manage these threatened habitats. Given that their eradication will have substantial negative consequences for the maintenance of planetary health, marine animal forests should be prioritized in conservation plans.

Sergio Rossi*, Lorenzo Bramanti, Paulo Horta, Louise Allcock, Marina Carreiro-Silva, Martina Coppari, Vianney Denis, Louis Hadjioannou, Enrique Isla, Carlos Jimenez, Mark Johnson, Christian Mohn, Covadonga Orejas, Andreja Ramšak, James Reimer, Baruch Rinkevich, Lucia Rizzo, Maria Salomidi, Toufiek Samaai, Nadine Schubert, Marcelo Soares, Ruth H. Thurstan, Paolo Vassallo, Patrizia Ziveri, Juanita Zorrilla-Pujana
*Corresponding author.
Email: sergio.rossi@unisalento.it
For full author affiliations, go to www.science.org/doi/10.1126/science.abq7583.

REFERENCES AND NOTES
1. UN Climate Change Conference UK 2021 (2021); https://ukcop26.org/.
2. Decade of Ocean Science for Sustainable Development (2021–2030); www.oceandecade.org.
3. UN Decade on Ecosystem Restoration 2021–2030; www.decadeonrestoration.org.
4. N. L. Bindoff *et al.*, in *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate*, H.-O. Pörtner *et al.*, Eds. (Cambridge University Press, 2019), pp. 447–587.
5. C. M. Duarte *et al.*, *Nature* **580**, 39 (2020).
6. S. Rossi *et al.*, in *Marine Animal Forests: The Ecology of Benthic Biodiversity Hotspots*, S. Rossi *et al.*, Eds. (Springer, 2017), pp. 1–28.
7. M. R. Clark *et al.*, *Front. Mar. Sci.* 10.3389/fmars.2019.00063 (2019).
8. J. S. Stark *et al.*, in *The IUCN Global Ecosystem Typology 2.0: Descriptive Profiles for Biomes and Ecosystem Functional Groups*, D. A. Keith *et al.*, Eds. (IUCN, Gland, Switzerland, 2020).
9. C. Paoli *et al.*, in *Marine Animal Forests: The Ecology of Benthic Biodiversity Hotspots*, S. Rossi *et al.*, Eds. (Springer, 2017), pp. 1271–1312.
10. S. Rossi, L. Rizzo, in *Perspectives on the Marine Animal Forests of the World*, S. Rossi, L. Bramanti, Eds. (Springer-Nature, 2020), pp. 333–400.
11. P. Rossi *et al.*, *Front. Mar. Sci.* 10.3389/fmars.2021.591292 (2021).
12. R. H. Thurstan, J. M. Pandolfi, P. S. Zu Ermgassen, in *Marine Animal Forests: The Ecology of Benthic Biodiversity Hotspots*, S. Rossi *et al.*, Eds. (Springer, 2017), pp. 947–964.

10.1126/science.abq7583

RESEARCH

GLACIAL CYCLES
Solar controller

Before about 1.25 million years ago, glacial cycles reflected the 40,000-year obliquity insolation cycle, whereas over the past 800 thousand years, glacial cycles have been paced by the 100,000-year eccentricity cycle. What about the role of the 23,000-year cycle of precession? Barker *et al*. present a 1.7-million-year record showing that glacial termination has depended mostly on precession for the past million years. That change seems to be a function of ice sheet size. —HJS   *Science*, abm4033, this issue p. 961

A record of North Atlantic ice rafting, as shown in the photograph, reveals the effect of precession on ice sheet variability in the recent geologic past.

DEVELOPMENT
Demethylation controls development

FTO was the first discovered RNA demethylase that reverses messenger RNA N^6-methyladenosine (m^6A) modification. Despite numerous past studies, the physiological substrates of FTO in mammalian development remain unclear. Wei *et al*. uncovered RNA transcribed from long-interspersed element-1 (LINE1), one of the most abundant mammalian retrotransposons, as a major substrate of FTO in mouse embryonic stem cells and mouse tissues. m^6A demethylation by FTO regulates the LINE1 RNA level, which shapes local and global chromatin state. Deletion of *Fto* in cells deactivates LINE1-containing genes by repressing intragenic LINE1 RNA. This FTO-LINE1 RNA axis also affects mouse oocyte and embryonic development. —BAP

Science, abe9582, this issue p. 968

ATMOSPHERE
Powerful oxidants in the atmosphere

Hydrotrioxides (ROOOHs) have intrigued the atmospheric chemistry community because of their strong oxidizing properties and theoretical predictions that they could form in atmospherically relevant RO_2 + OH reactions. Much of the work to date has focused on CH_3O_2, but this chemistry has been found to play a minor role. Using a mass spectrometry–based scheme for direct detection and ab initio calculations supplemented by global modeling, Berndt *et al*. showed that ROOOHs could form routinely for heavier RO_2 and have appreciable lifetimes. Potentially detectable steady-state concentrations in the atmosphere were established. This work draws attention to an important class of strong oxidizing agents previously disregarded in atmospheric kinetics models. —YS

Science, abn6012, this issue p. 979

CATALYSIS
Nanoparticle response to redox conditions

In heterogeneous catalysts, the interaction between transition metal nanoparticles and their oxide supports can be mainly that of absorbed species moving between metal and oxide surfaces. However, substantial restructuring can occur. For platinum nanoparticles on titania, reaction conditions can cause the oxide support to encapsulate and deactivate the metal. Frey *et al*. used transmission electron microscopy to image this system interacting with 1 bar of a hydrogen and oxygen mixture forming water. The effects of this redox-active chemical environment included destabilization of the encapsulation layer, as well as platinum particles forming twin planes and directionally migrating across the titania surface. —PDS

Science, abm3371, this issue p. 982

IMMUNOLOGY
CD8α keeps T cells humming along

In addition to acting as a T cell lineage marker, the CD8 heterodimer acts as a co-receptor to enhance T cell receptor (TCR) signaling after antigen stimulation. Whether CD8 plays any other antigen-independent functions is unclear. Zheng *et al*. found that inducible deletion of CD8α in mice resulted in a loss of both naïve and memory CD8 T cell quiescence, which is essential for their

survival and repertoire diversity. Quiescence is actively maintained by T cell–myeloid cell cross-talk through cell surface CD8α–PILRα interactions. Disruption of this interaction leads to spontaneous activation and subsequent loss of naïve and memory CD8+ T cell pools in peripheral lymphoid organs. —STS

Science, aaz8658, this issue p. 996

PHYSIOLOGY
Joints degenerate with a lncRNA

Nuclear factor κB (NF-κB) signaling contributes to the development and progression of osteoarthritis by promoting cartilage degradation and joint inflammation. Tang *et al.* found that the long noncoding RNA (lncRNA) PILA exacerbated inflammation-induced joint damage in osteoarthritis. In human articular chondrocytes, PILA stimulated the modification of the RNA helicase DHX9 by the protein arginine methyltransferase PRMT1, which resulted in increased levels of an upstream activator of NF-κB. Expression of PILA in mouse knee joints elicited spontaneous cartilage degradation and enhanced experimentally induced osteoarthritis. —AMV

Sci. Signal. **15**, eabm6265 (2022).

PLANETARY SCIENCE
Tracking the greenhouse effect on Mars

Mars was far warmer and wetter in the past than it is today, but the when and why of this transition is actively debated. One popular model holds that this transition was driven by a loss of atmospheric carbon dioxide (CO_2), lowering the radiation-forcing effect of that greenhouse gas. Previous studies indicated that other greenhouse gases may also have played a role, but their importance remains unclear. Kite *et al.* tracked the spatial distribution of rivers through time on Mars using remote sensing data as a proxy for how the greenhouse effect may have changed. The authors were able to constrain climate models that provide insights into the relative importance of CO_2 forcing and forcing by other gases during the wet-dry transition. Unexpectedly, their models suggest that a reduction in non-CO_2 gases may have played the primary role in cooling and drying the planet. —KVH

Sci. Adv. 10.1126/sciadv.abo5894 (2022).

MATERIALS SCIENCE
Untwisted heterobilayers

Stacking two-dimensional materials can lead to a range of interesting and useful properties. For example, the emergence of piezoelectricity and ferroelectricity can occur by twisting the layers relative to each other. Rogée *et al.* used a different strategy for obtaining these sorts of properties by stacking alternating layers of molybdenum disulfide and tungsten disulfide, which breaks symmetry without requiring twisting. The authors show that these vapor-deposited heterobilayers can be made into a ferroelectric tunnel junction device. —BG

Science, abm5734, this issue p. 973

Water shaped the surface geography of Mars, seen here in a high-resolution satellite image, in different ways as the planet cooled.

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

MORPHOGENESIS
From ordered cells to patterned tissues

In developing animal skin, cells self-organize to generate different structures such as sweat glands, hair, and feather follicles. But what comes first, patterning or gene expression? Palmquist *et al.* show that follicle formation in bird skin depends on mechanical self-organization. The authors made a reconstituted system of embryonic avian dermal cells on a substrate that mimics the extracellular matrix. In this model, dermal cells initially formed a ring but separated into regularly spaced aggregates because of feedback between cell contractility and rearrangements of the extracellular membrane that further align the cells. —VV and SMH

Cell 10.1016/j.cell.2022.04.023 (2022).

Mechanical forces in embryonic bird skin determine the formation of feather follicles.

TRANSCRIPTION
Lampbrush loops in the nucleus

RNA polymerase II in the nucleus transcribes the instruction in a gene into a messenger RNA. In any given moment, it has been proposed that thousands of genes are transcribed in transcription factories. In these factories, RNA polymerases II remain stationary, and active genes approach the enzyme to be transcribed. Contrary to this hypothesis, Leidescher *et al.* has found that very long but highly expressed genes form transcription loops decorated by RNA polymerase II. The enzyme moves along the gene axis, carrying newly synthesized messenger RNAs undergoing co-transcriptional splicing. These transcription loops look like lateral loops of lampbrush chromosomes observed more than 100 years ago in the oocytes of many non-mammalian animals. It is not yet clear how general this mode of transcription is or how it contributes to nuclear organization. —DJ

Nat. Cell Biol. **24**, 327 (2022).

CONSERVATION
A new hope

The vast negative impact of humanity on other species tends to lead to extinction. However, our manipulations have also led to a wide variety of less-recognized species alterations. Chinese giant salamanders (*Andrias* spp.) are the victims of a combination of incomplete knowledge, commercial enterprise, and poorly informed conservation actions. This genus is now mostly represented by thousands of individuals in farms (they are consumed for food and used in traditional Chinese medicine). Most of these animals are hybrids of as many as eight previously unrecognized species, and many of these hybrids

Edited by **Michael Funk**

CANCER GENOMICS
Grouping prostrate cancers

Therapeutic interventions are needed for prostate tumors that exhibit a loss of androgen receptor dependence in castration-resistant prostate cancer (CRPC). However, there is a scarcity of prostate cancer cell line models. Tang *et al.* performed a molecular study on a number of CRPC organoids, cell lines, and patient-derived xenographs by combining sequencing techniques and transcriptomics to identify four distinct genetic subgroups of tumors. Dominant transcription factors were identified for each subgroup, with a deeper analysis of subgroup 4 leading to the identification of a chromatin-opening positive-feedback loop. From these data, the authors propose that inhibitors of the transcriptional coactivators YAP and TAZ may be used to treat individuals with subgroup 4–like tumors, the second most common type of CRPC observed in this study. —LMZ

Science, abe1505, this issue p. 960

BIOMEDICINE
Better cardiac care

There is an unmet medical need for minimally invasive devices to aid patients with slow heart rates, particularly after cardiovascular surgery. This is because the lines required to connect internal electrodes with battery packs are a common cause of infection and extended hospitalization. Choi *et al.* developed a biodegradable, closed-loop, wireless microelectromechanical system (Bio-MEMS) for heart rate monitoring and control (see the Perspective by Zimmermann). Because the devices allow for both monitoring and control of the heart rate, they can be reprogrammed on demand in response to physiological conditions using the onboard electronics. Furthermore, the batteries in

the devices can be recharged wirelessly through the skin, fully eliminating the need for transcutaneous wires. —MSL

Science, abm1703, this issue p. 1006; see also abq0605, p. 917

ADAPTATION
Rapid change

Human impacts are leading to exceedingly rapid alteration of our world, from land conversion and habitat loss to climate change. Some have proposed that rapid adaptation could help some species persist in the face of these changes, but questions remain about whether adaptation could occur rapidly enough to make a difference. Bonnet *et al.* looked at additive genetic variance, which determines the contribution of selection to genetic change that increases fitness, in long-term data from 19 species and found it to be higher than expected—often substantially higher (see the Perspective by Walsh). These results suggest that many species may have some capacity to adapt to our changing world. —SNV

Science, abk0853, this issue p. 1012; see also abo4624, p. 920

ENVIRONMENTAL TOXINS
Nutrient control must include nitrogen

Lake Erie receives water from important agricultural areas of Canada and the United States and is subject to high levels of nitrogen and phosphorus in runoff. These nutrients can lead to rapid growth of photosynthetic organisms, some of which produce toxins that harm aquatic animals and compromise drinking water. Recent efforts have focused on reducing phosphorus loading. With support from a large literature meta-analysis, Hellweger *et al.* developed an agent-based model of cyanobacterial metabolism to determine how toxin production changed under a range of nutrient and

environmental conditions and defined the associated molecular mechanisms (see the Perspective by Ofiţeru and Picioreanu). The authors found that phosphorus reduction alone was potentially harmful, lowering total biomass but increasing toxin production. The proposed mechanism involves response to hydrogen peroxide stress and increased light transmission. —MAF

Science, abm6791, this issue p. 1001; see also abq0956, p. 914

ANTIBIOTIC DISCOVERY
Panning for antibiotics that stand out

Despite their importance in modern medicine, antibiotics are under constant threat as pathogens adapt, acquire, or evolve resistance with chilling regularity. Finding new molecules and mechanisms is one way that we can keep ahead. Panning genomes enriched in secondary metabolites, Wang *et al.* used a computational approach to predict the structure of a cyclic nonribosomal lipopeptide antibiotic, which they named "cilagicin" (see the Perspective by Seipke). Chemical synthesis of the predicted peptide revealed potent and broad antimicrobial activity against Gram-positive bacteria, including a number of drug-resistant pathogens, and a cilagicin derivative protected mice in an acute infection model. Mechanistic experiments suggested that the peptide binds to two closely related molecules involved in lipid biosynthesis, an ability that may help to prevent the rapid development of resistance. —MAF

Science, abn4213, this issue p. 991; see also abq3206, p. 919

METABOLIC DISEASE
Mitigating methylmalonylation

Methylmalonic acidemia (MMA) is a genetic metabolic disorder

that results in a buildup of methylmalonic acid in the body and can lead to severe symptoms. Examining liver tissue from patients and a mouse model of MMA, Head *et al.* found that the buildup of this metabolite resulted in a widespread posttranslational modification, methylmalonylation, on a variety of proteins. This included the sirtuin (SIRT5) that reverses methylmalonylation, dampening its ability to undo the protein modification damage. A SIRT5 gene engineered to be resistant to methylmalonylation reduced symptoms in the MMA mice, demonstrating a proof-of-concept gene therapy strategy. —CAC

Sci. Transl. Med. **14**, eabn4772 (2022).

GEOMORPHOLOGY
Abruptly changing course

River avulsions are places where a river abandons its channel, and are a common feature of geomorphological structures such as deltas. Brooke *et al.* used 50 years of satellite images to look at the location and change in river avulsions globally (see the Perspective by Passalacqua and Moodie) and found that avulsions are often tied to changes in channel slope or sedimentation just upstream of the river. However, in some cases, the river avulsion is farther upstream than expected, likely due to erosional processes. Understanding what controls avulsion location in the context of climate and land use changes is vital because avulsions are strongly tied to risks from flooding. —BG

Science, abm1215, this issue p. 987; see also abq1166, p. 916

IMMUNE REGULATION
A type 2 skew ages neutrophils

Immune responses in patients with atopy are skewed toward type 2, with overproduction of cytokines that drive allergic

responses to environmental antigens. Atopic individuals have increased susceptibility to bacterial and viral infections; however, the contribution of immune cells such as neutrophils to this susceptibility is not well defined. Egholm *et al.* studied neutrophil-mediated immunity to *Listeria monocytogenes* or *Staphylococcus aureus* infection in mice with genetic or induced predisposition to type 2 immune responses. Neutrophil-mediated responses were impaired in mice with type 2 predisposition because of neutrophil intrinsic interleukin-4 receptor (IL-4R) signaling. IL-4R signaling accelerated neutrophil aging, resulting in reduced effector functions, increased apoptosis, and an impaired ability to clear bacteria. These findings provide new insights into the interplay between type 2 immune predisposition and neutrophil function. —HMI

Sci. Immunol. **7**, eabi9733 (2022).

APPLIED PHYSICS
Ultra-broadband terahertz detection

Because of the low energies involved, terahertz detection has presented a difficult challenge. Ratchets, which use local asymmetries to rectify nondirectional signals, could be a promising solution but have mostly been a theoretical concept to date. Bai *et al.* successfully fabricated and demonstrated a gallium-arsenide–based ratchet structure that was able to efficiently detect radiation in the near-infrared to terahertz range. The device has an elusive combination of broad response (4 to 300 terahertz) and peak responsivity five orders of magnitude higher than that of existing broadband photon-type detectors. These devices will be promising for a range of applications, including biomedical, astrophysics, security, and communication systems. —SDS

Sci. Adv. 10.1126/sciadv.abn2031 (2022).

survival and repertoire diversity. Quiescence is actively maintained by T cell–myeloid cell cross-talk through cell surface CD8α–PILRα interactions. Disruption of this interaction leads to spontaneous activation and subsequent loss of naïve and memory $CD8^+$ T cell pools in peripheral lymphoid organs. —STS

Science, aaz8658, this issue p. 996

PHYSIOLOGY
Joints degenerate with a lncRNA

Nuclear factor κB (NF-κB) signaling contributes to the development and progression of osteoarthritis by promoting cartilage degradation and joint inflammation. Tang *et al.* found that the long noncoding RNA (lncRNA) PILA exacerbated inflammation-induced joint damage in osteoarthritis. In human articular chondrocytes, PILA stimulated the modification of the RNA helicase DHX9 by the protein arginine methyltransferase PRMT1, which resulted in increased levels of an upstream activator of NF-κB. Expression of PILA in mouse knee joints elicited spontaneous cartilage degradation and enhanced experimentally induced osteoarthritis. —AMV

Sci. Signal. **15**, eabm6265 (2022).

PLANETARY SCIENCE
Tracking the greenhouse effect on Mars

Mars was far warmer and wetter in the past than it is today, but the when and why of this transition is actively debated. One popular model holds that this transition was driven by a loss of atmospheric carbon dioxide (CO_2), lowering the radiation-forcing effect of that greenhouse gas. Previous studies indicated that other greenhouse gases may also have played a role, but their importance remains unclear. Kite *et al.* tracked the spatial distribution of rivers through time on Mars using remote sensing data as a proxy for how the greenhouse effect may have changed. The authors were able to constrain climate models that provide insights into the relative importance of CO_2 forcing and forcing by other gases during the wet-dry transition. Unexpectedly, their models suggest that a reduction in non-CO_2 gases may have played the primary role in cooling and drying the planet. —KVH

Sci. Adv. 10.1126/sciadv.abo5894 (2022).

MATERIALS SCIENCE
Untwisted heterobilayers

Stacking two-dimensional materials can lead to a range of interesting and useful properties. For example, the emergence of piezoelectricity and ferroelectricity can occur by twisting the layers relative to each other. Rogée *et al.* used a different strategy for obtaining these sorts of properties by stacking alternating layers of molybdenum disulfide and tungsten disulfide, which breaks symmetry without requiring twisting. The authors show that these vapor-deposited heterobilayers can be made into a ferroelectric tunnel junction device. —BG

Science, abm5734, this issue p. 973

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

MORPHOGENESIS
From ordered cells to patterned tissues

In developing animal skin, cells self-organize to generate different structures such as sweat glands, hair, and feather follicles. But what comes first, patterning or gene expression? Palmquist *et al.* show that follicle formation in bird skin depends on mechanical self-organization. The authors made a reconstituted system of embryonic avian dermal cells on a substrate that mimics the extracellular matrix. In this model, dermal cells initially formed a ring but separated into regularly spaced aggregates because of feedback between cell contractility and rearrangements of the extracellular membrane that further align the cells. —VV and SMH

Cell 10.1016/j.cell.2022.04.023 (2022).

Mechanical forces in embryonic bird skin determine the formation of feather follicles.

TRANSCRIPTION
Lampbrush loops in the nucleus

RNA polymerase II in the nucleus transcribes the instruction in a gene into a messenger RNA. In any given moment, it has been proposed that thousands of genes are transcribed in transcription factories. In these factories, RNA polymerases II remain stationary, and active genes approach the enzyme to be transcribed. Contrary to this hypothesis, Leidescher *et al.* has found that very long but highly expressed genes form transcription loops decorated by RNA polymerase II. The enzyme moves along the gene axis, carrying newly synthesized messenger RNAs undergoing co-transcriptional splicing. These transcription loops look like lateral loops of lampbrush chromosomes observed more than 100 years ago in the oocytes of many non-mammalian animals. It is not yet clear how general this mode of transcription is or how it contributes to nuclear organization. —DJ

Nat. Cell Biol. **24**, 327 (2022).

CONSERVATION
A new hope

The vast negative impact of humanity on other species tends to lead to extinction. However, our manipulations have also led to a wide variety of less-recognized species alterations. Chinese giant salamanders (*Andrias* spp.) are the victims of a combination of incomplete knowledge, commercial enterprise, and poorly informed conservation actions. This genus is now mostly represented by thousands of individuals in farms (they are consumed for food and used in traditional Chinese medicine). Most of these animals are hybrids of as many as eight previously unrecognized species, and many of these hybrids

Water shaped the surface geography of Mars, seen here in a high-resolution satellite image, in different ways as the planet cooled.

have escaped. It was believed that no breeding populations of the ancestors of these species existed in the wild. However, Chai *et al.* report the discovery of a new species in a remote protected area that is apparently not a hybrid, appears to be free from poaching, and is reproducing at healthy levels, realizing the value of exclusive protected areas and community action. —SNV

Zoo. Res. 10.24272/j.issn.2095-8137.2022.101 (2022).

BATTERIES
Foldable solid-state battery electrolyte

Covalent organic framework (COF) materials have large pore channels and high stability and thus can facilitate rapid ion transport. However, lithium ion conduction is often hampered by the size of the channels, because the solid-state diffusion occurs mainly along the pore walls. Guo *et al.* dissolved dimethylacrylamide into the electrolyte lithium bis(trifluoromethanesulfonyl) imide that is then deposited into a Li COF and subsequently polymerized. The net effect was to decouple the transport of Li$^+$ from both the COF walls and polymer chains, thus improving Li$^+$ conductivity by a factor of 100. The authors show that the electrolyte can be used to make solid-state pouch batteries that endure repeated bending or folding cycles. —MSL

Adv. Mater. 10.1002/adma.202201410 (2022).

IMMUNOLOGY
Mirror to T cell selection

During positive selection, immature CD4$^+$CD8$^+$ T cells in the thymus differentiate into either CD4$^+$ helper T cells or CD8$^+$ cytotoxic T cells after receiving signals through their T cell receptors (TCRs). Whether TCR and co-receptor signaling strength or duration decides the developing T cell's fate has been a matter of much debate. Shinzawa *et al.* generated FlipFlop mice in which *Cd4* and *Cd8* genes encoded co-receptor proteins opposite of wild-type mice. This switch resulted in a reversal of T cell immunity with MHC II–responsive cytotoxic CD4$^+$ T cells and MHC I–reactive CD8$^+$ helper T cells. These findings argue against TCR signal strength as the operational paradigm. Rather, *Cd4* and *Cd8* co-receptor gene loci appear to determine T cell lineage fate by regulating the kinetics and duration of TCR signaling. —STS

Nat. Immunol. **23**, 731 (2022).

SCIENTIFIC WORKFORCE
Fewer free rides, more collaboration

All collaborations have a "social loafer" who expends the least effort. With collaboration being increasingly required in today's academic and research contexts, understanding factors that shape social loafing in team tasks could lead to more efficient—and less stressful—group work. Gabelica *et al.* examined the degree to which goal orientation and changes in team learning affected social loafing. Using a sample of undergraduate business students, they implemented a multiwave design and conducted linear mixed-effects modeling to show that individual team members on teams scoring higher on team learning over 9 weeks of collaboration experienced a decrease in social loafing. Loafing tendencies were also found to be dynamic, suggesting that implementing strategies for maximizing team learning could lead to fewer free rides and more optimal collaborations. —MMc

J. Educ. Psychol. **114**, 716 (2021).

NANOMATERIALS
Tuning nanoparticle assembly

Polymers or peptides can form ordered arrays of nanostructures mainly through side-chain packing effects, but inorganic nanoparticles form superlattices mainly through physical effects such as solvent drying. Monahan *et al.* explored the effects of solvents on the self-assembly of peptoids (*N*-substituted glycine oligomers) bearing CdS or CdSe quantum dots (QDs). In aqueous solvent systems, the QDs helped to stabilize intermediate peptoid morphologies and ultimately formed multilayers of small peptoid sheets linked by QDs. In organic solvents, the QDs could either facilitate or inhibit formation, depending on the hydrophobicity of QD-capping ligands. When the QDs were larger than the peptoid monomers, large hybrid sheets formed that contained highly ordered square-packed arrays of QDs. —PDS

ACS Nano 10.1021/acsnano.2c01203 (2022).

CANCER GENOMICS

Chromatin profiles classify castration-resistant prostate cancers suggesting therapeutic targets

Fanying Tang†, Duo Xu†, Shangqian Wang†, Chen Khuan Wong†, Alexander Martinez-Fundichely, Cindy J. Lee, Sandra Cohen, Jane Park, Corinne E. Hill, Kenneth Eng, Rohan Bareja, Teng Han, Eric Minwei Liu, Ann Palladino, Wei Di, Dong Gao, Wassim Abida, Shaham Beg, Loredana Puca, Maximiliano Meneses, Elisa de Stanchina, Michael F. Berger, Anuradha Gopalan, Lukas E. Dow, Juan Miguel Mosquera, Himisha Beltran, Cora N. Sternberg, Ping Chi, Howard I. Scher, Andrea Sboner, Yu Chen*, Ekta Khurana*

INTRODUCTION: Untreated prostate cancers rely on androgen receptor (AR) signaling for growth and survival, forming the basis for the initial efficacy of androgen deprivation therapy (ADT). Yet the disease can relapse and progress to a lethal stage termed castration-resistant prostate cancer (CRPC). Reactivation of AR signaling represents the most common driver of CRPC growth, and next-generation AR signaling inhibitors (ARSIs) are now used in combination with ADT as a first-line therapy. However, ARSIs can result in selective pressure, thereby generating AR-independent tumors. The transition from AR dependence frequently accompanies a change in phenotype resembling developmental trans-differentiation or "lineage plasticity." Neuroendocrine prostate cancer, which lacks a defined pathologic classification, is the most studied type of lineage plasticity. However, most AR-null tumors do not exhibit neuroendocrine features and are classified as "double-negative prostate cancer," the drivers of which are poorly defined.

RATIONALE: Lineage plasticity studies in CRPC are limited by the lack of genetically defined patient-derived models that recapitulate the disease spectrum. To address this, we developed a biobank of organoids generated from patient biopsies to study the landscape of metastatic CRPC and allow for functional validation assays. Proteins called transcription factors (TFs) are drivers of tumor lineage plasticity. To identify the key TFs that drive the growth of AR-independent tumors, we integrated epigenetic and transcriptomic data generated from CRPC models.

RESULTS: We generated ATAC-seq (assay for transposase-accessible chromatin sequencing) and RNA-seq data from 22 metastatic human prostate cancer organoids, six patient-derived xenografts (PDXs), and 12 derived or traditional cell lines. We classified the 40 models into four subtypes and predicted key TFs of each subtype. We identified the well-characterized AR-dependent (CRPC-AR) and neuroendocrine subtypes (CRPC-NE) as well as two AR-negative/low groups, including a Wnt-dependent subtype (CRPC-WNT), driven by TCF/LEF TFs, and a stem cell–like (SCL) subtype (CRPC-SCL), driven by the AP-1 family

of TFs. We applied RNA-seq signatures derived from the organoids to 366 patient samples from two independent CRPC datasets, which recapitulated the four-subtype classification. We found that CRPC-SCL is the second most prevalent group and is associated with shorter time under ARSI treatment compared to CRPC-AR. Additional chromatin immunoprecipitation sequencing (ChIP-seq) analysis indicated that AP-1 works together with the proteins YAP, TAZ, and TEAD, revealing YAP/TAZ and AP-1 as potential actionable targets in CRPC-SCL. Using overexpression assays in AR-high cells, we revealed how AP-1 functions as a pioneering factor and master regulator for CRPC-SCL.

CONCLUSION: By using a diverse biobank of organoids, PDXs, and cell lines that recapitulate the heterogeneity of metastatic prostate cancer, we created a map of the chromatin accessibility and transcriptomic landscape of CRPC. We validated the CRPC-AR and CRPC-NE subtypes and report two subtypes of AR-negative/low samples as well as their respective key TFs. Additional analysis revealed a model in which YAP, TAZ, TEAD, and AP-1 function together and drive oncogenic growth in CRPC-SCL samples. Overall, our results show how stratification of CRPC patients into four subtypes using their transcriptomes can potentially inform appropriate clinical decisions. ∎

S **READ THE FULL ARTICLE AT**
https://doi.org/10.1126/science.abe1505
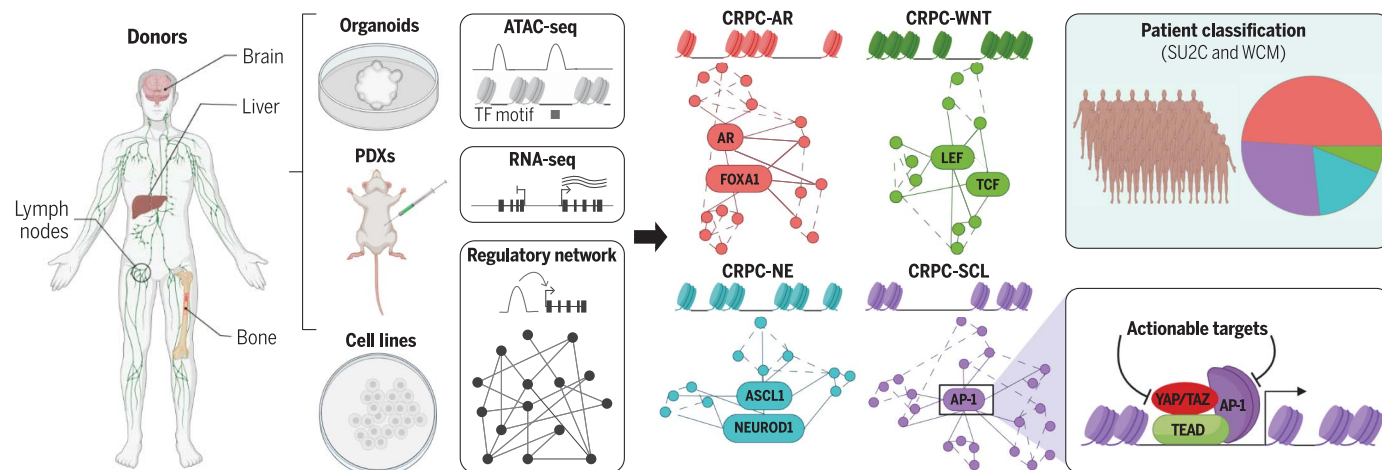
Identification of four subtypes of castration-resistant prostate cancer (CRPC) by integration of chromatin accessibility and transcriptomic data from organoids, patient-derived xenografts (PDXs), and cell lines. TF, transcription factor; AR, androgen receptor; NE, neuroendocrine; SCL, stem cell–like. YAP/TAZ/TEAD/AP-1 cooperation in CRPC-SCL suggests actionable targets. Application of RNA-seq signatures derived from the models to 366 patient samples recapitulates the four-subtype classification.

CANCER GENOMICS

Chromatin profiles classify castration-resistant prostate cancers suggesting therapeutic targets

Fanying Tang[1,2]†‡, Duo Xu[1,2,3,4]†, Shangqian Wang[5,6]†, Chen Khuan Wong[5]†,
Alexander Martinez-Fundichely[1,2,3,4], Cindy J. Lee[5], Sandra Cohen[1], Jane Park[7], Corinne E. Hill[7],
Kenneth Eng[4]§, Rohan Bareja[4], Teng Han[5], Eric Minwei Liu[1,2,8], Ann Palladino[1,2], Wei Di[5], Dong Gao[5,9],
Wassim Abida[10], Shaham Beg[4], Loredana Puca[1,4]¶, Maximiliano Meneses[11], Elisa de Stanchina[11],
Michael F. Berger[12], Anuradha Gopalan[12], Lukas E. Dow[1,13], Juan Miguel Mosquera[1,4,14],
Himisha Beltran[4,15], Cora N. Sternberg[1,4], Ping Chi[5,10,13], Howard I. Scher[10,16], Andrea Sboner[2,4,14],
Yu Chen[5,10,13]*, Ekta Khurana[1,2,3,4]*

In castration-resistant prostate cancer (CRPC), the loss of androgen receptor (AR) dependence leads to clinically aggressive tumors with few therapeutic options. We used ATAC-seq (assay for transposase-accessible chromatin sequencing), RNA-seq, and DNA sequencing to investigate 22 organoids, six patient-derived xenografts, and 12 cell lines. We identified the well-characterized AR-dependent and neuroendocrine subtypes, as well as two AR-negative/low groups: a Wnt-dependent subtype, and a stem cell–like (SCL) subtype driven by activator protein–1 (AP-1) transcription factors. We used transcriptomic signatures to classify 366 patients, which showed that SCL is the second most common subtype of CRPC after AR-dependent. Our data suggest that AP-1 interacts with the YAP/TAZ and TEAD proteins to maintain subtype-specific chromatin accessibility and transcriptomic landscapes in this group. Together, this molecular classification reveals drug targets and can potentially guide therapeutic decisions.

Untreated prostate cancers rely on androgen receptor (AR) signaling for growth and survival, forming the basis for the initial efficacy of androgen deprivation therapy (ADT). Yet the disease can relapse and progress to a lethal stage termed castration-resistant prostate cancer (CRPC). Reactivation of AR signaling represents the most common driver of CRPC growth, and next-generation AR signaling inhibitors (ARSIs) are now used in combination with ADT as a first-line therapy (1). However, ARSIs can also result in selective pressure, thereby generating AR-independent tumors. The transition from AR dependence frequently accompanies a change in phenotype that resembles developmental transdifferentiation or "lineage plasticity" (2). Neuroendocrine prostate cancer (NEPC), which lacks a defined pathologic classification, is the most studied type of lineage plasticity (3, 4). However, most AR-null tumors do not exhibit neuroendocrine features and

are classified as "double-negative prostate cancer" (DNPC), the drivers of which are poorly defined (5, 6).

Mechanistic studies in CRPC are limited by the lack of genetically defined patient-derived models that recapitulate the disease spectrum. To address this, we have developed a biobank of organoids generated from patient biopsies to study the landscape of metastatic CRPC and allow for functional validation assays (7, 8).

Biobank of patient-derived organoids of metastatic CRPC

We generated and characterized 15 organoids from specimens of patients with metastatic prostate cancer (MSKPCa8-MSKPCa20, MSKPCa22, MSKPCa24), adding to our biobank of seven organoids (7, 8). In general, the organoids were from patients with aggressive disease, short response to initial ADT, and rapid progression following second-line treatment with an ARSI (table S1). In culture, the organoids adopted

histology similar to the tissues from which they were developed (fig. S1, A and B), and the neuroendocrine samples maintained immunohistochemistry staining of synaptophysin (SYP) (fig. S1C).

We generated mutational and copy number profiles of each organoid, as well as 10 of 15 matching tumor biopsy specimens, using MSK-IMPACT (Memorial Sloan Kettering–Integrated Mutation Profiling of Actionable Cancer Targets) (9). The copy number landscape was similar between tumors and organoids and was representative of metastatic CRPC when compared to the Stand Up to Cancer (SU2C) cohort (10) (Fig. 1A). We observed a mean of 3.6 somatic mutations per patient, similar to the cohort of metastatic prostate cancer patients profiled using MSK-IMPACT (9) (Fig. 1B). The majority of organoids exhibited the same copy number variations (CNVs) and single-nucleotide variants (SNVs) as the original biopsies (fig. S1D and tables S2 and S3). In fact, organoids contained a higher fraction of tumor cells than the original biopsies, as shown by the increased allelic frequency of SNVs and CNVs (Fig. 1A and table S2).

Chromatin accessibility landscape reveals four molecular subtypes of metastatic prostate cancer

We performed ATAC-seq (assay for transposase-accessible chromatin sequencing) assays for 35 metastatic prostate cancer models, including 22 patient-derived organoids, six patient-derived xenografts (PDXs), and seven cell lines (two biological replicates for each) (fig. S2, A to E, and table S4). We also included published ATAC-seq data from five NEPC models (11) (Fig. 1C). Overall, we identified 861,195 reproducible peaks. The majority of the ATAC-seq peaks mapped to distal intergenic and intronic regions, similar to reports by other groups (12) (Fig. 1C). We identified four CRPC subtypes using consensus k-means clustering on the regions showing the most variable accessibility (fig. S2, F to H). We obtained the same four groups using other approaches, such as hierarchical clustering and UMAP (uniform manifold approximation and projection) (Fig. 1, D and E). There was no significant difference between the numbers of peaks among the four subtypes (fig. S2I, two-sided Wilcoxon rank-sum

Fig. 1. Classification of metastatic prostate cancer into four molecular subtypes from chromatin accessibility. (**A**) Top: Genomic aberrations in the prostate oncogenome from the SU2C CRPC patient samples (*10*) and MSKPCa organoids. Bottom: Copy number landscape of the 15 patient-derived organoid lines and 10 matching tumor tissues using MSK-IMPACT sequencing data. Shades of red and blue represent extent of gain and loss. (**B**) Number of mutations and fraction of copy number–altered genome of the 10 organoids and their matching patient tumor tissues. (**C**) Feature distribution of the mapped ATAC-seq peaks across all samples. (**D**) Correlation heatmap based on the normalized number of reads of the top 1% variable peaks across all samples. The ATAC-seq group and the sample source are indicated for each sample. The colors of the four ATAC-seq groups are kept consistent throughout the paper. (**E**) Unsupervised UMAP on the top 1% variable accessible peaks across all samples. (**F**) Immunoblot showing the expression of AR, SYP, and GAPDH (control) across the 35 organoids, PDXs, and cell lines.

test). The signature peaks for each subtype, defined as those with $\log_2$(fold change) > 2 and P < 0.01, were enriched at enhancers relative to promoters ($P = 4.18 \times 10^{-126}$, one-tailed Fisher's exact test), consistent with reports that the enhancer landscape better reflects cell identity than promoter accessibility (table S5) (*13*).

One of the four subtypes consisted of high AR-expressing samples, including the cell lines LNCaP, VCaP, 22Rv1, and C4-2; the organoids MSKPCa2, MSKPCa19, and MSKPCa22; and five PDXs (Fig. 1F). All these samples showed increased chromatin accessibility at AR target genes such as *KLK2* (fig. S2J). Another subtype included published NEPC samples H660, WCM154, PARCB1, PARCB3, PARCB6, PARCB8, and MSKEF1 (a derivative of published neuroendocrine organoid MSKPCa4); three new organoid lines, MSKPCa10, MSKPCa14, and MSKPCa24; and one PDX model with high SYP expression and small-cell carcinoma phenotype (Fig. 1F and fig. S1, A and C). The remaining two subtypes consisted of neuroendocrine-negative (NE-negative) and AR-negative/low samples (Fig. 1, E and F). For further biological characterization, we integrated the ATAC-seq data with RNA-seq and DNA sequencing data.

Transcriptomic profiles of the four CRPC subtypes

We next analyzed the transcriptomes of the 40 samples using UMAP and found that the clusters agree with the subtypes identified using ATAC-seq (Fig. 2A). Gene set enrichment analysis (GSEA) and selective marker gene expression (Fig. 2, B and C, and fig. S3, A to C) were used to name the four subtypes as follows: (i) CRPC-AR, which is enriched in the AR signature (*14*); (ii) CRPC-WNT, which is enriched in Wnt signaling and includes the organoids WCM1078, WCM1262, MSKPCa1, and MSKPCa16; (iii) CRPC-NE, which is enriched in the NE signature (*15*) in agreement with the pathology classification (fig. S1, A and C) and has high expression of NE markers, including SYP, CHGA, and DLL3; and (iv) CRPC-SCL, consisting of stem cell–like (SCL) samples, including 11 organoids and cell lines DU145 and PC3.

CRPC-SCL has not been previously identified. Samples in this subtype were enriched in the mammary stem cell signature, with high expression of the cancer stem cell markers CD44 and TACSTD2 (TROP2A) (Fig. 2, B and C, and fig. S3, A and B). Samples in CRPC-SCL were also enriched in pathways involving interleukin-6/Janus kinase/signal transducer and activator of transcription 3, transforming growth factor–β, tumor necrosis factor–α signaling, epithelial-mesenchymal transition, inflammation, and interferon response (fig. S4A).

Relative to CRPC-AR, the other three groups were enriched with a basal signature (*16*) and prostate basal stem cell signature (*17*), with

CRPC-SCL exhibiting the highest enrichment score (fig. S4B) and expression of basal cell markers (Fig. 2C). In addition, consistent with previous studies of AR-negative/low tumors, CRPC-WNT and CRPC-SCL showed enrichment of fibroblast growth factor receptor (FGFR) signaling and expression of selective FGF ligands and receptors compared to the other two groups (fig. S5, A and B) (*5*).

Genomic characterization and loss of tumor suppressors in the four CRPC subtypes

Samples in CRPC-AR were enriched for AR amplification and/or AR mutation (Fig. 2D; $P = 7.01 \times 10^{-6}$, one-tailed Fisher's exact test). In CRPC-WNT, all four samples showed alterations in the Wnt signaling pathway (*18*) (Fig. 2D and fig. S5C). Three CRPC-WNT samples showed hot spot mutations in CTNNB1 (β-catenin) (Fig. 2D and fig. S5C). The fourth sample had shallow deletion of APC and gain of RSPO2 (*18*) (fig. S5C).

Loss of the tumor suppressors TP53, PTEN, and RB1 is associated with lineage plasticity and aggressive disease in CRPC (*16*, *19*). We found that *TP53* was the most frequently mutated gene, with putative driver mutations or deep deletions in 23 of 35 samples (66%) across all four groups. *RB1* and *PTEN* had biallelic alterations in 20% and 43% of samples, respectively (Fig. 2D and fig. S5D). Using RNA-seq and immunoblot analysis, we found that an additional 11 of 35 samples (31.4%) exhibited RB1 loss, and 10 of 35 samples (28.5%) exhibited PTEN loss (table S6 and figs. S5E and S6). Overall, we found an enrichment of RB1 loss in AR-negative/low samples (14 of 24) compared to CRPC-AR (2 of 12) ($P = 0.0200$, one-tailed Fisher's exact test); there was no statistical difference in *PTEN* and *TP53* alterations between CRPC-AR and others. AR-independent CRPC has worse prognosis, and thus these results agree with recent studies indicating that *RB1* alterations, but not *TP53* and *PTEN* alterations, are associated with shorter survival in CRPC (*20*, *21*). It is notable that although 11 of 24 lines exhibited loss of both *TP53* and *RB1* in the AR-negative/low samples, only three were overtly NEPC (fig. S1, A to C), consistent with recent observations that loss of *TP53* and *RB1* in prostate carcinoma attenuates AR signaling but does not uniformly induce the neuroendocrine phenotype (*22*). This further highlights the importance of transcriptomic and epigenetic analysis in defining CRPC subtypes.

Construction of regulatory networks and identification of key TFs

To identify the key TFs that drive the subtype-specific transcriptome, we first identified the hubs in regulatory networks that target a large number of genes in a given sample (*23*). We constructed regulatory networks by integrat-

ing ATAC-seq and RNA-seq data, and built the peak-gene links based on the correlation between chromatin accessibility at ATAC-seq peaks and expression of genes within ±0.5 Mb (Fig. 3A, step 1) (*12*, *24*). In total, we identified at least one peak-gene link for 4752 protein-coding genes (table S7). We predicted that 75.2% of the peaks regulate only one gene, and on average the expression of one gene was correlated with the activity of three peaks (Fig. 3, B and C). To uncover TF-DNA binding sites in the accessible regions, we used a footprinting method called HINT-ATAC (Hmm-based identification of TF footprints using ATAC-seq) and a curated collection of sequence-binding motifs for 809 TFs from CIS-BP (Fig. 3A, step 2) (*25*). By combining peak-gene and TF-peak links, we constructed TF-gene links and generated sample-specific regulatory networks (Fig. 3A, steps 3 and 4). We define TF out-degree as the number of target genes a given TF regulates in the network (fig. S7).

Next, we identified the key TFs for each subtype as those at the top of the gene regulation hierarchy (Fig. 3D and table S8). Each TF is ranked according to a combination of three metrics: (i) its differential out-degree O_diff (fig. S8A), (ii) its differential chromatin accessibility at its motifs A_diff (fig. S8B), and (iii) its differential gene expression E_diff in a given subtype relative to others (Fig. 3D). We assigned ranks to the TFs independently on the basis of the three metrics and added up the three ranks to get the final TF_rank.

In CRPC-AR, AR and FOXA1 were the top two TFs, validating our approach (*26*, *27*). In CRPC-NE, the top two TFs were neurogenic differentiation factor 1 (NEUROD1) and achaete-scute homolog 1 (ASCL1). NEPC and small-cell lung cancer (SCLC) have been shown to be similar at the phenotypic and molecular level, and ASCL1 and NEUROD1 have been demonstrated to be the main drivers in SCLC (*11*, *28*). In CRPC-WNT, transcription factor 7–like 2 (TCF7L2) was the highest-ranked TF. Also known as TCF-4, TCF7L2 has been shown to be the key driver in colorectal cancer upon upstream Wnt pathway gene alterations such as APC mutations (*29*). Other TCF and lymphoid enhancer binding factor (LEF) TFs were also among the top candidates, including LEF1/LEF, TCF7/TCF-1, and TCF7L1/TCF-3. Upon Wnt pathway activation, β-catenin translocates to the nucleus and coactivates TCF/LEF to promote the expression of downstream genes (*29*).

In CRPC-SCL, we identified the AP-1 family among the top TFs, with FOSL1 having the highest rank. AP-1 is a TF complex assembled through homo- or heterodimerization of members of the Fos and Jun family (*30*). The Fos family includes FOSL1, FOSL2, FOS, and FOSB, whereas the Jun family includes JUN, JUNB, and JUND. AP-1 has been shown to be activated by multiple upstream signals, including growth
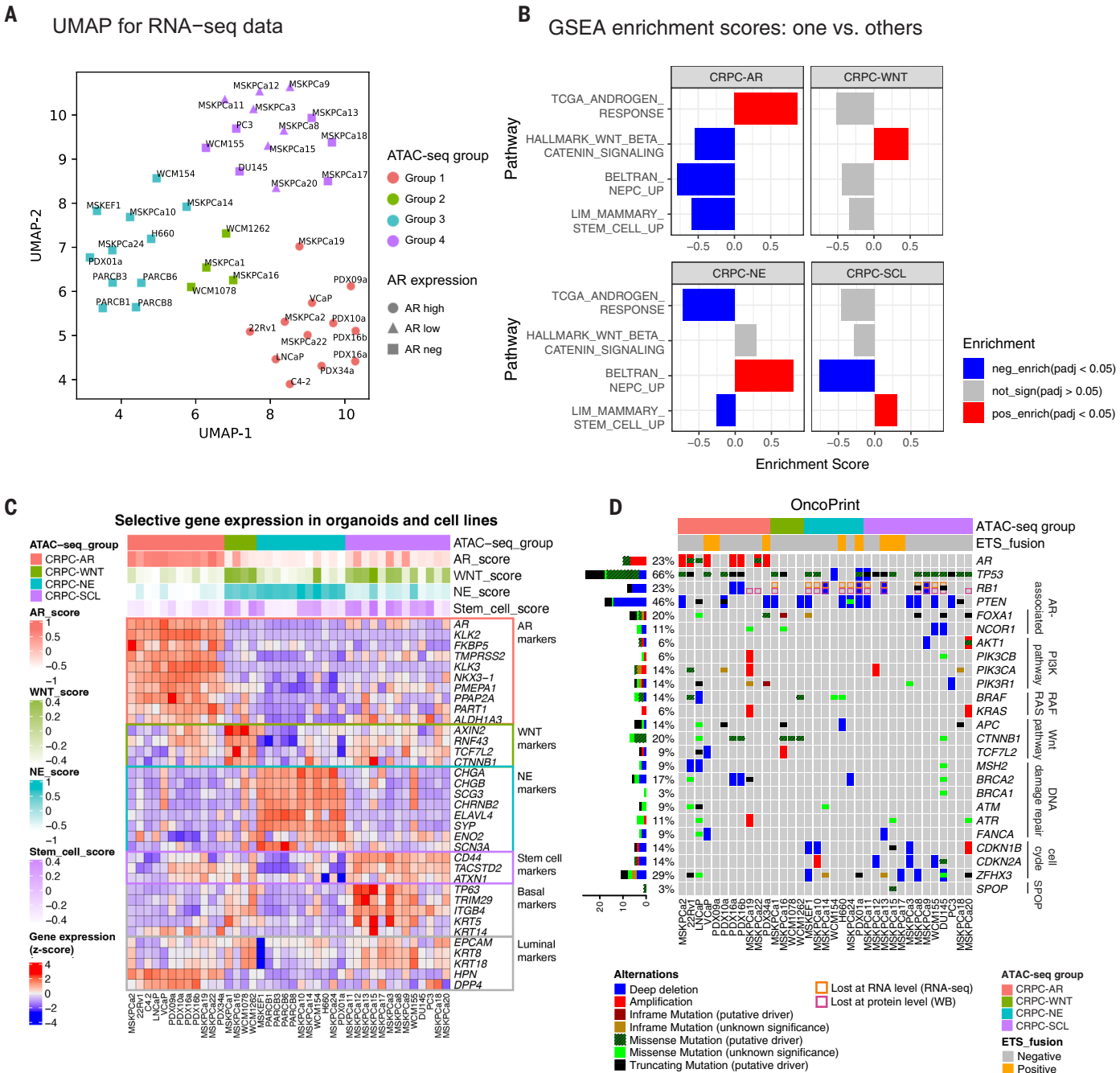
Fig. 2. Transcriptomic and genomic characterization of the four CRPC subtypes defined by ATAC-seq. (**A**) Unsupervised UMAP on the mRNA expression values of the 1000 most variably expressed genes across all samples. (**B**) Enrichment scores and *P* values from GSEA indicate that the four signals are significantly positively enriched in specific subtypes but not in others. (**C**) Heatmap shows the relative expression of subtype-specific marker genes and basal/luminal genes across all samples. The ATAC-seq group and signature scores of the four representative pathways for each sample are shown at the top. (**D**) OncoPrint shows the genomic alterations of the 35 samples with DNA-sequencing data. MSKPCa8 to 20, 22, and 24 were sequenced with MSK-IMPACT. The SNVs and CNVs for cell lines (LNCaP, 22Rv1, VCaP, H660, DU145, and PC3) were collected from CCLE (Cancer Cell Line Encyclopedia). Whole-exome sequencing (WES) was performed for other samples.

factors, hormones, cytokines, inflammation, and stresses. It controls the expression of many downstream genes related to cell division, apoptosis, cell migration, and immunity (*30*).

In addition to subtype-specific TF identification, we also predicted the important TFs

in a sample-specific manner. To do this, we analyzed all TFs' relative expression and chromatin accessibility at their motifs for samples in CRPC-WNT, CRPC-NE, and CRPC-SCL, and compared these to the average for CRPC-AR samples. This analysis demonstrated that the

key TFs identified in a subtype-specific manner agree with the sample-specific results (fig. S9). We also found significant correlation between the expression of key TFs and accessibility at their motifs, which suggests that they likely exhibit pioneering activity (fig. S10,

Fig. 3. Identification of the key transcription factors (TFs) of each subtype. (**A**) Schematic illustrating the construction of sample-specific regulatory networks using ATAC-seq and RNA-seq data. (**B**) Distribution of the number of genes linked per peak. (**C**) Distribution of the number of peaks linked per gene. (**D**) Rank order of the top 25 TFs for each of the four subtypes. For each TF, the relative contributions of three metrics to TF rank are shown (Expr, expression). In CRPC-SCL, FOSL1, BATF, FOSL1, JUNB, JDP2, FOS, MAFF, FOSB, and ATF3 belong to the AP-1 family.

adjusted $P < 0.05$) (*12*). Finally, the relative expression levels of predicted key TFs in each of the four groups were confirmed by a quantitative polymerase chain reaction (qPCR) (fig. S3C).

Classification of CRPC patients using transcriptomic signatures of the four subtypes

Next, we examined RNA-seq datasets from 366 CRPC patients to assign each patient to the four subtypes (*21*). We derived the sig-

nature genes for each of the four subtypes as the ones with higher expression in one group relative to others in organoids and cell lines, and filtered out genes with low expression or low variance in CRPC patient samples

(tables S5 and S9). A template was constructed for each subtype by combining the top signature genes. To assign the CRPC patient samples to a subtype, we used the nearest template prediction (NTP) algorithm (*31*, *32*), which involves computing the cosine distance (*d*) between each patient's RNA-seq data and each of the four templates and estimating the statistical significance by random resampling (Fig. 4A). We applied the method to two cohorts of CRPC patients with polyA-enriched RNA-seq data, including 266 published SU2C patients (*21*)
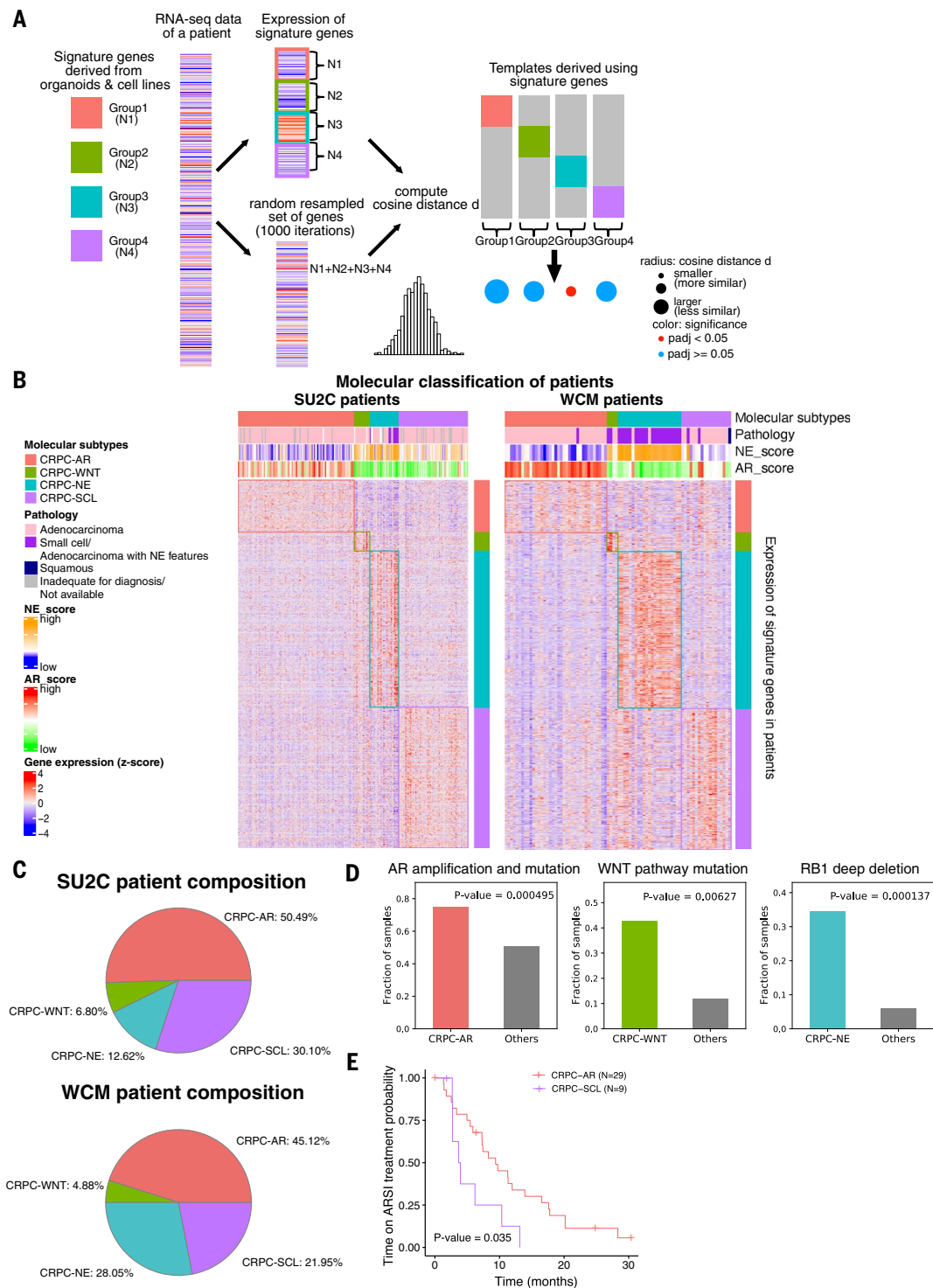
Fig. 4. Classification of CRPC patients using transcriptomic signatures for the four subtypes. (**A**) Schematic illustrating the assignment of each patient to one of the four groups using nearest template prediction. (**B**) Heatmaps showing relative expression of signature genes in SU2C and WCM patients. Top annotations indicate the AR score, NE score, pathology classification, and molecular subtypes of each patient. (**C**) Patient compositions for SU2C and WCM cohorts. (**D**) AR amplification and mutation, WNT pathway component mutations, and RB1 deep deletion are enriched in CRPC-AR, CRPC-WNT, and CRPC-NE SU2C patients, respectively (two-sided Fisher's exact test). (**E**) CRPC-SCL patients exhibit shorter time on ARSI treatment relative to CRPC-AR SU2C patients.

and 100 patients sequenced at Weill Cornell Medicine (WCM) (Fig. 4B). The majority of the patients (206 of 266 SU2C, 82 of 100 WCM) were assigned to one of the four subtypes (table S10). The relative ratios of the four subtypes of patients were similar between cohorts, with the largest group being CRPC-AR, then CRPC-SCL or CRPC-NE, and finally CRPC-WNT as the smallest (Fig. 4C). For WCM1078, WCM1262, WCM155, and MSKPCa2 organoids, RNA-seq data were available from matching tumor samples, which were assigned to the same subtype as the organoids (table S10).

As a complementary approach, we developed a linear SVM (support vector machine) using the gene expression data of the signature genes. Patients were assigned to the four subtypes on the basis of the highest probability from SVC (C-support vector classification). Application of the model to patients' RNA-seq data showed that for most subtypes, the mean probability of patient assignment to that subtype was ≥0.50 (table S10), indicating confident assignment to one group. The majority of samples were assigned the same subtype with either NTP or SVM (92% of WCM and 85% of SU2C samples). The samples unassigned by NTP tended to show more heterogeneity based on SVM (lower highest probabilities compared to the assigned, $P = 4.36 \times 10^{-6}$ for WCM and 1.73×10^{-6} for SU2C, Wilcoxon rank-sum test) and a tendency toward CRPC-SCL ($P = 0.033$ for WCM and 0.00029 for SU2C, Fisher's exact test), potentially pointing to transition to/via this subtype.

The genomic alterations, marker gene expression, and pathologic analysis provided validation of patient classification. CRPC-AR patients showed enrichment of *AR* amplification (Fig. 4D and fig. S11A) and had higher *AR* expression and AR score (Fig. 4B and fig. S11, B and C) compared to other groups. CRPC-NE patients had higher *SYP* expression and NE score (*21*) (fig. S11, B and C) compared to others, and their genotypes were enriched with *RB1* deep deletion (Fig. 4D and fig. S11A). The majority of patients in this class were also diagnosed as having either small-cell, NEPC, or adenocarcinoma with NE features on the basis of histology analysis (Fig. 4B). Patients in CRPC-WNT showed elevated expression of *AXIN2* (fig. S11B) and an enrichment of mutations of Wnt pathway components (Fig. 4D and figs. S11A and S12A). We observed increased expression of the stem cell marker *CD44* in CRPC-SCL patients (fig. S11B) compared to others, as expected, but no consistent enrichment of gene or pathway alterations at the genomic level (figs. S11A and S12B). Moreover, consistent with the results in the 40 models, we found an enrichment of basal signature in the three AR-low/negative groups relative to CRPC-AR (fig. S12C).

Among the 266 SU2C patients, 56 had time-on-treatment data for the next-generation ARSIs enzalutamide and abiraterone acetate. We found that patients classified as CRPC-SCL exhibited shorter time on ARSI treatment using Cox log-rank statistics (Fig. 4E), indicating that the ARSI treatments were less effective for CRPC-SCL patients. We could not compare the time on ARSI treatment for CRPC-AR or CRPC-SCL to other subtypes because there were fewer than five samples for CRPC-WNT and CRPC-NE (table S11).

AP-1 cooperates with YAP, TAZ, and TEAD in CRPC-SCL

The proportion of patients classified as CRPC-SCL was the second largest in the combined SU2C and WCM cohorts (28%) (Fig. 4C and table S10); thus, we further explored samples in this subtype. We focused on MSKPCa3, an AR-low organoid, and DU145, an AR-negative cell line, as CRPC-SCL models for experimental validations.

We identified an AP-1 family member, FOSL1, as the top candidate key TF for CRPC-SCL (Fig. 3D). Expression of various AP-1 components across the four subtypes confirmed *FOSL1* as the AP-1 gene with highest relative expression in CRPC-SCL samples compared to others (fig. S13A), whereas it was barely detectable in CRPC-AR samples at mRNA and protein levels (fig. S13, A and B). To directly assess the importance of FOSL1 for tumor growth, we performed cell competition assays in MSKPCa3 and DU145, with CRPC-AR organoid MSKPCa2 as a control. We transduced the cells with constructs containing green fluorescent protein (GFP), Cas9, and single guide RNAs (sgRNAs) against *FOSL1*, *RPA3* (positive control), or *Rosa26* (negative control) and monitored the relative proportion of GFP-positive sgRNA-expressing cells over time by fluorescence-activated cell sorting (FACS) (fig. S13C). Depletion of GFP-positive sg*FOSL1* was observed in both MSKPCa3 and DU145, but not in MSKPCa2, supporting our prediction that FOSL1 is important for tumor progression in CRPC-SCL (Fig. 5A and fig. S13, D and E).

TFs work together and bind cooperatively in a context-specific manner to achieve specificity and execute their functions (*33*). Thus, to further investigate the regulation of CRPC-SCL samples by AP-1, we investigated the other top TFs identified from chromatin accessibility profiles. We found that TEAD motifs were the second most enriched after AP-1 in the CRPC-SCL–specific accessible peaks (fig. S14A) and they ranked highly on the basis of the gain of chromatin accessibility and out-degree in CRPC-SCL samples (fig. S14B). TEAD TFs are activated by YAP and TAZ transcriptional coactivators (*34*). In fact, motif analysis of the chromatin immunoprecipitation sequencing (ChIP-seq) peaks of YAP and TAZ has revealed that TEAD TFs are the main platform by which these proteins interact with DNA (*34*). TEAD, YAP, and TAZ were reported to be associated with AP-1 genome-wide to jointly regulate the proliferation and motility in multiple cancers, including breast, colorectal, and lung (*35*). In addition, *TAZ* (*WWTR1*) is among the top genes codependent with *FOSL1* based on CRISPR (Avana) Public 20Q2 in DepMap (*36*) with a Pearson correlation of 0.33, further supporting the model in which FOSL1 functions together with YAP and TAZ. From ATAC-seq data and published literature, we hypothesized that YAP, TAZ, TEAD, and AP-1 (FOSL1) may function together to promote the oncogenic growth of CRPC-SCL tumors (Fig. 5B). This is supported by our observation that GSEA using the combined YAP and TAZ (YAP/TAZ) target signature as defined in (*37*) revealed strong enrichment [false discovery rate (FDR) < 0.001] in CRPC-SCL compared to other samples (Fig. 5C). CRPC-SCL also showed significantly higher expression of *YAP* and *TAZ* (fig. S14C; $P < 0.05$, Wilcoxon rank-sum test), and qPCR analysis of representative YAP/TAZ target genes across the 28 samples showed their high expression in this group (fig. S14D).

