Putting cosmic timestamps on historical events p. 124

Global warming speeds drought development pp. 130 & 187

Fix flawed reporting of AI system performance p. 136

Science

$15
14 APRIL 2023
SPECIAL ISSUE
science.org
AAAS

HUMAN REPRODUCTION

Progress and roadblocks ahead p. 148



REPRODUCTIVE MEDICINE

The quest for safer and better planned reproduction

REVIEWS

Barriers to progress in pregnancy research: How can we break through? p. 150
A brief history and future prospects of contraception p. 154
Understanding the genetics of human infertility p. 158
Assisted reproductive technologies at the nexus of fertility treatment and disease prevention p. 164

RELATED ITEMS

BOOKS *ET AL.* p. 141 PODCAST SERIES

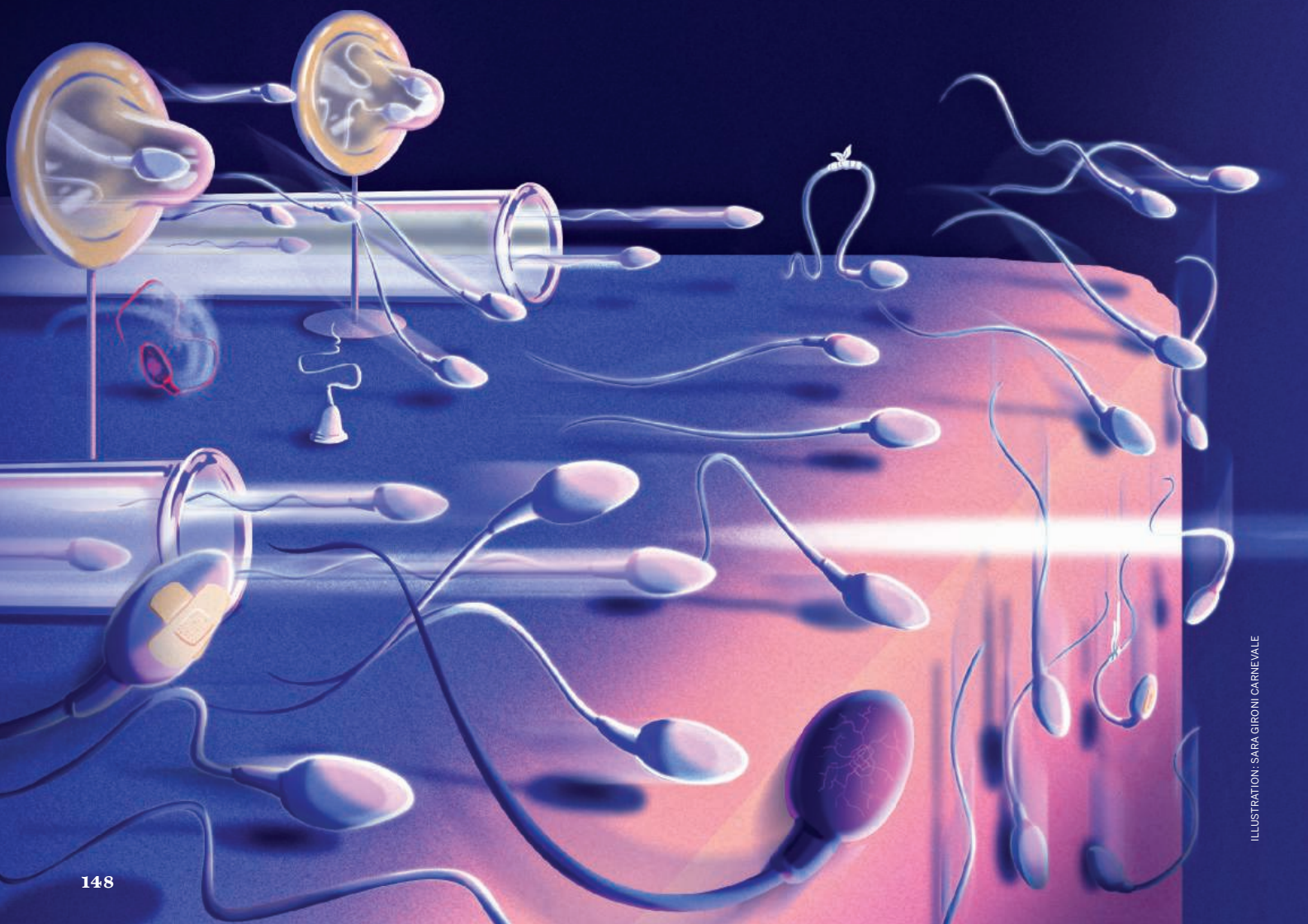


ILLUSTRATION: SARA GIRONI CARNEVALE

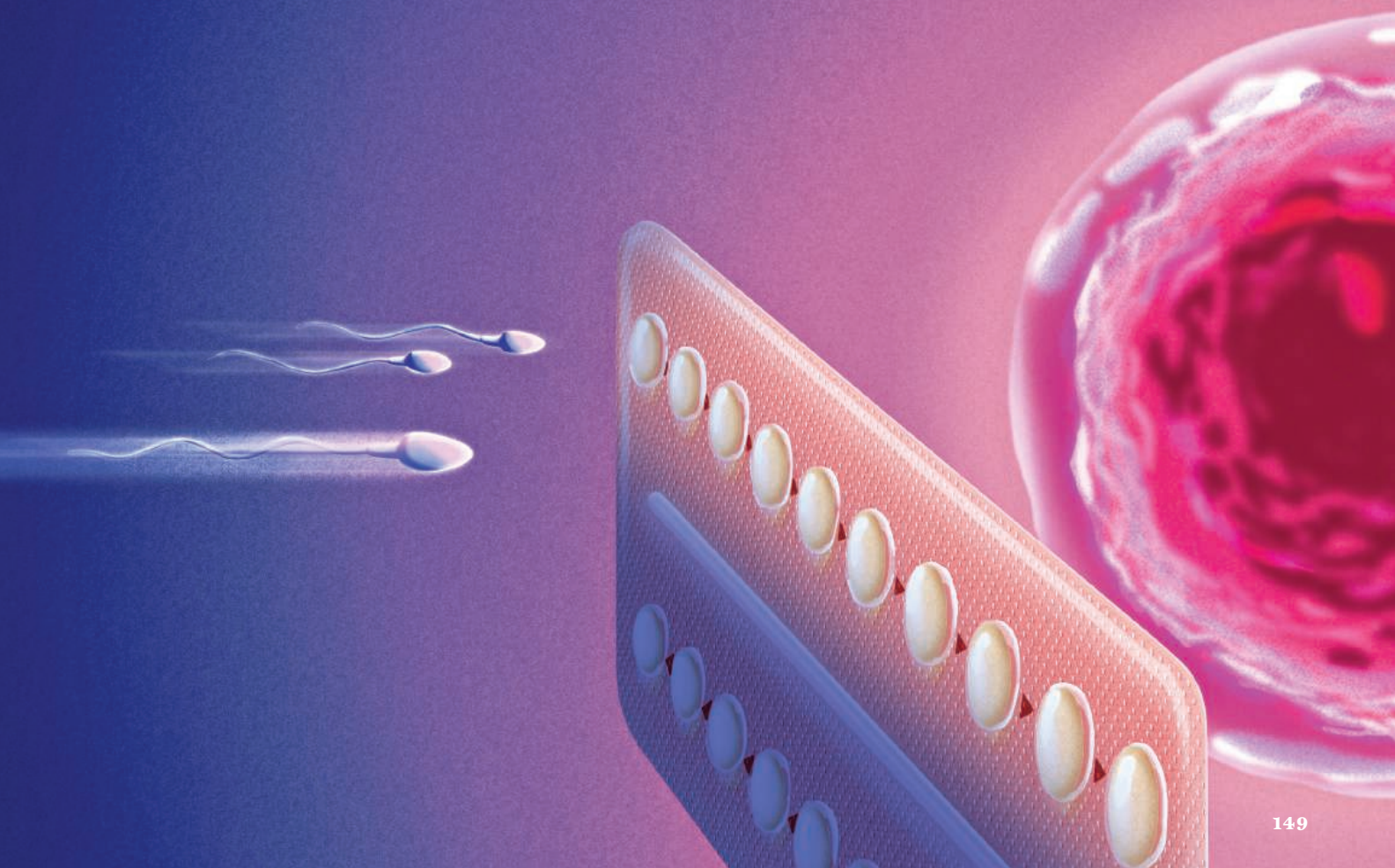
By **Yevgeniya Nusinovich**

Since ancient times, humans have been trying to exercise control over their reproductive decisions, whether to avoid undesired pregnancy or to improve their chances of conceiving. In addition, the risks of pregnancy and childbirth have always been a major challenge. Humanity has come a long way since the days of herbal poultices and linen condoms, and modern medicine allows for unprecedented control of reproductive decisions and health. However, major limitations remain in both health care access and scientific knowledge.

Despite the rapid progress that has been made over the past century, pregnancy continues to pose health risks. Recent research has provided insights into the biology of and interactions between the mother, fetus, and placenta and their roles in both healthy pregnancies and complications. For preventing pregnancy, recent research offers the promise of expanding contraceptive options for men and strategies to combine contraception with disease prevention. For people struggling with infertility, insights from modern genetics are increasingly being applied to make a diagnosis and identify the best approach to treatment. Nevertheless, specific interventions are limited, not least because it is unethical to genetically modify human germ cells. To treat infertility regardless of the biological cause or to help people avoid passing on genetic diseases to their offspring, assisted reproductive technology offers an ever-increasing array of options, despite important risks and limitations.

With continued discoveries in human genetics and cell biology, progress in our understanding of human reproduction is all but inevitable. However, improving worldwide access to state-of-the-art reproductive care will be key to ensuring that these discoveries benefit human health.

10.1126/science.adh8322



The illustration depicts sperm trying to reach an oocyte. Some sperm are damaged and receiving medical help, whereas others encounter contraceptive obstacles.

REVIEW

Barriers to progress in pregnancy research: How can we break through?

Sarah J. Stock[1,2,3]* and Catherine E. Aiken[4,5,6]

Healthy pregnancies are fundamental to healthy populations, but very few therapies to improve pregnancy outcomes are available. Fundamental concepts—for example, placentation or the mechanisms that control the onset of labor—remain understudied and incompletely understood. A key issue is that research efforts must capture the complexity of the tripartite maternal-placental-fetal system, the dynamics of which change throughout gestation. Studying pregnancy disorders is complicated by the difficulty of creating maternal-placental-fetal interfaces in vitro and the uncertain relevance of animal models to human pregnancy. However, newer approaches include trophoblast organoids to model the developing placenta and integrated data-science approaches to study longer-term outcomes. These approaches provide insights into the physiology of healthy pregnancy, which is the first step to identifying therapeutic targets in pregnancy disorders.

Pregnancy health is a key determinant of population health, with around 85% of women experiencing pregnancy during their lifetime (*1*). Pregnancy and childbirth have profound effects on immediate and life-course health for both infants and mothers. It is therefore crucial to develop effective methods for preventing and treating pregnancy complications. Pregnancy complications are common (Table 1) and increasing, as more women globally enter pregnancy with comorbidities, including obesity, advanced age, and coexisting medical disorders (*2*). At present, there are very few therapies designed specifically for pregnancy conditions, and there has been a virtual absence of new products coming to market in recent years (*3*, *4*). Excluding reformulations, in the past 30 years, there have been two drugs licensed for use in pregnancy. Atosiban is an oxytocin antagonist licensed as a preterm labor treatment in the United Kingdom but not the United States. Makena (hydroxyprogesterone caproate injection) was approved for use in pregnancy for preterm birth prevention in the United States, but approval has now been revoked (*5*). Another two drugs (carboprost, a synthetic prostaglandin; and carbetocin, an oxytocin agonist) have been licensed for postpartum use for bleeding control. The reasons for the lack of progress to safe and effective treatments for pregnancy complications are complex. It seems likely that systemic sexism and paternalism contribute, coupled with risk-averse or defensive practices that lead to purposive exclusion of pregnant women from research (*6*). Pregnancy is also particularly challenging to study. First, pregnancy complications result from three-way interactions between the pregnant woman, the developing baby, and the placenta, which are complex to model in vitro. Second, evolving changes over gestation need to be accounted for. Third, in vivo study of human pregnancy is restricted by the inaccessibility of samples from within the uterine environment. Finally, pronounced disparities in pregnancy outcomes exist between and within countries. This indicates that interactions between social determinants of health and biological pathways must be considered if progress is going to be made on a global scale.

[1]University of Edinburgh Usher Institute, Edinburgh EH16 4UX, UK. [2]University of Edinburgh MRC Centre for Reproductive Health, Edinburgh EH16 4TJ, UK. [3]Wellcome Leap In Utero Program, Wellcome Leap Inc., Culver City, CA 90232, USA. [4]The Rosie Hospital and NIHR Cambridge Biomedical Research Centre, Department of Obstetrics and Gynaecology, University of Cambridge, Cambridge CB2 0SW, UK. [5]Centre for Trophoblast Research, Department of Obstetrics and Gynaecology, University of Cambridge, Cambridge CB2 0SW, UK. [6]Wellcome-MRC Institute of Metabolic Science and Medical Research Council Metabolic Diseases Unit, Addenbrooke's Hospital, University of Cambridge, Cambridge CB2 0QQ, UK.
*Corresponding author. Email: sarah.stock@ed.ac.uk

We explore three examples of fundamental knowledge gaps, each of which is often considered to relate to one particular aspect of the maternal-placental-fetal triad: placenta (preeclampsia), mother (gestational diabetes), and fetus (preterm birth). We aim to illustrate how each is more usefully considered in the context of a dynamic tripartite system and highlight promising research directions.

Preeclampsia

The placenta is a pregnancy-specific organ that grows de novo from the extraembryonic tissues of the conceptus, interfacing with both the maternal and the fetal circulations to sustain fetal development. The placenta dynamically adopts the functions of the lungs, kidneys, gut, endocrine system, and many other body systems to maintain fetal homeostasis throughout pregnancy. Placental support for the fetus also extends to subverting maternal physiology, for example, altering cardiac output to meet fetal requirements (*7*) However, complete dependence on the placenta as a flexible support system also carries risk to the fetus. Placentas are unlike other human organs in that they grow rapidly—from a few cells to full size in a few short weeks—but also age rapidly (*8*). Tiny variations in placental development at the start of pregnancy can lead to catastrophic failure to sustain later fetal growth. In the earliest stages of pregnancy, optimal establishment of placentation, particularly interaction with maternal decidual cells (*9*), is crucial to reduce the risk of life-threatening disorders such as preeclampsia.

Preeclampsia affects 2 to 8% of all pregnancies worldwide (*10*) and remains a major cause of maternal and fetal mortality. Preeclampsia (at least in its early-onset form) results from

Table 1. Global estimates of the incidence of selected pregnancy complications. High-quality data on maternal and perinatal morbidity are not available in many settings, which is a barrier to pregnancy research. In this table, we present best available global estimates for selected outcomes. CI, confidence interval; UI, uncertainty interval.

Condition	Global estimate (%)	Reference
Preterm birth (before 37 weeks gestation)	10.6 (UI 8.7 to 11.9)*	(*42*)
Preeclampsia	2.16 (95% CI 2.11 to 2.22)†	(*10*)
Gestational diabetes	14.0 (95% CI 13.97 to 14.04)‡	(*26*)
Low birthweight (<2500 g)	14.6 (UI 12.4 to 17.1)*	(*43*)
Maternal near miss (maternal life-threatening complication)	1.4 (95% CI 0.4 to 2.5)§	(*44*)
Postpartum hemorrhage	6.09 (95% CI 6.06 to 6.11)¶	(*45*)
Chorioamnionitis	3.9 (95% CI 1.8 to 6.8)¶	(*46*)
Sepsis	0.05 (95% CI 0.03 to 0.07)¶	(*46*)
Birth asphyxia (hypoxic ischemic encephalopathy)	0.15 (95% CI 0.13 to 0.17*	(*47*)

*Per 100 live births. †Per 100 pregnancies. ‡Per 100 pregnancies ≥24 weeks gestation. §Per 100 deliveries or live births. ¶Per 100 births.

incomplete conversion of the uterine spiral arteries by a particular subset of placental cells, the extravillous trophoblast, in a process that is fundamental to healthy pregnancy yet incompletely understood (Fig. 1). The quality of the interaction between the trophoblast and the maternal decidua depends on a complex multitude of factors, including both maternal and paternal genetics (*11*). Suboptimal connections between the developing placenta and the uterus early in pregnancy result in a constellation of damaging sequelae as pregnancy progresses, including severe maternal hypertension and fetal growth restriction (*12*).

Despite the high global burden of morbidity associated with preeclampsia (*10*), strategies for treatment that are in use now remain crude and poorly targeted. Magnesium sulfate and antihypertensives are the mainstays of therapy. Maintaining control of maternal blood pressure is essential, but the antihypertensives used to treat preeclampsia are limited and old-fashioned. Labetalol, a combined α and β blocker, is commonly used. Labetalol is relatively inefficient at reducing blood pressure, and, despite decades of use, its pharmacokinetics in pregnancy remain poorly understood (*13*). Importantly, no drug now used in the treatment of preeclampsia modifies the underlying disease process, which is the culmination of events that begin months earlier during early placentation. There is an urgent need to understand more about these early stages of placental development with a view to identifying targets for the prevention of preeclampsia and treatment of severe early-onset preeclampsia.

For preeclampsia prevention, repurposed basic drugs are in clinical use now, including calcium and low-dose aspirin (<300 mg) (*14*). Although evidence from clinical trials suggests that aspirin is effective in reducing the risk of preeclampsia by ~60% in selected women whose risk of preterm preeclampsia was >1 in 100 using a biomarker and risk factor–based screening algorithm (*15*), there is no clarity on the underlying mechanism of action. Possible theories include a direct effect on the extravillous trophoblast cells, reduction of platelet aggregation that reduces the risk of placental infarction, or reduction of endothelial activation in the maternal circulation [reviewed in (*16*)]. Our inability to distinguish between these theories exemplifies the difficulties with placental drug targeting.

Human placentas have several distinctive features, meaning that results from animal models are less directly applicable than in other areas of medicine. There is an enormous amount of variability among eutherian mammals in the basic structure and function of the placenta, reflecting both placental evolution and the ecological niche of various species (*17*). For experiments that study many functions—for example, oxygen or glucose transport—mammals such as sheep are extremely useful, but crucially, only other great apes share human patterns of placental invasion in early pregnancy. In humans, the extravillous trophoblast invades both through the uterus toward the maternal vessels and down the lumen of the uterine spiral arteries (Fig. 1). Given the importance of trophoblast invasion as the initiating event for preeclampsia, results from models in which this does not occur are less readily translatable.

A recent advance that is likely to provide important insights into fundamental placental biology is the development of placental organoids. These differentiate in vitro into both the cellular and syncytial components of placental epithelial cells within matrigel droplets (*18*). In the context of preeclampsia research, they can be induced to differentiate into extravillous trophoblast. Having a stable, renewable, three-dimensional (3D) "placenta in a dish" that contains multiple fully differentiated epithelial lineages opens possibilities for the discovery of targeted therapeutics. At present, limitations include the "inside-outside" structure of the organoids (Fig. 2), although whether there are any functional implications of this structure is presently unclear. Alternative methods for investigating the biology of early trophoblast include cultured trophoblast stem cells (*19*, *20*), which are easier to culture in large numbers but lack 3D structure and cell interaction. A further approach, which avoids the use of immortalized cell lines, is to derive fresh syncytiotrophoblast cultures (*21*); however, this negates the advantages of studying the interplay between different trophoblast lineages.

Future horizons in this work include moving toward developing stable coculture systems to model the maternal-placental-fetal triad. It is also now possible to derive organoids that recapitulate single decidual cells from maternal uterine tissue, opening possibilities for combined model systems that include both trophoblast and decidual components (*22*). Preeclampsia is a disorder that arises from the quality of the interaction between the mother and the placenta, and therefore finding potential therapeutic targets relies on being able to study trophoblast that is behaving as it would in healthy human pregnancy.

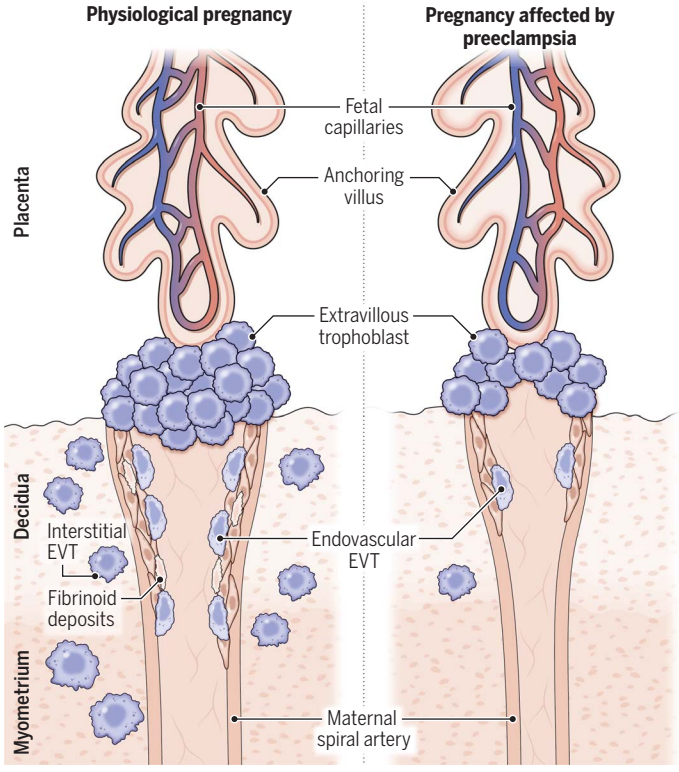


Fig. 1. Limited conversion of the spiral arteries in the context of preeclampsia. EVT, extravillous trophoblast.

Gestational diabetes

Integrating maternal aspects with fetal and placental model systems is also necessary to advance research into disorders that arise from maternal adaptation to pregnancy. Pregnancy involves the subversion of normal maternal homeostatic control for fetal benefit, encompassing virtually every physiological system. These changes are largely driven by placental hormone production, which changes over the course of pregnancy to reflect fetal requirements. Although these demands remain within the boundaries of physiological reserve in otherwise healthy mothers, underlying maternal disease tendencies are revealed when this capacity is exceeded. With global shifts in maternity populations toward later age at first birth and high rates of maternal obesity, the risk of pregnancy unmasking underlying physiological vulnerabilities increases. Compromised maternal adaptation to pregnancy carries serious consequences for the survival of both mother and baby.

A key example of maladaptation of maternal physiology is gestational diabetes. Pregnant women develop relative insulin resistance, driven

ILLUSTRATION: A. FISHER/*SCIENCE*

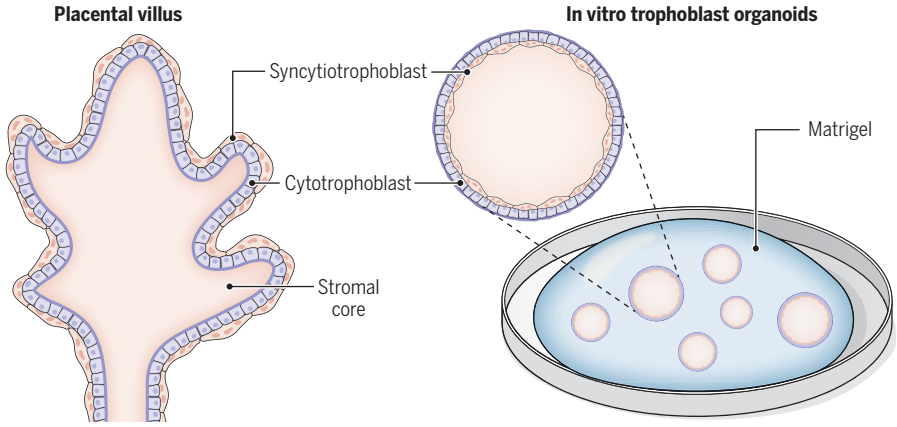


Fig. 2. "Inside-out" structure of trophoblast organoids derived from first-trimester placental villi.

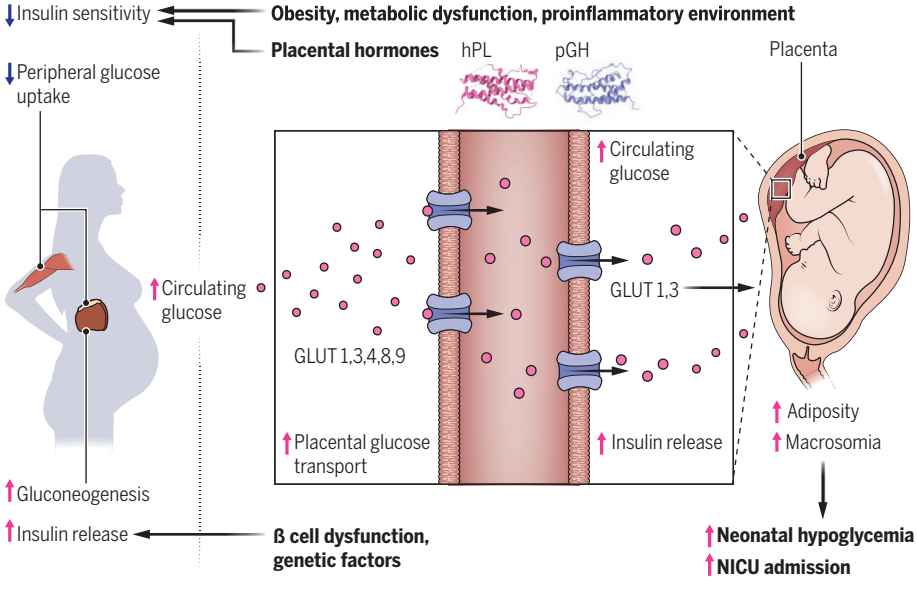


Fig. 3. Present understanding of the pathophysiology of gestational diabetes. The addition of placental hormones in a mother with one or more risk factors for diabetes (obesity, insulin resistance, other metabolic dysfunction, β cell dysfunction, genetic risk factors, and/or high proinflammatory cytokines) leads to high maternal circulating glucose. There is subsequent increased facilitated diffusion of glucose across the placenta, which provokes high fetal and neonatal insulin release. GLUT, glucose transporter; hPL, human placental lactogen; NICU, neonatal intensive care unit; pGH, placental growth hormone.

ILLUSTRATIONS (TOP TO BOTTOM): A. FISHER/*SCIENCE*; A. MASTIN/*SCIENCE*

by placental hormone synthesis (*23*), to ensure that more circulating glucose is available for fetal growth. Gestational diabetes arises when normal circulating glucose thresholds are exceeded (Fig. 3). In addition to unmasking an increased lifetime risk of type 2 diabetes in the mother, gestational diabetes carries risks for the fetus; fetal overgrowth and neonatal complications are problematic in the short term (*24*), in addition to longer-term metabolic dysregulation into childhood (*25*). With ~1 in 6 pregnancies globally affected by gestational diabetes (*26*), this is a problem that threatens population health. Defining optimal treatment strategies for gestational diabetes is thus a high priority.

At present, treatments used for gestational diabetes are repurposed directly from those used in type 2 diabetes; no gestational diabetes–specific therapies are available. Diet and lifestyle modification are first-line therapies but are only effective in ~50 to 66% of cases (*27*). Pharmacological options that are now recommended in various international guidelines include insulin, glyburide, and metformin, all of which are effective in controlling maternal hyperglycemia (*28*); however, repurposing drugs from standard diabetes care to pregnancy involves consideration of the potential impacts on all aspects of the maternal-placental-fetal triad. The placenta is a key consideration here; although assessing placental transfer of drugs between the maternal and fetal circulations is routine, the direct impact of drugs on the placenta itself is more often overlooked in assessing new therapies. An example is the impact of metformin, which is a known inhibitor of complex I of the electron transport chain that reduces placental adenosine triphosphate (ATP) production (*21*) and is associated with lower birthweight (*28*). These findings have led to considerable divergence in opinion regarding metformin as a first-line drug therapy in pregnancy. This example highlights not only the need for more rigorous evaluation of placental impacts of drug therapies in pregnancy but also the challenge of evaluating the longer-term impacts of intrauterine exposure to therapeutics on health into childhood or beyond. Follow-up of birth cohorts is expensive, slow, and difficult to achieve without introducing marked attrition bias. Studies that link electronic health records and other routinely collected data sources, especially those with enhanced methodologies such as sibling pairing, will be increasingly important to ensure that subtle alterations in long-term health outcomes that arise from suboptimal early life environments can be detected and monitored.

Spontaneous preterm labor

A further critical knowledge gap surrounds the mechanisms that control the onset of human labor, which must be appropriately timed for optimal birth outcomes. Preterm birth (<37 weeks gestation) affects ~1 in 10 pregnancies, resulting in sequelae ranging from neonatal death to impaired physical growth and neurodevelopment (*29*). By contrast, if birth is not achieved by 40 weeks, the risk of stillbirth increases (*30*).

Spontaneous labor at term is a coordinated process that involves maternal (cervix, myometrium, and decidua) and placental [placenta and fetal membranes (chorion and amnion)] tissues, accompanied by fetal maturation (*17*). Inflammation is a key feature, but the fundamental mechanisms that control the onset of parturition in humans remain unknown (*17*). In most mammals other than primates, labor is initiated by withdrawal of progesterone (*31*). In human pregnancy, there is no drop in circulating progesterone, but there is thought to be locally regulated "functional" progesterone withdrawal (*31*). However, the specific control mechanisms and transcriptional effects of this are not known, highlighting a key problem of using animal models for developing treatments for pregnancy conditions.