To validate the co-binding of AP-1 (FOSL1), TEAD, YAP, and TAZ, we performed ChIP-seq in MSKPCa3 and DU145. We found significant enrichment of overlaps between the ChIP-seq peaks of these proteins in both MSKPCa3 and DU145, pointing to their cooperation (Fig. 5, D and E, fig. S15, A and B, and table S12; $P < 0.001$, Fisher's exact test). We also found significant overlap between the target genes of AP-1 and TEAD predicted in our regulatory networks (fig. S15C; $P < 0.001$, Fisher's exact test).

We found that the ChIP-seq peaks of all these four proteins exhibit large overlap with CRPC-SCL ATAC-seq peaks, but barely any with CRPC-AR peaks (Fig. 5F and figs. S15D and S16). Correspondingly, we also observed a strong enrichment of the ChIP-seq signal over CRPC-SCL–specific ATAC-seq peaks relative to the other three subtypes (Fig. 5F and fig. S16; $P < 0.0001$, one-sided Fisher's exact test). As a negative control, the trend was opposite for AR ChIP-seq peaks, in which they showed much larger overlap and signal enrichment at CRPC-AR peaks compared to CRPC-SCL, as expected (Fig. 5F and fig. S16; $P < 0.0001$, one-sided Fisher's exact test) (*38*, *39*). For example, the ChIP-seq and ATAC-seq profiles illustrated open chromatin and binding of AP-1, YAP, TAZ, and TEAD at enhancers of the representative YAP/TAZ target genes, *CYR61* and *AXL*, in CRPC-SCL lines; the same loci were barely accessible in other groups (figs. S17 and S18).

To determine the role of YAP and TAZ in growth, we used small interfering RNA (siRNA) to knock down *YAP* and *TAZ* alone or together in MSKPCa3 and DU145 cells (Fig. 6, A and B, and fig. S19, A and B; $P < 0.05$, two-tailed unpaired *t* test). We observed a significant decrease of cell growth upon *TAZ* and *YAP*/*TAZ* double knockdown in both MSKPCa3 and DU145 but not in the AR-dependent lines MSKPCa2
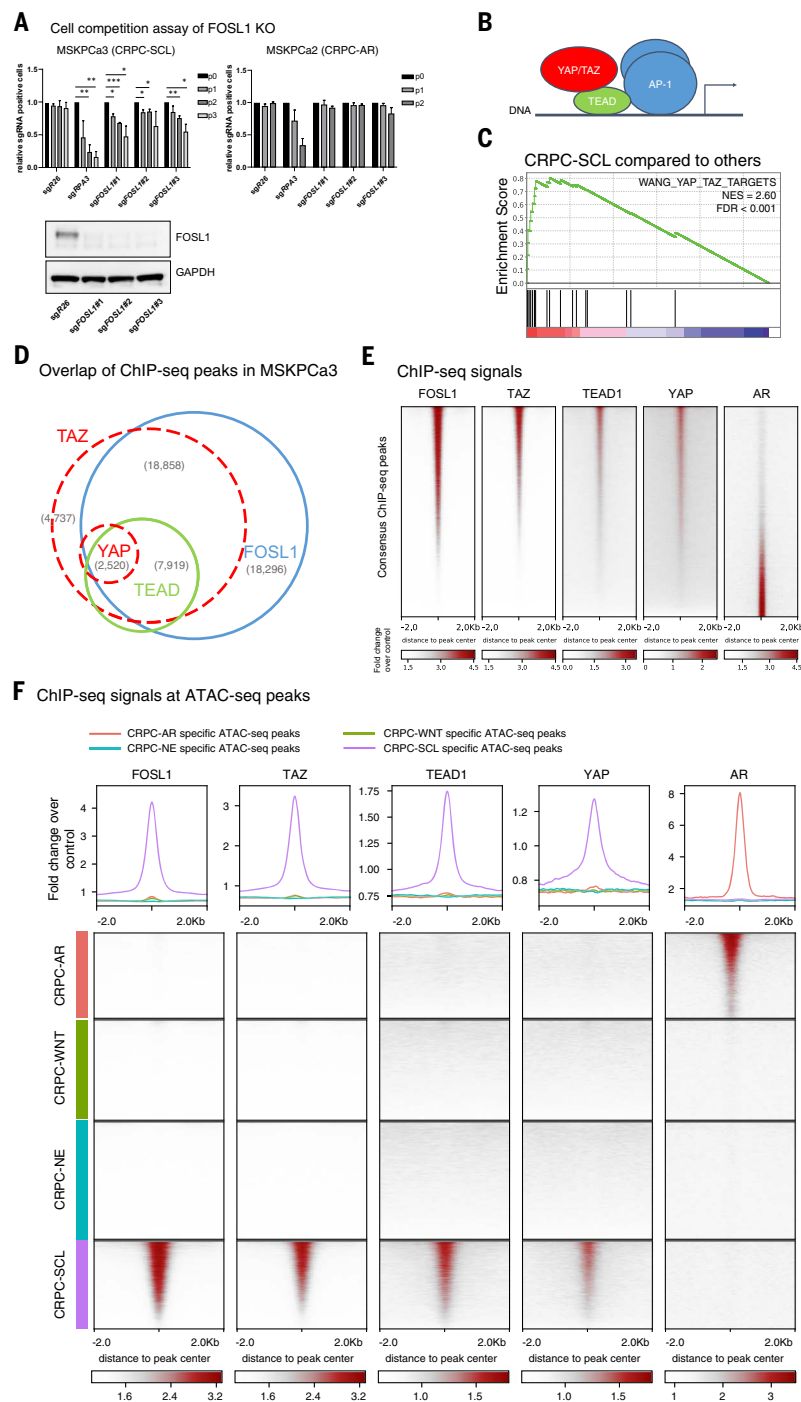
Fig. 5. AP-1 works together with YAP, TAZ, and TEAD in CRPC-SCL. (**A**) Percentage of GFP-positive MSKPCa3 (left) or MSKPCa2 (right) expressing CRISPR guides against FOSL1 or sgR26 (negative control) or sgRPA3 (positive control). Mean ± SEM, $n = 2$ for MSKPCa3, $n = 2$ for MSKPCa2. ***$P < 0.001$, **$P < 0.01$, *$P < 0.05$ [multiple unpaired t test comparing between passages (p1 versus p0, p2 versus p0, p3 versus p0)]. Knockout of FOSL1 was confirmed in MSKPCa3 by Western blot. (**B**) Schematic showing the cooperation of AP-1 with YAP/TAZ and TEAD in CRPC-SCL samples. (**C**) GSEA plot showing enrichment of YAP/TAZ signature in CRPC-SCL organoids and cell lines compared to other samples. (**D**) Venn diagram showing the overlaps of FOSL1, TEAD, YAP, and TAZ ChIP-seq peaks in MSKPCa3. Overlaps with more than 1000 peaks are marked, and those between YAP/FOSL1 (48 peaks) and YAP/FOSL1/TEAD (4 peaks) are not shown. (**E**) ChIP-seq signal of FOSL1, TAZ, TEAD1, and YAP in MSKPCa3, and AR (GSE61852) from LNCaP on the consensus peak set. (**F**) ChIP-seq signal of FOSL1, TAZ, TEAD1, and YAP in MSKPCa3, and AR (GSE61852) from LNCaP on subtype-specific ATAC-seq peaks. FOSL1, TEAD1, YAP, and TAZ ChIP-seq peaks for MSKPCa3 show stronger signal at CRPC-SCL–specific ATAC-seq peaks, whereas AR (GSE61852) ChIP-seq peaks have stronger signal in CRPC-AR–specific peaks.

and 22Rv1 (Fig. 6C and fig. S19C; $P < 0.0001$, two-tailed unpaired t test).

AP-1 and YAP/TAZ are important for subtype-specific chromatin accessibility and gene expression in CRPC-SCL and can be targeted by drugs

Knockdown of FOSL1 and YAP/TAZ shows impact on chromatin accessibility and gene expression

We performed ATAC-seq and RNA-seq in MSKPCa3 cells with knockdown of *FOSL1*, *YAP*, *TAZ*, and *YAP/TAZ* together. We observed a significant decrease of chromatin accessibility at the CRPC-SCL–specific open chromatin sites as well as at regions bound by FOSL1, TEAD, YAP, and TAZ from ChIP-seq upon *FOSL1*, *YAP*, and *YAP/TAZ* knockdown (Fig. 6, D and E, and fig. S20) ($P < 0.001$, permutation test). The signal was strongest upon *FOSL1* and double *YAP/TAZ* knockdown and highlights their important role in maintaining the chromatin accessibility landscape for the CRPC-SCL group. In agreement with the decrease in chromatin accessibility, we found that CRPC-SCL peaks are the most enriched among down-regulated ATAC-seq peaks upon *FOSL1*, *YAP*, and *YAP/TAZ* knockdown (table S13). Moreover, we found that the TEAD and AP-1 motifs are the most enriched at the regions where chromatin accessibility is reduced upon *YAP/TAZ* double knockdown, further confirming the model in which YAP and TAZ cooperate with AP-1 and TEAD (fig. S21A). Consistent with ATAC-seq results, we found down-regulation of YAP/TAZ targets in RNA-seq data in all knockdown assays (Fig. 6F and fig. S21B). Double knockdown of *YAP* and *TAZ* showed stronger inhibition of canonical downstream targets (*CTGF*, *CYR61*, *AJUBA*, and *ANKRD1*) and cell cycle–regulating gene *CCND1* (Fig. 6A and fig. S19A), which have been reported as YAP/TAZ targets in various model systems, relative to individual knockdown of *YAP* or *TAZ* (*35*, *40*). Although we also observed enrichment of CRPC-SCL genes among the down-regulated genes, their enrichment in the up-regulated gene set showed that ATAC-seq changes are likely a more meaningful marker to analyze lineage plasticity, as also demonstrated by previous studies (table S14) (*41*). Similar chromatin accessibility and gene expression changes were observed upon *YAP*, *TAZ*, and *YAP/TAZ* knockdown in DU145 (figs. S21, C and D, and S22).

Positive feedback loop between YAP/TAZ and FOSL1

We found that *YAP/TAZ* double knockdown caused robust depletion of FOSL1, the predicted master TF, at RNA and protein levels in both MSKPCa3 and DU145 (Fig. 6, A and B, and fig. S19, A and B). Moreover, in our regulatory networks for CRPC-SCL samples, *FOSL1* was predicted to be a target of FOS/JUN itself as
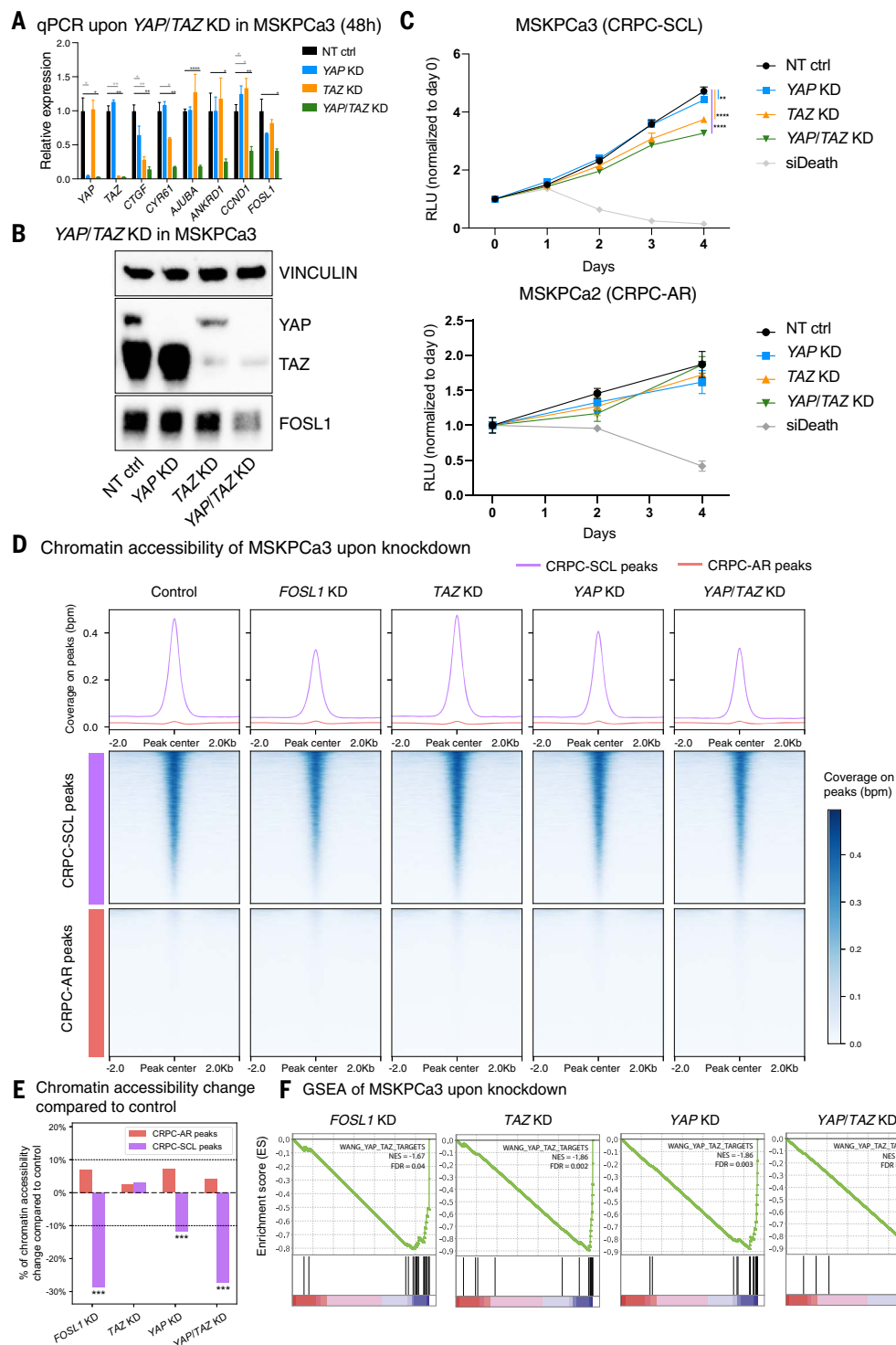
Fig. 6. Impact of *FOSL1*, *YAP*, *TAZ*, and *YAP/TAZ* knockdown. (**A**) qPCR showing that double knockdown of *YAP/TAZ* by siRNA in MSKPCa3 leads to decreased expression of downstream target genes and *FOSL1*. Mean ± SD, $n = 3$. Nontargeting (NT) siRNA serves as the negative control. ****$P <$ 0.0001, **$P <$ 0.01, *$P <$ 0.05 (two-tailed unpaired t test); KD, knockdown. (**B**) Western blot (WB) confirming the knockdown efficiency and the decrease of FOSL1 expression 72 hours after transfection in MSKPCa3. (**C**) Cell growth curves of MSKPCa3 and MSKPCa2 after siRNA knockdown. Mean ± SD, $n = 8$ for MSKPCa3, $n = 4$ for MSKPCa2. ****$P <$ 0.0001, **$P <$ 0.01 (two-tailed unpaired t test). (**D**) Chromatin accessibility changes upon *FOSL1*, *TAZ*, *YAP*, and *YAP/TAZ* knockdown at CRPC-AR and CRPC-SCL peaks. (**E**) Average percentage of chromatin accessibility changes upon *FOSL1*, *TAZ*, *YAP*, and *YAP/TAZ* knockdown compared to control at CRPC-AR and CRPC-SCL peaks. ***$P <$ 0.001 (permutation test). (**F**) GSEA plots showing negative enrichment of YAP/TAZ target genes in MSKPCa3 upon knockdown of *FOSL1*, *TAZ*, *YAP*, and *YAP/TAZ*.

well as TEAD, which cooperates with FOSL1 in our proposed model, similar to observations in other cancers (*35*). Using ChIP-seq data, we also found that the CRPC-SCL–specific *FOSL1* enhancer is bound by TEAD, YAP, TAZ, and FOSL1 (fig. S23). Together, the results suggest that YAP, TAZ, TEAD, and FOSL1 increase the expression of *FOSL1* itself, forming a positive feedback loop to further open chromatin (fig. S24A).

Exogenous expression of FOSL1 alters chromatin accessibility and gene expression from CRPC-AR toward CRPC-SCL

To determine whether FOSL1 could alter the chromatin accessibility landscape and activate the CRPC-SCL signature, we stably expressed *FOSL1* alone or in combination with *YAP* or *TAZ* in LNCaP cells. We observed an increase in chromatin accessibility at CRPC-SCL–specific open chromatin sites in all assays with exogenous expression of *FOSL1*, providing evidence of its role as a pioneering factor in potential lineage plasticity (Fig. 7, A and B; $P < 0.001$, permutation test). We also observed a decrease in chromatin accessibility at CRPC-AR–specific open chromatin sites in all *FOSL1* overexpression assays, further pointing toward lineage transformation (Fig. 7, A and B; $P < 0.001$, permutation test). The RNA-seq results were consistent with ATAC-seq results, and we observed significant up-regulation of CRPC-SCL signature genes with *FOSL1* overexpression, either alone or with *YAP* or *TAZ* (Fig. 7C and fig. S24B, FDR $< 10^{-5}$).

Small-molecule inhibitors for CRPC-SCL

We used two small-molecule inhibitors that act on the YAP/TAZ/AP-1 pathway for their potential use as therapeutics for CRPC-SCL tumors. Verteporfin is a benzoporphyrin derivative and a medication used as a photosensitizer approved by the FDA for the treatment of age-related macular degeneration. It has been widely reported to inhibit YAP/TAZ and cellular proliferation of multiple tumors (*42*). Consistent with the role of YAP/TAZ in CRPC-SCL, we found that MSKPCa3 and DU145 cells were more sensitive to verteporfin than were MSKPCa2 and 22Rv1, respectively (Fig. 7D and fig. S24, C and D). T-5224 is a c-Fos/AP-1 inhibitor, specifically affecting the DNA binding activity of c-Fos/c-Jun and under clinical trial for use in other cancers and diseases (*43*). We found that T-5224 inhibited MSKPCa3 and DU145 cell growth in a dose-dependent fashion, whereas it had no effect on MSKPCa2 and 22Rv1 (Fig. 7E and fig. S24, D and E).

The YAP/TAZ pathway is enriched in CRPC-SCL patients

Finally, we examined YAP/TAZ activity in transcriptomic data from CRPC patients from both SU2C and WCM. YAP/TAZ pathway activity (sum of z-scores) was significantly higher in CRPC-SCL patients relative to all samples (Fig. 7F; $P < 0.01$, one-tailed Wilcoxon rank-sum test), with higher expression of *YAP*, *TAZ*, and representative downstream genes (fig. S25A). We also observed a significant negative correlation between *AR* expression and YAP/TAZ pathway activity across all SU2C samples (Fig. 7G and fig. S25B; $P < 0.001$).

Discussion

We used a diverse biobank of organoids and PDXs that recapitulate the genotypic and phenotypic heterogeneity of metastatic prostate cancer to generate a map of the chromatin accessibility and transcriptomic landscape of CRPC. In so doing, we validated the CRPC-AR and CRPC-NE subtypes and identified two subtypes of AR-negative/low samples. Our integrated use of ATAC-seq and RNA-seq data allowed us to identify the master TFs driving AR-negative/low CRPCs. Previous studies using only RNA-seq data could not identify these drivers because GSEA identifies numerous biological processes that are enriched among CRPC-SCL samples (fig. S4A), complicating efforts to find driver events. Furthermore, our work shows that CRPC-SCL constitutes the second most prevalent group of CRPC patients, exhibits lower AR expression and AR transcriptional output, and is associated with shorter time under ARSI treatment compared to CRPC-AR.

Integrated analysis of ATAC-seq, RNA-seq, and ChIP-seq data revealed a model in which YAP, TAZ, TEAD, and AP-1 function together and drive oncogenic growth in CRPC-SCL samples. We validated this with CRISPR and depletion studies using siRNA knockdown. From overexpression assays in AR-high LNCaP cells, we showed how FOSL1 functions as a pioneering factor and master regulator for CRPC-SCL. This model reveals potential therapeutic vulnerabilities in CRPC-SCL tumors by inhibition of the YAP/TAZ/AP-1 pathway.

Prior studies support these conclusions. For example, the Wnt pathway has been identified as a driver of metastasis and resistance to AR-targeted therapies (*44*, *45*). Furthermore, knocking down *TAZ* in DU145 (CRPC-SCL from our study) decreased cell migration and metastasis, whereas overexpression of *TAZ* in RWPE (normal prostate cells) promoted cell migration, epithelial-mesenchymal transition, and anchorage-independent growth (*46*). Overexpression of *YAP* has been reported to promote cell proliferation, invasion, and castration-resistant growth in LNCaP and RWPE (*47*). In addition, YAP/TAZ activation has been found to be related to cell proliferation, therapy resistance, and metastasis in various other tumor types by extensive rewiring of the epigenome of differentiated cells, reprogramming them into stem-like cells and conferring lineage plasticity (*34*, *35*). These lines of evidence along with those from our study show the importance of AP-1, YAP, and TAZ in the generation and maintenance of the chromatin and transcriptomic landscape in a specific subtype of CRPC.

Enrichment of basal signature, such as we saw in CRPC-SCL organoids and patient samples, has been observed in prostate cancer cell lines after depletion of TP53 and RB1 (*16*, *19*) and is also observed in models of DNPC derived from AR knockout of luminal prostate cancer cells (*5*). This suggests that CRPC-SCL tumors acquire lineage plasticity similar to NEPC but are driven by different master TFs, resulting in a different phenotype. Because CRPC-SCL tumors are pathologically adenocarcinoma without neuroendocrine features, our study may guide the use of ARSIs in these cases. However, this requires further detailed mechanistic studies of lineage transformation and heterogeneity among the four subtypes. Moreover, future in vivo studies over longer time periods are needed to provide further insights about the efficacy and specificity of small molecules that target this subtype. Although we showed the differential impact of verteporfin on CRPC-SCL versus CRPC-AR samples, its YAP-independent effects may limit its clinical potential (*48*, *49*). Meanwhile, other AP-1 inhibitors not tested in this study may show high clinical potential for CRPC-SCL (*50*). Overall, we have shown how an approach to stratify CRPC patients into four subtypes using their transcriptomic signatures can potentially inform appropriate clinical decisions.

Methods summary

ATAC-seq and RNA-seq data were generated for 35 metastatic prostate cancer models, including 22 organoids, six PDXs, and seven cell lines. Together with five more derived CRPC cell lines from Park *et al.* (*11*), we uniformly processed ATAC-seq and RNA-seq data from 40 models. The ATAC-seq data were used to cluster the samples and revealed four subtypes. We generated gene signatures of the four epigenetically defined subgroups using RNA-seq data. These gene signatures were used to classify 366 CRPC patient samples. Moreover, we constructed regulatory networks for the models using a correlation-based method connecting ATAC-seq peaks to gene expression and a footprint-based method for TF to regulatory element connections. Key TFs for each subgroup were identified using a metric integrating regulatory network, ATAC-seq, and RNA-seq features. The proposed cooperation between key TFs and other proteins was validated using ChIP-seq. The impact of FOSL1, YAP, and TAZ on subtype-specific chromatin accessibility and gene expression
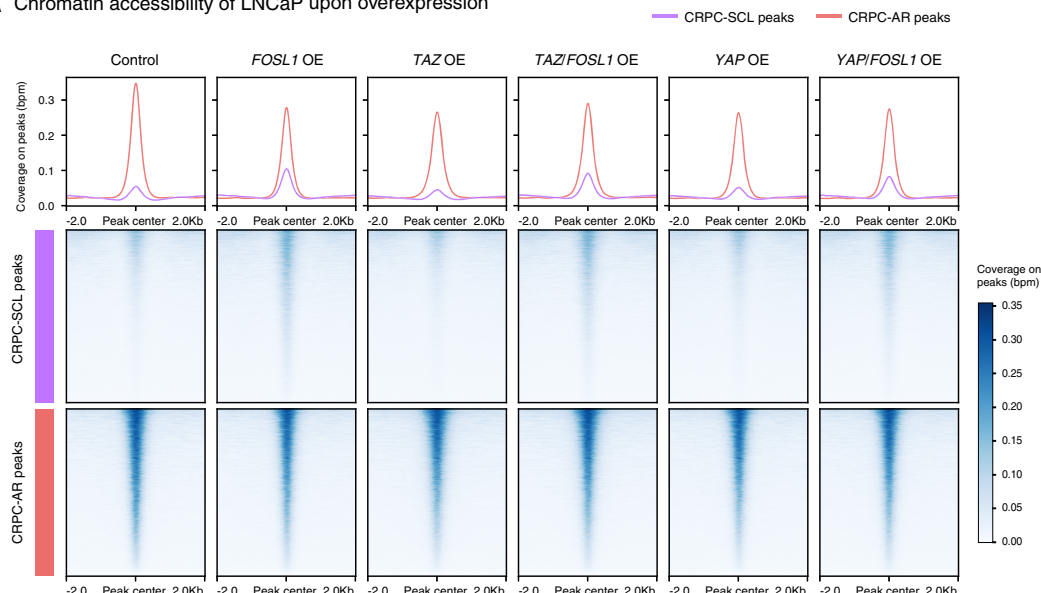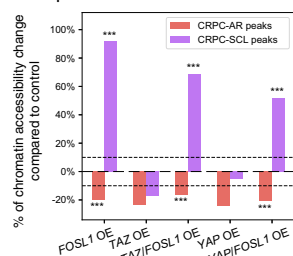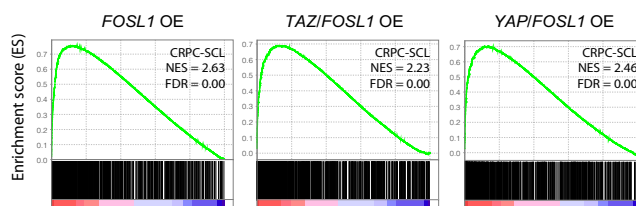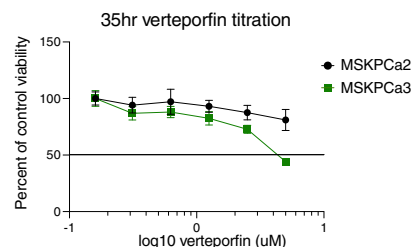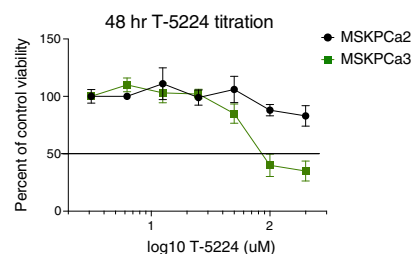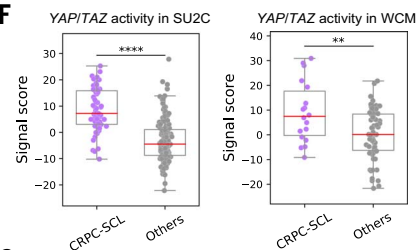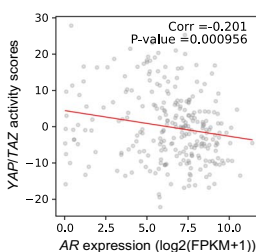
Fig. 7. Evidence of the role of AP-1, YAP, and TAZ in CRPC-SCL from overexpression, small-molecule inhibition, and patient transcriptome findings. (**A**) Chromatin accessibility changes upon *FOSL1*, *TAZ/FOSL1*, *TAZ*, *YAP/FOSL1*, and *YAP* overexpression in LNCaP cells at CRPC-AR and CRPC-SCL peaks. (**B**) Average percentage of chromatin accessibility changes upon *FOSL1*, *TAZ/FOSL1*, *TAZ*, *YAP/FOSL1*, and *YAP* overexpression in LNCaP compared to control at CRPC-AR and CRPC-SCL peaks. ***$P < 0.001$ (permutation test). (**C**) Overexpression of *FOSL1* in LNCaP cells shows up-regulation of CRPC-SCL signature genes. (**D**) Effect of verteporfin on MSKPCa2 (CRPC-AR) and MSKPCa3 (CRPC-SCL) cell growth. (**E**) Effect of T-5224 on MSKPCa2 and MSKPCa3 cell growth. (**F**) YAP/TAZ activity (sum of z-scores) is significantly higher in CRPC-SCL patients. ****$P < 0.0001$, **$P < 0.01$ (one-tailed Wilcoxon rank-sum test, CRPC-SCL compared to the other groups). (**G**) YAP/TAZ activity is negatively correlated with *AR* expression across the 266 SU2C patients with Corr = −0.201 and $P < 0.001$.

of the CRPC-SCL group was validated by knockdown and exogenous overexpression assays followed by ATAC-seq, RNA-seq, and qPCR. Lastly, we tested the effect of two compounds on proliferation and downstream gene expression in both organoids and cell lines.

REFERENCES AND NOTES

1. P. A. Watson, V. K. Arora, C. L. Sawyers, Emerging mechanisms of resistance to androgen receptor inhibitors in prostate cancer. *Nat. Rev. Cancer* **15**, 701–711 (2015). doi: 10.1038/nrc4016; pmid: 26563462
2. Á. Quintanal-Villalonga *et al.*, Lineage plasticity in cancer: A shared pathway of therapeutic resistance. *Nat. Rev. Clin. Oncol.* **17**, 360–371 (2020). doi: 10.1038/s41571-020-0340-z; pmid: 32152485
3. D. S. Rickman, H. Beltran, F. Demichelis, M. A. Rubin, Biology and evolution of poorly differentiated neuroendocrine tumors. *Nat. Med.* **23**, 1–10 (2017). doi: 10.1038/nm.4341; pmid: 28586335
4. H. Beltran *et al.*, The Role of Lineage Plasticity in Prostate Cancer Therapy Resistance. *Clin. Cancer Res.* **25**, 6916–6924 (2019). doi: 10.1158/1078-0432.CCR-19-1423; pmid: 31363002
5. E. G. Bluemn *et al.*, Androgen Receptor Pathway-Independent Prostate Cancer Is Sustained through FGF Signaling. *Cancer Cell* **32**, 474–489.e6 (2017). doi: 10.1016/j.ccell.2017.09.003; pmid: 29017058
6. M. P. Labrecque *et al.*, Molecular profiling stratifies diverse phenotypes of treatment-refractory metastatic castration-resistant prostate cancer. *J. Clin. Invest.* **129**, 4492–4505 (2019). doi: 10.1172/JCI128212; pmid: 31361600
7. D. Gao *et al.*, Organoid cultures derived from patients with advanced prostate cancer. *Cell* **159**, 176–187 (2014). doi: 10.1016/j.cell.2014.08.016; pmid: 25201530
8. L. Puca *et al.*, Patient derived organoids to model rare prostate cancer phenotypes. *Nat. Commun.* **9**, 2404 (2018). doi: 10.1038/s41467-018-04495-z; pmid: 29921838
9. W. Abida *et al.*, Prospective Genomic Profiling of Prostate Cancer Across Disease States Reveals Germline and Somatic Alterations That May Affect Clinical Decision Making. *JCO Precis. Oncol.* **2017**, 1–16 (2017). doi: 10.1200/PO.17.00029; pmid: 28825054
10. D. Robinson *et al.*, Integrative clinical genomics of advanced prostate cancer. *Cell* **161**, 1215–1228 (2015). doi: 10.1016/j.cell.2015.05.001; pmid: 26000489
11. J. W. Park *et al.*, Reprogramming normal human epithelial tissues to a common, lethal neuroendocrine cancer lineage. *Science* **362**, 91–95 (2018). doi: 10.1126/science.aat5749; pmid: 30287662
12. M. R. Corces *et al.*, The chromatin accessibility landscape of primary human cancers. *Science* **362**, eaav1898 (2018). doi: 10.1126/science.aat5749; pmid: 30287662
13. M. R. Corces *et al.*, Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. *Nat. Genet.* **48**, 1193–1203 (2016). doi: 10.1038/ng.3646; pmid: 27526324
14. Cancer Genome Atlas Research Network, The Molecular Taxonomy of Primary Prostate Cancer. *Cell* **163**, 1011–1025 (2015). doi: 10.1016/j.cell.2015.10.025; pmid: 26544944
15. H. Beltran *et al.*, Divergent clonal evolution of castration-resistant neuroendocrine prostate cancer. *Nat. Med.* **22**, 298–305 (2016). doi: 10.1038/nm.4045; pmid: 26855148
16. P. Mu *et al.*, SOX2 promotes lineage plasticity and antiandrogen resistance in TP53- and RB1-deficient prostate cancer. *Science* **355**, 84–88 (2017). doi: 10.1126/science.aah4307; pmid: 28059768
17. B. A. Smith *et al.*, A basal stem cell signature identifies aggressive prostate cancer phenotypes. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E6544–E6552 (2015). doi: 10.1073/pnas.1518007112; pmid: 26460041
18. V. Murillo-Garzón, R. Kypta, WNT signalling in prostate cancer. *Nat. Rev. Urol.* **14**, 683–696 (2017). doi: 10.1038/nrurol.2017.144; pmid: 28895566
19. S. Y. Ku *et al.*, Rb1 and Trp53 cooperate to suppress prostate cancer lineage plasticity, metastasis, and antiandrogen resistance. *Science* **355**, 78–83 (2017). doi: 10.1126/science.aah4199; pmid: 28059767
20. W. S. Chen *et al.*, Genomic Drivers of Poor Prognosis and Enzalutamide Resistance in Metastatic Castration-resistant Prostate Cancer. *Eur. Urol.* **76**, 562–571 (2019). doi: 10.1016/j.eururo.2019.03.020; pmid: 30928160

21. W. Abida *et al.*, Genomic correlates of clinical outcome in advanced prostate cancer. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 11428–11436 (2019). doi: 10.1073/pnas.1902651116; pmid: 31061129
22. M. D. Nyquist *et al.*, Combined TP53 and RB1 Loss Promotes Prostate Cancer Resistance to a Spectrum of Therapeutics and Confers Vulnerability to Replication Stress. *Cell Rep.* **31**, 107669 (2020). doi: 10.1016/j.celrep.2020.107669; pmid: 32460015
23. P. Dhingra *et al.*, Identification of novel prostate cancer drivers using RegNetDriver: A framework for integration of genetic and epigenetic alterations with tissue-specific regulatory network. *Genome Biol.* **18**, 141 (2017). doi: 10.1186/s13059-017-1266-3; pmid: 28750683
24. E. Khurana *et al.*, Role of non-coding sequence variants in cancer. *Nat. Rev. Genet.* **17**, 93–108 (2016). doi: 10.1038/nrg.2015.17; pmid: 26781813
25. Z. Li *et al.*, Identification of transcription factor binding sites using ATAC-seq. *Genome Biol.* **20**, 45 (2019). doi: 10.1186/s13059-019-1642-2; pmid: 30808370
26. E. J. Adams *et al.*, FOXA1 mutations alter pioneering activity, differentiation and prostate cancer phenotypes. *Nature* **571**, 408–412 (2019). doi: 10.1038/s41586-019-1318-9; pmid: 31243370
27. A. Parolia *et al.*, Distinct structural classes of activating FOXA1 alterations in advanced prostate cancer. *Nature* **571**, 413–418 (2019). doi: 10.1038/s41586-019-1347-4; pmid: 31243372
28. C. M. Rudin *et al.*, Molecular subtypes of small cell lung cancer: A synthesis of human and mouse model data. *Nat. Rev. Cancer* **19**, 289–297 (2019). doi: 10.1038/s41568-019-0133-9; pmid: 30926931
29. H. Clevers, Wnt breakers in colon cancer. *Cancer Cell* **5**, 5–6 (2004). doi: 10.1016/S1535-6108(03)00339-8; pmid: 14749120
30. R. Eferl, E. F. Wagner, AP-1: A double-edged sword in tumorigenesis. *Nat. Rev. Cancer* **3**, 859–868 (2003). doi: 10.1038/nrc1209; pmid: 14668816
31. J. Guinney *et al.*, The consensus molecular subtypes of colorectal cancer. *Nat. Med.* **21**, 1350–1356 (2015). doi: 10.1038/nm.3967; pmid: 26457759
32. Y. Hoshida, Nearest template prediction: A single-sample-based flexible class prediction with confidence assessment. *PLOS ONE* **5**, e15543 (2010). doi: 10.1371/journal.pone.0015543; pmid: 21124904
33. S. A. Lambert *et al.*, The Human Transcription Factors. *Cell* **172**, 650–665 (2018). doi: 10.1016/j.cell.2018.01.029; pmid: 29425488
34. S. Piccolo, S. Dupont, M. Cordenonsi, The biology of YAP/TAZ: Hippo signaling and beyond. *Physiol. Rev.* **94**, 1287–1312 (2014). doi: 10.1152/physrev.00005.2014; pmid: 25287865
35. F. Zanconato, M. Cordenonsi, S. Piccolo, YAP/TAZ at the Roots of Cancer. *Cancer Cell* **29**, 783–803 (2016). doi: 10.1016/j.ccell.2016.05.005; pmid: 27300434
36. A. Tsherniak *et al.*, Defining a Cancer Dependency Map. *Cell* **170**, 564–576.e16 (2017). doi: 10.1016/j.cell.2017.06.010; pmid: 28753430
37. Y. Wang *et al.*, Comprehensive Molecular Characterization of the Hippo Signaling Pathway in Cancer. *Cell Rep.* **25**, 1304–1317.e5 (2018). doi: 10.1016/j.celrep.2018.10.001; pmid: 30380420
38. S. Shukla *et al.*, Aberrant Activation of a Gastrointestinal Transcriptional Circuit in Prostate Cancer Mediates Castration Resistance. *Cancer Cell* **32**, 792–806.e7 (2017). doi: 10.1016/j.ccell.2017.10.008; pmid: 29153843
39. G. G. Galli *et al.*, YAP Drives Growth by Controlling Transcriptional Pause Release from Dynamic Enhancers. *Mol. Cell* **60**, 328–337 (2015). doi: 10.1016/j.molcel.2015.09.001; pmid: 26439301
40. F. Zanconato *et al.*, Genome-wide association between YAP/TAZ/TEAD and AP-1 at enhancers drives oncogenic growth. *Nat. Cell Biol.* **17**, 1218–1227 (2015). doi: 10.1038/ncb3216; pmid: 26258633
41. S. Ma *et al.*, Chromatin Potential Identified by Shared Single-Cell Profiling of RNA and Chromatin. *Cell* **183**, 1103–1116.e20 (2020). doi: 10.1016/j.cell.2020.09.056; pmid: 33098772
42. F. Gibault *et al.*, Molecular Features of the YAP Inhibitor Verteporfin: Synthesis of Hexasubstituted Dipyrrins as Potential Inhibitors of YAP/TAZ, the Downstream Effectors of the Hippo Pathway. *ChemMedChem* **12**, 954–961 (2017). doi: 10.1002/cmdc.201700063; pmid: 28334506
43. Y. Aikawa *et al.*, Treatment of arthritis with a selective inhibitor of c-Fos/activator protein-1. *Nat. Biotechnol.* **26**, 817–823 (2008). doi: 10.1038/nbt1412; pmid: 18587386
44. J. Leibold *et al.*, Somatic Tissue Engineering in Mouse Models Reveals an Actionable Role for WNT Pathway Alterations in Prostate Cancer Metastasis. *Cancer Discov.* **10**,

1038–1057 (2020). doi: 10.1158/2159-8290.CD-19-1242; pmid: 32376773
45. P. Isaacsson Velho *et al.*, Wnt-pathway Activating Mutations Are Associated with Resistance to First-line Abiraterone and Enzalutamide in Castration-resistant Prostate Cancer. *Eur. Urol.* **77**, 14–21 (2020). doi: 10.1016/j.eururo.2019.05.032; pmid: 31176623
46. C.-Y. Liu, T. Yu, Y. Huang, L. Cui, W. Hong, ETS (E26 transformation-specific) up-regulation of the transcriptional co-activator TAZ promotes cell migration and metastasis in prostate cancer. *J. Biol. Chem.* **292**, 9420–9430 (2017). doi: 10.1074/jbc.M117.783787; pmid: 28408625
47. L. Zhang *et al.*, The hippo pathway effector YAP regulates motility, invasion, and castration-resistant growth of prostate cancer cells. *Mol. Cell. Biol.* **35**, 1350–1362 (2015). doi: 10.1128/MCB.00102-15; pmid: 25645929
48. H. Zhang *et al.*, Tumor-selective proteotoxicity of verteporfin inhibits colon cancer progression independently of YAP1. *Sci. Signal.* **8**, ra98 (2015). doi: 10.1126/scisignal.aac5418; pmid: 26443705
49. V. R. Dasari *et al.*, Verteporfin exhibits YAP-independent anti-proliferative and cytotoxic effects in endometrial cancer cells. *Oncotarget* **8**, 28628–28640 (2017). doi: 10.18632/oncotarget.15614; pmid: 28404908
50. N. Ye, Y. Ding, C. Wild, Q. Shen, J. Zhou, Small molecule inhibitors targeting activator protein 1 (AP-1). *J. Med. Chem.* **57**, 6930–6948 (2014). doi: 10.1021/jm5004733; pmid: 24831826

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abe1505

Materials and Methods
Supplementary Text
Figs. S1 to S25
Tables S1 to S17
References (51–79)
MDAR Reproducibility Checklist

View/request a protocol for this paper from *Bio-protocol*.

GLACIAL CYCLES

Persistent influence of precession on northern ice sheet variability since the early Pleistocene

Stephen Barker[1]*, Aidan Starr[1], Jeroen van der Lubbe[2], Alice Doughty[3], Gregor Knorr[4], Stephen Conn[1], Sian Lordsmith[1], Lindsey Owen[1], Alexandra Nederbragt[1], Sidney Hemming[5], Ian Hall[1], Leah Levay[6], IODP Exp 361 Shipboard Scientific Party†

Prior to ~1 million years ago (Ma), variations in global ice volume were dominated by changes in obliquity; however, the role of precession remains unresolved. Using a record of North Atlantic ice rafting spanning the past 1.7 million years, we find that the onset of ice rafting within a given glacial cycle (reflecting ice sheet expansion) consistently occurred during times of decreasing obliquity whereas mass ice wasting (ablation) events were consistently tied to minima in precession. Furthermore, our results suggest that the ubiquitous association between precession-driven mass wasting events and glacial termination is a distinct feature of the mid to late Pleistocene. Before then (increasing), obliquity alone was sufficient to end a glacial cycle, before losing its dominant grip on deglaciation with the southward extension of Northern Hemisphere ice sheets since ~1 Ma.

G lacial cycles of the mid to late Pleistocene [approximately the last 0.7 million years (Myr)] were characterized by their long duration [~100 thousand years (kyr)] and relatively abrupt termination (~10 kyr) (1). Their ~100-kyr periodicity has stimulated debate as to which orbital parameters (if any) are most important in driving glacial-interglacial (G-IG) variability, given that direct orbital forcing at this frequency (eccentricity) is negligible (2) (Fig. 1). However, there is growing consensus that both precession and obliquity play a role (at least in glacial termination) through their combined influence on summer insolation across northern high latitudes (3–5). Before the Mid-Pleistocene Transition [MPT; 1.25 to 0.7 million years ago (Ma)], the situation was, at face value, more straightforward; G-IG variability was dominated by ~41-kyr cyclicity, reflecting the near-linear control of ice sheet growth and decay by changes in axial tilt (giving rise to stronger or weaker seasonality) (6). On the other hand, the lack of a clear precession signal in pre-MPT G-IG cyclicity makes little sense because precession plays a substantial role (dependent on the metric employed; Fig. 1) in modulating Northern Hemisphere summer insolation, which is often considered to be the most important factor in the growth and decay of large continental ice sheets (3, 7, 8).

Building on this premise, Raymo, Lisiecki, and Nisancioglu (9) proposed that the "missing" precession signal expected in pre-MPT records of benthic foraminiferal δ18O (a first-order proxy for global ice volume) might be obscured by the equal and opposite effects of Northern versus Southern Hemisphere ice sheet variability on mean ocean δ18O; essentially, the interhemispheric asymmetry of precession effectively cancelled out variations on this time scale in the record of global ice volume, whereas equivalent variations in the obliquity band (which are in-phase between north and south) were amplified. The elegant proposition of (9) is ultimately testable; records of northern (and southern) ice sheet variability should display strong fluctuations on precession time scales.

Searching for precession in early Pleistocene ice sheet variability

To this end we have produced a record of ice rafted debris (IRD) accumulation from NE Atlantic ODP Site 983 (60.4°N, 23.6°W, 1983 m water depth; fig. S1), extending the previous record (10–12) by 500 kyr back to 1.7 Ma (13). The complete record constitutes 9389 samples with an average temporal resolution of 181 years. We employ three independent approaches for age model construction (13): First, we use the age model constructed by Lisiecki and Raymo (14) for ODP Site 983 as part of their benthic δ18O stack (LR04). The LR04 age model was
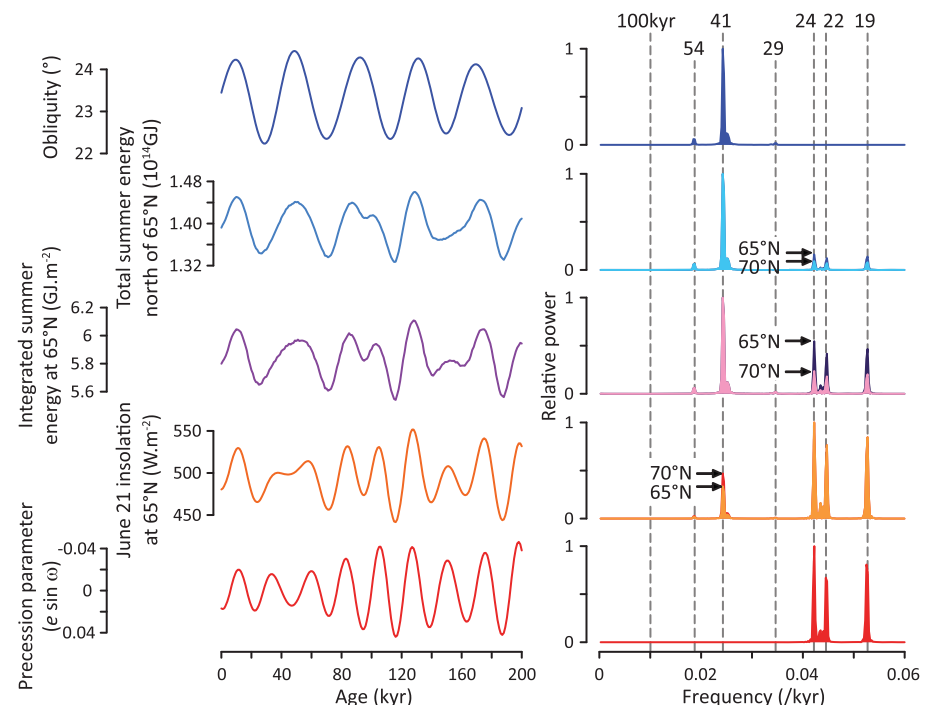
Fig. 1. Influence of obliquity and precession on the spatiotemporal distribution of incoming solar radiation. (Left) Variations in axial tilt (obliquity) and precession (26) with various metrics for assessing changes in northern summer insolation (13). Note that a negative value of the precession parameter is associated with a positive anomaly in Northern Hemisphere summer insolation. (Right) Power spectra for obliquity, precession, and insolation metrics at two different latitudes. All metrics display significant [>99% CL according to the various tests outlined by (18)] power in the obliquity and precession bands with increasing power in the precession band relative to obliquity at lower latitudes. Power spectra are normalized to the maximum power in each case. Vertical dotted lines in the right panel are expected orbital periods in kyr (note that no metric displays substantial power at ~100 kyr).

[1]School of Earth and Environmental Sciences, Cardiff University, Cardiff, UK. [2]Department of Earth Sciences, Vrije Universiteit Amsterdam, Amsterdam, Netherlands. [3]School of Earth and Climate Sciences, University of Maine, Orono, ME, USA. [4]Alfred Wegener Institute, Bremerhaven, Germany. [5]Lamont-Doherty Earth Observatory, Columbia University, New York, NY, USA. [6]International Ocean Discovery Program, Texas A&M University, College Station, TX, USA.
*Corresponding author. Email: barkers3@cf.ac.uk
†IODP Exp 361 Shipboard Scientific Party authors and affiliations are listed in the supplementary materials.
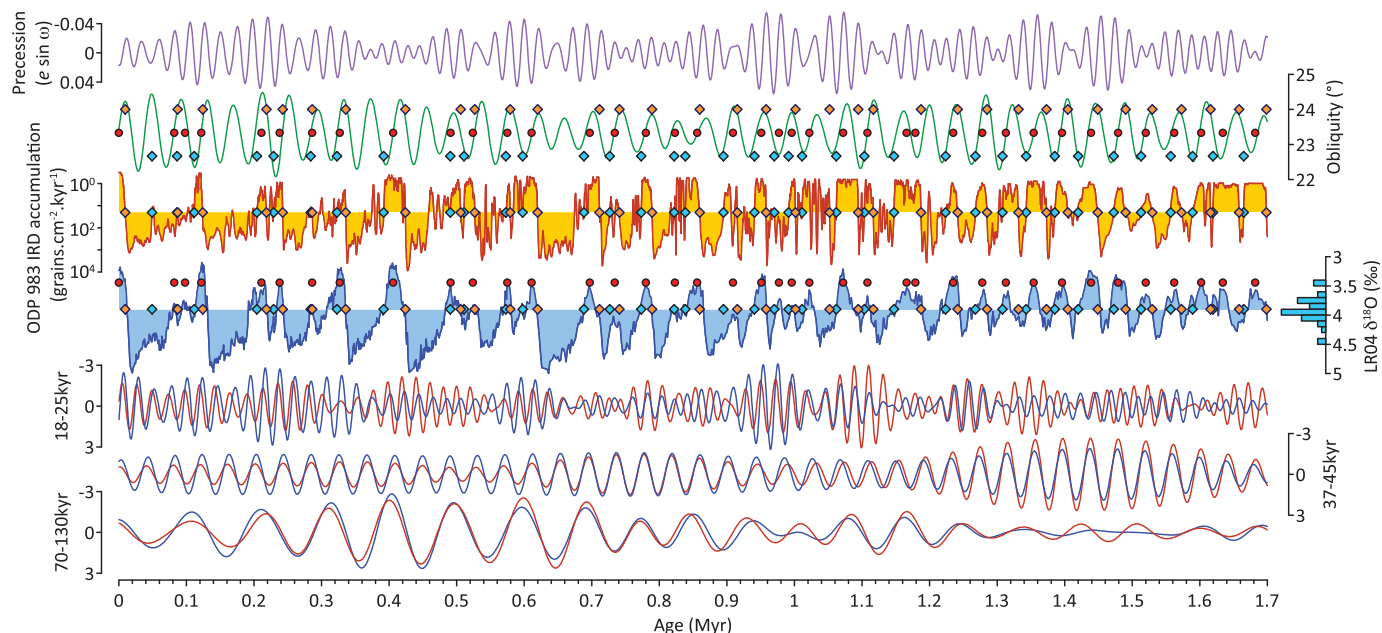
Fig. 2. 1.7 Myr of ice rafting across the NE Atlantic. Red circles represent interglacials (as determined by our algorithm) (*13*), blue diamonds represent onset of significant ice rafting (see orange-filled curve), and orange diamonds represent the end of TIR events (*13*). (Top to bottom) Precession (*26*), obliquity (*26*), IRD accumulation from ODP 983 on the LR04 age model (data have been smoothed and detrended to highlight intervals of significant ice rafting) (*13*), the LR04 benthic stack (histogram represents values of $\delta^{18}O$ at time of each IRD onset, mean = 3.9 ± 0.2‰ as indicated by the horizontal fill threshold of the LR04 curve), 18 to 25 kyr, 37 to 45 kyr, and 70 to 130 kyr filter ouputs of log IRD (red) and the LR04 stack (blue). Note coherence between the LR04 stack and log IRD on G-IG (41 kyr and subsequently ~100 kyr) time scales throughout the past 1.7 Myr (see also fig. S17).

constructed by tuning to a target derived from a simple ice sheet model driven by 21 June insolation at 65°N (a signal dominated by precession; Fig. 1). This implies that any orbital frequencies present in records using the LR04 time scale should be detectable, but also risks introducing frequencies that might not exist in nature. We therefore derive a second age model (U1476pMag; fig. S2) on the basis of the growing body of absolutely dated magnetic polarity reversals and magnetic excursions (*15*), as well as a new record of benthic $\delta^{18}O$ from IODP Site U1476 in the western Indian Ocean (15.8°S, 41.8°E, 2166 m water depth) (*13*). The frequency of age constraints available for this approach is relatively low (table S1), resulting in relatively large (mean 16 kyr 1σ) age uncertainties (fig. S3) (*13*). However, the calculated absolute offset between U1476pMag and LR04 over the past 1.8 Myr averages only 6.6 kyr (maximum 16 kyr), suggesting good agreement between these two completely independent approaches. Finally, we derive a third time scale for our record on the basis of the precession-tuned age model developed for "Shackleton" site IODP U1385 (37.6°N, 10.1°W, 2578 m water depth) (*13, 16*). Although this model is also orbitally tuned, it is independent of LR04 and as such provides a useful comparison (figs. S4 and S5).

The record of IRD accumulation from ODP Site 983 is shown in Fig. 2. IRD accumulation is plotted on a log scale to highlight variations

during periods of relatively low accumulation when ice sheets are small (i.e., interglacial periods). It has been suggested (*9*) that the apparent absence of precession (and dominance of obliquity) frequencies in pre-MPT records of ice rafting from the North Atlantic could be due to the fact that most ice rafting occurs during deglaciation, when rising sea level can destabilize marine-based ice sheets (i.e., iceberg calving rates on North Atlantic marine-based ice margins are controlled primarily by sea level and hence are expected to follow the sea level record even if land-based ice sheets themselves vary on different time scales). Indeed, we typically observe the highest levels of IRD accumulation during terminal events, when sea level is rising (fig. S6) (*13*). This leads to an apparent lag in IRD accumulation (non–log-transformed) behind sea level (in this case the LR04 stack) on G-IG time scales. By contrast, we observe no such lag when comparing log IRD with the LR04 stack (*13*). In fact, we observe coherency between log IRD and the LR04 stack on G-IG (at ~41-kyr and subsequently ~100-kyr) time scales throughout the past 1.7 Myr (Fig. 2 and fig. S17). We suggest that this reflects the fact that following peak interglacial conditions, ice rafting increases as ice sheets expand to form more extensive marine-based margins and we note [using an algorithm to identify the start and end of significant ice rafting during each glacial cycle (*13*)] that the onset

of significant ice rafting tends to occur within a narrow range of benthic $\delta^{18}O$ (3.9 ± 0.2‰; Fig. 2), and continues throughout much of the subsequent glacial period. We therefore conclude that our record of ice rafting (log IRD) reflects variation in the size and/or lateral extent of circum-NE Atlantic ice sheets rather than sea level per se. Specifically, although ice rafting always represents ice sheet ablation (through iceberg calving), we suggest that the increase in ice rafting following an interglacial reflects net growth or extension of ice sheets, whereas the end of ice rafting reflects ice sheet recession before the next interglacial.

We test for the presence of significant [>90% confidence level (CL)] frequencies in our records of IRD and benthic $\delta^{18}O$ (plus the LR04 stack) for different intervals and age models with three methods (*13*) (Fig. 3 and figs. S12 to S14): (i) a multitaper method (MTM) autoregressive [AR(1)] model (*17, 18*), (ii) an alternative spectral noise estimation method (LOWSPEC) (*18*), and (iii) an MTM harmonic F test, which is independent (to the first order) of the first two methods (*18*). We also perform continuous wavelet transforms on each dataset to allow visualization of its temporal evolution (fig. S16). In summary, although we find strong power in the obliquity band of NE Atlantic ice rafting before the MPT (1.7 to 1.25 Ma), we observe no significant (>90% CL) peaks in the precession band before 1.25 Ma. We note that some power in the precession band is observed

with the U1385 age model (which extends back to 1.43 Ma; Fig. 3) but no peaks in this range pass the various significance tests we apply (*13*) (fig. S15). Recent work (*19*) suggests that traditional spectral analysis techniques may be too insensitive for isolating precession frequencies in pre-MPT records but we note that the methods we employ do identify significant (>90% CL) precession-like peaks (albeit with relatively low power) within the LR04 stack and the benthic δ^{18}O record from U1476 even on its absolute (non-orbitally tuned) U1476pMag age model (Fig. 3), which suggests [in agreement with (*19*)] that precession probably played some role in early Pleistocene benthic δ^{18}O variability, either through an influence on ice volume or deep ocean temperature. By contrast, our spectral analysis results suggest that circum-NE Atlantic ice sheets did not fluctuate strongly at precession frequencies before the MPT. Moreover, the accumulation of IRD at our site rarely falls to zero for more than a few thousand years (figs. S8 and S9), suggesting that marine-proximal ice sheet margins existed even during interglacial periods throughout the past 1.7 Myr. Thus, it is unlikely that we are missing a precession signal derived from purely terrestrial ice sheets. In summary, although we cannot rule out the prediction (*9*) that pre-MPT ice sheets varied strongly at precession frequencies, we find little evidence to support it. However, closer visual inspection of the IRD record reveals an alternative picture, as described below.

Ice sheet ablation driven by precession

In Fig. 4, the record of IRD is plotted using its three independent age models, together with the beginning and end of significant ice rafting for each glacial cycle as determined by our algorithm (figs. S7 to S10) (*13*). Note that the end of ice rafting within a given cycle is defined as the end of the latest phase of significant ice rafting lasting 2800 years (*13*) or more before the subsequent interglacial (defined as a minimum in LR04 δ^{18}O with absolute value <4.25‰) (*13*). This is because we wish to identify the major ice rafting events (i.e., mass ablation events) that are characteristic of mid to late Pleistocene glacial terminations (*11, 20, 21*). For our purposes these are termed "terminal ice rafting" (TIR) events, meaning that they represent the last major episode of ice rafting within a glacial cycle, even if (as described below) they do not always coincide with the transition to interglacial conditions.

In Fig. 5, we assess the start and end of ice rafting with respect to the phase of precession and obliquity. We find a consistent link between TIR events and precession over the past 1.7 Myr (table S2), with the end of TIR events occurring, on average, 0.2 ± 2.4 kyr (95% CI, when combining events from the three age model approaches) (*13*) before minima in pre-
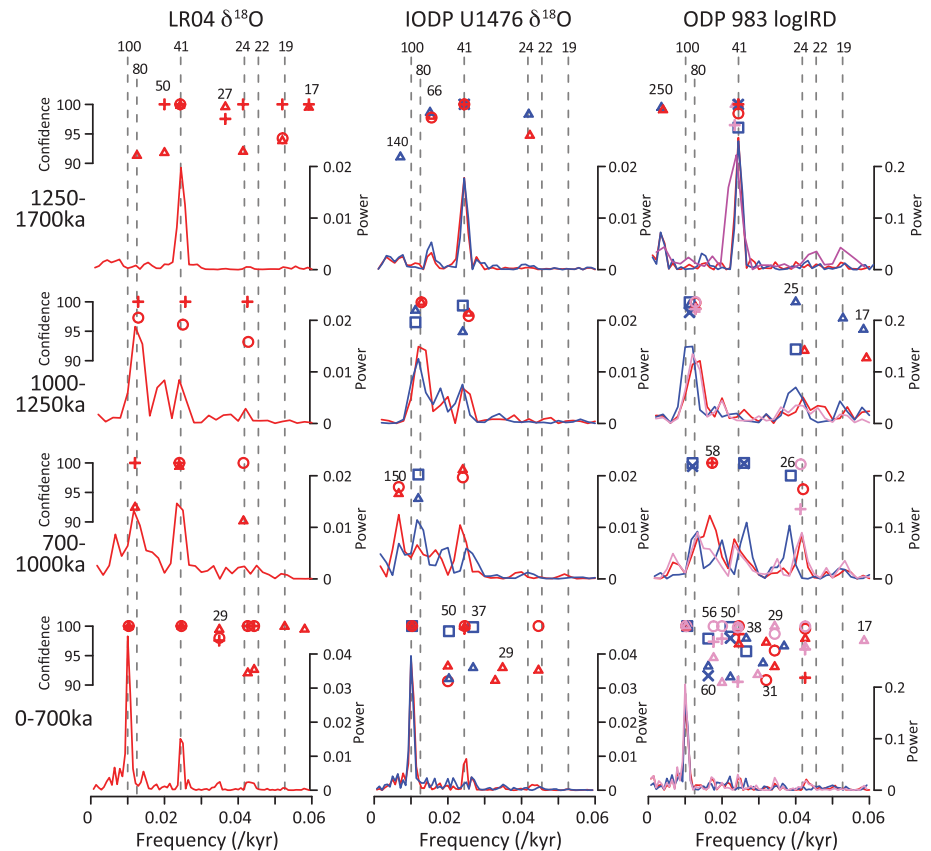
Fig. 3. No significant precession periodicity in pre-MPT ice rafting. Illustrative power spectra and significant spectral peaks identified in LR04, U1476 benthic δ^{18}O, and ODP 983 log IRD for four distinct time windows with three time scales and three different spectral methods. Red, blue, and magenta represent LR04, U1476pMag, and U1385 time scales, respectively. Circles and squares represent results with LOWSPEC (*13, 18*), crosses and exes represent the robust AR(1) method (*17, 18*), and triangles represent the MTM harmonic F test (*13*). No statistically significant (>90% CL) precession-like peaks are identified in the pre-MPT [1250 to 1700 thousand years ago (ka)] section of log IRD although they are present in records of benthic δ^{18}O.

cession prior to 1 Ma and 0.9 ± 2.0 kyr behind precession minima since 1 Ma. Combining all events and age models, we find that TIR events ended, on average, 0.5 ± 1.5 kyr behind minima in precession over the past 1.7 Myr. Because TIR events are (by our definition) 2.8 kyr or more in duration, our findings imply that large-scale ablation of northern ice sheets typically coincided with the increase to maximum summer insolation as a function of precession throughout the past 1.7 Myr. Notably, this conclusion could not be drawn by using traditional spectral analytical techniques, which are designed to detect periodic waveforms rather than the timing of discrete events.

A different relationship is observed with obliquity (table S2). Before 1 Ma, TIR events tended to end while obliquity was increasing (ending on average 8.3 ± 2.4 kyr before a peak in obliquity; Fig. 5) but since 1 Ma, their ending has been more closely aligned with maxima in obliquity (offset = 0.3 ± 3.4 kyr; Fig. 5). On the other hand, deglacial tran-

sitions in benthic δ^{18}O have been closely aligned with maxima in obliquity throughout the past 1.7 Myr (average offset between midpoint of deglaciation and maximum obliquity = 2.3 ± 1.7 kyr; fig. S11) (*13*). The association of post-MPT TIR events with insolation maxima as a function of both obliquity and precession (as described above) provides additional support for a dual role of obliquity and precession in driving mid to late Pleistocene glacial terminations (*4*). However, the situation seems to have been different before that time, as discussed below.

Terminal ice rafting events and deglaciation

Our results suggest that precession has played a persistent role in TIR events since the early Pleistocene. However, before 1 Ma these events were not always aligned with deglacial transitions according to benthic δ^{18}O (Fig. 6). Over the past 1 Myr, the end of TIR events corresponded closely with the latter half of deglaciation as defined by the corresponding decrease in benthic δ^{18}O (Fig. 6C); i.e., mass ice sheet

Fig. 4. Obliquity loses its grip on deglaciation. Red circles represent interglacials (numbered; IG28 is a minimum in δ^{18}O associated with MIS 28) (*12*), white diamonds are deglacial transitions with respect to δ^{18}O (*13*), blue diamonds and vertical dashed lines represent onset of significant ice rafting, and orange diamonds (and lines) represent the end of TIR events. Orange and blue double-headed arrows highlight lengthening of glacial cycles to approximate multiples of the obliquity period following the late occurrence of T17 (see text and Fig. 3, right-hand panel, 1000 to 1250 ka). (Top to bottom) Log (Ca/Ti) from U1385 (*16*), ODP 983 IRD accumulation on a linear scale (note cropped scale), obliquity and precession (*26*), integrated summer energy at 65°N, log IRD from ODP 983 on U1385, U1476pMag, and LR04 age models, benthic foraminiferal δ^{13}C and δ^{18}O from U1476 on its LR04 age model (*13*), the LR04 stack (*14*), and linear IRD on its LR04 age model. Black triangles (bottom) highlight "nonterminating" TIR events (identified as T14.1, T15.1, and so forth). Large black circles highlight shifts toward lighter values of benthic δ^{18}O in U1476 within glacial intervals. Both of these shifts are aligned with coincident features in the records of benthic δ^{13}C and IRD. Note that our algorithm does not assign a TIR event for T13 because IRD accumulation does not subside sufficiently before MIS 27.
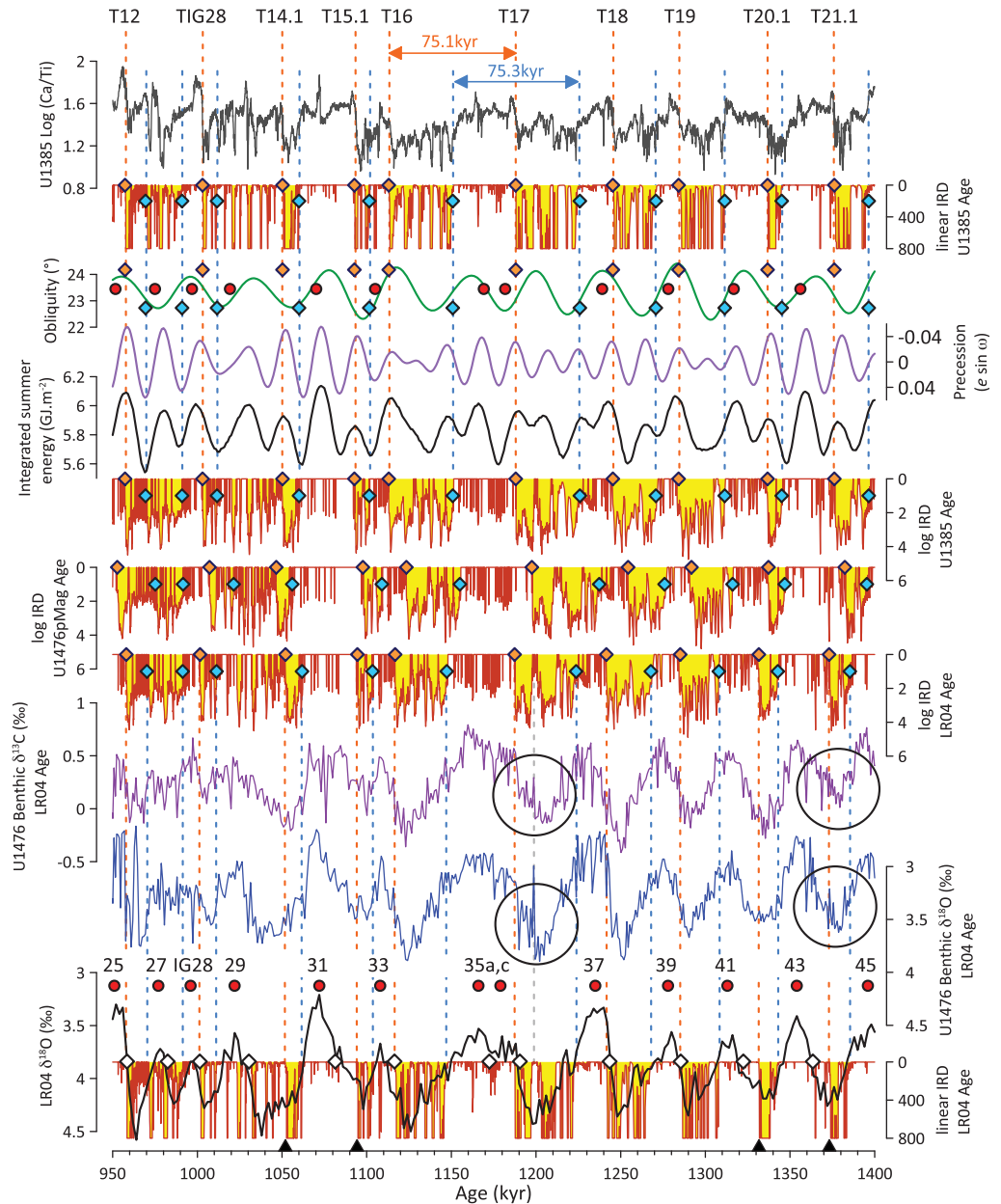
wasting coincided with the transition to interglacial conditions. Before 1 Ma, however, TIR events tended to end earlier with respect not only to obliquity (Fig. 5) but also to deglaciation as constrained by benthic δ^{18}O (Fig. 6C and table S2; note that this result is insensitive to the age model employed). Barker *et al.* (*12*) identified a termination-like event (T14.1 in Fig. 4) ~1.05 Ma within Marine Isotope Stage (MIS) 30. Although it was conceded that T14.1 could not be a true termination in the sense that it did not coincide with decreasing benthic δ^{18}O, the event notably bore other hallmarks of late Pleistocene terminations, including widespread ice rafting followed by abrupt warming across the NE Atlantic. It was also noted that TIR event T14.1 was aligned with a maximum in North-

ern Hemisphere summer insolation which was a function of precession (i.e., a minimum in the precession parameter) and not obliquity.

Our records suggest that this pattern was relatively common prior to 1 Ma. In fact, before that time TIR events commonly ended before the decrease in benthic δ^{18}O typically taken to represent deglaciation (Fig. 6C). Our results therefore suggest that before 1 Ma, TIR events were not synonymous with deglaciation (the transition to interglacial conditions). Furthermore, whether a TIR event coincided (or even overlapped) with deglaciation before 1 Ma depended on the phasing between obliquity and precession.

In Fig. 6D, we plot the temporal offset between the midpoint of each deglaciation and the end of its associated TIR event versus the

offset between the obliquity peak closest to that deglaciation and its nearest minimum in precession. Throughout the past 1.7 Myr, the end of a TIR event tended to coincide (within a few thousand years) with the midpoint of deglaciation when the closest precession minimum occurred prior to the peak in obliquity (a negative obliquity-precession offset in Fig. 6, events ii and iii). On the other hand, since 1 Ma, when the closest precession minimum occurred after the peak in obliquity (a positive obliquity-precession offset in Fig. 6, event i) the TIR event more often than not continued beyond the midpoint of deglaciation (Fig. 6D), presumably driven by the precession minimum directly following the maximum in obliquity (Fig. 6E). Moreover, before 1 Ma, the same positive obliquity-precession
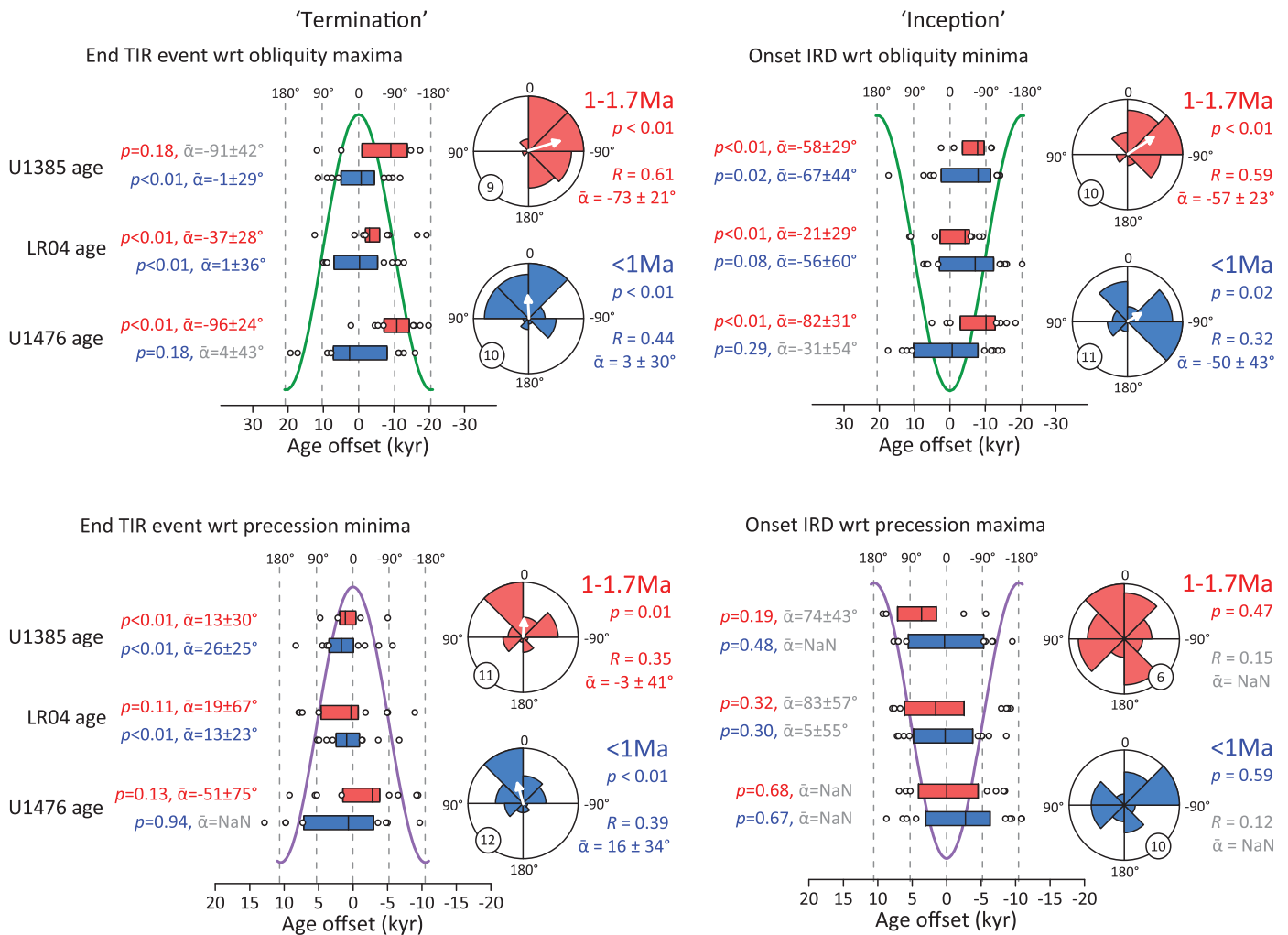
Fig. 5. Precession drives terminal ice rafting events throughout the past 1.7 Myr. Onset (end) of North Atlantic ice rafting across glacial cycles of the past 1.7 Myr with respect to insolation minima (maxima) as a function of obliquity and precession (note precession minima imply insolation maxima and vice versa in the Northern Hemisphere). All curves represent idealized cycles of obliquity (41-kyr period; green) or precession (21-kyr period; purple). Box plots for each age model represent the median and interquartile range; all other data points are shown. Red and blue colors represent pre- and post-1 Ma. Rose diagrams combine results from the three age models (13): lower values of P suggest a higher likelihood of a nonuniform distribution, R is mean resultant vector (R→1 as data converge), and ᾱ is mean direction with 95 or 80% CI (gray text) for P > 0.15 (NaN is returned for P > 0.4). Direction of white arrow, ᾱ; length, R*radial axis. Circled number is the length of the radial axis.

offset could produce a TIR event that ended ~10 to 20 kyr before the midpoint of deglaciation (Fig. 6D, events iv and v), apparently driven by the preceding minimum in precession (Fig. 6E). Thus, since 1 Ma TIR events have always coincided with deglacial transitions, but before 1 Ma a TIR event might start and end before deglaciation had even begun, if the obliquity-precession offset was positive.

The decoupling observed between TIR events and deglaciation before 1 Ma suggests that although obliquity may have been responsible for most deglacial ice sheet ablation prior to the MPT, the most conspicuous ice rafting events (at least across the NE Atlantic) were a result of precession forcing. We suggest this reflects the difference between precession and obliquity in their influence on the spatiotemporal distri-bution of insolation (Fig. 1); although obliquity has a greater effect on integrated summer energy over higher latitudes—which could explain its dominant control on the net waxing and waning of high latitude ice sheets prior to the MPT— precession drives larger changes in the intensity of peak summer warmth at lower latitudes, which could explain why it is so effective at driving massive iceberg calving events along the southern margins of large ice sheets (even if these had little effect on the overall volume of land-based ice before 1 Ma).

Our results therefore imply a change in the influence of obliquity over deglaciation across the MPT. Specifically, before 1 Ma ice sheets were apparently unable to survive a maximum in obliquity such that a subsequent minimum in precession would have little left to melt (in terms of marine-terminating ice sheets) and thus the preceding precession-driven ablation event (which occurred when ice sheets were still large; Fig. 6E) was the last within that cycle. After 1 Ma, sufficient ice apparently remained even after the peak in obliquity associated with deglaciation, such that a subsequent minimum in precession could drive further ablation and ice rafting. In summary, it seems that precession minima will always drive ice ablation events if sufficient ice exists and whether this is the case depends on the influence of obliquity over deglaciation, which weakened across the MPT. Such a change implies the growth of larger Northern Hemisphere ice sheets since ~1 Ma (which required more energy to melt) (5) and/or their net migration toward lower latitudes, where the
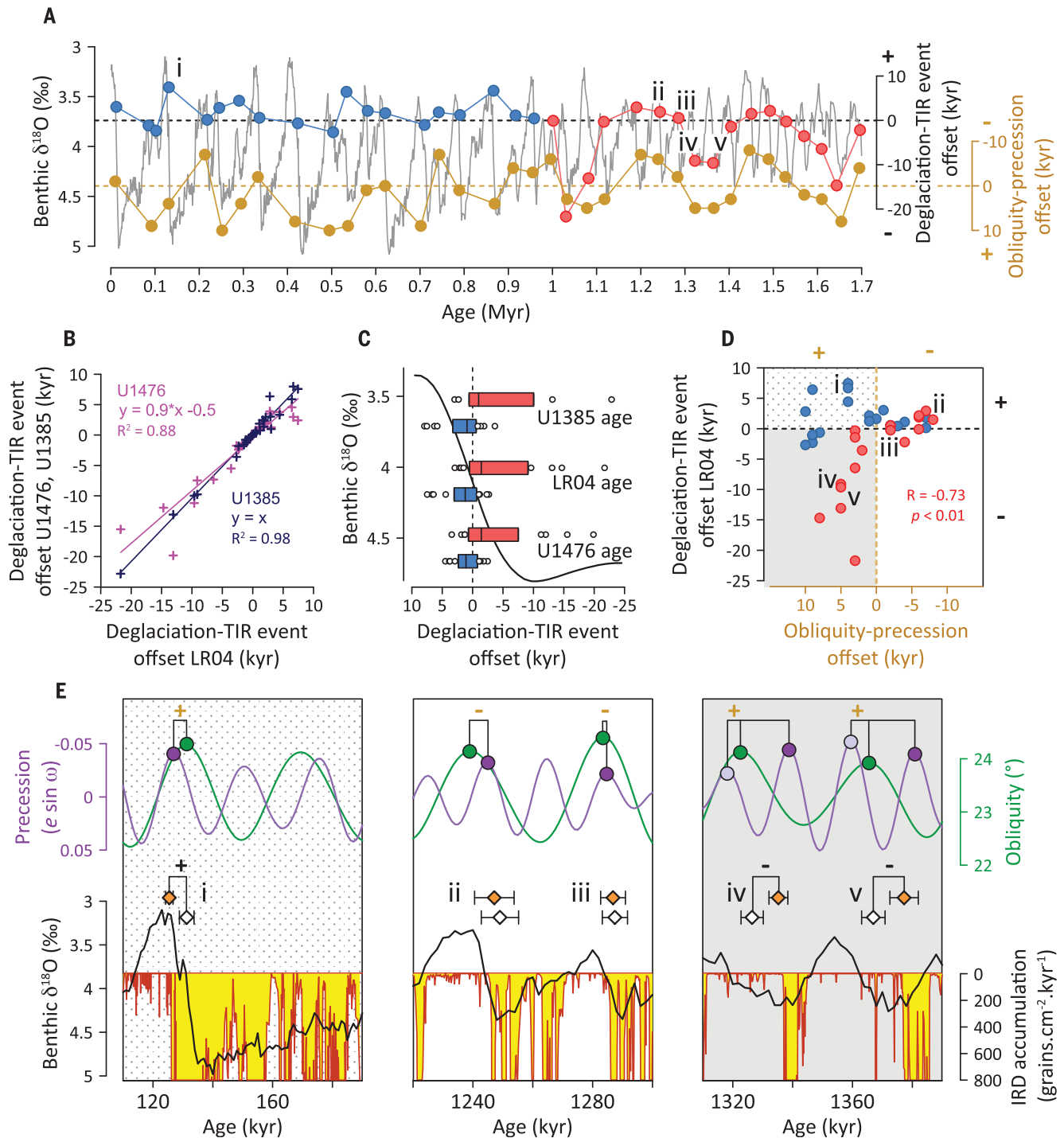
Fig. 6. Changing orbits control the relative timing of terminal ice rafting versus deglaciation. (**A**) Offsets between the end of TIR events and deglacial transitions as a function of benthic $\delta^{18}O$ [blue and red circles; a zero offset means that the end of TIR occurs in parallel with the midpoint of deglaciation—defined as the maximum rate of change in $\delta^{18}O$ (13); a negative offset indicates that TIR ends earlier than the midpoint] compared with offsets between maxima in obliquity and their nearest minima in precession (gold circles; negative offset indicates that the closest precession peak occurs before the peak in obliquity and vice versa). (**B**) Comparison of the three age models for calculating offsets.

(**C**) Box plots representing the calculated offsets before (red) and after (blue) 1 Ma. The black curve is the smoothed LR04 stack across termination 2, for context. (**D**) Calculated offsets before (red) and after (blue) 1 Ma versus the phasing between obliquity and precession [roman numerals refer to (A)]. For events >1 Ma the relationship is significant ($P < 0.005$, $R = -0.73$) (13). (**E**) Examples of the various phasing between obliquity and precession shown in (D) and associated offsets between TIR events and deglaciation. Orange diamonds represent the end of TIR events and white diamonds are deglacial transitions (each data point represents the mean of the three age model results ±1σ).

effects of obliquity are weaker and those of precession stronger (Fig. 1). In turn, this could reflect the proposed net increase in flux of Atlantic surface waters entering the Nordic Seas (the Atlantic inflow) since ~1.2 Ma, promoting enhanced moisture transport and the growth of larger ice sheets with an increasingly southern influence (12).

A notable exception to the "pre-1 Ma" pattern described above occurred ~1.2 Ma (Fig. 4). Although TIR event T17 coincided with the deglacial transition into MIS 35, neither event was associated with a maximum in obliquity (according to two of the three age models employed), occurring instead on the next downward limb of obliquity and coincident with a minimum in precession (~1.19 Ma). Moreover (and according to all three age models) the subsequent minimum in obliquity (~1.18 Ma) was the only minimum in the past 1.7 Myr not associated with the onset or continuation of significant ice rafting (Fig. 2), which we suggest could reflect the late occurrence of T17. Ultimately the result was the appearance of glacial cycles substantially exceeding ~41 kyr in duration (Figs. 3 and 4), marking the weakening grip of obliquity on G-IG variability and the beginning of the MPT.

We suggest that our observation of a direct link between the temporal offset of TIR events versus deglaciation and the phasing of obliquity versus precession (Fig. 6) provides the strongest evidence yet for the influence of precession—in addition to obliquity—on ice sheet variability during the early Pleistocene. This is highlighted by a test for significance of the relationship >1 Ma (Fig. 6D). Using the surrogateCor function in Astrochron (22), we obtained a correlation coefficient of −0.73 ($P < 0.005$) (13), meaning that we can state with confidence that the timing of TIR events versus deglaciation before 1 Ma was related to the phase of precession with respect to obliquity.

Pre-MPT precession signal obscured by obliquity

Based on a detailed study of the past 0.8 Myr, Tzedakis and colleagues (23) concluded that glacial inception (the start of a glacial period) is strongly tied to periods of decreasing obliquity, emphasizing the importance of milder Northern Hemisphere summers (and winters) for the growth and survival of terrestrial ice masses at high latitudes. Our results extend that conclusion to the past 1.7 Myr; the onset of significant ice rafting (implying that ice sheets have grown large enough to develop extensive marine-based margins) is aligned with decreasing to low obliquity in almost all cases (Fig. 5). Conversely, our results do not allow us to determine an unambiguous link between precession and the onset of ice rafting (Fig. 5).

Because decreasing obliquity appears to be critical for glacial inception, this places an upper limit of 1/41 kyr on the frequency of glacial cycles throughout the past 1.7 Myr [a possible exception is the interval of high-frequency precession-like variability (~1 Ma) when three "inception" events occurred across two obliquity cycles; Fig. 2]. Hence, although our results suggest that precession has played a persistent role in the ablation of marine terminating ice sheets since 1.7 Ma, we should not be surprised by the lack of a strong precession signal in the frequency domain of pre-MPT G-IG variability.

Emergence of glacial terminations across the MPT

Global impacts related to TIR events during the mid to late Pleistocene are well documented; for example, abrupt shifts in ocean circulation (11, 20), changing greenhouse gas concentrations (11, 24) and global precipitation patterns (24, 25) are all thought to occur in tandem with major North Atlantic ice rafting during glacial termination. Our results suggest that these impacts may also have been prevalent in the ~41 kyr world. For example, we observe features within the records of benthic foraminiferal $\delta^{18}O$ and $\delta^{13}C$ from IODP Site U1476 that match the timing of major IRD events at Site 983, including some that did not lead directly to deglaciation (Fig. 4). The record of $\delta^{18}O$ from Site U1476 itself displays an excellent correlation with the LR04 stack (13), suggesting that it represents a reasonable reflection of global mean ocean composition. Thus, the commonality between such distal and diverse proxy records (between IRD accumulation in the North Atlantic and proxies for deep ocean chemistry and circulation in the Indian Ocean; fig. S1) attests to the wider relevance of our record from Site 983; i.e., the events recorded are likely more than just local phenomena.

Our results therefore suggest that precession-driven mass ice ablation events have occurred since the early Pleistocene, with related effects on a global scale. On the other hand, their ubiquitous association with deglaciation is a phenomenon unique to the mid to late Pleistocene. Specifically, we suggest that prior to the MPT, obliquity alone was sufficient to end a glacial cycle (as evidenced by the fact that subsequent peaks in precession did not drive further ablation). However, since then, all deglaciations have been associated with precession-driven mass ablation events, which we attribute to the increased importance of precession for melting the larger proportion of Northern Hemisphere ice sheets situated at lower latitudes since the MPT. We therefore recommend that the term glacial termination (sensu stricto) be reserved for deglaciations of the mid to late Pleistocene, which involve the mutual interplay between maxima and minima in obliquity and precession, respectively.

REFERENCES AND NOTES

1. W. S. Broecker, J. van Donk, Rev. Geophys. 8, 169–198 (1970).
2. J. Imbrie et al., Paleoceanogr. Paleoclimatol. 8, 699–735 (1993).
3. M. Milankovitch, Kanon der Erdbestrahlung und seine Andwendung auf das Eiszeiten-problem (Royal Serbian Academy, 1941).
4. P. Huybers, Nature 480, 229–232 (2011).
5. P. C. Tzedakis, M. Crucifix, T. Mitsui, E. W. Wolff, Nature 542, 427–432 (2017).
6. J. Imbrie et al., Paleoceanogr. Paleoclimatol. 7, 701–738 (1992).
7. J. Imbrie, J. Z. Imbrie, Science 207, 943–953 (1980).
8. M. E. Raymo, K. Nisancioglu, Paleoceanogr. Paleoclimatol. 18, 2002PA000791 (2003).
9. M. E. Raymo, L. E. Lisiecki, K. H. Nisancioglu, Science 313, 492–495 (2006).
10. S. Barker et al., Nature 520, 333–336 (2015).
11. S. Barker et al., Paleoceanogr. Paleoclimatol. 34, 1455–1475 (2019).
12. S. Barker, X. Zhang, L. Jonkers, S. Lordsmith, S. Conn, G. Knorr, Paleoceanogr. Paleoclimatol. 36, e2020PA004200 (2021).
13. See supplementary materials.
14. L. E. Lisiecki, M. E. Raymo, Paleoceanogr. Paleoclimatol. 20, PA1003 (2005).
15. J. E. Channell, B. S. Singer, B. R. Jicha, Quat. Sci. Rev. 228, 106114 (2020).
16. D. Hodell et al., Global Planet. Change 133, 49–64 (2015).
17. M. E. Mann, J. M. Lees, Clim. Change 33, 409–445 (1996).
18. S. R. Meyers, Paleoceanogr Paleoclimatol. 27, 2012PA002307 (2012).
19. P. R. Liautaud, D. A. Hodell, P. J. Huybers, Earth Planet. Sci. Lett. 536, 116137 (2020).
20. K. A. Venz, D. A. Hodell, C. Stanton, D. A. Warnke, Paleoceanogr. Paleoclimatol. 14, 42–52 (1999).
21. S. Barker, G. Knorr, Nat. Commun. 12, 2273 (2021).
22. S. R. Meyers, Astrochron: An R Package for Astrochronology (2014); https://CRAN.R-project.org/package=astrochron.
23. P. Tzedakis et al., Clim. Past 8, 1473–1485 (2012).
24. L. Loulergue et al., Nature 453, 383–386 (2008).
25. H. Cheng et al., Nature 534, 640–646 (2016).
26. A. Berger, M. F. Loutre, Quat. Sci. Rev. 10, 297–317 (1991).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abm4033
Materials and Methods
Figs. S1 to S17
Tables S1 and S2
References (27–50)
Data S1
Matlab Files
IODP Exp 361 Shipboard Scientific Party Information

FTO mediates LINE1 m⁶A demethylation and chromatin regulation in mESCs and mouse development

Jiangbo Wei[1,2]†, Xianbin Yu[1,2]†, Lei Yang[3,4]†, Xuelian Liu[3]†, Boyang Gao[1,2], Boxian Huang[5], Xiaoyang Dou[1,2], Jun Liu[1,2]‡, Zhongyu Zou[1,2], Xiao-Long Cui[1,2], Li-Sheng Zhang[1,2], Xingsen Zhao[6,7], Qinzhe Liu[1,2], P. Cody He[1,2], Caraline Sepich-Poore[1,2], Nicole Zhong[1,2], Wenqiang Liu[4], Yanhe Li[4], Xiaochen Kou[3], Yanhong Zhao[3], You Wu[3], Xuejun Cheng[6,7], Chuan Chen[3], Yiming An[3], Xueyang Dong[1,2], Huanyu Wang[1,2], Qiang Shu[6], Ziyang Hao[1,2], Tao Duan[4], Yu-Ying He[8], Xuekun Li[6,7], Shaorong Gao[3,4]*, Yawei Gao[3]*, Chuan He[1,2]*

*N*⁶-methyladenosine (m⁶A) is the most abundant internal modification on mammalian messenger RNA. It is installed by a writer complex and can be reversed by erasers such as the fat mass and obesity-associated protein FTO. Despite extensive research, the primary physiological substrates of FTO in mammalian tissues and development remain elusive. Here, we show that FTO mediates m⁶A demethylation of long-interspersed element-1 (LINE1) RNA in mouse embryonic stem cells (mESCs), regulating LINE1 RNA abundance and the local chromatin state, which in turn modulates the transcription of LINE1-containing genes. FTO-mediated LINE1 RNA m⁶A demethylation also plays regulatory roles in shaping chromatin state and gene expression during mouse oocyte and embryonic development. Our results suggest broad effects of LINE1 RNA m⁶A demethylation by FTO in mammals.

N⁶-methyladenosine (m⁶A) is the most prevalent mammalian mRNA internal modification. It is regulated by writer and eraser proteins, thus affecting transcript fate through reader proteins (*1–3*). The fat mass and obesity-associated protein FTO was the first RNA demethylase shown to remove mRNA m⁶A (*4*). FTO is known to be involved in mammalian development and human diseases; for example, *Fto*⁻/⁻ mice display severe developmental defects (*5, 6*). Extensive functional characterizations in human cancer cells have shown that FTO can localize to the cell cytoplasm and remove m⁶A from mRNA transcripts that contribute to cancer progression (*7–12*); however, similar activity was not apparent across mouse and human tissues,

where FTO tends to exhibit nuclear localization (*13*). Another form of m⁶A, m⁶Aₘ, which is enriched at the cap of a portion of mRNA and certain small nuclear RNAs, was also identified as being a substrate of FTO (*14–16*). However, depletion of the cap-m⁶Aₘ methyltransferase PCIF1 only causes mild cellular effects (*17–19*) and has negligible impacts on mouse viability or fertility (*20*). These discordant findings indicate that the functionally relevant substrates of FTO during mammalian development remain elusive.

Unlike *Pcif1* knockout (KO), *Mettl3* depletion in mice causes early embryonic lethality (*21*). We recently found that chromatin-associated regulatory RNAs (carRNAs) can be m⁶A methylated by METTL3, which regulates chromatin state and transcription in mouse embryonic stem cells (mESCs) (*22*). Independent reports confirmed the chromatin-regulating role of carRNA m⁶A and further showed notable effects of m⁶A on the expression of endogenous retroviruses (ERVs) and heterochromatin formation (*23–25*). Thus, we speculated that a subset of m⁶A-marked carRNAs could be the physiological substrates of FTO and that FTO-mediated m⁶A demethylation may regulate chromatin state in mammalian tissues and during development.

Long-Interspersed Element-1 (LINE1) RNA is a major substrate of FTO in mESCs

To uncover the major substrates of FTO in these contexts, we derived *Fto*⁻/⁻ and control wild-type (WT) mESCs (fig. S1, A to C). We quantified m⁶A-level changes of RNAs isolated from different subcellular fractions between *Fto*⁻/⁻ and WT mESCs (fig. S1D). The m⁶A levels of RNA isolated from chromatin-associated and soluble nuclear fractions were increased (fig. S1, E to G), consistent with the nuclear

localization of FTO (fig. S1H). We performed m⁶A methylated RNA immunoprecipitation sequencing (MeRIP-seq) to examine the chromatin-associated RNA (caRNA) methylome of WT and *Fto*⁻/⁻ mESCs (fig. S1, I to L) and detected more hypermethylated peaks with *Fto* depletion (fig. S1M), accompanied by an increased overall caRNA m⁶A level (fig. S1N).

We annotated carRNAs as promoter-associated RNA, enhancer RNA, and RNA transcribed from transposable elements (repeat RNA) (*22*). Most carRNAs exhibited more hypermethylated m⁶A peaks (fig. S2A) and elevated m⁶A levels (fig. S2B) upon *Fto* KO. Compared with other carRNAs, *Fto* depletion led to more pronounced hypermethylation of repeat RNAs (Fig. 1A), more down-regulated m⁶A-marked repeat RNAs (fig. S2C), and greater down-regulation of hypermethylated repeat RNAs (Fig. 1B). Upon *Fto* KO, carRNA expression changes negatively correlated with m⁶A-level changes (fig. S2D), with repeat RNAs showing the strongest correlation between transcript down-regulation and m⁶A hypermethylation (fig. S2, E and F).

Among m⁶A-marked repeat RNAs, LINE1 RNA emerged as a major substrate of FTO in mESCs. It showed the highest number of hypermethylated peaks, the most increased m⁶A levels, the most reduced abundance (Fig. 1C), and a reduced overall expression (fig. S3, A and B) upon *Fto* KO. LINE1 RNA mainly associates with chromatin in mESCs (fig. S3, C and D) (*26, 27*). We observed colocalization of LINE1 RNA and FTO protein (fig. S3D) and validated the binding of LINE1 RNA by FTO (fig. S3E). The m⁶A level and expression of whole-cell LINE1 RNA exhibited changes similar to those on chromatin upon *Fto* KO (fig. S3, F and G), accompanied by reduced L1ORF1p expression (fig. S3H). Treating WT mESCs with an FTO inhibitor (*9*) recapitulated the effect of *Fto* KO on LINE1 RNA (fig. S3I).

Upon *Fto* KO, we observed that greater m⁶A level increases correlated with greater LINE1 RNA abundance reductions (Fig. 1D). Across published mouse and human tissue m⁶A methylomes (*13*), LINE1 RNA m⁶A level also negatively correlated with its expression, and high FTO expression was associated with low m⁶A level and high LINE1 RNA expression (fig. S4), supporting LINE1 RNA as a substrate of FTO in most tissues.

LINE elements are one of the most abundant retrotransposons in mammalian genomes, and LINE1 RNA is known to play critical roles during mammalian early development (*26, 27*). In mESCs, LINE1 RNA can function as a nuclear RNA scaffold for *trans*-regulation, with LINE1 RNA knockdown by morpholino antisense oligo (ASO) causing activated two-cell (2C) program and repressed ESC-high genes (*27, 28*). *Fto* KO largely recapitulated these transcriptomic changes (Fig. 1E), with

[1]Department of Chemistry, Department of Biochemistry and Molecular Biology, and Institute for Biophysical Dynamics, The University of Chicago, Chicago, IL 60637, USA. [2]Howard Hughes Medical Institute, The University of Chicago, Chicago, Chicago, IL 60637, USA. [3]Institute for Regenerative Medicine, Shanghai East Hospital, Shanghai Key Laboratory of Signaling and Disease Research, Frontier Science Center for Stem Cell Research, School of Life Sciences and Technology, Tongji University, Shanghai 200120, China. [4]Clinical and Translation Research Center of Shanghai First Maternity & Infant Hospital, Shanghai Key Laboratory of Signaling and Disease Research, Frontier Science Center for Stem Cell Research, School of Life Sciences and Technology, Tongji University, Shanghai 200092, China. [5]State Key Laboratory of Reproductive Medicine, Suzhou Affiliated Hospital of Nanjing Medical University, Suzhou Municipal Hospital, Gusu School, Nanjing Medical University, Suzhou 215002, China. [6]The Children's Hospital, School of Medicine, Zhejiang University, National Clinical Research Center for Child Health, Hangzhou 310052, China. [7]The Institute of Translational Medicine, School of Medicine, Zhejiang University, Hangzhou 310029, China. [8]Department of Medicine, Section of Dermatology, University of Chicago, Chicago, IL 60637, USA.
*Corresponding author. Email: chuanhe@uchicago.edu (C.H.); gaoyawei@tongji.edu.cn (Y.G.); gaoshaorong@tongji.edu.cn (S.G.).
†These authors contributed equally to this work.
‡Present address: State Key Laboratory of Protein and Plant Gene Research, School of Life Sciences, Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China.

lower 2C gene activation but greater down-regulation of ESC-high genes (fig. S5, A to E). *Fto* KO also caused several phenotypic changes similar to those which occurred after LINE1 ASO treatment in mESCs, including cell cycle dysregulation, self-renewal impairment, and induction of capacity to form embryoid bodies (EBs) (fig. S5, F to L). Moreover, *Fto*[-/-] mESCs exhibited a reduced ability to integrate to chimeric mice compared with WT mESCs (fig. S5, M and N).

Evolutionarily young LINE1s display higher RNA m6A levels and greater m6A increases upon *Fto* KO (fig. S6, A and B). Longer and less divergent LINE1 RNA also tends to exhibit a higher m6A level (fig. S6, C and D). We observed elevated m6A density near the 5′ end with *Fto* depletion (Fig. 2A and fig. S6E), which also responded to *Mettl3* KO (fig. S6F). Our analysis suggested that several young LINE1 subfamilies were down-regulated and hypermethylated upon *Fto* KO (Fig. 2B and fig. S6, G to J). We used a CRISPR dCas13b system fused with WT FTO or a catalytically inactive mutant (fig. S6K) (*22*), and observed that delivery of dCas13b-*wt*FTO by guide RNA targeting LINE1 RNA reversed its m6A level and expression changes (fig. S6, L to N).

Consistent with the reported YTHDC1-mediated destabilization of m6A-marked carRNAs (*22*), we detected accelerated decay of LINE1 RNA upon *Fto* KO (Fig. 2C), accompanied by elevated binding of LINE1 RNA by YTHDC1 (Fig. 2D). Both changes were recovered by targeting LINE1 RNA with dCas13b-*wt*FTO in *Fto*[-/-] mESCs (fig. S7, A and B).

LINE1 transcription was markedly reduced with *Fto* depletion (Fig. 2E and fig. S7, C to E). Moreover, *Fto* KO led to greater decreases in the nascent synthesis of m6A-marked compared with unmarked LINE1 RNAs (Fig. 2F), but not ERVK or Alu transcripts (fig. S7F). We also observed reduced DNA association of LINE1 RNA and decreased R-loop formation around LINE1 loci with *Fto* depletion (fig. S7, G and H). These effects caused by *Fto* KO could all be reversed by targeting dCas13b-*wt*FTO to LINE1 RNA (fig. S7, I to K). Therefore, FTO appears to mediate m6A demethylation of a subset of LINE1 RNAs, maintaining their levels on chromatin.

Fto KO leads to closed chromatin in mESCs

LINE1 RNA and m6A on carRNAs have been shown to regulate chromatin state and transcription (*22–27*). Indeed, we observed decreased nascent RNA synthesis (fig. S8, A and B) accompanied by more closed chromatin (Fig. 3A) upon *Fto* KO. Similar effects were observed when treating WT mESCs with an FTO inhibitor (fig. S8, C and D). Additionally, LINE1 ASO treatment in WT mESCs led to more closed chromatin (fig. S8E), whereas delivering dCas13b-*wt*FTO to LINE1 RNA largely res-

cued chromatin closure observed in *Fto*[-/-] mESCs (Fig. 3B and fig. S8F). We next monitored LINE1 RNA m6A demethylation and chromatin accessibility in a time-course of FTO inhibition in WT mESCs (fig. S8, G to J). Assay for transposase-accessible chromatin sequencing (ATAC-seq) results also validated reduced chromatin accessibility upon *Fto* KO (fig. S9, A to C), with gained-closed regions enriching gene ontology (GO) terms relevant to development (fig. S9D). Slightly decreased H3K4me3 and H3K27ac and increased H3K9me3 and H3K27me3 were observed with *Fto* depletion (fig. S9E).

YY1 and EP300 can be recruited by carRNA to promote transcription (*22, 29, 30*). We found notable enrichment at gained-closed regions caused by *Fto* KO for H3K4me1, H3K4me3, and H3K27ac, as well as YY1, EP300, and Pol II binding, but not repressive histone marks (fig. S10A). Consistently, chromatin immunoprecipitation sequencing experiments confirmed reduced chromatin accessibility of these regions upon *Fto* depletion (fig. S10B). *Fto* KO–induced gained-closed regions were also affected by *Mettl3* KO (fig. S10C) (*22*). In turn,

chromatin closure upon *Fto* KO could reduce access of METTL3 to certain genomic regions, potentially explaining m6A hypomethylation within them.

We observed distinct profiles of H3K4me3 and H3K9me3 between loci with m6A-marked and -unmarked LINE1 RNA (Fig. 3C). At m6A-marked LINE1 RNA loci, we observed decreased H3K4me3 and H3K27ac levels and Pol II binding upon *Fto* KO, accompanied by increased H3K9me3 levels (Fig. 3C and fig. S11, A to C). Similar patterns were observed for FTO-targeted LINE1 subfamilies but not *L1M2b*, an unmarked LINE1 subfamily, or *IAPEz-int*, an ERVK subfamily regulated by METTL3 (*23–25*) (fig. S11D). Targeting LINE1 RNA with dCas13b-*wt*FTO reversed the dysregulated histone marks (fig. S11E). Moreover, *Fto* KO led to decreased chromatin association and LINE1 RNA binding of YY1 and EP300 (fig. S11, F to H).

Altogether, our data reveal an interplay between LINE1 RNA m6A demethylation by FTO and chromatin state. After *Fto* depletion, increased m6A on LINE1 RNA could promote
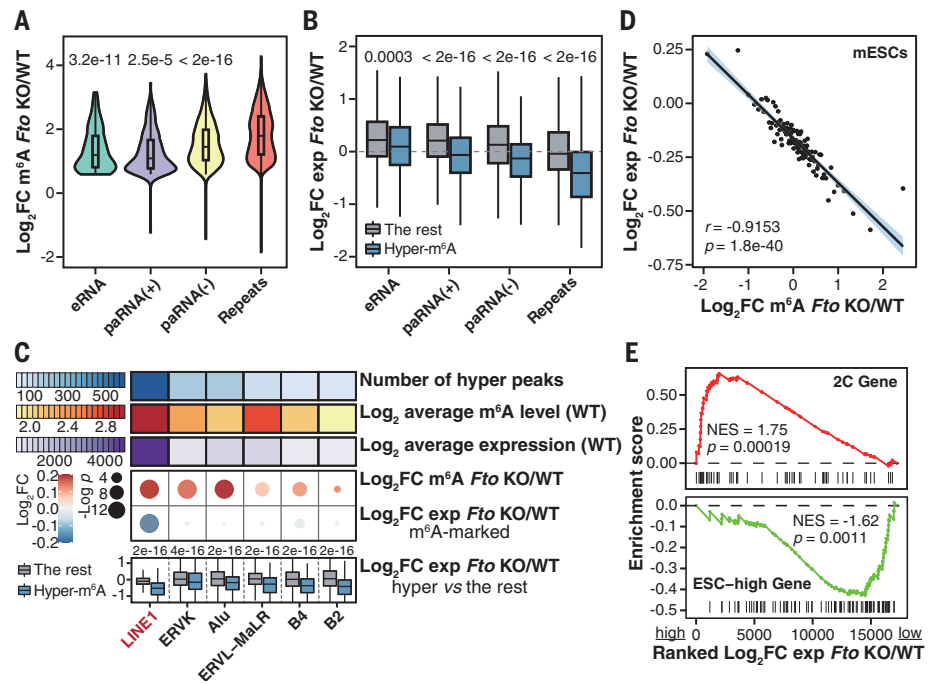
Fig. 1. m6A on LINE1 RNA is a major substrate of FTO in mESCs. (**A**) Violin plots showing m6A-level fold changes of hypermethylated m6A peaks on carRNAs upon *Fto* KO. *P* values were determined using Wilcoxon rank sum tests. (**B**) Boxplots showing expression fold changes of hypermethylated carRNAs versus other m6A-marked carRNAs upon *Fto* KO. *P* values were determined using Wilcoxon signed-rank tests. (**C**) Summary of repeat RNAs on chromatin upon *Fto* KO. Top: number of hypermethylated peaks, average m6A level, and expression. Middle: m6A-level and expression fold changes. *P* values were determined using Wilcoxon signed-rank tests. Bottom: expression fold changes of hypermethylated versus non-hypermethylated repeat RNAs. *P* values were determined using Wilcoxon rank sum tests. (**D**) Scatter plot showing the negative correlation of fold changes between m6A level and expression of LINE1 RNA upon *Fto* KO. LINE1 RNAs were categorized into 100 bins on the basis of their ranked m6A-level fold changes upon *Fto* KO. *r* refers to Pearson's correlation coefficient. *P* value was determined using *t* distribution. (**E**) Gene set enrichment analysis showing up-regulated 2C genes (top) and down-regulated ESC-high genes (bottom) from mRNA-seq upon *Fto* KO. NES, normalized enrichment score.

YTHDC1 binding, which could recruit nuclear decay machinery (*22*) and histone modifiers that add repressive histone marks (*23*, *24*), resulting in reduced LINE1 RNA abundance and more closed chromatin (Fig. 3D).

FTO regulates gene expression mostly through intragenic LINE1 RNA

We observed that *Fto* and LINE1 RNA abundance increased after EB differentiation induction, but LINE1 RNA increase was diminished upon *Fto* KO (fig. S12A). Targeting dCas13b-*wt*FTO to LINE1 RNA in *Fto*[−/−] mESCs partially rescued aberrantly expressed differentiation markers associated with induced EB differentiation and impaired self-renewal (fig. S12, B to G).

Therefore, we investigated the mechanism underlying the impact of *Fto* KO on mESC differentiation. We found that genes containing LINE1 RNA were down-regulated upon *Fto* KO compared with genes beyond 10 kb from LINE1 RNA (Fig. 4A and fig. S13A) (*26*). The expression of LINE1-containing genes and intragenic LINE1 RNA tended to decrease with *Fto* depletion, accompanied by increased intragenic LINE RNA m⁶A level (Fig. 4B and fig. S13B). LINE1-containing genes also displayed greater transcription rate reductions compared with other genes upon *Fto* KO (fig. S13, C to E).

Both intragenic and intergenic LINE1 RNAs showed suppressed expression upon *Fto* KO, but intragenic RNAs were generally more

Fig. 2. FTO regulates LINE1 RNA level through m⁶A demethylation. (**A**) m⁶A peak profiles on m⁶A-marked LINE1 RNA in WT and *Fto*[−/−] mESCs. (**B**) Volcano plot showing differentially expressed subfamilies of m⁶A-marked repeat RNAs upon *Fto* KO. (**C**) Cumulative distribution and boxplots (inset) showing nuclear LINE1 RNA lifetime in WT and *Fto*[−/−] mESCs. *P* value was determined using Wilcoxon signed-rank test. (**D**) Relative enrichment of LINE1 RNA by YTHDC1 from cross-linking and immunoprecipitation quantitative polymerase chain reaction (qPCR) in WT and *Fto*[−/−] mESCs. *P* values were determined using unpaired two-tailed *t* tests. Error bars and means ± SD are shown for *n* = 3 experiments. (**E**) Cumulative distribution and boxplots (inset) showing LINE1 RNA transcription rate in WT and *Fto*[−/−] mESCs. *P* value was determined using Wilcoxon signed-rank test. (**F**) Cumulative distribution and boxplots (inset) showing the difference in transcription rate changes between m⁶A-marked and -unmarked LINE1 RNA upon *Fto* KO. *P* value was determined using Wilcoxon rank sum test.
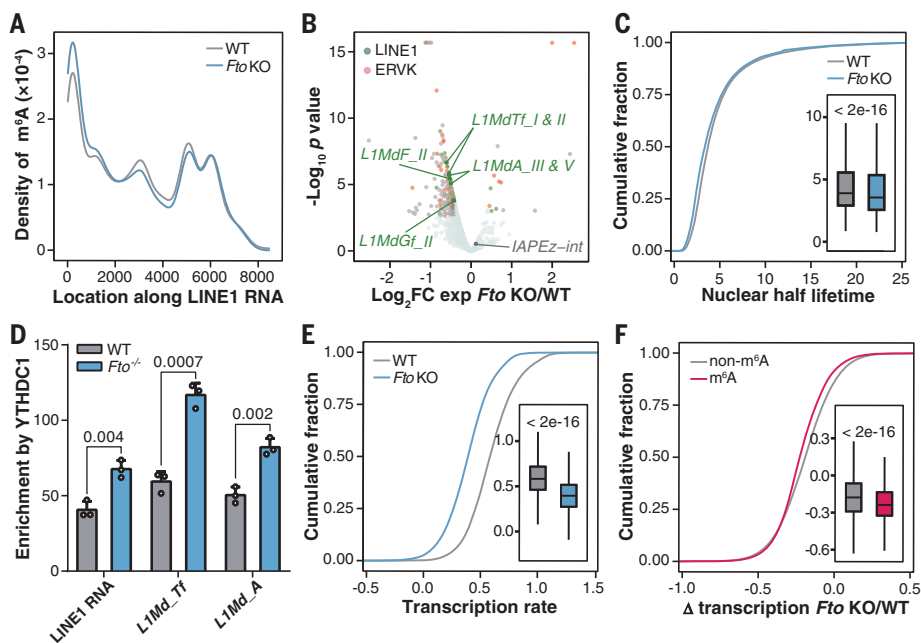
Fig. 3. FTO affects chromatin state through LINE1 RNA m⁶A demethylation. (**A**) Left: DNase I–terminal deoxynucleotidyl transferase–mediated deoxyuridine triphosphate nick end labeling (TUNEL) assay showing more closed chromatin upon *Fto* KO. Scale bars, 20 μm. The nucleus was counterstained with 4′,6-diamidino-2-phenylindole (DAPI). Representative images were selected from three independent experiments. Right: relative TUNEL intensity in WT and *Fto*[−/−] mESCs. *P* value was determined using an unpaired two-tailed *t* test. TUNEL intensity was quantified by ImageJ. (**B**) TUNEL signals of WT (gray) and *Fto*[−/−] mESCs rescued by dCas13b-*wt*FTO with guide RNA (gRNA) targeting LINE1 RNA (red) or control gRNA (blue). (**C**) Profiles of H3K4me3 and H3K9me3 levels on LINE1 RNA loci from 3.0 kb upstream of the transcription start site (TSS) to 3.0 kb downstream of the transcription end site (TES) in WT and *Fto*[−/−] mESCs. (**D**) Schematic model showing how *Fto* KO affects LINE1 RNA abundance and local chromatin state.
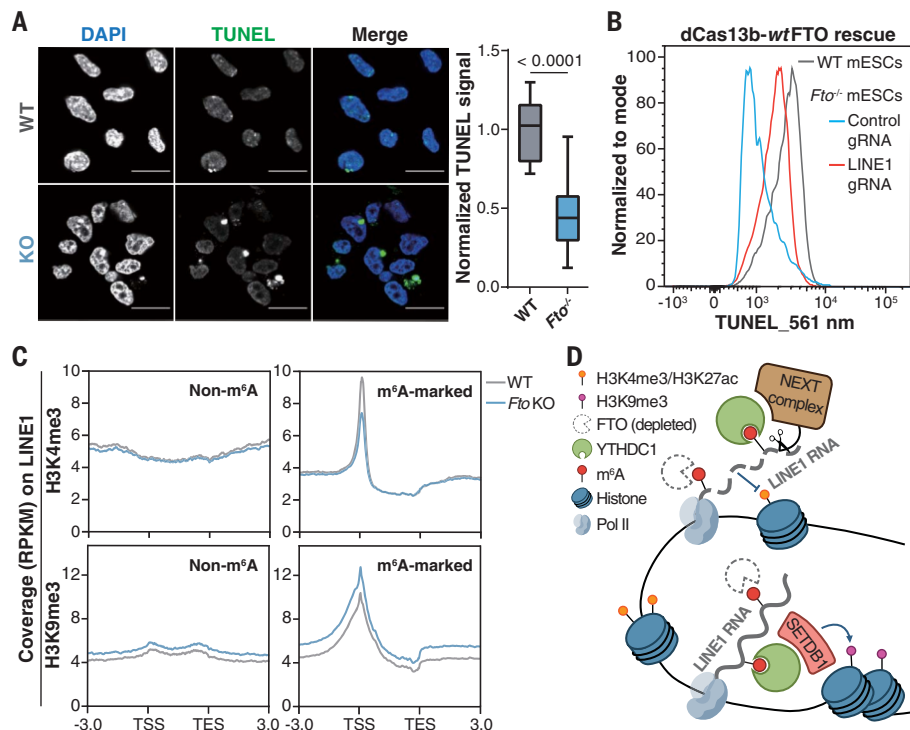
responsive (fig. S13, F and G). We further explored the *cis*-regulatory role of intragenic LINE1 RNAs suppressed by *Fto* KO and observed significantly decreased expression and reduced transcription rate for genes containing down-regulated LINE1 RNA compared with other genes (Fig. 4C). Accordingly, reduced H3K4me3 and H3K27ac levels and Pol II binding, accompanied by elevated H3K9me3 levels, were observed for these gene loci (Fig. 4D) and those containing FTO-targeted LINE1 subfamilies (fig. S13H).

Most genes containing down-regulated LINE1 RNA were also down-regulated upon *Fto* KO (fig. S14A), and these genes enriched pathways involved in differentiation and development (fig. S14B). We examined several key genes in ESC pluripotency and early development (*Esrrb*, *Nek5*, *Phf3*, and *Zfp982*) and found that all of them contained intragenic LINE1 RNA with increased m^6A level and decreased expression upon *Fto* KO, associated with more closed local chromatin and reduced gene transcription rate (fig. S14, C to G), which could all be reversed by targeting LINE1 RNA with dCas13b-*wt*FTO (fig. S15, A to I).

Certain 2C genes, such as *Dub1* and *Zscan4*, do not contain LINE1 RNA. These genes and MERVL RNA, a 2C-specific transposon, could be repressed by LINE1 RNA through *trans*-regulation (27). For these transcripts, we observed reduced expression when applying dCas13b-*wt*FTO to LINE1 RNA in *Fto*$^{-/-}$ mESCs (fig. S15J).

The FTO-LINE1 RNA axis is functionally relevant in mouse development

We examined the mouse cerebellum, hippocampus, and adult neural stem cells because

Fto is highly expressed in mouse brain tissues (*31*). We observed increased LINE1 RNA m^6A levels, decreased LINE1 RNA expressions, and more closed chromatin from samples derived from *Fto*$^{-/-}$ mice compared with WT controls (fig. S16, A to J). Similar trends of *Fto* and LINE1 RNA abundance during mESC differentiation were also observed for adult neural stem cells (fig. S16, K and L).

We next investigated the effects of FTO during early development. *Fto*$^{-/-}$ pups were born at a lower rate than the expected Mendelian ratio (fig. S17A). *Fto*$^{-/-}$ female mice showed ovarian defects and impaired fertility (fig. S17, B and C). Previous studies revealed that LINE1 is activated before the sex determination of primordial germ cells and the meiotic entry of oocytes (*32*, *33*). Indeed, both *Fto* and LINE1 RNA expression decreased from primordial germ cell to the metaphase II (MII) stage (fig. S17D). The number of germinal vesicle (GV) oocytes and the ratio of surrounded nucleolus GV oocytes from *Fto*$^{-/-}$ mice were lower compared with WT controls (Fig. 5, A and B). *Fto* depletion led to significantly reduced LINE1 RNA expression in GV and MII oocytes (Fig. 5C and fig. S17E), with down-regulated LINE1 subfamilies largely resembling those observed in mESCs (fig. S17F). We observed more closed chromatin for GV oocytes from *Fto*$^{-/-}$ mice compared with WT controls (Fig. 5D). *Fto*$^{-/-}$ GV oocytes could mature to the MII stage (fig. S17G), but *Fto*$^{-/-}$ MII oocytes showed increased chromosome misalignment and spindle collapse (fig. S17H). RNA-sequencing (RNA-seq) results revealed greater expression decreases of genes containing down-regulated LINE1 RNA compared with other genes in *Fto*$^{-/-}$ GV and MII oocytes,

respectively (fig. S17, I and J). GO analysis suggested that *Fto* KO caused the observed defects in oocyte development, likely through LINE1 RNA, which regulates LINE1-containing genes (fig. S17K).

We further generated WT (*Fto*$^{P+/M+}$), maternal *Fto*-deficient (*Fto*$^{P+/M-}$), paternal *Fto*-deficient (*Fto*$^{P-/M+}$), and homozygous KO (*Fto*$^{P-/M-}$) embryos (fig. S18, A and B). Embryos from all four groups could reach the blastocyst stage [embryonic day 3.5 (E3.5)] and hatch out of the zona pellucida at E4.5, but *Fto*-deficient embryos showed a slightly weakened ability to do so (fig. S18C). Moreover, maternal loss of *Fto* severely impeded decidua formation and the generation of E7.5 embryos, with no E7.5 embryos produced upon homozygous *Fto* depletion (Fig. 5E and fig. S18D). We examined the transcriptome differences between *Fto*$^{P-/M-}$ and *Fto*$^{P+/M+}$ morulae, in which we detected repressed LINE1 RNA and down-regulated LINE1 subfamilies similar to those in mESCs (Fig. 5F and fig. S18, E and F). We again observed induced expression of *Zscan4* and MERVL RNA (fig. S18, G and H) and greater expression decreases of genes containing down-regulated LINE1 RNA, including regulators essential during early embryonic development, such as *Lin28b*, *Tet2*, and *Gsk3b* (fig. S18, I to L).

Finally, we applied dCas13b-*wt*FTO in *Fto*$^{-/-}$ MII oocytes and fertilized them with *Fto*$^{-/-}$ sperm. Embryos were developed in vitro to the morula stage for subsequent analyses (fig. S18, M and N). Induced LINE1 RNA associated with more open chromatin was observed in *Fto*$^{P-/M-}$ morulae by targeting dCas13b-*wt*FTO to LINE1 RNA in *Fto*$^{-/-}$ MII oocytes (fig. S18, O and P). The expression of *Zscan4*, MERVL RNA, and selected genes containing down-regulated LINE1

Fig. 4. *Fto* KO deactivates LINE1-containing genes by repressing intragenic LINE1 RNA. (**A**) Boxplots showing gene expression fold changes upon *Fto* KO from caRNA sequencing. Genes were categorized according to their genomic distance to the nearest LINE1 RNA with at least 10 reads. *P* value was determined using Wilcoxon rank sum test. (**B**) Scatter plot showing the negative correlation of fold changes between expression of LINE1-containing genes and m^6A level of corresponding intragenic LINE1 RNA upon *Fto* KO. Intragenic LINE1 RNAs were categorized into 100 bins on the basis of their ranked expression fold changes upon *Fto* KO. *r* refers to Pearson's correlation coefficient. *P* value was calculated based on *t* distribution. (**C**) Boxplots showing fold changes of gene expression from caRNA sequencing (left)

and gene transcription rate (right) upon *Fto* KO. Genes were categorized into three groups: genes containing down-regulated LINE1 RNA, genes near (<1 Mb) down-regulated LINE1 RNA, and genes containing other LINE1 RNA. *P* values were determined using Wilcoxon rank sum tests. (**D**) Profiles of H3K4me3, H3K9me3, H3K27ac, and Pol II levels on loci of genes containing down-regulated LINE RNA from 3.0 kb upstream of the TSS and 3.0 kb downstream of the TES in WT and *Fto*$^{-/-}$ mESCs.
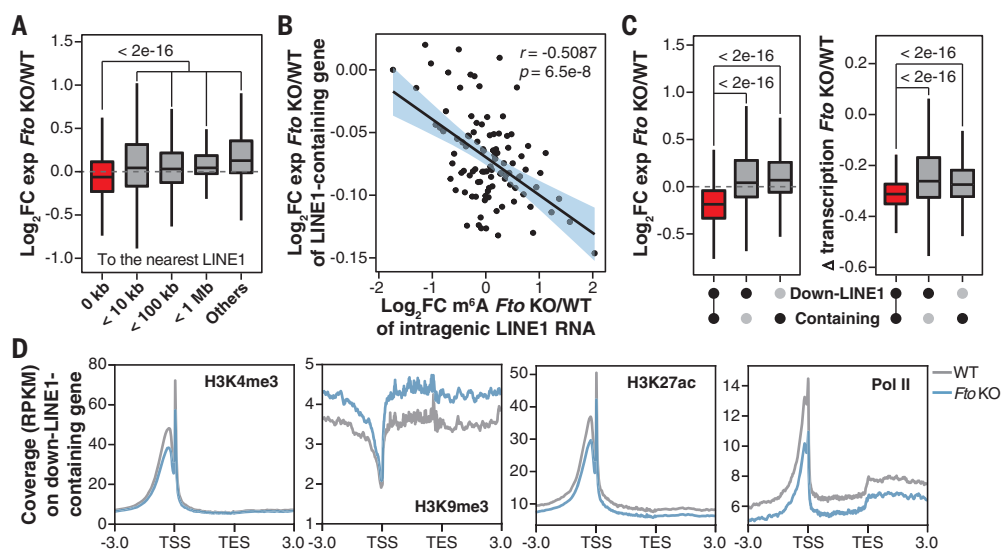
Fig. 5. The FTO-LINE1 RNA axis plays critical roles during early development.
(**A**) Number of GV oocytes from 4-week-old WT and $Fto^{-/-}$ mice. Error bars and means ± SEM are shown for n = 4 WT mice and n = 6 $Fto^{-/-}$ mice. (**B**) Ratio of surrounded nucleolus (SN)/nonsurrounded nucleolus (NSN) oocytes from 4-week-old WT mice (n = 5) and $Fto^{-/-}$ mice (n = 4). (**C**) Relative LINE1 RNA expression measured by reverse transcription qPCR in WT and $Fto^{-/-}$ oocytes. (**D**) Left: DNase I-TUNEL assay showing more closed chromatin in oocytes upon Fto KO. Scale bars, 50 μm. The nucleus was counterstained by DAPI. Representative images were selected from three independent experiments. Right: relative TUNEL intensity in WT and $Fto^{-/-}$ oocytes (n = 12 each). pSN, partly surrounded nucleolus. (**E**) Implantation rate (left) and E7.5 embryo rate (right) of $Fto^{P+/M+}$, $Fto^{P+/M-}$, $Fto^{P-/M+}$, and $Fto^{P-/M-}$ zygotes. (**F**) Relative LINE1 RNA expression measured by reverse transcription qPCR in $Fto^{P+/M+}$ and $Fto^{P-/M-}$ morulae. For (A), (C), (D), and (F), P values were determined using unpaired two-tailed t tests; error bars and means ± SD are shown for n = 3 experiments in (C) and (F). For (B) and (E), P values were determined using two-tailed z tests.
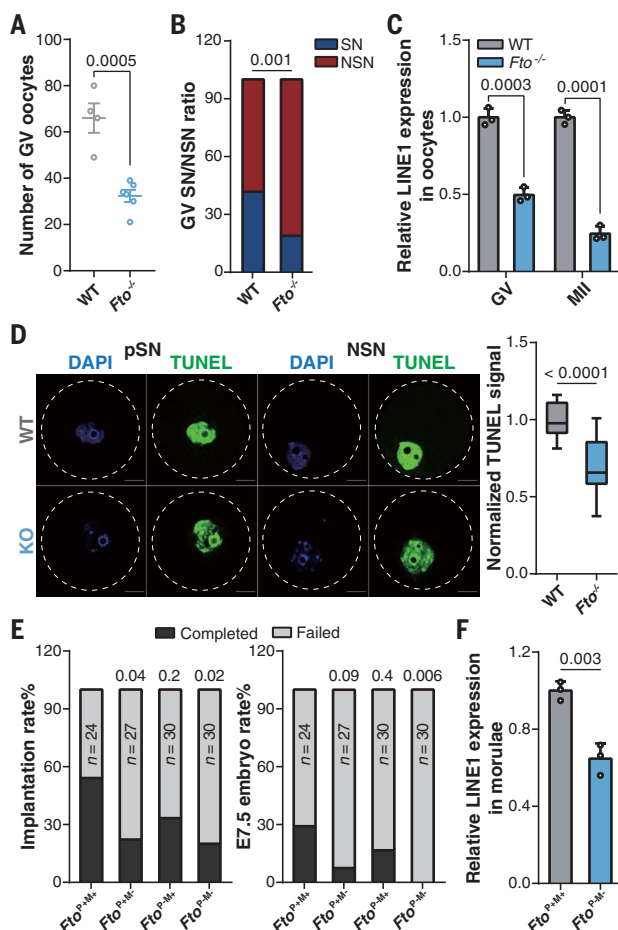
RNA were also reversed in the rescued embryos (fig. S18, Q and R).

Discussion

In this work, we have shown functionally relevant substrates and proposed mechanisms of RNA m⁶A demethylation through FTO in mammalian development. Our findings support LINE1 RNA as a major substrate of FTO in mESCs, although FTO may additionally mediate m⁶A demethylation of other carRNAs to affect gene expression (supplementary text and figs. S19 to S21). In contrast to certain cancer cells in which FTO can be hijacked to mediate mRNA m⁶A demethylation (7–12), which may dominate chromatin state regulation (34) (supplementary text and fig. S22), we found that FTO-mediated m⁶A demethylation maintained LINE1 RNA abundance in mESCs, which contributes to promoting local chromatin openness and activating LINE1-containing genes. We further showed that the FTO-LINE1 RNA axis is functionally relevant in mouse oocyte and embryonic development. How FTO achieves target selectivity in different cellular contexts still needs future investigation. In addition to m⁶A, RNA 5-methylcytosine oxidation is also known to affect transcription of ERVL and ERVL-associated genes in mESCs (35), suggesting a potential widespread presence of regulation through retrotransposon RNA modifications (36).

Materials and methods summary

$Fto^{-/-}$ and control WT mESCs were derived from the inner cell mass of E3.5 blastocysts. m⁶A immunoprecipitation was performed for nonribosomal RNA isolated from the chromatin-associated fraction or from whole cell, as indicated, using the EpiMark N^6-Methyladenosine Enrichment Kit (New England Biolabs). All RNA-seq libraries were prepared using SMARTer Stranded Total RNA-Seq Kit version 2, Pico Input Mammalian (TaKaRa). For most samples, libraries were sequenced on an Illumina NovaSeq 6000 in a 100-base-pair paired-end mode. For RNA-seq data, trimmomatic trimmed reads were aligned to the mm10 reference genome using HISAT2. Read counts were calculated by featureCounts, and differential expression was analyzed using DESeq2. For the generation of preimplantation embryos, MII oocytes were subjected to intracytoplasmic sperm injection and embryo culture. Detailed materials and methods are available in the supplementary materials.