During gestation, exogenous stimuli, such as infection or trauma, or endogenous stimuli (including placental insufficiency, hemorrhage, hypoxia, or uterine overdistension) can prematurely stimulate inflammation in various reproductive tract tissues, propagating spontaneous

preterm labor (*32*). To date, many studies of preterm labor have focused on the responses of one tissue type. Recently, systems biology approaches applied to multiple tissues in tandem have shown promise (*33*). However, technological challenges remain to enable a whole-system approach that captures dynamic changes in the maternal, fetal, and placental compartments across gestation. This has contributed to a failure to convert promising preclinical targets to effective clinical treatments (*33*).

At present, there are very few effective drugs for the prevention of preterm labor. Progesterone supplementation can sometimes extend gestation in women who are thought to be at high risk of preterm birth. However, injectable 17-α-hydroxyprogesterone caproate (Makena) has had US Food and Drug Administration (FDA) approval withdrawn because postmarket studies failed to confirm clinical benefit (*5*). Vaginal progesterone is widely recommended and used off-label for preterm birth prevention instead. However, this also lacks FDA approval, and there is conflicting evidence as to its efficacy (*34*). Given the already-high concentrations of circulating progesterone, the mode of action of additional supplementary progesterone is not known.

An alternative, widely used preventive treatment in women with an increased risk of preterm birth is cervical cerclage. This is where a mesh or monofilament suture is inserted through the body of the cervix to help keep it closed, which carries risk of complications, including damage to the cervix, bleeding, and membrane rupture (*35*). A problem with both progesterone and cervical cerclage is the identification of individuals who would most benefit from treatment. At present, the risk of preterm birth is usually based on nonspecific clinical features such as a history of preterm birth or emerging signs of preterm labor such as a short cervix on ultrasound. Efforts to better understand both the genotype and phenotype of preterm birth, which has multiple underlying etiologies, may help to determine groups that would benefit from these treatments, as well as identifying much-needed new treatment targets.

In women with signs and symptoms of spontaneous preterm labor, tocolytic agents may temporarily slow or stop contractions, buying time to administer corticosteroids to promote fetal maturation or to access a higher-level health care facility. Recommended agents vary by country but include calcium channel blockers, betamimetics, nitric oxide donors, cyclooxygenase inhibitors, magnesium sulfate, and oxytocin receptor antagonists. In general, there is evidence that these agents may prolong pregnancy for 2 days to 1 week, but they have maternal side effects and evidence of neonatal or longer-term benefit is absent (*36*).

Much of the effort to develop preterm-birth treatments focuses on identifying candidates that prolong gestation (*37*). This assumes that advancing gestation will benefit infant outcomes, given that the severity of preterm morbidity is inversely related to gestation at birth. However, preterm labor is the final common pathway of multiple intrauterine challenges that can threaten the developing fetus (*32*). Infection, placental insufficiency, hemorrhage, and hypoxia are examples of direct risks to fetal life that are resolved by spontaneous or iatrogenic preterm birth. Injudicious treatments to keep a baby in a toxic uterine environment may thus be damaging, with potential long-lasting effects on childhood development. To optimize infant outcomes, we need to move to an approach that includes (i) early identification and treatment of pathologies that can lead to preterm labor, (ii) individualized evaluation of both the fetal and maternal responses to such threats, and (iii) timing birth to when the risks of continuing pregnancy outweigh the risks of prematurity. Focusing only on modulating the onset, or progress, of preterm labor, as most trials to date have done (*38*), only addresses the last part of this paradigm. In parallel, we need better methods of assessing fetal well-being and the effects on development, with therapeutics aimed at modifying underlying causes. The coordinated application of advances in in utero imaging (*39*), the development of wearable monitors (*40*), and the identification of circulating biomarkers such as nucleic acids (*41*) is a promising strategy toward this goal.

Conclusion

Development of therapies to treat pregnancy-related conditions is, at present, inadequate. For all the conditions we discussed, progress toward new therapies would be enhanced by the identification of improved biomarkers to stratify risk and target therapeutics. As of now, the biomarkers in clinical use either have poor predictive power [the ratio of soluble fms-like tyrosine kinase 1 (sFlt) to placental growth factor (PlGF) in preeclampsia] or are the manifestations of the conditions themselves (blood glucose in gestational diabetes mellitus and fetal fibronectin in preterm birth). The sequencing of extracellular RNA that is circulating in maternal blood is an exciting new development with potential for predicting pregnancy complications (*41*).

In the absence of safe and effective treatments, large numbers of women and their babies will continue to get sick or die. More research that takes an integrated approach to the dynamic interactions of the maternal-placental-fetal triad is urgently required. This may involve looking outside the traditional methodologies of laboratory animal and observational human cohort studies used in pregnancy research toward techniques used in other disciplines, for example, the development of digital twins and 3D tissue-architecture models. Addressing major gaps in our fundamental understanding of the physiology of pregnancy, development, and birth is necessary to identify candidate therapeutic targets for the benefit of future population health.

REFERENCES AND NOTES

1. T. Valerio, B. Knop, R. M. Kreider, W. He, in *Current Population Reports*, US Census Bureau, Ed. (Government Printing Office, 2021), pp. 70–173.
2. C. C. Brown, C. E. Adams, K. E. George, J. E. Moore, *Obstet. Gynecol.* **136**, 892–901 (2020).
3. N. M. Fisk, R. Atun, *PLOS Med.* **5**, e22 (2008).
4. L. C. Chappell, A. L. David, *PLOS Med.* **13**, e1002161 (2016).
5. D. G. Aaron, I. G. Cohen, E. Y. Adashi, *JAMA* **328**, 2394–2395 (2022).
6. S. J. Stock, J. E. Norman, *F1000Res.* **8**, F1000 Faculty Rev-911 (2019).
7. K. P. Conrad *et al*., *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **321**, R454–R468 (2021).
8. K. Maiti *et al*., *Am. J. Obstet. Gynecol.* **217**, 441.e1–441.e14 (2017).
9. O. Huhn *et al*., *Nat. Commun.* **11**, 381 (2020).
10. E. Abalos *et al*., *BJOG* **121**, 14–24 (2014).
11. S. E. Hiby *et al*., *J. Exp. Med.* **200**, 957–965 (2004).
12. P. Von Dadelszen, A. Syngelaki, R. Akolekar, L. A. Magee, K. H. Nicolaides, *BJOG* **130**, 524–530 (2023).
13. D. van de Vusse *et al*., *Eur. J. Clin. Pharmacol.* **78**, 1763–1776 (2022).
14. T. Liabsuetrakul *et al*., *Syst. Rev.* **11**, 135 (2022).
15. D. L. Rolnik *et al*., *N. Engl. J. Med.* **377**, 613–622 (2017).
16. S. W. Walsh, J. F. StraussIII, *Int. J. Mol. Sci.* **22**, 6985 (2021).
17. R. Smith, *N. Engl. J. Med.* **356**, 271–283 (2007).
18. M. Y. Turco *et al*., *Nature* **564**, 263–267 (2018).
19. H. Okae *et al*., *Cell Stem Cell* **22**, 50–63.e6 (2018).
20. C. Dong, T. W. Theunissen, *Methods Mol. Biol.* **2416**, 91–104 (2022).
21. J. L. Tarry-Adkins *et al*., *Front. Cell Dev. Biol.* **10**, 935403 (2022).
22. M. Y. Turco *et al*., *Nat. Cell Biol.* **19**, 568–577 (2017).
23. H. D. McIntyre *et al*., *J. Clin. Endocrinol. Metab.* **85**, 1143–1150 (2000).
24. W. Ye *et al*., *BMJ* **377**, e067946 (2022).
25. W. L. Lowe Jr. *et al*., *Diabetologia* **62**, 598–610 (2019).
26. H. Wang *et al*., *Diabetes Res. Clin. Pract.* **183**, 109050 (2022).
27. B. E. Metzger *et al*., *N. Engl. J. Med.* **358**, 1991–2002 (2008).
28. J. L. Tarry-Adkins, C. E. Aiken, S. E. Ozanne, *PLOS Med.* **17**, e1003126 (2020).
29. R. L. Goldenberg, J. F. Culhane, J. D. Iams, R. Romero, *Lancet* **371**, 75–84 (2008).
30. J. Muglu *et al*., *PLOS Med.* **16**, e1002838 (2019).
31. B. F. Mitchell, M. J. Taggart, *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **297**, R525–R545 (2009).
32. R. Romero, S. K. Dey, S. J. Fisher, *Science* **345**, 760–765 (2014).
33. M. S. Vidal Jr., R. C. V. Lintao, M. E. L. Severino, O. A. G. Tantengco, R. Menon, *Front. Endocrinol.* **13**, 1015622 (2022).
34. L. A. Stewart *et al*., *Lancet* **397**, 1183–1194 (2021).
35. Z. Alfirevic, T. Stampalija, N. Medley, *Cochrane Database Syst. Rev.* **6**, CD008991 (2017).
36. A. Wilson *et al*., *Cochrane Database Syst. Rev.* **8**, CD014978 (2022).
37. A. R. A. McDougall *et al*., *BJOG*1471-0528.17314 (2022).
38. N. Medley, J. P. Vogel, A. Care, Z. Alfirevic, *Cochrane Database Syst. Rev.* **11**, CD012505 (2018).
39. N. Andescavage, C. Limperopoulos, *Exp. Neurol.* **347**, 113868 (2022).
40. D. Ryu *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2100466118 (2021).
41. M. Rasmussen *et al*., *Nature* **601**, 422–427 (2022).
42. S. Chawanpaiboon *et al*., *Lancet Glob. Health* **7**, e37–e46 (2019).
43. H. Blencowe *et al*., *Lancet Glob. Health* **7**, e849–e860 (2019).
44. T. Firoz, C. L. Trigo Romero, C. Leung, J. P. Souza, Ö. Tunçalp, *BMJ Glob. Health* **7**, e007077 (2022).
45. G. Carroli, C. Cuesta, E. Abalos, A. M. Gulmezoglu, *Best Pract. Res. Clin. Obstet. Gynaecol.* **22**, 999–1012 (2008).
46. S. L. Woodd *et al*., *PLOS Med.* **16**, e1002984 (2019).
47. J. J. Kurinczuk, M. White-Koning, N. Badawi, *Early Hum. Dev.* **86**, 329–338 (2010).

ACKNOWLEDGMENTS

Competing interests: S.J.S. is the program director for the Wellcome Leap In Utero program. She reports research grants (paid to her institution) from the Wellcome Trust, the National Institute of Healthcare Research, Health Data Research UK, Tommy's Charity, and the Chief Scientist Office Scotland. She has received consultancy fees (paid to her institution) from Natera and honoraria (paid to her institution) from Hologic for scientific talks. C.E.A. has no conflicts of interest to declare.

License information: Copyright © 2023 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

Submitted 12 January 2023; accepted 24 February 2023
10.1126/science.adf9347

REVIEW

A brief history and future prospects of contraception

Deborah J. Anderson[1]* and Daniel S. Johnston[2,3,4]

Modern contraception ushered in an era of improved family planning, but more than 60 years after approval of "the pill," product gaps and unmet needs still exist. Nearly 250 million women worldwide who want to delay or avoid pregnancy do so ineffectively or not at all, and the principal mechanism of male contraception, condoms, has not changed in 100 years. As a result, about half of the pregnancies that occur globally each year are unintended. Increasing contraceptive options and uptake will curtail abortions, empower women and men, promote healthy families, and moderate population growth that overtaxes the environment. This Review addresses the history of contraception, shortcomings in contraceptive methods, promising approaches for male and female contraception, and simultaneous protection against unintended pregnancy and sexually transmitted infections.

Whereas fewer than 1 billion people inhabited Earth 200 years ago, the human population recently surpassed 8 billion and is on track to exceed 10 billion by 2100 (*1*). Population growth was fueled by the Industrial Revolution, which introduced more efficient agricultural practices and the refinement of fossil fuels and minerals, which led to global economic growth. Other major influences were advancements in sanitation practices, and new drugs and vaccines that reduced human disease. In 1800, the global child mortality rate was 43%, and the average human life span was 49 years, whereas in 2022, these figures were 3.8% and 73 years, respectively (*2*, *3*). The global fertility rate (the average number of children born per woman) decreased from 5.0 in 1960 to 2.5 today in part because of new contraception practices and burgeoning education and empowerment of women. Even so, the population continues to grow because the global fertility rate (2.5) exceeds the level of sustainable replacement (2.1), and because of "population momentum" (*1*, *4*).

The era of modern contraception arose amid the exponential population growth of the 20th century (Fig. 1). Contraception is defined as the act of intentionally preventing pregnancy by means of devices, practices, medications, or surgery. Modern contraceptives have a sound basis in reproductive biology and evidence of efficacy by using a precise protocol for correct use (*1*). Examples of modern options now available for women include tubal ligation, barrier methods, hormonal methods ["the pill", implants and injectables, progestin intrauterine devices (IUDs), or emergency contraception], and nonhormonal IUDs. Only two choices are available for men: vasectomy and condoms (Fig. 2).

Women now have the potential for reproductive autonomy, and most will use at least one contraception method during their lifetime. However, an estimated 170 million women worldwide who want to avoid pregnancy are not using any contraceptive method at all, and another 80 million women are not using safe and effective modern contraceptives (*5*). Consequently, of the 210 million pregnancies that occur globally each year, about half are unintended (*6*). Unintended pregnancies occur disproportionately in young and poor women and are associated with increased incidences of maternal and child mortality and cycles of intergenerational poverty. Worldwide, more than 60% of unintended pregnancies end in abortion (*7*).

The contraception gap is multifaceted. Many women have limited access to comprehensive sexual education and modern contracept… Of those with access, some express concerns over side effects or social stigma, whereas others experience opposition for social and cultural reasons (*1*). These hurdles contribute to inconsistent use and discontinuation. At the same time, many men report willingness to help close the gap and assume greater responsibility for family planning if provided with innovative male contraceptive methods (*8*).

The importance of contraception extends beyond preventing unintended pregnancies. Contraception plays an essential role in family planning and the ability of individuals and couples to influence the timing of initial and successive pregnancies. Child spacing has clear implications on the health of the child and the mother. Family planning also has economic benefits; it allows families to determine how many children they can support and affords women greater opportunity to pursue education and jobs. Arguably, contraceptive choices are as important during and after the childbearing stages as they are beforehand.

This Review identifies current problems and gaps in existing contraception options and details three focus areas for future generations of contraceptives: (i) more male methods; (ii) nonhormonal on-demand female contraceptives that have fewer side effects; and (iii) contraception that also protects against sexually transmissible infections (STIs), also called multipurpose prevention technology (MPT).

Contraception past and present

Most women are fertile for at least 30 years and are theoretically capable of giving birth annually during this phase of their life span. Men produce 70 million to 150 million sperm per day, can maintain their fertility well into their latter decades of life, and thus are capable

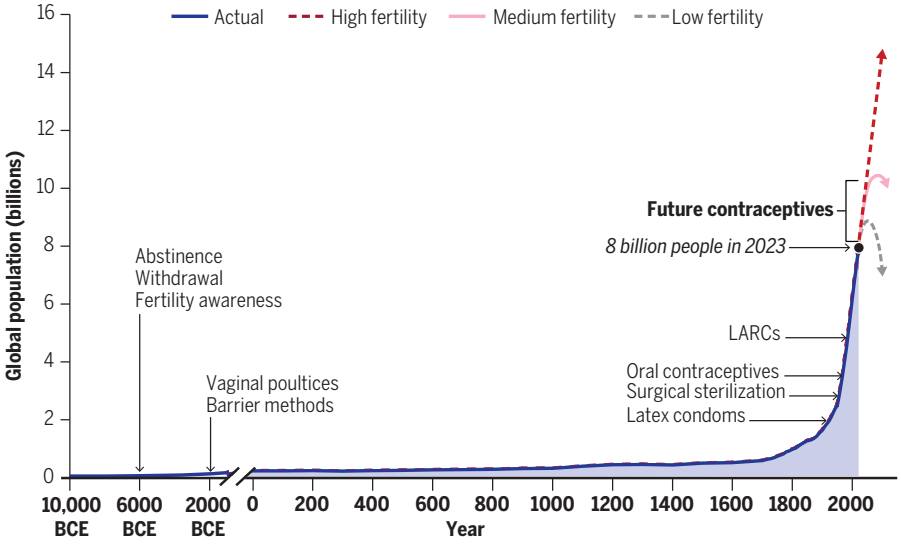


Fig. 1. Contraception and human population growth.

[1]Department of Medicine, Boston University Chobanian and Avedisian School of Medicine, Boston, MA 02118, USA.
[2]Contraception Research Branch, Eunice Kennedy Shriver National Institutes of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20817, USA.
[3]Fertility and Infertility Branch, Eunice Kennedy Shriver National Institutes of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20817, USA.
[4]Gynecologic Health and Disease Branch, Eunice Kennedy Shriver National Institutes of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20817, USA.
*Corresponding author. Email: deborah.anderson@bmc.org

of siring hundreds of offspring. Historical contraceptive options have included abstinence, fertility awareness (monitoring physical signs to estimate ovulation), and withdrawal. In present times, fertility awareness and withdrawal methods are practiced by about 8% of couples desiring contraception (*9*). These methods have annual failure rates of more than 20% (*1*).

Women have used vaginal poultices for contraception since at least 2000 BCE, using soft materials such as cotton or wool soaked in various mixtures of oils, fruit juice, herbs, and other substances to form a tampon-like plug. Today, user-applied vaginal methods are limited. Nonoxynol-9 (N-9)—a spermicidal surfactant formulated as a gel, cream, or film—has been available since the 1960s. Phexxi, an acidic vaginal gel that impairs sperm motility by buffering the vaginal pH after intercourse, was introduced in 2020. As with all user-controlled methods, efficacy depends on adherence. The annual failure rate of these vaginal contraceptive methods can be high (N-9, 18 to 28%; Phexxi, 14%) (*10*, *11*).

Male barrier methods (penile sheaths) were used in ancient Egypt and Rome, and a female condom was described in the Greek literature from 3000 BCE (*12*). Early condoms were made from animal bladders and intestines, or from chemically treated linen (*13*). During the 1700s, condoms were widely used by men in Europe to prevent sexually transmitted infections. In the 1920s, synthetic latex was invented and led to the manufacture of thin male condoms with high flexibility and increased tensile strength. Synthetic latex condoms remain the predominant male barrier method used today. Modern female condoms are made of latex or nitrile. Condoms offer advantages: they are relatively inexpensive, widely available directly to the end user, provide effective contraception, and protect against most STIs. Disadvantages are the time needed to apply a condom (interrupts intimacy) and decreased sexual sensation. Under perfect use, condoms are 98% effective, but in actual use, they are 87% effective. Diaphragms and cervical caps, two female barrier methods, do not protect users from most STIs and have typical failure rates of 18% (*1*). Globally, nearly one-quarter of contracepting couples use barrier methods as their primary means of contraception: 21% male condoms, 2% diaphragms and cervical caps, and 1% female condoms (*9*).

Vasectomy is a surgical procedure that severs and seals the vas deferens, the duct that transports sperm during ejaculation. Female tubal ligation is a surgical procedure that prevents fertilization by cutting and cauterizing the fallopian tubes, which connect the ovaries to the uterus. Both vasectomy and tubal ligation procedures are safe and highly (>99%) effective, but there is a large gender disparity in their uptake. Worldwide, 26.6% of contracepting women have had tubal ligation surgery, whereas only 2.8% of men are vasectomized (*5*). These procedures are referred to as sterilization, although this designation is somewhat misleading. Men and women continue to produce gametes, and both procedures can be successfully reversed with surgery (vasovasostomy or tubal reanastomosis), with fertility restored in 40 to 70% of cases (*14*, *15*). Alternatively, many vasectomized men and ligated women can achieve pregnancy through in vitro fertilization technologies, with fertility rates ranging from 30 to 70% (*16*, *17*). However, these methods are expensive and inaccessible to many individuals. Nonsurgical methods to block the vas deferens and readily reversible vasectomies are currently being explored with the hope that they will increase the popularity of vasectomies (*18*).

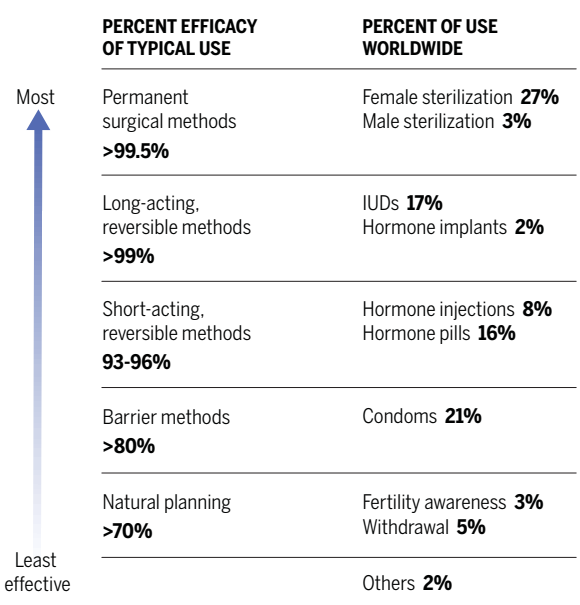


Fig. 2. Modern contraception methods: Types, usage, and efficacy. See (*5*, *9*).

IUDs were invented in the early 1900s but were not widely used for contraception in the United States until the late 1950s; by 1974, there were more than 2.5 million users. Unfortunately, one of the IUDs marketed at that time, the Dalkon Shield, was associated with pelvic infections and mortality because of a multifilamentous tail string that allowed bacteria to ascend from the vagina into the uterus. It was removed from the market in 1974, but the damage it caused set the IUD field back several years (*19*). Two classes of IUDs are currently widely available: the nonhormonal copper IUD (Paragard) and hormonal IUDs (Mirena and others). These newer IUDs are safe, are 99% effective, and produce long-acting reversible contraception (LARC) for 8 to 10 years. IUDs primarily work by preventing fertilization. The copper IUD emits copper ions that are toxic to gametes, and hormonal IUDs block ovulation and thicken cervical mucus, preventing sperm from reaching the fallopian tubes. The effectiveness of IUDs as emergency contraceptives implies that they may also act by preventing implantation of the blastocyst (*20*). The copper IUD causes heavier menstrual bleeding, which is a serious side effect for some women; hormonal IUDs generally reduce menstrual bleeding, which is often considered an advantage. IUDs earn the highest rate of satisfaction and continuation of all reversible contraceptives: 17% of women using contraception today use IUDs (*9*).

The first oral hormonal contraceptive pill, Enovid, was approved by the US Food and Drug Administration (FDA) for contraception in 1960. Other oral contraceptives soon followed, and by 1963, 2.3 million women were using some form of oral birth control (*21*). Serious adverse effects associated with these early pills were soon reported. Epidemiological studies concluded that pill users were more susceptible to thromboembolism than were non-pill users (*22*). Estrogens, and ethinyl estradiol in particular, were implicated. Over time, both estrogen and progestin components of combination pills were markedly reduced, and safety increased.

In combined hormonal contraceptives, both progestins and estrogens inhibit the hypothalamic-pituitary-ovarian axis, which controls the reproductive cycle. Progestins prevent pregnancy by inhibiting the luteinizing hormone (LH) surge, thus suppressing ovulation, thickening the cervical mucus, and lowering fallopian tube motility. Estrogens prevent pregnancy by suppressing follicle-stimulating hormone (FSH) production, which prevents egg maturation. The addition of estrogen also helps prevent irregular bleeding (*23*).

Collectively, about 30% of women worldwide currently use contraceptives with hormonal mechanisms of action. The estrogen and/or progestin-only birth control pills, when taken as directed, are >98% effective (*24*). Furthermore, substantial modifications have allowed for reduced dosages and expanded routes of administration (implants, injectables, intrauterine devices, vaginal rings, and dermal patches), which has increased safety and product choices. However, the estrogens and progestogens in female hormonal methods are still associated with side effects and potential health risks. Estrogens are contraindicated in

ILLUSTRATION: A. MASTIN/*SCIENCE*

women at risk for thrombosis (from smoking, hypertension, obesity, or history of thrombosis) and after hormone-dependent cancer. Progestins are associated with abnormal bleeding patterns, headaches, mood swings, and other side effects that can cause end-user dissatisfaction and discontinuation (*25*). Such risks and drawbacks are reasons for developing additional methods of contraception.

Emerging contraceptives

Male contraceptives

Hormonal

Reversible oligospermia (<15 million sperm/ml of ejaculate) was achieved by the exogenous administration of androgen in 1939 (*26*), but it was a 1972 publication (*27*) that ignited interest that continues to the present day (*28*–*30*). "Hormonal male contraception" functions by suppressing FSH and LH secretion from the pituitary gland. LH stimulates biosynthesis of testosterone by testicular Leydig cells, which is critical to sustaining spermatogenesis. Suppression of LH results in decreased intratesticular testosterone, spermatogenic failure, and reduced ejaculatory sperm concentrations. Concentrations of <1 million sperm/ml are often achieved, a level deemed compatible with effective contraception. Cessation of male hormonal treatments leads to a return of normal LH and FSH secretion, normal intratesticular testosterone concentrations, and a resumption of spermatogenesis.

Administration of estrogen or progestin along with androgen provides markedly greater suppression of the hypothalamic-pituitary-gonadal (HPG) axis as compared with use of androgen alone (*31*). Gonadotropin-releasing hormone (GnRH) agonists and antagonists also suppress FSH and LH secretion. To date, many androgenic molecules have been clinically evaluated alone or in combination with estrogens, progestins, GnRH agonists, or GnRH antagonists in a variety of delivery systems, including oral pills, intramuscular injections, subdermal injections, implants, and transdermal gels (*28*, *32*–*34*).

Nestorone + testosterone (NES+TES) combination gel is a promising product that functions through the HPG axis and has advanced into clinical trials (ClinicalTrials.gov identifier: NCT03452111). NES+TES is a single transdermal gel product that combines nestorone, a progestin, and testosterone and is applied daily by men to the shoulder and upper arm. In a previous trial, NES+TES formulations reduced sperm counts to <1 million sperm/ml in 88 to 89% of subjects, whereas only 23% of men in the testosterone-only group had reductions to that level. No serious adverse effects were reported. The most common side effect in the NES+TES groups was acne (about 25% of subjects). Secondary transfer of NES+TES gel to others is possible but was minimized by washing and covering the site after application (*35*).

Dimethandrolone undecanoate (DMAU) is a prodrug that is converted into dimethandrolone, which has both androgenic and progestational activity (*36*). A phase I clinical trial is enrolling participants to evaluate single injections of either intramuscular or subcutaneous DMAU in castor oil–benzyl benzoate over a 5- to 7-month period (ClinicalTrials.gov identifier: NCT02927210).

Nonhormonal molecularly targeted male contraceptives

The past 25 years have been an active period for the development of small-molecule male contraceptive methods that do not function through the HPG axis. This approach identifies molecules that modulate a specific molecular target such as a kinase required for the formation and/or maturation of functional sperm. For the development of a contraceptive that must have minimal side effects, selecting targets exclusively or predominantly expressed in the reproductive system will minimize potential systemic toxicity.

A leading example of a nonhormonal molecularly targeted contraceptive is soluble adenyl cyclase (sAC; ADCY10). Targeted deletions of ADCY10 in mice result in male infertility (*37*), and two infertile men were identified that have identical frameshift mutations that disrupt sAC's active site (*38*). A recent study demonstrated inhibition of sperm motility and short-term infertility in mice after intraperitoneal administration of a sAC inhibitor. Mating behavior was not affected, and fertility was restored within 24 hours, demonstrating reversible on-demand male contraceptive activity (*39*). Additional research to improve drug-like properties, such as oral bioavailability and compound residence time, may extend the duration of effectiveness and may lead to a clinical trial for this potential on-demand male contraceptive.

Innovative nonhormonal contraceptives for women

Ovaprene is an investigational hormone-free, monthly contraceptive intravaginal ring being developed by Daré Bioscience. Sperm transport through the cervix is prevented by the release of ferrous gluconate (a spermiostatic agent) and a knitted polymer barrier that spans the ring (*40*). Early clinical trials indicate that the device is safe and highly effective at lowering numbers of progressively motile sperm in the post-coital test. The product is intended to be used continuously for up to one menstrual cycle. A pivotal clinical trial is expected to begin in 2023.