REFERENCES AND NOTES

1. I. A. Roundtree, M. E. Evans, T. Pan, C. He, *Cell* **169**, 1187–1200 (2017).
2. M. Frye, B. T. Harada, M. Behm, C. He, *Science* **361**, 1346–1349 (2018).
3. H. Shi, J. Wei, C. He, *Mol. Cell* **74**, 640–650 (2019).
4. G. Jia et al., *Nat. Chem. Biol.* **7**, 885–887 (2011).
5. F. van der Hoeven et al., *Development* **120**, 2601–2607 (1994).
6. J. Fischer et al., *Nature* **458**, 894–898 (2009).
7. Z. Li et al., *Cancer Cell* **31**, 127–141 (2017).
8. R. Su et al., *Cell* **172**, 90–105.e23 (2018).
9. Y. Huang et al., *Cancer Cell* **35**, 677–691.e10 (2019).
10. R. Su et al., *Cancer Cell* **38**, 79–96.e11 (2020).
11. Y. Qing et al., *Mol. Cell* **81**, 922–939.e9 (2021).
12. Y. Liu et al., *Cell Metab.* **33**, 1221–1233.e11 (2021).
13. J. Liu et al., *Mol. Cell* **77**, 426–440.e6 (2020).
14. J. Mauer et al., *Nature* **541**, 371–375 (2017).
15. J. Mauer et al., *Nat. Chem. Biol.* **15**, 340–347 (2019).
16. J. Wei et al., *Mol. Cell* **71**, 973–985.e5 (2018).
17. K. Boulias et al., *Mol. Cell* **75**, 631–643.e8 (2019).
18. E. Sendinc et al., *Mol. Cell* **75**, 620–630.e9 (2019).
19. S. Akichika et al., *Science* **363**, eaav0080 (2019).
20. R. R. Pandey et al., *Cell Rep.* **32**, 108038 (2020).
21. P. J. Batista et al., *Cell Stem Cell* **15**, 707–719 (2014).
22. J. Liu et al., *Science* **367**, 580–586 (2020).
23. J. Liu et al., *Nature* **591**, 322–326 (2021).
24. W. Xu et al., *Nature* **591**, 317–321 (2021).
25. T. Chelmicki et al., *Nature* **591**, 312–316 (2021).
26. J. W. Jachowicz et al., *Nat. Genet.* **49**, 1502–1510 (2017).
27. M. Percharde et al., *Cell* **174**, 391–405.e19 (2018).
28. J. Y. Lu et al., *Cell Rep.* **30**, 3296–3311.e5 (2020).
29. A. A. Sigova et al., *Science* **350**, 978–981 (2015).
30. D. A. Bose et al., *Cell* **168**, 135–149.e22 (2017).
31. T. Gerken et al., *Science* **318**, 1469–1472 (2007).
32. S. Malki, G. W. van der Heijden, K. A. O'Donnell, S. L. Martin, A. Bortvin, *Dev. Cell* **29**, 521–533 (2014).
33. S. Seisenberger et al., *Mol. Cell* **48**, 849–862 (2012).
34. Y. Wang et al., *Nat. Neurosci.* **21**, 195–206 (2018).
35. D. Guallar et al., *Nat. Genet.* **50**, 443–451 (2018).
36. J. Wei, C. He, *Curr. Opin. Cell Biol.* **70**, 109–115 (2021).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abe9582

Materials and Methods
Supplementary Text
Figs. S1 to S22
References (*37–70*)
MDAR Reproducibility Checklist
Data S1 to S3

Ferroelectricity in untwisted heterobilayers of transition metal dichalcogenides

Lukas Rogée[1]†, Lvjin Wang[2]†, Yi Zhang[1], Songhua Cai[1], Peng Wang[3], Manish Chhowalla[4]*, Wei Ji[2]*, Shu Ping Lau[1]*

Two-dimensional materials with out-of-plane (OOP) ferroelectric and piezoelectric properties are highly desirable for the realization of ultrathin ferro- and piezoelectronic devices. We demonstrate unexpected OOP ferroelectricity and piezoelectricity in untwisted, commensurate, and epitaxial MoS_2/WS_2 heterobilayers synthesized by scalable one-step chemical vapor deposition. We show d_{33} piezoelectric constants of 1.95 to 2.09 picometers per volt that are larger than the natural OOP piezoelectric constant of monolayer In_2Se_3 by a factor of ~6. We demonstrate the modulation of tunneling current by about three orders of magnitude in ferroelectric tunnel junction devices by changing the polarization state of MoS_2/WS_2 heterobilayers. Our results are consistent with density functional theory, which shows that both symmetry breaking and interlayer sliding give rise to the unexpected properties without the need for invoking twist angles or moiré domains.

The rational vertical integration of two-dimensional (2D) materials has led to exciting condensed matter effects that have opened different avenues of research. These interesting effects are a consequence of the interactions between the layers of atomically thin materials that give rise to moiré superlattices, hybrid electronic structures, and breaking of the usual crystal symmetries (1). Materials such as graphene and bilayer 2H MoS_2 are centrosymmetric (2). In contrast, odd numbers of layers of 2D materials such as MoS_2 are noncentrosymmetric, belonging to the $\bar{6}m2$ point group (or D_{3h}), and therefore exhibit in-plane (IP) piezoelectricity. Noncentrosymmetric 2D materials also generate second harmonic emission that can be used to confirm the absence of inversion symmetry. The magnitude of the IP piezoelectric component, referred to as d_{11} (or d_{22} if the armchair direction of the lattice is indexed as 2), has been estimated to be ~2.5 to 4 pm V^{-1} for single-layer MoS_2 (3). Materials of the $\bar{6}m2$ point group do not exhibit out-of-plane (OOP) piezoelectricity (4).

OOP piezoelectricity in 2D materials has been reported in few-layered In_2Se_3 (5) and by introducing chalcogen vacancies in $MoTe_2$ (6). Theoretical studies have explored the piezoelectric properties of transition metal dichalcogenide (TMDC) alloys when assembled into vertical heterostructures (7). Recently, ferroelectricity has been observed in twisted layers of hexagonal boron nitride (h-BN) and TMDCs (8, 9). The origin of ferro- and piezoelectricity in twisted bilayers arises from the formation of moiré lattices and interlayer sliding (10). Ferro- and piezoelectricity have also been observed in rhombohedral homobilayer TMDCs (11). However, OOP piezoelectric and ferroelectric effects in epitaxially grown, untwisted, commensurately stacked, laterally large vertical heterostructures of 2D TMDCs have not been experimentally reported.

We have developed a simple one-step chemical vapor deposition (CVD) process to grow commensurate MoS_2/WS_2 heterobilayers on SiO_2 substrates that possess measurable OOP ferroelectricity and an OOP piezoelectric component d_{33}, even though individual layers of WS_2 and MoS_2 have $d_{33} = 0$. We explain this observation by taking the heterobilayer to be one crystal system with its own point group. In the case of the CVD-grown MoS_2/WS_2 heterobilayers we studied, the point group is $3m$ (or C_{3v}), which lacks the vertical symmetry to nullify OOP strain effects and thus possesses a nonzero d_{33} component that has a magnitude of up to 2.09 pm V^{-1}. A special subgroup of piezoelectrics are also ferroelectric; that is, their internal electric polarization can be switched between two stable states via an external electric field. Ordinary 2D TMDCs are not known to exhibit any ferroelectric characteristics (12). The classification of MoS_2/WS_2 heterobilayers as $3m$ point group materials suggests that they could be ferroelectric. We confirm this via piezoresponse measurements at room temperature. We demonstrate ferroelectric tunnel junctions (FTJs) based on MoS_2/WS_2 heterobilayers, which use the switchability of the ferroelectric to control the tunneling current density through the device (13).

We show an example of our CVD-grown heterobilayers (Fig. 1), which shows smaller WS_2 triangles (lateral dimensions of ~10 μm) draped by a larger MoS_2 monolayer (lateral dimensions of up to 200 μm). The size and shape of the triangles can be changed through variations in the growth recipe [see texts S1 and S2 (14)]. We performed detailed Raman analysis from different regions on the sample (Fig. 1B), which shows pure single-layer MoS_2 (region labeled as α, IP vibrational mode E' at ~383 cm^{-1} and OOP vibrational mode A'_1 at ~403 cm^{-1}). The triangles labeled with β show Raman signals from both MoS_2 and WS_2 (E' mode of WS_2 at ~355 cm^{-1} and its A'_1 mode at ~417 cm^{-1} along with the MoS_2 peaks). We also show a scanning electron microscope (SEM) image of a large MoS_2 layer covering a smaller WS_2 triangle (Fig. 1C). We obtained cross-sectional high annular angle dark field scanning transmission electron microscope (HAADF-STEM) images from two regions (Fig. 1C, labeled d and e). MoS_2 and WS_2 can be easily distinguished in our cross-sectional STEM images by the higher contrast of the W atoms that make the WS_2 layer noticeably brighter than the MoS_2 layer. Bright-field STEM (BF-STEM) images provide additional evidence of the bilayer structure (see text S3). We performed chemical analysis of the heterobilayers using energy-dispersive x-ray spectroscopy (EDS) to confirm the chemical composition of the heterobilayers (see text S4). The larger MoS_2 layer draping over the edge of the WS_2 layer is clearly visible in the cross-sectional image (Fig. 1D). The interior region of the bilayer (Fig. 1E) clearly shows an MoS_2 layer on top of WS_2. We also collected additional photoluminescence (PL) and selected-area electron diffraction (SAED) data about the CVD-grown materials (see text S5).

We studied the stacking angle between MoS_2 and WS_2 by second harmonic generation (SHG) emission (Fig. 1G) (texts S6 and S7), which depends directly on the interlayer rotation angle θ (15). MoS_2 and WS_2 exhibit broad absorption at energies above 2.5 eV (16); thus, incident photons with a wavelength of 900 nm (1.37 eV) readily induce SHG emissions of 450 nm (2.74 eV) in both layers. In short, SHG emissions interfere entirely constructively (bright signal) when θ = 0° where the stacking sequence is similar to the 3R stacking in TMDC crystals (Fig. 1H). Conversely, when the stacking angle is θ = 60° (or 180°, 300° and so on because of three-fold rotation symmetry of TMDCs around the c axis) as in the 2H-phase TMDCs, the layers interfere entirely destructively and produce a dark signal.

Vertical bilayer heterostructures are often associated with the appearance of moiré patterns, which can have a substantial impact

[1]Department of Applied Physics, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong, P. R. China. [2]Department of Physics and Beijing Key Laboratory of Optoelectronic Functional Materials and Micro-Nano Devices, Renmin University of China, Beijing 100872, P. R. China. [3]College of Engineering and Applied Sciences and Collaborative Innovation Center of Advanced Microstructures, Nanjing University, Nanjing 210093, P. R. China. [4]Department of Materials Science and Metallurgy, University of Cambridge, Cambridge, UK.
*Corresponding author. Email: apsplau@polyu.edu.hk (S.P.L.); wji@ruc.edu.cn (W.J.); mc209@cam.ac.uk (M.C.)
†These authors contributed equally to this work.
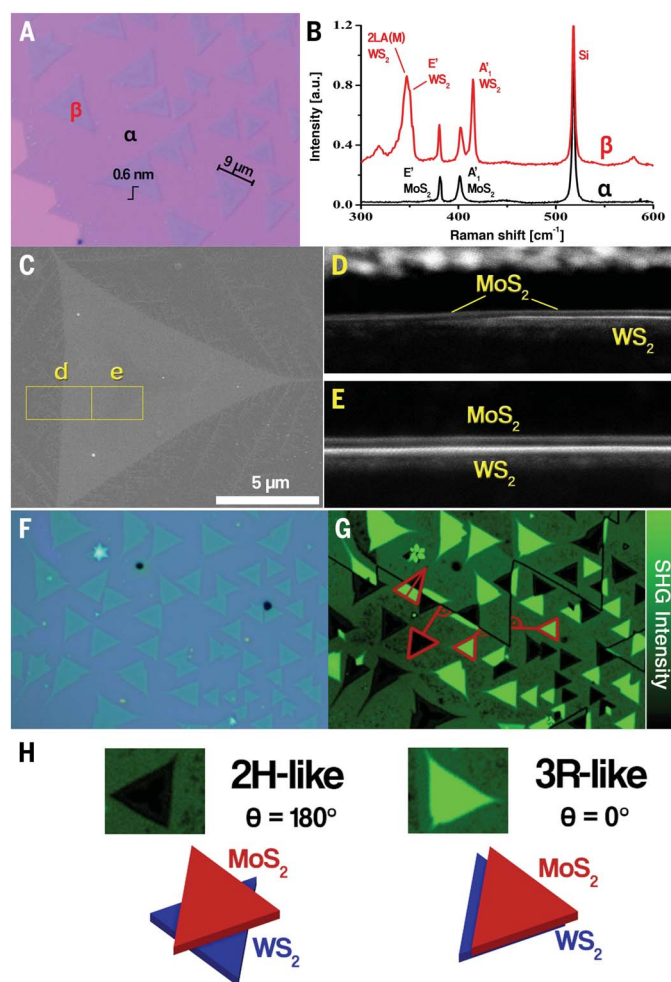
Fig. 1. CVD-grown MoS₂/WS₂ heterobilayers. (**A**) Optical microscope image of the as-grown heterobilayers showing smaller (~10 μm) WS₂ triangles draped over by a larger (~200 μm) MoS₂ layer. The thicknesses of the triangle edges were measured by AFM to be 0.6 nm—the thickness of a TMDC monolayer. (**B**) Raman spectra from α and β regions labeled in (A). The α regions show typical MoS₂ signals, whereas the β region shows both WS₂ and MoS₂ peaks. (**C**) Planar SEM image of a typical as-grown triangle. The two yellow rectangles *d* and *e* indicate regions selected for cross-sectional STEM imaging as shown in (D) and (E), respectively. (**D** and **E**) In region D, the MoS₂ layer draping over the brighter WS₂ layer is observed. In region E, a uniform bilayer consisting of MoS₂ on top of WS₂ can be seen. (**F**) Optical microscopy image of MoS₂/WS₂ triangles across a large MoS₂ cluster. (**G**) The corresponding unfiltered SHG intensity map. Note that there is no bare SiO₂ substrate visible in the image. The MoS₂/WS₂ triangles appear either very bright or very dark across the map. The black lines are single-crystal domain boundaries of the large MoS₂ monolayer. Bright triangles always point toward the nearest domain boundaries; dark triangles point away from them, as indicated by the red lines. (**H**) Relationship between SHG intensity and vertical stacking angle θ. Dark triangles are labeled 2H-like and bright triangles 3R-like.

on their piezo- and ferroelectric properties. In our case, the epitaxial heterobilayers do not show any moiré patterns (see text S8). Briefly, naturally grown bilayer TMDCs in either 2H or 3R stacking do not show moiré superlattices because their layers are commensurate. This property is not exclusive to homobilayers. MoS₂ and WS₂ have virtually identical lattice parameters (*16*, *17*) and as a consequence, heterobilayers of MoS₂ and WS₂ with a twist angle of $n \cdot 60°$ (with n being an integer) also do not show moiré superlattices. Our heterobilayers grown by CVD naturally align with each other in energetically ideal arrangements (i.e., epitaxially), consistent with epitaxially grown WSe₂/MoSe₂ heterobilayers (*18*) that also show an absence of moiré patterns because they also possess identical lattice constants. Commensurate stacking of different TMDC monolayers cannot yet be achieved by artificial stacking. A small difference in lattice constant between MoS₂ and WS₂ of 0.01 Å has been reported, but studies have shown that heterobilayers with lattice constant differences of up to 2% can become commensurate during the CVD growth. An example is CVD-grown monolayer graphene on Ni(111), which, despite a small lattice mismatch, reproduces a 1×1 commensurate heterolattice (*19*).

We show (fig. S9A) the edge of two heterobilayers with the same orientation. At this scale, outlines of moiré patterns would become apparent in artificially stacked bilayers with small stacking angles. However, this is not the case for our CVD samples. The corresponding fast Fourier transform (FFT) spectrum of the image shows a single hexagonal crystal pattern similar to the SAED pattern (fig. S6D). Moiré patterns are also absent in the STEM image of another sample (fig. S9B) of 3R-like and 2H-like heterobilayers. We collected the FFT spectra for the entire image, the individual 3R-like and 2H-like stacked heterobilayers as well as that of only MoS₂. All four FFT spectra are the same, which suggests no misalignment or twisting between the two layers. We therefore do not need to invoke twisting or moiré lattices to describe the origins of piezo- and ferroelectricity in our CVD-grown heterobilayers.

Piezoelectric mapping

Both 3R-like and 2H-like MoS₂/WS₂ heterobilayers can be treated as materials belonging to the $3m$ (or C_{3v}) point group that should exhibit a nonzero OOP piezoelectric constant d_{33} and potentially be ferroelectric. We have therefore investigated both properties using piezoresponse force microscopy (PFM) (DART-SS-PFM mode, Asylum Research) (*14*, *20*). In short, we created an alternating electric field (at voltage V_{AC}) locally using a conductive atomic force microscope (AFM) tip, which causes piezoelectric materials to deform and the magnitude of the deformation is measured and mapped. For PFM measurements, we transferred heterobilayers onto conductive substrates to avoid electrical charging during the measurements. We show SHG maps of triangles with different vertical stacking arrangements (Fig. 2A) and their corresponding AFM images (Fig. 2B). The two different stacking arrangements of MoS₂/WS₂ heterobilayers appear very similar in the AFM. We performed resonance-amplified PFM to obtain the OOP piezoelectric constant, mapping the results at different voltages between 1.2 and 2.0 V (Fig. 2C). We set the color scale of the PFM maps such that the mean height change of the pure MoS₂ is zero, and its color is green. This helps to account for possible electrostrictive effects from the background because MoS₂ by itself does not exhibit OOP piezoelectricity. The results show that with increasing voltage, the color contrast between the MoS₂/WS₂ heterobilayer and the pure MoS₂ monolayer significantly increases (Fig. 2C). That is, notable OOP piezoelectricity is indicated by the red color in the PFM maps.
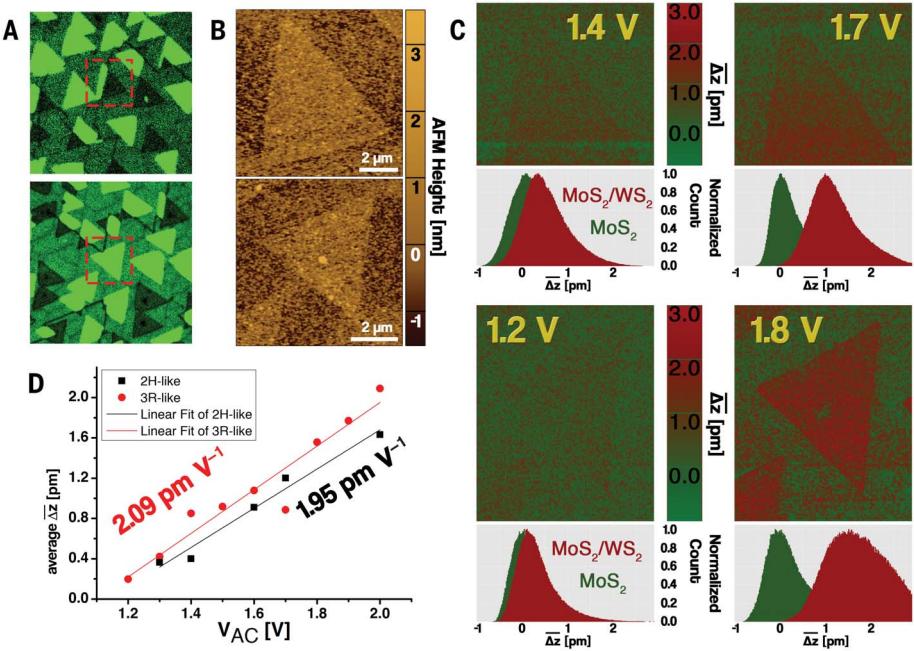
Fig. 2. PFM data of MoS$_2$/WS$_2$ heterobilayers on a conductive Pt-coated substrate. (**A**) SHG maps showing the triangles that were chosen for the PFM measurements. The upper triangle is largely 2H-like. The lower triangle is 3R-like. (**B**) AFM maps of the 2H-like and 3R-like triangles. The materials are atomically smooth with RMS roughness of ~0.1 nm. (**C**) Real piezoelectric height change $\overline{\Delta z}$ maps of both triangles, measured at different drive voltages V_{AC}. The distributions below each map show the piezoelectric height change for the MoS$_2$/WS$_2$ triangles (red) and the surrounding monolayer MoS$_2$ (green). (**D**) Plot of the average values of $\overline{\Delta z}$ as a function of V_{AC}.

Table 1. Overview of several 2D materials and their measured piezoelectric constants.

Material	Piezoelectric constant	Experimental value (pm V^{-1})
2H-like MoS$_2$/WS$_2$	d_{33}	1.95
3R-like MoS$_2$/WS$_2$	d_{33}	2.09
Monolayer MoS$_2$ (35)	d_{11}	3.78
Monolayer WSe$_2$ (36)	d_{11}	5.2
Monolayer α-In$_2$Se$_3$ (5)	d_{33}	0.34

The average of the vertical piezoelectric deformation $\overline{\Delta z}$ of multiple 2H-like and 3R-like triangles was obtained by rigorous statistical analysis (see text S9) to find the piezoelectric height change on the vertical heterostructures relative to the background. We plotted the $\overline{\Delta z}$ distributions for both the background and heterobilayers below each PFM map to show that the distance between the background and triangle distributions increases with increasing voltage, as we expected for OOP piezoelectricity. The distribution for the triangles also gets broader. This trend is because the real value of d_{33} varies slightly across the area of a triangle, and these slight differences multiplied by an increasing voltage V_{AC} result in ever greater contrast be-

tween the extremes of $\overline{\Delta z}$. We have plotted the average of $\overline{\Delta z}$ as a function of V_{AC} for both stacking types along with their respective linear fits (Fig. 2D).

The OOP piezoelectric component can be calculated using $d_{33} = \partial(\overline{\Delta z})/\partial V_{AC}$. We find $d_{33} = 1.95 \pm 0.2$ pm V^{-1} and 2.09 ± 0.2 pm V^{-1} for 2H-like and 3R-like stacked MoS$_2$/WS$_2$, respectively. We compare this to other experimentally obtained piezoelectric constants of other 2D materials (Table 1). We find that our measured data are similar in magnitude to the IP d_{11} piezoelectric constant of monolayer MoS$_2$ and substantially larger than the OOP d_{33} constant of monolayer α-In$_2$Se$_3$ (5). We also see that the stacking orientation of the heterostructure has a slight influence on its piezo-

electric constant. As a result of the difference in stacking, the relative positions of W, Mo, and S atoms are different, which ultimately influences the magnitude and direction of the internal polarization. Spurious OOP piezoelectric effects can arise when the root mean square (RMS) roughness is higher than the thickness of TMDCs (>1.5 nm) (21). The AFM images (Fig. 2B) show that this is not the case in our samples because the heterobilayers are atomically smooth with RMS surface roughness of ~0.1 nm. This value is considerably less than the thickness of the heterobilayer, hence the OOP deformation we reported is intrinsic.

Ferroelectric hysteresis

The observation of a piezoelectric response in MoS$_2$/WS$_2$ heterobilayers does not necessarily imply the presence of ferroelectricity; however, the $3m$ point group classification indicates that it is possible (12, 22). Room-temperature ferroelectricity in stacked large-area CVD-grown TMDCs could open up possibilities for exciting electronics applications. We thus investigated the ferroelectric response of the heterobilayers (Fig. 3). First, we applied the DART-SS-PFM hysteresis method to our sample (14). The OFF-field phase loop (Fig. 3A) shows the typical shape obtained from domain switching in ferroelectric materials. The polarization switching occurs at the coercive voltage of ~$V_{DC} = \pm 3$ V in the heterobilayers. The corresponding OFF-field $\overline{\Delta z}$ loop exhibits the typical ferroelectric butterfly shape. Generally, the OFF-field piezoresponse hysteresis loops are used to investigate the ferroelectric performance to avoid spurious electrostrictive and electrochemical forces that can otherwise also cause piezoresponse loops that appear similar to ferroelectric ones (23). The butterfly loop is slightly offset toward the negative voltage direction. This behavior is indicative of small influences from nonferroelectric artifacts, such as charge injection, which is a common feature in ultrathin ferroelectrics (24). A phase loop is also apparent in the corresponding ON-field hysteresis loops (Fig. 3B), although it is more abrupt. The ON-field $\overline{\Delta z}$ loop appears as a large V-shape with a small butterfly pattern. These two shapes show that the ferroelectric and electrostrictive deformation coexist as long as a strong unidirectional electric field is present. Some nonferroelectric materials such as Al$_2$O$_3$ exhibit piezoresponse hysteresis loops even in OFF-field loops, which can be mistaken to be of ferroelectric origin (25, 26). We show through variation of the drive voltage V_{AC} that the ferroelectricity we observed is intrinsic (see text S10). We provide further evidence of ferroelectricity through domain writing (14). We poled a large area of the heterobilayer (with –8 V tip bias) and

Fig. 3. Ferroelectricity in MoS$_2$/WS$_2$ heterobilayers. Piezoelectric hysteresis loops were measured by applying DC voltage sweeps from −5.5 to +5.5 V at 2.7 V AC drive voltage using DART-SS-PFM. (**A**) Phase and $\overline{\Delta z}$ hysteresis loops with the DC field OFF. (**B**) Phase and $\overline{\Delta z}$ hysteresis loops with the DC field ON. (**C** and **D**) Two square-shaped domains were written onto a heterobilayer triangle with −8 V for the outer square and 8 V for the inner square. The entire triangle is outlined by a dotted white line for clarity. The phase map (C) and PFM amplitude map (D) show that the poled areas, marked by dashed yellow outlines, do not extend beyond the boundaries. (**E**) Polarization-dependent tunneling current versus applied DC field through an FTJ device. The tunneling current is strongly dependent on the previous poling voltage. The ratio between the low- and high-resistance states (LRS and HRS) is 10^2 to 10^3. (**F**) Schematic of the tunneling potential barrier change for the LRS and HRS states for an FTJ. Electron tunneling is reduced when the device is poled with a negative voltage, and increased with a positive poling voltage. (**G**) Switching voltage program applied to an FTJ for more in-depth analysis. A negative poling voltage pulse of −3.5 V is applied, followed by three positive voltage pulses with increasing magnitudes of 1 V, 3 V, and 3.5 V. (**H**) The corresponding I-V loop. The arrows indicate the direction of the current change.
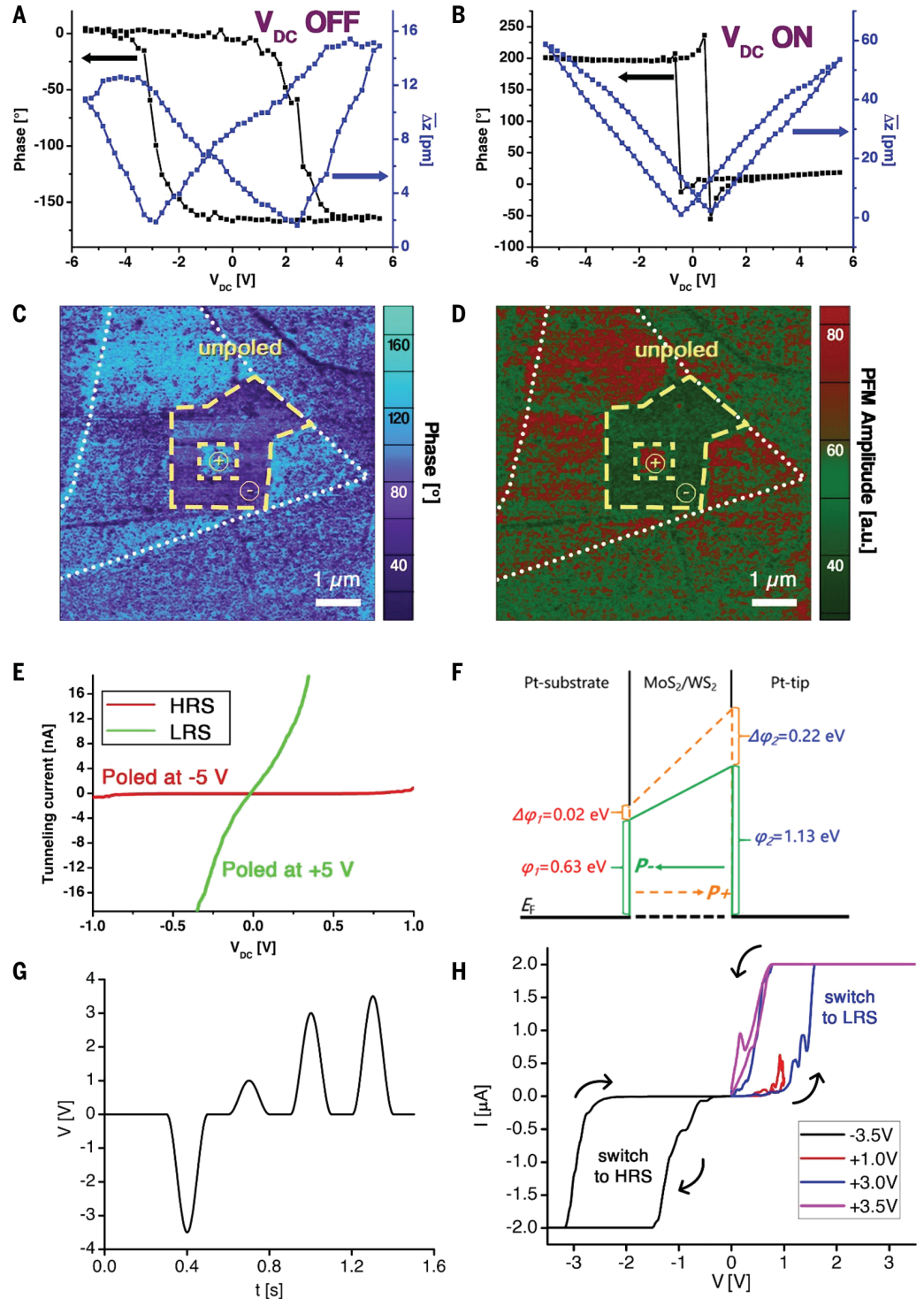
a smaller square within that area (+8 V tip bias). These shapes are outlined in the phase and amplitude maps performed after poling (Fig. 3, C and D). They show a strong change in both the phase direction as well as the PFM amplitude in the designated areas.

To translate the fundamental ferroelectric properties into a practical demonstration, we measured the properties of FTJ devices based on the heterobilayers. The V_{DC}-dependent tunneling current in a FTJ can be substantially modified by poling the device prior to measurements (Fig. 3E). The device characteristics are also similar to ferrodiode behavior as reported by Liu *et al.* (*27*). Thus, although we use the term FTJ to describe our device, distinguishing between the two requires additional observations. In this case, negative poling with −5 V diminishes the electron tunneling current and puts the device in a high-resistance state (HRS), whereas positive poling of +5 V puts it in a low-resistance state (LRS). The ratio of the tunneling resistance between the two states reaches a value of 10^2 to 10^3. We show schematically how modification of the tunnel barrier with polarization of the ferroelectric (*13, 28–30*) increases or decreases the
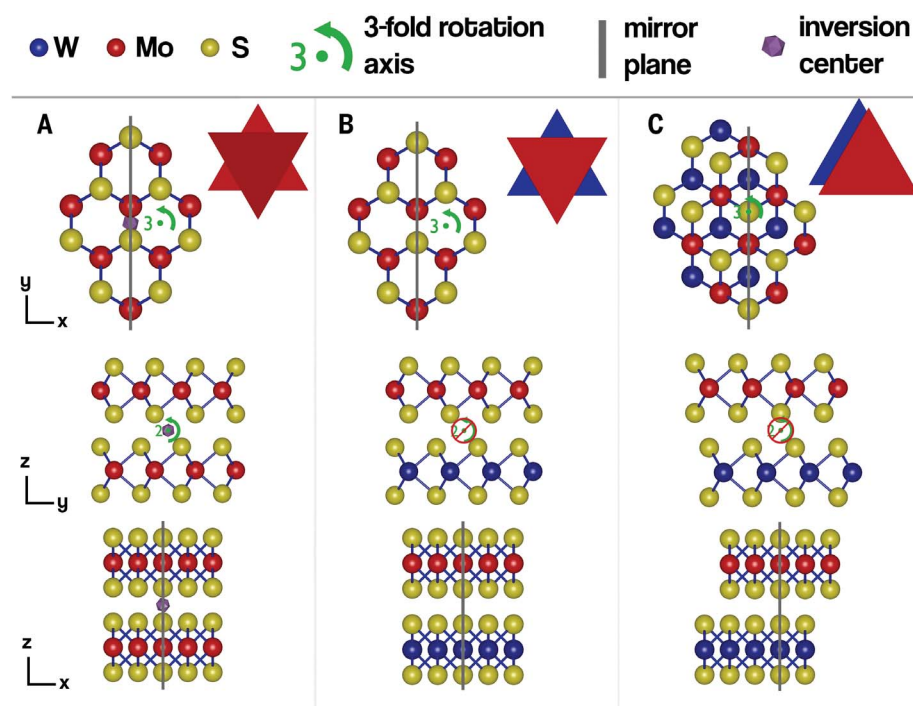
Fig. 4. Crystal symmetry models. (A to C) Schematic representation of bilayer 2H MoS$_2$ (A), 2H-like (B), and 3R-like heterostructure (C) of MoS$_2$/WS$_2$ from three different perspectives as indicated by the coordinate axes. The c direction (out-of-plane direction) is along the z axis, the zigzag direction is along the x axis, and the armchair direction is along the y axis.

tunneling current (Fig. 3F) (*31, 32*). We have modeled our FTJ as a dual-slab made up of the MoS$_2$/WS$_2$ bilayer sandwiched by platinum slabs (see text S11). To reflect the real experimental setup, we increased the distance between the bilayer and the top electrode by 2 Å to model their weak interaction (fig. S11, E to H). Further, we have repeatedly applied a resistance-switching voltage program to an FTJ (Fig. 3G) to test for robustness. For each measurement cycle, we applied a strong negative voltage pulse (–3.5 V) to the sample to ensure that it is poled in the HRS. Then, we applied three positive voltage pulses of 1 V, 3 V, and 3.5 V. The resulting tunneling current measurements (Fig. 3H) show that the negative voltage induces HRS. The application of the 1-V pulse increases and decreases the current along the HRS curve without forming an open loop. Once the 3-V pulse is applied, an open loop is created where the sample switches from HRS to LRS, and the 3.5-V pulse simply follows the LRS curve without any open loop because resistance switching has already occurred. These measurements confirm that a voltage of at least the coercive voltage is needed to switch between the FTJ states, in accordance with the PFM data from Fig. 3A. More detailed poling experimental data on the same and other devices can be found in (*14*). Our overall hysteresis results indicate that MoS$_2$/WS$_2$ heterobilayers, as a

3*m* point group material, exhibit ferroelectric properties at room temperature.

Crystal symmetries

TMDCs are known to show IP piezoelectric properties but no OOP piezoelectric properties (*4, 33*). In accordance with the rules of group theory (see text S13), bilayer 2H MoS$_2$ belongs to the $\bar{3}\frac{2}{m}$ (or D_{3d}) point group. We depict a schematic model of its crystal structure (Fig. 4A), which includes an inversion center. Ferroelectricity and piezoelectricity (also SHG) do not occur if an inversion center is present in a crystal. The 2H-like MoS$_2$/WS$_2$ heterobilayer (Fig. 4B) crystal structure is like bilayer 2H MoS$_2$ with the exception that Mo atoms are replaced with W atoms in the bottom layer. This makes a substantial difference to the crystal symmetry but does not lead to the appearance of moiré patterns because the lattice parameters of the two materials and their epitaxial growth are similar. As a result, all symmetry transformations that exchange atoms between the top and bottom layers become invalid, including the inversion center. The symmetry transformations that are left put the heterobilayer into the 3*m* (or C_{3v}) point group. The same symmetry transformations also apply for 3R-like MoS$_2$/WS$_2$ (Fig. 4C) Hence, both stacking types belong to the same point group. The 3*m* point group has exactly one nonzero

OOP piezoelectric constant, d_{33}. 3*m* point group materials also classify as polar materials because they have a unique rotation axis, no inversion center, and no mirror plane perpendicular to the rotation axis (*12*). This allows ferroelectricity to be possible from a geometric standpoint.

Theoretical derivation of strain-piezoelectric constant and mechanism of ferroelectric switching

We measured the strain-piezoelectric constant (d), which cannot be directly obtained from density functional theory (DFT) calculations. However, we can derive it from the stress-piezoelectric constant (e) and the elastic constant tensor (C) using the relation $d = eC^{-1}$, which can be obtained from DFT calculations. The OOP component of the strain piezoelectric constant tensor d_{33}, as measured in our PFM experiment, is theoretically derived from $d_{33} = [(C_{11} + C_{12})e_{33} - 2C_{13}e_{31}]/[(C_{11} + C_{12})C_{33} - 2C_{13}^2]$ (see text S14).

We determined the vertical strain piezoelectric constants of 2H-like and 3R-like MoS$_2$/WS$_2$ heterobilayers to be 2.28 pm V^{-1} and 2.40 pm V^{-1}, respectively. Both the absolute values of the two d_{33} constants, 2.28 and 2.40 pm V^{-1}, and their difference, 0.12 pm V^{-1}, are close to experimentally measured values (i.e., 1.95, 2.09, and 0.14 pm V^{-1}, respectively).

According to our calculations, both heterobilayers show spontaneous nonzero OOP electric polarizations, namely P$_{out}$[2H-like] = 0.44 pC m^{-1} and P$_{out}$[3R-like] = 0.60 pC m^{-1}. Switching their OOP polarization directions is not likely to be accessible by vertically moving any atoms, but could be achieved by a lateral sliding between the two monolayers in each of the heterobilayers, showing ferroelectric behavior. We show that the atomic structures of two stacking configurations (i.e., AA-up and AA-down) of the 3R-like heterobilayer differ from one another by a 1.83 Å lateral sliding along the armchair direction (see text S15). Configuration AA-up has an OOP polarization of 0.60 pC m^{-1} that is 1.9 meV/f.u. more stable than configuration AA-down with a negative value of –1.45 pC m^{-1}. An external electric field over 2.4 V/nm could switch their relative stability and thus, together with thermal excitation at finite temperatures, trigger the sliding occurrence surmounting a 16 meV/f.u. barrier, which accompanies reversal of polarization direction, displaying an explicit ferroelectric switching behavior.

The interfacial differential charge densities (DCD, a measure of charge variation at the interface) of two related AA-like (i.e., 3R-like for our particular case) stacking configurations of the heterobilayers explicitly show charge redistribution between the top and bottom layers that are illustrated as separation of the red (electron accumulation) and green (electron
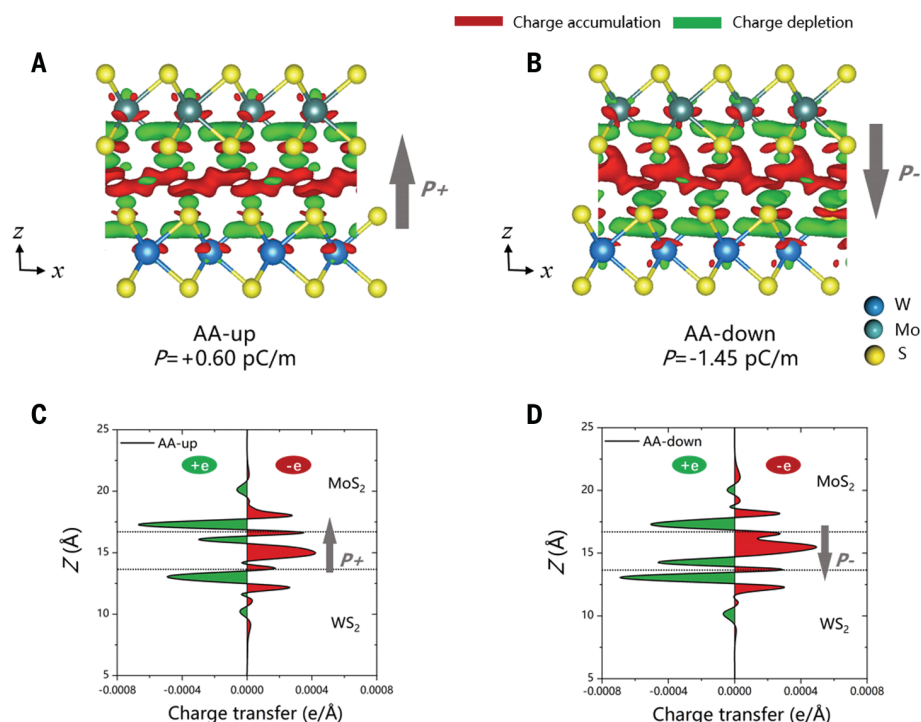
Fig. 5. Charge density plots. (**A** and **B**) Interlayer differential charge density for the up (AA-up) and down (AA-down) polarizations, respectively. An iso-surface value of 7×10^{-5} e/bohr^3 was used. (**C** and **D**) Their line profiles along z, respectively.

depletion) regions (Fig. 5). Their line profiles show explicit electric polarization at the interfaces and the direction of polarization is switchable under lateral sliding of one layer across roughly one-third of the unit cell. Our calculations give a switching barrier of 16 meV/f.u. This is comparable to the value of 9 meV/f.u. predicted by Li *et al.* (*34*) in BN bilayers where the authors used the sliding mechanism. This mechanism was also used to explain those experimentally observed ferroelectric effects by Stern *et al.* (*10*) and Yasuda *et al.* (*9*) in a temperature range from 4.2 K to 300 K.

Conclusion

This work demonstrates that ferro- and piezoelectricity can be found in untwisted commensurate bilayers consisting of monolayers of MoS$_2$ and WS$_2$. The heterobilayer is easy to grow in large quantities through a one-step CVD process, which does not require any precision transfer method or setup to realize, and can be scaled through variations in the growth recipe. Previous research on heterostructure ferroelectricity relied on

the local moiré structures. In contrast, our material is free from any periodic superstructure and can be explained by a group theory approach for each entire bilayer type. This approach could be applied to other bottom-up heterostructures.

REFERENCES AND NOTES

1. A. Weston *et al.*, *Nat. Nanotechnol.* **15**, 592–597 (2020).
2. S. Islam, S. S. Z. Ashraf, *Resonance* **24**, 445–457 (2019).
3. C. Cui, F. Xue, W.-J. Hu, L.-J. Li, *npj 2D Mater. Appl.* **2**, 18 (2018).
4. A. H. Meitzler, IEEE Standard on Piezoelectricity. *ANSI/IEEE Std 176-1987*, 0_1 (1988).
5. F. Xue *et al.*, *ACS Nano* **12**, 4976–4983 (2018).
6. D. Seol *et al.*, *Nano Energy* **79**, 105451 (2021).
7. S. Yu, Q. Rice, B. Tabibi, Q. Li, F. J. Seo, *Nanoscale* **10**, 12472–12479 (2018).
8. A. Weston *et al.*, arXiv 2108.06489 (2021).
9. K. Yasuda, X. Wang, K. Watanabe, T. Taniguchi, P. Jarillo-Herrero, *Science* **372**, 1458–1462 (2021).
10. M. Vizner Stern *et al.*, *Science* **372**, 1462–1466 (2021).
11. X. Wang *et al.*, *Nat. Nanotechnol.* **17**, 367–371 (2022).
12. P.-P. Shi *et al.*, *Chem. Soc. Rev.* **45**, 3811–3827 (2016).
13. P. Hou *et al.*, *J. Mater. Chem. C* **6**, 5193–5198 (2018).
14. See supplementary materials.
15. W.-T. Hsu *et al.*, *ACS Nano* **8**, 2951–2958 (2014).
16. R. A. Bromley, R. B. Murray, A. D. Yoffe, *J. Phys. C* **5**, 759–778 (1972).
17. A. Molina-Sánchez, L. Wirtz, *Phys. Rev. B* **84**, 155413 (2011).
18. W.-T. Hsu *et al.*, *Nat. Commun.* **9**, 1356 (2018).
19. Y. Dedkov, E. Voloshina, M. Fonin, *Phys. Status Solidi B* **252**, 451–468 (2015).
20. Oxford Instruments, *Piezoresponse Force Microscopy with Asylum Research AFMs*; https://afm.oxinst.com/assets/uploads/products/asylum/documents/Piezoresponse-Force-Microscopy-AFM-web.pdf.
21. S. Kang *et al.*, *ACS Appl. Mater. Interfaces* **10**, 27424–27431 (2018).
22. D. Litvin, Ferroelectric Space Groups, *Acta Crystallogr. A* **42**, 44–47 (1986).
23. B. Kim, D. Seol, S. Lee, H. N. Lee, Y. Kim, *Appl. Phys. Lett.* **109**, 102901 (2016).
24. S. S. Cheema *et al.*, *Nature* **580**, 478–482 (2020).
25. Z. Guan *et al.*, *AIP Adv.* **7**, 095116 (2017).
26. H. Qiao, O. Kwon, Y. Kim, *Appl. Phys. Lett.* **116**, 172901 (2020).
27. X. Liu *et al.*, *Appl. Phys. Lett.* **118**, 202901 (2021).
28. R. Li *et al.*, *Appl. Phys. Lett.* **116**, 222904 (2020).
29. J. Wu *et al.*, *Nat. Electron.* **3**, 466–472 (2020).
30. S. Yuan *et al.*, *Nat. Commun.* **10**, 1775 (2019).
31. E. Y. Tsymbal, H. Kohlstedt, *Science* **313**, 181–183 (2006).
32. A. Gruverman, M. Alexe, D. Meier, *Nat. Commun.* **10**, 1661 (2019).
33. R. Hinchet, U. Khan, C. Falconi, S.-W. Kim, *Mater. Today* **21**, 611–630 (2018).
34. L. Li, M. Wu, *ACS Nano* **11**, 6382–6388 (2017).
35. S. K. Kim *et al.*, *Nano Energy* **22**, 483–489 (2016).
36. E. Nasr Esfahani, T. Li, B. Huang, X. Xu, J. Li, *Nano Energy* **52**, 117–122 (2018).

ATMOSPHERE

Hydrotrioxide (ROOOH) formation in the atmosphere

Torsten Berndt[1]*, Jing Chen[2], Eva R. Kjærgaard[2]†, Kristian H. Møller[2]‡, Andreas Tilgner[1], Erik H. Hoffmann[1], Hartmut Herrmann[1], John D. Crounse[3], Paul O. Wennberg[3,4], Henrik G. Kjaergaard[2]*

Organic hydrotrioxides (ROOOH) are known to be strong oxidants used in organic synthesis. Previously, it has been speculated that they are formed in the atmosphere through the gas-phase reaction of organic peroxy radicals (RO_2) with hydroxyl radicals (OH). Here, we report direct observation of ROOOH formation from several atmospherically relevant RO_2 radicals. Kinetic analysis confirmed rapid RO_2 + OH reactions forming ROOOH, with rate coefficients close to the collision limit. For the OH-initiated degradation of isoprene, global modeling predicts molar hydrotrioxide formation yields of up to 1%, which represents an annual ROOOH formation of about 10 million metric tons. The atmospheric lifetime of ROOOH is estimated to be minutes to hours. Hydrotrioxides represent a previously omitted substance class in the atmosphere, the impact of which needs to be examined.

ydrotrioxides (ROOOH) are known, thermally unstable products formed in the low-temperature ozonolysis of saturated organic compounds in organic solvents. They are a chemical source of the powerful oxidant singlet molecular oxygen ($^1O_2$) released during their decomposition (1, 2). Accordingly, hydrotrioxides are used in preparative chemistry to form the corresponding oxetane and carbonyl products in the reaction with alkenes mostly carried out at dry-ice temperature (3).

In atmospheric gas-phase chemistry, theoretical calculations have proposed the formation of hydrotrioxides as intermediates in the reaction of RO_2 radicals with OH, as shown in pathway 1 below (4, 5). The rapid radical recombination reaction is exothermic by ~130 kJ mol^{-1}, nearly independent of the RO_2 radical, initially forming the energy-rich ROOOH*. This chemically excited species can decompose, leading to the corresponding alkoxy radical RO and HO_2 (pathway 2) or, to a lesser extent, to an alcohol and O_2 (pathway 3). In competition with decomposition, collisions with bath gas molecules, M, result in thermalized ROOOH (pathway 4) (4–6).

$$RO_2 + OH \rightarrow ROOOH^* \qquad (1)$$

$$ROOOH^* \rightarrow RO + HO_2 \qquad (2)$$

$$ROOOH^* \rightarrow ROH + O_2 \qquad (3)$$

$$ROOOH^* + M \rightarrow ROOOH + M \qquad (4)$$
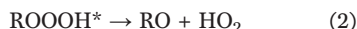
Previous experimental investigations of the RO_2 + OH reaction at 298 K and 50 torr He found a decreasing HO_2 yield, with the RO_2 radical size increasing from C_1 to C_4. For CH_3O_2, the HO_2 yield was 0.90 ± 0.10, which decreased to 0.15 ± 0.03 for n-$C_4H_9O_2$ (6). Calculations supported by these experimental findings suggested that at 298 K and 1 bar of N_2, pathway 4 was the dominant fate of ROOOH* in the case of $C_2H_5O_2$ (78%) and larger RO_2 radicals (>95%). For CH_3O_2 radicals, decomposition into CH_3O and HO_2 via pathway 2 still dominates (6). Thus, with the exception of CH_3O_2 radicals, formation of the thermalized ROOOH is the expected dominant product from RO_2 + OH reactions in the atmosphere.

There has been a lot of speculation in the literature about the physical chemistry of the reactions of RO_2 radicals with OH in the atmosphere. All evidence that hydrotrioxides are formed has, to date, been indirect, and an experimental proof of hydrotrioxides has been missing up to now (4–7).

In this work, we conclusively demonstrate, through their direct detection, that hydrotrioxide formation takes place from RO_2 + OH reactions under atmospheric conditions. The investigations were conducted in a free-jet flow system at 295 ± 2 K, a pressure of 1 bar of air, and a reaction time of 7.5 s using product monitoring by chemical ionization mass spectrometry (8, 9). Quantum chemical calculations (10–12) were carried out in support of the reaction mechanisms as well as the thermal stability and photostability of hydrotrioxides (supplementary materials, section S4).

ROOOH from trimethylamine oxidation

We observed a strong signal consistent with ROOOH formation in the reaction of OH radicals with trimethylamine [$N(CH_3)_3$] using iodide for product ionization in the mass spectrometric detection (fig. S1). In this reaction system, an efficient autoxidation mechanism (13) (repeated RO_2 isomerization and O_2 addition) rapidly formed the RO_2 radical $(HOOCH_2)_2NCH_2O_2$ (**I**) as a main product (10, 14). The RO_2 radical (**I**) can react with OH to form the hydrotrioxide (**III**); however, this is in competition with unimolecular RO_2 isomerization forming the dihydroperoxy amide (**II**) and with the RO_2 self-reaction (**I** + **I**) forming the accretion product (**IV**), as illustrated in Fig. 1.

The signal with the mass of the hydrotrioxide $(HOOCH_2)_2NCH_2OOOH$ (**III**) steeply increased with increasing concentrations of the OH precursor, isopropyl nitrite (IPN), i.e., for rising OH and RO_2 radical concentrations in the experiment (Fig. 2A). The signal of **III** behaved similarly to that of the accretion product **IV** formed in the self-reaction of the RO_2 radical **I**. Both followed second-order kinetics, in clear contrast to the first-order kinetics of amide $(HOOCH_2)_2NCHO$ (**II**) formation, arising from the RO_2 isomerization of **I** (10). Because OH and RO_2 radical concentrations increased in a similar way with increasing IPN concentrations, the product of the RO_2 + OH reaction increased almost parallel to the accretion product **IV** (Fig. 2A).

[1]Atmospheric Chemistry Department (ACD), Leibniz Institute for Tropospheric Research (TROPOS), 04318 Leipzig, Germany. [2]Department of Chemistry, University of Copenhagen, DK-2100 Copenhagen Ø, Denmark. [3]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125, USA. [4]Division of Engineering and Applied Science, California Institute of Technology, Pasadena, CA 91125, USA.
*Corresponding author. Email: berndt@tropos.de (T.B.); hgk@chem.ku.dk (H.G.K.)
†Present address: Department of Chemistry, Aarhus University, DK-8000 Aarhus C, Denmark.
‡Present address: Niels Bohr Institute, University of Copenhagen, DK-1350 Copenhagen, Denmark.

Fig. 1. Product formation starting from the $(HOOCH_2)_2NCH_2O_2$ radical (I).

Fig. 2. Product formation from OH + trimethylamine for increasing OH levels.
(**A**) Rising OH and RO_2 radical levels were the result of increasing IPN for otherwise constant reactant concentrations, [NO] = 1.0×10^{10} and [trimethylamine] = 5.2×10^{11} molecules cm^{-3}.
(**B**) Rising OH levels resulted from constant OH production, i.e., constant IPN and NO concentrations ([IPN] = 1.5×10^{11} and [NO] = 1.0×10^{10} molecules cm^{-3}), and lowering of the main OH consumer trimethylamine. In both experiments, OH radicals were produced from IPN photolysis in air whereby the OH generation finally proceeded via NO + $HO_2 \rightarrow$ OH + NO_2. Product signals were measured with a standard deviation of <20%. Stated OH concentrations were taken from modeling with an assumed uncertainty of a factor of 2.
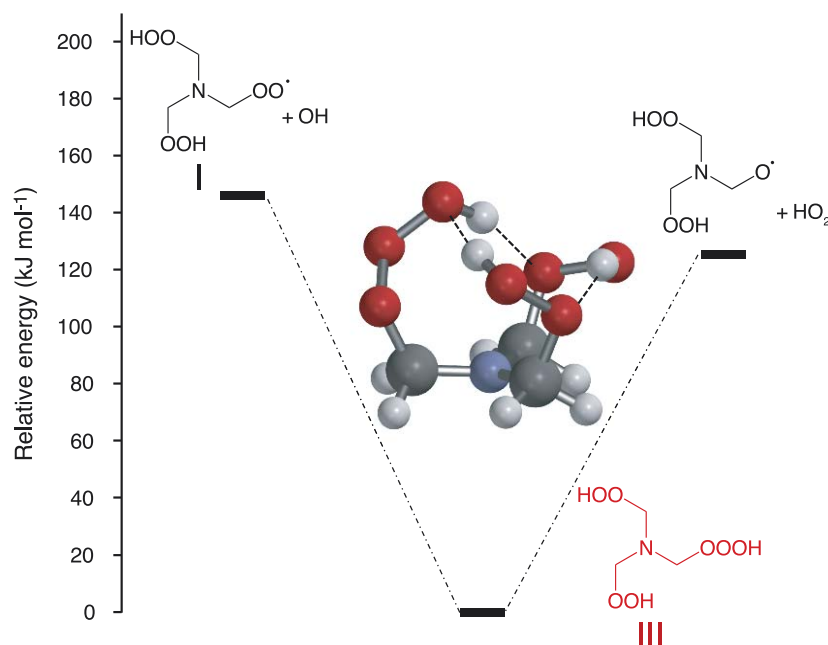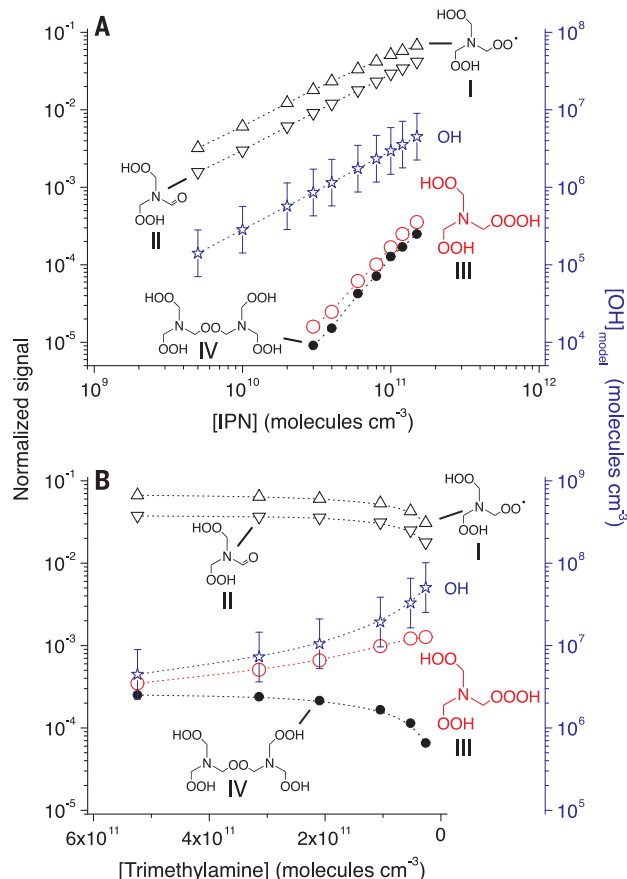
Fig. 3. Energy diagram for the formation of hydrotrioxide ($HOOCH_2)_2NCH_2OOOH$ (III). Its formation from ($HOOCH_2)_2NCH_2OO$ (**I**) + OH was exothermic by 146 kJ mol^{-1}, and its decomposition into the corresponding alkoxy radical + HO_2 was endothermic by 126 kJ mol^{-1}. The lowest-energy conformer of the hydrotrioxide is shown, and dashed lines illustrate the three hydrogen bonds. The zero-point vibrational energy–corrected electronic energies were calculated at the UCCSD(T)-F12a/cc-pVDZ-F12// M06-2X/aug-cc-pVTZ level.

For further mechanistic validation of ROOOH formation, we increased the OH level in the experiment by lowering the trimethylamine concentration for a constant OH production rate using unchanged IPN and nitric oxide (NO) in the photolysis (Fig. 2B). The OH level grew as a result of the decreasing OH loss rate with the lowering of the main OH consumer trimethylamine. At sufficiently small trimethylamine concentrations, OH started to react substantially with CO, methane, and other trace gases in the air, resulting in clear weakening of the RO_2 radical **I** production (Fig. 2B). Simultaneously, the signals of the RO_2 isomerization product **II** and the accretion product **IV** declined. By contrast, the signal attributed to the hydrotrioxide **III** showed a clear increase with increasing OH levels, which emphasizes that **III** has to be formed in a second OH reaction subsequent to the initial OH + trimethylamine reaction that forms **I**. Additional tests revealed that the hydrotrioxide formation was not influenced by interfering processes during product ionization and photolysis (fig. S2), or by the presence of water vapor (fig. S3), or by elevated NO concentrations (fig. S4). Using modeled OH concentrations (supplementary materials, section S1.5), we found a linear dependence of signal(ROOOH) versus signal(RO_2) × [OH]$_{model}$ in accord with the expected formation mechanism of ROOOH (fig. S5). OH radical measurements in amine systems were not possible with our technique.

The same product formation from OH + trimethylamine, including hydrotrioxide production, was also measured using nitrate as the reagent ion (fig. S6). H/D exchange experiments in the presence of heavy water (8, 15) (to determine the number of weakly bound, exchangeable H atoms) showed a signal shift by 3 mass units in the mass spectrum, in accord with the presence of two OOH groups and the OOOH group further supporting the structure of **III** (fig. S7). All experimental findings were consistent with the formation of the hydrotrioxide **III** through a RO_2 + OH reaction.

In Fig. 3, we show the calculated energy diagram of reactions 1 and 2 for the observed hydrotrioxide ($HOOCH_2)_2NCH_2OOOH$). The decomposition of the hydrotrioxide leading to the alkoxy radical and HO_2 was ~20 kJ mol^{-1} lower in energy than that leading to the RO_2 radical and OH, which is in overall agreement with previous calculations performed on small C_1 to C_4 hydrotrioxides (4, 6). Formation of three strong internal hydrogen bonds in ($HOOCH_2)_2NCH_2OOOH$ increased its thermostability compared with that of less-functionalized hydrotrioxides (supplementary materials, section S4.2).

ROOOH from isoprene oxidation

ROOOH formation was also probed in the reaction of OH radicals with isoprene (C_5H_8), one

Fig. 4. Product formation and ROOOH formation kinetics in the OH + isoprene reaction. (**A**) Product concentrations were obtained using calibration factors from former investigations (supplementary materials, section S1.3). Corresponding error bars have been determined from the uncertainty in the calibration factors. OH concentrations were derived from detection of SO_3 formed via OH + SO_2. Increasing OH concentrations (blue stars) for constant OH production in the IPN photolysis ([IPN] = 2.15×10^{11} and [NO] = 1.0×10^{10} molecules cm^{-3}) were the result of lowering of the OH loss rate by reducing the main consumer isoprene. Iodide was used as the reagent ion. The uncertainty of OH concentrations has been estimated to be ~30% considering the uncertainties of the SO_3 calibration and in k(OH + SO_2). (**B**) Data were taken from the experiments depicted in (A). Deduced rate coefficients of HO-C_5H_8OOOH and HO-$C_5H_8(O_2)OOOH$ formation from RO_2 + OH were 5.1×10^{-11} and 1.1×10^{-10} cm^3 molecule^{-1} s^{-1}, respectively. Error bars are not shown for clarity.
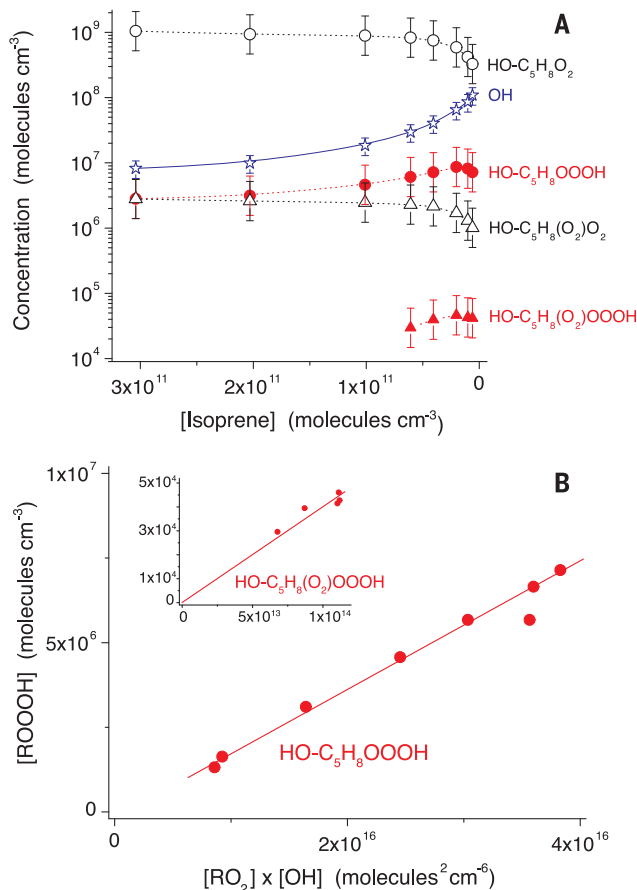
of the most important nonmethane hydrocarbons in the atmosphere (fig. S8) (*16*). In this case, different isomeric RO_2 radicals were formed as a result of isoprene's structure and the different positions of OH and O_2 addition as well as the RO_2 interconversion in this system (*17*, *18*). Moreover, RO_2 autoxidation led to a suite of different oxidized RO_2 radicals HO-$C_5H_8(O_2)_xO_2$, where $x = 0$, 1, and 2 (*9*, *12*, *17*–*19*). In Fig. 4A, we show the results from an experiment with rising OH level resulting from the lowering of isoprene in the reaction gas for constant photolysis conditions. The OH concentration was determined indirectly by monitoring SO_3 formation from OH + SO_2. The small SO_2 addition did not disturb the OH + isoprene reaction (fig. S9). Signals consistent with hydrotrioxides HO-$C_5H_8(O_2)_xOOOH$, where $x = 0$ and 1, emerged beside those from the corresponding RO_2 radicals. Both substance classes showed the expected behavior, i.e., the RO_2 radical concentrations decreased with the lowering of the isoprene concentration, and the ROOOH compounds increased proportional to the product [RO_2] × [OH]. Experiments with rising OH and RO_2 radical concentrations,

as a result of increasing IPN concentrations, confirmed these findings (fig. S10).

The ROOOH formation kinetics were assessed on the basis of the measured ROOOH, RO_2, and OH radical concentrations (Fig. 4B). From the linear dependence of [ROOOH] versus [RO_2] × [OH] in the case of HO-C_5H_8OOOH, a rate coefficient of k(HO-$C_5H_8O_2$ + OH) = 5.1×10^{-11} cm^3 molecule^{-1} s^{-1} follows. For the higher oxidized HO-$C_5H_8(O_2)OOOH$, the analysis yields k(HO-$C_5H_8(O_2)O_2$ + OH) = 1.1×10^{-10} cm^3 molecule^{-1} s^{-1}, assuming a regression line through zero (Fig. 4B, inset). These rate coefficients have an uncertainty of a factor of 3 to 4. Previously, high rate coefficients of RO_2 + OH reactions were reported for C_1 to C_4 RO_2 radicals through detection of the OH decay, e.g., k($C_4H_9O_2$ + OH) = (1.5 ± 0.3) × 10^{-10} cm^3 molecule^{-1} s^{-1} at 298 K (*20*), which supports our findings.

Universal ROOOH formation

We tested the general validity of hydrotrioxide formation from RO_2 + OH reactions, especially for RO_2 radicals with atmospheric relevance. In the OH radical–initiated oxidation of dimethyl sulfide (DMS), α-pinene, toluene, and

1-butene, the corresponding hydrotrioxide formation from the principal RO_2 radicals in the respective reaction system was clearly detectable in the flow experiment (figs. S11 to S18). Data analysis revealed a linear dependence of signal(ROOOH) versus signal(RO_2) × [OH] for each system (figs. S19 to S22). In the cases where RO_2 radicals and ROOOH could be detected with close-to-maximum sensitivity, the rate coefficients k(RO_2 + OH) were estimated (figs. S5 and S20), further supporting that RO_2 + OH reactions proceed with rate coefficients close to the collision limit. Table S2 summarizes the rate coefficients determined in this study. Moreover, signals consistent with hydrotrioxide formation from the reaction of OH radicals with 2-methylpropene were observed in separate experiments conducted in a 1-m^3 Teflon (fluorinated ethylene propylene) chamber in air using CF_3O^- chemical ionization mass spectrometry (supplementary materials, section S1.7, and fig. S23). Consequently, it could be inferred that RO_2 + OH reactions represent a universal pathway of hydrotrioxide formation under atmospheric conditions.

Atmospheric perspective of ROOOH

The hydroxy hydrotrioxide formed in the OH + 2-methylpropene reaction, as observed in the environmental chamber experiment (fig. S23), suggested an experimental ROOOH lifetime of ~20 min at 296 K including thermal gas-phase decomposition and losses on the chamber wall. Thus, 20 min can be regarded as a lower bound of its thermal lifetime. Theoretical calculations on the thermal decomposition energies (Fig. 3) for several hydrotrioxides pointed to similar or longer thermal lifetimes (supplementary materials, section S4.2). In addition, the calculations did not indicate any fast photolysis pathways for ROOOH (supplementary materials, section S4.3).

We estimated the atmospheric lifetime of hydrotrioxides against the OH reaction to be ~2 hours to a few days assuming [OH] = (5 to 20) × 10^5 molecules cm^{-3} and k(OH + ROOOH) = (1 to 7.5) × 10^{-11} cm^3 molecule^{-1} s^{-1}, relevant for saturated and unsaturated hydrotrioxides assuming an OH reactivity similar to that of the corresponding hydroperoxides (*21*, *22*). Thus, hydrotrioxides, once formed, would be present in the atmosphere for minutes to hours before further processing. Theoretical calculations favored the formation of alkoxy radicals RO from the reaction of OH radicals with saturated ROOOH. In the isoprene system, however, OH + HO-C_5H_8OOOH mainly forms the dihydroxy epoxide IEPOX (*18*) and HO_2 (supplementary materials, section S4.5).

Global simulations with the chemistry climate model ECHAM-HAMMOZ (*23*) permitted an assessment of ROOOH production from

the OH radical–initiated oxidation of isoprene (figs. S24 and S25). We lumped all isoprene-derived RO_2 radicals and calculated that up to 1% of these could react with OH, forming ROOOH. Given the large emission of isoprene (16), an annual ROOOH production of up to ~10 million metric tons is calculated. Furthermore, the modeling revealed that isoprene-derived hydrotrioxides can reach atmospheric concentrations of ~10^7 molecules cm^{-3} (fig. S24).

From the knowledge in preparative chemistry (1–3), we deduced that hydrotrioxides could act as oxidants. In the atmosphere, this reactivity could manifest itself in surface reactions after lung inhalation and reactions on and within aerosol particles forming $^1O_2$. Hydrotrioxides generated from highly oxidized RO_2 radicals represent highly oxygenated organic molecules (HOMs) (24) with a very large oxygen content, e.g., HO-$C_{10}H_{16}(O_2)_2OOOH$ (fig. S14). These ROOOH-HOMs would be relevant for atmospheric aerosol formation and thus for Earth's radiation budget. Further research is needed to ascertain the role of hydrotrioxides for health and the environment. We illustrated the direct observation of hydrotrioxides using mass spectrometry, which should open up opportunities to measure these compounds in different systems, including in the atmosphere, after further optimization of the analytical techniques.

REFERENCES AND NOTES

1. F. E. Stary, D. E. Emge, R. W. Murray, *J. Am. Chem. Soc.* **98**, 1880–1884 (1976).
2. M. Zarth, A. De Meijere, *Chem. Ber.* **118**, 2429–2449 (1985).
3. G. H. Posner, K. S. Webb, W. M. Nelson, T. Kishimoto, H. H. Seliger, *J. Org. Chem.* **54**, 3252–3254 (1989).
4. J. F. Müller et al., *Nat. Commun.* **7**, 13213 (2016).
5. Y. Liu et al., *Chem. Res. Chin. Univ.* **33**, 623–630 (2017).
6. E. Assaf, C. Schoemaecker, L. Vereecken, C. Fittschen, *Int. J. Chem. Kinet.* **50**, 670–680 (2018).
7. A. T. Archibald, A. S. Petit, C. J. Percival, J. N. Harvey, D. E. Shallcross, *Atmos. Sci. Lett.* **10**, 102–108 (2009).
8. T. Berndt et al., *Nat. Commun.* **7**, 13677 (2016).
9. T. Berndt, N. Hyttinen, H. Herrmann, A. Hansel, *Commun. Chem.* **2**, 21 (2019).
10. K. H. Møller, T. Berndt, H. G. Kjaergaard, *Environ. Sci. Technol.* **54**, 11087–11099 (2020).
11. E. Praske et al., *Proc. Natl. Acad. Sci. U.S.A.* **115**, 64–69 (2018).
12. K. H. Møller, K. H. Bates, H. G. Kjaergaard, *J. Phys. Chem. A* **123**, 920–932 (2019).
13. J. D. Crounse, L. B. Nielsen, S. Jørgensen, H. G. Kjaergaard, P. O. Wennberg, *J. Phys. Chem. Lett.* **4**, 3513–3520 (2013).
14. T. Berndt, K. H. Møller, H. Herrmann, H. G. Kjaergaard, *J. Phys. Chem. A* **125**, 4454–4466 (2021).
15. M. P. Rissanen et al., *J. Am. Chem. Soc.* **136**, 15596–15606 (2014).
16. K. Sindelarova et al., *Atmos. Chem. Phys.* **14**, 9317–9341 (2014).
17. J. Peeters, T. L. Nguyen, L. Vereecken, *Phys. Chem. Chem. Phys.* **11**, 5935–5939 (2009).
18. P. O. Wennberg et al., *Chem. Rev.* **118**, 3337–3390 (2018).
19. S. Wang, M. Riva, C. Yan, M. Ehn, L. Wang, *Environ. Sci. Technol.* **52**, 12255–12264 (2018).
20. E. Assaf, S. Tanaka, Y. Kajii, C. Schoemaecker, C. Fittschen, *Chem. Phys. Lett.* **684**, 245–249 (2017).
21. J. M. St. Clair et al., *J. Phys. Chem. A* **120**, 1441–1451 (2016).
22. C. Fittschen et al., *Atmos. Chem. Phys.* **19**, 349–362 (2019).
23. M. G. Schultz et al., *Geosci. Model Dev.* **11**, 1695–1723 (2018).
24. F. Bianchi et al., *Chem. Rev.* **119**, 3472–3509 (2019).
25. J. Chen, K. H. Møller, E. R. Kjaergaard, H. G. Kjaergaard, ROOOH-archive, Electronic Research Data Archive (University of Copenhagen, 2022); https://doi.org/10.17894/UCPH.99BF9F39-EC48-4AFE-B426-CF6EF0D3EA70.
26. E. H. Hoffmann, A. Tilgner, H. Herrmann, ROOOH modeling data — archive, Zenodo (2022); https://doi.org/10.5281/zenodo.6373045.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abn6012
Materials and Methods
Figs. S1 to S47
Tables S1 to S10
References (27–67)

CATALYSIS

Dynamic interplay between metal nanoparticles and oxide support under redox conditions

H. Frey[1,2]†, A. Beck[2,3]†, X. Huang[1,4]*, J. A. van Bokhoven[2,3]*, M. G. Willinger[1]*‡

The dynamic interactions between noble metal particles and reducible metal-oxide supports can depend on redox reactions with ambient gases. Transmission electron microscopy revealed that the strong metal-support interaction (SMSI)–induced encapsulation of platinum particles on titania observed under reducing conditions is lost once the system is exposed to a redox-reactive environment containing oxygen and hydrogen at a total pressure of ~1 bar. Destabilization of the metal–oxide interface and redox-mediated reconstructions of titania lead to particle dynamics and directed particle migration that depend on nanoparticle orientation. A static encapsulated SMSI state was reestablished when switching back to purely oxidizing conditions. This work highlights the difference between reactive and nonreactive states and demonstrates that manifestations of the metal-support interaction strongly depend on the chemical environment.

Noble metal nanoparticles (NPs) are used as active constituents of catalysts and thus play an important role in the sustainable production of chemicals and fuels and the mitigation of pollutants. The interesting catalytic properties of noble metals emerge when dispersed as nanometer-sized particles on high–surface area oxide supports. Although simple models attribute catalytic properties mainly to size effects, with the metal being the active phase on an inert oxide support, synergistic interactions between metal and support can be critical (1, 2).

A prominent example of metal-support interactions is the loss of hydrogen chemisorption of titania-supported platinum (Pt–TiO$_2$) upon high-temperature reductive treatment (HTR). The effect, which was first observed by Tauster et al. (3), can be reverted through a series of high-temperature oxidation and subsequent low-temperature reduction steps (3). Originally, the chemisorption suppression was attributed to an electronic perturbation of the system, that is, the bonding between Pt and Ti cations under reducing atmospheres. Subsequent studies showed that in situ reductive activation leads to encapsulation of Pt NPs in a thin, partially reduced layer of TiO$_2$ (4–10). The driving force for this so-called strong metal-support interaction (SMSI) state was attributed to surface energy minimization (11).

Similar encapsulation effects have also been reported for Co (12), Ni (13, 14), Au (15), and Cu

(16) NPs. More recently, it was shown that high-temperature oxidative treatment can also lead to NP encapsulation (15, 17–20). Potential benefits of SMSI encapsulation include enhanced resistivity against both NP coalescence and Ostwald ripening (21–23). Furthermore, the selectivity of the catalyst can be tuned by modifying the adsorption strength of molecules on the catalyst surface and by blocking specific active sites (24–28). Hence, the ability to exploit synergistic interactions between metal NPs and their support in a controlled way is of high interest with regard to the development of improved catalysts and processes.

Early studies on the SMSI were mostly based on indirect, integral spectroscopic observations and provided evidence of NP encapsulation either through an altered chemical composition of the surface or through SMSI-induced changes in the chemisorption capacity. High-resolution real-space methods, such as scanning tunneling microscopy (STM) and transmission electron microscopy (TEM), can provide direct atomic-scale imaging to confirm the formation of the encapsulating SMSI state after high-temperature reductive (29–31) or, respectively, oxidizing treatment (4). However, because of methodological constraints, direct imaging of the encapsulated state has mostly been achieved ex situ, in experiments in which the SMSI state was preserved during cooling and sample transfer from the catalytic reactor to the high-vacuum environment of a microscope. Despite detailed characterization of preserved SMSI states, little is known about metal-support interactions under catalytic reaction conditions.

Because the chemical states of metal particles and reducible support are influenced by the chemical environment (32, 33), their mutual interaction and possible synergistic effects should be investigated under catalytic working conditions. Direct real-space observations by in situ STM and environmental TEM have shown gas-phase–induced reconstruc-

tion of NPs (34–37) and changes at the metal-support interface (38). Environmental TEM is generally limited to chamber pressures of ~20 mbar (39). Depending on the system studied, the pressure-dependent chemical potential of reactive species might thus be too low to trigger specific processes and reactions that are relevant for catalytic function at higher pressure. However, the development and commercial availability of microelectromechanical-based reactors for in situ electron microscopy has extended the accessible pressure range by roughly two orders of magnitude (11, 40, 41) and enabled partially bridging the so-called pressure gap (42).

This study builds on earlier investigations of interfacial dynamics on oxide-supported noble metal NPs (38, 43, 44). Working at higher pressure allows us to visualize gas phase–induced processes that are directly linked to metal-support interactions and revealed the dynamic interplay between NP, support, and gas phase under working conditions. We selected a Pt–TiO$_2$ catalyst because it is the archetype system for which the SMSI state was first described (3) and for which we have recently shown that a static encapsulated state of Pt NPs exists not only under hydrogen but also under purely oxidizing conditions, however, with a specific overlayer structure that depends on the gas environment (4). We studied hydrogen oxidation because it is the most elementary redox reaction that can also be viewed as a representation of many reactions, such as partial hydrocarbon oxidations. Furthermore, this catalytic system only yields water in gas phase reactions at high temperatures (45). Our in situ TEM-based study revealed how the classical encapsulated SMSI state of Pt can be lost through destabilization of the overlayer once the system is exposed to a redox-active regime in which H$_2$ and O$_2$ are simultaneously interacting with the catalyst.

Under reaction conditions, we observed interfacial dynamics that are characterized by local structural collapse and rebuilding of TiO$_2$, i.e., redox processes involving the reduction and subsequent reoxidation of the support underneath Pt NPs. In this process, the inherent lattice mismatch between Pt and TiO$_2$ and the associated interfacial strain lowers the barrier for vacancy formation in the reducible oxide. The resulting interfacial reconstructions give rise to pronounced changes in particle morphology and, eventually, particle mobility. Previous studies in which purely reducing or oxidizing environments were used did not show such dynamics and were not representative of reaction conditions (4, 11, 46, 47). Water, which is formed as a reaction product, is not the cause of the observed phenomena. Instead, addition of water to the feed suppresses particle dynamics. Relevant for the NP behavior is the configuration of the metal-support interface, as

[1]Scientific Center of Optical and Electron Microscopy (ScopeM), ETH Zürich, 8093 Zürich, Switzerland. [2]Institute for Chemical and Bioengineering, ETH Zürich, 8093 Zürich, Switzerland. [3]Paul Scherrer Institute, 5232 Villigen, Switzerland. [4]College of Chemistry, Fuzhou University, Fuzhou 350116, P. R. China.
*Corresponding author. Email: xinghuang@fzu.edu.cn (X.H.); jeroen.vanbokhoven@chem.ethz.ch (J.A.vB.); marc.willinger@tum.de (M.G.W.)
†These authors contributed equally to this work.
‡Department of Chemistry, Technical University of Munich, Lichtenbergstrasse 4, D-85748 Garching, Germany.

shown for three selected cases of differently oriented NPs. This work provides an explanation for the observed particle restructuring and directed particle migration on the support. Furthermore, it provides additional evidence for a nonclassical SMSI state that is observed in oxidizing conditions (*17, 20, 48*).

Results

Switching to a redox-active H_2–O_2 mixture

We investigated the metal-support interactions and the relevance of an SMSI encapsulated state under redox conditions. TiO_2-supported Pt NPs were first heated in H_2 to induce the classical SMSI state and then transferred under inert gas purging into an O_2 atmosphere. As described in (*4*) and detailed in the experimental part (supporting materials and methods), this treatment led to Pt NPs in a nonclassical oxidized SMSI state. Once prepared, the system was transited into the relevant redox-reactive regime through addition of H_2 into the O_2 flow. This sequence led to a gradual change in the encapsulated state of the Pt NPs and, finally, complete removal of the overlayer (Fig. 1, A to F). The first observable effect that could be attributed to an increasing partial pressure of H_2 in the reactor was the onset of overlayer reduction. It was detected as an instant change in the overlayer structure on the (001) facet (movie S1), followed by retraction of the overlayer on the (111) plane (Fig. 1). The latter initiated near an uncovered Pt{110}-type microfacet, as indicated by the propagating reduction front in the image sequence of Fig. 1, G to I (see also movie S2). Within seconds, the overlayer vanished from the Pt{111} surfaces, whereas some transient patches, which selectively decorated Pt{100} planes, could still be observed (Fig. 1, C and D, and movies S1 and S2). Similar transiently existing nanopatches were observed for the Pt–TiO_2 system in field ion microscopy (*49*) and were reported for iron oxide–supported Pt NPs (*50*).

Once the gas composition reached a set mixture (60 mbar H_2 and 700 mbar O_2) after ~180 s, the encapsulating overgrowth layer was fully retracted from all particles. Thus, stable configurations of static Pt particles exhibiting encapsulating layers existed either in pure H_2 (the classical SMSI state) or in pure O_2 (the nonclassical SMSI state), but not in a regime in which both gases were simultaneously present. Signs of the structural incompatibility between the reduced and oxidized overlayers on the Pt particles could also be detected at the remaining NP-support interface (see next section). With the removal of the overlayer, Pt NPs furthermore underwent a shape change through a slight expansion of {100} facets (see Fig. 1, A, G, and F, and structure model in fig. S3). Once the overlayer was fully removed, the onset of pronounced parti-

cle dynamics involving restructuring and migration was observed (see image sequence in Fig. 1, E and F, and movie S1).

Particle and interfacial dynamics in the redox-active regime

Movies recorded at lower magnification showed the response of a collection of NPs to reaction conditions. As shown in movie S3, the degree of structural dynamics and mobility differed between NPs. Some NPs remained static, whereas others underwent structural fluctuations; some remained stationary, whereas others migrated across the substrate surface. These individual dynamics indicate that each NP responded according to local surface topological features and the configuration of the interface.

Because rutile TiO_2 preferentially exposes low-energy (110) facets (*51*) and Pt NPs were generally present in the form of truncated cuboctahedra, exposing mostly {100} and {111} facets, only a limited number of interface configurations need to be considered for a general description of the observed behavior. Indeed, a preferential orientation relationship between the Pt NP and the support is evident from the analysis of lattice fringes of isolated Pt NPs that are attached to the same TiO_2 particle. As shown in Fig. 2E and the corresponding co-

lorized lattice fringes in the Fourier-filtered image in Fig. 2F and fig. S4, the particles show identical orientation of their (111) and (100) planes, and thus, preferential orientation driven by a minimization of the lattice misfit induced interfacial strain.

The first NP we considered (Fig. 2, A to C) was oriented with Pt(111) planes perpendicular to the Pt–TiO_2 interface. Images that were recorded after overlayer retraction showed occasional slight rotations of the NP. Such slight rotations could be induced by reconstructions at the interface, which is similar to the recently reported case of gas phase–induced rotation of gold NPs on a TiO_2 support, which was observed by low-pressure environmental TEM (*43*). With increasing H_2 partial pressure, the Pt NP developed pronounced structural dynamics that involved twin formation and shearing along Pt(111) planes in an up-down motion, perpendicular to the Pt–TiO_2 interface (see arrows in Fig. 2C). This up-down motion can be seen in Fig. 2D, which was generated by cutting a plane through the recorded image stack along the time axis, and is also shown in movie S4.

These structural dynamics at the interface imply that the TiO_2 substrate locally collapsed and at a later point in time was rebuilt. The
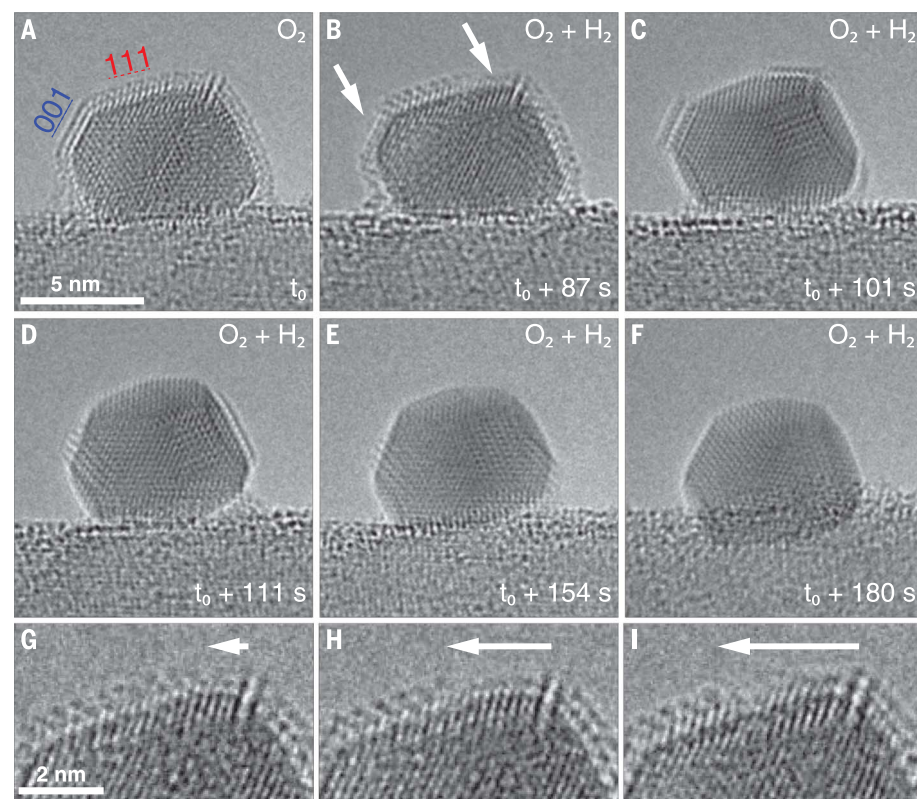
Fig. 1. Morphological change upon transition into the redox-active regime. (**A** to **F**) Images that were recorded while the composition of the gas phase in the in situ gas cell gradually changed from 700 mbar O_2 to a mixture of 60 mbar H_2 plus 700 mbar O_2. (**G** to **I**) Magnified images of the particle surface that show the gradual reduction of the overlayer between frames A and B. t_0 is the time at which the H_2 flow was turned on.

cross-sectional view provided by TEM imaging showed a substantial reshuffling of material at the interface, with Pt (111) planes moving up and down in response. Characteristic for this process is the pronounced formation of (111)-twinning planes, which has also been observed for the case of zinc oxide–supported copper NPs in a redox environment (*52*).

The second case, shown in Fig. 2, G to J, and movie S5, represents NPs in which Pt(111) planes were oriented almost parallel to the Pt–TiO$_2$ interface. Here, a repetitive forward-and-backward step flow–like motion of Pt{111} planes was observed. Similar to the case of Pt(111)-planes moving up- and downwards in Fig. 2, A to C, this motion could be caused by similar redox processes, however, rotated by 90° and without the involvement of twinning planes running through the Pt NP.

The third case we considered was a Pt NP oriented with a {001} plane parallel to the Pt–TiO$_2$ interface. Because TEM images only show a two-dimensional (2D) projection of a 3D object, the precise location of the interface was not clear in Fig. 2, K to N. Nevertheless, we could follow the shape evolution of the moving NP with time and abstract information about the ongoing processes. The NP exposed Pt(111) planes that were inclined with respect to the interface. At times, microfaceting of the (111) plane was observed at the right side of the particle, which, in effect, tilted it down toward the substrate (see Fig. 2K). Moments later, the right side reconstructed until a planar (111) facet was restored. This downward inclination and subsequent retraction occurred repetitively, whereas the opposing (111) plane on the left side, which faced away from the substrate, did not show any change. The reshuffling of Pt at one end was responsible for a net transport of Pt from the right (back side) to the left (front side) and resulted in a propagating motion. This redox chemistry–driven directional migration of the Pt NP is shown in movie S6.

These three cases show how the relative orientation of Pt NP and the TiO$_2$ support can be linked to the behavior of individual NPs. Depending on the configuration of the interface, the underlying redox processes can give rise either to Pt NPs that restructure and do not move or to NPs that restructure and migrate on the surface in a directed manner.

Retraction of H$_2$ and reformation of the oxidic SMSI overlayer

When we switched the gas composition back from a reactive to a purely oxidizing regime by turning off the H$_2$ flow, an encapsulated state of Pt NPs was reestablished. The first apparent effect of a reduced H$_2$ concentration was a sudden morphological change of the Pt NPs toward a more spherical shape, typical for the effect of O$_2$ (*33*) (see Fig. 3, A to C, and movie S7). Subsequently, the migration of support material onto the Pt NPs, and thus reformation of the overgrowth layer, was observed. The image sequence in Fig. 3, D to F, showed that NP coverage started at the Pt–TiO$_2$ interface and propagated upward. The support underneath showed a weakening of image contrast in the vicinity of the Pt NP (see Fig. 3D), indicating that the material forming the overlayer originated from there. Electron energy loss spectroscopy measurements confirmed that the overlayer consisted of TiO$_2$ (fig. S5).

Once the coverage was restored, all of the dynamical effects ceased, and the system reached a static state. Because water was formed as reaction product and could influence the above-described particle dynamics, further experiments were performed in which water vapor was co-fed first to an O$_2$-containing atmosphere [up to a partial pressure of water vapor (p_{H_2O}) ~20 mbar; supporting materials and methods] and subsequently to a mixture of O$_2$ (700 mbar) and H$_2$ (120 mbar) at 600°C. Movie S8 showed that neither particle dynamics nor migration was observed as a consequence of the added water vapor. Only after retracting water from the feed gas did particle dynamics reemerge (movie S8 and fig. S6).

Discussion

This work was motivated by questions regarding the relevance of the classical SMSI state
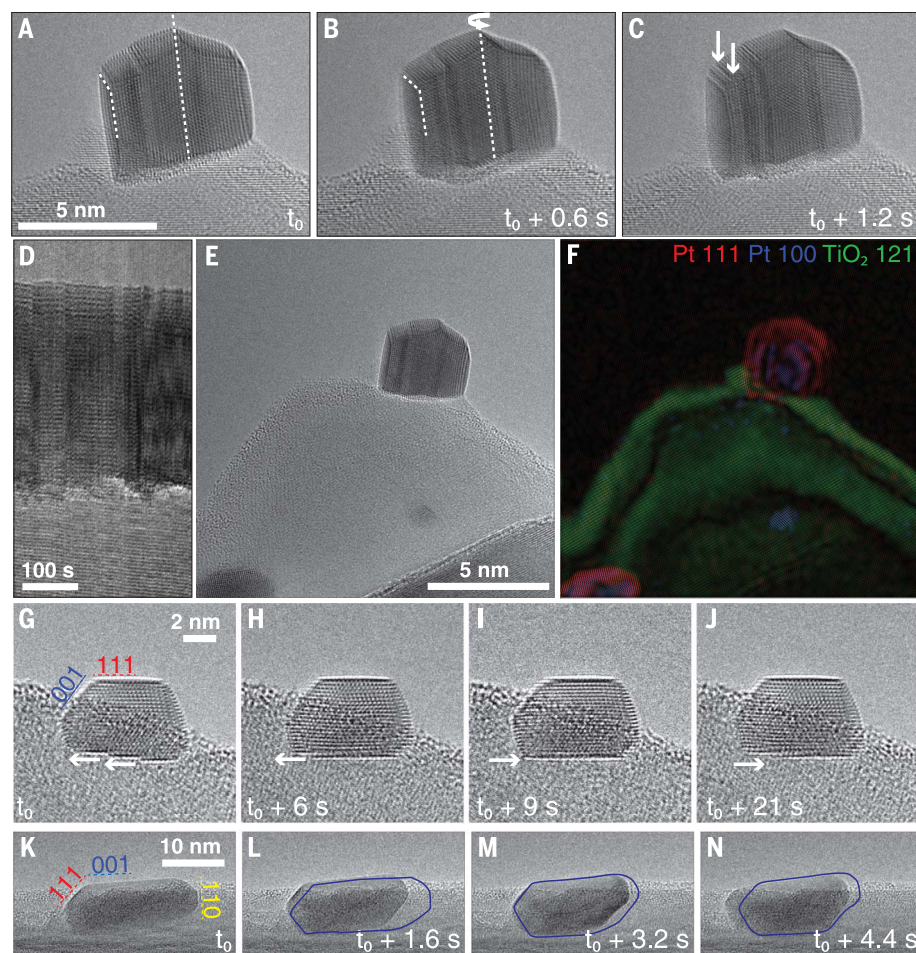
Fig. 2. Redox chemistry at the interface induces structural dynamics of Pt NPs and is the driving force for particle reconstruction and migration. (**A** to **C**) Individual frames of a movie recorded from a Pt NP that is oriented with (111) planes perpendicular to the interface. (**D**) View generated from the same movie by cutting a plane through the image stack along the time axis. It shows the up- and downward shifting of Pt(111) planes and the associated collapse and reconstruction of the underlaying TiO$_2$. (**E** and **F**) Image and its Fourier-filtered counterpart in which planes of identical orientation appear in the same color: Green indicates (121) planes of TiO$_2$ and red and blue indicate Pt(111) and Pt(100) planes, respectively. (**G** to **I**) Image sequence of a particle that is seemingly oriented with the Pt planes parallel to the TiO$_2$ interface. (**K** to **N**) NP that has its Pt{111} planes inclined toward the interface. The blue shapes indicate the respective positions of the Pt NP in the previous frames.

beyond reductive activation and thus the behavior of the system under reactive conditions. Encapsulation is often used to explain the altered behavior of supported metal NPs, although it represents a state obtained after HTR or, as recently shown, treatment in O_2. Little is known about the SMSI state under reaction conditions and its relation to catalytic function.

Starting with encapsulated NPs in an oxidizing atmosphere, the key finding of this work is that exposure to a redox-active environment led to the removal of the overlayer and subsequently the emergence of particle dynamics. Thus, the classical SMSI state was lost as soon as the surface was exposed to a reactive atmosphere. The absence of an overlayer indicated that neither the reduced TiO_2 overlayer found in H_2 nor the oxidized version found in O_2 was stable under redox conditions. The simultaneous presence of reducing and oxidizing agents induced redox processes that lead to destabilization and overlayer retraction. Hydrogen is easily activated on platinum, but activation on TiO_2 is much more unlikely (53–55). Its addition to the feed gas triggered oxygen abstraction from the overlayer, thereby destabilizing and finally stripping it from the particle surface.

The areal increase of Pt{100} facets upon overlayer retraction suggested that H_2, despite a presumably low equilibrium coverage at 600°C and 60 mbar (56), dominated the NP shape. Indirect evidence for activation of H_2 on Pt is provided by Fig. 1, G to I, and in movie S2, in which overlayer retraction was initiated at a kink in the NP and propagated from there across the Pt surface. A similar mechanism was observed for the reduction of an O-Fe-O trilayer on an extended Pt{111} surface during CO oxidation (57).

Redox-induced processes not only destabilized the encapsulating layer but also acted at the remaining Pt–TiO_2 interface and led to pronounced particle dynamics. Structural incoherence of Pt and TiO_2 (58) created inhomogeneous strain at the NP-support interface (59). The strain was modulated with a periodicity of the moiré or coincidence lattice (60). Strain modulation locally reduced the energetic costs for oxygen vacancy formation in the reducible support (60, 61). Experimentally, similar strain-facilitated introduction of oxygen vacancies at the NP-support interface under reducing conditions was observed for a cerium oxide–rhodium model catalyst (60) and a gold ceria interface (62), and furthermore was predicted by density functional theory (DFT) for Pt on TiO_2 (63).

In rutile, aggregation of oxygen vacancies is energetically favored and leads to shear plane formation. These so-called Wadsley defects (64–68) can form already at low oxygen deficiencies (69). The introduction of shear planes
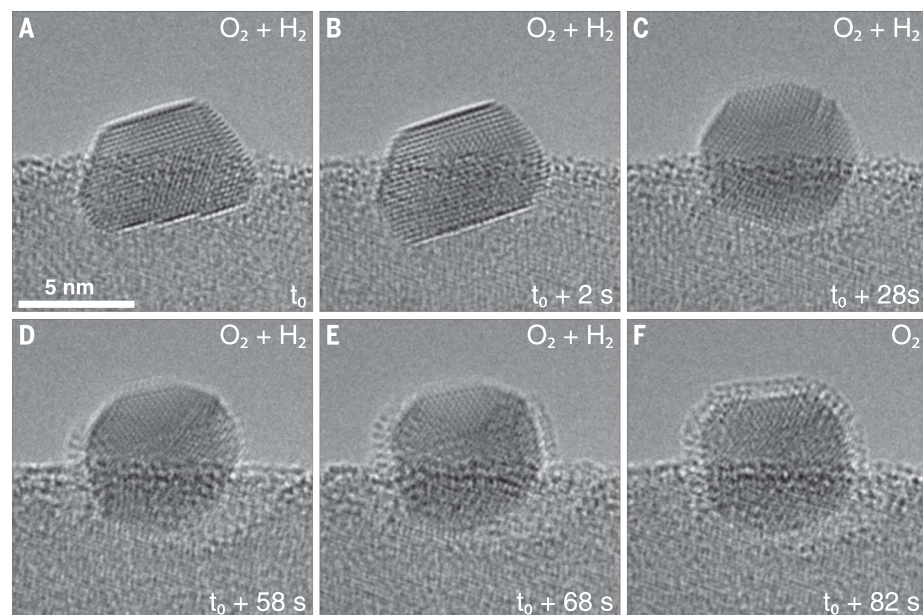
Fig. 3. Morphological change of a Pt NP upon leaving the redox-active regime. Switching of the gas atmosphere from 60 mbar H_2 plus 700 mbar O_2 back to 700 mbar O_2 at 600°C leads to reformation of a classical particle overgrowth. (**A** to **C**) The NP first adopted a spherical morphology. (**D** to **F**) As soon as H_2 was fully removed from the reactor cell, the overlayer reformed from support material from the vicinity of the NP; t_0 is the time at which the H_2 flow was set to zero.

reduces the elastic energy density stored in the interface (59) and prompts a morphology adaptation of the supported particle. With increasing extent of reduction, the local affinity of TiO_x toward reoxidation increases (70) up to the point at which reoxidation takes place and induces, once again, morphological adaptation of the Pt NPs. This redox-mediated reconstruction can be seen in movie S4, which shows the periodic collapse and rebuilding of the TiO_2 structure underneath a Pt NP. In a simplified model, the process can be viewed as an oscillator in which energy is transferred back and forth between strain or misfit and chemical energy. The process is driven by competing actions of H_2 and O_2 and enabled by a sufficiently high temperature and chemical potential of the constituent gases. Such atmosphere-induced reconstructions have been revealed by high-resolution transmission electron microscopy (HRTEM) in dilute atmospheres and were reported for the case of gold on gamma-iron oxide, gold on TiO_2, Au on CeO_2, and Pt on Fe_3O_4 (37, 43, 71, 72). In all of these cases, a change in interface configuration, expressed as either a rotational movement of the particle relative to the interface (43) or a stepwise translational motion (37), was detected.

Our work bridges the pressure gap from the low-pressure, quasistatic regime to a regime in which the chemical potential of the involved species (H_2 and O_2) is sufficient to induce more substantial redox dynamics. At the applied pressure, activated hydrogen can easily

remove oxygen from the underlying TiO_2 support (53). Overall, this process resembles the Mars van Krevelen mechanism, as oxygen from the catalyst ends up in the product (73). However, it is not restricted to surface reduction and oxidation. Our in situ observations show that, depending on the relative orientation of particle and support, interfacial reconstructions result in different particle dynamics. The three representative cases presented here show that interface reconstructions promoted the formation of twin boundaries and step-flow growth and retraction at Pt{111} gliding planes and can give rise to directional surface migration of Pt NPs. Under redox conditions, particle dynamics and surface migration are thus largely driven by chemical processes and determined by orientation relationships, surface structure, and topology. In this way, the migration behavior of particles can be rationalized.

The oscillatory behavior that we observed reflected the bistability associated with redox processes and the inability of the system to settle as long as a reactive state was maintained. Water, when present in excess, did not induce any dynamics, but rather quenched them. Water preferably dissociates at oxygen vacancies (74) and, when added in excess, worked against TiO_2 reduction by shifting the redox regime to higher temperatures. When the feed is switched back to dry H_2 plus O_2, particle dynamics re-emerged until finally, the dynamics were stopped when H_2 was removed from the feed. At that moment, a static

overlayer reformed. The system thus equilibrated by reverting to an oxygen-induced SMSI state, which was formed in this case directly, without prior initialization through the well-known high-temperature reduction route. This nonclassical, oxygen-induced SMSI state was recently observed by Tang *et al.* (*20*).

The dynamic behavior described in this work remains unnoticed in ex situ and postmortem studies, highlighting the importance of in situ observations. The complex interplay between metal and support under reaction conditions that was described here for the model reaction of H_2 oxidation shows how synergistic interactions emerge at strained interfaces between metals and reducible supports and provide a means for lowering oxidation barriers. Because these effects are influenced by the local structure of the interface of each individual particle, those insights are even hidden for integral in situ methods that average over a large fraction of a sample. The reported individual behavior and its correlation to structure have to be included in models that aim to describe the overall structural transformations of a catalyst under redox conditions. Theory may then provide insight into how the dynamic behavior and interfacial redox processes lead to the formation and disappearance of active sites that link the restructuring and motion directly to the reaction kinetics.

REFERENCES AND NOTES

1. M. Haruta, *Catal. Today* **36**, 153–166 (1997).
2. T. W. van Deelen, C. Hernández Mejía, K. P. de Jong, *Nat. Catal.* **2**, 955–970 (2019).
3. S. J. Tauster, S. C. Fung, R. L. Garten, *J. Am. Chem. Soc.* **100**, 170–175 (1978).
4. A. Beck *et al.*, *Nat. Commun.* **11**, 3220 (2020).
5. W. Zhao *et al.*, *J. Phys. Chem. C* **125**, 10386–10396 (2021).
6. X. Wang, A. Beck, J. A. van Bokhoven, D. Palagin, *J. Mater. Chem. A Mater. Energy Sustain.* **9**, 4044–4054 (2021).
7. D. N. Belton, Y. M. Sun, J. M. White, *J. Phys. Chem.* **88**, 1690–1695 (1984).
8. C. S. Ko, R. J. Gorte, *J. Catal.* **90**, 59–64 (1984).
9. R. A. Demmin, C. S. Ko, R. J. Gorte, *J. Phys. Chem.* **89**, 1151–1154 (1985).
10. S. Bernal *et al.*, *J. Chem. Soc., Faraday Trans.* **92**, 2799–2809 (1996).
11. S. Zhang *et al.*, *Nano Lett.* **16**, 4528–4534 (2016).
12. J. Lee *et al.*, *J. Catal.* **330**, 19–27 (2015).
13. D. Liu *et al.*, *J. Catal.* **266**, 380–390 (2009).
14. J. Li *et al.*, *ACS Catal.* **9**, 6342–6348 (2019).
15. X. Liu *et al.*, *J. Am. Chem. Soc.* **134**, 10251–10258 (2012).
16. T. Lunkenbein, J. Schumann, M. Behrens, R. Schlögl, M. G. Willinger, *Angew. Chem.* **127**, 4627–4631 (2015).
17. H. Tang *et al.*, *Chem. Sci.* **9**, 6679–6684 (2018).
18. S. Liu *et al.*, *Nat. Commun.* **10**, 5790 (2019).
19. S. Liu *et al.*, *ACS Catal.* **11**, 6081–6090 (2021).
20. M. Tang *et al.*, *Angew. Chem. Int. Ed.* **60**, 22339–22344 (2021).
21. H. Tang *et al.*, *J. Am. Chem. Soc.* **138**, 56–59 (2016).
22. H. Tang *et al.*, *Angew. Chem. Int. Ed.* **55**, 10606–10611 (2016).
23. A. K. Datye, M. Votsmeier, *Nat. Mater.* **20**, 1049–1059 (2021).
24. G. L. Haller, D. E. Resasco, in *Advances in Catalysis* (Elsevier, 1989), vol. 36, pp. 173–235; https://linkinghub.elsevier.com/retrieve/pii/S0360056408600188.
25. A. Corma, P. Serna, P. Concepción, J. J. Calvino, *J. Am. Chem. Soc.* **130**, 8748–8753 (2008).
26. J. C. Matsubu *et al.*, *Nat. Chem.* **9**, 120–127 (2017).
27. R. M. Kennedy *et al.*, *Catal. Lett.* **148**, 2223–2232 (2018).
28. M. Macino *et al.*, *Nat. Catal.* **2**, 873–881 (2019).
29. M. G. Willinger *et al.*, *Angew. Chem. Int. Ed.* **53**, 5998–6001 (2014).
30. S. Shaikhutdinov, *Catal. Lett.* **148**, 2627–2635 (2018).
31. O. Dulub, W. Hebenstreit, U. Diebold, *Phys. Rev. Lett.* **84**, 3646–3649 (2000).
32. P. L. Hansen *et al.*, *Science* **295**, 2053–2055 (2002).
33. T. Altantzis *et al.*, *Nano Lett.* **19**, 477–481 (2019).
34. H. Yoshida *et al.*, *Appl. Phys. Express* **4**, 065001 (2011).
35. Y. Kuwauchi, H. Yoshida, T. Akita, M. Haruta, S. Takeda, *Angew. Chem. Int. Ed.* **51**, 7729–7733 (2012).
36. T. W. Hansen *et al.*, *Science* **294**, 1508–1510 (2001).
37. Y. Kuwauchi *et al.*, *Nano Lett.* **13**, 3073–3077 (2013).
38. Y. Li *et al.*, *Nat. Commun.* **12**, 914 (2021).
39. J. B. Wagner, F. Cavalca, C. D. Damsgaard, L. D. L. Duchstein, T. W. Hansen, *Micron* **43**, 1169–1175 (2012).
40. M. Plodinec *et al.*, *Microsc. Microanal.* **26**, 220–228 (2020).
41. M. Boniface, M. Plodinec, R. Schlögl, T. Lunkenbein, *Top. Catal.* **63**, 1623–1643 (2020).
42. A. Beck *et al.*, *Nat. Catal.* **4**, 488–497 (2021).
43. W. Yuan *et al.*, *Science* **371**, 517–521 (2021).
44. J. Vincent, P. Crozier, *Microsc. Microanal.* **26** (S2), 1694–1695 (2020).
45. A. Akram *et al.*, *Chem. Sci.* **7**, 5833–5837 (2016).
46. M. Chi *et al.*, *Nat. Commun.* **6**, 8925 (2015).
47. J. Resasco, S. Dai, G. Graham, X. Pan, P. Christopher, *J. Phys. Chem. C* **122**, 25143–25157 (2018).
48. X. Du, H. Tang, B. Qiao, *Catalysts* **11**, 896 (2021).
49. R. Vanselow, M. Mundschau, *J. Catal.* **103**, 426–435 (1987).
50. X. Xu, Q. Fu, L. Gan, J. Zhu, X. Bao, *J. Phys. Chem. B* **122**, 984–990 (2018).
51. U. Diebold, *Surf. Sci. Rep.* **48**, 53–229 (2003).
52. T. Kandemir *et al.*, *Angew. Chem. Int. Ed.* **52**, 5166–5170 (2013).
53. A. Beck *et al.*, *J. Phys. Chem. C* **125**, 22531–22538 (2021).
54. J. M. Herrmann, M. Gravelle-Rumeau-Maillot, P. C. Gravelle, *J. Catal.* **104**, 136–146 (1987).
55. G. B. Raupp, J. A. Dumesic, *J. Phys. Chem.* **89**, 5240–5246 (1985).
56. C. Spreafico, W. Karim, Y. Ekinci, J. A. van Bokhoven, J. VandeVondele, *J. Phys. Chem. C* **121**, 17862–17872 (2017).
57. K. Zhang, L. Li, S. Shaikhutdinov, H.-J. Freund, *Angew. Chem. Int. Ed.* **57**, 1261–1265 (2018).
58. A. J. Fox, A. Drawl, G. R. Fox, B. J. Gibbons, S. Trolier-McKinstry, *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **62**, 56–61 (2015).
59. P. Müller, R. Kern, *Surf. Sci.* **457**, 229–253 (2000).
60. C. Castellarin-Cudia *et al.*, *Surf. Sci.* **554**, L120–L126 (2004).
61. A. Ruiz Puigdollers, P. Schlexer, S. Tosoni, G. Pacchioni, *ACS Catal.* **7**, 6493–6513 (2017).
62. J. L. Vincent, P. A. Crozier, *Microsc. Microanal.* **25** (S2), 1508–1509 (2019).
63. S. C. Ammal, A. Heyden, *J. Chem. Phys.* **133**, 164703 (2010).
64. C. Bäumer, R. Dittmann, in *Metal Oxide-Based Thin Film Structures* (Elsevier, 2018), pp. 489–522; https://linkinghub.elsevier.com/retrieve/pii/B9780128111666000200.
65. R. J. Kamaladasa *et al.*, *Microsc. Microanal.* **21**, 140–153 (2015).
66. D. S. Jeong, H. Schroeder, U. Breuer, R. Waser, *J. Appl. Phys.* **104**, 123716 (2008).
67. L. A. Bursill, B. G. Hyde, *Philos. Mag.* **23**, 3–15 (1971).
68. L. A. Bursill, B. G. Hyde, D. K. Philp, *Philos. Mag.* **23**, 1501–1513 (1971).
69. B. F. Donovan *et al.*, *Acta Mater.* **127**, 491–497 (2017).
70. M. D. Rasmussen, L. M. Molina, B. Hammer, *J. Chem. Phys.* **120**, 988–997 (2004).
71. P. Liu *et al.*, *Nanoscale* **11**, 11885–11891 (2019).
72. W. Gao, Z. D. Hood, M. Chi, *Acc. Chem. Res.* **50**, 787–795 (2017).
73. P. Mars, D. W. van Krevelen, *Chem. Eng. Sci.* **3**, 41–59 (1954).
74. R. Schaub *et al.*, *Phys. Rev. Lett.* **87**, 266104 (2001).

GEOMORPHOLOGY

Where rivers jump course

Sam Brooke[1], Austin J. Chadwick[2], Jose Silvestre[3], Michael P. Lamb[4],
Douglas A. Edmonds[5], Vamsi Ganti[1,6]*

Rivers can abruptly shift pathways in rare events called avulsions, which cause devastating floods. The controls on avulsion locations are poorly understood as a result of sparse data on such features. We analyzed nearly 50 years of satellite imagery and documented 113 avulsions across the globe that indicate three distinct controls on avulsion location. Avulsions on fans coincide with valley-confinement change, whereas avulsions on deltas are primarily clustered within the backwater zone, indicating a control by spatial flow deceleration or acceleration during floods. However, 38% of avulsions on deltas occurred upstream of backwater effects. These events occurred in steep, sediment-rich rivers in tropical and desert environments. Our results indicate that avulsion location on deltas is set by the upstream extent of flood-driven erosion, which is typically limited to the backwater zone but can extend far upstream in steep, sediment-laden rivers. Our findings elucidate how avulsion hazards might respond to land use and climate change.

T he gradual migration of rivers across floodplains is punctuated by episodic shifts in river course called avulsions (1). Avulsions are natural phenomena; river relocation nourishes floodplains with water, sediment, and nutrients and results in vast plains of fertile farmlands and biodiverse ecosystems that support the most populous places on Earth (2–4). Avulsions are also responsible for devastating historical floods (1, 5, 6) and linked to the decline of early urban settlements (7). Avulsions occur quasi-periodically and persistently at the apex of alluvial fans and river deltas (1, 8), resulting in their triangular shape. The time between successive avulsions can range from decades to millennia for different rivers (1, 8, 9), and direct observations of natural avulsions are rare. Consequently, the controls on where avulsions occur and how these sites will shift in response to climate change and human activity is poorly understood. Our understanding of the processes that control avulsion location are primarily based on physical experiments, numerical models, ancient river deposits (1, 8–14), and the analysis of river bifurcations (15), which can occur because of mechanisms other than avulsions (16). A time series of satellite imagery dating back to 1973 C.E. provides an opportunity to directly observe and characterize the global distribution of river avulsions to test classical and emerging theories (17).

We leveraged nearly 50 years of global satellite observations of river planform changes to locate avulsions. We focused on lobe-scale avulsions that occurred at the apex of river deltas and fans, including alluvial fans and

[1]Department of Geography, University of California Santa Barbara, Santa Barbara, CA, USA. [2]Department of Earth and Environmental Sciences, St. Anthony Falls Laboratory, University of Minnesota, Minneapolis, MN, USA. [3]Department of Earth and Environmental Sciences, Tulane University, New Orleans, LA, USA. [4]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA. [5]Department of Earth and Atmospheric Sciences, Indiana University, Bloomington, IN, USA. [6]Department of Earth Science, University of California Santa Barbara, Santa Barbara, CA, USA.
*Corresponding author. Email: vganti@ucsb.edu
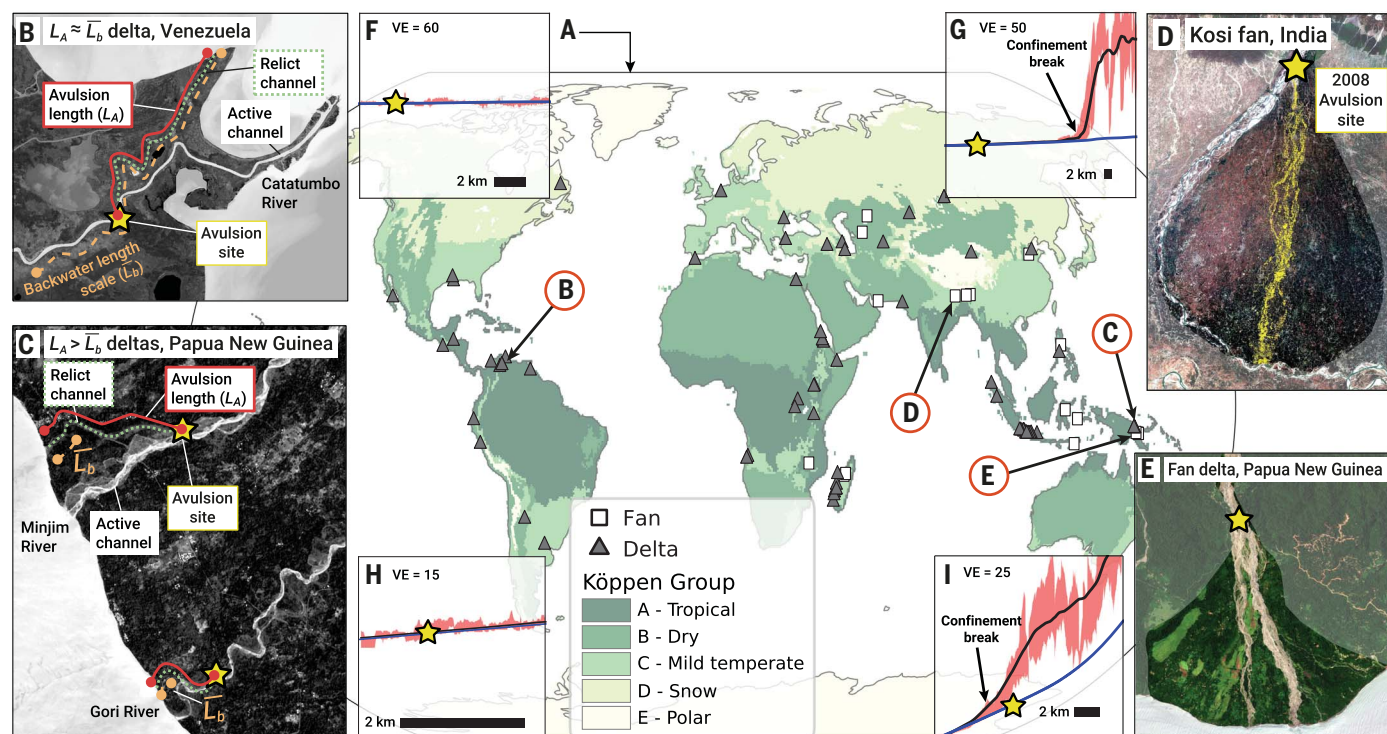
Fig. 1. Global avulsion sites on fans and river deltas. (**A**) World map of avulsion sites on fans (white squares) and deltas (gray triangles) with varying shades of green indicating the Köppen climate classification. Example avulsions on (**B**) the Catatumbo River delta in Venezuela and (**C**) the Minjim and Gori river deltas in Papua New Guinea; red lines denote avulsion length (L_A); orange dashed lines, estimated backwater length scale ($\overline{L_b}$); green dashed lines, relict channels; yellow stars, avulsion sites. Example avulsions on the (**D**) Kosi fan and a (**E**) fan delta in Papua New Guinea. (**F** to **I**) Topographic swath profiles across the example avulsion sites, where the black line indicates the median elevation, the shaded red area denotes its interquartile range along the cross-stream direction, and the blue line indicates the channel-floodplain elevation. Deviation of adjacent topography away from the floodplain elevations denotes a change in valley confinement (VE, vertical exaggeration).

fan deltas that build into a standing body of water from an adjacent highland. We defined avulsions as an abrupt and persistent change in the river course from the apex to the axial river or shoreline (Fig. 1). We identified avulsions from surface water maps derived from 30-m per pixel Landsat multispectral data from 1984 to 2020 C.E. (18), and from 60-m per pixel to 30-m per pixel Landsat imagery between 1973 and 2020 C.E. (19). We analyzed 113 historical river avulsions on fans and deltas, including 36 previously reported occurrences (Fig. 1 and table S1) (19).

Our results revealed 80 avulsions on coastal and inland river deltas and 33 avulsions on fans, captured by satellite imagery and historical maps (table S1). Snow and cloud cover in high-latitude regions and the spatial resolution of the satellite imagery affected avulsion documentation (19). Our compilation is therefore a representative—rather than exhaustive—global sample of avulsions. Avulsion sites covered 33°S to 54°N latitude, and were observed in tropical, temperate, and arid climates (Fig. 1). We found high avulsion density in the tropical islands of Papua New Guinea, Indonesia, and Madagascar (Fig. 1 and table S1) but did not observe avulsions in polar and snow climate zones. Rivers with avulsions covered a wide spectrum in modeled long-term water discharge (0.4 to 36,702 m^3/s), suspended sediment discharge (2 to 38,101 kg/s), and estimated riverbed slopes (4×10^{-5} to 2.6×10^{-2}) (Fig. 1 and table S1) (19). Study reaches on fans are steeper than on deltas, with median estimated riverbed slope (and interquartile range) of 4.3×10^{-4} (0.001) on deltas and 2.7×10^{-3} (0.011) on fans (Fig. 2A).

Avulsions on fans are thought to occur at the mouths of bedrock canyons or valleys, where rivers become unconfined (1, 20, 21) (fig. S4). Avulsions on deltas, however, lack a clear association with a canyon or valley and the processes that control avulsion location are debated (8, 11, 12, 22, 23). To assess the controls on avulsion sites, we extracted topographic swath profiles from 30-m spatial resolution global digital elevation models (Fig. 1) (19). Our results demonstrate that the avulsion sites on fans are always associated with at least a threefold slope drop in the topographic swath profiles (median of 6.5, interquartile range of 1.7; Fig. 2B), which is indicative of an abrupt valley-confinement change. This abrupt change can lead to a loss in fluvial sediment transport capacity (1, 9), causing focused sedimentation or present a steeper and more favorable path to the axial river, consistent with classical ideas (1, 21). Enhanced riverbed aggradation perches the water surface above the surrounding floodplain (10, 14, 24), and later floods trigger an avulsion (1, 10, 14, 24).

By contrast, the 80 avulsions on deltas were not coincident with an abrupt topographic

change (median slope break in swath profiles of 1.28, interquartile range of 0.65; Figs. 1 and 2B and table S1). We used these observations to test emerging ideas about the controls of avulsion location on deltas. Theory, physical experiments, and limited field observations indicate that avulsions on deltas cluster within the backwater zone (8, 22, 25)—the downstream reach of rivers characterized by nonuniform flows (26, 27). We estimated the backwater length scale, which approximates the upstream extent to which nonuniform flows prevail in alluvial rivers, defined as $\overline{L_b} = h_{bf}/S$, where h_{bf} and S are the bankfull flow depth and riverbed slope upstream of the avulsion site, respectively (19, 27, 28). We also estimated the avulsion length (L_A)—the streamwise distance from the avulsion site to the river mouth of the parent channel—from the satellite image that best captured the time of avulsion for each event (Fig. 1 and table S1) (19). Our measured L_A on deltas ranged from 0.5 to 490 km globally (Figs. 1 and 3A and table S1). Our results reveal that 62.5% of avulsions on deltas ($n = 50$) have a backwater-scaled avulsion

node with $L_A \approx \overline{L_b}$ (Fig. 3A). For these cases, the dimensionless avulsion length, defined as $L_A^* = L_A/\overline{L_b}$ (11), is 0.87 ± 0.38 [mean ± standard deviation (SD)], which is consistent with backwater-controlled avulsions (11, 12, 24). We interpreted that deltaic avulsions occur within the backwater zone because rivers during low flow decelerate in approach to the receiving basin, which causes sedimentation in the upstream part of the backwater zone (26, 27, 29)

Fig. 2. Topographic controls on avulsion sites of fans versus deltas. (**A**) Cumulative density function of the estimated riverbed slope, S, upstream of the avulsion site for river deltas (dark gray) versus fans (light gray). (**B**) Boxplots comparing the estimated slope drop in the topographic swath profile across the avulsion site for fans and deltas (19). The horizontal line denotes the median, the edges of the box indicate the first and third quartiles, and the edges of whiskers denote the 9th and 91st percentiles.

Fig. 3. Flood variability and backwater hydrodynamics control the avulsion location on deltas. (**A**) Measured avulsion length (L_A) as a function of the estimated backwater length scale ($\overline{L_b}$) for river deltas. The percentage of dry months (color of the markers) quantifies the degree to which a given river is ephemeral. (**B**) The dimensionless avulsion length ($L_A^* = L_A/\overline{L_b}$) as a function of the dimensionless flood duration (T_e^*) (19). The solid blue line and fill show LOESS regression and 95% confidence intervals, respectively, performed in log-log space. Markers of same color denote deltas with multiple avulsions. (**C**) The dimensionless avulsion length as a function of the probability of bankfull exceedance (F_{bf}) (19), for deltas with multiple avulsions. Error bars in panels denote standard deviation around the mean (19).

(fig. S4). Sedimentation is periodically interrupted by large floods, which cause an erosional wave to propagate from the river mouth to a distance upstream that is generally within the backwater zone (27, 29) (fig. S4). Over time, delta-lobe progradation causes riverbed aggradation and the sedimentation within the backwater zone during low flows—coupled with preferential riverbed scouring in the downstream accelerating part of the backwater zone during floods—causes a peak in riverbed sedimentation in the upstream part of the backwater zone (fig. S4), resulting in avulsions there (11, 12).

Our results also reveal a separate class of deltaic avulsions (n = 30) that correlate with neither the backwater length scale ($L_A \gg \overline{L_b}$; Fig. 3A) nor a valley-confinement break (Figs. 1 and 2B). These avulsions have $L_A^* \gg 1$ (mean and SD of 13.4 ± 13.0) and correspond with steep, sediment-laden rivers in deserts and tropical islands (figs. S9 and S10). In accordance with emerging theory (30), we hypothesized that the longitudinal extent of flood-driven scours in these rivers is more pronounced than in backwater-scaled deltas (fig. S4), which diminishes sedimentation within the backwater reach and thus causes $L_A \gg \overline{L_b}$ (31). To test this hypothesis, we estimated the dimensionless flood duration, defined as $T_e^* = t_{scour}/t_{adj}$, where t_{scour} is the typical bankfull-overtopping flood duration and t_{adj} is a bed-adjustment time scale (11), globally by simulated monthly water and suspended sediment discharges from 1980 to 2020 C.E. (19, 32) (fig. S6). For rivers with $T_e^* > 1$, the longitudinal extent of flood-driven erosion is expected to extend upstream of the backwater zone to a distance approximated by $\overline{L_b}\sqrt{T_e^*}$ (31, 33) (fig. S4). By contrast, for $T_e^* < 1$, flood-driven scours diminish within the backwater zone (11, 29) (fig. S4).

The dimensionless flood duration (T_e^*) separated the data into two distinct classes of deltaic avulsions (Fig. 3B and fig. S7) (19). Avulsions with $L_A^* \gg 1$ are associated with $T_e^* \gg 1$ (mean and SD of 168 ± 505). These rivers had t_{adj} on the order of days to weeks such that flood-driven scours propagated beyond the backwater zone during typical floods and caused avulsions with $L_A^* \gg 1$ (Fig. 3B and figs. S4 and S8). In addition, for rivers with $T_e^* \gtrsim 1$, the theoretical flood-driven scour length scale, $\overline{L_b}\sqrt{T_e^*}$, appears to control the avulsion length, L_A (Fig. 3B), rather than the backwater length scale. In comparison, backwater-scaled avulsions occurred only if $T_e^* \lesssim 1$, which is typical of lowland rivers with a t_{adj} of months to centuries (figs. S7 to S9). For these rivers, flood-driven scours were limited to the backwater zone during typical floods, resulting in backwater-scaled avulsion sites ($L_A \approx \overline{L_b}$) (Fig. 3, A and B, and fig. S4).

Physics-based numerical models demonstrate that historical avulsion lengths are a diagnostic indicator of future avulsion sites even in the face of anthropogenic climate change and interference (30), which suggests that our global database provides a first-order prediction of future avulsion sites on deltas. However, numerical and field studies also indicate that considerable shoreline encroachment from accelerated sea level rise can shift the deltaic avulsion nodes upstream (34, 35). In addition, our analysis implies that $T_e^* \approx 1$ is a transition point between two scaling regimes for avulsions on deltas (Fig. 3B and fig. S8), where a further reduction in t_{adj} can drive avulsion sites upstream. Agriculture and land use have enhanced the sediment loads of most global coastal rivers (36), and extensive dam infrastructure can reduce sediment caliber—changes likely to increase T_e^*. These changes can result in coastal rivers to transition beyond $T_e^* \approx 1$, causing an upstream shift of their avulsion nodes. Our analysis indicates that inland river deltas and small, low-gradient coastal deltas in tropical islands (e.g., Indonesia) are most susceptible to transitioning from the backwater-scaled avulsion regime to the high–sediment-load modulated avulsion regime with changes in magnitude and duration of floods as well as sediment supply (19) (fig. S8), which may expose previously unaffected upstream communities to the risks of avulsion hazards.

Our results also highlight that changes in flood frequency caused by differing climates (37) or engineering (e.g., dams) can substantially affect the avulsion location on lowland deltas (Fig. 3C). The backwater-scaled avulsions in our dataset are variable with $L_A^* \in [0.24, 1.62]$ (Fig. 3, A and B). We investigated the causes of this variability on four deltas with multiple recorded avulsions: the Huanghe (n = 7) (9, 13) and Sulengguole rivers (n = 6) (35) in China, and the Cisanggarung (n = 3) and Cipunagara (n = 3) rivers in Java, Indonesia (Fig. 4, A and B). The mean L_A^* of
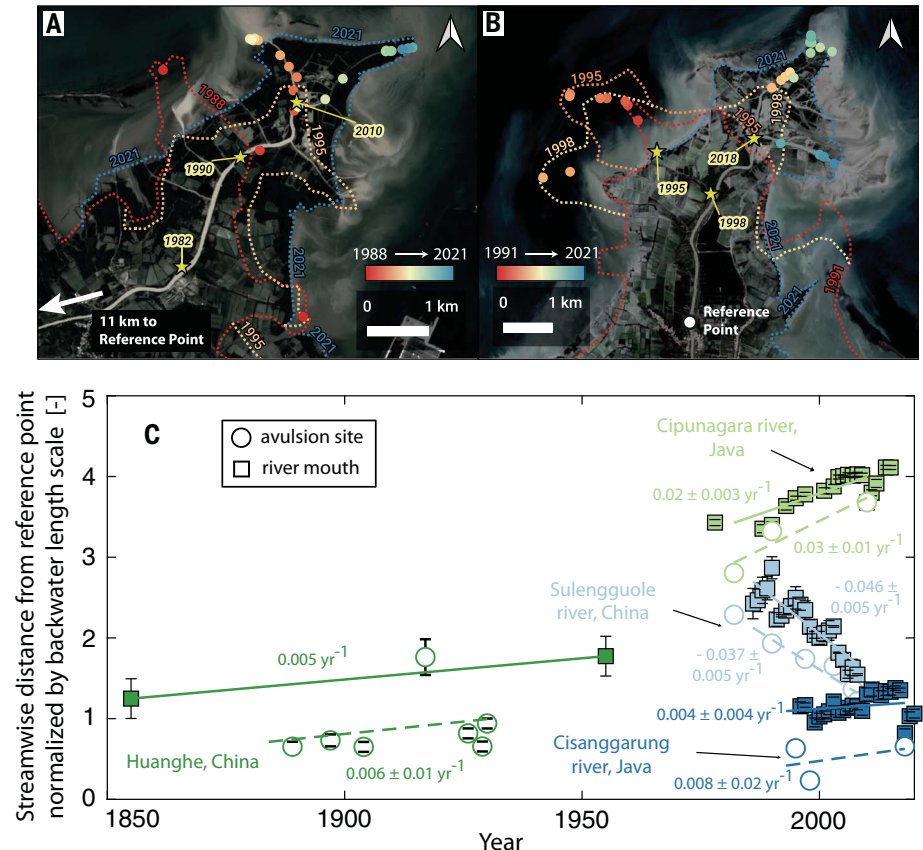
Fig. 4. Mobility of backwater-scaled avulsion sites on deltas. Modified ESA and Copernicus Sentinel 2 images of the (**A**) Cipunagara and (**B**) Cisanggarung river deltas in Java, Indonesia. Colored dashed lines denote the shoreline position at different times, derived from USGS and NASA Landsat imagery, and the circular markers show the river mouth position color coded by time. Yellow stars indicate avulsion sites. (**C**) Temporal evolution of the river mouth and avulsion site measured along the streamwise direction from a fixed reference point (19), normalized by the backwater length scale, for the Cipunagara, Cisanggarung, Huanghe (9, 13), and Sulengguole river deltas (35). The solid and dashed lines indicate the best-fitting linear regression for the river mouth evolution and the avulsion site evolution, respectively. The similarity of slopes between solid and dashed lines indicates that L_A remained consistent despite river mouth evolution.

these rivers is statistically different (one-way ANOVA test, $P = 0.001$), indicating that the observed L_A^* variability is not driven by chance. Instead, mean L_A^* increases with the frequency of bankfull-overtopping floods (Fig. 3C). This trend is consistent with numerical models (11), which show that more frequent floods enhance the longitudinal extent of flood-driven scours within the backwater zone, thereby increasing L_A^*.