Specific molecular targets from the male reproductive tract (testis, epididymis, and prostate) are being developed as user-controlled

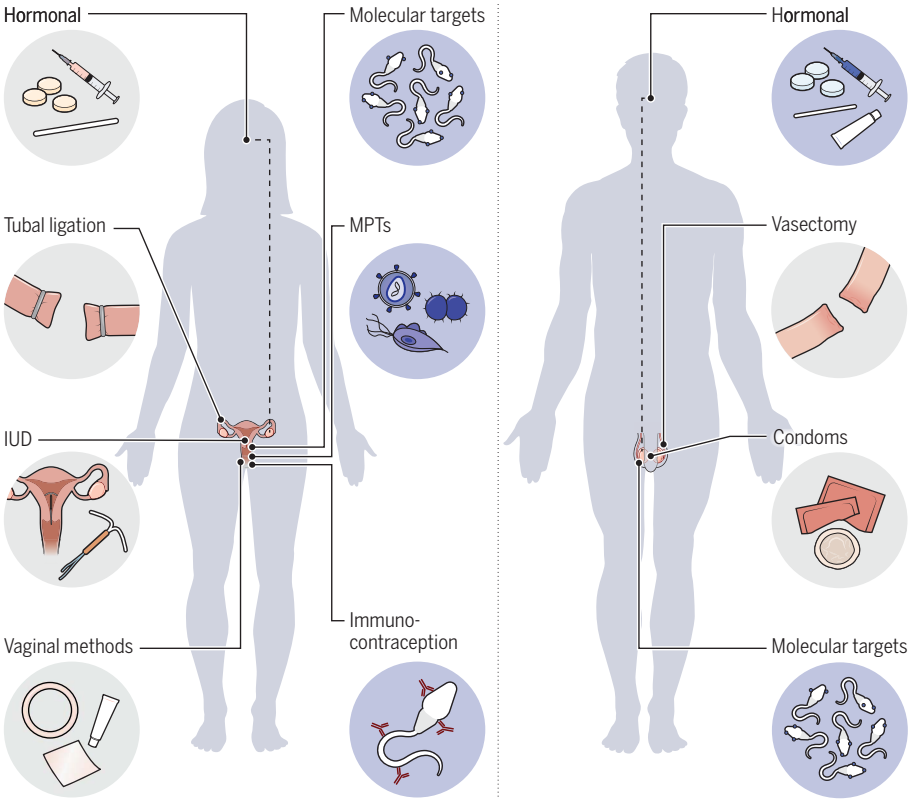


ILLUSTRATION: A. MASTIN/*SCIENCE*

Fig. 3. Current and future contraceptive methods. Current contraceptive methods are indicated with gray backgrounds; new contraceptive methods have blue backgrounds.

vaginal contraceptive products. Inhibitory molecules that target proteins required for sperm function "lie in wait" in the vagina and prevent the movement of sperm into the upper regions of the female reproductive tract, where fertilization occurs.

Immunocontraception

The field of immunocontraception began in the 1900s with discoveries that sperm can elicit immune responses in animals and humans, leading to infertility, and that many infertile men and women have antisperm antibodies (*41*). Several international agencies started contraceptive vaccine programs directed at inhibiting sperm, oocytes, placenta, and endocrine processes. In the 1980s and 90s, clinical trials were conducted to test the safety and efficacy of a contraceptive vaccine that targets human chorionic gonadotropin (hCG), a hormone secreted by the blastocyst early in pregnancy to ensure uterine receptivity. The vaccine trials demonstrated safety, but antibody levels were highly variable, indicating that some women would not be protected from pregnancy (*42*).

Today, the immunocontraception approach is undergoing a renaissance owing to breakthroughs in bioengineering and the ability to produce clinical-grade human monoclonal antibodies. The use of defined manufactured monoclonal antibodies for passive immunization bypasses the concerns raised with contraceptive vaccines based on active immunization.

One promising example is an antisperm monoclonal antibody, the "human contraceptive antibody" (HCA). This antibody recognizes a sperm surface carbohydrate epitope, CD52 g, and rapidly agglutinates and immobilizes sperm (*43*). Formulated into a vaginal film, it was recently tested in women in a phase I clinical trial. The antibody film was safe and effectively eliminated progressively motile sperm in cervical mucus in the post-coital test (*44*). Second-generation multivalent HCA constructs and alternate antibody delivery systems such as intravaginal rings and topical mRNA application are being explored (*45*).

Multipurpose prevention technologies

MPTs formally emerged from the microbicide field in 2009 and combine protection against multiple reproductive health risks such as unintended pregnancy, HIV, herpes simplex virus (HSV), and other pathogenic STIs. More than 350 million treatable STIs and 1.5 million new HIV cases are reported each year globally, and 500 million adults are infected with and can potentially transmit HSV (*46*). For this reason, many sexually active women and men would prefer products that are not only contraceptive but also protective against STIs. Condoms are the only MPT products currently available.

N-9, a vaginal spermicide introduced in the 1960s for contraception, showed potent efficacy against HIV and other STIs in preclinical research (*47*) and was tested in a series of clinical trials for protection against HIV-1 transmission. Unfortunately, vaginal administration of N-9 was not only ineffective against HIV (*48*) but enhanced transmission in high-risk populations because of its inflammatory effects (*49*). Five other nonspecific microbicides also failed HIV clinical trials (*50*).

A newer class of MPTs combines potent antiretroviral drugs (ARVs) with previously approved contraceptive hormones. Levonorgestrel was added to a monthly dapivirine intravaginal ring developed to prevent HIV sexual transmission, and this contraceptive-ARV ring demonstrated safety and effective drug-release profiles in phase I clinical trials. However, the dapivirine ring failed to gain FDA approval for HIV prevention because of poor efficacy (29%) in phase III clinical trials (*51*, *52*). This may be revisited because the World Health Organization (WHO) recently approved the dapivirine ring for HIV prevention (*53*). Three other potent ARVs—tenofovir alafenamide, islatravir, and dolutegravir—have been formulated into long-acting implants that release effective drug concentrations for 6 months to 1 year (*54*); plans are underway to add a contraceptive hormone such as levonorgestrel to these implants to develop long-acting contraceptive-MPTs. Oral and injectable anti-HIV–contraception combinations are also being explored. Long-acting cabotegravir-rilpivirine injections were recently approved for HIV prevention and could be combined with injectable contraceptives such as DepoProvera or Net-EN.

Several programs are also pursuing combinations of anti-STI and anti-sperm antibodies as MPTs. A vaginal film containing anti-HIV-1 and anti-HSV monoclonal antibodies was shown to be safe and effective (ex vivo) in a phase I clinical trial (*55*); plans are underway to combine these antibodies with the antisperm antibody HCA to produce an on-demand MPT vaginal film or an intravaginal ring (*56*). HCA alone may also have a future as a MPT because the sperm antigen recognized by HCA, a glycoprotein produced in the epididymis, inserts into the plasma membrane of sperm and other cells in semen by means of a glycosylphosphatidylinositol (GPI) anchor. HCA has been shown to bind to sperm, HIV-infected leukocytes, and certain pathogens such as Trichomonas vaginalis and cell-free HIV in semen and may clear infectious cells and organisms by trapping them in sperm agglutinates or mucus (*57*).

Industry is critical to bring products to market

The primary funders for contraception development are the National Institutes of Health, the US Agency for International Development, and the Bill and Melinda Gates Foundation. These entities are not equipped to bring products to market. Commitment and action are needed from the pharmaceutical industry, which exited from preclinical contraceptive development around 2007 after decades of activity and innovation. However, the past 5 years have seen renewed interest by pharma in emerging contraceptives. Bayer and Organon have entered into agreements with small biotech companies focused on preclinical-stage and early clinical-stage nonhormonal female contraception programs. Hopefully, this is the start of sustained recommitment to the field.

Conclusions and outlook

The long history of contraception is interesting and varied, from futile attempts to magnificent breakthroughs. A notable advance was achieved when hormonally acting female contraceptives were approved more than 60 years ago. Despite this accomplishment, valid concerns persist about the safety and side effects associated with hormonal methods. Although efficacious alternatives exist, these methods do not meet the full range of needs of men and women throughout their reproductive life spans. A contraception "gap" persists, which modern-day research and industry can bridge.

Exciting programs have emerged that demonstrate proof of concept for hormonal and on-demand male contraception. Targets in the male reproductive tract are being pursued as female contraceptives. Moreover, a newer class of contraceptives designated as MPTs are under development to provide both contraception and protection against STIs (Fig. 3).

This renaissance in contraception research has been made possible by renewed commitment from scientists, private foundations, and government institutions. These promising developments await industry involvement to usher them to market and into the hands of people around the world who have long waited for more tools to accomplish their personal goals.

REFERENCES AND NOTES

1. I. McFarlane, Ed., *State of the World Population 2022* (United Nations Population Fund, 2022).
2. United Nations, "UN Inter-agency Group for Child Mortality Estimation" (United Nations, 2022).
3. WHO, "The Global Health Observatory (Life expectancy at birth)" (WHO, 2023).
4. D. J. Anderson, *N. Engl. J. Med.* **381**, 397–399 (2019).
5. A. Haakenstad *et al*., *Lancet* **400**, 295–327 (2022).
6. J. Bearak *et al*., *Lancet Glob. Health* **8**, e1152–e1161 (2020).
7. WHO, "World Health Organization-Adolescent Pregnancy" (WHO, 2022).
8. MCI, "Potential 17 million men interested in male contraception, consumer survey finds" (MCI, 2019).
9. *Contraceptive Use by Method 2019* (United Nations Population Division, 2019).
10. M. A. Thomas *et al*., *Contracept. X* **2**, 100031 (2020).
11. J. Trussell, *Contraception* **83**, 397–404 (2011).
12. B. E. Finch, H. Green, in *Contraception Through the Ages*, C. Charles, Ed., (Thomas, 1964).
13. F. Khan, S. Mukhtar, I. K. Dickinson, S. Sriprasad, *Indian J. Urol.* **29**, 12–15 (2013).
14. X. Deffieux *et al*., *Arch. Gynecol. Obstet.* **283**, 1149–1158 (2011).
15. M. Duijn, J. A. van der Zee, Y. Bachour, *SN Compr. Clin. Med.* **3**, 2193–2203 (2021).

16. E. Malacova, A. Kemp, R. Hart, K. Jama-Alol, D. B. Preen, *Contraception* **91**, 240–244 (2015).
17. L. Witherspoon, R. Flannigan, *BC Med. J.* **63**, 62–66 (2021).
18. A. Colagross-Schouten, M. J. Lemoy, R. I. Keesler, E. Lissner, C. A. VandeVoort, *Basic Clin. Androl.* **27**, 4 (2017).
19. D. Hubacher, M. Kavanaugh, *Contraception* **98**, 467–470 (2018).
20. D. K. Turok *et al*., *N. Engl. J. Med.* **384**, 335–344 (2021).
21. B. Asbell, *The Pill: A Biography of the Drug that Changed the World* (Random House, 1995).
22. Risk of thromboembolic disease in women taking oral contraceptives. A preliminary communication to the Medical Research Council by a Subcommittee. *BMJ* **2**, 355–359 (1967).
23. L. E. Britton, A. Alspaugh, M. Z. Greene, M. R. McLemore, *Am. J. Nurs.* **120**, 22–33 (2020).
24. S. Christin-Maitre, *Best Pract. Res. Clin. Endocrinol. Metab.* **27**, 3–12 (2013).
25. C. L. Westhoff *et al*., *Am. J. Obstet. Gynecol.* **196**, 412.e1–412.e6, discussion 412.e6–412.e7 (2007).
26. N. J. Heckel, *Exp. Biol. Med.* **40**, 658–659 (1939).
27. P. R. Reddy, J. M. Rao, *Contraception* **5**, 295–301 (1972).
28. P. Kogan, M. Wald, *Urol. Clin. North Am.* **41**, 145–161 (2014).
29. A. Thirumalai, S. T. Page, *Annu. Rev. Med.* **71**, 17–31 (2020).
30. G. M. Waites, *Fertil. Steril.* **80**, 1–15 (2003).
31. L. L. Ewing, C. Desjardins, D. C. Irby, B. Robaire, *Nature* **269**, 409–411 (1977).
32. J. H. Chao, S. T. Page, *Pharmacol. Ther.* **163**, 109–117 (2016).
33. K. Piotrowska, C. Wang, R. S. Swerdloff, P. Y. Liu, *Lancet Diabetes Endocrinol.* **5**, 214–223 (2017).
34. C. Wang, R. S. Swerdloff, *Curr. Opin. Urol.* **20**, 520–524 (2010).
35. F. Yuen *et al*., *Andrology* **7**, 235–243 (2019).
36. B. J. Attardi, S. A. Hild, J. R. Reel, *Endocrinology* **147**, 3016–3026 (2006).
37. G. Esposito *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **101**, 2993–2998 (2004).
38. A. Akbari *et al*., *Hum. Reprod.* **34**, 1155–1164 (2019).
39. M. Balbach *et al*., *Nat. Comm.* **14**, 637 (2023).
40. DareBioscience, "An investigational hormone-free monthly intravaginal contraceptive" (DareBioscience, 2023).
41. J. Frayne, L. Hall, *J. Reprod. Immunol.* **43**, 1–33 (1999).
42. W. R. Jones *et al*., *Lancet* **1**, 1295–1298 (1988).
43. G. Baldeon-Vaca *et al*., *EBioMedicine* **69**, 103478 (2021).
44. A. R. Thurman *et al*., *Am. J. Obstet. Gynecol.* 10.1016/j.ajog.2023.02.024 (2023).
45. D. J. Anderson *et al*., *Biol. Reprod.* **103**, 275–285 (2020).
46. WHO, "STIs in 2022: Emerging and re-emerging outbreaks" (WHO, 2022).
47. P. J. Feldblum, J. A. Fortney, *Am. J. Public Health* **78**, 52–54 (1988).
48. D. Wilkinson, M. Tholandi, G. Ramjee, G. W. Rutherford, *Lancet Infect. Dis.* **2**, 613–617 (2002).
49. L. Van Damme *et al*., *Lancet* **360**, 971–977 (2002).
50. F. Notario-Pérez, R. Ruiz-Caro, M. D. Veiga-Ochoa, *Drug Des. Devel. Ther.* **11**, 1767–1787 (2017).
51. J. M. Baeten *et al*., *N. Engl. J. Med.* **375**, 2121–2132 (2016).
52. A. Nel *et al*., *PLOS ONE* **11**, e0147743 (2016).
53. E. R. Brown *et al*., *J. Int. AIDS Soc.* **23**, e25634 (2020).
54. P. Thoueille, E. Choong, M. Cavassini, T. Buclin, L. A. Decosterd, *J. Antimicrob. Chemother.* **77**, 290–302 (2022).
55. J. A. Politch *et al*., *PLOS Med.* **18**, e1003495 (2021).
56. CAMIHealth, MPT Product Development Database (2023).
57. G. Baldeon-Vaca *et al*., paper presented at the NICHD Contraceptive Research Centers Annual Meeting, Houston, 2019.

ACKNOWLEDGMENTS

We thank our colleagues—in particular, G. Kopf, W. Wright, K. Whaley, and K. Mayer—for helpful input in early drafts of this manuscript; the anonymous referees for their constructive comments; and J. Politch, E. Mausser, and E. Nador for assistance with the preparation of the manuscript and figures. **Funding:** D.J.A. receives support from P50 HD096957, a contraceptive research center grant from the Eunice Kennedy Shriver National Institutes of Child Health and Human Development (NICHD), National Institutes of Health (NIH). D.S.J. is an employee of NIH. Comments and views of the authors do not necessarily represent the views of the NICHD, NIH, or US federal government. **License information:** Copyright © 2023 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

Submitted 23 January 2023; accepted 1 March 2023
10.1126/science.adf9341

REVIEW

Understanding the genetics of human infertility

Qing Sang[1], Pierre F. Ray[2], Lei Wang[1]*

Reproduction involves a wide range of biological processes, including organ formation and development, neuroendocrine regulation, hormone production, and meiosis and mitosis. Infertility, the failure of reproduction, has become a major issue for human reproductive health and affects up to one in seven couples worldwide. Here, we review various aspects of human infertility, including etiology, mechanisms, and treatments, with a particular emphasis on genetics. We focus on gamete production and gamete quality, which is the core of successful reproduction. We also discuss future research opportunities and challenges to further expand our understanding of human infertility and improve patient care by providing precision diagnosis and personalized treatments.

Infertility is a major health issue, estimated to affect ~15% of the global population (*1*). Human reproduction requires normal developmental processes consisting of the production of competent gametes, fertilization, preimplantation embryonic development, implantation, pregnancy, and birth. Several biological processes, including organ formation and development, neuroendocrine regulation, hormone production, and meiosis and mitosis contribute to the high frequency and complex etiology of human infertility. Our understanding of the etiology of infertility has vastly improved in recent years, which will help to elucidate the pathophysiology of aberrant human reproduction and improve precision diagnosis and personalized treatment.

In this Review, we highlight recent insights into different factors impairing reproduction, with a particular emphasis on the genetic factors, and discuss the corresponding potential treatment strategies. Because the quality of gametes is at the core of successful reproduction, we mainly focuses on two interconnected aspects: factors that affect gonadal function in producing gametes and factors that affect gamete competence. In addition, we introduce potential treatment strategies for improving gamete competence. Although other causes of infertility will be introduced, this Review mainly emphasizes advances in well-established nonsyndrome phenotypes with monogenic inheritance.

Hypothalamic–pituitary–gonadal axis disturbance and gonad dysfunction

Under physiological conditions, the precise orchestration of hypothalamic kisspeptin neurons and gonadotropin-releasing hormone (GnRH) neurons induces the secretion of follicle-stimulating hormone (FSH) and luteinizing hormone (LH) from the anterior pituitary, which in turn initiates gonadal steroid production that regulates testicular development, follicular development, spermatogenesis, and oogenesis (*2*). Hormones produced by the gonads can then exert feedback effects on the hypothalamus and/or pituitary to adjust the concentrations of FSH and LH in a timely manner (Fig. 1).

Defects at one or more levels of the hypothalamic–pituitary–gonadal (HPG) axis can cause hypogonadism, which manifests as a reduction in sex hormones, absent or delayed puberty, and abnormal spermatogenesis or ovulation. Hypogonadotropic hypogonadism refers to the condition of insufficient GnRH and/or gonadotropin release caused by dysfunction of the pituitary gland or the hypothalamus. Congenital hypogonadotropic hypogonadism (CHH) has a male predominance (male:female ratio, 3 to 5:1) and is called Kallmann syndrome (KS) when combined with anosmia (in 50% of cases), which results from deficient migration of GnRH neurons from the olfactory placode to the forebrain (*3*). More than 40 genes have been associated with the pathogenesis of the isolated form, the syndromic form, or both forms of CHH (*4*, *5*). For example, mutant *FGF8* and *CHD7* affect the neurogenic niche of the nasal area, whereas mutant *KAL1* (the first causal gene for KS, also known as *ANOS1*) and *PROK2* disrupt GnRH neuron migration (*3*). Mutations in *GNRH1*, *KISS1*, and *KISS1R* influence GnRH neuron homeostasis, and *GNRHR* leads to a gonadotrope defect (*4*). Several other mutant genes that cause CHH were identified recently, including *NDNF*, which alters GnRH neuron migration and olfactory axonal projections (*6*); *PRDM13*, which disrupts kisspeptin neuron development (*7*); and *NOS1*, which leads to GnRH deficiency (*8*). Nitric oxide treatment can reverse the phenotype of *Nos1* mutant mice, suggesting a potential therapy (*8*). In addition to point mutations in these genes, copy number variants and the long noncoding RNA RMST may also contribute to CHH (*9*, *10*).

Hypergonadotropic hypogonadism usually results from gonadal failure and the subsequent loss of negative feedback from sex hormones. In males, Klinefelter syndrome (also known as 47,XXY) is the most common cause of hypergonadotropic hypogonadism and azoospermia, affecting one in 500 to 1000 males (*11*). Klinefelter syndrome was believed to be caused by rare, random nondisjunction events occurring during meiosis with largely unknown molecular mechanisms. A recent study demonstrated that mutations in *USP26* could interfere with the XY body (a specialized meiotic chromatin domain), resulting in the production of a high proportion of XY aneuploid spermatozoa and offspring with Klinefelter syndrome (*12*). Turner syndrome (also known as 45,X), affecting 1/2000 to 1/2500 female newborns, is characterized by hypergonadotropic hypogonadism and a wide range of clinical symptoms, including infertility (*13*). The missing X chromosome induces haploinsufficiency in genes normally escaping X inactivation and epigenetic changes in others, leading to the vast array of symptoms observed in patients (*13*). In addition, disorders of sex development include gonadal dysgenesis, sex reversal, and infertility, which is mainly associated with mutations in the genes regulating sex determination and differentiation, including *SRY*, *NR5A1*, *WT1*, and *SOX9* (*14*). Digenic or oligogenic effects have also be proposed for phenotypic variations of disorders of sex development (*15*).

Premature ovarian insufficiency (POI), defined by deficient ovarian function before the age of 40 years, is also usually present with hypergonadotropic hypogonadism. Chromosome abnormalities account for ~10 to 13% of POI cases, whereas a premutation in the well-studied *FMR1* gene is among the most common genetic cause of 46,XX POI (*16*). To date, >60 genes have been implicated in the development of POI with or without syndromic traits (*5*, *17*). Most genes are involved in homologous recombination and DNA damage repair (e.g., *STAG3*, *MCM8*, *MCM9*, *RAD51*, and *BRCA1/2*), follicle activation (e.g., *NOBOX* and *SOHLH1*), or follicle development and maturation (e.g., *GDF9*, *BMP5*, *FSHR*, and *FOXL2*) (*16*, *18*). Recent cohort analyses identified 29 POI-associated genes and several pathways, such as nuclear factor-κB, post-translational regulation, and mitophagy (*19*, *20*), expanding the genetic understanding of POI.

Polycystic ovary syndrome (PCOS), another kind of ovarian dysfunction, is the most common cause of anovulation infertility. The complex etiology of PCOS likely involves both genetic and environmental factors. Although no monogenic Mendelian transmission has been found in PCOS, genetic association studies suggest that common genetic variants account for ~10% of its heritability (*21*, *22*). A growing body of evidence suggests that the daughters of women with PCOS who have an excess of androgen and/or anti-Müllerian hormone during pregnancy present with a higher risk of developing PCOS themselves (*23*, *24*). PCOS inheritance thus likely occurs through the alteration of the genetic or epigenetic landscape (*25*), but whether these

[1]Institute of Pediatrics, Children's Hospital of Fudan University and Institutes of Biomedical Sciences, the State Key Laboratory of Genetic Engineering, Shanghai Key Laboratory of Medical Epigenetics, Fudan University, Shanghai 200032, China. [2]Université Grenoble Alpes, INSERM U1209, CNRS UMR 5309, Team Genetics Epigenetics and Therapies of Infertility, Institute for Advanced Biosciences, 380000 Grenoble, France.
*Corresponding author. Email: wangleiwanglei@fudan.edu.cn

changes are established in somatic or germ cells remains to be investigated. Recent findings also suggest that *Bacteroides vulgatus* in the gut microbiota could induce PCOS by altering the bile acid–interleukin-22 (IL-22) axis, suggesting an important role of gut microbes in fertility (*26*).

Developmentally and functionally incompetent gametes

Sperm defects

The testis is a highly specialized and efficient organ that normally assembles dozens of millions of spermatozoa every day. Overall, it takes 74 days to obtain the final product after going through four different production phases, including mitotic replication of germ cells, meiosis I and II, and spermiogenesis (Fig. 2).

Congenital cryptorchidism (undescended testes), one of most common congenital malformations in boys, is associated with hypogonadism, poor semen quality, and an increased risk of testicular germ cell tumors. Although diseases and conditions affecting testicular descent are associated with cryptorchidism, the specific underlying causes remain to be elucidated (*27*).

One type of deficit in sperm production relates to the quantity of sperm produced and includes patients with low sperm number (oligozoospermia) or no sperm present in the ejaculate (azoospermia) (Fig. 2). Anomalies in the cystic fibrosis transmembrane conductance regulator (CFTR) protein induce a thickening of the extracellular mucus that leads to clogging of the vas deferens and obstructive azoospermia (OA) (*28*). Genetic defects in the *CFTR* gene are found in >80% of OA patients, and a second gene, *ADGRG2,* encoding an epididymis-specific transmembrane protein, was recently found to be mutated in a few subjects (*29*). Klinefelter syndrome and microdeletions of the Y chromosome are the most frequent causes of non-obstructive azoospermia (NOA), accounting for 15 and 6.4% of patients with azoospermia, respectively, whereas chromosomal translocations are also correlated with reduced sperm count (*30*). Karyotype analysis is therefore recommended for all couples undergoing intracytoplasmic sperm injection (ICSI) (*31*). Whole-exome sequencing coupled with functional studies has recently permitted the identification of numerous genes responsible for NOA, including synaptonemal complex component–encoding genes (e.g., *SYCE1*, *C14ORF39*, and *STAG3*), meiotic recombination–associated genes (e.g., *TEX11* and *MSH4*), and DNA repair genes (e.g., *MEIOB* and *SPATA22*) (*32*). In addition, mutations in *PNLDC1* and *HIWI* have been shown to be responsible for azoospermia by affecting Piwi-interacting RNA (piRNA) processing. These findings further support the emerging role of piRNA in male infertility (*33*, *34*).

Patients with abnormally formed spermatozoa (teratozoospermia) have been extensively studied, and many morphological anomalies of the sperm have been correlated with specific gene defects (Fig. 2). *AURKC* was the first mutant gene to be formally associated with a well-defined sperm phenotype, macrozoospermia, which is defined by the presence of spermatozoa with a very enlarged head and usually four flagella. All spermatozoa were shown to be blocked in meiosis I and to be tetraploid (*35*). Deficiency in DPY19L2, an anchor between the acrosome and the sperm nucleus, is the main cause for globozoospermia, which is characterized by round spermatozoa devoid of acrosomes (*36*). Gene defects leading to the presence of headless or acephalic spermatozoa were also identified and mainly attributed to *SUN5* and *PMFBP1*, which play a role in connecting the head to the flagella (*37*, *38*).

The other main deficit in sperm concerns its motility (asthenozoospermia) (Fig. 2). Abnormalities in cation channel proteins that localize to the flagellum of spermatozoa (*CATSPER1-4*) were found to induce asthenozoospermia (*39*). Other genes also coding for channel proteins, such as *SLC26A8* (*40*), *SLC9C1* (*41*), and adenylyl cyclase *ADCY10* (*42*), were shown to be associated with impaired motility with no obvious morphological defect of the sperm. In recent years, many genes have been associated with asthenozoospermia caused by multiple morphological anomalies of the flagella (the MMAF phenotype), including short, coiled, or absent flagella. Flagella and motile cilia are evolutionary conserved organelles constructed around the axoneme, a structure formed mainly by nine peripheral microtubule doublets connected by outer and inner dynein molecules surrounding two central microtubule singlets. *DNAH1*, encoding an axonemal dynein of the inner dynein arms, was the first gene to be formally correlated with the MMAF (*43*). Since then, mutations in >20 genes have been reported to induce MMAF, including genes encoding the tether–tether head complex (e.g., *CFAP43* and *CFAP44*), the radial spoke complex (e.g., *WDR66*), the intraflagellar transport complex (e.g., *TTC29* and *CFAP69*), and centrosomal protein (e.g., *CEP135*) (*44*, *45*).

In addition, some rare infertile men have no apparent defects in sperm morphology or number, but still show fertilization failure even when ICSI is performed (Fig. 2). One mechanism that is well studied in these cases relates to defects in PLCZ1, a sperm-borne protein that triggers calcium release from the oocyte endoplasmic reticulum, followed by a cascade of molecular events referred to as oocyte activation (*46*). Similarly, defects in *ACTL7A*, coding

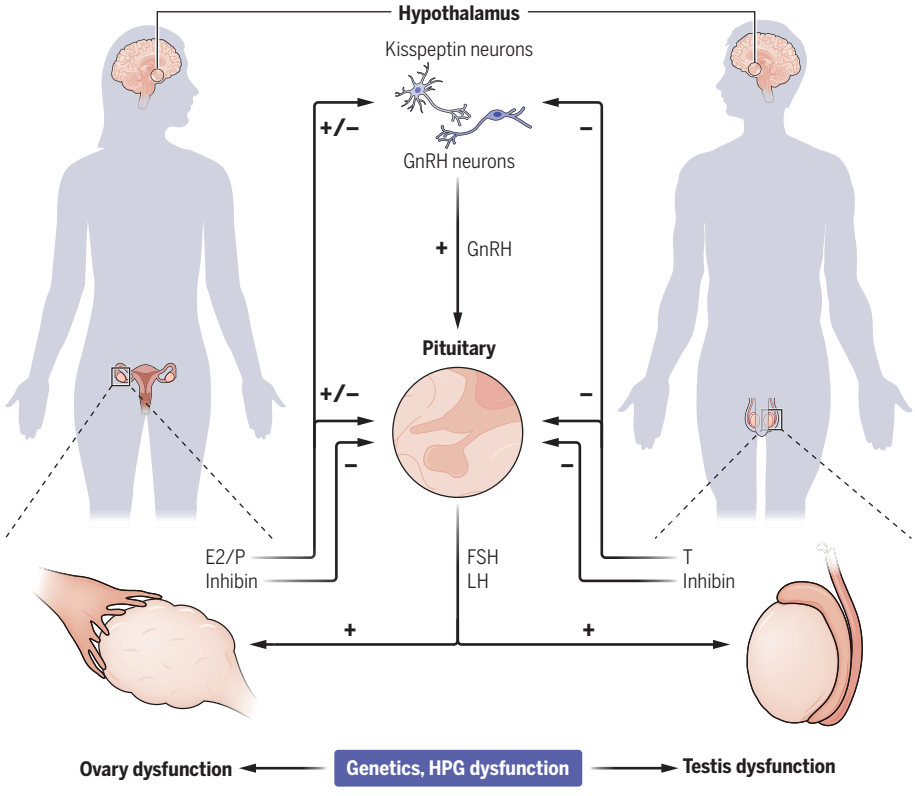


Fig. 1. The HPG axis and gonad function. Kisspeptin stimulates GnRH neurons to release GnRH, which induces the secretion of FSH and LH in the anterior pituitary gland, and these further regulate the normal function of ovary and testis. Hormones produced by the ovary [estradiol (E2), progesterone (P), and inhibin] or testis [testosterone (T) and inhibin] can exert feedback effects to adjust the amounts of FSH and LH. Genetic factors and hypothalamic or pituitary dysfunction can cause gonadal dysfunction by disrupting the HPG axis.