River deltas with multiple recorded avulsions also demonstrate the mobility of backwater-scaled avulsion sites with river mouth evolution. These examples provide a template for how backwater-scaled avulsion sites on deltas may respond to future changes in sediment supply and relative sea level. Backwater avulsion theory indicates that avulsions should shift upstream or downstream in tandem with river mouth evolution such that the avulsion node remains within the backwater zone (9, 12, 30, 34). Our data support this idea wherein progradation of the Cisanggarung and the Cipunagara river deltas caused downstream shifts in their avulsion sites (Fig. 4) (19), similar to previous observations on the Huanghe (9, 13); further, the Sulengguole river mouth retreat caused the avulsion sites to migrate inland (Fig. 4C) (35). In all cases, the episodic avulsion-site migration was commensurate with the direction and magnitude of river mouth evolution such that L_A^* for each delta remained consistent with time (Fig. 4C). These results suggest that accelerated delta progradation associated with global agriculture and deforestation (36) will lead to the downstream migration of avulsion sites; however, drowning of river deltas from accelerated relative sea-level rise will shift avulsion sites upstream (30, 34).

Our global analysis reveals distinct controls on avulsion locations on fans and deltas: fan avulsions are tied to an abrupt valley-confinement change; however, the longitudinal extent of flood-driven scours sets the avulsion location on deltas, and dimensionless flood duration separates deltaic avulsions between those that are backwater-scaled and those that occur farther upstream. For most rivers, avulsions reoccur on time scales longer than the 50-yr record we analyzed, such that most avulsion locations have not been documented historically, and the devastating consequences of flooding following an avulsion have not been realized. Our work provides a predictive framework to assess future avulsion locations on fans and deltas, and their response to land use and climate change.

REFERENCES AND NOTES

1. R. Slingerland, N. D. Smith, *Annu. Rev. Earth Planet. Sci.* **32**, 257–285 (2004).
2. C. J. Vörösmarty *et al.*, *Bull. At. Sci.* **65**, 31–43 (2009).
3. P. H. Gleick, *Science* **302**, 1524–1528 (2003).
4. D. A. Edmonds, R. L. Caldwell, E. S. Brondizio, S. M. O. Siani, *Nat. Commun.* **11**, 4741 (2020).
5. T. R. Kidder, H. Liu, *Archaeol. Anthropol. Sci.* **9**, 1585–1602 (2017).
6. R. Sinha, *Curr. Sci.* **97**, 429–433 (2009).
7. G. S. Morozova, *Geoarchaeology* **20**, 401–423 (2005).
8. D. J. Jerolmack, *Quat. Sci. Rev.* **28**, 1786–1800 (2009).
9. V. Ganti, Z. Chu, M. P. Lamb, J. A. Nittrouer, G. Parker, *Geophys. Res. Lett.* **41**, 7882–7890 (2014).
10. A. Aslan, W. J. Autin, M. D. Blum, *J. Sediment. Res.* **75**, 650–664 (2005).
11. A. J. Chadwick, M. P. Lamb, A. J. Moodie, G. Parker, J. A. Nittrouer, *Geophys. Res. Lett.* **46**, 4267–4277 (2019).
12. V. Ganti, A. J. Chadwick, H. J. Hassenruck-Gudipati, B. M. Fuller, M. P. Lamb, *Sci. Adv.* **2**, e1501768 (2016).
13. A. J. Moodie *et al.*, *J. Geophys. Res. Earth Surf.* **124**, 2438–2462 (2019).
14. D. Mohrig, P. L. Heller, C. Paola, W. J. Lyons, *Geol. Soc. Am. Bull.* **112**, 1787–1803 (2000).
15. O. A. Prasojo, T. B. Hoey, A. Owen, R. D. Williams, *Geophys. Res. Lett.*, **49**, e2021GL093656 (2022).
16. D. A. Edmonds *et al.*, *ESSOAr10507512.1* [Preprint] (2021).
17. J. M. Valenza, D. A. Edmonds, T. Hwang, S. Roy, *Nat. Commun.* **11**, 2116 (2020).
18. J.-F. Pekel, A. Cottam, N. Gorelick, A. S. Belward, *Nature* **540**, 418–422 (2016).
19. Materials and methods are available as supplementary materials.
20. T. C. Blair, J. G. McPherson, *J. Sediment. Res.* **64**, 450–489 (1994).
21. L. S. Jones, S. A. Schumm, *Spec. Publ. Int. Assoc. Sedimentol.* **28**, 171–178 (1999).
22. P. Chatanantavet, M. P. Lamb, J. A. Nittrouer, *Geophys. Res. Lett.* **39**, n/a (2012).
23. K. M. Ratliff, E. W. H. Hutton, A. B. Murray, *Earth Planet. Sci. Lett.* **559**, 116786 (2021).
24. V. Ganti, A. J. Chadwick, H. J. Hassenruck-Gudipati, M. P. Lamb, *J. Geophys. Res. Earth Surf.* **121**, 1651–1675 (2016).
25. D. J. Jerolmack, J. B. Swenson, *Geophys. Res. Lett.* **34**, n/a (2007).
26. J. A. Nittrouer, J. Shaw, M. P. Lamb, D. Mohrig, *Geol. Soc. Am. Bull.* **124**, 400–415 (2012).
27. M. P. Lamb, J. A. Nittrouer, D. Mohrig, J. Shaw, *J. Geophys. Res.* **117**, 2011JF002079 (2012).
28. C. Paola, D. Mohrig, *Basin Res.* **8**, 243–254 (1996).
29. P. Chatanantavet, M. P. Lamb, *J. Geophys. Res. Earth Surf.* **119**, 1263–1282 (2014).
30. A. J. Chadwick, M. P. Lamb, *J. Geophys. Res. Earth Surf.* **126**, e2020JF005950 (2021).
31. S. A. S. Brooke, V. Ganti, A. J. Chadwick, M. P. Lamb, *Geophys. Res. Lett.* **47**, (2020).
32. S. Cohen, A. J. Kettner, J. P. M. Syvitski, *Global Planet. Change* **115**, 44–58 (2014).
33. V. Ganti, M. P. Lamb, A. J. Chadwick, *J. Sediment. Res.* **89**, 815–832 (2019).
34. A. J. Chadwick, M. P. Lamb, V. Ganti, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 17584–17590 (2020).
35. J. Li, V. Ganti, C. Li, H. Wei, *Earth Planet. Sci. Lett.* **577**, 117270 (2022).
36. J. H. Nienhuis *et al.*, *Nature* **577**, 514–518 (2020).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abm1215
Supplementary Text
Figs. S1 to S12
Table S1
References (37–56)
Movies S1 and S2
Data S1

Submitted 26 August 2021; accepted 1 April 2022
10.1126/science.abm1215

ENVIRONMENTAL TOXINS

Models predict planned phosphorus load reduction will make Lake Erie more toxic

Ferdi L. Hellweger[1]*, Robbie M. Martin[2], Falk Eigemann[1], Derek J. Smith[3],
Gregory J. Dick[3,4], Steven W. Wilhelm[2]*

Harmful cyanobacteria are a global environmental problem, yet we lack actionable understanding of toxigenic versus nontoxigenic strain ecology and toxin production. We performed a large-scale meta-analysis including 103 papers and used it to develop a mechanistic, agent-based model of *Microcystis* growth and microcystin production. Simulations for Lake Erie suggest that the observed toxigenic-to-nontoxigenic strain succession during the 2014 Toledo drinking water crisis was controlled by different cellular oxidative stress mitigation strategies (protection by microcystin versus degradation by enzymes) and the different susceptibility of those mechanisms to nitrogen limitation. This model, as well as a simpler empirical one, predicts that the planned phosphorus load reduction will lower biomass but make nitrogen and light more available, which will increase toxin production, favor toxigenic cells, and increase toxin concentrations.

Harmful cyanobacteria and their toxins constitute one of the most important global environmental challenges faced by humanity, which is expected to get worse in a warmer climate (*1, 2*). The problem is exemplified by *Microcystis*, which can produce the potent hepatotoxin microcystin (MC), a class of cyclic nonribosomal peptides originally known as "fast death factor" that has already disrupted the drinking water supplies of Toledo, Ohio on Lake Erie and those of other cities (*3*).

In fresh waters, phytoplankton growth is often limited by the availability of phosphorous (P), and that concept has been applied in mathematical models and used to control bulk biomass—i.e., eutrophication—in many systems (*4*). It is also the basis for a costly binational agreement aimed at controlling toxic cyanobacteria in Lake Erie using a 40% P load reduction (*5*). However, this simple model does not address or explain the ecology of toxigenic versus nontoxigenic strains or the production of toxins, where nitrogen (N), temperature, and reactive oxygen species [e.g., hydrogen peroxide (H_2O_2)] are important factors (*6–10*). Advances in our understanding and management of cyanobacteria necessitate the development of new conceptual and quantitative models that incorporate relevant mechanisms.

The biology of *Microcystis*, including toxin production, has been extensively investigated in the laboratory, and a natural first step in the development of a next-generation model is to summarize and synthesize this information.

We performed a broad literature meta-analysis, including 103 papers published from 1958 and totaling 708 experiments (i.e., cultures, all cataloged and discussed individually in the supplementary materials). Experiments were conducted with 67 strains using various methods. Consequently, the database is heterogeneous, but some consistent and ecologically relevant patterns emerge (Fig. 1; model results discussed subsequently). Across 20 experiments, the optimum T for MC production is not 6.3°C, it is 6.3°C less than that for growth (Fig. 1A). As expected from the chemical formula of MC, which includes ~10 N atoms per molecule, lower N availability reduces MC content (Fig. 1B). The observed MC content can be higher or lower at increased light, which is also affected by binding to proteins (Fig. 1, C and F) (*9, 11*). These patterns show that the catalog of observations is a useful resource, even without model analysis.

Building on this large catalog of observations and existing cyanobacteria models (*12*) and following a pattern-oriented modeling approach (*13*), we developed a dynamic, mechanistic, and molecular-level model of *Microcystis* growth and toxin production. The agent-based model (ABM) simulates individual cells (*14*), with explicit representation of select representative genes with corresponding transcripts, enzymes, and metabolite pools (Fig. 2 shows a subset of the model). For example, *mcyD* is used as a proxy for all 10 genes in the MC synthesis cluster. The model includes a single gene, *t2prx*, as a representative of all H_2O_2-degrading enzymes [e.g., *katG* and *trxA* (*10, 15*)]. GLU and G3P represent labile N and C pools.

We repeated each experiment in the database in silico using the model. The ability of the model to reproduce observations is quantified using a pattern-oriented approach, where we identify patterns in the observations and compare them with the model (*12*) (supple-mentary materials). In total, there are 897 patterns, and the model reproduces 87% of them. Mechanistic modeling thus provides a natural and intuitive way to summarize and interpret observations for *Microcystis*, as has been found for other organisms (*12, 16*).

The model can reproduce the relatively simple temperature optima, but also the more complex effect of N on MC content (Fig. 1B). It also predicts the decrease in free or measurable MC content at higher light intensities, which is the result of increased MC binding to proteins (*9, 11*). In some cases, the model proposes mechanisms underlying previously unexplained observed patterns, like the transient increase in MC content upon light downshift (Fig. 1D). In the model, this pattern is related to the dynamics of G3P and GLU, which are the limiting substrates for biomass synthesis and MC synthesis, respectively, in this experiment (fig. S109). When the light intensity decreases abruptly, photosynthesis and G3P content drop rapidly, and biomass decreases. However, N assimilation continues, and the biomass-based GLU content increases. Consequently, biomass-based MC synthesis increases. The *mcyD* gene is down-regulated rapidly upon light-downshift, but it takes some time for the enzyme level to respond. Once this occurs, the MC synthesis and content also decrease.

For some experiments, there can be substantial differences between observations and model predictions (Fig. 1E). This can be partially attributed to the constraint of calibrating the model with one parameter set (for each strain) to multiple datasets. There are experiments from 28 papers for this strain in the database. However, the main purpose of the model application to the database is to test its structure, i.e., mechanisms; differences in magnitude are less relevant than patterns because they can be calibrated for any field application of the model. In this example (Fig. 1E), the main observed pattern, the increase in MC content upon temperature decrease and vice versa, is reproduced by the model.

The key to understanding the differential ecology of toxigenic and nontoxigenic strains lies in the biological role of MC. There is increasing evidence that MC binds to enzymes and protects them from damage by reactive oxygen species, such as H_2O_2 (*7, 9*). Experiments with toxigenic wild-type and nontoxigenic Δ*mcyB* mutant cells show that, when H_2O_2 is added at environmentally relevant concentrations, the MC producer is less vulnerable than the non–MC-producing mutant (*7*) (Fig. 3). By contrast, when H_2O_2 is added at very high concentrations—levels corresponding to algicide or cyanocide treatment—the MC producer is more vulnerable (*15*). These observed patterns are relevant to the strain-level ecology and test the structural realism of the model.

[1]Water Quality Engineering, Technical University of Berlin, Berlin, Germany. [2]Department of Microbiology, University of Tennessee, Knoxville, TN, USA. [3]Department of Earth and Environmental Sciences, University of Michigan, Ann Arbor, MI, USA. [4]Cooperative Institute for Great Lakes Research, University of Michigan, Ann Arbor, MI, USA.
*Corresponding author. Email: ferdi.hellweger@tu-berlin.de (F.L.H.); wilhelm@utk.edu (S.W.W.)

Fig. 1. Patterns of toxin production in *Microcystis* and comparison with model. (**A**) Temperature optima for MC production and growth (*n* = 20); error bars are 95% Cls. (**B**) MC content under lower N (*n* = 41) and P (*n* = 24) relative to control; values are log₂ ratios. (**C**) MC content (solid symbols) or *mcy* transcripts (open symbols) at high relative to low light (HL/LL) (*n* = 16); values are log₂ ratios. Diagonal solid line is 1:1 (indicating perfect model-data agreement), and the dashed line is linear regression; R^2 = 0.77. (**D** and **E**) Transient response of MC content to changes in light (D) and temperature (E) in continuous culture. Data are from (*8, 27*). (**F**) MC content versus light. Data are from (*11*). Symbols are data, and lines are models in (D) to (F). (**G**) Relative fitness of toxigenic and nontoxigenic strains in mono- and coculture under various light, temperature, and nutrient conditions (*n* = 13). Growth rate difference (Δ Gr. rate) indicates the toxigenic − nontoxigenic growth rate. Co-H_2O_2 is a coculture simulation with the H_2O_2 damage turned off to illustrate that the advantage of the toxigenic strain in coculture is the result of interaction through H_2O_2. Data are from (*18*). nd, no data.
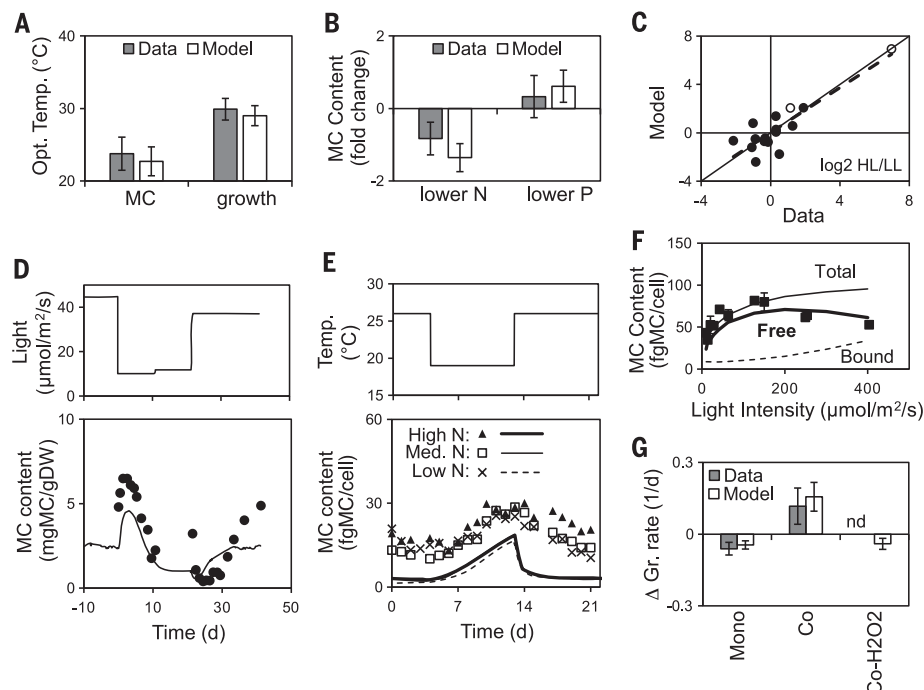
Fig. 2. H_2O_2 generation, damage to enzymes and protection by MC, or degradation by T2prx in the model. Only select components and processes are shown; see supplementary materials for full model details. H_2O_2 is generated by photosynthesis and respiration, diffuses across the membrane, and inhibits enzymes, including PSURbcL and RptMH (ribosome). MC is synthesized from G3P and GLU and binds to and protects enzymes. T2prx (peroxiredoxin, used as a proxy for all H_2O_2 degradation enzymes) degrades H_2O_2.
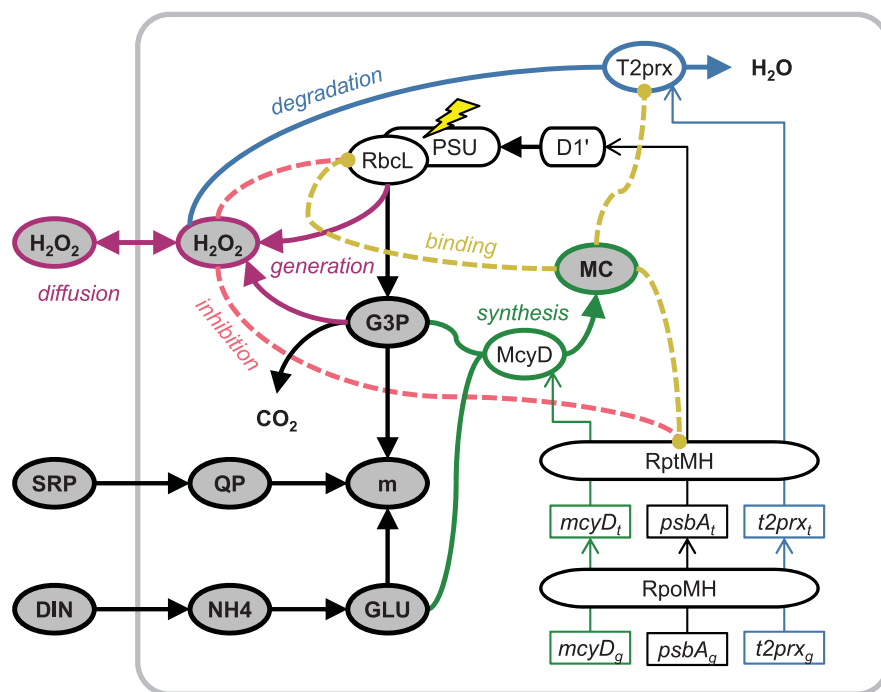
The model includes generation of H_2O_2, damage to enzymes by H_2O_2, and two H_2O_2-management systems, including protection by MC and degradation by T2prx (which represents all H_2O_2-degrading enzymes) (Fig. 2), and it reproduces the observations (Fig. 3). The pattern at low-H_2O_2 levels can simply be attributed to protection by MC. The pattern at high-H_2O_2 levels is more complex. In the model, before the H_2O_2 addition, the wild-type strain relies on the MC system for H_2O_2 management and has the T2prx system down-regulated. When hit with H_2O_2, the MC system is overwhelmed. The cells express *t2prx*, but by this time, the ribosomes are damaged and the cells cannot synthesize T2prx enzymes and do not recover. The mutant, however, has the T2prx system active before H_2O_2 addition and rapidly degrades the H_2O_2 and recovers. These are the mechanisms underlying the pattern in the model, which is consistent with the observed pattern. The model thus constitutes a viable mechanistic explanation or hypothesis for the mechanisms responsible for the observed pattern.

In the high-H_2O_2 experiment, the toxigenic strain down-regulated H_2O_2-degrading enzymes under ambient conditions. This general strategy of protection against H_2O_2 by MC over degradation with enzymes may also be reflected in the gene repertoire of *Microcystis* strains—e.g., *katG* genes are less frequently found in toxigenic genotypes (*10*).

H2O2 readily diffuses across cell membranes, and the model predicts that degradation by the nontoxigenic strain leads to lower extracellular H2O2 levels, which also benefits the toxigenic strain—like the interaction between marine cyanobacteria and heterotrophic bacteria demonstrated previously (17). This interaction mechanism can explain observations where toxigenic strains outcompete nontoxigenic strains in coculture, despite equal or lower growth rate in monoculture (18, 19) (Fig. 1G).

The success of the model in reproducing *Microcystis* biology suggests that it may provide useful insights into ecology at the field scale. To test this, we simulate the water column around the Toledo drinking water intake during the 2014 growing season, when MC was detected in the drinking water (Fig. 4A). We use a simplified approach and simulate a completely mixed box [continuous stirred tank reactor (CSTR)] with dissolved inorganic nitrogen (DIN) and soluble reactive phosphorus (SRP) input rates estimated from observed in situ DIN, SRP, and phycocyanin (PCN) concentrations and including estimates of photochemical H2O2 production (20) (details in section S3). The simulation includes toxigenic and nontoxigenic strains that differ only in their H2O2 management strategy—i.e., the toxigenic strain has *mcyD* and the nontoxigenic has *t2prx*—so any differences in their behavior can be directly attributed to these mechanisms. The parameters of the Lake Erie strains (same for toxigenic and nontoxigenic) were calibrated within the range of the laboratory strains, except that a lower H2O2 membrane permeability is needed, which may be associated with colony formation in the field.

The succession from toxigenic to nontoxigenic strains in the model is the result of differences in H2O2 management strategies that have different susceptibilities to N limitation (fig. S111). In June and July, the DIN concentration is high, and the toxigenic strain can synthesize sufficient MC to protect its enzymes—it incurs less damage and outcompetes the nontoxigenic strain. In August and September, DIN is depleted, curtailing the production of MC by the toxic strain, which increases damage and lowers its growth rate. The *t2prx* system of the nontoxic strain is not affected by the lower DIN, and it outcompetes the toxic strain at that time.

Laboratory experiments show that N limitation results in lower MC levels (Fig. 1B) and that MC helps protect against H2O2 at ambient concentrations (Fig. 3) (7, 9). Together, these observations (and the model) suggest that toxigenic *Microcystis* is more vulnerable to H2O2 under N limitation, although that hypothesis has not yet been tested experimentally at environmental H2O2 levels.

Although our model does not consider all factors expected to affect strain-level ecology and toxin production (10, 21), it is based on mechanisms and reproduces the laboratory and field observations. It therefore represents a step forward in the mechanistic understanding of toxic cyanobacteria ecology and can inform lake management.

We used the model to evaluate load reduction scenarios, including 40% reduction in N, P, and both N and P (Fig. 4B). The largest biomass decrease is predicted for the N and P scenario, but all scenarios produce a decrease and none reach 40%, pointing to N, P, and light limitation. For the P-only reduction scenario, total *Microcystis* biomass decreases, but the increased N and light availability increase MC synthesis by the toxigenic strain (Fig. 1, B, C, D, and F), which lowers H2O2 damage and increases the toxigenic fraction. The toxigenic cells have more MC, and there are more of them. These two factors counteract the decrease in biomass and lead to increased MC concentration. When the effect of N and light on MC production is removed in the model, it predicts that MC concentration will decrease also for the P-only reduction scenario (Fig. 4B, part 1). Simulations where the P load reduction is focused earlier, when P is limiting (Fig. 4A, part 5) (22), are more effective at controlling biomass but will further increase MC concentrations through the same mechanisms as those for the even reduction (fig. S117). This pattern emerges in the relatively complex model, but the causal chain is simple and is predicted using a simple calculation or model that builds on mass balance and previous models and is parameterized directly from laboratory experiments (23, 24) (Fig. 4C and section S4).

In addition to changes in nutrient loads, global warming is expected to affect the lake (1–3, 25). For present loading, the model predicts cyanobacteria biomass increases and
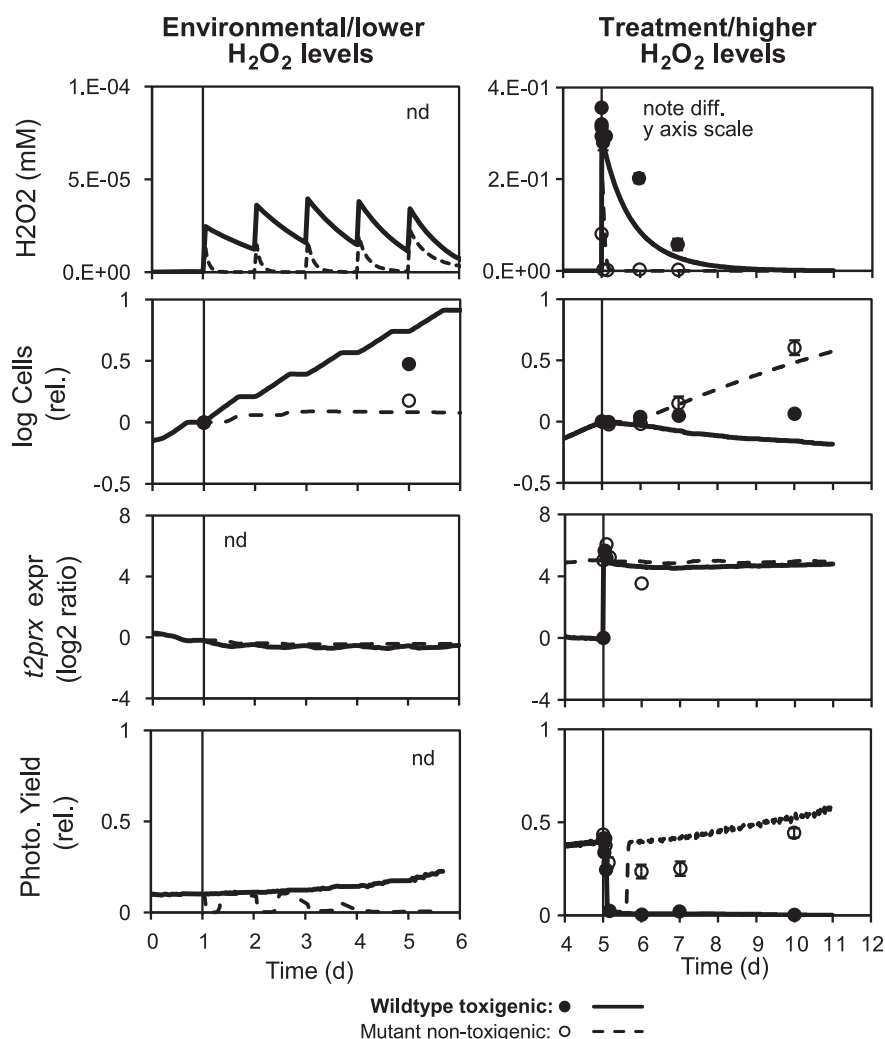
Fig. 3. H2O2 management and vulnerability of toxigenic and nontoxigenic strains. At environmental H2O2 concentrations (left), protection of enzymes by MC is advantageous. Data are from (7). At very high H2O2 concentrations, i.e., treatment levels (right), degradation is advantageous. Symbols are data, and lines are model. Data are from (15).
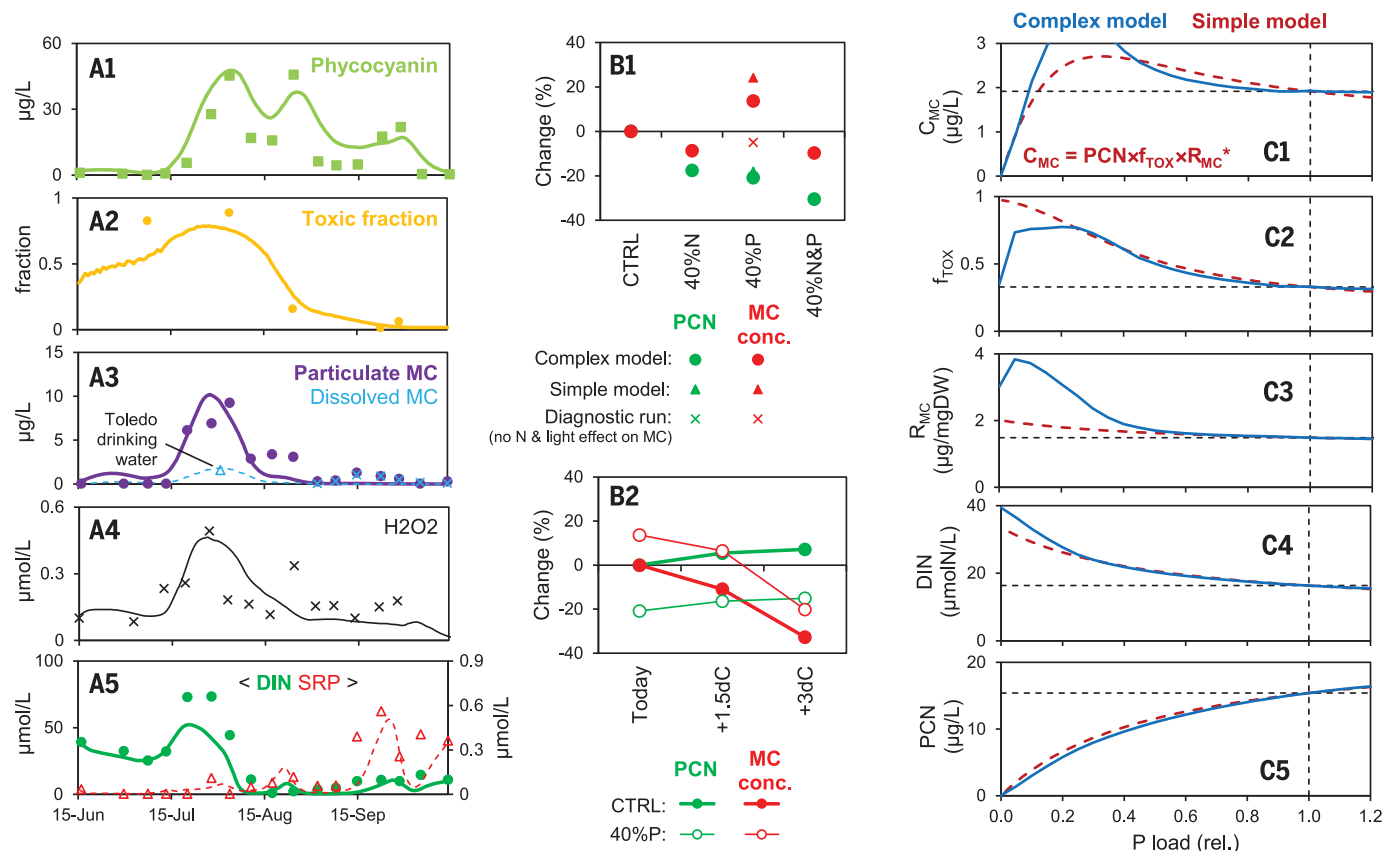
Fig. 4. Lake Erie simulation. Station W12 in 2014. (**A**) BaseCase model (lines) and data (symbols) versus time. From top to bottom, cyanobacteria biomass (PCN), toxigenic fraction, particulate and dissolved MC concentration (methanol-extractable fraction), H_2O_2 concentration, and nutrient concentrations are shown. See section S3 for additional results and discussion. Data are from this study and other sources (section S3). (**B**) Management and temperature scenarios. Shown are reductions in average PCN and observable MC relative to the current loading (CTRL) and temperature. Part 1 (top) is for present temperature and shows all management scenarios, as well as results from the simple model and a diagnostic run to illustrate the effect of N and light. Part 2 (bottom) is for various temperature scenarios and shows current loading (CTRL) and 40% P-only reduction (40%P). (**C**) Response of full complex and a simpler model to reductions in P loading. From top to bottom, volume-based MC concentration (C_{MC}), toxigenic fraction (f_{TOX}), biomass-based MC content (R_{MC}), DIN concentration, and cyanobacteria biomass (PCN) are shown. The asterisk indicates units converted using 0.35 gC/gDW and 89 gC/gPCN, see table S4.

MC concentration decreases with rising temperature (Fig. 4B, part 2, CTRL). These results and observations of lower temperature optima for MC synthesis (Fig. 1A) suggest that toxin concentrations will not go up with the expected increase in biomass and brightens the otherwise bleak outlook for harmful cyanobacteria blooms. The model predicts that the temperature effect superimposes those of nutrient load reductions, and for the 40%P scenario, the net effect is a reduction in MC concentration for the higher temperature increase evaluated (Fig. 4B, part 2, 40%P). However, the decrease in MC concentration for this scenario is the result of the warmer temperature and not the P load reduction—i.e., the load reduction still increases the MC concentration relative to the warmer BaseCase scenario. These results suggest that P-only management is counterproductive for reducing MC concentration under all climate sce-narios evaluated, and they support a dual N and P management strategy.

Our results suggest that future management efforts limited to P will increase relative availability of N and light, promote toxigenic strains, and increase toxin concentrations. This mechanism may be in part responsible for the presently observed resurgence of toxic cyanobacteria after historical P load reductions to Lake Erie and many other systems (*26*). Lake health is endangered by climate change and can be threatened by management actions that are well intended but based on an incomplete understanding of *Microcystis* biology and biochemistry. We may presently be witnessing the consequences of both threats

REFERENCES AND NOTES

1. H. W. Paerl, J. Huisman, *Science* **320**, 57–58 (2008).
2. S. C. Chapra *et al.*, *Environ. Sci. Technol.* **51**, 8933–8943 (2017).
3. G. S. Bullerjahn *et al.*, *Harmful Algae* **54**, 223–238 (2016).
4. D. W. Schindler, S. R. Carpenter, S. C. Chapra, R. E. Hecky, D. M. Orihel, *Environ. Sci. Technol.* **50**, 8923–8929 (2016).
5. United States Environmental Protection Agency, "U.S. Action Plan for Lake Erie" (2018); https://www.epa.gov/glwqa/us-action-plan-lake-erie.
6. C. J. Gobler *et al.*, *Harmful Algae* **54**, 87–97 (2016).
7. C. Dziallas, H.-P. Grossart, *PLOS ONE* **6**, e25569 (2011).
8. R. M. Martin *et al.*, *Front. Microbiol.* **11**, 601864 (2020).
9. Y. Zilliges *et al.*, *PLOS ONE* **6**, e17615 (2011).
10. G. J. Dick *et al.*, *Environ. Microbiol.* **23**, 7278–7313 (2021).
11. C. Wiedner *et al.*, *Appl. Environ. Microbiol.* **69**, 1475–1481 (2003).
12. F. L. Hellweger *et al.*, *Environ. Microbiol.* **18**, 2721–2731 (2016).
13. V. Grimm *et al.*, *Science* **310**, 987–991 (2005).
14. F. L. Hellweger, R. J. Clegg, J. R. Clark, C. M. Plugge, J.-U. Kreft, *Nat. Rev. Microbiol.* **14**, 461–471 (2016).
15. J. M. Schuurmans *et al.*, *Harmful Algae* **78**, 47–55 (2018).
16. D. N. Macklin *et al.*, *Science* **369**, eaav3751 (2020).

17. J. J. Morris, Z. I. Johnson, M. J. Szul, M. Keller, E. R. Zinser, . *PLOS ONE* **6**, e16805 (2011).
18. L. Lei, C. Li, L. Peng, B.-P. Han, *Ecotoxicology* **24**, 1411–1418 (2015).
19. D. Schatz *et al*., *Environ. Microbiol.* **7**, 798–805 (2005).
20. R. M. Cory *et al*., *Front. Mar. Sci.* **4**, 377 (2017).
21. S. W. Wilhelm, G. S. Bullerjahn, R. M. L. McKay, *mBio* **11**, e00529-20 (2020).
22. R. P. Stumpf, T. T. Wynne, D. B. Baker, G. L. Fahnenstiel, *PLOS ONE* **7**, e42444 (2012).
23. B. M. Long, G. J. Jones, P. T. Orr, *Appl. Environ. Microbiol.* **67**, 278–283 (2001).
24. C. Vézie, J. Rapala, J. Vaitomaa, J. Seitsonen, K. Sivonen, *Microb. Ecol.* **43**, 443–454 (2002).
25. J. Trumpickas, B. J. Shuter, C. K. Minns, *J. Great Lakes Res.* **35**, 454–463 (2009).
26. H. W. Paerl *et al*., *Hydrobiologia* **847**, 4359–4375 (2020).
27. H. Utkilen, N. Gjølme, *Appl. Environ. Microbiol.* **58**, 1321–1325 (1992).
28. F. L. Hellweger, fhellweger/Microcystis: version 1, version 1.0, Zenodo (2022); https://doi.org/10.5281/zenodo.6383659.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abm6791
Supplementary Text
Figs. S1 to S131
Tables S1 to S37
References (*29–180*)
MDAR Reproducibility Checklist

ANTIBIOTIC DISCOVERY

Bioinformatic prospecting and synthesis of a bifunctional lipopeptide antibiotic that evades resistance

Zongqiang Wang[1]†, Bimal Koirala[1]†, Yozen Hernandez[1], Matthew Zimmerman[2], Sean F. Brady[1]*

Emerging resistance to currently used antibiotics is a global public health crisis. Because most of the biosynthetic capacity within the bacterial kingdom has remained silent in previous antibiotic discovery efforts, uncharacterized biosynthetic gene clusters found in bacterial genome–sequencing studies remain an appealing source of antibiotics with distinctive modes of action. Here, we report the discovery of a naturally inspired lipopeptide antibiotic called cilagicin, which we chemically synthesized on the basis of a detailed bioinformatic analysis of the *cil* biosynthetic gene cluster. Cilagicin's ability to sequester two distinct, indispensable undecaprenyl phosphates used in cell wall biosynthesis, together with the absence of detectable resistance in laboratory tests and among multidrug-resistant clinical isolates, makes it an appealing candidate for combating antibiotic-resistant pathogens.

The discovery and therapeutic development of natural product antibiotics, especially those produced by microbes, has substantially reduced mortality caused by bacterial infections (*1*). Unfortunately, this situation is challenged by antibiotic resistance, which is rising at a faster rate than the introduction of molecules with modes of action (MOAs) capable of circumventing existing resistance mechanisms (*2*, *3*). From a clinical development standpoint, nonribosomal peptide synthetase (NRPS)–encoded lipopeptides are an appealing potential source of future antibiotics because they have a history of inhibiting bacterial growth by diverse MOAs (*4*, *5*). Bacterial genome–sequencing efforts have uncovered a large number of biosynthetic gene clusters (BGCs) that do not appear to encode for known natural products, including many that are predicted to encode undescribed lipopeptides. These BGCs likely contain genetic instructions for the biosynthesis of antibiotics with diverse MOAs that could help to replenish antibiotic discovery pipelines. Unfortunately, most sequenced BGCs remain silent in the laboratory, and the molecules they encode remain a mystery. Here, we used a phylogenetic analysis of condensation starter (Cs) domain sequences, which introduce the acyl substituent into lipopeptides, to identify the cryptic *cil* BGC as a potential source of an uncharacterized lipopeptide antibiotic.

To identify BGCs that might encode lipopeptide antibiotics with distinctive MOAs, we collected NRPS BGCs from ~10,000 sequenced bacterial genomes. Clinically relevant lipopeptide antibiotics (e.g., polymyxin, daptomycin, etc.) have historically tended to be larger macrocyclic structures, and therefore BGCs predicted to encode peptides with fewer than five amino acids (i.e., containing fewer than five adenylation domains) were removed from this collection. A recent screen of this collection for BGCs that were predicted to encode congeners of known antibiotics guided our

[1]Laboratory of Genetically Encoded Small Molecules, The Rockefeller University, New York, NY 10065, USA. [2]Center for Discovery and Innovation, Hackensack Meridian Health, Nutley, NJ 07110, USA.
*Corresponding author. Email: sbrady@rockefeller.edu
†These authors contributed equally to this work.
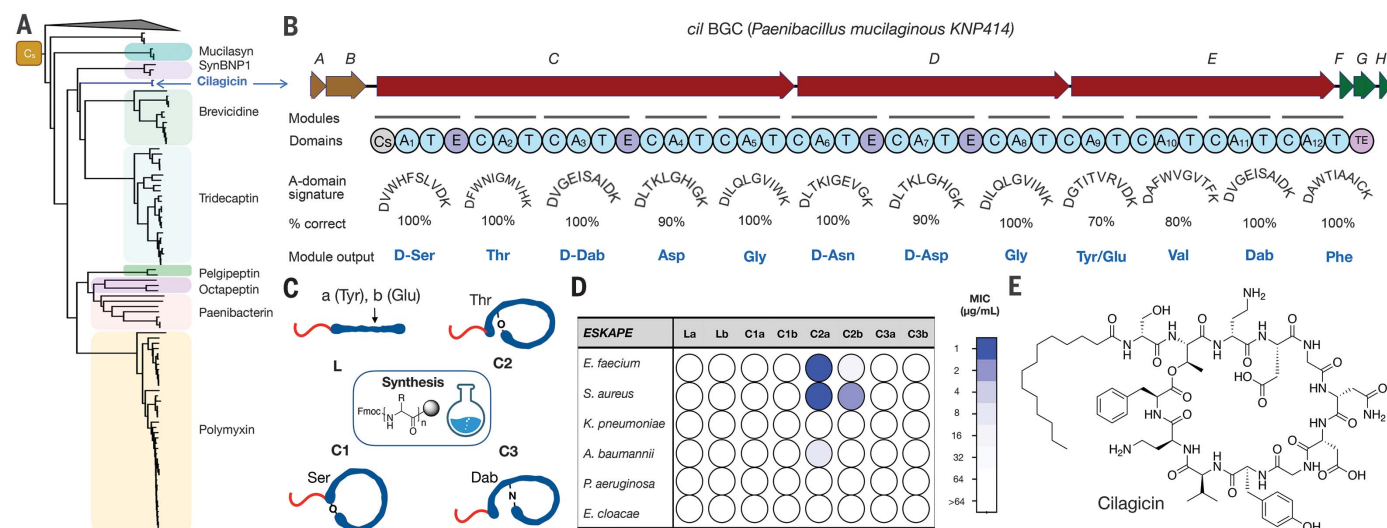
Fig. 1. Discovery of cilagicin. (**A**) Cs domains from sequenced NRPS BGCs were used to construct a phylogenetic tree. Clades associated with characterized antibiotic BGCs are labeled. The "orphan" cilagicin clade is labeled in blue. (**B**) The *cil* BGC contains three NRPS open reading frames (*cil* C to E). Biosynthesis of cilagicin is predicted to start from a Cs domain in CilC and then proceed using one A (adenylation)- and T (thiolation)-domain-containing initiation module, followed by 11 C (condensation)-, A-, and T-domain-containing extender modules. The substrate specificity of the *cil* BGC A domains was predicted on the basis of a comparison of each A domain's substrate-binding pocket against the 10–amino acid A-domain signature sequences found in functional characterized BGCs. E (epimerization) domains in modules 1, 3, 6, and 7 are predicted to result in the incorporation of D amino acids. The thioesterase (TE) domain at the end of *cil* E releases the mature structure from the final T domain as either a linear or cyclic product. (**C**) Diagrams of the four different peptide topologies that were synthesized from the linear peptide predicted to arise from the *cil* BGC. Position 9 was either Tyr (a) or Glu (b). L is a linear peptide. C1, C2, and C3 are cyclized through the C-terminal carboxyl group and Ser-1, Thr-2, or Dab-3, respectively. Dab, 2,4-diaminobutyric acid. (**D**) MIC data against the ESKAPE pathogens for the eight predicted synthetic structures depicted in (C). Concentrations tested ranged from 1 μg/ml (blue) to 64 μg/ml (white). Data are representative of three independent experiments. (**E**) Structure of the antibiotic cilagicin, which corresponds to C2a in (C).

synthesis of macolacin, a colistin analog that is effective against pathogens that express the *mcr-1* resistance gene (6). In the current analysis, we searched for uncharacterized BGC families. The key conserved feature across lipopeptides is the presence of an N-terminal lipid that is installed by a Cs domain (7–9). Among the sequenced large NRPS BGCs that we collected, we identified 3426 that contained a Cs domain. Cs domain sequences from these BGCs were used to construct a phylogenetic tree that guided our discovery efforts. As we have seen with other biosynthetic genes, sequences arising from BGCs sharing close common ancestors, and thus the same MOA, are likely to group together into the same clade (10–12). By extension, clades that do not contain any sequences from characterized BGCs would represent candidates for identifying structurally and mechanistically new antibiotics. The Cs domain phylogenetic tree contained a number of clades that were not associated with any characterized lipopeptides; however, one was particularly intriguing because it fell into a larger group of sequences in which most other clades were associated with antibiotic biosynthesis. These included BGCs for a number of clinically used, as well as clinically appealing, antibiotics (e.g., polymyxins, tridecaptins, and brevicidines). This "orphan" Cs clade that we identified contained three closely related sequences that arose from the same BGC found in two different sequenced *Paenibacillus mucilaginosus* strains (*KNP414* and *K02*) (Fig. 1A). On the basis of gene content and gene organization, this BGC, which we have called the *cil* BGC, did not resemble any characterized BGCs. Most sequenced BGCs remain silent in the laboratory, even when examined with advanced synthetic biology tools (13). With the power of modern synthetic organic chemistry and the increasing accuracy of natural product structure prediction algorithms, it is now possible to generate a bioactive molecule from the genetic instructions found in the primary sequence of a BGC. This was done by first bioinformatically predicting the encoded structure and then chemically synthesizing the predicted structure, i.e., producing a synthetic-bioinformatic natural product, syn-BNP (14–16). In this study, we used a syn-BNP approach to generate lipopeptide structures on the basis of the *cil* BGC and then tested these small molecules for antibacterial activity.

The *cil* BGC contains three NRPS open reading frames (*cil* C to E) that encode 12 distinct modules (Fig. 1B and table S1). The biosynthesis of a 12-mer lipopeptide is predicted to begin with the Cs domain at the N terminus of CilC and end with the thioesterase at the C terminus of CilE. The composition of each module's A-domain substrate-binding pocket (i.e., the substrate signature based on positions 235, 236, 239, 278, 299, 301, 322, 330, 331, and 517 of the A domain) was used to predict the 12 monomers used by this BGC (17). Eleven A domains had perfect or near perfect (80%) matches to characterized A domains, so we could make high-confidence predictions for the amino acid incorporated by these modules. The A-domain substrate signature from module 9 had equally good matches (70%) to two amino acids, Tyr and Glu. This analysis gave us two potential predicted linear lipopeptide products for the *cil* BGC: La and Lb (Fig. 1C). Epimerization domains found in modules 1, 3, 6, and 7 indicated that these amino acids appear in the D configuration in the *cil* BGC product. The absence of any genes predicted to encode tailoring enzymes (i.e., methyltransferase, hydroxylation, amino transferase, glycosyl transferase, etc.) within 10 kB of the *cil* NRPS genes suggested that the product of the *cil* BGC was not modified beyond the NRPS-produced lipopeptide (18).

Naturally occurring lipopeptides appear as either linear or cyclic structures. The predicted *cil* linear peptide contains three amino acids, D-Ser-1, Thr-2, and D-Dab-3, that could serve as nucleophiles for cyclization through the C-terminal carboxyl. Bringing together our linear peptide prediction and three potential cyclization sites yielded eight structures (two linear and six cyclic) that we predicted could arise from the *cil* BGC (Fig. 1C). Each of the eight potential BGC products was generated by solid-phase peptide synthesis (table S2). The *cil* BGC does not contain any lipid biosynthetic genes, so we predicted that the lipid found on the product of the *cil* BGC would arise directly from native fatty acid biosynthesis. Among the characterized bacterial lipopeptides, myristic acid is one of the most frequently seen simple lipids, so all synthetic peptides were N-terminal acylated with myristic acid.

All eight synthetic structures were assayed for activity against the ESKAPE pathogens (*Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa*, and *Enterobacter* species) (Fig. 1D and table S3). The 11–amino acid macrocycle that was cyclized through the hydroxyl of Thr-2 and contained Tyr at position 9 (compound C2a) showed potent antibacterial activity against the two Gram-positive ESKAPE pathogens [minimum inhibitory concentration (MIC), 1 μg/ml]. None of the remaining close analogs that we synthesized showed more potent activity against

Table 1. Activity of cilagicin against microorganisms and human cells.

Pathogens/human cells	Cilagicin MIC (μg/ml)*
Gram-positive	
Staphylococcus aureus USA300 (MRSA)	1
Staphylococcus aureus BAA1717(BRSA)	0.5
Enterococcus faecium EF18 (VRE)	0.5
Enterococcus faecalis AR785 (VRE)	0.5
Enterococcus gallinarum AR784 (VRE)	0.5
Enterococcus casseliflavus AR798 (VRE)	0.125
Streptococcus pneumoniae R[†]	0.5
Streptococcus pneumoniae Tigr4[†]	0.25
Clostridium difficile HM89[‡]	2
Clostridium difficile HM746[‡]	2
Streptococcus pyrogens ATCC19615	0.125
Streptococcus agalactiae BAA2675	1
Streptococcus agalactiae BAA1176	1
Bacillus subtilis 168A1	2
Gram-negative	
Acinetobacter baumannii ATCC17978	8
Escherichia coli BAS849	4
Escherichia coli ATCC25922	>64
Klebsiella pneumoniae ATCC13833	>64
Pseudomonas aeruginosa PAO1	>64
Enterobacter cloacae ATCC13047	>64
Human cell line	
HEK293	>64[‡]

*The MIC was tested by broth microdilution. †Bacteria were cultured under 5% CO_2. ‡Bacteria were cultured under anaerobic conditions. MRSA, methicillin-resistant *S. aureus*; BRSA, bacitracin-resistant *S. aureus*; VRE, vancomycin-resistant Enterococci.

these strains. We have called the active structure cilagicin (Fig. 1E).

Cilagicin was active against all Gram-positive pathogens that we tested (Table 1). It was also active against a number of difficult-to-treat vancomycin-resistant *Enterococci* pathogens, as well as *Clostridioides difficile*, both of which are considered urgent and serious threat pathogens by the US Centers for Disease Control and Prevention (CDC) (*19*). It was also active against all antibiotic-resistant Gram-positive pathogens that we tested. It

maintained potent activity against a panel of 19 *S. aureus* strains that showed different patterns of resistance to clinically relevant families of antibiotics (table S4). Further, in stark contrast to US Food and Drug Administration (FDA)–approved antibiotics, cilagicin maintained potent activity against all strains found in a panel of 30 vancomycin-resistant *Enterococci* clinical isolates (table S5). This collection is highly enriched in multidrug-resistant (MDR) isolates, with more than half exhibiting resistance to between five and eight different clinically used antibiotics. Cilagicin was largely inactive against Gram-negative bacteria, with the exception of *A. baumannii* (table S6) and outer membrane–permeabilized *Escherichia coli* BAS849, suggesting that the outer membrane of Gram-negative bacteria blocks cilagicin's access to its target. Even at the highest concentration we tested, cilagicin did not show human cell line cytotoxicity (Table 1).

In a time-dependent killing curve analysis, cilagicin was found to be bactericidal and to reduce the number of viable bacteria by more than four orders of magnitude after 4 hours (Fig. 2A). Electron microscopy images of cilagicin-treated cells showed cell collapse over time (Fig. 2B). In an effort to elucidate cilagicin's MOA, we tried to raise mutants by direct plating of *S. aureus* on cilagicin. In these direct plating experiments, we never observed any colonies that showed more than a onefold increase in MIC. To explore the possibility of cilagicin having a detergent-like activity, we tested it for membrane depolarization and cell lytic activities using 3,3'-dipropylthiadicarbocyanine iodide [DiSC$_3$(5)]– and SYTOX–based fluorescence assays, respectively (Fig. 2, C and D) (*20–22*). No response was detected in either assay when *S. aureus* was exposed to even eightfold the MIC of cilagicin, ruling out membrane disruption as its MOA.

Cilagicin is a zwitterion with two positively charged residues, 3-D-Dab and 11-D-Dab, and two negatively charged residues, 4-Asp and 7-D-Asp. Charged lipopeptide antibiotics often do not enter the cell and instead function by disrupting synthesis of the cell wall outside the cell membrane (*11, 23, 24*). Antibiotics that block peptidoglycan biosynthesis lead to the accumulation of the lipid II precursor UDP-MurNAc-pentapeptide, which is easily detected by liquid chromatography–mass spectroscopy (LC-MS) in antibiotic-exposed cultures (*11, 25–27*). LC-MS analysis of *S. aureus* cultures exposed to cilagicin (1× MIC) showed an obvious accumulation of UDP-MurNAc-pentapeptide (Fig. 2E). Because it is often much more difficult to alter a small-molecule target than a protein target through genomic mutations, our inability to identify cilagicin-resistant mutants hinted at the binding of a small molecule instead of a protein as the mode of inhibiting cell wall biosynthesis. To identify the
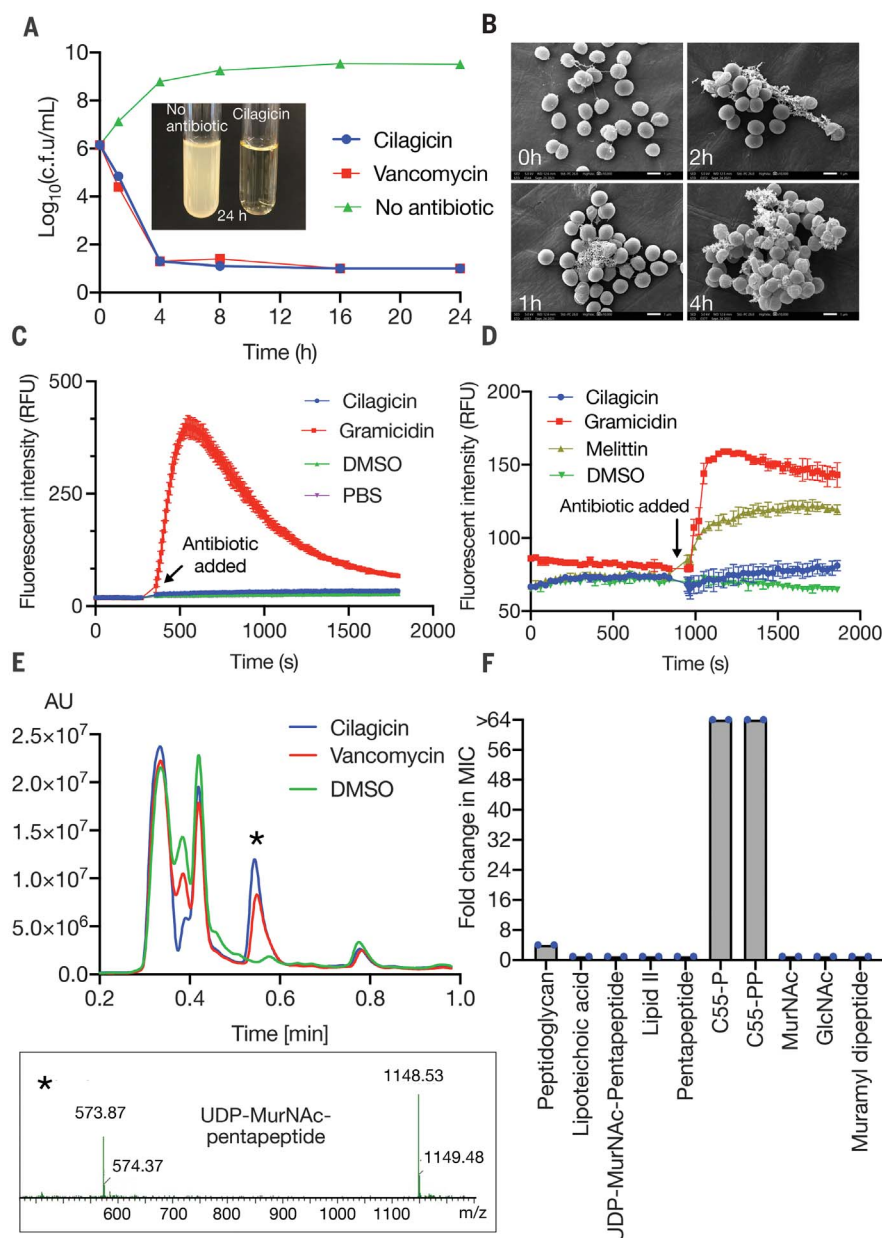
Fig. 2. Cilagicin mode of action. (**A**) Survival of *S. aureus* USA300 after timed exposure to 10× the MIC of cilagicin. Dimethyl sulfoxide (DMSO) and vancomycin (10× MIC) were included as controls. CFUs were counted three independent times and are plotted as mean ± SD. (**B**) Scanning electron microscopy image of *S. aureus* USA300 cultures treated with cilagicin. Conditions were the same as in (A). (**C** and **D**) Cell lysis (C) or membrane depolarization (D) in cilagicin-treated *S. aureus* cultures was monitored using SYTOX and DiSC3(5) dyes, respectively. Data are presented as the mean of three independent experiments ± SD. (**E**) Accumulation of UDP-MurNAc-pentapeptide after treating *S. aureus* cultures with cilagicin (1× MIC) was monitored by LC-MS. DMSO- and vancomycin (10× MIC)–treated cultures were examined as controls. UDP-MurNAc-pentapeptide corresponds to [M-H]$^-$ = 1148.53 and [M-2H]$^{2-}$ = 573.87. (**F**) Fold change in cilagicin MIC upon treatment of *S. aureus* with fivefold molar excess of different lipid II intermediates. The peptidoglycan mixture was added at 100 μg/ml. Data are representative of two independent experiments.

metabolite(s) that interacts with cilagicin, we screened a series of lipid II intermediates for their ability to suppress cilagicin's antibacterial activity (Fig. 2F). In these studies, the MIC of cilagicin against *S. aureus* USA300 was determined in the presence of a fivefold molar excess of each metabolite. Two of the compounds that we tested, undecaprenyl phosphate (C55-P) and undecaprenyl pyrophosphate (C55-PP), showed dose-dependent inhibition of cilagicin's antibacterial activity (Fig. 3, A and B).

C55-P is a monophosphorylated, 55-carbon-long isoprene that is essential for transporting intermediates in the biosynthesis of cell wall carbohydrate polymers (e.g., peptidoglycan, O antigen, teichoic acids, etc.) across the bacteria cell membrane (*28, 29*). C55-PP is the diphosphorylated version of the same 55-carbon isoprene. It is both produced de novo and recycled from C55-P during the biosynthesis of the cell wall. Its dephosphorylation by membrane-embedded pyrophosphatases generate the cellular pool of C55-P that is required for cell wall synthesis (*30*). Using isothermal titration calorimetry, we observed that cilagicin bound both C55-P and C55-PP (Fig. 3, C and D), but a representative inactive analog from our initial synthesis studies, cilagicin-3b, did not bind either compound. Collectively, our MOA studies are consistent with cilagicin being able to sequester both C55-P and C55-PP and thus acting as a bifunctional antimicrobial.

Bacteria only have a small pool of free undecaprenyl carrier lipids (~10^5 molecules per cell) to use in the transfer of critical biosynthetic intermediates across the cell membrane (*31*). Although disruption of this process is an appealing antibacterial MOA, it remains underexploited clinically because only a few antibiotics have been identified that bind even one undecaprenyl phosphate. These include the antibiotics bacitracin and tripropeptin, which specifically bind C55-PP in a zinc- and calcium-dependent manner, respectively (*32, 33*). The only known antibiotics that bind C55-P are the calcium-dependent lipopeptide friulimicin and its close congeners (e.g., amphomycin and laspartomycin) (*26, 27, 34*). Binding C55-P directly reduces the amount of available C55-P, whereas sequestering C55-PP indirectly reduces C55-P by preventing C55-PP dephosphorylation. In either case, this disrupts the flow of peptidoglycan precursors into the cell wall, ultimately leading to cell death (Fig. 3E).

Bacitracin is used clinically as a topical antibiotic, and friulimicin is in development for use in animal health. Unfortunately, bacteria exposed to antibiotics that bind a single undecaprenyl phosphate are reported to readily develop resistance. Antibiotics with multiple molecular targets tend to have reduced rates of resistance because of the difficulty associated with altering multiple targets simultaneously. Therefore, we predicted that in
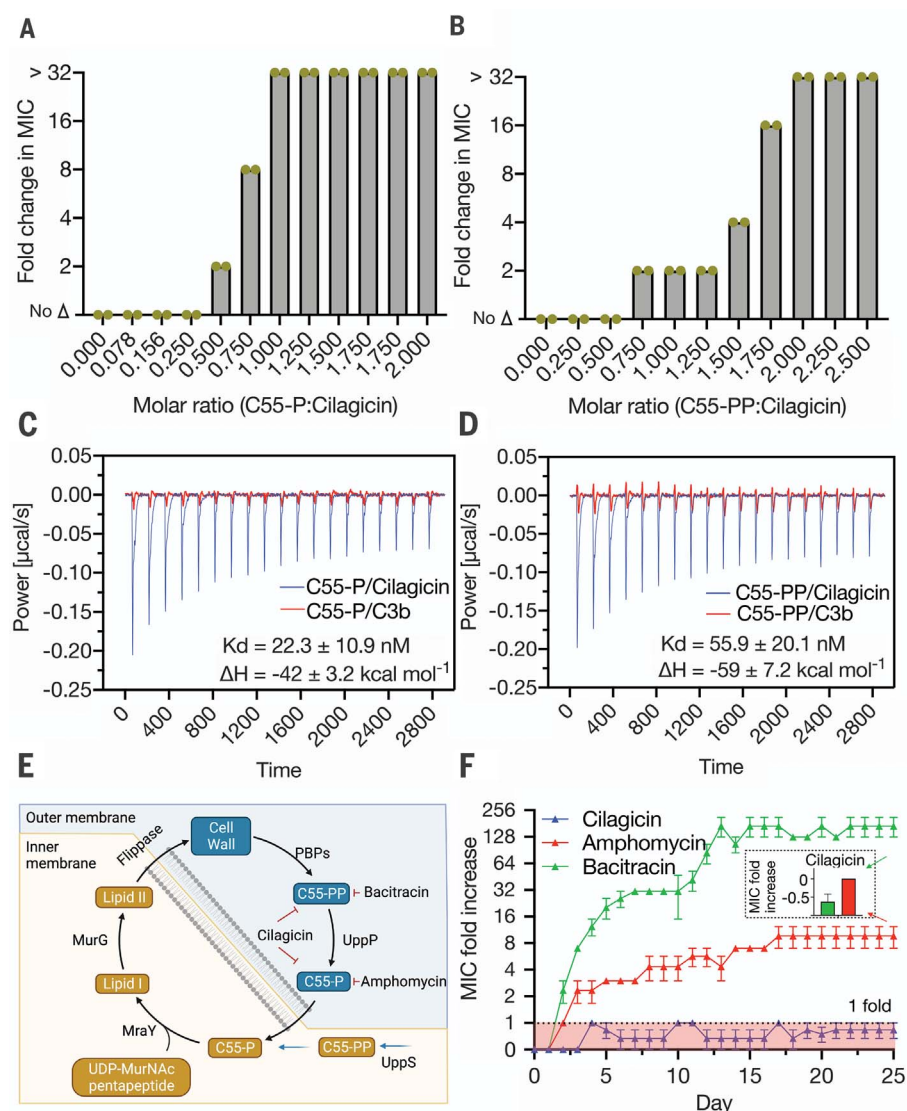
Fig. 3. Interaction of cilagicin with C55-P and C55-PP. (**A** and **B**) Fold change in MIC of cilagicin-treated cultures of *S. aureus* USA300 in the presence of different concentrations of C55-P (A) or C55-PP (B). The highest concentration tested was 32× the MIC. Data from two independent experiments are presented. (**C** and **D**) Isothermal titration calorimetry data for cilagicin or its inactive analog C3b interacting with either C55-P (C) or C55-PP (D). Two independent experiments were performed with similar results. (**E**) Diagram of the role of C55-P and C55-PP in Gram-positive cell wall biosynthesis. (**F**) Resistance acquisition during serial passaging of *S. aureus* USA300 in the presence of sub-MIC levels of cilagicin, bacitracin, or amphomycin. Data shown represent the mean of three independent experiments ± SEM. Inset: Fold increase in the MIC of cilagicin against bacitracin (green)– and amphomycin (red)–resistant strains at 25 days.

the case of cilagicin, its ability to bind both undecaprenyl phosphates (i.e., two distinct small molecules) would lead to a reduced resistance rate compared with antibiotics that bind a single phosphorylated undecaprenyl moiety. Because we had failed to identify mutants resistant to cilagicin by direct plating on antibiotic-containing media, we attempted to raise *S. aureus*–resistant mutants by daily serial passage in the presence of sub-MIC levels of antibiotic using cilagicin, bacitracin, or the friulimicin congener amphomycin to allow a direct comparison of resistance rates for antibiotics

that bind either one or two phosphorylated undecaprenyl moieties. *S. aureus* rapidly developed resistance to both bacitracin and amphomycin. MICs for these antibiotics increased by eightfold to 256-fold, respectively, during the course of the serial passage experiment. By contrast, after 25 days of constant exposure to cilagicin, we observed no higher than a doubling of the original MIC (Fig. 3F). In addition, neither the highly bacitracin-resistant mutants nor the highly amphomycin-resistant mutants that we generated showed cross-resistance to cilagicin. The *cil* BGC is found in the genome of
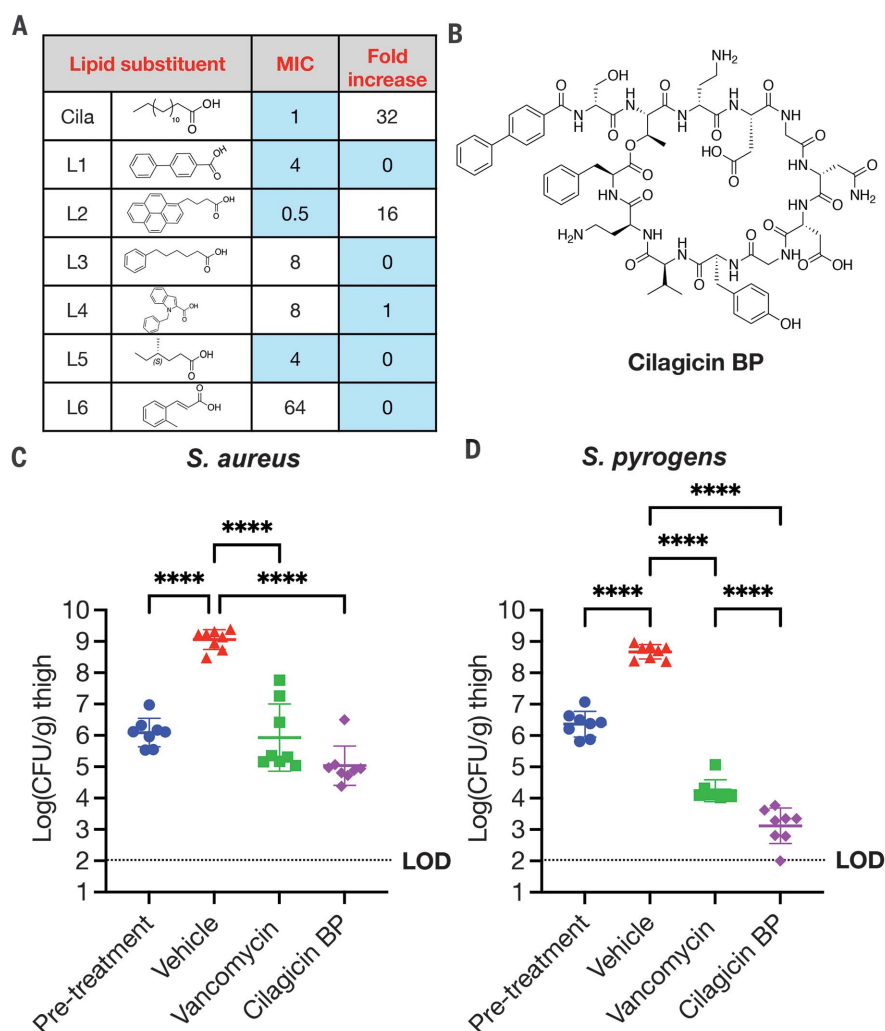
Fig. 4. Cilagicin BP activity in a murine neutropenic thigh infection model. (**A**) Anti–*S. aureus* activity of cilagicin analogs with different lipid substituents in the presence of 10% serum. Blue indicates MIC <4 µg/ml or no change in MIC in the presence of serum. (**B**) Structure of cilagicin BP (L1). (**C**) Neutropenic thigh infection model using *S. aureus* USA300. (**D**) Neutropenic thigh infection model using *S. pyrogens* ATCC19615. Two hours after infection with a fresh bacterial suspension (1 × 10⁶ CFUs), vehicle (10% DMSO, TID), vancomycin (40 mg/kg, TID), or cilagicin BP (40 mg/kg, TID) was delivered by intraperitoneal injection. Twenty-four hours after infection, CFUs were determined from homogenized thigh tissue samples. Significant differences between groups were analyzed by one-way analysis of variance (ANOVA) (***P = 0.0001, ****P < 0.0001) (n = 4 mice, n = 8 thighs). Mean CFU counts and SDs are shown.

development of resistance has often proved difficult in laboratory experiments using individual pathogens (*39*). Historically, once these antibiotics are exposed to the global pool of resistance determinants present in the clinical setting, resistance might appear, albeit often at a much slower rate than is seen for other MOAs. Although we have not yet observed cilagicin resistance in MDR clinical isolates, this does not rule out the eventual appearance of resistance with broader environmental exposure.