ILLUSTRATION: A. MASTIN/*SCIENCE*

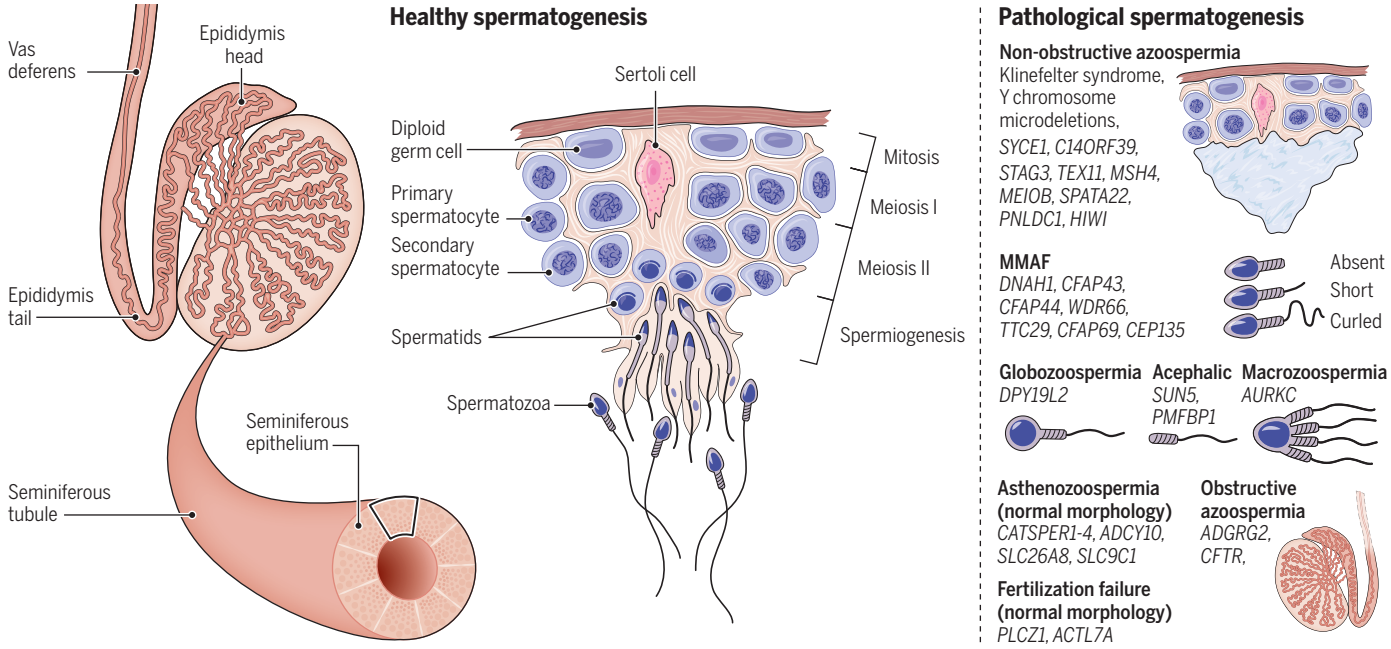


Fig. 2. Representation of physiological and pathological spermatogenesis. Shown is a representation of healthy and pathological spermatogenesis. Obstructive azoospermia can be caused by blockage or atresia of the vas deferens, whereas abnormal spermatogenesis can induce a variety of sperm anomalies such as Non-obstructive azoospermia, MMAF, globozoospermia, acephalic spermatozoa syndrome, macrozoospermia, asthenozoospermia, and fertilization failure without morphological anomalies. Because of the lack of space, only some examples of pathological genes are indicated for each anomaly.

for an actin-like related protein, interfere with PLCZ1 localization and prevent fertilization of in vitro fertilization (IVF). ICSI could achieve fertilization, but was followed by early embryonic arrest (*47*). With widely used sequencing technology and careful phenotyping of sperm defects, additional causal genes will be identified.

Oocyte defects

Aging is a prominent physiological factor that affects the developmental competence of oocytes. It has been suggested that meiotic segregation errors in oocytes follow a U curve according to female age, which reflects women's natural fertility curve peaking at ~27 years of age (*48*). In addition, ovarian aging caused by increased DNA damage, oxidative stress, mitochondrial dysfunction, and telomere shortening in oocytes and/or surrounding granulosa cells have also been related to the age-related decline in women's fertility (*49*, *50*).

The acquisition of developmental competence in oocytes requires normal oocyte nuclear and cytoplasmic maturation (*51*). In each menstrual cycle, in response to the LH surge, some germinal vesicle oocytes undergo germinal vesicle breakdown, microtubule nucleation, and spindle assembly to achieve chromosome segregation and the extrusion of the first polar body to complete nuclear maturation. Cytoplasmic maturation involves changes in organelles and cytoskeleton and other molecular events, preparing oocytes for fertilization and embryo development (Fig. 3).

An increasing number of genetic studies have recently focused on abnormalities in human oocytes (Fig. 3). In 2016, mutations in *TUBB8*, a primate-specific beta-tubulin, were shown to be responsible for oocyte nuclear maturation arrest by impairing spindle assembly. *TUBB8* accounts for the etiology of oocyte meiosis I arrest in up to 30% of patients (*52*). Biallelic variants in *PATL2*, coding for a putative RNA-binding protein regulating maternal mRNA homeostasis, also induce oocyte nuclear maturation arrest and female infertility (*53*–*55*). Recently, the human oocyte microtubule organizing center was found to be a driver for spindle assembly in human oocytes, and deficiencies in TACC3, the key component of this center, were shown to impair microtubule nucleation and meiotic spindle assembly, leading to nuclear maturation arrest (*56*). In addition, aberrant PANX1 channel activity induced by *PANX1* mutations results in abnormal adenosine triphosphate release and oocyte death, whereas mutations in genes encoding the zona pellucida (ZP) proteins ZP1 to ZP3 cause abnormal ZP formation or oocyte degeneration and empty follicle syndrome (*57*, *58*).

The clinical phenotypes of abnormal oocyte cytoplasmic maturation often manifest in different forms, including fertilization failure, zygotic cleavage failure, and early embryonic arrest (Fig. 3). Inactivating mutations in *WEE2* lead to fertilization failure by disrupting CDC2 phosphorylation, and account for ~40% of cases of fertilization failure resulting from oocyte defects (*59*). *BTG4* and *CHEK1* are two main causative genes responsible for zygotic cleavage failure, impairing the programmed decay of maternal mRNA and arresting the G_2/M transition in zygotes, respectively (*60*, *61*). *PADI6*, encoding a subcortical maternal complex protein, is the first mutant gene reported to cause early embryonic arrest (*62*), and a new mechanism was recently identified in which karyopherin α (KPNA7) deficiency contributes to the phenotype by impairing nuclear protein transport (*63*). Mutations in *NLRP7* (occurring in 55% of patients), *TOP6BL*, *MEI1*, or *REC114* can result in hydatidiform mole, characterized by abnormal embryonic cleavage and trophoblastic hyperplasia (*64*, *65*). There are currently >20 genes reported to cause female infertility from various kinds of oocyte defects, but these only account for a small proportion of patients (*5*, *58*), indicating that additional genes necessary to produce competent human oocytes are yet to be discovered.

Tubal and uterine abnormalities and embryo implantation failure

Favorable tubal and uterine environments are essential for embryo implantation and for the final establishment of pregnancy (Fig. 3). Tubal factor infertility can result from the blockage of fallopian tubes or the disruption of oocyte collection from the ovary caused by pelvic adhesions (*66*). Salpingitis, the most common cause of tubal defects, is a chronic inflammation

ILLUSTRATION: K. HOLOSKI/*SCIENCE*

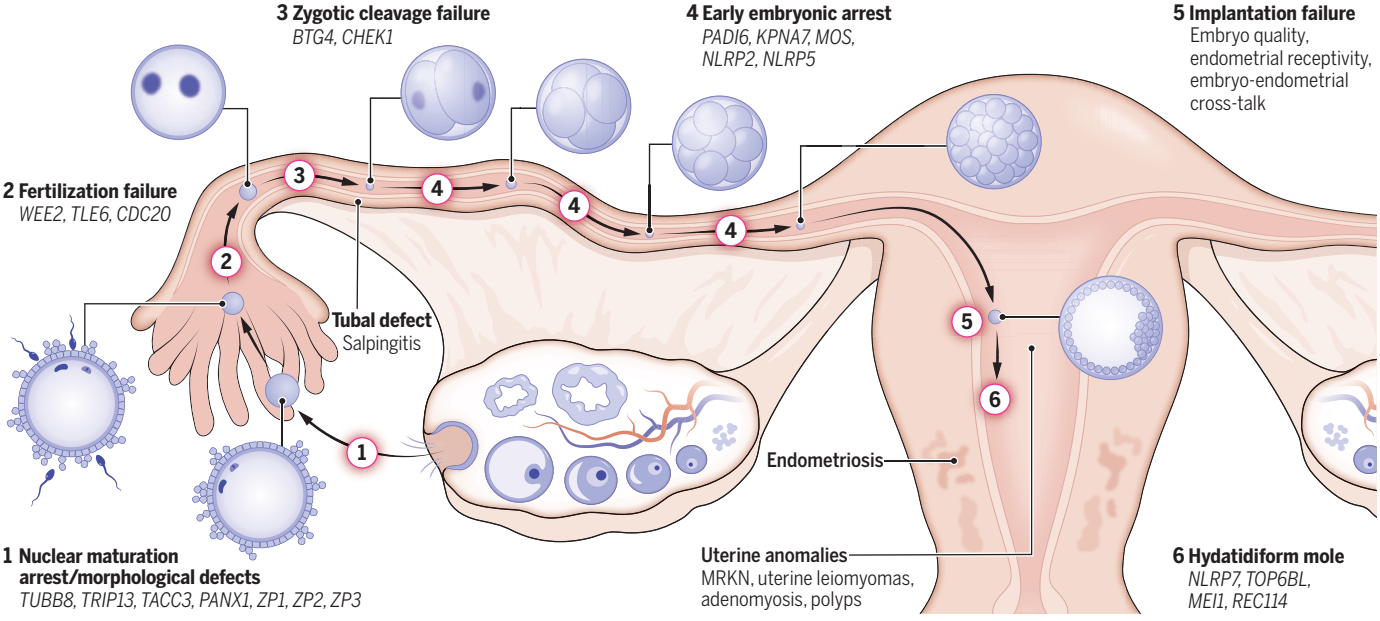


Fig. 3. Process of oocyte maturation, fertilization, embryonic development, and establishment of pregnancy. The oocyte undergoes nuclear maturation to complete meiosis I and develop into a mature egg, which can be fertilized in the fallopian tube. The zygote then undergoes several rounds of cleavage and gradually moves toward the uterus, where the embryo can develop into a blastocyst and implant. Abnormalities in oocyte or embryonic development can result in infertility, including nuclear maturation arrest, fertilization failure, zygotic cleavage failure, early embryonic developmental arrest, and hydatidiform mole. Other associated factors include tubal defects, uterine anomalies, implantation failure, and endometriosis. Some examples of pathological genes are indicated for each anomaly.

induced by infections, which can lead to subsequent damage of the fallopian tubes or pelvic tissues such as hydrosalpinx or pelvic-peritoneal adhesions (*67*). Compelling evidence has shown that the pathogenic bacteria *Chlamydia trachomatis* and *Neisseria gonorrhoeae* are involved in reproductive tract infections that are mainly sexually transmitted, and in these cases, galectins may act as endogenous mediators and potential therapeutic targets (*67*, *68*).

Uterine factor infertility is caused by either the absence of a uterus or the presence of a nonfunctional uterus. Mayer-Rokitansky-Küster-Hauser (MRKH) syndrome, the most common type of congenital uterine malformations (1/4500 to 5000), results from agenesis or atresia of the uterus and/or vagina (*69*). Although a few MRKH candidate genes with incomplete penetrance have been reported, causal genetic contribution in MRKH syndrome is still underestimated (*70*). Acquired factors associated with uterine abnormalities include uterine leiomyomas, adenomyosis, and polyps. Somatic mutations in *MED12* and *FH* and chromosomal rearrangements of *HMGA2* account for 90% of uterine leiomyoma cases (*71*, *72*). Recent studies described that somatic or germline mutations in genes encoding SRCAP complex proteins contribute to the genesis of uterine leiomyomas by disturbing H2A.Z deposition and leading to epigenetic stability (*72*, *73*).

Endometriosis is defined by the presence of endometrial-like tissue outside of the uterine cavity, mainly on the pelvic organs and tissues (*74*) (Fig. 3). One proposed mechanism of endometriosis is the retrograde menstruation of endometrial cells through the fallopian tubes into the peritoneal cavity, where they adhere to the ovaries, ligaments, and peritoneal surfaces (*75*). Population-based genetic association studies and meta-analyses have identified 14 repeated loci of common variations in women of European and Japanese ancestry, and genes related to these loci are implicated in the etiology of endometriosis (*76*). Uncovering the functional evidence for the pathogenic role of these loci remains a challenge, and studies involving subjects of various ancestries should be performed.

Recurrent implantation failure is attributed to low embryo quality (mainly caused by aneuploidies), reduced endometrial receptivity, and impaired embryo-endometrial cross-talk (*77*) (Fig. 3). Aberrant endometrial expression of *BMI1* (*78*) or *HDAC3* (*79*) is associated with abnormal uterine receptivity. Proper combinations of genetically diverse maternal natural killer cell immunoglobulin-like receptors and fetal human leukocyte antigen C molecules are important for embryo invasion (*80*). Recent single-cell transcriptomic analyses of human peri-implantation embryos (*81*) and endometrium across the menstrual cycle (*82*) provide insights into embryo-endometrial cross-talk and embryo implantation, and help in our understanding of the pathological aspects of recurrent implantation failure.

Precision diagnostics and potential treatment strategies for human infertility

Fertility assessment includes general physical examination and various specific tests such as hormone analysis, semen analysis, ovulation tests, and tubal and uterine patency evaluation (*83*, *84*) (Fig. 4). Commonly applied infertility treatments include hormone therapy to induce ovulation and gonad development, surgery to correct tubal or uterine anomalies, and IVF/ICSI and testicular sperm extraction (TESE) to overcome some sperm defects, tubal or uterine defects, and unexplained infertility (*66*, *84*) (Fig. 4). With the proliferation of genetic studies, targeted gene analysis and whole-exome or whole-genome sequencing may greatly improve precision diagnostics. A recent study found that defects in known meiotic genes are always correlated with a negative TESE, indicating that a precise genetic diagnosis can avoid unnecessary TESE (*85*). Additional genetic diagnostics for specific reproductive diseases and phenotypes are expected in the near future (Fig. 4).

More personalized fertility treatment strategies can also be applied. For example, when infertility and fertilization failure are caused by *PLCZ1* mutations in males, calcium ionophores and ICSI can be used directly, and prolonging the interval between ovulation triggering and oocyte pickup is effective for infertile females with *LHCGR* mutations characterized by empty follicle syndrome (*86*, *87*). In addition, drug-free in vitro activation and

ILLUSTRATION: A. FISHER/*SCIENCE*

laparoscopic ovarian incision, although not widely applied, have been used successfully for treating patients with ovarian infertility (*88*).

Considering that there is still a large number of infertile patients who cannot be successfully treated, exploring new treatment strategies is imperative. The rescue of infertile mice lacking different genes necessary for sperm production has been achieved using different gene delivery systems, resulting in live offspring (*89*). A more complex procedure involving in vitro culture of spermatogonial stem cells, the correction of their *Tex11* genetic defect by CRISPR/Cas9, and successful transplantation of the repaired spermatogonial stem cells into the testes of knockout recipients to restore spermatogenesis has also been tested (*90*). Compared with other approaches, CRISPR/Cas9 has potential off-target effects that need to be further investigated. In contrast to editing of somatic cells, gene editing in human gametes opens up important ethical concerns (*91*). No countries permit germline modifications of humans, and the use of germline editing for research purposes is guided by different policies in different countries (*91*).

Although there are no practicable clinical strategies for restoring oocyte quality, some possibilities have been explored (Fig. 4). For example, oocyte quality, embryonic development, and fertility can be restored by nicotinamide mononucleotide treatment in aged mice (*92*). Germline nuclear transfer, including germinal vesicle, pronuclear, and spindle transfer, is a potential strategy for improving the meiotic competence of aged mouse and human oocytes (*93*, *94*). Recent advances in human in vitro gametogenesis provide possibilities for the future treatment of patients with gamete abnormalities, although there are multiple challenges in terms of both methods and ethics (*95*). A genetic correction could be achieved by the injection of *WEE2* mRNA in oocytes from WEE2-deficient infertile women, allowing them to regain normal fertilization and embryonic development (*59*). For oocytes with mitochondrial DNA mutations, mitochondrial replacement therapy by spindle transfer or a forced mitophagy approach provide potential strategies to prevent mitochondrial DNA diseases (*96*, *97*). However, further investigations are required to alleviate all ethical concerns and to evaluate the safety of these methods for the offspring. Treatment with the methyl group donor *S*-adenosylmethionine can correct transcriptomic, neuroendocrine, and metabolic defects in a mouse model of PCOS, thus providing potential candidates for epigenetic-based therapy for patients. Modifying the gut microbiota or the kisspeptin receptor agonist MVT-602 were shown to be possible treatments for PCOS (*26*, *98*). In addition, an anti–IL-8 antibody improved inflammation and fibrosis in a monkey model of endometriosis (*99*), suggesting a potential therapeutical strategy.

Summary and future outlook

Despite recent advances in our knowledge of the etiology of human infertility, our basic understanding of its complex mechanisms remains limited. In the era of multiple-omics studies, it is likely that an increasing number of candidate genes, proteins, and metabolites related to fertility will be identified by integrating genome, transcriptome, and epigenome sequencing; proteomics; and metabolic analysis. A variety of disease-causing variants, such as point mutations, structural variations, and noncoding RNA modification and inheritance patterns including monogenic, digenic and oligogenic, and microbial-host interactions will be uncovered. By using various models, including cell culture, in vitro embryo culture, genetically modified animals, and patient in situ tissues, new mechanisms contributing to human infertility will be systematically elucidated (Fig. 4).

Some specific challenges remain to be overcome. Acquiring diseased reproductive organs and in situ tissues for analysis is critical to comprehend the pathophysiological mechanisms underlying each phenotype. The increasing numbers of IVF and ICSI attempts will provide great opportunities to overcome the constraints in acquiring gametes, cells in ovarian follicles (granulosa cells, cumulus cells, and thecal cells), and early embryos. Moreover, studies in other research fields such as organoids, pluripotent stem cell differentiation, and synthetic biology will certainly provide alternative ways to further improve our mechanistic understanding of infertility. Currently, only limited numbers of infertile patients benefit from precision molecular diagnosis, which hinders disease classification and exploration of targeted treatment. With the design of large databases of multiple omics data with clinically well-characterized infertility phenotypes and artificial intelligence applications, it can be expected that clinical infertility diagnosis will be more precise and personalized. It is therefore expected that comprehensive sequencing and other molecular methodologies will be widely applicable for the diagnosis of infertility. In addition, only very limited treatment strategies targeting certain kinds

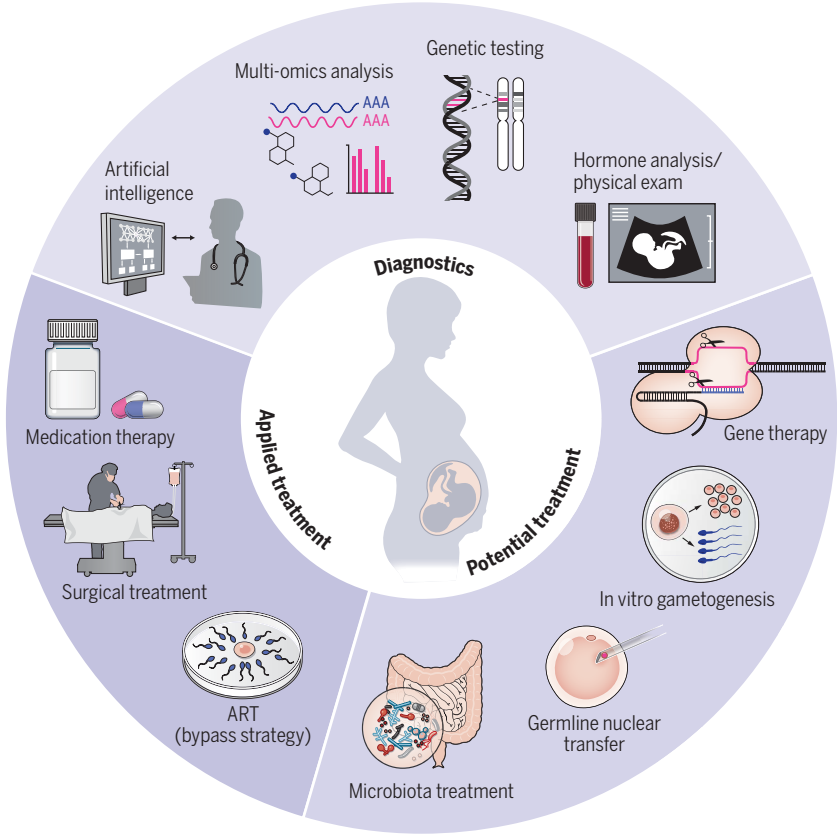


Fig. 4. Precision diagnostics and potential treatment strategies for human infertility. Current diagnostics, including physical examinations, hormone analysis, and applied treatments, and bypass strategies, including medications, surgery, and assisted reproductive technology (ART), have greatly benefited many patients with infertility. Future diagnostic methods should help to optimize clinical practice for infertility with more precision diagnostics and targeted treatment in patients.

ILLUSTRATION: N. CARY/*SCIENCE*

of infertility are currently applicable in clinical practice. The establishment of well-phenotyped longitudinal infertility cohorts of different ethnic populations integrating cutting-edge knowledge on drug screening, gene therapy, and in vitro gametogenesis will help to translate new knowledge into treatment strategies. It should be noted that any molecular manipulation of human germline cells should strictly abide by ethics policies for research purposes. At the same time, suitable policies need to be developed for the genetic manipulation of human germ cells and embryos for research purposes to develop personalized treatment strategies (Fig. 4).

With these advances and endeavors, it is expected that many factors responsible for infertility will be identified, pathological mechanisms of human infertility will be better understood, the clinical practice of infertility diagnosis will be greatly improved, and infertility patients will finally benefit from precision diagnosis and targeted treatments.

REFERENCES AND NOTES

1. J. L. Evers, *Lancet* **360**, 151–159 (2002).
2. L. Pinilla, E. Aguilar, C. Dieguez, R. P. Millar, M. Tena-Sempere, *Physiol. Rev.* **92**, 1235–1316 (2012).
3. U. Boehm *et al*., *Nat. Rev. Endocrinol.* **11**, 547–564 (2015).
4. J. Young *et al*., *Endocr. Rev.* **40**, 669–710 (2019).
5. A. Van Der Kelen *et al*., *Hum. Reprod. Update* **29**, 218–232 (2023).
6. A. Messina *et al*., *Am. J. Hum. Genet.* **106**, 58–70 (2020).
7. D. E. Whittaker *et al*., *J. Clin. Invest.* **131**, e141587 (2021).
8. K. Chachlaki *et al*., *Sci. Transl. Med.* **14**, eabh2369 (2022).
9. M. I. Stamou *et al*., *J. Clin. Endocrinol. Metab.* **107**, 2228–2242 (2022).
10. M. Stamou *et al*., *J. Clin. Endocrinol. Metab.* **105**, e231–e244 (2020).
11. A. Bojesen, S. Juul, C. H. Gravholt, *J. Clin. Endocrinol. Metab.* **88**, 622–626 (2003).
12. C. Liu *et al*., *EMBO J.* **40**, e106864 (2021).
13. C. H. Gravholt, M. H. Viuff, S. Brun, K. Stochholm, N. H. Andersen, *Nat. Rev. Endocrinol.* **15**, 601–614 (2019).
14. M. Cools *et al*., *Nat. Rev. Endocrinol.* **14**, 415–429 (2018).
15. N. Camats, M. Fernández-Cancio, L. Audí, A. Schaller, C. E. Flück, *Eur. J. Hum. Genet.* **26**, 1329–1338 (2018).
16. X. Jiao, H. Ke, Y. Qin, Z. J. Chen, *Trends Endocrinol. Metab.* **29**, 795–807 (2018).
17. I. Huhtaniemi *et al*., *Trends Endocrinol. Metab.* **29**, 400–419 (2018).
18. V. Wesevich, A. N. Kellen, L. Pal, *F1000Res.* **9**, 1101 (2020).
19. H. Ke *et al*., *Nat. Med.* **29**, 483–492 (2023).
20. A. Heddar *et al*., *EBioMedicine* **84**, 104246 (2022).
21. Y. Shi *et al*., *Nat. Genet.* **44**, 1020–1025 (2012).
22. K. M. Hoeger, A. Dokras, T. Piltonen, *J. Clin. Endocrinol. Metab.* **106**, e1071–e1083 (2021).
23. S. Risal *et al*., *Nat. Med.* **25**, 1894–1904 (2019).
24. B. Tata *et al*., *Nat. Med.* **24**, 834–846 (2018).
25. N. E. H. Mimouni *et al*., *Cell Metab.* **33**, 513–530.e8 (2021).
26. X. Qi *et al*., *Nat. Med.* **25**, 1225–1233 (2019).
27. W. Rodprasert, H. E. Virtanen, J. A. Mäkelä, J. Toppari, *Front. Endocrinol. (Lausanne)* **10**, 906 (2020).
28. D. Le Lannou *et al*., *Hum. Reprod.* **10**, 338–341 (1995).
29. O. Patat *et al*., *Am. J. Hum. Genet.* **99**, 437–442 (2016).
30. M. Punab *et al*., *Hum. Reprod.* **32**, 18–31 (2017).
31. J. Gekas *et al*., *Hum. Reprod.* **16**, 82–90 (2001).
32. S. Ozturk, *Cell Cycle*10.1080/15384101.2023.2171544 (2023).
33. L. Nagirnaja *et al*., *N. Engl. J. Med.* **385**, 707–719 (2021).
34. L. T. Gou *et al*., *Cell* **169**, 1090–1104.e13 (2017).
35. K. Dieterich *et al*., *Nat. Genet.* **39**, 661–665 (2007).
36. R. Harbuz *et al*., *Am. J. Hum. Genet.* **88**, 351–361 (2011).
37. F. Zhu *et al*., *Am. J. Hum. Genet.* **99**, 942–949 (2016).
38. F. Zhu *et al*., *Am. J. Hum. Genet.* **103**, 188–199 (2018).
39. L. Visser *et al*., *Fertil. Steril.* **95**, 1020–4.e1, 9 (2011).
40. T. Dirami *et al*., *Am. J. Hum. Genet.* **92**, 760–766 (2013).
41. E. Cavarocchi *et al*., *Clin. Genet.* **99**, 684–693 (2021).
42. A. Akbari *et al*., *Hum. Reprod.* **34**, 1155–1164 (2019).
43. M. Ben Khelifa *et al*., *Am. J. Hum. Genet.* **94**, 95–104 (2014).
44. C. Coutton *et al*., *Nat. Commun.* **9**, 686 (2018).
45. A. Touré *et al*., *Hum. Genet.* **140**, 21–42 (2021).
46. J. Escoffier *et al*., *Hum. Mol. Genet.* **25**, 878–891 (2016).
47. A. Xin *et al*., *Sci. Adv.* **6**, eaaz4796 (2020).
48. J. R. Gruhn *et al*., *Science* **365**, 1466–1469 (2019).
49. L. Secomandi, M. Borghesan, M. Velarde, M. Demaria, *Hum. Reprod. Update* **28**, 172–189 (2022).
50. K. S. Ruth *et al*., *Nature* **596**, 393–397 (2021).
51. E. Voronina, G. M. Wessel, *Curr. Top. Dev. Biol.* **58**, 53–110 (2003).
52. R. Feng *et al*., *N. Engl. J. Med.* **374**, 223–232 (2016).
53. B. Chen *et al*., *Am. J. Hum. Genet.* **101**, 609–615 (2017).
54. M. Christou-Kent *et al*., *EMBO Mol. Med.* **10**, e8515 (2018).
55. S. Maddirevula *et al*., *Am. J. Hum. Genet.* **101**, 603–608 (2017).
56. T. Wu *et al*., *Science* **378**, eabq7361 (2022).
57. Q. Sang *et al*., *Sci. Transl. Med.* **11**, eaav8731 (2019).
58. Q. Sang, Z. Zhou, J. Mu, L. Wang, *J. Assist. Reprod. Genet.* **38**, 993–1002 (2021).
59. Q. Sang *et al*., *Am. J. Hum. Genet.* **102**, 649–657 (2018).
60. W. Zheng *et al*., *Am. J. Hum. Genet.* **107**, 24–33 (2020).
61. H. Zhang *et al*., *Cell Res.* **31**, 814–817 (2021).
62. Y. Xu *et al*., *Am. J. Hum. Genet.* **99**, 744–752 (2016).
63. W. Wang *et al*., *J. Clin. Invest.* **133**, e159951 (2023).
64. S. Murdoch *et al*., *Nat. Genet.* **38**, 300–302 (2006).
65. N. M. P. Nguyen *et al*., *Am. J. Hum. Genet.* **103**, 740–751 (2018).
66. S. A. Carson, A. N. Kallen, *JAMA* **326**, 65–76 (2021).
67. D. G. Tsevat, H. C. Wiesenfeld, C. Parks, J. F. Peipert, *Am. J. Obstet. Gynecol.* **216**, 1–9 (2017).
68. A. L. Lujan, D. O. Croci, G. A. Rabinovich, M. T. Damiani, *Nat. Rev. Urol.* **19**, 240–252 (2022).
69. Committee on Adolescent Health Care, *Obstet. Gynecol.* **131**, e35–e42 (2018).
70. M. K. Herlin, M. B. Petersen, M. Brännström, *Orphanet J. Rare Dis.* **15**, 214 (2020).
71. N. Mäkinen *et al*., *Science* **334**, 252–255 (2011).
72. N. Välimäki *et al*., *Am. J. Hum. Genet.* **110**, 460–474 (2023).
73. D. G. Berta *et al*., *Nature* **596**, 398–403 (2021).
74. C. Chapron, L. Marcellin, B. Borghese, P. Santulli, *Nat. Rev. Endocrinol.* **15**, 666–682 (2019).
75. K. T. Zondervan *et al*., *Nat. Rev. Dis. Primers* **4**, 9 (2018).
76. K. T. Zondervan, C. M. Becker, S. A. Missmer, *N. Engl. J. Med.* **382**, 1244–1256 (2020).
77. L. Craciunas *et al*., *Hum. Reprod. Update* **25**, 202–223 (2019).
78. Q. Xin *et al*., *J. Clin. Invest.* **128**, 175–189 (2018).
79. T. H. Kim *et al*., *Sci. Transl. Med.* **11**, eaaf7533 (2019).
80. J. M. Franasiak *et al*., *Fertil. Steril.* **116**, 1436–1448 (2021).
81. D. Liu *et al*., *Sci. Adv.* **8**, eabj3725 (2022).
82. W. Wang *et al*., *Nat. Med.* **26**, 1644–1653 (2020).
83. C. M. Farquhar *et al*., *Nat. Rev. Dis. Primers* **5**, 7 (2019).
84. National Collaborating Centre for Women's and Children's Health (UK), *Fertility: Assessment and Treatment for People with Fertility Problems* (Royal College of Obstetricians & Gynaecologists, 2013).
85. Z. E. Kherraf *et al*., *Am. J. Hum. Genet.* **109**, 508–517 (2022).
86. J. Kashir *et al*., *Hum. Reprod.* **27**, 222–231 (2012).
87. X. Lu *et al*., *J. Clin. Endocrinol. Metab.* **104**, 5877–5892 (2019).
88. Y. Tanaka, A. J. Hsueh, K. Kawamura, *Fertil. Steril.* **114**, 1355–1357 (2020).
89. J. P. Alves-Lopes, J. B. Stukenborg, *Hum. Reprod. Update* **24**, 176–191 (2018).
90. S. Wang *et al*., *Asian J. Androl.* **23**, 562–571 (2021).
91. J. Turocy, E. Y. Adashi, D. Egli, *Cell* **184**, 1561–1574 (2021).
92. M. J. Bertoldo *et al*., *Cell Rep.* **30**, 1670–1681.e7 (2020).
93. M. Tang *et al*., *Hum. Reprod.* **35**, 1562–1577 (2020).
94. J. Zhang, *J. Assist. Reprod. Genet.* **32**, 313–317 (2015).
95. M. Saitou, K. Hayashi, *Science* **374**, eaaz6830 (2021).
96. E. Kang *et al*., *Nature* **540**, 270–275 (2016).
97. X. Y. Fan *et al*., *Nat. Biomed. Eng.* **6**, 339–350 (2022).
98. A. Abbara *et al*., *J. Clin. Invest.* **130**, 6739–6753 (2020).
99. A. Nishimoto-Kakiuchi *et al*., *Sci. Transl. Med.* **15**, eabq5858 (2023).