Our initial pharmacological assessment in mice showed that cilagicin had high plasma bioavailability when delivered by intraperitoneal injection (Fig. S2). However, it did not reduce bacterial burden in an animal infection model. We subsequently observed that cilagicin's antibacterial activity was significantly suppressed in the presence of serum, suggesting that high serum binding might have also limited its in vivo activity (Fig. 4A). We therefore attempted to generate a cilagicin analog with reduced serum binding. To do this, we created a collection of cilagicin analogs with different N-terminal lipids and compared their MICs in the presence and absence of serum. An analog containing a biphenyl N-terminal substituent, cilagicin BP, maintained good antimicrobial activity and showed no increase in MIC in the presence of serum (Fig. 4B and table S7). Similar biphenyl substituents are found in several synthetically optimized natural product antibiotics, including glycopeptides (oritavancin) and other lipopeptides (macolacin) (*6*, *40*). This change of lipid substituent did not alter the antibiotics' MOA (Fig. S3A), and cilagicin BP continued to show no hemolytic activity and no human cell cytotoxicity (Fig. S3, B and C). We assessed the in vivo efficacy of cilagicin BP using a neutropenic mouse thigh model. At 24 hours after infection, cilagicin BP showed significant antibacterial activity against *S. aureus* USA300 at 40 mg/kg given three times a day (TID), resulting in an almost 4 log₁₀ reduction in colony-forming units (CFUs) compared with the vehicle control (Fig. 4C). Next, we evaluated cilagicin BP against *S. pyrogens* ATCC19615 in the same neutropenic thigh model. In this case, cilagicin BP showed an even more impressive reduction (>5 log₁₀) in bacterial burden compared with the vehicle control, which was consistent with the lower MIC values for this pathogen in vitro (Fig. 4D). Cilagicin BP resulted in more than a log greater reduction of bacterial burden than vancomycin against *S. pyrogens*.

Cilagicin BP's mode of action, absence of detectable resistance, and in vivo activity make it an appealing lead structure for the development of a next-generation antibiotic that may help to address the growing antibiotic resistance crisis. As seen with the characterization of biologically produced natural products, the

P. mucilaginosus. The genus *Paenibacillus* is Gram variable. *P. mucilaginosus*'s reported negative Gram staining suggests that it contains an outer membrane that could protect it from cilagicin's toxicity, thus potentially eliminating the need for self-resistance elements to have evolved in nature (*35*, *36*).

Bacitracin resistance arises from the release of C55-PP from bacitracin by the ABC transporter BceAB, increased production of C55-PP by the undecaprenyl-pyrophosphate phosphatase BcrC, or by a still ill-defined mechanism associated with phage shock protein–like protein expression (LiaI and LiaH) (*37*). In the

case of C55-P binding, antibiotic resistance is associated with the expression of the cell envelope stress response regulator (*38*), which provides only low-level resistance compared with what is seen for bacitracin-resistance mechanisms (less than eightfold versus more than 254-fold MIC increases, respectively). Because the binding of C55-P alone appears to be difficult to overcome, it was not surprising that sequestration of the entire pool of undecaprenyl phosphates by cilagicin would further reduce the propensity for resistance to develop. In general, the sequestration of an essential extracellular metabolite is an appealing MOA because the

study of syn-BNPs that are inspired by un-explored BGC families should prove to be a productive strategy for identifying distinctive scaffolds that can serve as lead structures for developing antibiotics with diverse MOAs.

REFERENCES AND NOTES

1. S. E. Rossiter, M. H. Fletcher, W. M. Wuest, *Chem. Rev.* **117**, 12415–12474 (2017).
2. J. Davies, D. Davies, *Microbiol. Mol. Biol. Rev.* **74**, 417–433 (2010).
3. B. Aslam *et al.*, *Infect. Drug Resist.* **11**, 1645–1658 (2018).
4. I. W. Hamley, *Chem. Commun.* **51**, 8574–8583 (2015).
5. J. M. Raaijmakers, I. De Bruijn, O. Nybroe, M. Ongena, *FEMS Microbiol. Rev.* **34**, 1037–1062 (2010).
6. Z. Wang *et al.*, *Nature* **601**, 606–611 (2022).
7. K. Bloudoff, T. M. Schmeing, *Biochim. Biophys. Acta. Proteins Proteomics* **1865** (11 Pt B), 1587–1604 (2017).
8. C. Rausch, I. Hoof, T. Weber, W. Wohlleben, D. H. Huson, *BMC Evol. Biol.* **7**, 78 (2007).
9. N. Roongsawang *et al.*, *FEMS Microbiol. Lett.* **252**, 143–151 (2005).
10. H. S. Kang, S. F. Brady, *ACS Chem. Biol.* **9**, 1267–1272 (2014).
11. B. M. Hover *et al.*, *Nat. Microbiol.* **3**, 415–422 (2018).
12. Z. Wang *et al.*, *Angew. Chem.* **133**, 22346–22351 (2021).
13. H. A. Tomm, L. Ucciferri, A. C. Ross, *J. Ind. Microbiol. Biotechnol.* **46**, 1381–1400 (2019).
14. J. Chu *et al.*, *Nat. Chem. Biol.* **12**, 1004–1006 (2016).
15. J. Chu *et al.*, *J. Am. Chem. Soc.* **142**, 14158–14168 (2020).
16. J. Chu, X. Vila-Farres, S. F. Brady, *J. Am. Chem. Soc.* **141**, 15737–15741 (2019).
17. T. Stachelhaus, H. D. Mootz, M. A. Marahiel, *Chem. Biol.* **6**, 493–505 (1999).
18. C. T. Walsh *et al.*, *Curr. Opin. Chem. Biol.* **5**, 525–534 (2001).
19. Centers for Disease Control and Prevention, "Antibiotic resistance threats in the United States 2019" (CDC, 2019); https://www.cdc.gov/drugresistance/pdf/threats-report/2019-ar-threats-report-508.pdf.
20. B. L. Roth, M. Poot, S. T. Yue, P. J. Millard, *Appl. Environ. Microbiol.* **63**, 2421–2431 (1997).
21. G. Cabrini, A. S. Verkman, *J. Membr. Biol.* **92**, 171–182 (1986).
22. M. Wu, E. Maier, R. Benz, R. E. Hancock, *Biochemistry* **38**, 7235–7242 (1999).
23. C. Wu, Z. Shang, C. Lemetre, M. A. Ternei, S. F. Brady, *J. Am. Chem. Soc.* **141**, 3910–3919 (2019).
24. S. A. Cochrane *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 11561–11566 (2016).
25. L. L. Ling *et al.*, *Nature* **517**, 455–459 (2015).
26. T. Schneider *et al.*, *Antimicrob. Agents Chemother.* **53**, 1610–1618 (2009).
27. L. H. Kleijn *et al.*, *J. Med. Chem.* **59**, 3569–3574 (2016).
28. T. Touzé, D. Mengin-Lecreulx, *Ecosal Plus* **3**, 10.1128/ecosalplus.4.7.1.7 (2008).
29. J. van Heijenoort, *Nat. Prod. Rep.* **18**, 503–519 (2001).
30. M. El Ghachi *et al.*, *Nat. Commun.* **9**, 1078 (2018).
31. V. M. Hernández-Rocamora *et al.*, *Cell Surf.* **2**, 1–13 (2018).
32. H. Hashizume *et al.*, *Antimicrob. Agents Chemother.* **55**, 3821–3828 (2011).
33. K. J. Stone, J. L. Strominger, *Proc. Natl. Acad. Sci. U.S.A.* **68**, 3223–3227 (1971).
34. M. Singh, J. Chang, L. Coffman, S. J. Kim, *Sci. Rep.* **6**, 31757 (2016).
35. X. F. Hu *et al.*, *Int. J. Syst. Evol. Microbiol.* **60**, 8–14 (2010).
36. D. Goswami, S. Parmar, H. Vaghela, P. Dhandhukia, J. N. Thakker, *Cogent Food Agric.* **1**, 1000714 (2015).
37. J. Radeck *et al.*, *Mol. Microbiol.* **100**, 607–620 (2016).
38. T. Wecke *et al.*, *Antimicrob. Agents Chemother.* **53**, 1619–1623 (2009).
39. D. G. J. Larsson, C. F. Flach, *Nat. Rev. Microbiol.* **20**, 257–269 (2021).
40. K. D. Brade, J. M. Rybak, M. J. Rybak, *Infect. Dis. Ther.* **5**, 1–15 (2016).
41. Z. Wang, B. Koirala, Y. Hernandez, M. Zimmerman, S. F. Brady, Code for: Bioinformatic prospecting and synthesis of a bifunctional lipopeptide antibiotic that evades resistance, Zenodo (2022); https://doi.org/10.5281/zenodo.5793911.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abn4213
Materials and Methods
Figs. S1 to S9
Tables S1 to S7
References (*42–45*)
MDAR Reproducibility Checklist

View/request a protocol for this paper from *Bio-protocol*.

IMMUNOLOGY

The CD8α–PILRα interaction maintains CD8+ T cell quiescence

Linghua Zheng[1], Xue Han[1], Sheng Yao[1], Yuwen Zhu[1], John Klement[2], Shirley Wu[2], Lan Ji[1], Gefeng Zhu[1], Xiaoxiao Cheng[1], Zuzana Tobiasova[1], Weiwei Yu[1], Baozhu Huang[1], Matthew D. Vesely[1], Jun Wang[1], Jianping Zhang[1], Edward Quinlan[1], Lieping Chen[1]*

T cell quiescence is essential for maintaining a broad repertoire against a large pool of diverse antigens from microbes and tumors, but the underlying molecular mechanisms remain largely unknown. We show here that CD8α is critical for the maintenance of CD8+ T cells in a physiologically quiescent state in peripheral lymphoid organs. Upon inducible deletion of CD8α, both naïve and memory CD8+ T cells spontaneously acquired activation phenotypes and subsequently died without exposure to specific antigens. PILRα was identified as a ligand for CD8α in both mice and humans, and disruption of this interaction was able to break CD8+ T cell quiescence. Thus, peripheral T cell pool size is actively maintained by the CD8α–PILRα interaction in the absence of antigen exposure.

CD8+ T cells are major effector cells within the immune system for eliminating infected and malignant cells. Therefore, it is important to maintain a large and diverse repertoire of CD8+ T cells within lymphoid organs prior to their enlistment. T cells are actively maintained in a metabolically inactive or quiescent state to ensure survival (1). Genetic deficiency of the intracellular proteins FOXP1, TSC1, SLFN2, and BTG1/BTG2, as well as inducible loss of regulatory T cells, has been shown to abrogate T cell quiescence (1–3). However, the extracellular signals that sustain CD8+ T cell quiescence remain unknown.

The transmembrane protein CD8 is a heterodimer composed of α and β subunits linked by disulfide bonds and serves as a lineage marker to identify a major subset of T cells with cytotoxic capacity (4, 5). CD8α interacts with the major histocompatibility complex class I (MHC I) via its extracellular domain and stabilizes lymphocyte-specific protein tyrosine kinase (Lck) via its intracellular domain (6, 7). CD8 is thought to primarily function by delivering Lck to the T cell receptor (TCR) to strengthen TCR signaling upon antigen stimulation (8). CD8α is also required for CD8+ T cell development in the thymus because genetic deletion of CD8α leads to a nearly complete elimination of cytotoxic T cells (9).

To study the potential functions of CD8α in the periphery, we generated $Cd8a^{loxp/loxp}$ mice, which were subsequently backcrossed with tamoxifen-inducible Cre-ER^{T2} transgenic mice (fig. S1, A and B). CD8+ T cells purified from $Cre^{+/+}Cd8a^{loxp/loxp}$ mice were adoptively transferred into wild-type (WT) mice followed by tamoxifen treatment. This allowed the function of CD8α on peripheral CD8+ T cells to be studied while avoiding any effects on thymic development. As expected, tamoxifen treatment abolished the expression of both CD8α and CD8β on adoptively transferred CD8+ T cells (fig. S1C). Thus, CD8β surface expression depends on CD8α expression, consistent with a previous report (10). Both TCR and CD3 expression remained unchanged after CD8α deletion (fig. S1C). Tamoxifen-induced deletion of CD8α led to the progressive loss of CD8+ T cells, especially those proliferating cells in the blood, lymph node, and spleen (fig. S2, A to D). Thirty-four days after tamoxifen treatment, memory CD8+ T cell survival was decreased by 30 to 40% (Fig. 1, A and B). Similarly, naïve CD8+ T cell survival was decreased by 30 to 40% (Fig. 1, C and D) and 80% (fig. S3), respectively, 36 and 55 days after tamoxifen treatment. By contrast, without tamoxifen treatment, CD8+ T cells ($Cre^{+/+}Cd8a^{+/+}$ or $Cre^{+/+}Cd8a^{loxp/loxp}$) showed equivalent survival (fig. S4, A and B). There was also increased expression of the T cell activation markers CD69 and Fas (fig. S5, A and B, and Fig. 2, A and B) and decreased expression of CD127, CD122, and CD5 (fig. S5, C to E, and Fig. 2, C to E) on $Cd8a^{-/-}$ memory T cells. Similarly, inducible CD8α deletion resulted in the up-regulation of CD69 and Fas (fig. S6, A and B) and the down-regulation of CD127, CD122, and CD5 (fig. S6, C to E) on naïve CD8+ T cells. Moreover, the inducible deletion of CD8α increased the proportion of naïve T cells that exited the G0 phase (fig. S7), indicative of T cell activation. Thus, upon inducible CD8α deletion, CD8+ T cells acquire a loss-of-quiescence phenotype without specific antigen stimulation. Because cell surface CD5 transmits a tuning signal for TCR (11), decreased CD5 expression on $Cd8a^{-/-}$ T cells suggested dampened TCR signaling. Similarly, Nur77, an immediate early protein downstream of TCR

pathway (12), was also decreased in naïve and memory $Cd8a^{-/-}$ T cells (fig. S8, A and B). Thus, the activation or loss of quiescence induced by CD8α deletion is not caused by TCR triggering.

The decrease of $Cd8a^{-/-}$ T cells observed in lymphoid organs was not due to their migration or redistribution to other tissues because there was no detectable increase of $Cd8a^{-/-}$ T cells in nonlymphoid tissues (fig. S9, A and B). Cleaved caspase-3 expression was increased in $Cd8a^{-/-}$ T cells (fig. S10, A and B), indicating activation of caspase-3. Furthermore, deletion of CD8α resulted in elevated Fas expression by T cells (Fig. 2B and fig. S6B), and anti-Fas antibody treatment ex vivo induced significantly more apoptosis of $Cd8a^{-/-}$ naïve or memory T cells (fig. S11, A and B). Thus, the decrease of CD8+ T cells upon CD8α deletion may be caused by programmed cell death.

To validate these findings, we generated an anti-mouse CD8α monoclonal antibody (mAb, clone 3D9). Mice were thymectomized before 3D9 inoculation to specifically study the effect of 3D9 on mature peripheral T cells. 3D9 decreased the survival and increased the activation of CD8+ T cells in thymectomized mice (fig. S12, A to C). 3D9 also enhanced the cell cycle transition from G0 to G1 by transferred naïve CD8+ T cells (fig. S13). Thus, 3D9 induces a loss-of-quiescence phenotype similar to that caused by CD8α deletion. To exclude an agonistic or depletion effect of 3D9, the genes encoding the Fab fragment of 3D9 were cloned and expressed in vivo by intravenous hydrodynamic gene delivery (13). 3D9 Fab treatment reduced the survival of transferred memory (fig. S14, A and B) and naïve (fig. S14, A and C) CD8+ T cells and increased the activation of memory (fig. S15, A and B) and naïve (fig. S16, A and B) CD8+ T cells.

The loss of quiescence by genetic deletion or antibody blockade suggested that CD8α may interact with an unknown ligand to mediate this function. We used a high-throughput genome-scale receptor array (14–16) to search this ligand by testing the binding of human CD8α to around 6000 human transmembrane proteins that were individually displayed on the surface of 293T cells. The paired immunoglobulin-like type 2 receptor alpha (PILRα) was identified as a binding partner (Fig. 3A), and this interaction was confirmed by flow cytometry and enzyme-linked immunosorbent assay (ELISA) (Fig. 3, B to D). This CD8α–PILRα interaction was also observed in mice (Fig. 3, E to G). The CD8α–PILRα interaction appeared to be specific because CD8α did not bind PILRβ, a close homolog of PILRα (17), in either mice or humans (fig. S17, A and B). Both 3D9 mAb and 3D9 Fab could compete with CD8α binding to PILRα (Fig. 3G and fig. S17C), suggesting that the effect of 3D9 or its Fab fragment in vivo may have been due to a blockade of the CD8α–PILRα interaction.

[1]Department of Immunobiology, Yale University School of Medicine, New Haven, CT, USA. [2]Yale College, Yale University, New Haven, CT, USA.
*Corresponding author. Email: lieping.chen@yale.edu
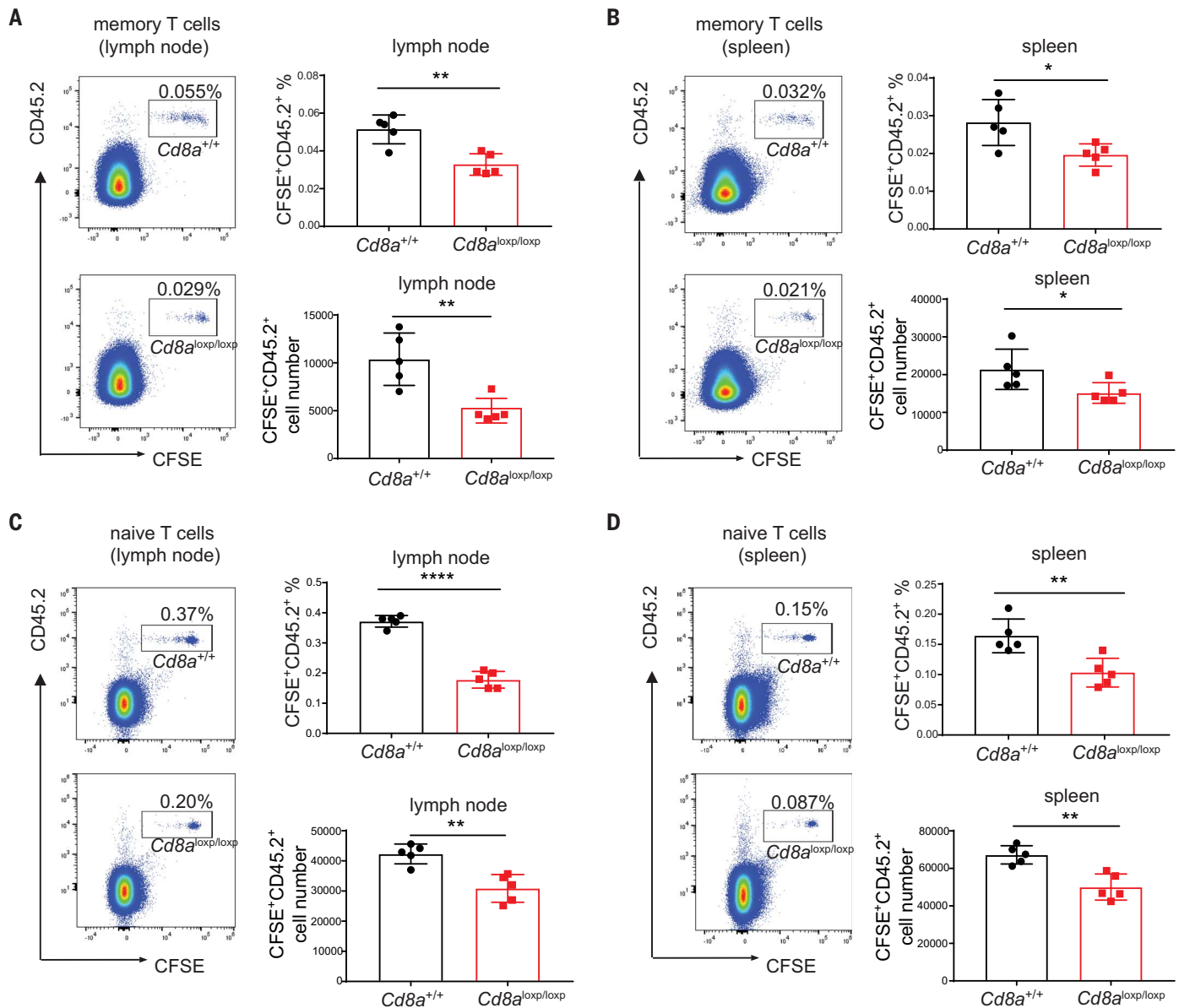
Fig. 1. Inducible genetic deletion of *Cd8a* disrupts the homeostasis of memory and naive CD8⁺ T cells in the periphery. (**A** and **B**) Memory CD44^hiCD8⁺ T cells isolated by fluorescence-activated cell sorting (FACS) from *Cre^+/+Cd8a^+/+* (*Cd8a^+/+*) or *Cre^+/+Cd8a^loxp/loxp* (*Cd8a^loxp/loxp*) mice were CFSE labeled and transferred into B6 congenic CD45.1 mice followed by tamoxifen treatment. The percentage and absolute number of CD45.2⁺CFSE⁺ cells in the lymph node (A) and spleen (B) cells were examined 34 days after tamoxifen treatment. (**C** and **D**) FACS-isolated naïve CD44^loCD8⁺ T cells from *Cre^+/+Cd8a^+/+* (*Cd8a^+/+*) or *Cre^+/+Cd8a^loxp/loxp* (*Cd8a^loxp/loxp*) mice were CFSE labeled and transferred into B6 congenic CD45.1 mice followed by tamoxifen treatment. The percentage and absolute number of CD45.2⁺CFSE⁺ cells in the lymph node (C) and spleen (D) cells were examined 36 days after tamoxifen treatment. (A to D) (left panel) Representative flow cytometry plots of CD45.2⁺CFSE⁺ percentage in the lymph node or spleen of each group 34 (A and B) and 36 (C and D) days after tamoxifen treatment; (right panel): bar graphs including all mice from each group. *n* = 5 mice per group in (A) to (D). Data shown in (A) to (D) are representative of at least three independent experiments. Mean ± SD is shown. *P < 0.05, **P < 0.01, ****P < 0.0001 by unpaired Student's *t* test.

PILRα is a transmembrane protein broadly expressed on myeloid cells, including macrophages, dendritic cells, and neutrophils (*18, 19*). To evaluate its role in maintaining CD8⁺ T cell quiescence, we generated a specific mAb against mouse PILRα (clone 9B12) that could block the PILRα–CD8α interaction (Fig. 3G). Administration of 9B12 to B6 WT mice decreased CD8⁺ T cell survival in the lymph node as early as day 1 after antibody treatment, whereas de-creased CD8⁺ T cell survival was not observed in the spleen until day 15 (fig. S18A and Fig. 4A). Similar results were obtained with thymectomized mice (fig. S18B and Fig. 4B), indicating that the effect of 9B12 was not dependent on the thymus. The decrease of CD8⁺ T cells was not due to mAb-mediated depletion because PILRα is not expressed by T cells (*19*). In contrast to the decrease in CD8⁺ T cells, CD4⁺ T cell survival appeared to be unaffected by 9B12 treatment (fig. S18C and Fig. 4C). In addition, blockade of PILRα with 9B12 resulted in increased CD8⁺ T cell activation with increased expression of CD69 and interferon-γ (IFN-γ) (Fig. 4, D and E). CD8⁺ T cells from β-actin luciferase transgenic mice were purified and transferred into B6 albino mice to facilitate real-time imaging analysis in vivo. Transferred CD8⁺ T cells remained in the lymphoid organs, and their total cell number
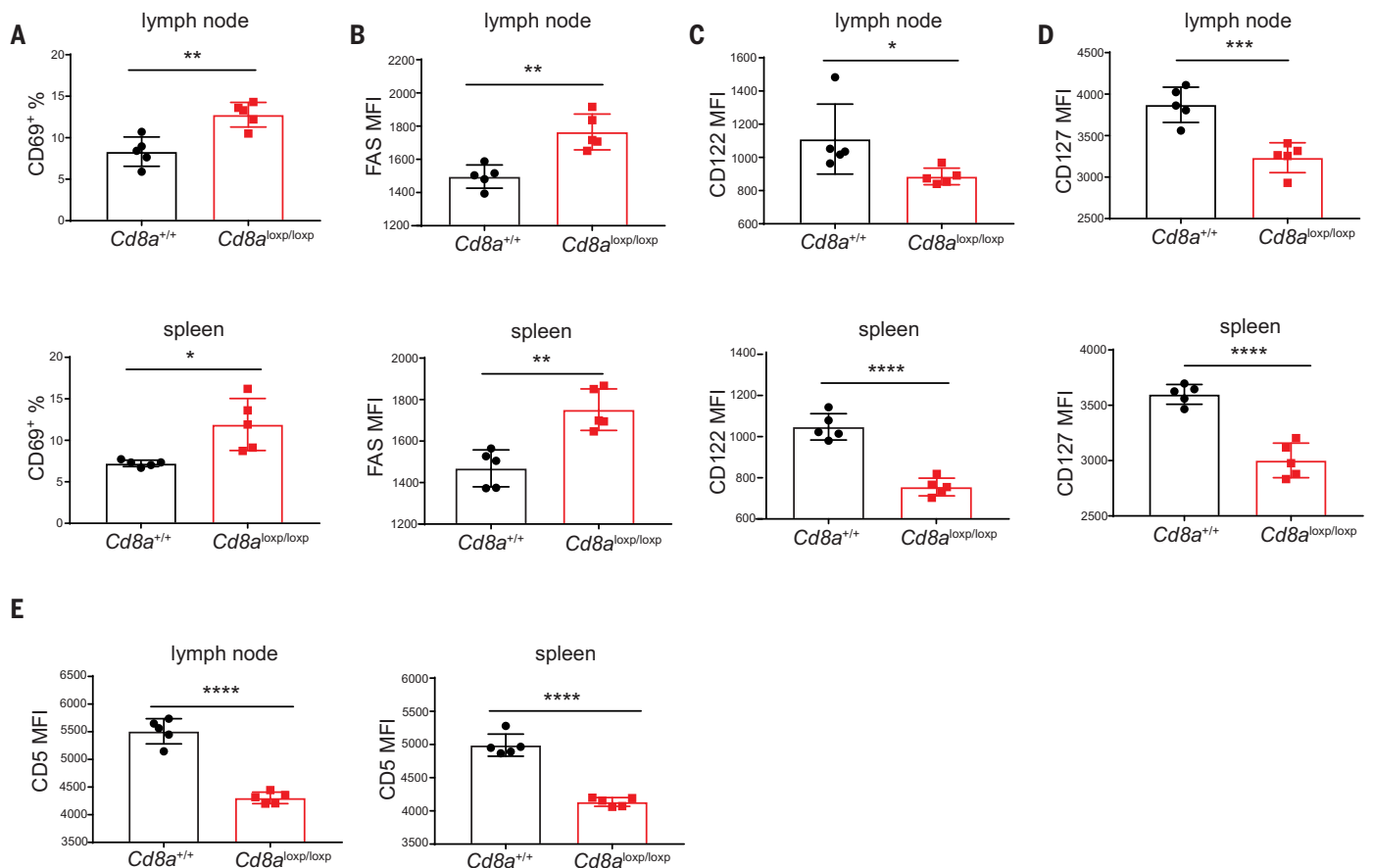
Fig. 2. CD8α maintains memory CD8+ T cell quiescence and CD127 and CD122 expression in the periphery. The lymph node and spleen cells analyzed here were from the same populations as those used in Fig. 1, A and B. As described in Fig. 1, A and B, memory CD45.2+CFSE+ T cells were gated and analyzed for (**A**) CD69+ percentage, (**B**) FAS MFI, (**C**) CD122 MFI, (**D**) CD127 MFI, and (**E**) CD5 MFI. See also fig. S5 for representative flow cytometry plots and histograms. MFI: mean fluorescence intensity. Mean ± SD is shown. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$ by unpaired Student's t test. $n = 5$ mice per group. Data are representative of three independent experiments.

decreased 2 days after 9B12 treatment (fig. S18D). In a complementary approach, treatment with 9B12 reduced the absolute number of carboxyfluorescein diacetate succinimidyl ester (CFSE)–labeled and transferred CD8+ T cells, but not CD4+ T cells, in both lymph nodes and the spleen 32 days after treatment (fig. S19). To determine the effect of 9B12 on individual CD8+ T cell subsets, we transferred naïve or memory CD8+ T cells into mice. 9B12 treatment reduced the survival and broke the quiescence of both naïve (fig. S20, A and B) and memory (fig. S21, A and B) CD8+ T cells. 9B12 had minimal effects on Fas, CD127, and CD122 expression (fig. S20, C to E, and fig. S21, C to E). Moreover, 9B12 enhanced G_0 phase exit in transferred naïve CD8+ T cells (fig. S22). Soluble PILRα could not be detected in mouse sera (fig. S23), suggesting that cell–cell interactions were important in maintaining CD8+ T cell quiescence. We cannot explain the rapid effect of 9B12 on the loss of quiescence of CD8+ T cells in the lymph node compared with its delayed effect in the spleen, but the contribution of tissue-

specific signals in these organs cannot presently be excluded.

We next induced bona fide monoclonal memory CD8+ T cells from OT-I TCR transgenic mice using ovalbumin peptide (OVA) antigen. As described above for polyclonal memory CD8+ T cells, the quiescence of antigen-specific memory OT-I CD8+ T cells was also broken after 9B12 (fig. S24, A and B) or 3D9 (fig. S25, A and B) antibody treatment. To study naïve OT-I T cell quiescence, we backcrossed OT-I transgenic mice to the $Cre^{+/+}Cd8a^{loxp/loxp}$ strain. The inducible deletion of CD8α disrupted naïve OT-I CD8+ T cell quiescence and significantly reduced OT-I CD8+ T cell survival in the absence of OVA antigen stimulation (fig. S26, A to D). Similarly, treatment with anti-PILRα 9B12 resulted in up-regulated CD69 expression on naïve OT-I T-cells, which was associated with decreased survival (fig. S27, A and B).

The ability of 9B12 to alter CD8+ T cell quiescence was dependent on PILRα, as this effect was abolished in $Pilra^{-/-}$ mice (figs. S28, A and B, and S29D). Consistent with a previous report (19), we did not observe any changes

in the frequency of memory or naïve CD8+ or CD4+ T cells in peripheral lymphoid organs in 3-month-old $Pilra^{-/-}$ mice (fig. S29, A to C). CD8+ T cell quiescence in $Pilra^{-/-}$ mice also remained unchanged (fig. S29D). Moreover, purified memory and naïve CD8+ T cells from WT mice did not show decreased survival after transfer into $Pilra^{-/-}$ mice (fig. S30, A and B). Although the differences were not pronounced in 6-month-old mice (fig. S31A), the frequency of memory CD8+ T cells was decreased significantly by 20 to 30% in both lymph nodes and spleens of $Pilra^{-/-}$ mice in comparison with controls at 1 year of age (fig. S31B). By contrast, the frequency of naïve CD8+ T cells or CD4+ T cells remained unchanged in these mice (fig. S31B). Thus, the genetic ablation of PILRα in mice results in a partial phenotype as compared to the mAb against CD8α or PILRα. One possible interpretation for this discrepancy is that while interacting with CD8α to maintain T cell quiescence, PILRα may also engage other receptors promoting T cell activation, which may partially or completely counter the effect of the PILRα–CD8α interaction.
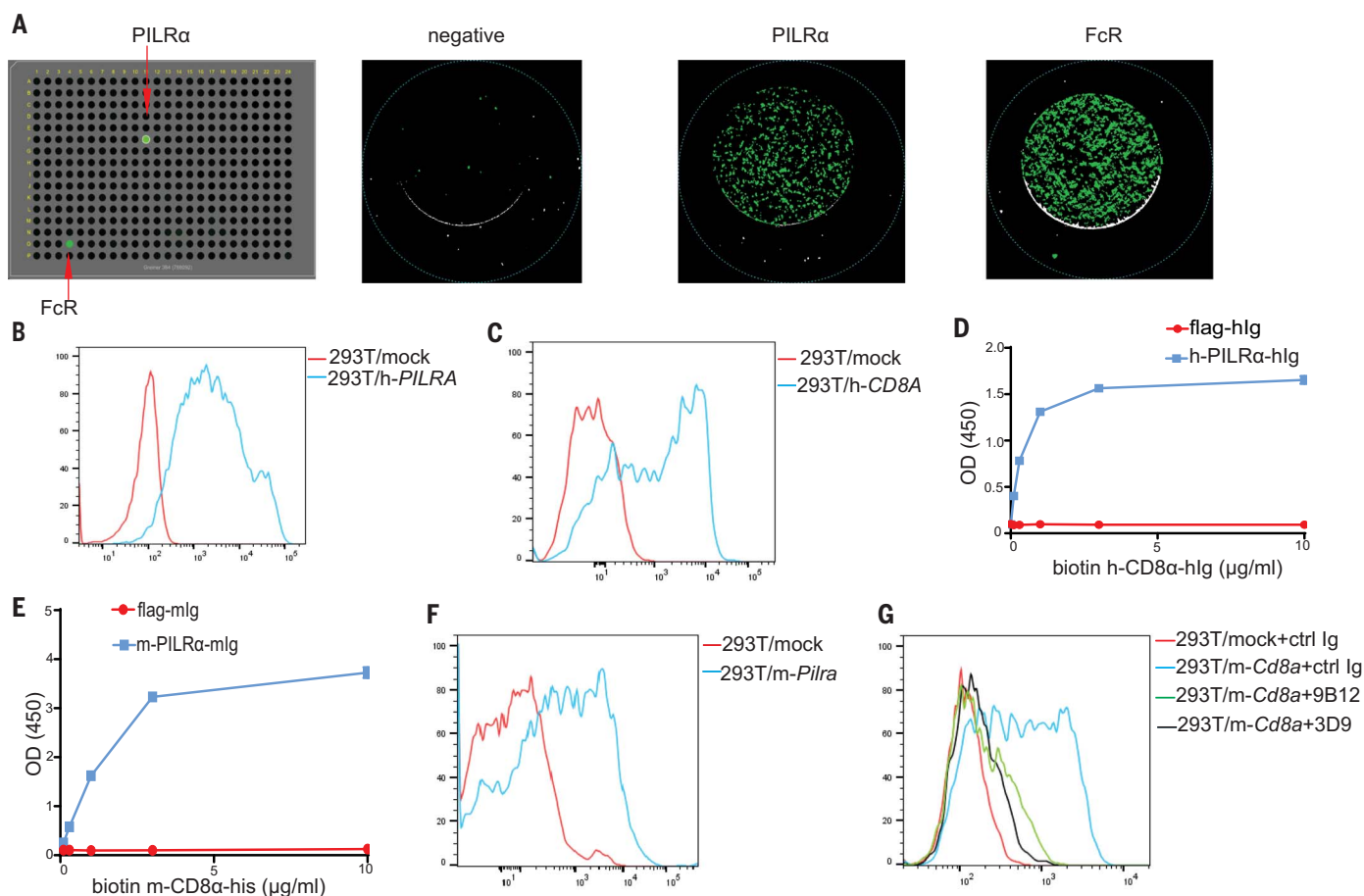
Fig. 3. Identification of PILRα as a ligand for CD8α in both mice and humans.
(**A**) Screening for the binding partner of human CD8α-hIg fusion protein against a human transmembrane protein library on the Mirrorball Fluorescence Cytometer. Representative images of 384-well plates (left) and single wells (right) are shown. (**B**) Flow cytometry analysis of human CD8α-hIg fusion protein binding to mock or human *PILRA*-transfected 293T cells. (**C**) Flow cytometry analysis of human PILRα-mIg binding to mock or human *CD8A*-transfected 293T cells. (**D**) ELISA of biotin-labeled human CD8α-hIg binding to immobilized flag-hIg or human PILRα-hIg. (**E**) ELISA of biotin-labeled mouse CD8-his binding to immobilized flag-mIg or mouse PILRα-mIg. (**F**) Flow cytometry analysis of mouse CD8α-hIg fusion protein binding to mock or mouse *Pilra*-transfected 293T cells. (**G**) Flow cytometry analysis of allophycocyanin (APC)-labeled mouse PILRα-hIg protein binding to mock or mouse *Cd8a*-transfected 293T cells in the presence of anti-mouse CD8α mAb 3D9 or anti-mouse PILRα mAb 9B12. In (D) and (E), values are mean ± SD, and three replicates are included in each concentration. Representative data of two (D), three (E), and at least four [(B), (C), (F), and (G)] independent experiments are shown.

This is reminiscent of B7's interactions with both CD28 (promoting T cell activation) and CTLA-4 (suppressing T cell activation) (*20*). PILRα has been shown to interact with several transmembrane proteins such as NPDC1, COLEC12 (*21*), and PANP (*22*), although these proteins are not thought to be expressed by T cells. PILRα may also bind CD99 (*17*), which is highly expressed on B cells, T cells, and macrophages (*23*) and which has been shown to act as a costimulatory receptor on T cells (*24*). Future studies characterizing PILRα's possible interactions with other non-CD8α receptors will be necessary to test this hypothesis.

The effect of CD8α deletion on CD8+ T cell development may be partial because the quantity of *Cd8a*−/− cytotoxic lymphocytes (CTLs) in the spleen was decreased by only half in the absence of MHC II (*25*). We established a role for the PILRα–CD8α interaction in mature

CD8+ T cell quiescence, but its function in thymic CD8+ T cell development was unclear. To evaluate this, we treated pregnant mice and subsequent neonates with anti-PILRα mAb 9B12. Thymic CD8+ T cell development was normal at postnatal day 14 (fig. S32A), whereas CD8+ T cells in the spleens of 9B12-treated neonates were decreased significantly as expected (fig. S32B). We also tested whether the PILRα–CD8α interaction could affect thymic negative selection. BALB/c mice have been widely used as a model of T cell negative selection owing to this strain's integration of mouse mammary tumor virus types 6, 8, and 9, resulting in the deletion of thymocytes bearing TCR Vβ3, Vβ5, Vβ11, and Vβ12 (*26–28*). CD8+ T cell negative selection in the thymus was unaffected by 9B12 mAb treatment of BALB/c mice (fig. S32C). Thus, the PILRα–CD8α interaction functions in peripheral lymphoid organs but not in the thymus.

Our results support the notion that the quiescent state of T cells, especially CD8+ T cells, is actively maintained by constant receptor–ligand interactions on the cell surface even in the absence of antigen stimulation. PILRα is broadly expressed on various types of myeloid cells (*18*, *19*) and CD8α is highly expressed on naïve and memory CD8+ T cells, providing ample opportunity for this interaction to occur in peripheral lymphoid organs. Similarly, a recent study shows that cell surface PD-1H/VISTA on naïve CD4+ T cells may regulate T cell quiescence (*29*). Memory CD8+ T cells require the PILRα–CD8α interaction to maintain this quiescent status, suggesting that this interaction may contribute to detuning memory CD8+ T cells during antigen-induced activation and help them return to a normal state after the resolution of an immune response. Our findings therefore
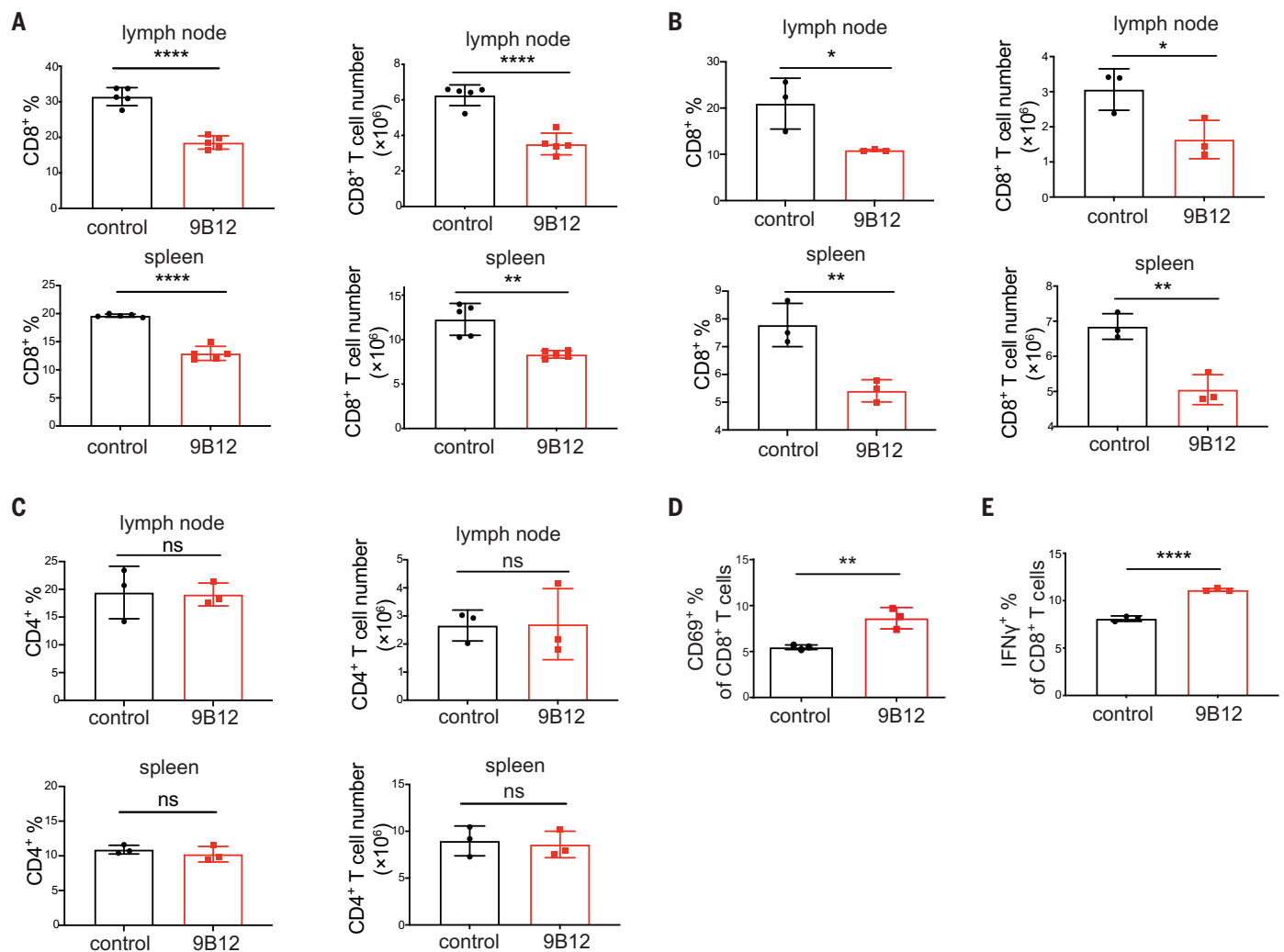
Fig. 4. Blockade of PILRα–CD8α interactions disrupts CD8⁺ T cell homeostasis and quiescence. (**A**) B6 WT mice were administered 200 μg of control or 9B12 antibody intraperitoneally (i.p.) every 3 to 4 days starting on day 0. The frequency and the absolute number of CD8⁺ T cells among lymph node and spleen cells were examined on day 15. (**B** and **C**) Thymectomized B6 WT mice were administered 200 μg of control or 9B12 antibody i.p. every 3 days starting from day 0. CD8⁺ T cell frequency and cell number (B), and CD4⁺ T cell frequency and cell number (C) in the lymph node and spleen were examined on day 14. (**D**) One day after 200 μg of control or 9B12 antibody was administered to B6 WT mice i.p., lymph node CD8⁺ T cells were analyzed for CD69 expression by flow cytometry. (**E**) One day after 200 μg of control or 9B12 antibody was administered to B6 WT mice i.p., CD8⁺ T cells purified from lymph node cells were stimulated by phorbol 12-myristate 13-acetate (PMA), ionomycin, and brefeldin A and analyzed for IFN-γ expression by flow cytometry. Representative data of two [(B) and (C)], three [(A) and (E)], and eight (D) independent experiments are shown. $n = 3$ mice per group in (B), (C), (D), and (E). $n = 5$ mice per group in (A). Mean ± SD is shown. *$P < 0.05$, **$P < 0.01$, ****$P < 0.0001$ by unpaired Student's t test. ns, not significant.

provide greater understanding of how naïve and memory T cell repertoires are maintained and persist in normal and pathological conditions.

REFERENCES AND NOTES

1. S. E. Hamilton, S. C. Jameson, *Trends Immunol.* **33**, 224–230 (2012).
2. S. S. Hwang *et al.*, *Science* **367**, 1255–1260 (2020).
3. V. Kalia, L. A. Penny, Y. Yuzefpolskiy, F. M. Baumann, S. Sarkar, *Immunity* **42**, 1116–1129 (2015).
4. E. A. Boyse, *J. Exp. Med.* **141**, 1376–1389 (1975).
5. J. A. Ledbetter, W. E. Seaman, T. T. Tsu, L. A. Herzenberg, *J. Exp. Med.* **153**, 1503–1516 (1981).
6. A. M. Norment, R. D. Salter, P. Parham, V. H. Engelhard, D. R. Littman, *Nature* **336**, 79–81 (1988).
7. A. Veillette, M. A. Bookman, E. M. Horak, J. B. Bolen, *Cell* **55**, 301–308 (1988).
8. M. N. Artyomov, M. Lis, S. Devadas, M. M. Davis, A. K. Chakraborty, *Proc. Natl. Acad. Sci. U.S.A.* **107**, 16916–16921 (2010).
9. W. P. Fung-Leung *et al.*, *Cell* **65**, 443–449 (1991).
10. S. D. Gorman, Y. H. Sun, R. Zamoyska, J. R. Parnes, *J. Immunol.* **140**, 3646–3653 (1988).
11. H. S. Azzam *et al.*, *J. Exp. Med.* **188**, 2301–2311 (1998).
12. A. E. Moran *et al.*, *J. Exp. Med.* **208**, 1279–1289 (2011).
13. J. E. Knapp, D. Liu, *Methods Mol. Biol.* **245**, 245–250 (2004).
14. S. Yao *et al.*, *Immunity* **34**, 729–740 (2011).
15. Y. Zhu *et al.*, *Nat. Commun.* **4**, 2043 (2013).
16. J. Wang *et al.*, *Cell* **176**, 334–347.e12 (2019).
17. I. Shiratori, K. Ogasawara, T. Saito, L. L. Lanier, H. Arase, *J. Exp. Med.* **199**, 525–533 (2004).
18. J. Wang, I. Shiratori, J. Uehori, M. Ikawa, H. Arase, *Nat. Immunol.* **14**, 34–40 (2013).
19. Y. Sun *et al.*, *J. Immunol.* **193**, 860–870 (2014).
20. J. L. Riley, C. H. June, *Blood* **105**, 13–21 (2005).
21. Y. Sun *et al.*, *J. Biol. Chem.* **287**, 15837–15850 (2012).
22. A. Kogure, I. Shiratori, J. Wang, L. L. Lanier, H. Arase, *Biochem. Biophys. Res. Commun.* **405**, 428–433 (2011).
23. H. J. Park *et al.*, *Mol. Cells* **33**, 259–267 (2012).
24. K. I. Oh *et al.*, *Exp. Mol. Med.* **39**, 176–184 (2007).
25. D. S. Riddle *et al.*, *Eur. J. Immunol.* **38**, 1511–1521 (2008).
26. M. T. Scherer, L. Ignatowicz, G. M. Winslow, J. W. Kappler, P. Marrack, *Annu. Rev. Cell Biol.* **9**, 101–128 (1993).
27. J. X. Gao *et al.*, *J. Exp. Med.* **195**, 959–971 (2002).
28. A. C. Davalos-Misslitz, T. Worbs, S. Willenzon, G. Bernhardt, R. Förster, *Blood* **110**, 4351–4359 (2007).
29. M. A. ElTanbouly *et al.*, *Science* **367**, eaay0524 (2020).

SUPPLEMENTARY MATERIALS

A transient, closed-loop network of wireless, body-integrated devices for autonomous electrotherapy

Yeon Sik Choi[1,2,3]†, Hyoyoung Jeong[1,2]†, Rose T. Yin[4]†, Raudel Avila[5], Anna Pfenniger[6], Jaeyoung Yoo[1,2], Jong Yoon Lee[1,2,7], Andreas Tzavelis[1,2,8,9], Young Joong Lee[1,2], Sheena W. Chen[10,11], Helen S. Knight[4], Seungyeob Kim[1,2,12], Hak-Young Ahn[1,2,3], Grace Wickerson[1,2,13], Abraham Vázquez-Guardado[1,2], Elizabeth Higbee-Dempsey[14], Bender A. Russo[4], Michael A. Napolitano[10,11], Timothy J. Holleran[10,11], Leen Abdul Razzak[1,2,8], Alana N. Miniovich[4], Geumbee Lee[1,2], Beth Geist[6], Brandon Kim[7], Shuling Han[15,16], Jaclyn A. Brennan[4], Kedar Aras[4], Sung Soo Kwak[1,2]‡, Joohee Kim[1,2], Emily Alexandria Waters[8,17], Xiangxing Yang[18], Amy Burrell[6], Keum San Chun[18], Claire Liu[1,2,8], Changsheng Wu[1,2], Alina Y. Rwei[19], Alisha N. Spann[17], Anthony Banks[1,2], David Johnson[6], Zheng Jenny Zhang[15,16], Chad R. Haney[8,17], Sung Hun Jin[1,2,12], Alan Varteres Sahakian[8,20], Yonggang Huang[1,3,5,21], Gregory D. Trachiotis[11], Bradley P. Knight[6], Rishi K. Arora[6]*, Igor R. Efimov[2,4]*§¶, John A. Rogers[1,2,5,8,13,22]*

Temporary postoperative cardiac pacing requires devices with percutaneous leads and external wired power and control systems. This hardware introduces risks for infection, limitations on patient mobility, and requirements for surgical extraction procedures. Bioresorbable pacemakers mitigate some of these disadvantages, but they demand pairing with external, wired systems and secondary mechanisms for control. We present a transient closed-loop system that combines a time-synchronized, wireless network of skin-integrated devices with an advanced bioresorbable pacemaker to control cardiac rhythms, track cardiopulmonary status, provide multihaptic feedback, and enable transient operation with minimal patient burden. The result provides a range of autonomous, rate-adaptive cardiac pacing capabilities, as demonstrated in rat, canine, and human heart studies. This work establishes an engineering framework for closed-loop temporary electrotherapy using wirelessly linked, body-integrated bioelectronic devices.

All living systems function through the interaction of complex networks of physiological feedback loops to maintain homeostasis. Engineering approaches to treat disorders, such as those based on cardiac pacemakers, exploit conceptually similar methods for closed-loop control to enable autonomous, adaptive regulation of one or more essential physiological parameters to target set points without human intervention (1–3). These and other existing platforms have key limitations that follow from their reliance on conventional electronic hardware, monitoring schemes, and interfaces to the body. First, such systems often require physical tethers and percutaneous access points that may lead to systemic infections (4–7). Second, connections to external modules for power supply, sensing, control, and other essential functions constrain patient mobility and impede

clinical care. Third, removal or replacement of electronic components (e.g., leads or batteries) demands surgical procedures that impose additional risks and burdens on patients (8, 9). These features can extend the duration of hospitalization, often in intensive care units. For example, short-term bradyarrhythmias that commonly occur in the 5 to 7 days after cardiac surgery must be treated with temporary percutaneous pacing systems, typically prolonging hospital stays with limited ability to initiate physical therapy (supplementary text 1). Recently reported wireless, bioresorbable electronic implants for temporary therapies address some of these challenges, but they still require external, wall-plugged equipment for monitoring, power, and control (10–16).

We introduce a transient, closed-loop system that incorporates a time-synchronized, wireless network with seven key components:

(i) a temporary, bioresorbable, stretchable epicardial pacemaker; (ii) a bioresorbable steroid-eluting interface that minimizes local inflammation and fibrosis (17); (iii) a subcutaneous, bioresorbable power harvesting unit; (iv) a set of soft, skin-interfaced sensors that capture electrocardiograms (ECGs), heart rate (HR), respiratory information, physical activity, and cerebral hemodynamics for physiological monitoring of the patient; (v) a wireless radiofrequency (RF) module that transfers power to the harvesting unit; (vi) a soft, skin-interfaced haptic actuator that communicates via mechanical vibrations; and (vii) a handheld device with a software application for real-time visualization, storage, and analysis of data for automated adaptive control. These components integrate into a fully implantable, bioresorbable module [(i) to (iii)]; a set of skin-interfaced modules [(iv) to (vi)]; and an external control module (vii).

Figure 1A illustrates the use of this system for temporary cardiac pacing. The bioresorbable module wirelessly receives power for epicardial pacing. A network of skin-interfaced modules transmits diverse physiological data to the control module via Bluetooth low energy (BLE) protocols for real-time data visualization and algorithmic control. A haptic module provides tactile feedback to the patient. After a period of therapy, the bioresorbable module dissolves in the body, and the skin-interfaced modules are removed by peeling them off the skin. These "transient" characteristics of the system eliminate the need for surgical removal and allow ambulatory end of treatment. Figure 1B illustrates the closed-loop scheme that interconnects these modules into a wireless network (table S1). Soft, flexible designs (Fig. 1C) enable placement of the modules onto various target locations of the body.

Figure 1D shows that the constituent materials of the bioresorbable module completely disappear in simulated biofluid consisting of phosphate-buffered saline (PBS). Results of in vivo studies are provided in fig. S1. As shown in Fig. 2A, the bioresorbable module consists of an RF power harvester, which includes an inductive receiver (Rx) coil [molybdenum (Mo)] and a PIN diode [silicon nanomembrane

[1]Center for Bio-Integrated Electronics, Northwestern University, Evanston, IL 60208, USA. [2]Querrey Simpson Institute for Bioelectronics, Northwestern University, Evanston, IL 60208, USA. [3]Precision Biology Research Center, Sungkyunkwan University, Suwon 16419, Republic of Korea. [4]Department of Biomedical Engineering, The George Washington University, Washington, DC 20052, USA. [5]Department of Mechanical Engineering, Northwestern University, Evanston, IL 60208, USA. [6]Feinberg School of Medicine, Cardiology, Northwestern University, Chicago, IL 60611, USA. [7]Sibel Health, Niles, IL 60714, USA. [8]Department of Biomedical Engineering, Northwestern University, Evanston, IL 60208, USA. [9]Medical Scientist Training Program, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA. [10]Department of General Surgery, The George Washington University, Washington, DC 20052, USA. [11]Department of Cardiothoracic Surgery, Veteran Affairs Medical Center, Washington, DC 20422, USA. [12]Department of Electronic Engineering, Incheon National University, 119 Academy-ro, Yeonsu-gu, Incheon 406-772, Republic of Korea. [13]Department of Materials Science and Engineering, Northwestern University, Evanston, IL 60208, USA. [14]Developmental Therapeutics Core, Northwestern University, Evanston, IL 60208, USA. [15]Comprehensive Transplant Center, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA. [16]Department of Surgery, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA. [17]Center for Advanced Molecular Imaging, Northwestern University, Evanston, IL 60208, USA. [18]Department of Electrical and Computer Engineering, University of Texas at Austin, Austin, TX 78712, USA. [19]Department of Chemical Engineering, Delft University of Technology, Van der Maasweg 9, 2629 HZ Delft, Netherlands. [20]Department of Electrical and Computer Engineering, Northwestern University, Evanston, IL 60208, USA. [21]Department of Civil and Environmental Engineering, Northwestern University, Evanston, IL 60208, USA. [22]Department of Neurological Surgery, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA.
*Corresponding author. Email: r-arora@northwestern.edu (R.K.A.); efimov@gwu.edu (I.R.E.); jrogers@northwestern.edu (J.A.R.)
†These authors contributed equally to this work.
‡Present address: Center for Bionics of Biomedical Research Institute, Korea Institute of Science and Technology, Seoul 02792, Korea.
§Present address: Department of Biomedical Engineering, Northwestern University, Evanston, IL 60208, USA.
¶Present address: Department of Medicine, Division of Cardiology, Feinberg School of Medicine, Northwestern University, Chicago, IL 60611, USA.
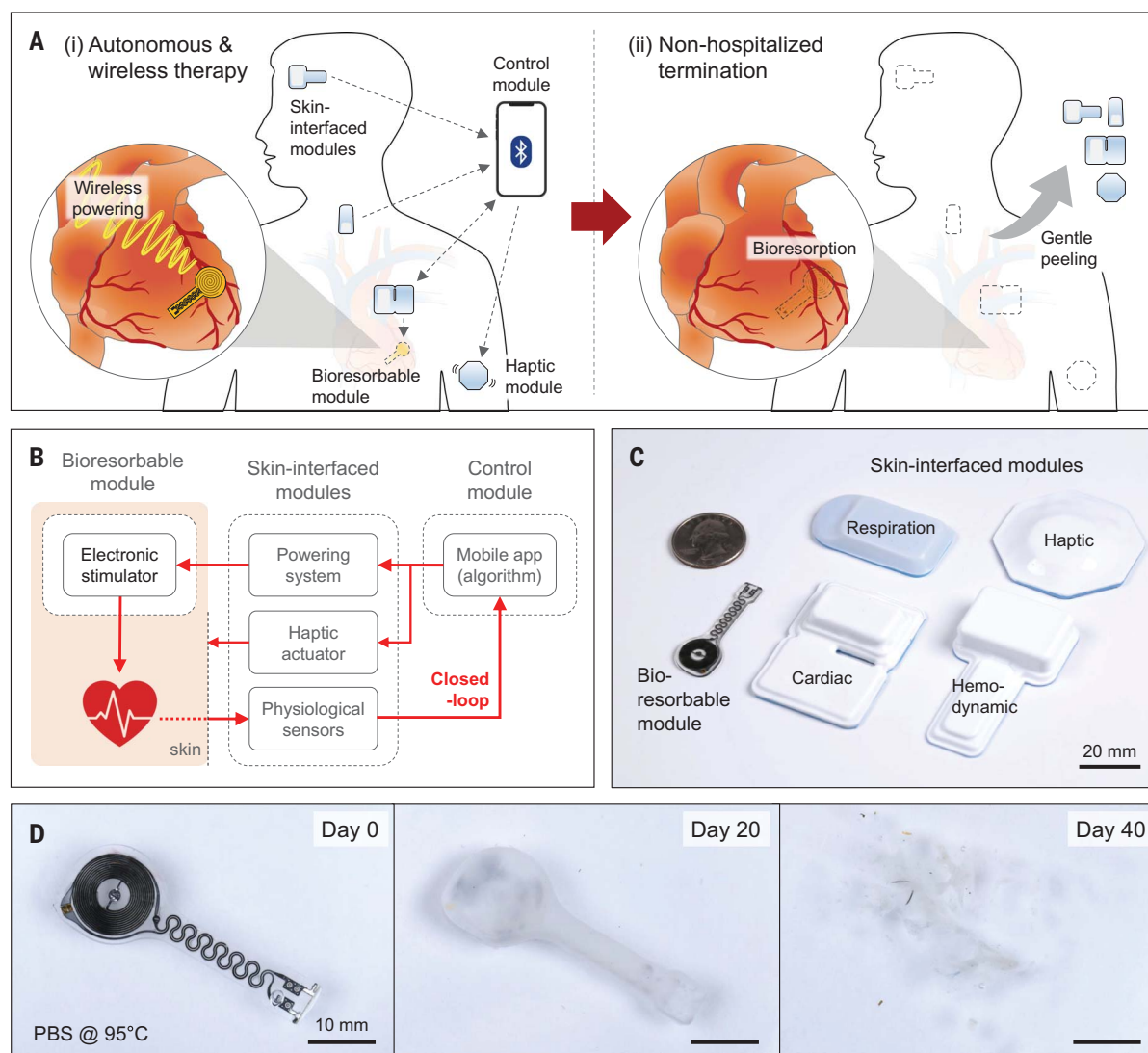
Fig. 1. Transient closed-loop system for temporary cardiac pacing. (**A**) Schematic illustration of a system for (i) autonomous and wireless pacing therapy and (ii) nonhospitalized termination. (**B**) Operational diagram of the closed-loop system for continuous monitoring, autonomous treatment, and haptic feedback. (**C**) Photographs showing the sizes of the various modules, relative to a US quarter. (**D**) Photographs of a bioresorbable module at different time points during immersion in a simulated biofluid (in PBS at 95°C).