ACKNOWLEDGMENTS

We apologize for not citing any studies that were relevant to the theme but were left out because of article length limitations. **Funding:** This work was supported by the National Natural Science Foundation of China (grants 82288102 and 32130029) and the National Key Research and Development Program of China (grant 2021YFC2700100). **Competing interests:** The authors declare no competing interests.

License information: Copyright © 2023 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

Submitted 10 January 2023; accepted 28 February 2023
10.1126/science.adf7760

REVIEW

Assisted reproductive technologies at the nexus of fertility treatment and disease prevention

Mary Herbert[1,2,3]*, Meenakshi Choudhary[2], Deidre Zander-Fox[4,5]

Assisted reproductive technology (ART) refers to processing gametes in vitro and usually involves in vitro fertilization. Originally developed for the treatment of infertility, culture of human embryos in vitro also provides an opportunity to screen embryos for inherited genetic disorders of the nuclear and mitochondrial genomes. Progress in identifying causative genetic variants has massively increased the scope of preimplantation genetic testing in preventing genetic disorders. However, because ART procedures are not without risk of adverse maternal and child outcomes, careful consideration of the balance of risks and benefits is warranted. Further research on early human development will help to minimize risks while maximizing the benefits of ART.

Since the groundbreaking work that led to the birth of the first in vitro fertilization (IVF) baby in 1978, an estimated 0.1% of the global population have been conceived by assisted reproductive technology (ART) (*1*). ART in combination with genetic testing of preimplantation embryos has also enabled the birth of healthy children who would otherwise be at risk of serious inherited congenital disorders (*2*). Advances in ART procedures combined with whole-genome amplification and genome-wide analysis (*3*) have greatly facilitated the clinical application of preimplantation genetic testing (PGT). In this Review, we focus on recent developments and current limitations of ART procedures in reducing the risk of genetic disorders, including trisomy and diseases associated with pathogenic variants in the mitochondrial genome. We also highlight the importance of research on human embryos to better understand the effects of ART interventions on early development and on the health of ART-conceived children.

Preimplantation genetic testing

Until the advent of PGT, interventions to reduce the risk of congenital disorders were limited to prenatal diagnosis followed by termination of an affected pregnancy. PGT enables identification of unaffected embryos from an ART-generated cohort and can be used to detect disease-associated nuclear genetic variants, inherited structural abnormalities such as Robertsonian translocations, and numerical chromosome abnormalities (aneuploidy) (*4*). In a more recent application, PGT has also been used to reduce the risk of disease associated with pathogenic variants in the mitochondrial genome (*5*).

PGT is based on the principle that the genetic constitution of cells biopsied from the preimplantation embryo is representative of the population of cells that will eventually give rise to the fetus. Following fertilization, the single-cell zygote undergoes successive rounds of cell division to form totipotent blastomeres, which, upon differentiation, give rise to the embryonic and extraembryonic lineages of the blastocyst. The epiblast lineage contained within the inner cell mass contributes to all fetal tissues, and the outer layer of trophectoderm cells forms the placenta (Fig. 1). PGT was initially performed on one or two blastomeres removed from the embryo at the six- to eight-cell stage. Improvements in embryo culture conditions and blastocyst cryopreservation facilitated a shift toward trophectoderm-based genetic analysis (*4*).

> **"...an estimated 0.1% of the global population have been conceived by assisted reproductive technology."**

Trophectoderm biopsy offers several advantages over cleavage-stage biopsy, including providing more DNA for genetic analysis and enhanced embryo viability (*6*). Despite evidence that trophectoderm biopsy is associated with an increased incidence of hypertensive disorders during pregnancy, adverse neonatal outcomes are not increased compared with standard intracytoplasmic sperm injection (ICSI) (*7*). However, it should be noted that ICSI is associated with an increased risk of adverse maternal and child outcomes compared with natural conception (*8*, *9*). Thus, use of PGT for otherwise fertile people warrants careful consideration of the balance of risks and benefits.

PGT for nuclear-encoded genetic disorders

Initially used to detect sex-linked disorders, PGT is now widely used to screen for monogenic diseases, most prominently cystic fibrosis, sickle cell anemia, thalassemia, and Huntington's disease (*2*, *10*). Rapid progress in identifying causative variants for a large fraction of the ~7000 known monogenic disorders (*11*) has massively increased the scope of PGT in preventing rare monogenic diseases estimated to affect 1 in 50 individuals (*11*). Beyond testing in cases with a family history of disease, PGT can be further extended by proactive approaches, such as a recent Australian initiative, called Mackenzie's Mission, to perform carrier screening for 750 serious childhood-onset genetic conditions. Of the 5000 couples screened thus far through Mackenzie's Mission, 1.9% were determined to be at risk of having a child with a serious genetic condition (*12*).

Although the logic for PGT in preventing monogenic disorders is very clear, the recent use of PGT to predict the risk of common polygenic diseases such as type 2 diabetes and cardiovascular disease (*13*) is controversial. Genetic predisposition to such conditions can be reflected in polygenic risk scores encompassing combinations of many variants across the genome (*13*). Whereas polygenic risk scoring models are considered useful in informing clinical care and public health policy, their use for embryo selection has raised several concerns (*14*, *15*). Chief among these is the fact that an individual's risk of developing a condition such as cardiovascular disease or type 2 diabetes is determined by environmental and lifestyle factors in combination with genetic variation. In addition, current polygenic risk scoring models are based largely on data from people of European ancestry, and their validity across diverse populations is questionable (*14*, *15*).

PGT for chromosomal abnormalities

PGT for aneuploidy screening (PGT-A) was originally used to reduce the risk of miscarriage and congenital disorders (such as Down syndrome and Edwards syndrome) due to trisomy (*2*). Trisomy in the fertilized egg is predominantly of maternal origin and arises during the meiotic divisions, which generate a haploid maternal genome from the diploid oocyte (*16*). Because the incidence of trisomy increases as women approach their late 30s (*16*), PGT-A was primarily targeted at older women undergoing ART treatment. This screening was initially based on the use of fluorescence-labeled probes to detect error-prone chromosomes such as 13, 16, 18, 21, and X (*2*), but the advent of comparative genomic hybridization

[1]Biosciences Institute, Newcastle University, Newcastle upon Tyne NE1 4EP, UK. [2]Newcastle Fertility Centre, Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne NE1 4EP, UK. [3]Development and Stem Cells Program, Department of Anatomy and Developmental Biology, Monash Biomedicine Discovery Institute, Monash University, Melbourne, VIC 3800, Australia. [4]Monash IVF, Melbourne, VIC 3800, Australia. [5]Department of Obstetrics and Gynaecology, School of Clinical Science, Monash University, Melbourne, VIC 3800, Australia.
*Corresponding author. Email: mary.herbert@ncl.ac.uk

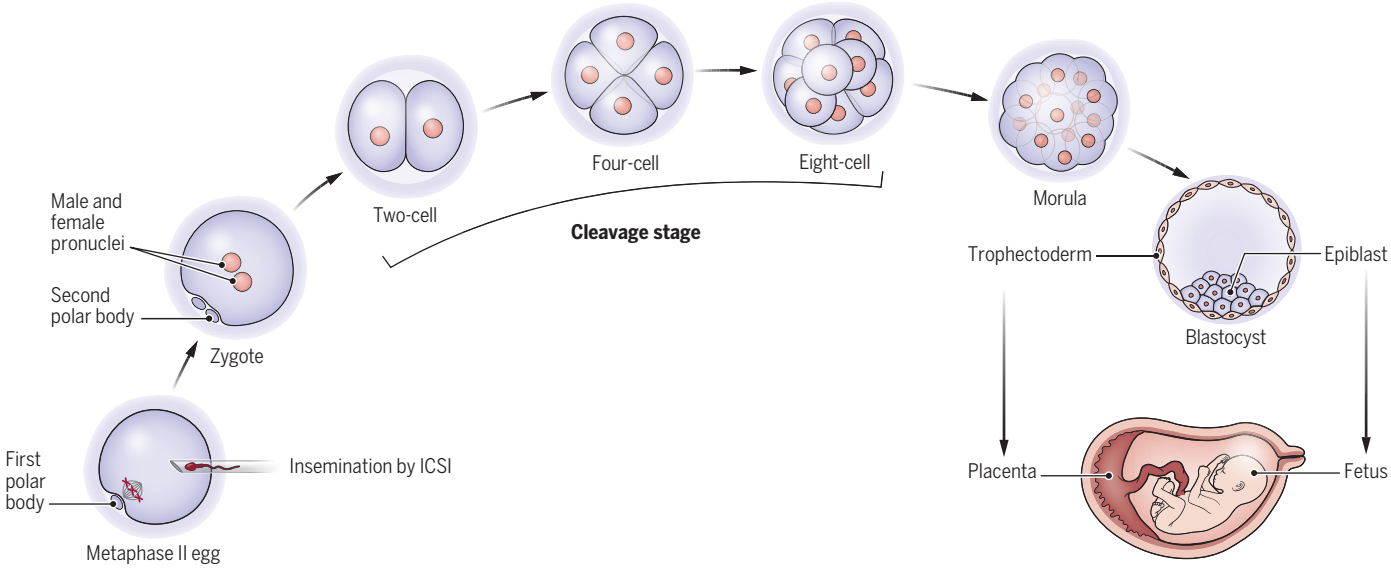


Fig. 1. Preimplantation development of human embryos and the fate of the trophectoderm and epiblast cells.

(CGH) and, more recently, next-generation sequencing (NGS) made it possible to detect copy number variation across all chromosomes (*3*, *10*). This advance heralded the use of PGT-A to improve ART outcomes, irrespective of a woman's age, by identifying euploid embryos. Although there is a lack of consensus as to the benefits (*10*), a recurring theme, including from a multicenter clinical trial, is that while PGT-A reduces the time to pregnancy for older women, it also reduces the chance of live birth per ART cycle for younger women (*17*).

Key to understanding the limitations of PGT-A is the question of whether biopsied trophectoderm cells accurately reflect the chromosomal constitution of the epiblast cells destined to give rise to the fetus. Because aneuploidy of meiotic origin is expected to be present in all cells, a diagnosis of euploidy in the sampled trophectoderm cells is indicative of an error-free meiosis. However, interpretation of aneuploidy in some or all of the sampled trophectoderm cells is complicated by a high incidence of chromosome segregation errors during the mitotic divisions of the embryo (*18*). The number of cells affected by aneuploidy of mitotic origin depends on when during development the missegregation event occurred. Moreover, concordance between sampled trophectoderm cells and the epiblast lineage depends on a shared progenitor without subsequent chromosome segregation errors. Thus, detection of aneuploidy in a sample of trophectoderm cells does not necessarily imply aneuploidy in cells of the epiblast lineage. This raises the question of whether the ~30% of PGT-A cases deemed to have no embryos suitable for transfer (*19*) may be deprived of a chance at pregnancy owing to misdiagnosis. Consistent with this, a growing body of evidence indicates that the presence of mosaicism (a mix of euploid and aneuploid cells) in the trophectoderm biopsy is compatible with pregnancy and euploid live births (*20*, *21*).

Although the efficacy of PGT-A in improving pregnancy and live birth rates is contentious, the rationale for reducing the risk of maternal-origin trisomy remains compelling, especially in view of the ever-increasing numbers of older women undergoing ART treatment. Before the practice of trophectoderm biopsy came to prominence, a number of studies reported promising findings from genetic analysis of polar bodies (*22*–*24*). Because polar bodies are the nonviable products of female meiosis (Fig. 1), they offer a proximal readout of errors during both meiotic divisions with the potential added advantage of avoiding adverse outcomes associated with removal of cells at later stages of development (*7*). A clinical trial to test the efficacy of polar body analysis indicated that the miscarriage rate was halved (*25*). The live birth rate was equivalent to that of untested controls; however, success was hampered by failure to obtain a genetic diagnosis in 21% of cases using array-CGH (*25*). A subsequent study reported promising findings on diagnostic yield and accuracy using NGS combined with a modified procedure to reduce DNA amplification errors (*24*). Polar body–based analysis could be further optimized by research to determine the optimal timing of removal of the first and second polar bodies. Genetic analysis of polar bodies therefore offers a minimally invasive approach to detecting maternal-origin trisomy for at-risk groups, without the confounding effects of chromosome segregation errors arising during the mitotic divisions of the embryo.

PGT to prevent inherited mitochondrial disease

Whereas PGT is conventionally used to detect abnormalities of the nuclear genome, it has more recently been used to detect pathogenic variants in the mitochondrial genome. Pathogenic mitochondrial DNA (mtDNA) variants are associated with a broad spectrum of disease symptoms estimated to affect ~1 in 5000 individuals, which in the most severe cases result in death in the first years of life (*26*). MtDNA is maternally inherited, and in most cases pathogenic variants coexist with wild-type mtDNA, resulting in heteroplasmy. Wide variation in variant load between oocytes from the same woman makes it very difficult to predict the risk of disease. In such cases, PGT can be used to identify embryos with variant loads below the threshold for disease symptoms (*5*).

The success of PGT in reducing risk of mtDNA disease depends on the mutation load in sampled cells being representative of the rest of the embryo. Evidence across a range of inherited mtDNA mutations indicates little variation in heteroplasmy levels between blastomeres of cleavage-stage embryos (*5*, *27*). PGT for mtDNA disease is largely based on testing blastomeres biopsied at the six- to eight-cell stage. Whereas embryo viability may be improved by performing trophectoderm-based testing (*6*), a discrepancy in child heteroplasmy levels between two different laboratories (*28*, *29*) raised some doubt about the reliability of trophectoderm

ILLUSTRATION: K. HOLOSKI/*SCIENCE*

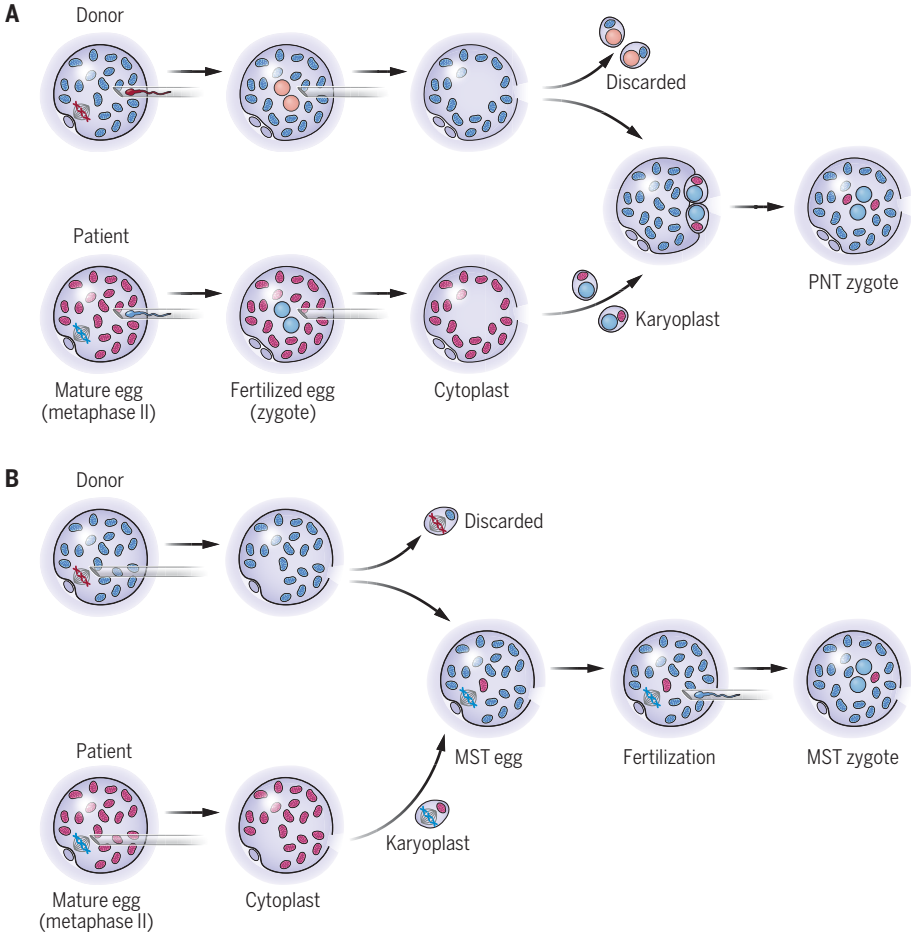


Fig. 2. Illustration of two mitochondrial replacement techniques currently in use. (**A**) Pronuclear transfer (PNT) involves transplantation of the nuclear genome from a fertilized patient egg into an enucleated donor egg after the parental genomes become encapsulated in pronuclei. (**B**) Metaphase II spindle transfer (MST) involves transplantation of the second meiotic spindle from the patient egg into the enucleated donor egg before fertilization.

biopsy. Further investigation will be needed to ensure that access to PGT treatment for affected families is not hindered by technical issues related to the biopsy stage or quantification of mtDNA heteroplasmy.

Mitochondrial replacement to reduce transmission of mtDNA mutations

Mitochondrial replacement technology (MRT) has the potential to reduce transmission of pathogenic mtDNA variants in cases where all available embryos contain high variant loads. Notably, some pathogenic variants are typically homoplasmic (present in all copies of mtDNA). Because homoplasmic variants are characterized by variable disease penetrance (*30*), an unaffected woman may have a child who develops severe disease.

MRT involves removing the nuclear genome from the egg of a woman who carries a pathogenic variant and fusing it with an enucleated egg from an unaffected donor (*31*). The procedure can be performed before fertilization when the egg is arrested at metaphase of meiosis II, or after sperm entry when the parental genomes become encapsulated in pronuclei (Fig. 2). MRT effectively replaces maternal mitochondria together with all other cytoplasmic components. However, some maternal mitochondria are unavoidably cotransferred along with the nuclear genome. Despite accounting for a small fraction of the mtDNA of MRT-generated embryos, cotransferred mitochondria can come to prominence, resulting in complete reversion to the maternal mitochondrial genome in a minority of embryonic stem cell lines derived from MRT embryos (*32*). Although complete mitochondrial genome reversion has not been observed during development in vivo, animal studies indicate that MRT-induced heteroplasmy can vary between different organ systems (*33*). Moreover, a recent report in which MRT was used for infertility treatment (see the section below) indicates that one in six babies had elevated (40%) heteroplasmy for maternal mtDNA in peripheral blood, despite very low levels in tested trophectoderm cells. These findings suggest unequal segregation and/or preferential replication of maternal mtDNA during development.

Bridging the gap between risk reduction and disease prevention will require the development of adjunct technologies to minimize MRT-induced heteroplasmy. In principle, this could be accomplished using targeted nucleases to linearize the circular molecule of mtDNA, which results in its rapid degradation (*34*). An alternative approach would be to deploy the molecular machinery of autophagy or mitophagy to induce selective degradation of maternal mitochondria. In support of this, ectopic expression of an autophagy receptor fused to an outer mitochondrial membrane–targeting sequence has been reported to reduce MRT-induced heteroplasmy (*35*). However, safe clinical application will require strategies to avoid unintended degradation of donor mitochondria.

While ongoing research is focused on addressing the technical limitations, a clinical study to determine the efficacy of MRT is currently underway in the UK, and a live birth with low levels of heteroplasmy for a pathogenic mtDNA variant has been reported (*32*). Follow-up studies in children born after MRT will be essential for continued evaluation of safety and efficacy. In the meantime, a cautious approach is warranted, and MRT should be offered only in cases where PGT in unlikely to be effective.

Mitochondrial replacement for the treatment of infertility

MRT has been legalized solely for the purpose of reducing transmission of mtDNA disease in the UK and, more recently, in Australia. However, a recent report on the use of MRT for infertility treatment indicated that MRT by spindle transfer resulted in successful outcomes in cases with a history of numerous failed ART cycles (*36*). Thus, transplantation of the egg's nuclear genome to a donor cytoplasm may compensate for deficiencies in organelle function and/or in the stockpile of maternal transcripts on which the early embryo relies. Consistent with this idea, it has been reported that spindle transfer rescues developmental arrest in mouse embryos (*37*). With the exception of one case in which 40% heteroplasmy for maternal mtDNA was detected in peripheral blood (discussed in the previous section), babies born after MRT for infertility treatment had low levels of heteroplasmy for maternal mtDNA, and no developmental abnormalities were detected during follow-up (*36*).

It should be noted that MRT is unlikely to be effective in reversing the female age–related increase in oocyte aneuploidy. Aneuploidy due

ILLUSTRATION: K. HOLOSKI/*SCIENCE*

to errors during meiosis I is already present by the time the spindle transfer is performed. Moreover, age-related errors during the second meiotic division predominantly involve prematurely separated sister chromatids (*16*). Cohesion between sister chromatids cannot be restored by performing spindle transfer. Indeed, the presence of prematurely separated chromatids may render oocytes more vulnerable to chromosome loss during the spindle transfer procedure.

Risk of adverse outcomes after ART treatment

While ART treatments continue to be of enormous benefit to human health and wellbeing, the procedures themselves are not without risk. Progress in understanding ART-related risk has been hampered by confounding effects of subfertility on maternal and child outcomes. However, a clearer picture of the ART-specific effects is emerging from real-world population studies in which pregnancy and birth outcomes are compared among cases of subfertility in which conception was achieved with or without ART (*8*, *9*). While there are caveats related to the severity of subfertility, the evidence suggests that ART procedures contribute independently to placentation abnormalities (*9*) and a marginal increase in adverse child outcomes (*8*). The reported incidence of nonchromosomal birth anomalies after natural conception in cases of subfertility is 2.0% compared with up to 2.5% after ART treatment (*8*), and it has been proposed that this may be related to ART-associated placentation abnormalities (*38*). Further analysis of real-world population data to investigate ART procedure–specific effects is needed to inform the assessment of risk relative to proposed benefits of complex ART interventions.

Outlook

ART has enabled millions of people across the globe to have children and to avoid transmission of genetic disorders. The potential for preventing diseases is greatly enhanced by recent progress in identifying causative variants of thousands of monogenic diseases (*11*). In relation to difficulty in interpreting trophectoderm-based aneuploidy screening for maternal-origin trisomy, the validated technique of polar body testing (*22*–*25*) avoids confounding effects of mitotic errors by directly analyzing the products of female meiosis. The benefit of an unambiguous diagnosis for at-risk women should not be underestimated when considering appropriate metrics of success. In cases where all eggs are aneuploid, a clear diagnosis would help women to avoid undergoing multiple futile and costly ART treatment cycles.

The use of ART procedures to reduce transmission of pathogenic variants in the mitochondrial genome offers hope to families affected by otherwise incurable disease. PGT has been used successfully to reduce the risk of mtDNA disease by identifying embryos with low variant loads (*39*). In cases where all embryos carry high levels of a pathogenic variant, recently developed MRT procedures can be used to reduce transmission (*32*). However, MRT may not be effective in all cases owing to the potential for resurgence of the maternal mitochondrial genome (*32*). Further research is required to understand the underlying causes and develop mitigation strategies. In the meantime, it will be important to counsel prospective patients that MRT cannot guarantee disease prevention. Moreover, given the invasive nature of MRT treatments, it will be important to define criteria for the use of MRT in the treatment of infertility. The rapidly advancing fields of single-cell transcriptomic and proteomic analysis provide an opportunity for improved diagnostics to underpin targeted clinical application of MRT for the treatment of infertility.

Finally, research on the earliest stages of human development will help to maximize the considerable benefits of ART while minimizing the risks. Exciting developments in stem cell–derived embryo models offer a powerful tool for investigating human embryogenesis, pregnancy loss, and early developmental anomalies (*40*). However, these models do not recapitulate the very earliest stages of human development. Progress on this front necessitates creation of human embryos specifically for research. Thanks to the brave and pioneering work of Anne McLaren and Mary Warnock (*41*), the legal and regulatory framework in the UK demonstrates that this type of research can be conducted within a robust regulatory environment. However, the creation of embryos for research is permitted in surprisingly few countries. In addition, many major research funding bodies have a blanket ban on funding research involving human embryos. Given progress toward validation of low-cost ART approaches to increase global access (*42*), it is estimated that ART-conceived children will account for 3.5% of the global population by the end of the century (*1*). With this in mind, it is perhaps time to reframe the ethical debate in terms of safeguarding the health of future generations.

REFERENCES AND NOTES

1. M. J. Faddy, M. D. Gosden, R. G. Gosden, *Reprod. Biomed. Online* **36**, 455–458 (2018).
2. D. Cimadomo *et al*., *Hum. Reprod. Update* **26**, 453–473 (2020).
3. O. Tšuiko, E. Fernandez Gallardo, T. Voet, J. R. Vermeesch, *Reproduction* **160**, A19–A31 (2020).
4. M. De Rycke *et al*., *Reproduction* **160**, A45–A58 (2020).
5. H. J. M. Smeets, S. C. E. H. Sallevelt, J. C. F. M. Dreesen, C. E. M. de Die-Smulders, I. F. M. de Coo, *Ann. N. Y. Acad. Sci.* **1350**, 29–36 (2015).
6. R. T. Scott Jr., K. M. Upham, E. J. Forman, T. Zhao, N. R. Treff, *Fertil. Steril.* **100**, 624–630 (2013).
7. W. Y. Zhang *et al*., *Fertil. Steril.* **112**, 283–290.e2 (2019).
8. B. Luke *et al*., *Hum. Reprod.* **37**, 2672–2689 (2022).
9. J. E. Stern *et al*., *Fertil. Steril.* **117**, 1223–1234 (2022).
10. I. T. Lee, M. Kappy, E. J. Forman, A. Dokras, *Fertil. Steril.* S0015-0282(23)00162-0 (2023).
11. K. M. Boycott *et al*., *Am. J. Hum. Genet.* **100**, 695–705 (2017).
12. About Mackenzie's Mission, https://www.mackenziesmission.org.au/what-is-mackenzies-mission/.
13. C. M. Lewis, E. Vassos, *Genome Med.* **12**, 44 (2020).
14. P. Turley *et al*., *N. Engl. J. Med.* **385**, 78–86 (2021).
15. E. Karavani *et al*., *Cell* **179**, 1424–1435.e8 (2019).
16. M. Herbert, D. Kalleas, D. Cooney, M. Lamb, L. Lister, *Cold Spring Harb. Perspect. Biol.* **7**, a017970 (2015).
17. S. Munné *et al*., *Fertil. Steril.* **112**, 1071–1079.e7 (2019).
18. E. Vanneste *et al*., *Nat. Med.* **15**, 577–583 (2009).
19. A. van Montfoort *et al*., *Hum. Reprod. Open* **2021**, hoab024 (2021).
20. S. Munné *et al*., *Eur. J. Med. Genet.* **63**, 103741 (2020).
21. A. Capalbo *et al*., *Am. J. Hum. Genet.* **108**, 2238–2247 (2021).
22. J. Geraedts *et al*., *Hum. Reprod.* **26**, 3173–3180 (2011).
23. A. H. Handyside *et al*., *Eur. J. Hum. Genet.* **20**, 742–747 (2012).
24. Y. Hou *et al*., *Cell* **155**, 1492–1506 (2013).
25. W. Verpoest *et al*., *Hum. Reprod.* **33**, 1767–1776 (2018).
26. G. S. Gorman *et al*., *Nat. Rev. Dis. Primers* **2**, 16080 (2016).
27. J. Steffann *et al*., *Cell Rep.* **7**, 933–934 (2014).
28. S. Mitalipov, P. Amato, S. Parry, M. J. Falk, *Cell Rep.* **7**, 935–937 (2014).
29. N. R. Treff *et al*., *Fertil. Steril.* **98**, 1236–1240 (2012).
30. L. Caporali *et al*., *Mitochondrion* **36**, 130–137 (2017).
31. J. Richardson *et al*., *Stem Cells* **33**, 639–645 (2015).
32. A. Greenfield *et al*., *Nat. Biotechnol.* **35**, 1059–1068 (2017).
33. H. Ma *et al*., *Hum. Reprod.* **36**, 493–505 (2021).
34. V. Peeva *et al*., *Nat. Commun.* **9**, 1727 (2018).
35. X.-Y. Fan *et al*., *Nat. Biomed. Eng.* **6**, 339–350 (2022).
36. N. Costa-Borges *et al*., *Fertil. Steril.* S0015-0282(23)00136-X (2023).
37. N. Costa-Borges *et al*., *eLife* **9**, e48591 (2020).
38. J. E. Stern *et al*., *F S Rev.* **3**, 242–255 (2022).
39. D. M. E. I. Hellebrekers *et al*., *Hum. Reprod. Update* **18**, 341–349 (2012).
40. A. T. Clark *et al*., *Stem Cell Reports* **16**, 1416–1424 (2021).
41. S. Franklin, *Comp. Stud. Soc. Hist.* **61**, 743–773 (2019).
42. W. Ombelet *et al*., *Reprod. Biomed. Online* **45**, 481–490 (2022).

ACKNOWLEDGMENTS

We thank M. Aushev for preparing the draft figures from which the finalized figures were adapted and members of the Newcastle Fertility Center for helpful discussions. **Funding:** Funding was provided by Wellcome Trust Human Developmental Biology Initiative grant 215116/Z/18/Z. **Competing interests:** M.H. is a member of the scientific advisory board of Gameto Ltd. The other authors declare that they have no competing interests.

License information: Copyright © 2023 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

Submitted 3 February 2023; accepted 7 March 2023
10.1126/science.adh0073

EDITORIAL

UK science during rapid change



H. Holden Thorp
Editor-in-Chief,
Science journals.
hthorp@aaas.org;
@hholdenthorp

When Ottoline Leyser was appointed as the chief executive of UK Research and Innovation (UKRI) in 2020, she acquired a front-row seat for some extraordinary events in United Kingdom and European politics. She took charge of UKRI—which was forged from multiple agencies to unify government-funded research in all fields—after Brexit and during a time of bold moves in the reorganization of science in the UK, continuous turnover in the UK government, and formidable challenges in coordinating with European science. She sat down with me for a candid talk about these matters and a refreshing willingness to explain the issues.

With all of the turmoil in the UK government in recent years, Leyser already has worked for five secretaries of state and has been part of yet another reorganization earlier this year to form a new Department for Science, Innovation and Technology (DSIT) led by Michelle Donelan as DSIT Secretary of State. Like the Director of the Office of Science, Technology, and Policy (OSTP) in the United States, Donelan fills a cabinet position. This change generally has been well received by the scientific community, although it is unknown whether it will survive the next general election if the Labour Party wins control of the government. Fortunately, unlike the Senate-approved position of OSTP director in the United States, Leyser's appointment is through a public process. Therefore, she retains her position even when a new government takes over. "A key part of my role," she said, "is to ensure that the UK's research and innovation system is healthy in 20 years' time, as well as in 1 year's time." That's an enviable contrast to the "what-have-you-done-for-me-lately" cycle of congressional hearings and political churn in the United States.

The hottest issue for Leyser and her colleagues in government science roles is the association of the UK with Horizon Europe, the main source of research funding in the European Union (EU). Prior to Brexit, UK scientists could acquire funding for collaborations with scientists in the predecessor to Horizon Europe, Horizon 2020. The UK set aside funding, including £6 billion over its first 3 years, to continue these collaborations post-Brexit, but the association was precluded because the EU decided not to explore this relationship until the EU-UK Trade and Cooperation Agreement governing Brexit included terms outlining trade with Northern Ireland. The 2021–2022 funds were returned to the treasury, eliciting howls from scientists who worried that this support for UK science was lost. Fortunately, in February, UK Prime Minister Rishi Sunak and President of the European Commission Ursula von der Leyen reached an agreement on trade, called the Windsor Framework.

This agreement has opened the door for continuing EU-UK research interaction. Still, according to Leyser, hurdles remain, primarily regarding the amount the UK should pay to associate, given that the programme has been running for more than 2 years. Scientific leaders are hoping for the UK's entry into Horizon Europe, but even with the start of negotiations in Brussels last week over the UK's participation in EU research, Michelle Donelan said that there must be a "Plan B" in case of a breakdown. The UK's Plan B, called the Pioneer program, was also unveiled last week, and allots £14.6 billion over 7 years to fund research in the UK as well as support international collaborations.

> **"There is...a huge desire to maintain...strong collaborations with Europe..."**
> **Ottoline Leyser,**
> Chief executive of UKRI

Despite lingering uncertainties, Leyser sees a silver lining to all of this. "I've been pleasantly surprised by how much wide press interest there's been in the UK's association to Horizon Europe," she said, "and I think it signals a much wider, very positive trend of real and deep interest in research and innovation across the UK system."

Leyser is not afraid to be outspoken about the benefits of a strong EU-UK research association, despite the fact that her first two prime ministers and many current members of the UK government are hard-core Brexiteers who proclaimed that the UK would be a scientific superpower while they simultaneously opposed a Northern Ireland agreement that would have allowed scientific collaboration. She told me, "There is...a huge desire to maintain those really close strong collaborations with Europe to be part of the world's biggest research and innovation program and to be able to collaborate and work in a common framework with our European colleagues." At a time when research around the world is becoming dangerously and increasingly nationalized, these are welcome words.

–H. Holden Thorp

PHOTO: CAMERON DAVIDSON

10.1126/science.adi2171

> “It opens the door to another politically motivated lawsuit to get a safe and effective drug off the market.”
>
> **University of Pittsburgh law professor Greer Donley**, in *STAT*, about a U.S. District Court ruling invalidating the 2000 decision by the U.S. Food and Drug Administration to approve the abortion pill mifepristone. The ruling cited safety concerns, although multiple studies have found the pill safe.

IN BRIEF

Edited by **Jeffrey Brainard**



A space-based sensor will provide a fine-grained portrait of North America's smog, such as in Mexico City.

REMOTE SENSING

Satellites will offer details of air pollution

North America will soon get a near-constant, close-up view of its air pollution, courtesy of satellite-borne instruments that include one launched last week. NASA's Tropospheric Emissions Monitoring of Pollution (TEMPO) instrument, riding aboard a commercial communications satellite, will hover 36,000 kilometers above Earth. From its geostationary orbit, matching Earth's rotation, the instrument will enable scientists to make detailed hourly scans of air pollution across the continent, from the oil fields of northern Canada to the skyscrapers of Mexico City. TEMPO can identify the distinct, telltale pattern of light reflected by each of several types of air pollutants, including gases such as ozone, nitrogen dioxide, sulfur dioxide, and bromine, as well as tiny bits of soot. Scientists say the instrument, along with a sibling satellite launched by South Korea in 2020 and another scheduled for launch by European countries in 2024, will be a substantial improvement over the once-a-day images they get now from other satellites. The new instruments could help researchers and regulators zero in on key pollution sources, better model how pollution moves through a region, and identify overlooked hot spots.

U.S. reboots coronavirus research

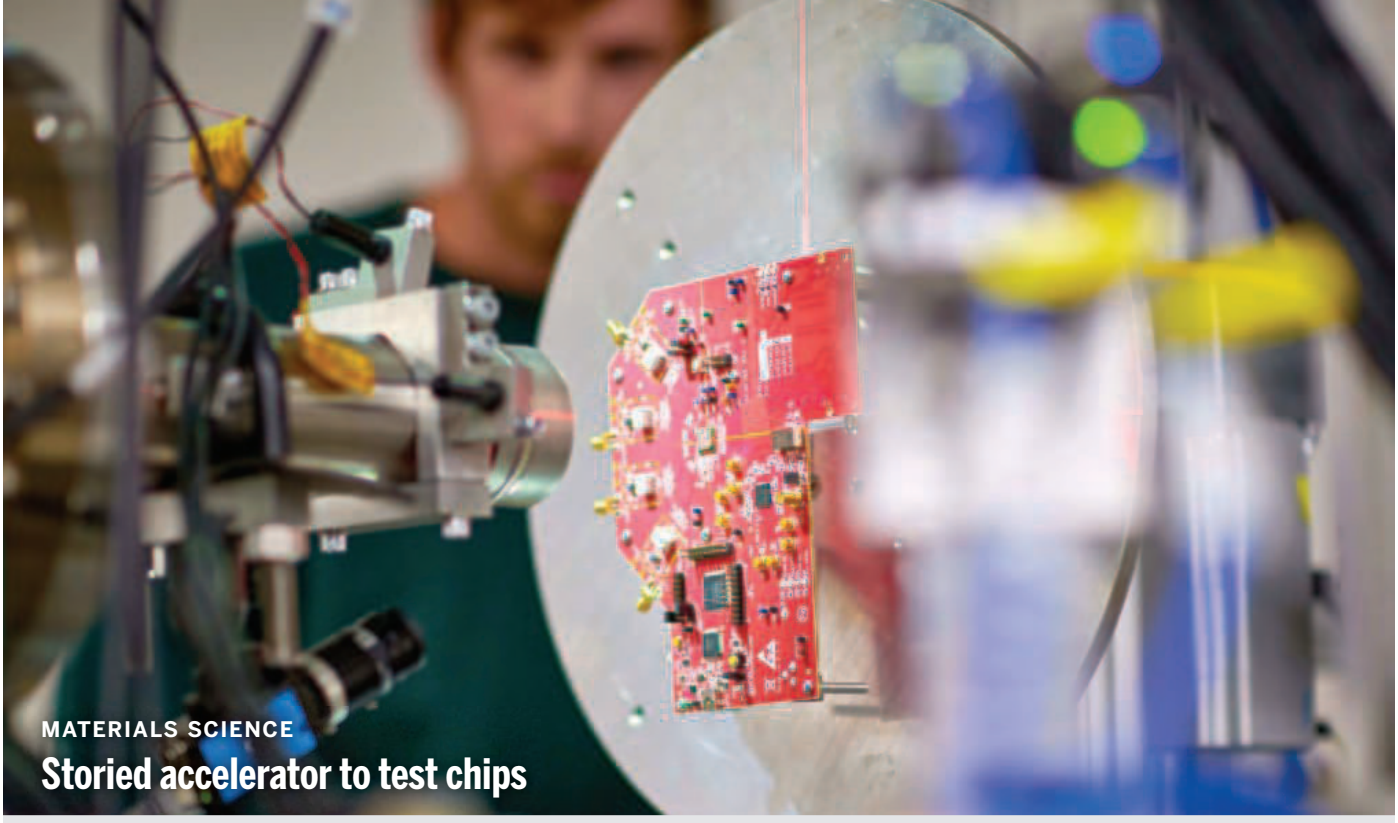
PUBLIC HEALTH | President Joe Biden's administration said this week it plans to spend more than $5 billion to stoke development of better coronavirus vaccines and treatments to curb future pandemics. Like Operation Warp Speed, its predecessor during the administration of former President Donald Trump, the new program, Project NextGen, will rely on public-private partnerships, *The Washington Post* reported. Its top goals include devising upgraded monoclonal antibodies to replace once-potent varieties whose effectiveness against the latest SARS-CoV-2 variants has waned. Another goal is to produce nasal vaccines that prompt immune responses within the body's mucosal linings, potentially eliciting a stronger defense than shots in the arm. The program will also aim to create vaccines against multiple types of coronaviruses.

Lab gear has a heavy carbon cost

SCIENTIFIC COMMUNITY | The energy required to make and transport lab equipment, from centrifuges to mass spectrometers, and provide support services accounts for more than half the carbon emissions associated with research laboratories, according to a preprint study. The resulting carbon footprint was three times greater than that linked with conference travel, normal commuting, or heating, says the analysis, posted last week on the bioRxiv server. Its authors examined purchases by more than 100 laboratories in France and estimated the per capita emissions connected with various items. They also suggest ways to reduce a research group's carbon footprint. Staff members could repair existing equipment to maximize its lifespan, for example. Duplication of equipment could be avoided by concentrating it at regional or national research centers. Researchers could use glassware instead of disposable plastics.

China blocks key databases

RESEARCH SECURITY | A Chinese company has suspended access to four databases widely used by China scholars worldwide because of a new security law that restricts the dissemination of data outside the



MATERIALS SCIENCE

Storied accelerator to test chips

The world's first superconducting cyclotron will receive a new lease on life testing next-generation microchips, Michigan State University's Facility for Rare Isotope Beams (FRIB) announced last week. From 1982 to 2020, the K-500 cyclotron produced beams of atomic nuclei ranging from hydrogen to uranium for experiments in nuclear physics, relying on superconducting magnets to confine the particles. Last year, the cyclotron was replaced by FRIB's new, more powerful, $730 million linear accelerator. Typically, old particle accelerators are demolished or employed to feed their bigger successors. But thanks to a $14 million grant from the Department of Defense, the refurbished K-500 will irradiate chips to see how they perform under challenging conditions resembling those in space.

A refurbished accelerator at Michigan State University will help ease a shortage in U.S. facilities for testing microelectronics.

PHOTO: GENNARA PHOTOGRAPHY, COURTESY OF THE FACILITY FOR RARE ISOTOPE BEAMS

country. The move by the China National Knowledge Infrastructure (CNKI), which took effect on 1 April, covers databases holding national census information, dissertations, governmental statistical yearbooks, and conference proceedings. Its separate Academic Journals Database, which offers the full text of millions of papers in both Chinese and English representing more than 90% of the country's scholarly publications, remains available. In December 2022, Chinese authorities slapped CNKI with an 87.6 million yuan ($13 million) fine for charging unfairly high prices. The company was instructed to reform its monopolistic behavior.

Hydro authority to return remains

ANTHROPOLOGY | The Tennessee Valley Authority (TVA) says it intends to repatriate the remains of nearly 5000 Native Americans whose graves were excavated during dam-building projects in several states from the 1930s to the 1970s. The federally owned hydroelectric utility also said in last month's announcement that it will return to tribal nations thousands of funerary objects also taken during excavations. Many tribal advocates have complained that institutions across the United States have been too slow in returning Native American remains and artifacts, in violation of the 1990 Native American Graves Protection and Repatriation Act. As of 2022, 108,000 sets of remains were being held by museums, universities, and federal agencies, according to the National Park Service. Since 2011, TVA has returned more than 9000 sets of remains.

Genentech finds no fraud inquiry

RESEARCH INTEGRITY | Genentech has found no evidence that its employees alleged or investigated possible research fraud involving a 2009 *Nature* paper coauthored by Stanford University's president, neuroscientist Marc Tessier-Lavigne, contrary to allegations in a February article in the university's student newspaper. Tessier-Lavigne worked at the company at the time; he has led Stanford since 2016. Genentech last week summarized the results of an investigation by its legal team, which interviewed more than 35 current and former employees. The probe did find that the *Nature* paper contains duplicate or composite images. Genentech's summary also describes a frustrating and failed 3-year effort by Genentech scientists to replicate key results from the paper, which suggested a new mechanism for neurodegeneration in Alzheimer's disease. One senior Genentech official wanted the paper retracted or corrected. Stanford's board has been investigating possible misconduct in other papers authored by Tessier-Lavigne before he became Stanford's president.

BY THE NUMBERS

500

Home runs in Major League Baseball since 2010 attributed to anthropogenic global warming. Warmer air is less dense, allowing balls to travel further. But changes in hitters' strategies and balls' stitching account for most of the big increase in home runs during that period. (*Bulletin of the American Meteorological Society*)



OCEANOGRAPHY

Dwindling sea ice may speed melting of Antarctic glaciers

Expanded Ross Gyre would unleash warm water and disrupt ocean "conveyor belt"

By **Paul Voosen**

In February, on an icebreaker off the coast of West Antarctica, Robert Larter, a marine geophysicist with the British Antarctic Survey (BAS), came on deck to a startling sight: open gray water as far as the eye could see. There was no ice at all for the ship to break. The next day, satellite surveys would find sea ice around the continent hitting a record low.

Unlike fast-shrinking Arctic sea ice, the sea ice ringing Antarctica seemed more resistant to climate change—until recently. But now a long-term decline may have set in, and it could have unexpected and ominous domino effects, according to several recent studies. Dwindling sea ice could strengthen a whirling current called the Ross Gyre, bringing warm waters closer to land and hastening the collapse of the West Antarctic ice sheet, which locks up enough water to raise global sea levels by 3.3 meters. The warmer water and glacial melt expected from a stronger gyre already show hints of slowing part of the global ocean's overturning circulation, a critical "conveyor belt" of currents that distributes heat and removes carbon dioxide from the atmosphere. "These are pretty grim predictions," Larter says. "There are a lot of bad consequences if we really are on a downward sea ice trend."

Although climate models predict an eventual decline in Antarctic sea ice, the ice cover was actually expanding slightly until a decade ago. Then from 2014 to 2017, the ice began to vanish rapidly, losing more in 3 years than the Arctic had lost in 3 decades, says Claire Parkinson, a climatologist at NASA's Goddard Space Flight Center. "It was just astonishing." The ice rebounded for a few years, then resumed its downward march, reaching record lows last year and again this year.

This erratic behavior makes it difficult to know whether the ice's long-term decline has begun in earnest, says Clare Eayrs, a sea ice researcher at the Korea Polar Research Institute. "There has been substantial variability in total Antarctic sea ice over the last decade," she says. "We don't know yet why this is." Short-term climate swings like El Niño and fluctuations in the Southern Ocean's roaring winds are probably playing a role, she says. But in some regions—including the seas off West Antarctica—a decline is already unmistakable.

In these waters, hidden beneath sea ice, spins the 1000-kilometer-wide Ross Gyre, fed by warm currents from farther north. Felipe Gomez Valdivia, a physical oceanographer at BAS, and his colleagues wanted to find out how declining sea ice might affect the gyre. Using the UK Earth System Model, a next-generation climate model, they found that as sea ice thinned, the force of surface winds on newly exposed water strengthened the gyre and expanded it, according to a study published last month in *Geophysical Research Letters*. The larger gyre brought warm water closer to the glaciers of West Antarctica, which spill out into the ocean as ice shelves and are vulnerable to melting from below. Within 30 years of its expansion, the gyre could warm the waters beneath the glaciers by 1°C, Valdivia says—an unprecedented heat wave that would accelerate erosion of the ice.

In February, the seas west of the Antarctic Peninsula held almost no ice—a sign of a long-term decline.

Most alarming, the gyre expanded even in scenarios where greenhouse emissions dropped and climate warming slowed. As soon as the sea ice started to decline in the model, the expansion occurred. "It's kind of out of control now because of what humans have already done," Larter says.

Although it is only a single model result, the expansion makes physical sense, says Tiago Dotto, a physical oceanographer at the U.K.'s National Oceanography Centre. This "could play a significant role on the ice shelf melting in West Antarctica," he says. It's quite possible the expansion has already begun: Work from several years ago showed that a strengthening gyre could have con-

tributed to part of the 0.3°C of warming recorded in waters under the ice shelves over the past 20 years.

As the gyre expands, it would also send warm water and freshwater from the melting glaciers away from the ice shelves and toward the Ross Sea, where they could disrupt a process called bottom water formation. Sea ice continually forms and blows away in the Ross Sea. Enriched in salt expelled by the ice and exposed to the frigid wind, the surface waters grow heavy enough to sink into the abyss. This bottom water "factory" is one of the engines for the global overturning circulation, the large-scale flows that shuttle heat and nutrients throughout the world's ocean basins, while also sequestering carbon dioxide in the deep sea.

Arrays of robotic Deep Argo floats, which sink into the depths and then surface to transmit data, have found that these bottom waters have started to warm, according to a study published last year. Moreover, a 63-year record from the Ross Sea, also published in 2022, found that waters close to the surface had steadily become less dense. Both effects could contribute to the faltering of the bottom-water factory and the overturning circulation, Dotto says.

Oceanographers have similar worries about a more famous part of the overturning circulation, in the Atlantic Ocean, but they have not yet detected any weakening in modern observations. But the change in the Southern Ocean is clear, says Sang-Ki Lee, a physical oceanographer at the National Oceanic and Atmospheric Administration's Atlantic Oceanographic and Meteorological Laboratory. Using a climate model calibrated by historical ocean measurements, Lee and his colleagues found that the lower part of the overturning circulation has weakened by up to 20% since the 1970s, according to a paper published last month in *Communications Earth & Environment*. "This is not something that happened suddenly," Lee says. "It gradually happened."

As sea ice declines and the gyre expands, Antarctic bottom water formation could slow further. Late last month in *Nature*, a group led by Qian Li, a physical oceanographer at the Massachusetts Institute of Technology, published a new high-resolution ocean model predicting that, by 2050, the lower cell of the overturning circulation will have weakened by 40% from present times due to increased meltwater. And that means the ocean's ability to mitigate climate change by soaking up nearly one-third of humanity's annual carbon emissions could falter as well, Lee says. "That capacity is reducing." ■

BIOLOGY

An ancient bacterial gene set the stage for human sight

Some 500 million years ago, early vertebrates acquired bacterial DNA that gave rise to a key vision gene

By **Elizabeth Pennisi**

The eye is so complex that even Charles Darwin was at a loss to explain how it could have arisen. Now, it turns out that the evolution of the vertebrate eye got an unexpected boost—from bacteria, which contributed a key gene involved in the retina's response to light. The work, reported this week in the *Proceedings of the National Academy of Sciences*, drives home the evolutionary importance of genes borrowed from other species.

Bacteria are known to readily swap genes with one another, but vertebrates, too, can incorporate microbial genes. When the human genome was first sequenced in 2001, scientists thought it contained about 200 bacteria-derived genes, though many of these turned out to be the result of contamination.

Hoping to improve on those earlier efforts, Matthew Daugherty, a biochemist at the University of California, San Diego, and colleagues used computer software to trace the evolution of hundreds of human genes by searching for similar sequences in other species. Genes that seemed to have appeared first in vertebrates and had no predecessors in earlier animals were good candidates for having jumped from bacteria, particularly if they had counterparts in modern microbes. Among the dozens of potentially alien genes, one "blew me away," Daugherty recalls.

The gene, called *IRBP* (interphotoreceptor retinoid-binding protein), was already known to be important for seeing. The protein it encodes resides in the space between the retina and the retinal pigment epithelium, a thin layer of cells overlying the retina.

When light hits a light-sensitive photoreceptor in the retina of the vertebrate eye, vitamin A complexes become kinked, setting off an electrical pulse that activates the optic nerve. *IRBP* then shifts these molecules to the epithelium to be unkinked. Finally, it shuttles the restored molecules back to the photoreceptor. "*IRBP* is essential for the vision of all vertebrates," explains Ling Zhu, a retinal biologist at University of Sydney's Save Sight Institute who was not involved with the work.



A bacterial gene helps make vertebrate vision possible.

PHOTO: INGO ARNDT/SCIENCE SOURCE

Vertebrate *IRBP* most closely resembles a class of bacterial genes called pepsidases, whose proteins recycle other proteins. Since *IRBP* is found in all vertebrates but generally not in their closest invertebrate relatives, Daugherty and his colleagues propose that more than 500 million years ago microbes transferred a pepsidase gene into an ancestor of all living vertebrates.

Once the gene was in place, the protein's recycling function was lost and the gene duplicated itself twice, making *IRBP* bigger, with four copies of the original DNA strung together. Even in its microbial forebears, this protein may have had some ability to bind to light-sensing molecules, Daugherty suggests. Other mutations completed its transformation into a molecule that could escape from cells and serve as a shuttle.

Not everyone agrees that the evolution of *IRBP* was crucial for vertebrate vision. "I don't think it had to happen" in order for vertebrates to see well, says Sönke Johnsen, a biologist at Duke University. Invertebrate eyes make do without *IRBP*, he notes.

Daugherty agrees that vertebrates' reliance on *IRBP* could simply be a historical accident. "We are sort of stuck with it," he says.

Either way, the work supports the idea that horizontal gene transfer can help to endow organisms with new functions, says Julie Dunning Hotopp, a genome biologist at the University of Maryland School of Medicine's Institute for Genome Sciences. Once these genes take root in a new species, evolution can tinker with them to produce totally new abilities or enhance existing ones. "It is the biological equivalent of upcycling that happens in my Buy Nothing Group." ■

tributed to part of the 0.3°C of warming recorded in waters under the ice shelves over the past 20 years.

As the gyre expands, it would also send warm water and freshwater from the melting glaciers away from the ice shelves and toward the Ross Sea, where they could disrupt a process called bottom water formation. Sea ice continually forms and blows away in the Ross Sea. Enriched in salt expelled by the ice and exposed to the frigid wind, the surface waters grow heavy enough to sink into the abyss. This bottom water "factory" is one of the engines for the global overturning circulation, the large-scale flows that shuttle heat and nutrients throughout the world's ocean basins, while also sequestering carbon dioxide in the deep sea.

Arrays of robotic Deep Argo floats, which sink into the depths and then surface to transmit data, have found that these bottom waters have started to warm, according to a study published last year. Moreover, a 63-year record from the Ross Sea, also published in 2022, found that waters close to the surface had steadily become less dense. Both effects could contribute to the faltering of the bottom-water factory and the overturning circulation, Dotto says.

Oceanographers have similar worries about a more famous part of the overturning circulation, in the Atlantic Ocean, but they have not yet detected any weakening in modern observations. But the change in the Southern Ocean is clear, says Sang-Ki Lee, a physical oceanographer at the National Oceanic and Atmospheric Administration's Atlantic Oceanographic and Meteorological Laboratory. Using a climate model calibrated by historical ocean measurements, Lee and his colleagues found that the lower part of the overturning circulation has weakened by up to 20% since the 1970s, according to a paper published last month in *Communications Earth & Environment*. "This is not something that happened suddenly," Lee says. "It gradually happened."

As sea ice declines and the gyre expands, Antarctic bottom water formation could slow further. Late last month in *Nature*, a group led by Qian Li, a physical oceanographer at the Massachusetts Institute of Technology, published a new high-resolution ocean model predicting that, by 2050, the lower cell of the overturning circulation will have weakened by 40% from present times due to increased meltwater. And that means the ocean's ability to mitigate climate change by soaking up nearly one-third of humanity's annual carbon emissions could falter as well, Lee says. "That capacity is reducing." ■

BIOLOGY

An ancient bacterial gene set the stage for human sight

Some 500 million years ago, early vertebrates acquired bacterial DNA that gave rise to a key vision gene

By **Elizabeth Pennisi**

The eye is so complex that even Charles Darwin was at a loss to explain how it could have arisen. Now, it turns out that the evolution of the vertebrate eye got an unexpected boost—from bacteria, which contributed a key gene involved in the retina's response to light. The work, reported this week in the *Proceedings of the National Academy of Sciences*, drives home the evolutionary importance of genes borrowed from other species.

Bacteria are known to readily swap genes with one another, but vertebrates, too, can incorporate microbial genes. When the human genome was first sequenced in 2001, scientists thought it contained about 200 bacteria-derived genes, though many of these turned out to be the result of contamination.



A bacterial gene helps make vertebrate vision possible.

Hoping to improve on those earlier efforts, Matthew Daugherty, a biochemist at the University of California, San Diego, and colleagues used computer software to trace the evolution of hundreds of human genes by searching for similar sequences in other species. Genes that seemed to have appeared first in vertebrates and had no predecessors in earlier animals were good candidates for having jumped from bacteria, particularly if they had counterparts in modern microbes. Among the dozens of potentially alien genes, one "blew me away," Daugherty recalls.

The gene, called *IRBP* (interphotoreceptor retinoid-binding protein), was already known to be important for seeing. The protein it encodes resides in the space between the retina and the retinal pigment epithelium, a thin layer of cells overlying the retina.

When light hits a light-sensitive photoreceptor in the retina of the vertebrate eye, vitamin A complexes become kinked, setting off an electrical pulse that activates the optic nerve. *IRBP* then shifts these molecules to the epithelium to be unkinked. Finally, it shuttles the restored molecules back to the photoreceptor. "*IRBP* is essential for the vision of all vertebrates," explains Ling Zhu, a retinal biologist at University of Sydney's Save Sight Institute who was not involved with the work.