(Si NM)], a pair of stretchable interconnects (Mo), and stimulation electrodes that integrate a steroid-eluting patch at the myocardial interface. The thin, lightweight, and stretchable design minimizes the possibility for irritation or damage at the tissue interface, with geometries that can be tailored to the anatomy of the patient (fig. S2). Figure 2B shows scattering parameters (S_{11}) of power harvesters with three different sizes of Rx coils (supplementary text 2). Continuous alternating current applied to a transmission (Tx) coil wirelessly delivers power to the Rx coil via magnetic induction and induces an approximately direct current monophasic output defined by the diode rectifier (Fig. 2C). The magnetic resonance imaging (MRI) compatibility of this wireless system is discussed in supplementary

text 3. Top and bottom encapsulating layers of a bioresorbable dynamic covalent polyurethane (b-DCPU) and stretchable electrodes (*11*) ensure reliable pacing against the mechanically dynamic surface of the heart (*18*). Figure 2D shows negligible differences in output voltage during mechanical deformation, consistent with modeling results (fig. S7). Because the wireless energy transfer is inversely proportional to the coil-to-coil distance (fig. S8), the power harvester resides subcutaneously to maximize the efficiency. Poly(lactic-co-glycolic acid) (PLGA)–based steroid-eluting patches release dexamethasone acetate (DMA) over the course of several months to minimize local inflammation and fibrosis during cardiac pacing (Fig. 2E and fig. S9). The slow rate of dissolution of the bioresorbable conductor

(Mo) enables >1 month of functional lifetime under simulated physiological conditions (Fig. 2F and supplementary text 4).

A network of skin-interfaced modules placed on various locations of the body acquires diverse data relevant to patient status. These collective data streams form the basis for closed-loop control. As the essential component, the cardiac module mounts on the chest to collect physiological information and to provide RF power to the bioresorbable module. Its materials and architectures (Fig. 2G and fig. S12) follow design principles of soft electronics to ensure robust, irritation-free coupling to the skin (fig. S13) at relevant locations (fig. S14) (*19*). The multihaptic module on the mid-medial forearm provides information on patient status and device operation through up to 625 patterns
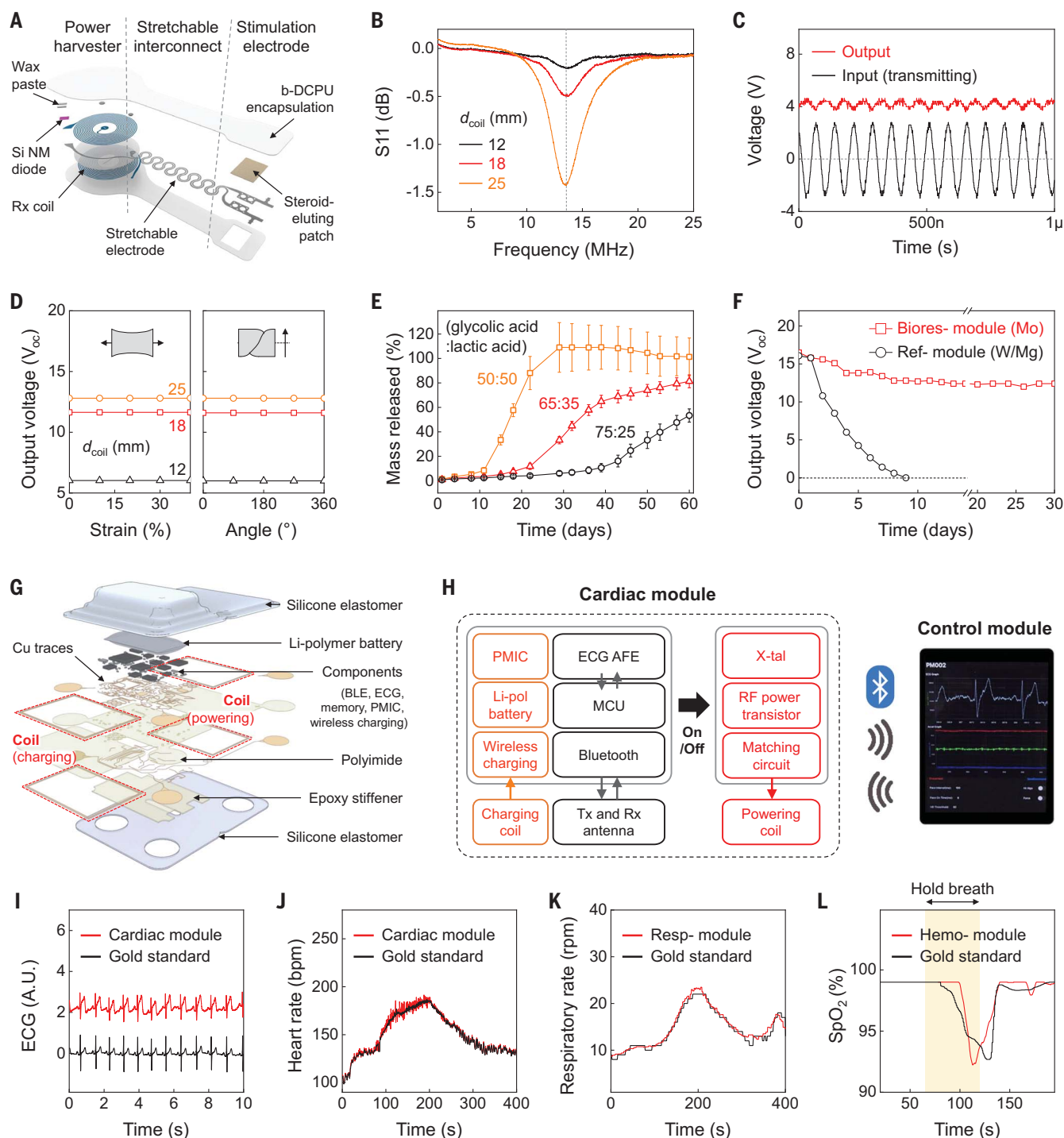
Fig. 2. Materials, design features. (**A**) Schematic illustration of a bioresorbable module. (**B**) S_{11} values of the Rx coils with different diameters (d_{coil}). (**C**) Example output waveform (red; d_{coil} = 12 mm) wirelessly generated by an alternating current (black; ~6 V_{pp}; 13.56 MHz) applied to the Tx coil. (**D**) Output open circuit voltage (V_{OC}) of devices as a function of tensile strain (left) and twist angle (right) at a fixed transmitting voltage (8 V_{pp}) and frequency (13.56 MHz). (**E**) Drug-release behaviors of steroid-eluting patches with three different ratios of base polymer. Error bars represent standard deviation. (**F**) Measurements of V_{OC} of the bioresorbable module (red squares; 10-μm-

thick Mo) and a reference module (black circles; 700-nm-thick W coated 50-μm-thick Mg) immersed in PBS (37°C). (**G**) Schematic illustration of a skin-interfaced cardiac module. PMIC, power management integrated circuit. (**H**) System block diagram of the cardiac module. (**I** to **L**) Comparisons of ECG, HR, respiratory rate, and SpO$_2$ levels determined by the skin-interfaced modules [red; cardiac module in (I) and (J); respiratory module in (K); hemodynamic module in (L)] and a reference device (black). In (L), data were collected from a healthy subject who held their breath for 60 s (yellow background). A.U., arbitrary units; rpm, respirations per minute.

of vibrotactile input (*20*). The respiratory module mounts at the suprasternal notch to capture physical activity, body temperature, and respiratory behavior in a dual-sensing design for accurate operation (*21*). The hemodynamic module on the forehead measures peripheral blood oxygen saturation (SpO_2) (*22*).

Figure 2H shows a block diagram of the skin-interfaced cardiac module. An ECG analog front end (AFE) and a microcontroller unit (MCU) process measured data in real time to calculate the HR (fig. S15). A BLE-enabled user interface serves as a control unit that stores and displays ECG tracings and three-axis acceleration data associated with cardiac and
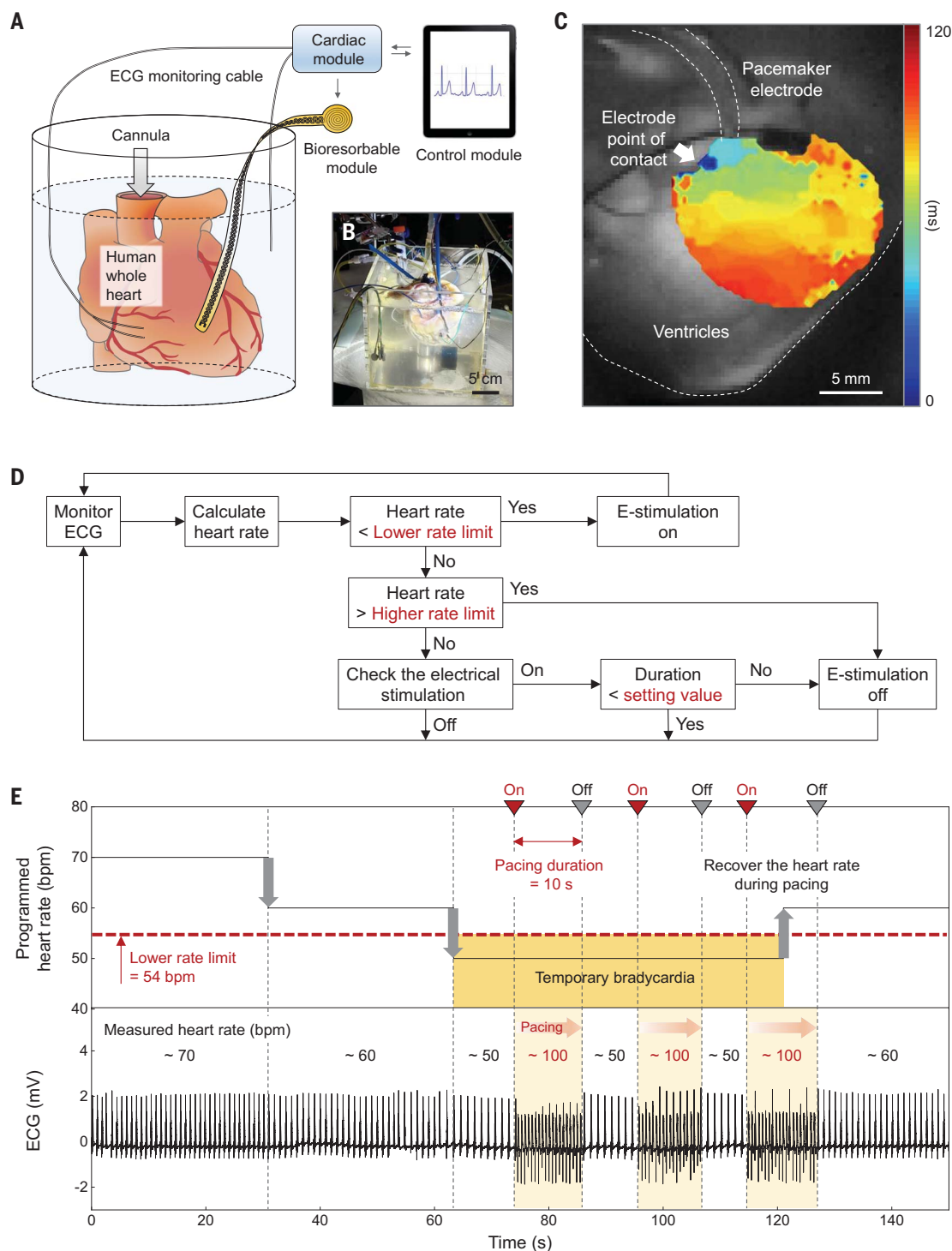
Fig. 3. Treatment of temporary bradycardia. (**A** and **B**) Schematic illustration (A) and photograph (B) of a Langendorff-perfused human whole-heart model with a transient closed-loop system (d_{coil} = 25 mm). (**C**) Action potential maps obtained by optical mapping of the human epicardium. (**D**) Flow chart of closed-loop hysteresis pacing to activate the pacemaker upon automatic detection of temporary bradycardia (supplementary text 8). (**E**) Programmed HR (top) and measured ECG (bottom) of a human whole heart. Set parameters are as follows: The lower rate limit is 54 bpm, pacing duration is 10 s, and pacing rate is 100 bpm.

respiratory activity (fig. S16). Figure 2, I to K, shows that the skin-interfaced modules and data analytics approaches accurately determine HR and respiratory rate (fig. S17). The hemodynamic module yields SpO$_2$ data comparable to that recorded by a medical-grade finger probe (Fig. 2L). These systems use current best practices to protect health data, from the sensor, BLE link, phone, cloud, and beyond. To ensure secure medical data storage and processing, the interface application is compatible with hypertext transfer
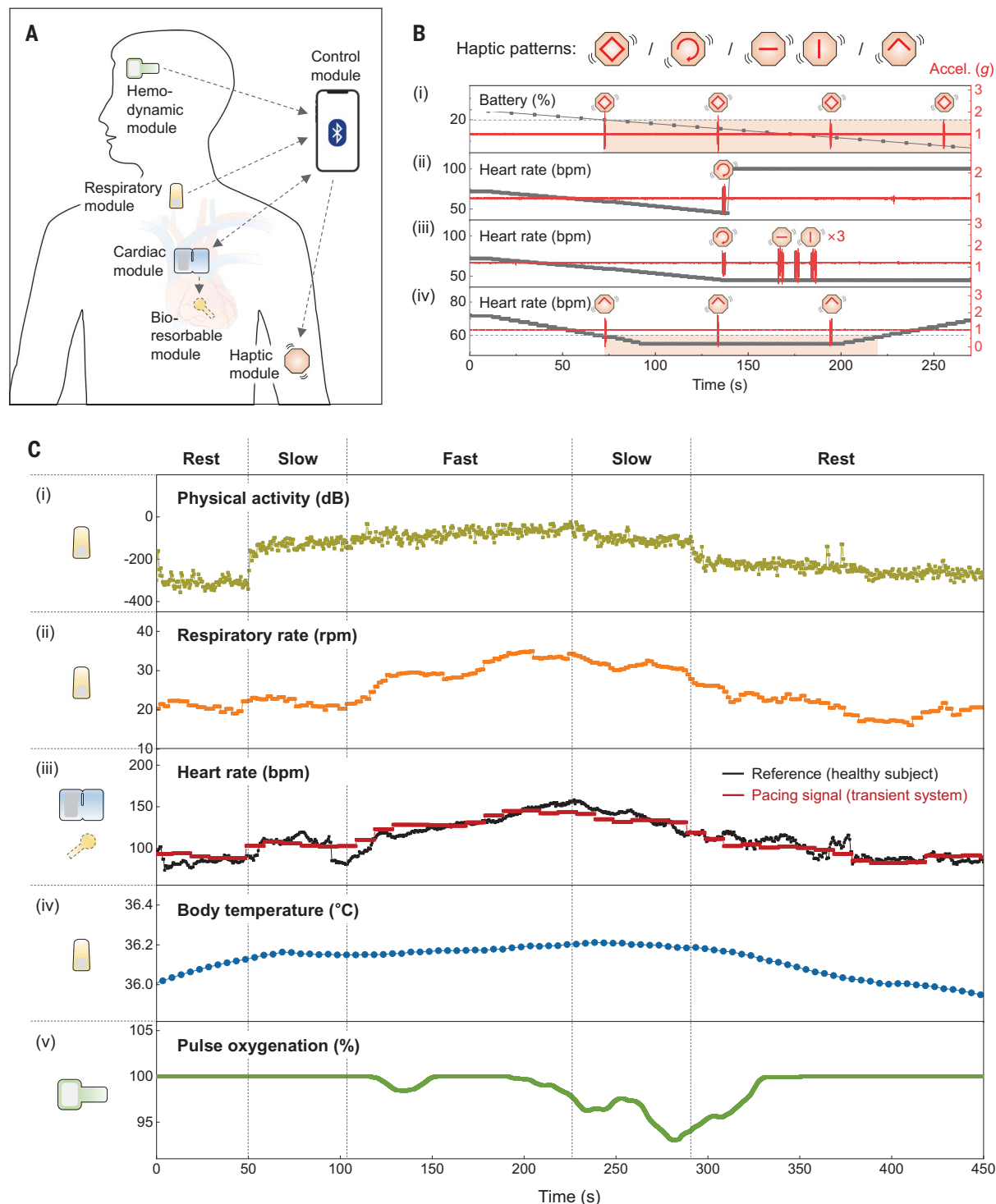
Fig. 4. Patient feedback and adaptive pacing functions. (**A**) Schematic illustration of a transient closed-loop system with the full collection of skin-interfaced modules. (**B**) Demonstration of the patient-awareness function using a multihaptic module. Accelerometer data (*g*) corresponds to vibrations (*z* axis) of the haptic actuators. (**C**) Results of clinical tests with a healthy human subject: (i) calculated physical activity and (ii) respiratory rate using data from the respiratory module, (iii) comparison of the HR (black) of a healthy human subject monitored by the cardiac module and rate-adaptive pacing signals (red) processed from the algorithm, (iv) calibrated and measured changes in core body temperatures using data from the respiratory module, and (v) representative SpO$_2$ measurements from the hemodynamic module.

protocol secure (HTTPS) transport layer security (TLS 1.2) and with algorithms for encryption and decryption (fig. S18). In-sensor encryption [advanced encryption standard–128 (AES-128)] and Health Insurance Portability and Accountability Act (HIPAA)–compliant cloud data storage further protect patient data.

One of the key features of this transient closed-loop system (10–15) is that the skin-interfaced cardiac module eliminates requirements for wall-plugged external hardware for power transfer and control of the implanted pacemaker (fig. S19). In vivo studies with a canine whole-heart model demonstrate its capabilities (fig. S20). When the wireless cardiac module generates pulsed alternating currents [6 peak-to-peak voltage (V_{pp})], the bioresorbable module rectifies the received waveform and delivers it to the myocardium-interface as a cathodic monophasic pulse (~4 mW) (supplementary text 5). Investigations using rodent models demonstrate continuous, long-term pacing and biocompatibility (supplementary text 6 and 7).

An additional capability of this system is in autonomous treatment based on algorithmic identification of ECG signatures of abnormal cardiac activity. For example, hysteresis pacing delivers programmed electrical stimuli if the intrinsic rate falls below a certain threshold (23). Ex vivo human whole-heart studies demonstrate this type of treatment for temporary bradycardia (Fig. 3). Anisotropic activation of the membrane potential confirms that the bioresorbable module is the driving source of cardiac activation (Fig. 3C).

A flow chart of the feedback control system (Fig. 3D) implemented in the mobile application describes the hysteresis pacing scheme by which the system recognizes bradycardia and activates pacing during the programmed period of treatment. A separate pacing electrode enables manual control of the HR to mimic bradycardia (fig. S40). Figure 3E shows that the transient closed-loop system detects bradycardia [in this case, the bradycardic threshold is set to 54 beats per minute (bpm)] and automatically initiates pacing (~100 bpm). After a predetermined pacing duration (10 s), the system automatically stops pacing and evaluates the underlying intrinsic ECG signals to determine the need for additional pacing treatment. When the heart recovers from temporary bradycardia, the system detects the normal HR (~60 bpm) and ceases to deliver on-demand pacing.

For advanced forms of operation, the control module wirelessly communicates with the full collection of skin-interfaced modules via BLE protocols in a manner that is expandable and customizable to accommodate wide-ranging types of devices with various actuation, feedback, and/or monitoring capabilities. The schematic illustrations in Fig. 4A and fig. S41 summarize the most sophisticated system configuration reported here. This network of modules also includes the option to deliver tactile inputs through different patterns of vibration (fig. S42 and movie S1) to inform the patient of (i) the remaining battery life, (ii) the proper operation of the cardiac module, (iii) instances of malfunction of the other modules, and (iv) symptoms of bradycardia (Fig. 4B). The haptic module can also be activated to facilitate positioning of the cardiac module during mounting, of particular importance in the course of device replacement for recharging (fig. S43).

Real-time monitoring of cardiopulmonary status and physical activity, along with other essential parameters enables elaborate schemes for rate-adaptive pacing (supplementary text 9). Exercise tests of healthy human subjects on stationary bicycles demonstrate this rate-adaptive function (fig. S44). Figure 4C shows a strong qualitative correspondence (i) between measured physical activity and exercise intensity (e.g., rest, slow, fast). The respiratory rate (ii) shows a time-delayed correlation to physical activity and has gradual changes at the transition of exercise intensities. The pacing signal (iii), calculated by (i) and (ii), shows good agreement with the HR of the healthy subject because the metabolic demand is consistent with the level of exercise intensity and respiration. Results from different human subjects ($n = 8$) confirm the reliability of this algorithm (fig. S46), and supplementary text 10 describes strategies for stable and reliable pacing. Other physiological parameters, such as body temperature (iv) and blood oxygen saturation level (v), provide additional information that is postoperatively useful for patients with limited cardiopulmonary reserve, slowly resolving pneumonia, or persistent supplemental oxygen requirements.

This transient, closed-loop system represents a distributed, wireless bioelectronics technology that provides autonomous electrotherapy over a time frame that matches postoperative needs. The operation involves coordinated operation of a network of skin-interfaced modules and a bioresorbable device in time-synchronized communication with a control platform. Data captured from various locations of the body yield detailed information on cardiopulmonary health and physical activity. The results define autonomous, rate-adaptive pacing parameters to match metabolic demand through wireless powering of the bioresorbable module; they also support feedback on device and physiological status through a multihaptic interface. The bioresorbable module for cardiac pacing undergoes complete dissolution by natural biological processes after a defined operating time frame. The skin-interfaced devices can be easily removed after patient recovery. This system provides a framework for closed-loop technologies to treat various diseases and temporary patient conditions in a way that can complement traditional biomedical devices and pharmacological approaches.

REFERENCES AND NOTES

1. K. Kaszala, K. A. Ellenbogen, *Circulation* **122**, 1328–1340 (2010).
2. R. Hovorka, *Diabet. Med.* **23**, 1–12 (2006).
3. A. D. Mickle et al., *Nature* **565**, 361–365 (2019).
4. J. L. Austin, L. K. Preis, R. S. Crampton, G. A. Beller, R. P. Martin, *Am. J. Cardiol.* **49**, 301–306 (1982).
5. F. J. Lumia, J. C. Rios, *Chest* **64**, 604–608 (1973).
6. K. D. Donovan, K. Y. Lee, *Anaesth. Intensive Care* **13**, 63–70 (1985).
7. M. U. Braun et al., *Pacing Clin. Electrophysiol.* **29**, 875–879 (2006).
8. K. A. McLeod, *Heart* **96**, 1502–1508 (2010).
9. E. Buch, N. G. Boyle, P. H. Belott, *Circulation* **123**, e378–e380 (2011).
10. Y. S. Choi et al., *Nat. Biotechnol.* **39**, 1228–1238 (2021).
11. Y. S. Choi et al., *Nat. Commun.* **11**, 5990 (2020).
12. J. Koo et al., *Sci. Adv.* **6**, eabb1093 (2020).
13. Y. S. Choi, J. Koo, J. A. Rogers, *MRS Bull.* **45**, 103–112 (2020).
14. Y. S. Choi et al., *Adv. Funct. Mater.* **30**, 2000941 (2020).
15. S. K. Kang et al., *Nature* **530**, 71–76 (2016).
16. S.-W. Hwang et al., *Science* **337**, 1640–1644 (2012).
17. H. G. Mond, J. R. Helland, K. Stokes, G. A. Bornzin, R. McVenes, *Pacing Clin. Electrophysiol.* **37**, 1232–1249 (2014).
18. E. K. S. Espe et al., *J. Cardiovasc. Magn. Reson.* **15**, 82 (2013).
19. H. U. Chung et al., *Nat. Med.* **26**, 418–429 (2020).
20. X. Yu et al., *Nature* **575**, 473–479 (2019).
21. H. Jeong et al., *Sci. Adv.* **7**, eabg3092 (2021).
22. A. Y. Rwei et al., *Proc. Natl. Acad. Sci. U.S.A.* **117**, 31674–31684 (2020).
23. E. García-Izquierdo, S. Vilches, V. Castro, *Circulation* **135**, 711–713 (2017).

and AliveCor. **Data and materials availability:** All data are available in the main text or the supplementary materials. **License information:** Copyright © 2022 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.science.org/about/science-licenses-journal-article-reuse

ADAPTATION

Genetic variance in fitness indicates rapid contemporary adaptive evolution in wild animals

Timothée Bonnet[1]*, Michael B. Morrissey[2], Pierre de Villemereuil[3,4], Susan C. Alberts[5], Peter Arcese[6], Liam D. Bailey[7], Stan Boutin[8], Patricia Brekke[9], Lauren J. N. Brent[10], Glauco Camenisch[11], Anne Charmantier[12], Tim H. Clutton-Brock[13,14], Andrew Cockburn[1], David W. Coltman[8], Alexandre Courtiol[7], Eve Davidian[7], Simon R. Evans[15,16,17], John G. Ewen[9], Marco Festa-Bianchet[18], Christophe de Franceschi[12], Lars Gustafsson[16], Oliver P. Höner[7], Thomas M. Houslay[13,17], Lukas F. Keller[11,19], Marta Manser[11,14], Andrew G. McAdam[20], Emily McLean[21], Pirmin Nietlisbach[22], Helen L. Osmond[1], Josephine M. Pemberton[23], Erik Postma[17], Jane M. Reid[24,25], Alexis Rutschmann[4], Anna W. Santure[4], Ben C. Sheldon[15], Jon Slate[26], Céline Teplitsky[12], Marcel E. Visser[27], Bettina Wachter[7], Loeske E. B. Kruuk[1,23]

The rate of adaptive evolution, the contribution of selection to genetic changes that increase mean fitness, is determined by the additive genetic variance in individual relative fitness. To date, there are few robust estimates of this parameter for natural populations, and it is therefore unclear whether adaptive evolution can play a meaningful role in short-term population dynamics. We developed and applied quantitative genetic methods to long-term datasets from 19 wild bird and mammal populations and found that, while estimates vary between populations, additive genetic variance in relative fitness is often substantial and, on average, twice that of previous estimates. We show that these rates of contemporary adaptive evolution can affect population dynamics and hence that natural selection has the potential to partly mitigate effects of current environmental change.

How fast are wild populations currently evolving in response to natural selection? The rate of adaptive evolution in nature is both of fundamental theoretical importance and of increasing practical relevance given the clear impact of human activities on the environments that wild organisms inhabit (1). There are numerous examples of phenotypic and genetic changes for traits under selection (2–5), which suggests that adaptive evolution can occur in wild populations over contemporary time scales. At the same time, however, many studies have found that trait changes do not correspond to adaptive expectations or suggest evolutionary stasis (6, 7). However, estimates of the rate of evolution of specific traits are unlikely to represent the overall rate of adaptation of a population, as natural selection acts on many traits concurrently. Instead, a comprehensive assessment of the rate of adaptive evolution in a population needs to integrate adaptive genetic changes across all traits that determine individual fitness, that is, the contribution of

an individual to the gene pool of the next generation.

According to Fisher's fundamental theorem of natural selection, the per-generation proportional change in mean absolute fitness caused by natural selection is given by the additive genetic variance in relative fitness, $V_A(w)$ (8–10). In nontechnical terms, $V_A(w)$ is the extent of heritable (transmitted from parents to offspring) genetic differences in the ability to reproduce. The realized change in mean fitness between generations may deviate from $V_A(w)$ because of concurrent effects of genetic mutations, gene flow, environmental change, or gene–environment interactions (8, 9, 11). Nonetheless, a nonzero value of $V_A(w)$ indicates that, all else being equal, natural selection contributes to an increase in mean fitness (8, 9). It also indicates that at least some of the traits that determine individual fitness are currently evolving in response to selection. Thus, $V_A(w)$ is arguably the most important evolutionary parameter in any population (9, 12).

Robust estimation of $V_A(w)$ requires accurate measures both of individual fitness and pairwise genetic relatedness for large numbers of individuals. Such data are difficult to collect for wild populations of animals or plants (13). Moreover, their analysis is made challenging by the distribution of individual fitness, which generally does not conform well to common statistical methods (14). Consequently, our knowledge of $V_A(w)$ in natural populations is currently limited: two reviews (12, 14) report estimates of $V_A(w)$ from 16 populations of 13 plant and (nonhuman) animal species with fitness measured over complete lifetimes (we discuss these results alongside our own). However, notwithstanding possible issues specific to each analysis (such as the omission of important nongenetic sources of similarity between relatives), most of these estimates were obtained from Gaussian models [for exceptions, see (10)], which generally do not fit the distribution of fitness well. In natural populations, the distribution of fitness of all individuals is typically both highly right-skewed, with most individuals having low values but a few having very high values, and zero-inflated, with an excess of zeros over and above what would otherwise be expected (zero inflation may, for example, be generated by high levels of juvenile mortality). Estimates of $V_A(w)$ from Gaussian models, and their associated uncertainty, may thus be unreliable (14, 15).

In this study, we addressed the gap in our knowledge of the value of $V_A(w)$ in the wild and its implications in terms of adaptation, trait evolution, and population dynamics. We developed and applied Bayesian quantitative genetic methods to data from long-term studies of 19 free-living vertebrate populations with high-quality lifetime reproduction and multigenerational relatedness data. Covering more populations and species than all previous studies combined, these 19 populations of 15 different species (six bird and nine mammal species) have contrasting ecologies, life histories, and social systems (10) (tables S1 and S2) and are located in diverse terrestrial biomes and continents (Fig. 1). Our analysis is restricted to birds and mammals because of their predominance among long-term studies with suitable data (13). The populations

[1]Research School of Biology, Australian National University, Canberra, ACT, Australia. [2]School of Biology, University of St Andrews, St Andrews, Fife, UK. [3]Institut de Systématique, Évolution, Biodiversité (ISYEB), École Pratique des Hautes Études, PSL, MNHN, CNRS, SU, UA, Paris, France. [4]School of Biological Sciences, University of Auckland, Auckland, New Zealand. [5]Departments of Biology and Evolutionary Anthropology, Duke University, Durham, NC, USA. [6]Forest and Conservation Sciences, University of British Columbia, Vancouver, British Columbia, Canada. [7]Departments of Evolutionary Ecology and Evolutionary Genetics, Leibniz Institute for Zoo and Wildlife Research, Berlin, Germany. [8]Department of Biological Sciences, University of Alberta, Edmonton, Alberta, Canada. [9]Institute of Zoology, Zoological Society of London, Regents Park, London, UK. [10]Centre for Research in Animal Behaviour, University of Exeter, Penryn, UK. [11]Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland. [12]Centre d'Écologie Fonctionnelle et Évolutive, Université de Montpellier, CNRS, EPHE, IRD, Montpellier, France. [13]Department of Zoology, University of Cambridge, Cambridge, UK. [14]Mammal Research Institute, University of Pretoria, Pretoria, South Africa. [15]Edward Grey Institute, Department of Zoology, University of Oxford, Oxford, UK. [16]Department of Ecology and Genetics, Uppsala University, Uppsala, Sweden. [17]Centre for Ecology and Conservation, University of Exeter, Penryn, UK. [18]Département de Biologie, Université de Sherbrooke, Sherbrooke, Québec, Canada. [19]Zoological Museum, University of Zurich,, Zurich, Switzerland. [20]Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, CO, USA. [21]Biology Department, Oxford College, Emory University, Oxford, GA, USA. [22]School of Biological Sciences, Illinois State University, Normal, IL, USA. [23]Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, UK. [24]Centre for Biodiversity Dynamics, Norwegian University of Science and Technology (NTNU), Trondheim, Norway. [25]School of Biological Sciences, University of Aberdeen, Aberdeen, UK. [26]Ecology and Evolutionary Biology, School of Biosciences, University of Sheffield, Sheffield, UK. [27]Department of Animal Ecology, Netherlands Institute of Ecology (NIOO-KNAW), Wageningen, Netherlands.
*Corresponding author. Email: timotheebonnetc@gmail.com
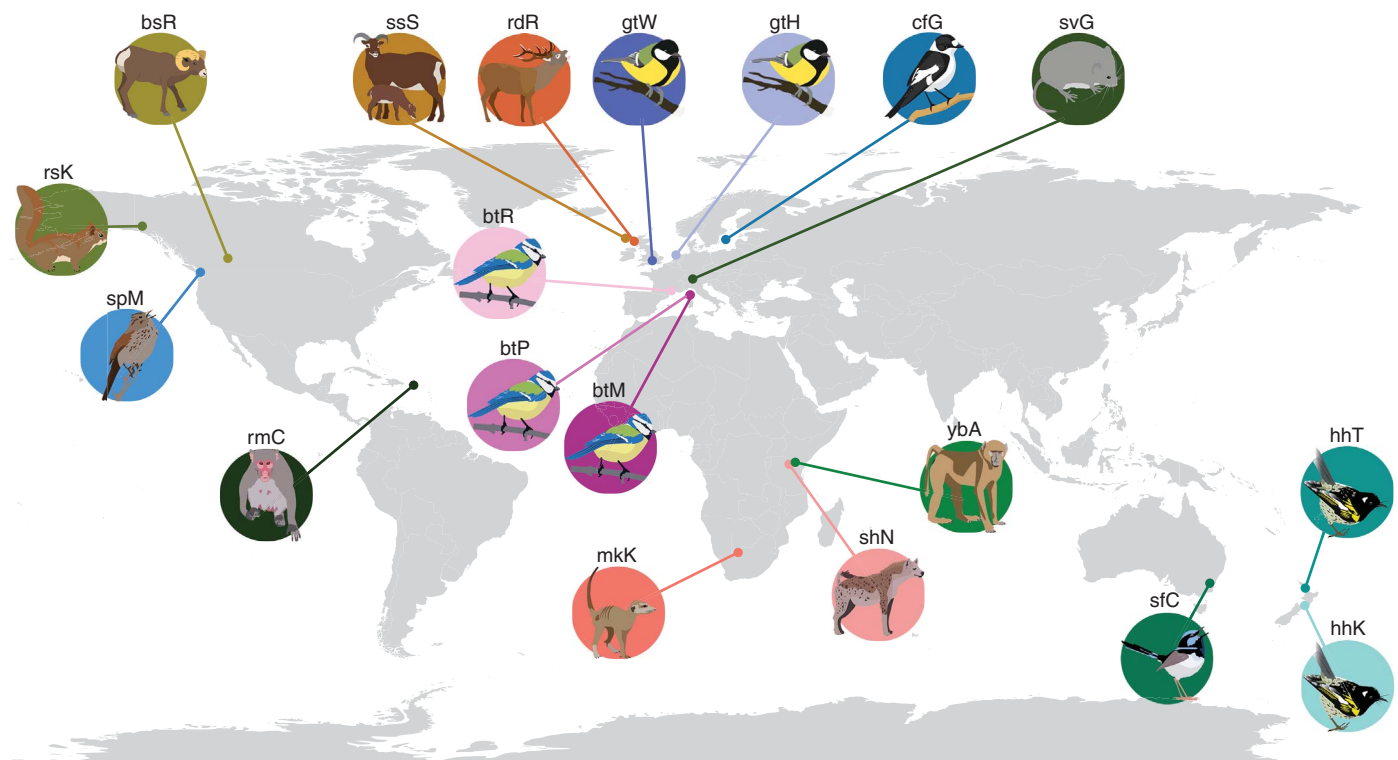
Fig. 1. Locations of the 19 long-term population studies. Abbreviations are as follows: bsR, bighorn sheep on Ram Mountain in Canada; ssS, Soay sheep on St Kilda, UK; rdR, red deer on the Isle of Rum, UK; gtW, great tits in Wytham Woods, UK; gtH, great tits in Hoge Veluwe, the Netherlands; cfG, collared flycatchers on Gotland, Sweden; svG, snow voles in Graubünden, Switzerland; rsK, red squirrels in Kluane, Canada; btR, blue tits at la Rouvière, France; spM, song sparrows on Mandarte Island, Canada; btP, blue tits at Pirio, France; btM, blue tits at Muro, France; rmC, rhesus macaques on Cayo Santiago, Puerto Rico; ybA, yellow baboons at Amboseli National Park in Kenya; hhT, hihi on Tiritiri Matangi Island, New Zealand; shN, spotted hyenas in the Ngorongoro Crater, Tanzania; mkK, meerkats in the Kalahari, South Africa; sfC, superb fairy-wrens in Canberra, Australia; hhK, hihi in Karori, New Zealand.

have been monitored for between 11 and 63 years, providing fitness records for 561 fully monitored cohorts totaling 249,430 individuals of both sexes (*10*). For all datasets used here, an individual's fitness was measured as "lifetime breeding success," or the total number of offspring produced over its lifetime, irrespective of offspring survival. While there are numerous definitions of fitness, each motivated by different theoretical frameworks (*16*), measuring fitness as lifetime breeding success corresponds most closely to a life cycle–calibrated "zygote-to-zygote" definition of individual fitness, consistent with quantitative genetic theory (*17*). Individuals were identified soon after birth or hatching, and fitness was estimated for all known individuals in each population, including the often-large proportion that died as juveniles (*10*). We modeled absolute lifetime breeding success using a quantitative genetic form of a mixed-effects model known as an "animal model" (*18*), assuming that lifetime breeding success follows zero-inflated overdispersed Poisson distributions and including relevant covariates (such as inbreeding, genetic group, sex, and cohort; see tables S3 and S4, supplementary text S1 for model details, figs. S1 and S2 for eval-

uation of model goodness of fit, supplementary text S2, and fig. S3 for prior distribution). The zero-inflated Poisson models were fitted to absolute fitness data, and the resulting parameter estimates, obtained on link-function scales, were then back-transformed to derive estimates of $V_A(w)$ and other components of variances for relative fitness on the scale of the data (*15*). We first ran one model for each study population and subsequently combined results into a meta-analysis (*10*).

We found evidence for additive genetic variance in relative fitness in multiple populations. Our models provided estimates of $V_A(w)$ with posterior modes ranging from 0.003 to 0.497 (Fig. 2A). The 95% credible intervals (95% CI) for $V_A(w)$ excluded values below 0.001 in 10 of the 19 populations and excluded values below 0.01 in eight (thresholds explained in caption of Fig. 2A and supplementary text S2 and S3). Therefore, there was clear evidence that selection contributed to genetic changes, and hence a predicted increase in fitness, in roughly half of the study populations (*9*, *19*). Across populations, the median of the posterior modes for $V_A(w)$ was 0.100, and the meta-analytic mean of $V_A(w)$ was 0.185, 95% CI [0.088; 0.303]. There was

also considerable variation among populations, with a meta-analytic among-population standard deviation in $V_A(w)$ of 0.11, 95% CI [0.01; 0.26]. The median and mean values of $V_A(w)$ were about four and two times larger than those of previous estimates (previous median: 0.023; previous mean: 0.092) (*12*, *14*). Our values can be considered large given theoretical considerations (supplementary text S3 and fig. S4), and they were robust to the modeling of possible confounders: inbreeding, sex, linear environmental changes in mean fitness, gene flow due to immigration, and variance among cohorts and among mothers (*10*) as well as mother-by-cohort interactions, social group effects (supplementary text S4, table S5, and fig. S5), and the social inheritance of social dominance within families (supplementary text S5 and figs. S6 and S7). For completeness, we also present estimates relating to an alternative formulation of Fisher's fundamental theorem expressing change in terms of absolute fitness leading to the same conclusions (supplementary text S6 and fig. S8).

Previous work on adaptive evolution has often focused on the heritability of fitness, $h^2(w) = V_A(w)/V_p(w)$, where $V_p(w)$ is the phenotypic variance in relative fitness, or

the "opportunity for selection" (20). However, $h^2(w)$ may be a poor measure of the overall rate of adaptive evolution (20). In natural conditions, stochastic or unaccounted environmen-

tal variation is expected to dominate variation in individual fitness, even in the presence of large deterministic sources of variation in fitness (21), so that $h^2(w)$ may be small even

when $V_A(w)$ is large (21, 22). In line with this expectation, we found that $h^2(w)$ was generally small, with a meta-analytic average of 2.99%, 95% CI [0.80; 6.60%] and a value of <1% in
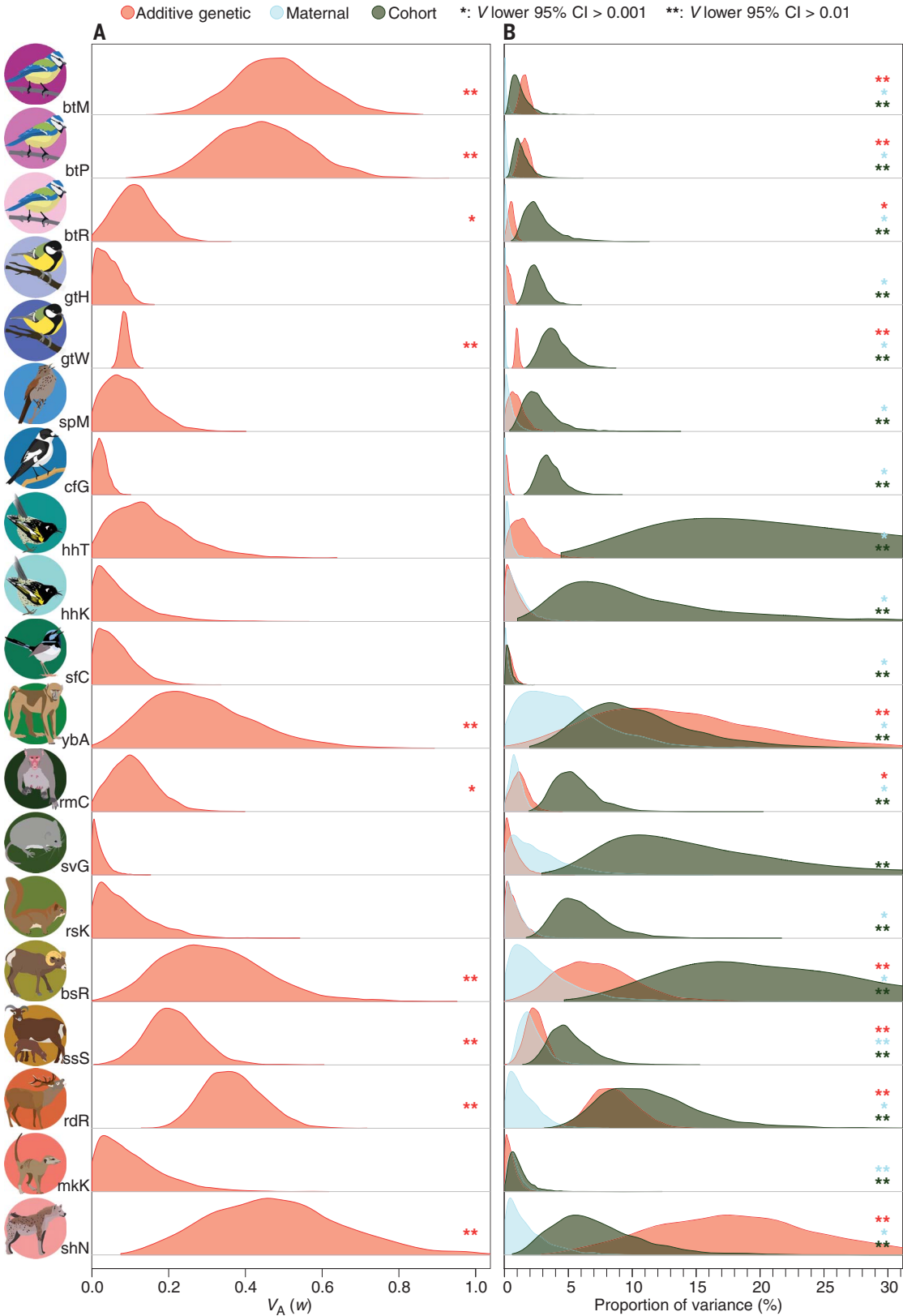
Fig. 2. Additive genetic variance and other components of variance in relative fitness. Panels show posterior distributions of (**A**) additive genetic variance in relative fitness, $V_A(w)$, and (**B**) proportion of phenotypic variance in fitness due to different variance components: additive genetic variance (i.e., heritability; red), maternal effect variance (light blue), and cohort variance (dark green). Species are ordered by phylogenetic proximity. Each distribution has an area of 1 but is scaled arbitrarily on the y axis to aid comparison. Single asterisk indicates that the 95% CI of a variance component does not overlap 0.001 [approximately the mode of the prior distribution for $V_A(w)$; supplementary text S2]. Double asterisk indicates that the 95% CI does not overlap 0.01 (the approximate threshold between small and moderate rates of adaptive evolution; supplementary text S3). Asterisks indicate absolute variance values, not proportions of variance. Abbreviations of population names are as in Fig. 1.

11 populations (Fig. 2B), similar to previous estimates of $h^2(w)$ (14). Nevertheless, estimates of $h^2(w)$ were of similar magnitude to the proportion of variance explained by maternal effect and cohort variances (Fig. 2B, supplementary text S7, and tables S6 to S10 for parameter estimates on different scales). Furthermore, $h^2(w)$ was highly variable between populations and was sometimes substantial, with posterior modes ranging from 0.019 to 17.1%.

What do our estimates of $V_A(w)$ imply about the evolution of traits in our study populations? $V_A(w)$ is the partial increase in fitness expected to result from the combined responses to selection across heritable traits (23). Therefore, a nonzero $V_A(w)$, as was found for at least half of our study populations, implies that for one or several traits, the responses to selection tend to cause adaptive change, although the total change may be affected by mutations or environmental change (19). The value of $V_A(w)$ sets an upper bound for the possible per-generation response to selection of any trait (19). Given the meta-analytic estimate of $V_A(w)$ = 0.185 and a trait with a heritability of 0.3 [an average value for trait heritability in wild populations (24)], the maximal rate of response to selection is 0.24 standard deviations per generation (10, 19). Across our 19 populations, the upper bound of response to selection for a trait with a heritability of 0.30 varies from 0.05, 95% CI [0.01; 0.13], to 0.39, 95% CI [0.29; 0.50], standard deviations. These upper bounds are substantial: For comparison, in natural populations the rates of phenotypic change, irrespective of whether the change is known to be adaptive, are rarely above 0.03 standard deviations (~10% of estimates) and only very rarely above 0.13 standard deviations (~5% of estimates) (2). Furthermore, evolutionary studies of wild populations, including several conducted in our study populations, have often failed to detect phenotypic change in response to current selection (5, 6, 25). Our results may therefore appear at odds with these observations. However, attempts to estimate genetic evolution of traits, as opposed to just phenotypic trends, remain rare and underpowered (25). Genetic evolution of traits may be masked at the phenotypic level, either because phenotypic plasticity hides genetic change (6) or because direct evolution is counterbalanced by the evolution of "indirect genetic effects," that is, the effect of other individuals' genotypes (26). Moreover, approaches to estimating genetic change for a trait, such as estimation of trends in individual genetic merit ("breeding values") (27) or by estimation of polygenic scores (28), may have limited statistical power. Finally, if $V_A(w)$ is ultimately driven by the cumulative effects of many traits evolving in response to selection, the evolutionary change in each trait will be small and even more difficult to identify statistically. Any or all of these scenarios could prevent observed rates of phenotypic change in single traits from reaching the upper bound of what might be possible given the observed levels of $V_A(w)$.

Irrespective of the rates of adaptive evolution in the potentially many traits that contribute to $V_A(w)$, our estimates of their combined effect, summarized in $V_A(w)$, indicate that adaptive evolution may have substantially affected recent population dynamics (see supplementary text S6 and S8 and fig. S8). For instance, in a thought experiment assuming that no forces oppose adaptive evolution and that $V_A(w)$ remains constant, 11 of our 19 populations would recover from an arbitrary one-third reduction in fitness in fewer than 10 generations (supplementary text S8). Moreover, the median $V_A(w)$ of 0.10 means that in half the populations, natural selection tends to increase mean absolute fitness by at least 10% every generation. Whereas such a change would lead to exponential population growth if not counterbalanced, none of our study populations showed any exponential increase in population size such as that predicted by the thought experiment (supplementary text S9). This indicates that any adaptive evolution was countered by simultaneous deleterious effects of other processes such as mutation, gene flow, or environmental changes (19). The presence of these counterbalancing forces, as well as potential changes in future selective pressures and the potential instability of $V_A(w)$ in future environments, makes it impossible to project whether the contemporary adaptive evolution that our results indicate is sufficiently fast and lasting to ensure population persistence. Other studies that focused on specific traits, rather than on the net effect of selection on fitness, suggest that short-term phenotypic changes in response to climate change are overall insufficient to ensure the persistence of populations (29, 30). Crucially, however, our finding that most populations harbor biologically meaningful levels of additive genetic variance in fitness indicates that the machinery of adaptive evolution often operates at a substantial pace on generation-to-generation time scales. Without ongoing adaptive genetic changes, these populations would presumably have had (often substantially) lower growth rates over recent generations.

REFERENCES AND NOTES

1. F. Pelletier, D. W. Coltman, *BMC Biol.* **16**, 7 (2018).
2. A. P. Hendry, in *Eco-evolutionary Dynamics* (Princeton Univ. Press, 2017), chap. 3.
3. J. N. Thompson, *Relentless Evolution* (Univ. of Chicago Press, 2013).
4. P. Karell, K. Ahola, T. Karstinen, J. Valkama, J. E. Brommer, *Nat. Commun.* **2**, 208 (2011).
5. T. Bonnet *et al.*, *PLOS Biol.* **17**, e3000493 (2019).
6. J. Merilä, B. C. Sheldon, L. E. B. Kruuk, *Genetica* **112**, 199–222 (2001).
7. B. Pujol *et al.*, *Trends Ecol. Evol.* **33**, 337–346 (2018).
8. G. R. Price, *Ann. Hum. Genet.* **36**, 129–140 (1972).
9. A. Grafen, *J. Theor. Biol.* **456**, 175–189 (2018).
10. Material and methods are provided as supplementary materials.
11. D. N. Fisher, A. G. McAdam, *Evol. Lett.* **3**, 4–14 (2019).
12. A. Burt, *Evolution* **49**, 1–8 (1995).
13. T. Clutton-Brock, B. C. Sheldon, *Trends Ecol. Evol.* **25**, 562–573 (2010).
14. A. P. Hendry, D. J. Schoen, M. E. Wolak, J. M. Reid, *Annu. Rev. Ecol. Evol. Syst.* **49**, 457–476 (2018).
15. T. Bonnet, M. B. Morrissey, L. E. B. Kruuk, *J. Hered.* **110**, 383–395 (2019).
16. B.-E. Sæther, S. Engen, *Trends Ecol. Evol.* **30**, 273–281 (2015).
17. J. D. Hadfield, in *The Evolution of Parental Care*, N. J. Royle, P. T. Smiseth, M. Kölliker, Eds. (Oxford Univ. Press, 2012), pp. 267–284.
18. L. E. B. Kruuk, *Philos. Trans. R. Soc. London Ser. B* **359**, 873–890 (2004).
19. T. F. Hansen, C. Pélabon, D. Houle, *Evol. Biol.* **38**, 258–277 (2011).
20. R. E. Snyder, S. P. Ellner, *Am. Nat.* **191**, E90–E107 (2018).
21. T. Price, D. Schluter, *Evolution* **45**, 853–861 (1991).
22. R. Fisher, *The Genetical Theory of Natural Selection* (Clarendon Press, ed. 1, 1930).
23. J. Merilä, A. P. Hendry, *Evol. Appl.* **7**, 1–14 (2014).
24. P. Bijma, *Heredity* **112**, 61–69 (2014).
25. J. D. Hadfield, A. J. Wilson, D. Garant, B. C. Sheldon, L. E. B. Kruuk, *Am. Nat.* **175**, 116–125 (2010).
26. J. P. Beauchamp, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 7774–7779 (2016).
27. B. Walsh, M. Lynch, in *Evolution and Selection of Quantitative Traits* (Oxford Univ. Press, 2018), chap. 6.
28. E. Postma, in *Quantitative Genetics in the Wild*, A. Charmantier, D. Garant, L. E. B. Kruuk, Eds. (Oxford Univ. Press, ed. 1, 2014), chap. 2.
29. L. Browne, J. W. Wright, S. Fitz-Gibbon, P. F. Gugger, V. L. Sork, *Proc. Natl. Acad. Sci. U.S.A.* **116**, 25179–25185 (2019).
30. V. Radchuk *et al.*, *Nat. Commun.* **10**, 3109 (2019).

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abk0853
Materials and Methods
Supplementary Text S1 to S10
Figs. S1 to S10
Tables S1 to S10
References (*31–242*)
MDAR Reproducibility Checklist
Data S1
Code S1

By **Zahira Jaser**

Setting myself free

After 15 years working in banking, my body had given up. I was bedridden with pneumonia, and for the first time in years I had a reason to stop—and to reflect. I hadn't set out to be a banker. As an undergraduate studying politics and economics, I loved academic research, and I was told I excelled at it. My undergraduate supervisor even asked me to stay and work as a researcher after I graduated—but I wouldn't be paid. I grew up in Italy as a child of a Palestinian uprooted father and an Italian civil servant mother, and my family had spent much of my life wrestling with citizenship challenges and aiming just to make ends meet. Unpaid work was not a viable option.

When a U.S. investment bank offered me a job in London, I jumped at the opportunity. The trading floor was exhilarating, full of supersharp high achievers. There was a sense of having made it into "the club." But this came at a cost. The demands were relentless—to be "on" all the time, to constantly prove myself and meet ever higher sales targets, to fit in with the elite, white, masculine environment. My earnings freed me from the financial difficulties of previous generations, but as the years passed, I lost a sense of who I was. I pushed myself so hard for so long—until my body forced me to take time off.

To my surprise, my bedbound month was one of the happiest times of my life. I took the time to explore why I felt so demotivated to return to work, reading research papers on motivational theories, psychology papers on career change, sociologists' conversations on identity. The more I read, the more certain I became that I wanted to return to academic research. Banking had given me financial stability, and I now had the opportunity to choose my path anew.

I started a master's in organizational behavior, and I felt life come back to me. When I finished, I decided to go for a Ph.D. Some enlightened mentors valued and supported me as a nontraditional student, but returning to academia as an "old" early-career researcher wasn't easy. Some senior academic colleagues discouraged me, discounting the life and work experience I had already accumulated; they did not realize that the resilience, discipline, and organizational skills I had learned on the trading floor were massive assets in my academic pursuits. Others thought I was too old and that my previous work tainted my ability to conduct "objective" research. I was even asked whether I was sure this was not a midlife diversion, a chance to spend time at home with my kids and study flexibly.

> **"Through my research, I started to make sense of what I had experienced."**

The thing that kept me going was the joy of rediscovering myself. Through my research, I started to make sense of what I had experienced in the blur of the previous 15 years in banking: toxic managers, poor leadership decisions, clumsy intercultural exchanges. I started to appreciate the emotional labor I undertook to hide hardship and struggle, and I found my voice. I could be critical of the environment that had brought opportunities but had also gagged me in conformity. I could speak about inequalities, bias, and discrimination. I could be who I am, an Italian-Palestinian woman interested in justice, equality, and how people connect. After completing my Ph.D., I decided to stay in academia and secured a faculty position.

Academia is the right professional home for me, but I know it is far from perfect and not right for everyone. Some find it as crushing as I found banking. Individuals with nontraditional career paths or backgrounds may be dismissed by frustratingly close-minded academics. These biases and many more can make academic careers unwelcoming to many scholars, especially those from minority backgrounds. And precarious, even unpaid, work excludes too many, as it excluded me at the beginning.

But I have found, through my life experience and my research, that we need not remain professionally trapped. It may be easier said than done, but it is often possible and preferable to find work that responds to our needs and offers fulfilment. There is no need to wait, as I did, for your body to crumble. ∎

Zahira Jaser is an assistant professor at the University of Sussex Business School. Send your career story to SciCareerEditor@aaas.org.

ILLUSTRATION: ROBERT NEUBECKER
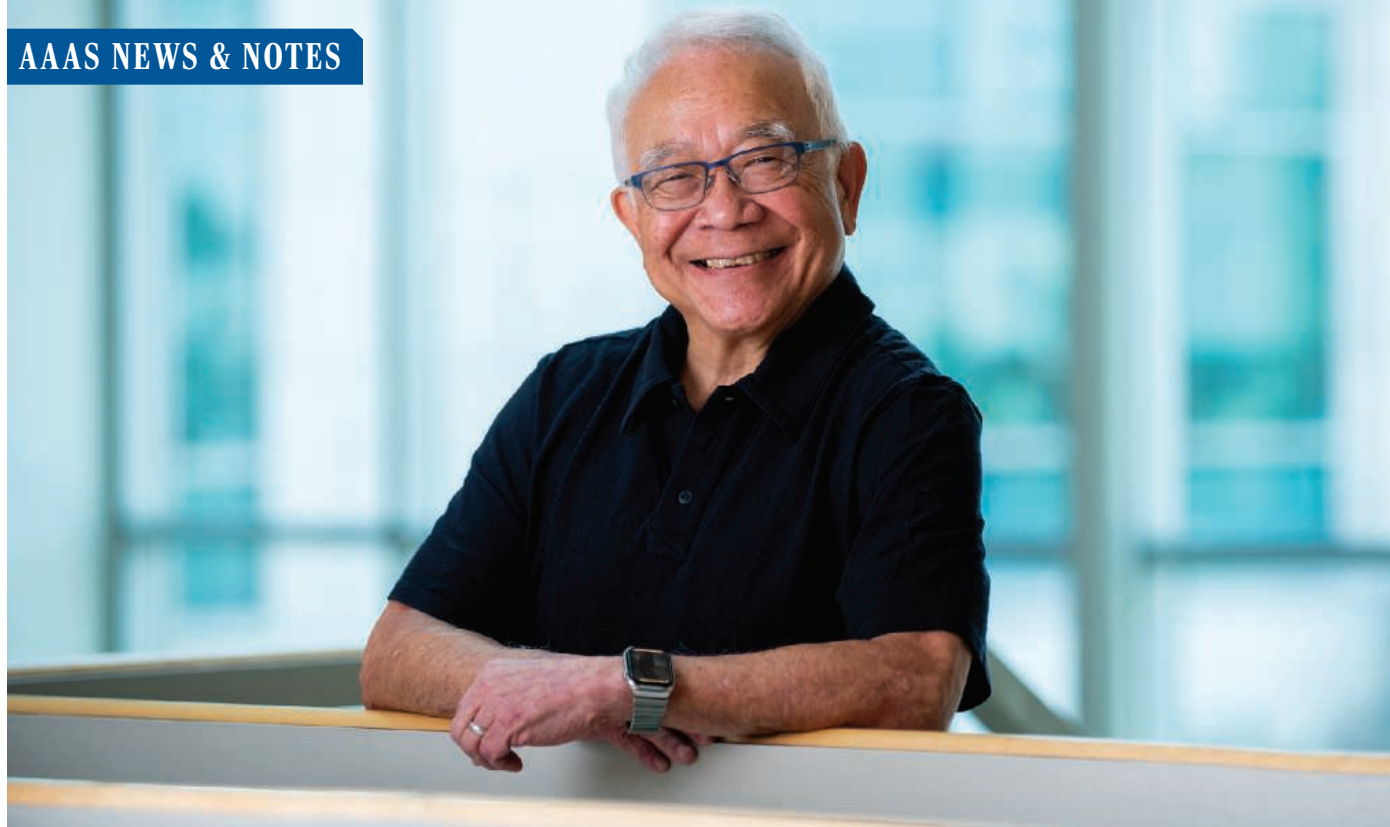
Keith Yamamoto to serve as AAAS president-elect

Susan Rosenberg and Jane Maienschein elected to AAAS Board of Directors

By **Andrea Korte**

Keith Yamamoto, a cellular and molecular pharmacologist and biologist and the vice chancellor for science policy and strategy at the University of California, San Francisco (UCSF), was chosen by the membership of the American Association for the Advancement of Science (AAAS) to serve as the organization's president-elect. Yamamoto's term begins immediately, and after serving as president-elect, he will serve for 1 year as AAAS president, followed by 1 year as immediate past-president.

During the annual election, held 7 to 21 April 2022, AAAS members also elected two new members of the AAAS Board of Directors: Susan Rosenberg, the Ben F. Love Chair in Cancer Research at the Baylor College of Medicine, and Jane Maienschein, University Professor and director of the Center for Biology and Society at Arizona State University. Rosenberg and Maienschein's terms begin immediately, and each will serve for 4 years on the board.

"I'm excited for AAAS to benefit from the expertise of Keith, Susan, and Jane as the organization's newest president-elect and board members. Their perspectives will be instrumental as AAAS continues its essential work to advance science and serve society," said Sudip S. Parikh, chief executive officer of AAAS and executive publisher of the *Science* family of journals.

Susan Amara, chair of the AAAS Board of Directors, added: "I look forward to collaborating with Keith, Susan, and Jane as they join the AAAS Board of Directors. The diversity of disciplines and backgrounds they represent echoes the diversity of AAAS's membership and will be invaluable as we work together in service of the association's mission."

Yamamoto calls for focus on scientific literacy, diversity, impact

Outreach to promote scientific literacy and an appreciation for using evidence to understand and solve the world's problems is "not just for scientists. It's for everyone," Yamamoto shared. This is one key area that he believes AAAS must continue to promote.

He recounted his early experiences with public outreach while pursuing his PhD in biochemical sciences at Princeton University under his thesis adviser, Bruce Alberts—who later became editor-in-chief of *Science*. Alberts brought area high schoolers onto campus to learn more about science and recruited assistant professors and students, including Yamamoto, to create experiments to share how science can be "interesting and fun and important"—a rewarding experience for all involved.

But when Alberts sought to expand the program, he received pushback. Alberts was told he should be working in the lab. Undeterred, he persisted in his efforts.

The experience, Yamamoto said, "convinced me that scientists have responsibilities that go beyond their work in their laboratories to do things that advance the scientific enterprise."

In his candidacy statement distributed to AAAS membership in advance of the annual election, he also identified two other priorities for AAAS.

First, AAAS can promote policies and practices that ensure a diverse, equitable, and inclusive scientific enterprise. "Science is a global enterprise that's going to move forward best if it's practiced by a diverse workforce that approaches scientific problems from different perspectives and points of view," Yamamoto said.

Second, AAAS can focus on building "a continuum from fundamental discovery to societal impact," so that scientists can visual-

ize the real effects of their work on societal issues and, in so doing, may alter the work we do or the ways we do it, he said.

Because of the breadth and scope of its transdisciplinary, international membership, AAAS should be a leading voice in these conversations, now more than ever, Yamamoto noted.

"Keith will contribute significant experience in science policy to AAAS, both at the federal and state levels," said UCSF Chancellor Sam Hawgood. "He brings a deep knowledge of biology and the transdisciplinary push toward precision medicine. He understands the many roles that scientists play in society. And he is committed to advancing diversity in science."

Yamamoto brings to AAAS a significant background in science policy. He became UCSF's first vice chancellor for science policy and strategy in 2015, but he has held a range of leadership roles there since joining the institution as a faculty member in 1976. In addition to his groundbreaking research on signaling and transcriptional regulation by nuclear receptors, Yamamoto has served as chair of the Department of Cellular and Molecular Pharmacology, vice dean for research in the School of Medicine, and vice chancellor for research.

Outside of UCSF, he cochairs the Science & Technology Action Committee (along with Parikh and others), which brings together nonprofit, academic, foundation, and corporate leaders to encourage US investments in science and technology research, development, and education.

A member of AAAS since 1977, he was elected as Fellow of AAAS in 2002—recognized for his scientifically or socially distinguished efforts on behalf of the advancement of science or its applications. He has served two terms on the AAAS Committee on Nominations and has participated in several AAAS forums and roundtables for science and public policy.

Rosenberg and Maienschein bring expertise to Board of Directors

Newly elected members of the AAAS Board of Directors Susan Rosenberg and Jane Maienschein bring a wealth of scientific expertise to the governing body, which is responsible for the affairs of the association.

Rosenberg's research at the Baylor College of Medicine focuses on molecular mechanisms of genome instability in evolution, antibiotic resistance, and cancer. Rosenberg served as a council delegate for the AAAS Section on Biological Sciences and has conducted extensive work on the AAAS Governance Modernization Working Group. She was elevated to the rank of AAAS Fellow in 2010.

In addition to serving as a University Professor of History of Science at ASU, Maienschein leads the university's Center for Biology and Society, which promotes research, education, and engagement related to the study of the life sciences and their interconnections with society. She also serves as a Fellow and runs the history program at the Marine Biological Laboratory. Maienschein has served as chair and council delegate for the AAAS Section on History and Philosophy of Science, as well as chair of the Section on Societal Impacts on Science and Engineering. She has also served on the History Committee for the 150-year anniversary and on the Program Committee. She joined the ranks of AAAS Fellows in 1996.

Maienschein noted that her scientist father gave her a AAAS gift membership in 1976, and she has remained a lifelong member committed to the vision and values. "Instead of advocating for science in a vacuum, AAAS promotes science 'for the benefit of all people,' emphasizing education and communication, and recognizing that science exists in a complex and messy society and often has to deal with uncertainty. As AAAS approaches its 175th anniversary, I look forward to helping launch the next 175," she said.