Vertebrate *IRBP* most closely resembles a class of bacterial genes called pepsidases, whose proteins recycle other proteins. Since *IRBP* is found in all vertebrates but generally not in their closest invertebrate relatives, Daugherty and his colleagues propose that more than 500 million years ago microbes transferred a pepsidase gene into an ancestor of all living vertebrates.

Once the gene was in place, the protein's recycling function was lost and the gene duplicated itself twice, making *IRBP* bigger, with four copies of the original DNA strung together. Even in its microbial forebears, this protein may have had some ability to bind to light-sensing molecules, Daugherty suggests. Other mutations completed its transformation into a molecule that could escape from cells and serve as a shuttle.

Not everyone agrees that the evolution of *IRBP* was crucial for vertebrate vision. "I don't think it had to happen" in order for vertebrates to see well, says Sönke Johnsen, a biologist at Duke University. Invertebrate eyes make do without *IRBP*, he notes.

Daugherty agrees that vertebrates' reliance on *IRBP* could simply be a historical accident. "We are sort of stuck with it," he says.

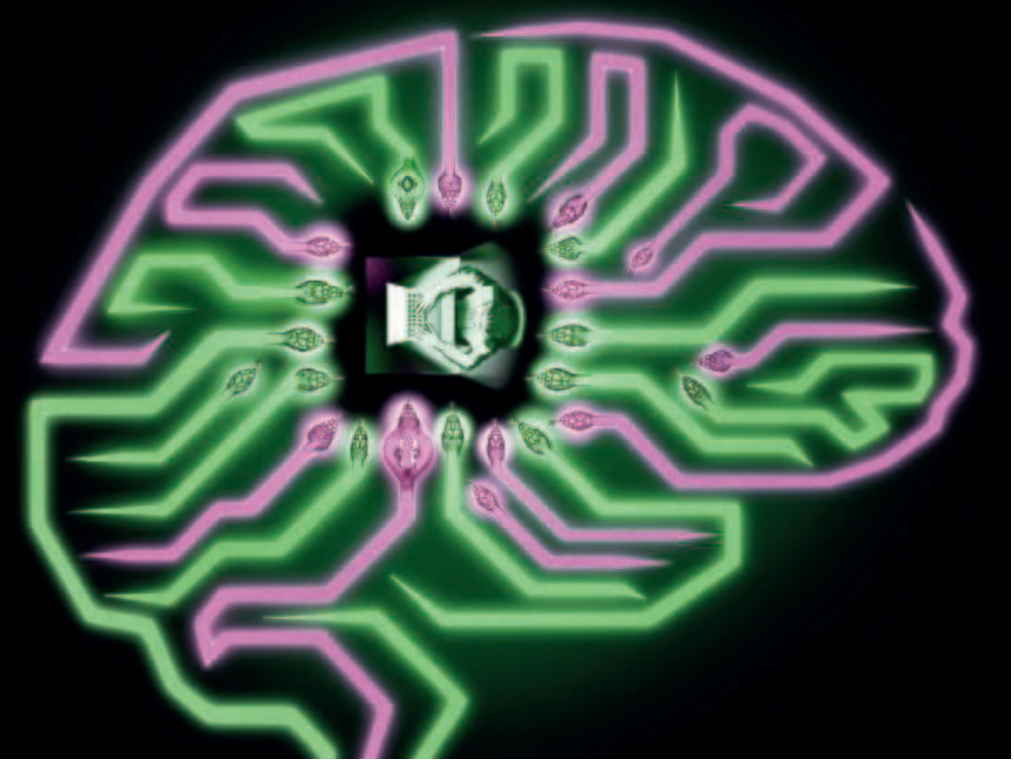
Either way, the work supports the idea that horizontal gene transfer can help to endow organisms with new functions, says Julie Dunning Hotopp, a genome biologist at the University of Maryland School of Medicine's Institute for Genome Sciences. Once these genes take root in a new species, evolution can tinker with them to produce totally new abilities or enhance existing ones. "It is the biological equivalent of upcycling that happens in my Buy Nothing Group." ■

PHOTO: INGO ARNDT/SCIENCE SOURCE

COMPUTER SCIENCE

Call for AI pause highlights potential dangers

As systems dazzle, researchers worry about lack of safeguards and regulation



By **Laurie Clarke**

An open letter calling for a pause on the development of advanced artificial intelligence (AI) systems has divided researchers. Attracting signatures from the likes of Tesla CEO Elon Musk and Apple co-founder Steve Wozniak, the letter, released early last week, advocates for a 6-month moratorium to give AI companies and regulators time to formulate safeguards to protect society from potential risks of the technology.

AI has galloped along since the launch last year of the image generator DALL-E 2, from the Microsoft-backed company OpenAI. The company has since released ChatGPT and GPT-4, two text-generating chatbots, to frenzied acclaim. The ability of these so-called "generative" models to mimic human outputs, combined with the speed of adoption—ChatGPT reportedly reached more than 100 million users by January and major tech companies are racing to build generative AI into their products—have caught many off guard.

"I think many people's intuitions about the impact of technology aren't well calibrated to the pace and scale of [these] AI models," says letter signatory Michael Osborne, a machine learning researcher and co-founder of AI company Mind Foundry. He is worried about the societal impacts of the new tools, including their potential to put people out of work and proliferate disinformation. "I feel that a 6-month pause would … give regulators enough time to catch up with the rapid pace of advances," he says.

The letter, released by a nonprofit organization called the Future of Life Initiative, rankles some researchers by invoking far-off, speculative harms. It asks, "Should we develop nonhuman minds that might eventually outnumber, outsmart, obsolete and replace us? Should we risk loss of control of our civilization?" Sandra Wachter, an expert in technology regulation at the University of Oxford, says there are many known harms that need addressing today. Wachter, who didn't sign the letter, says the focus should be on how AI systems can be disinformation engines, persuading people of incorrect, potentially libelous information; how they perpetuate systemic bias in the information they surface to people; and how they rely on the invisible labor of workers, often toiling under poor conditions, to label data and train the systems.

Privacy is another emerging concern, as critics worry that systems could be prompted to exactly reproduce personally identifiable information from their training sets. Italy's data protection authority banned ChatGPT on 31 March over concerns that Italians' personal data are being used to train OpenAI's models. (An OpenAI blog post says, "We work to remove personal information from the training dataset where feasible, fine-tune models to reject requests for personal information of private individuals, and respond to requests from individuals to delete their personal information from our systems.")

Some technologists warn of deeper security threats. Planned ChatGPT-based digital assistants that can interface with the web and read and write emails could offer new opportunities for hackers, says Florian Tramèr, a computer scientist at ETH Zürich. Already, hackers rely on a tactic called "prompt injection" to trick AI models into saying things they shouldn't, like offering advice on how to carry out illegal activities. Some methods involve asking the tool to roleplay as an evil confidant, or act as a translator between different languages, which can confuse the model and prompt it to disregard its safety restrictions.

Tramèr worries the practice could evolve into a way for hackers to trick the digital assistants through "indirect prompt injection"—by, for example, sending someone a calendar invitation with instructions for the assistant to export the recipient's data and send it to the hacker. "These models are just going to get exploited left and right to leak people's private information or to destroy their data," he says. He says AI companies need to start warning users of the security and privacy risks, and do more to address them.

OpenAI seems to be becoming more alert to security risks. OpenAI President and co-founder Greg Brockman tweeted last month that the company is "considering starting a bounty program" for hackers who flag weaknesses in its AI systems, acknowledging that the stakes "will go up a *lot* over time."

However, many of the problems inherent in today's AI models don't have easy

solutions. One vexing issue is how to make AI-generated content identifiable. Some researchers are working on "watermarking"—creating an imperceptible digital signature in the AI's output. Others are trying to devise means of detecting patterns that only AI produces. However, recent research found that tools that slightly rephrase AI-produced text can significantly undermine both approaches. As AI begins to sound more human, the authors say, its output will only become harder to detect.

Other elusive safeguards include ones to prevent systems from generating violent or pornographic images. Tramèr says most researchers are simply applying after-the-fact filters, teaching the AI to avoid "bad" outputs. He believes these issues need to be remedied before training, at the data level. "We need to find better ways of curating the training sets of these generative models to remove sensitive data altogether," he says.

The pause itself seems unlikely to happen. OpenAI CEO Sam Altman didn't sign the letter, telling *The Wall Street Journal* that the company has always taken safety seriously, and regularly collaborates with the industry on safety standards. Microsoft co-founder Bill Gates told Reuters the proposed pause won't "solve the challenges" ahead.

Osborne believes governments will need to step in. "We can't rely on the tech giants to self-regulate," he says. The Biden administration has proposed an AI "Bill of Rights" designed to help businesses develop safe AI systems that protect the rights of U.S. citizens—but the principles are voluntary and nonbinding. The European Union's AI Act, which is expected to come into force this year, will apply different levels of regulation depending on the level of risk. For example, policing systems that aim to predict individual crimes are considered unacceptably risky, and are therefore banned.

Wachter says a 6-month pause appears arbitrary, and she is leery of banning research. Instead, "we need to go back and think about responsible research and embed that type of thinking very early on," she says. As a part of this, she says companies should invite independent experts to hack and stress test their systems before rolling them out.

She notes the people behind the letter are heavily immersed in the tech world, which she thinks gives them a narrow perspective on the potential risks. "You really need to talk to lawyers, to people who do ethics, to people who understand economics and politics," she says. "The most important thing is that those questions are not decided among tech people alone." ■

Laurie Clarke is a journalist in the United Kingdom.

ARCHAEOLOGY

In Miami, a modern clash over a 2000-year-old settlement

High-rises planned for "extraordinary" Native American site

By **David Malakoff**, *in Miami*

For nearly 2 years, bustling teams of archaeologists—totaling more than 120 researchers on some days—have excavated a sprawling waterfront lot here in one of this seaside city's toniest neighborhoods. The diggers have uncovered ancient human remains and some 1 million artifacts that are providing rare insight into a major Native American settlement, known as Tequesta, which flourished at the mouth of the Miami River some 2000 years ago.

"It's a truly extraordinary site," says archaeologist Traci Ardren of the University of Miami (UM), who is not part of the dig. Notable discoveries, she says, include ancient wooden tools and plant materials that rarely survive in Florida's subtropical climate, as well as copper and stone artifacts that indicate Tequesta conducted a vibrant trade with distant tribes. Researchers know relatively little about Miami's prehistoric founders, says Ardren, who expects the finds to "add a great deal to our understanding."

But the dig—which city rules required a developer to conduct before building three luxury high-rises—has also fueled a fierce conflict over how Miami, which boosters tout as a "new city," should recognize and preserve its ancient past. A loose alliance of archaeologists, Indigenous people, and preservation advocates want the city to require the developer to transfer the vast artifact trove to a university or museum, and help pay for efforts to study the collection and share findings with the public. "We can't just let this important collection languish in some warehouse," says UM archaeologist William Pestle, one of the leaders of the campaign.

For its part, the Related Group, the site's developer, says it has already spent about $20 million on the dig and is willing to spend millions more—but only if the city doesn't derail its plan to build some 1400 residential units.

Last week, the issue came to a head at a contentious 5-hour hearing before Miami's historic preservation board, which was considering proposals to place historic designations on the site that could hamper construction. An overflow crowd clapped and jeered as some 50 speakers supported or opposed designation. Members of Native American tribes expressed concerns that the developer had not adequately consulted with Indigenous people about its plans, as well as about Florida's secretive process for deciding what to do with ancient human remains uncovered by archaeologists, such as those found on the site. "There's too little transparency around how our ancestors are being treated," says Betty Osceola, a member of the Panther Clan of the Miccosukee Tribe.

In the end, both the preservationists and the developer emerged from the 4 April hearing claiming partial victory. But all sides agree the story is far from over.

"What happens next will be the real test of whether Miami is ready to stop ignoring its past," Pestle says.

The flatlands surrounding the Miami River's junction with Biscayne Bay have long been prime real estate. Native Americans began settling in the area at least 4000 years ago, researchers say, attracted by abundant fresh water and food. "It's a classic location for an early Florida occupation," says archaeologist Robert Carr, who has been studying the region for more than 40 years and is head of the Archaeological and Historical Conservancy, a Florida nonprofit that is conducting the current dig. Archaeologists believe Tequesta, which straddled both banks of the Miami River, became the capital of a chiefdom that stretched across southeast Florida from roughly 500 B.C.E. to the 1500s C.E.

Today, Tequesta's remains sit beneath some of Miami's most valuable land, and over the past 25 years a string of building projects have sparked controversy. In 1998, archaeologists conducting a predevelopment dig for a luxury high-rise just a few blocks from the current site discovered a large circle of post holes in the limestone bedrock—likely the remains of a ceremonial structure built about 2000 years ago. After an outcry, state and local officials spent $27 million to buy and protect what is now known as the Miami Circle.

Since then, preconstruction digs at other nearby sites have uncovered hundreds of human burials and archaeological features

solutions. One vexing issue is how to make AI-generated content identifiable. Some researchers are working on "watermarking"—creating an imperceptible digital signature in the AI's output. Others are trying to devise means of detecting patterns that only AI produces. However, recent research found that tools that slightly rephrase AI-produced text can significantly undermine both approaches. As AI begins to sound more human, the authors say, its output will only become harder to detect.

Other elusive safeguards include ones to prevent systems from generating violent or pornographic images. Tramèr says most researchers are simply applying after-the-fact filters, teaching the AI to avoid "bad" outputs. He believes these issues need to be remedied before training, at the data level. "We need to find better ways of curating the training sets of these generative models to remove sensitive data altogether," he says.

The pause itself seems unlikely to happen. OpenAI CEO Sam Altman didn't sign the letter, telling *The Wall Street Journal* that the company has always taken safety seriously, and regularly collaborates with the industry on safety standards. Microsoft co-founder Bill Gates told Reuters the proposed pause won't "solve the challenges" ahead.

Osborne believes governments will need to step in. "We can't rely on the tech giants to self-regulate," he says. The Biden administration has proposed an AI "Bill of Rights" designed to help businesses develop safe AI systems that protect the rights of U.S. citizens—but the principles are voluntary and nonbinding. The European Union's AI Act, which is expected to come into force this year, will apply different levels of regulation depending on the level of risk. For example, policing systems that aim to predict individual crimes are considered unacceptably risky, and are therefore banned.

Wachter says a 6-month pause appears arbitrary, and she is leery of banning research. Instead, "we need to go back and think about responsible research and embed that type of thinking very early on," she says. As a part of this, she says companies should invite independent experts to hack and stress test their systems before rolling them out.

She notes the people behind the letter are heavily immersed in the tech world, which she thinks gives them a narrow perspective on the potential risks. "You really need to talk to lawyers, to people who do ethics, to people who understand economics and politics," she says. "The most important thing is that those questions are not decided among tech people alone." ■

Laurie Clarke is a journalist in the United Kingdom.

ARCHAEOLOGY

In Miami, a modern clash over a 2000-year-old settlement

High-rises planned for "extraordinary" Native American site

By **David Malakoff**, *in Miami*

For nearly 2 years, bustling teams of archaeologists—totaling more than 120 researchers on some days—have excavated a sprawling waterfront lot here in one of this seaside city's toniest neighborhoods. The diggers have uncovered ancient human remains and some 1 million artifacts that are providing rare insight into a major Native American settlement, known as Tequesta, which flourished at the mouth of the Miami River some 2000 years ago.

"It's a truly extraordinary site," says archaeologist Traci Ardren of the University of Miami (UM), who is not part of the dig. Notable discoveries, she says, include ancient wooden tools and plant materials that rarely survive in Florida's subtropical climate, as well as copper and stone artifacts that indicate Tequesta conducted a vibrant trade with distant tribes. Researchers know relatively little about Miami's prehistoric founders, says Ardren, who expects the finds to "add a great deal to our understanding."

But the dig—which city rules required a developer to conduct before building three luxury high-rises—has also fueled a fierce conflict over how Miami, which boosters tout as a "new city," should recognize and preserve its ancient past. A loose alliance of archaeologists, Indigenous people, and preservation advocates want the city to require the developer to transfer the vast artifact trove to a university or museum, and help pay for efforts to study the collection and share findings with the public. "We can't just let this important collection languish in some warehouse," says UM archaeologist William Pestle, one of the leaders of the campaign.

For its part, the Related Group, the site's developer, says it has already spent about $20 million on the dig and is willing to spend millions more—but only if the city doesn't derail its plan to build some 1400 residential units.

Last week, the issue came to a head at a contentious 5-hour hearing before Miami's historic preservation board, which was considering proposals to place historic designations on the site that could hamper construction. An overflow crowd clapped and jeered as some 50 speakers supported or opposed designation. Members of Native American tribes expressed concerns that the developer had not adequately consulted with Indigenous people about its plans, as well as about Florida's secretive process for deciding what to do with ancient human remains uncovered by archaeologists, such as those found on the site. "There's too little transparency around how our ancestors are being treated," says Betty Osceola, a member of the Panther Clan of the Miccosukee Tribe.

In the end, both the preservationists and the developer emerged from the 4 April hearing claiming partial victory. But all sides agree the story is far from over.

"What happens next will be the real test of whether Miami is ready to stop ignoring its past," Pestle says.

The flatlands surrounding the Miami River's junction with Biscayne Bay have long been prime real estate. Native Americans began settling in the area at least 4000 years ago, researchers say, attracted by abundant fresh water and food. "It's a classic location for an early Florida occupation," says archaeologist Robert Carr, who has been studying the region for more than 40 years and is head of the Archaeological and Historical Conservancy, a Florida nonprofit that is conducting the current dig. Archaeologists believe Tequesta, which straddled both banks of the Miami River, became the capital of a chiefdom that stretched across southeast Florida from roughly 500 B.C.E. to the 1500s C.E.

Today, Tequesta's remains sit beneath some of Miami's most valuable land, and over the past 25 years a string of building projects have sparked controversy. In 1998, archaeologists conducting a predevelopment dig for a luxury high-rise just a few blocks from the current site discovered a large circle of post holes in the limestone bedrock—likely the remains of a ceremonial structure built about 2000 years ago. After an outcry, state and local officials spent $27 million to buy and protect what is now known as the Miami Circle.

Since then, preconstruction digs at other nearby sites have uncovered hundreds of human burials and archaeological features



Archaeologists have found extensive evidence of the ancient settlement of Tequesta at this Miami building site.

related to Tequesta, but preservation advocates complain that the city did relatively little to protect or publicize these finds. "Miami has a troubling history of allowing developers to pave over its indigenous history," says archaeologist Sara Ayers-Rigsby of the Florida Public Archaeology Network.

The current controversy involves a 1.5-hectare parcel divided into three lots; two have been cleared for construction and one still hosts an older office building. In early 2021, when Carr began examining the cleared lots, he fully expected to find signs of Tequesta. But the discoveries proved "more extensive than anything we've ever seen," he said at the hearing.

Below the water table, where low oxygen levels slowed decay, diggers found strips of cord woven from fragile plant material and three wooden artifacts, including a 20-centimeter-long pestle or club and a bowl-shaped object used for starting fires. Researchers also uncovered celts, or ax heads, made from shell and stone, including some made from basalt that likely came from the distant Appalachians. And they recovered decorative beads made from copper and galena that likely originated in central North America. Such exotic items might have been used for ceremonies or tribute, and suggest Tequesta was a trade hub with "a very complicated, stratified society," Carr said.

The site also has yielded nearly 1 ton of animal bone, including the remains of extinct Caribbean monk seals and whales. They suggest the Tequestans "had some pretty sophisticated fishing technologies," Ardren says, perhaps even hunting large marine mammals from their dugout canoes.

Several hundred post holes suggest the site was a major residential area, with people living on platforms suspended above the damp ground. Dating studies indicate a peak of activity between 100 C.E. and 400 C.E., around the time the nearby Miami Circle was constructed. That period "might have been the golden age of Tequesta," which likely had thousands of residents, Carr said.

At last week's hearing, some speakers called on the city to simply preserve the site. But legal experts say that is unlikely, especially given that construction has already begun on one lot. Others, including archaeologists, asked the board to designate the remaining two lots as historic. They argued that would give the city greater legal leverage to require the developer to care for artifacts, develop interpretive programs, and possibly even modify construction plans.

The Related Group and its allies opposed designation, asserting it could cause "immense" financial harm to the firm—and hinting it would fight the move in court. But after hours of discussion, the developer didn't object to the board's 8-0 vote to consider designation for the undeveloped lot. The board, in turn, dropped a proposal to consider designating the cleared lots, voting instead to give the developer time to work with archaeologists, universities, and others to devise an enforceable "action plan" for preservation and interpretation.

Board members also urged the developer to do more to consult with Native American tribes. The Related Group says it complied with Florida rules that empower a panel of Indigenous representatives to approve plans for ancient human remains found at digs; in this case, they will be reburied on the site. But those consultations were too limited and opaque, Osceola and others say.

The outcomes didn't please everyone. But Pestle says the packed hearing "at least showed that people do care about preserving Miami's history—and served notice that we'll be watching what happens next." ■

BIOMEDICINE

Report on trial death stokes Alzheimer's drug fears

Amyloid-clearing antibody lecanemab faces key FDA hearing in June

By **Charles Piller**

A full autopsy and detailed examination of the brain of a 79-year-old Florida woman who died after receiving lecanemab, an experimental Alzheimer's therapy, in a pivotal clinical trial has deepened some researchers' concerns that it poses serious risks for patients who share the woman's hard-to-diagnose, preexisting condition.

The patient's history and autopsy "strongly suggests that lecanemab infusions were a catalyst leading to the events resulting in her death," says Vanderbilt University pathologist Hannah Harmsen, co-author of a recently completed case report, which *Science* has obtained. The authors say the woman's fatal brain swelling and hemorrhaging likely resulted when the drug, a monoclonal antibody, attacked the abnormal protein deposits that had built up in her blood vessels.

At the end of March, worries about such side effects led some prominent Alzheimer's scientists to urge more stringent prescribing criteria for lecanemab than the Food and Drug Administration (FDA) signed off on in 2022 when giving it provisional approval based on biomarker evidence. But some authors of the new case report and others who have studied lecanemab's side effects argue that the proposed "Appropriate Use Recommendations," from researchers who all have current or recent ties to the antibody's makers, still don't go far enough to protect potential recipients.

The case report and the experts' recommendations are likely to be hot topics at an FDA advisory panel meeting on the drug, which the agency this week said would be on 9 June. The panel will assess the drug's pros and cons before FDA makes its final decision on full approval, likely in July.

Lecanemab and another recently ap-



Archaeologists have found extensive evidence of the ancient settlement of Tequesta at this Miami building site.

related to Tequesta, but preservation advocates complain that the city did relatively little to protect or publicize these finds. "Miami has a troubling history of allowing developers to pave over its indigenous history," says archaeologist Sara Ayers-Rigsby of the Florida Public Archaeology Network.

The current controversy involves a 1.5-hectare parcel divided into three lots; two have been cleared for construction and one still hosts an older office building. In early 2021, when Carr began examining the cleared lots, he fully expected to find signs of Tequesta. But the discoveries proved "more extensive than anything we've ever seen," he said at the hearing.

Below the water table, where low oxygen levels slowed decay, diggers found strips of cord woven from fragile plant material and three wooden artifacts, including a 20-centimeter-long pestle or club and a bowl-shaped object used for starting fires. Researchers also uncovered celts, or ax heads, made from shell and stone, including some made from basalt that likely came from the distant Appalachians. And they recovered decorative beads made from copper and galena that likely originated in central North America. Such exotic items might have been used for ceremonies or tribute, and suggest Tequesta was a trade hub with "a very complicated, stratified society," Carr said.

The site also has yielded nearly 1 ton of animal bone, including the remains of extinct Caribbean monk seals and whales. They suggest the Tequestans "had some pretty sophisticated fishing technologies," Ardren says, perhaps even hunting large marine mammals from their dugout canoes.

Several hundred post holes suggest the site was a major residential area, with people living on platforms suspended above the damp ground. Dating studies indicate a peak of activity between 100 C.E. and 400 C.E., around the time the nearby Miami Circle was constructed. That period "might have been the golden age of Tequesta," which likely had thousands of residents, Carr said.

At last week's hearing, some speakers called on the city to simply preserve the site. But legal experts say that is unlikely, especially given that construction has already begun on one lot. Others, including archaeologists, asked the board to designate the remaining two lots as historic. They argued that would give the city greater legal leverage to require the developer to care for artifacts, develop interpretive programs, and possibly even modify construction plans.

The Related Group and its allies opposed designation, asserting it could cause "immense" financial harm to the firm—and hinting it would fight the move in court. But after hours of discussion, the developer didn't object to the board's 8-0 vote to consider designation for the undeveloped lot. The board, in turn, dropped a proposal to consider designating the cleared lots, voting instead to give the developer time to work with archaeologists, universities, and others to devise an enforceable "action plan" for preservation and interpretation.

Board members also urged the developer to do more to consult with Native American tribes. The Related Group says it complied with Florida rules that empower a panel of Indigenous representatives to approve plans for ancient human remains found at digs; in this case, they will be reburied on the site. But those consultations were too limited and opaque, Osceola and others say.

The outcomes didn't please everyone. But Pestle says the packed hearing "at least showed that people do care about preserving Miami's history—and served notice that we'll be watching what happens next." ■

BIOMEDICINE

Report on trial death stokes Alzheimer's drug fears

Amyloid-clearing antibody lecanemab faces key FDA hearing in June

By **Charles Piller**

A full autopsy and detailed examination of the brain of a 79-year-old Florida woman who died after receiving lecanemab, an experimental Alzheimer's therapy, in a pivotal clinical trial has deepened some researchers' concerns that it poses serious risks for patients who share the woman's hard-to-diagnose, preexisting condition.

The patient's history and autopsy "strongly suggests that lecanemab infusions were a catalyst leading to the events resulting in her death," says Vanderbilt University pathologist Hannah Harmsen, co-author of a recently completed case report, which *Science* has obtained. The authors say the woman's fatal brain swelling and hemorrhaging likely resulted when the drug, a monoclonal antibody, attacked the abnormal protein deposits that had built up in her blood vessels.

At the end of March, worries about such side effects led some prominent Alzheimer's scientists to urge more stringent prescribing criteria for lecanemab than the Food and Drug Administration (FDA) signed off on in 2022 when giving it provisional approval based on biomarker evidence. But some authors of the new case report and others who have studied lecanemab's side effects argue that the proposed "Appropriate Use Recommendations," from researchers who all have current or recent ties to the antibody's makers, still don't go far enough to protect potential recipients.

The case report and the experts' recommendations are likely to be hot topics at an FDA advisory panel meeting on the drug, which the agency this week said would be on 9 June. The panel will assess the drug's pros and cons before FDA makes its final decision on full approval, likely in July.

Lecanemab and another recently ap-

proved drug called aducanumab belong to a class of antibodies that bind to beta amyloid, a protein that is widely thought to drive the progression of Alzheimer's disease. In January, FDA approved lecanemab for use after clinical trials showed it removed brain amyloid deposits, an effect the agency said was "reasonably likely to predict a clinical benefit." The agency relied on the same reasoning in its 2021 approval of aducanumab.

For its "accelerated" approval of lecanemab, FDA did not consider what many saw as the key evidence from the antibody's phase 3 trial. After 18 months of use, lecanemab modestly slowed the devastating cognitive deterioration of Alzheimer's—the first drug to clearly do so.

But the drug, now marketed as Leqembi by its maker Eisai Co., has also been linked to several deaths and brain injuries, as well as to a reduction in brain volume, during its trials or their extensions. The case report, currently under review at a journal, expands on an earlier preliminary analysis of the Florida woman's death, which took place in the phase 3 trial's extension and was first reported by *Science*.

That initial analysis showed the woman ultimately died from multiorgan failure and pneumonia, 5 days after she began having seizures, brain swelling, and bleeding. The woman had early Alzheimer's disease and moderate, asymptomatic coronary artery disease, but had otherwise been in good health when she joined the trial.

Like all Alzheimer's patients, the woman was also at risk of a condition called cerebral amyloid angiopathy (CAA). In Alzheimer's disease, beta amyloid accumulates between brain cells in hallmark deposits called plaques. But in roughly half of patients, the protein also extensively replaces the smooth muscle of blood vessel walls, rendering them brittle. Antibodies such as lecanemab often strip away this amyloid, making the vessels weak, inflamed, and prone to dangerous bleeding. Other people who died during the lecanemab trial were taking blood thinners, which likely increased their risk. But the Florida woman was not.

It takes sophisticated, expensive brain imaging to detect signs of CAA; only an autopsy can provide a definitive diagnosis. The case report, prepared by neuroscientists and pathologists at Vanderbilt, Boston University, the University of Florida, and a private Florida autopsy company, showed clear evidence of CAA: many brain hemorrhages, and cerebral blood vessels lined with amyloid plaques. The examination of the woman's brain revealed "remarkable" inflammation and degeneration of blood vessels, says Matthew Schrag, a Vanderbilt neurologist and neuroscientist and the case report's senior author.

Besides weakening blood vessel walls, lecanemab and other antiamyloid antibodies may pose another hazard to people with CAA, Harmsen, Schrag, and their colleagues say. As the antibodies remove clumps of amyloid from brain tissue, the agglomerations may clog blood vessels and lead to "an inflammatory cascade" that increases the risk of life-threatening swelling and bleeding.

According to Northwestern University pathologist Rudolph Castellani, the report provides definitive evidence that lecanemab started the chain of events that resulted in the woman's death. "Leqembi is toxic to the cerebral vasculature in people with CAA and potentially lethal," says Castellani, who previously reported on a separate lecanemab clinical trial death in an Illinois woman who had CAA and had been given tissue plasminogen activator (tPA), a powerful clot-busting treatment for stoke.

Nicolas Villain, a neurologist and neuroscientist at the Sorbonne University and an investigator on the lecanemab phase 3 trial, also reviewed the case report for *Science*. The drug triggered the Florida death "beyond any reasonable doubt," says Villain, whose team has found two other cases of severe brain injuries they believe are linked to the antibody.

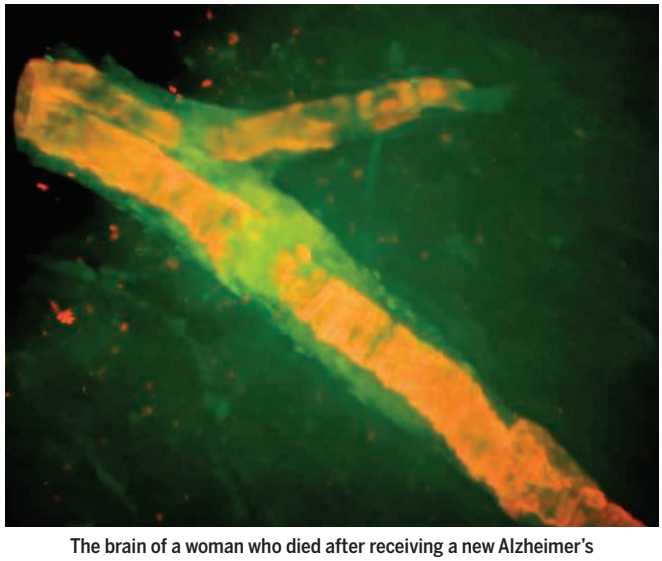
Eisai declined to comment about the specifics of the case report, which *Science* provided to it. In a written statement, a spokesperson said the company would "report any necessary and relevant information to appropriate health authorities." Eisai also referred to the current FDA-approved prescribing information, which urges doctors to review brain imaging data from a potential recipient for signs of inflammation or bleeding that might indicate CAA.

The authors of the case report say that's not enough. "This case should prompt a careful evaluation of the safety of lecanemab treatment and a refinement of the approach to pre-screening potential recipients of lecanemab for [CAA]," they conclude.

Even supporters of lecanemab share some of those concerns. On 27 March, 11 Alzheimer's researchers issued recommendations for the drug's use in a commentary in *The Journal of Prevention of Alzheimer's Disease*. They said Alzheimer's patients on blood thinners—a significant fraction—should not receive the antibody, and those taking the drug should not be given tPA. The group added that all treatment candidates should undergo genetic analysis to be sure they do not carry two copies of a gene variant called *APOE4*, which raises the risk of brain swelling and bleeding in people given antiamyloid antibodies.

Those restrictions are stronger than FDA's own prescribing criteria, and they are notable because the authors all have had consulting or advisory relationships with Eisai, its lecanemab partner Biogen, or both. Most, including Harvard University's Dennis Selkoe; Jeffrey Cummings of the University of Nevada, Las Vegas; Stephen Salloway of Brown University, have strongly advocated for the drug's approval.

The group also says Alzheimer's patients with probable CAA and related brain inflammation, as indicated by brain imaging, should not receive lecanemab. Villain calls that proposal too weak. He and colleagues wrote in an alternate set of recommendations for the use of antiamyloid antibodies in France that any Alzheimer's patients with probable CAA, even those without evidence of brain inflammation, should not receive lecanemab. By that standard, the Florida woman who died would have been excluded from the extension of the drug's phase 3 trial. ■



The brain of a woman who died after receiving a new Alzheimer's disease drug shows amyloid (orange) lining blood vessels and a site where a vessel ruptured and bled (yellow-green).

IMAGE: LISSA VENTURA-ANTUNES

This story was supported by the *Science* Fund for Investigative Reporting.



A Japanese ceda… this one offer… carbon key to unlocking historical mysteries.

MARKING TIME

Radiocarbon timestamps left in ancient tree rings by cosmic ray bombardments can date historical events with unprecedented precision

By **Michael Price**

One thousand years ago, according to Viking sagas, intrepid seafarers sailed west from Greenland to a coast they called Vinland. There they set up camps, harvested wild grapes, and skirmished with local people. At an archaeological site in a coastal town in Newfoundland called L'Anse aux Meadows, remnants of structures laid out like Viking longhouses and artifacts such as a bronze cloak pin and iron nails record their presence—likely the first Europeans to set foot in North America. Yet pinning down exactly when Vikings came to Vinland was impossible, until researchers discovered a kind of cosmic timestamp preceding their arrival.

In 2021, Margot Kuitems and Michael Dee of the University of Groningen obtained pieces of timber from a longhouse with visible tree rings and intact bark. Kuitems, an archaeologist, and Dee, a radiocarbon expert, were hoping to find a tree ring with an unusually large amount of carbon-14 (^{14}C): the

signature of a barrage of high-energy particles from outer space known to have occurred in 993–94 C.E. "It was a bit of a shot in the dark," Kuitems says. But 28 rings in from the bark, they found the telltale ^{14}C spike: persuasive evidence that Vikings had felled the fir and juniper trees used to build the longhouse in 1021 C.E.

The timing of the Viking foothold at L'Anse aux Meadows heralds a revolution in archaeology: a newfound ability to whittle the age of wooden artifacts from decades down to single years. The breakthrough paving the way for such precision came in 2012, when Japanese physicist Fusa Miyake revealed that a massive influx of cosmic rays caused a big uptick in ^{14}C in a tree ring dated to 774–75 C.E. Since then, at least seven more confirmed spikes, known as Miyake events, have been found so far. The earliest well-supported spike dates to 7176 B.C.E.

These chronological lighthouses are now guiding a growing cadre of scientists as they work to date ancient ruins, natural disasters, and other historic turning points. "If we're able to start pinning things down to the year," Dee says, "we can start analyzing early history, perhaps even prehistory, with the sort of rigor that previously we could only apply to modern history."

The technique "is phenomenally cool," says Charlotte Pearson, a dendrochronologist and radiocarbon scientist at the University of Arizona's Laboratory of Tree-Ring Research. "These are eureka moments, and we're about to have a lot more of them."

EVEN SCRIBES in the early Middle Ages noticed the cosmic assault that produced Miyake's first event. In 774 C.E., the Anglo-Saxon Chronicle, a collection of manuscripts that recounts key moments in Anglo-Saxon history, recorded the appearance in the sky, after sunset, of a "red crucifix." Astronomers have speculated that the sighting may have been atmospheric dust scattering light from a nearby supernova, or an especially vivid aurora borealis triggered by a giant solar flare.

That same year, thousands of kilometers to the east, as Japan's Emperor Kōnin struggled with a succession crisis, a Japanese cedar on Yaku Island off the nation's southern coast was sucking up carbon dioxide (CO_2) through its short, bristled needles. The tree converted some of the CO_2 into sugar and deposited that in the outermost layer of its wood. As empires rose and fell, the cedar added ring after ring of annual growth—succumbing only in 1956, at the venerable age of 1900 or so, to the blades of Yaku forestry managers. A half-century later, Miyake, then a graduate student at Nagoya University, cut a cross-section of the cedar's stump.

Each of its rings held a trace of ^{14}C. The radioactive isotope forms continuously in the upper atmosphere as cosmic rays—high-energy particles from space—collide with gas molecules, spawning neutrons. When one of these neutrons knocks out a proton in a nitrogen atom, that nitrogen is transformed into ^{14}C. As it inhaled CO_2, the cedar had incorporated the ^{14}C into its wood.

All green plants take up ^{14}C and pass it on to the animals they sustain; it can be detected in fossils and other preserved tissues going back tens of thousands of years. Its mostly predictable decay rate is the basis of standard radiocarbon dating. Using methods such as mass spectrometry, scientists can analyze organic remains to determine how much ^{14}C has decayed since the life form that absorbed it died, giving the sample's age to within a few decades under ideal circumstances. (Because the isotope has a half-life of about 5700 years, samples from earlier than about 50,000 years ago contain too little for dating.)

But Miyake wasn't trying to date the tree. She was looking for anomalous ^{14}C spikes left by violent space weather events—solar flares, other outbursts from the Sun, and exploding stars—that unleash short, exceptionally intense showers of high energy particles.

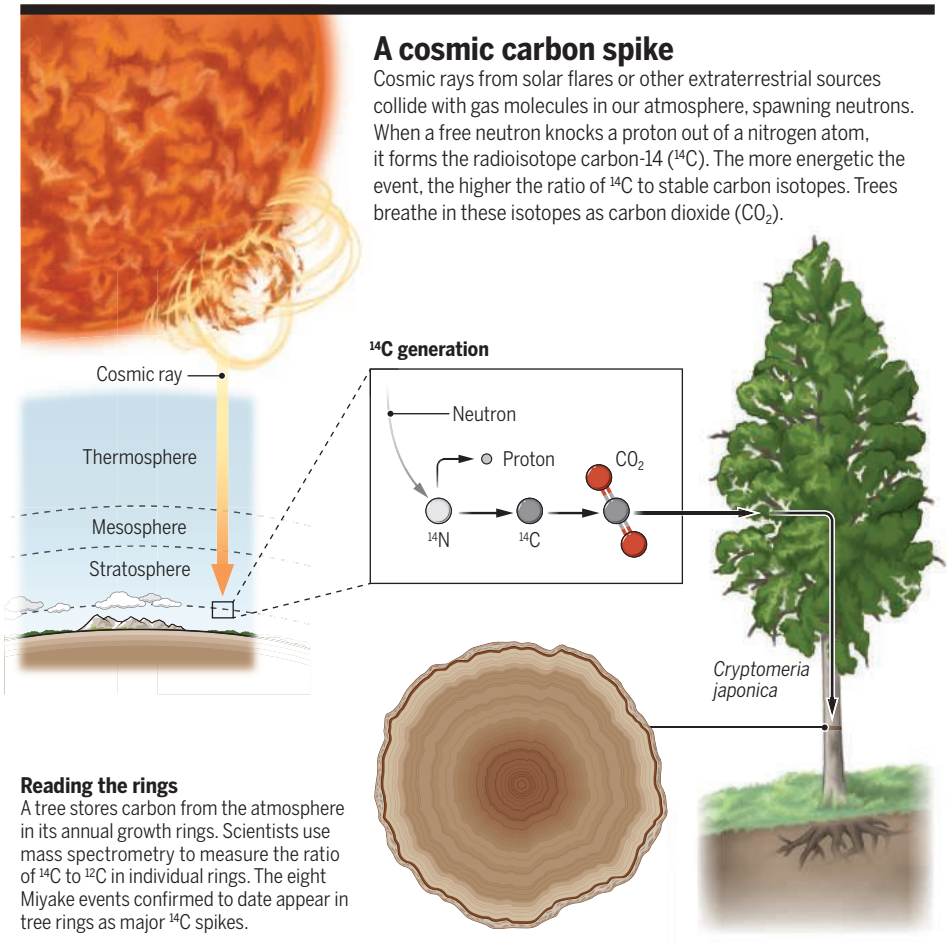
Using techniques developed in part by Lukas Wacker, a physicist at ETH Zürich, she zeroed in on a period toward the end of the 8th century C.E., where previous radiocarbon scientists had detected an anomalous hump. Working ring by ring, she sliced out tiny wood chips from the cedar and ran them through an accelerator mass spectrometer to determine the ratio of ^{14}C to stable carbon isotopes. In the ring corresponding to 774–75 C.E., she saw a 12% jump in ^{14}C: an increase 20 times larger than ordinary cosmic ray oscillations produce. Other teams confirmed the spike in samples from a German oak and a New Zealand kauri tree.

Later that year, Miyake and her colleagues published a paper in *Nature* attributing the ^{14}C spike to a massive bombardment of cosmic rays, perhaps from the Sun or a gamma ray burst from a distant star. In 2013, they found a second, slightly smaller ^{14}C spike in the same Yaku cedar at 993–94 C.E. Other scientists started calling the phenomena Miyake events.

When Wacker read about Miyake's 774–75 C.E. and 993–94 C.E. spikes, he recognized the radiocarbon beacon's untapped po-

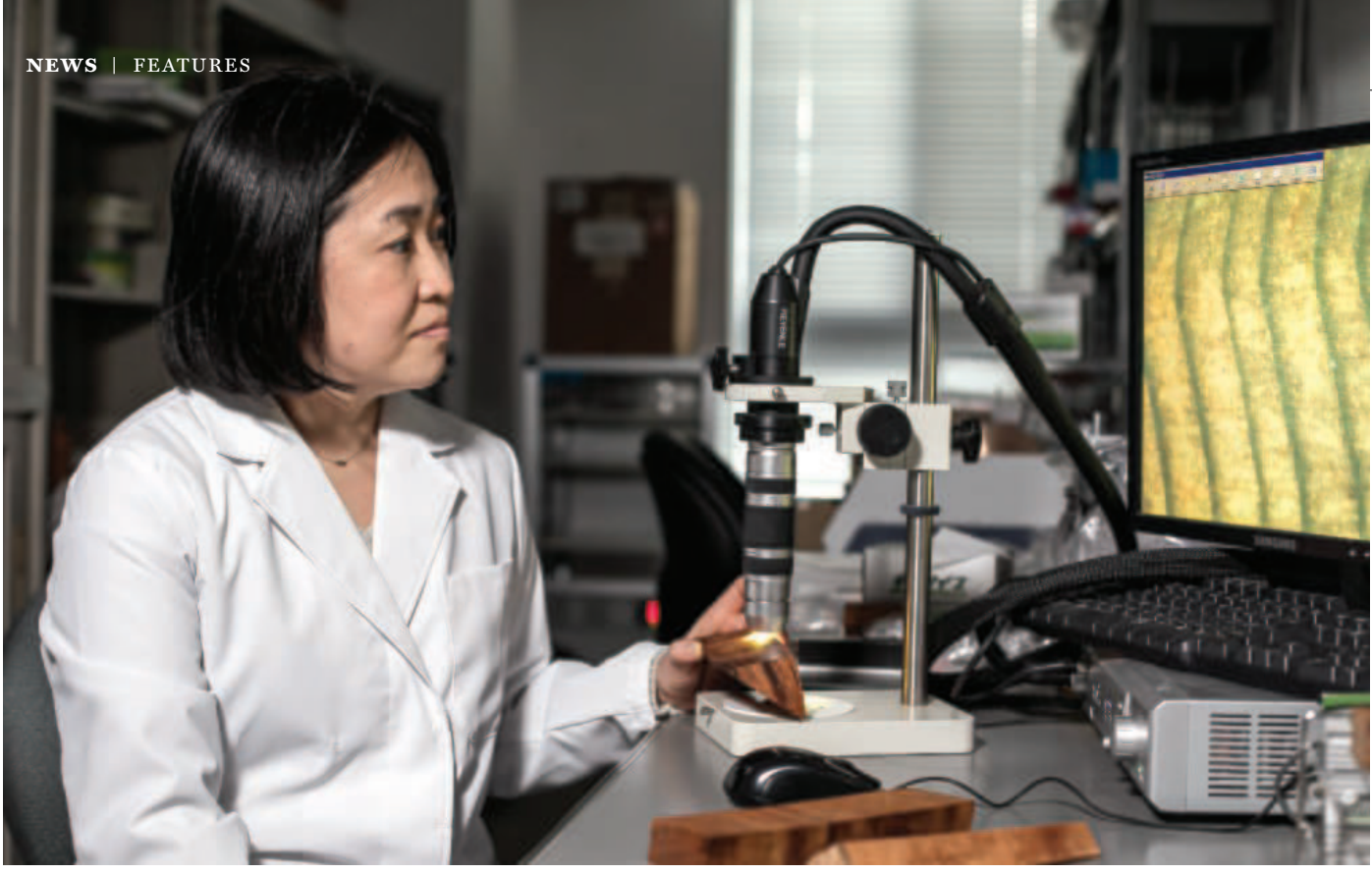
A cosmic carbon spike

Cosmic rays from solar flares or other extraterrestrial sources collide with gas molecules in our atmosphere, spawning neutrons. When a free neutron knocks a proton out of a nitrogen atom, it forms the radioisotope carbon-14 (^{14}C). The more energetic the event, the higher the ratio of ^{14}C to stable carbon isotopes. Trees breathe in these isotopes as carbon dioxide (CO_2).



Reading the rings
A tree stores carbon from the atmosphere in its annual growth rings. Scientists use mass spectrometry to measure the ratio of ^{14}C to ^{12}C in individual rings. The eight Miyake events confirmed to date appear in tree rings as major ^{14}C spikes.

GRAPHIC: A. FISHER/*SCIENCE*



A decade ago, Fusa Miyake discovered two precise markers of time in the rings of an ancient cedar: carbon-14 spikes now called Miyake events.

tential. He trained it on a chapel in Müstair, Switzerland, supposedly built by the first Holy Roman Emperor, Charlemagne, on a site where he and his party had survived a horrendous blizzard. Another team of scientists had dated the chapel to 785 C.E. by meticulously analyzing the widths of tree rings in a wooden beam—wider rings often indicate wetter years—then matching the patterns with more recent timber from the area until they had an unbroken record stretching from the present back to the beam.

Wacker knew that if his team could find the ^{14}C spike from the 774–75 C.E. Miyake event in the beam, they could simply count to its outer edge to obtain a precise date. They did just that, confirming in 2016 the 785 C.E. date. "That was pioneering work," Miyake says. "I was really happy to see that kind of application was made possible because of our findings."

THE HOLY CROSS CHAPEL dating helped persuade Dee and colleagues that Miyake events might illuminate historical timelines confused or obscured by the passage of time. Reliable written records—chronicles of empires and wars, and more mundane archives of taxes, trade receipts, births and deaths—go back only some 2500 years, Dee says, and only in societies that kept written records. Radiocarbon dates, with their decades or centuries of uncertainty, may help reconstruct history in broad strokes, but key historical moments often play out over shorter timescales. "One couldn't examine the political events of the 20th century if you only had decadal resolution," Dee says. "If the First World War and the Second World War had been in 1914 B.C. and 1939 B.C., we wouldn't be able to tell those two things apart."

In 2019, Kuitems used the technique to clear up a longstanding mystery surrounding rectangular ruins on a lake island in southern Siberia. Despite decades of archaeological work, nobody knew for sure how old the Por-Bazhyn site is, who built it, or what its purpose was. Por-Bazhyn lacked any sign of occupation such as trash heaps or charcoal from hearths, suggesting it was abandoned soon after construction. But the 774–75 C.E. spike was present in an outer ring of a beam, indicating the source tree had been felled in 777 C.E. That was during the reign of Tengri Bögü Qaghan, a Uyghur monarch who converted to Manichaeism, a religion that views existence as an epic struggle between good and evil. Most likely, Kuitems says, Bögü Qaghan built Por-Bazhyn as a Manichaean monastery.

Por-Bazhyn's precise age points to the apparent reason it was never occupied. In 779 C.E., rebels opposed to Manichaeism killed Bögü Qaghan. "So at the moment this whole building was finished, it immediately became useless," Kuitems says. Dee and other colleagues used the same Miyake spike to date another ancient fortress, on an island in Lake Āraiši in central Latvia. "These were questions that people couldn't answer for decades, and now we could with this method," Kuitems says.



A Miyake event in timber from Viking structures in Newfoundland gave them a precise date of 1021 C.E. (These are modern reconstructions.)

In the wake of their headline-grabbing Vinland success, the Groningen team is tackling another alluring set of historical puzzles: "floating" chronologies, or ancient, internally consistent timelines that are untethered to specific dates. For instance, Dee has long hoped to moor the floating timelines of ancient Egypt's pharaohs to our modern calendar. Radiocarbon dating has pegged the Egyptian chronology to within a couple hundred years; scientists have tried to use ancient observations of known astronomical events,

such as star and planet alignments, to narrow dates further. But timelines remain blurry. Egypt's Old Kingdom era, for example, encompasses dozens of pharaohs who reigned from approximately 2700 to 2200 B.C.E. "The calendrical dates are not even really known to the century," Dee says.

Many native Egyptian tree species don't produce annual growth rings, complicating the quest for Miyake events there. But the Egyptians imported conifers—which builders often preferred for their straightness and strength—from Lebanon and Syria. The pandemic and laws restricting the export of Egyptian artifacts have slowed his team's work, Dee says. But he has his sights set on discovering a precise date for the construction of the Great Pyramid, which was built by the Old Kingdom Pharaoh Khufu. "There is an old Arab proverb which goes, 'Man fears time; time fears the Pyramids,'" Dee says. "If this structure, and his reign, could be precisely dated, that would solve a mystery that has endured since antiquity itself."

Another floating chronology Dee's team hopes to anchor is the Mesoamerican Long Count timekeeping system. Used by the Maya, Aztec, and other Mesoamerican civilizations, the system counts the days linearly from a supposed day of the world's creation. While there have been several attempts to correlate the Long Count with the Gregorian calendar, academics are still debating the precise creation date. Finding a Miyake event in wood from a Mesoamerican structure—such as a lintel in the Maya temple Tikal in Guatemala, whose construction is recorded in the Long Count—would settle the matter.

MIYAKE EVENTS ALSO PROMISE to date natural disasters that have altered the course of human history. "This technique can help answer questions about the rise and fall of civilizations," Pearson says.

In 2017, Clive Oppenheimer, a volcanologist at University of Cambridge, teamed up with Wacker and others to use a Miyake event to date a larch tree buried in ash by the so-called Millennium eruption of Mount Paektu, a volcano straddling the border of China and North Korea. Radiocarbon dates pegged the eruption—one of the biggest in the past 10,000 years—to sometime during the ninth to 11th centuries. Some historians hypothesized that the blast led to the fall in 926 C.E. of the Bohai Kingdom, which encompassed parts of the Korean Peninsula, Northeast China, and Russia's Far East. Oppenheimer and colleagues used the 774–75 C.E. Miyake event to date the larch's demise to 946 C.E., exonerating the eruption in the kingdom's collapse. "Thanks to the Miyake event, we could really nail it," Oppenheimer says. "With anything that kills trees—earthquakes, fires, biological pests—



A contemporary magazine illustration depicted the radiant aura of 1859.

Past boon, future menace

Miyake events, blasts from the Sun or other stars that deposit spikes of radioactive carbon in tree rings, illuminate the past. But such a blast could seriously disrupt the present.

The largest solar storm to hit our planet in recorded history was the Carrington Event from 1–2 September 1859, named after a British astronomer who linked a massive solar flare others had spotted with the bizarre phenomena that soon followed. In the Northern Hemisphere, auroral displays were seen as far south as the Caribbean and were so bright in the northeastern United States that people could read newspapers at night. Telegraph stations caught fire as transmission wires sizzled under a barrage of electrons. Newspapers reported that birds began chirping in the night, and gold miners in California woke up in the wee hours to prepare breakfast, confusing the bright aurora for daybreak.

Yet compared with a Miyake event, those two scintillating days in 1859 are "barely a blip," says Charlotte Pearson, a dendrochronologist and radcarbon scientist at the University of Arizona. Carbon-14 (^{14}C) records in tree rings from 1859 show virtually no upticks—nothing approaching the massive spike caused by Miyake events.

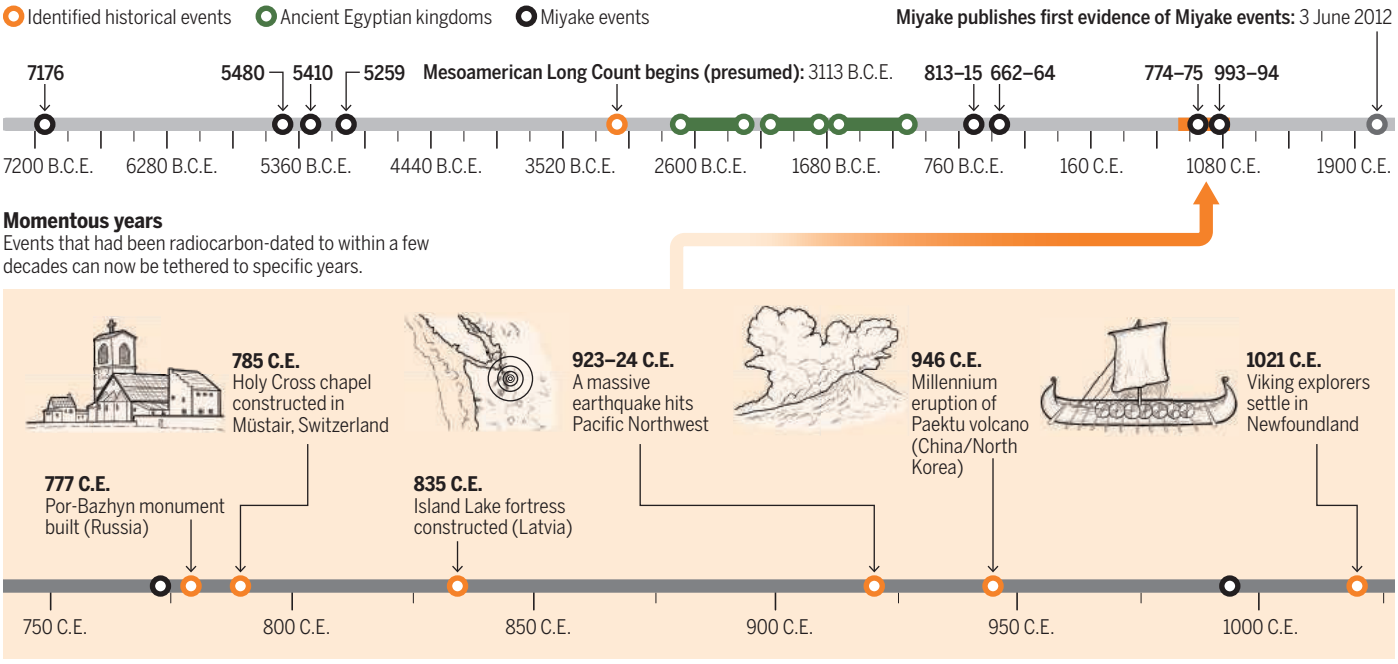
A solar storm or gamma ray assault on the scale of the 774–75 C.E. event would similarly turn the night sky into an awe-provoking swirl of crackling, shimmering light. It would also do catastrophic damage to technologies and infrastructure. "We're massively vulnerable," Pearson says. The surge of high-energy particles would fry satellite electronics, and astronauts in orbit likely would absorb lethal doses of radiation; even airplane passengers might absorb dangerous levels, Miyake says. As Earth's magnetic field convulsed under the assault, surges of current would destroy power grids. Nearly anything running on electricity—from electric vehicles to cellphones to ventilators—would short out. At the very least, we'd experience widespread communications failures, she says.

The grave consequences have scientists trying to discern whether past Miyake events follow a pattern or cycle, or whether a future event would blindside us. "The goal is to try and understand the frequency and severity of these events with a view to actually do some kind of risk assessment," says physicist Andrew Smith of the Australian Nuclear Science and Technology Organisation. That includes working out how to better shield people and equipment on Earth and in space. "Obviously, a big [Miyake] event today would be a disaster," Smith says. "All of our tech would be in a smoking heap overnight."—*M. P.*

ILLUSTRATION: *HARPER'S NEW MONTHLY MAGAZINE*

Pinning down history

Thanks to Miyake events inscribed in ancient wood, scientists are arriving at precise dates for a growing number of archaeological sites—from enigmatic ruins to high-profile monuments like a Viking settlement in North America—as well as natural disasters such as volcanic eruptions and earthquakes.



there's now the potential to date that very, very precisely."

Pearson is hoping to replicate a version of that success in a tree that was downwind of the volcano Thera when it erupted around 3500 years ago on the island of Santorini, possibly contributing to the fall of the Minoan civilization. The widespread devastation would have disrupted commerce throughout the Mediterranean and sent refugees scrambling for new homes. A ^{14}C spike centered on 1528 B.C.E.—so far unconfirmed as a Miyake event—confirms that in 1562 B.C.E, plus or minus 1 year, an abrupt chemical signal shows up in the tree's rings, possibly caused by sulfate from the eruption.

Miyake events may even shed light on potential future cataclysms. Dendrochronologist Bryan Black, also at the Arizona tree-ring lab, recently used the 774–75 C.E. Miyake event to date two mass tree deaths separated by more than 80 kilometers in the Puget Sound region of the U.S. Pacific Northwest. The forests grew along two distinct shallow faults near present-day Seattle, and geological evidence shows they died more than 1000 years ago. Sediments record the upheaval, as do oral traditions of the indigenous Salish people. Earth scientists have wondered whether both faults ruptured separately, or at once in a single massive earthquake. Black found these trees all died during the same season between 923 and 924 C.E., suggesting the faults could conspire again to produce a massive earthquake. "It takes the current worst-case scenario of these shallow faults and raises it to the next level," he asserts.

RADIOCARBON SCIENTISTS are scouring tree ring archives for more Miyake events. Kuitems hopes to firm up one in the sixth millennium B.C.E., which may allow her to precisely date an undisclosed Neolithic site in Europe. Other researchers have spotted a possible Miyake cluster in 1261 C.E., 1268 C.E., and 1279 C.E. Several different labs are aiming to verify these and other ^{14}C spikes in trees from around the world.

"It's like a sort of jigsaw puzzle," Dee says. "In the end, if we have enough spikes, we'll be able to put together a lot of this history of the Holocene in absolute time."

Ice core data can also confirm Miyake events. Cosmic rays spawn two other radioisotopes—beryllium-10 and chlorine-36—in the upper atmosphere. Rather than lodging in living tissue, these isotopes are washed to the surface in rain and snow, and they can end up in the annual layers of polar ice sheets. Like ^{14}C in tree rings, sharp spikes of beryllium-10 and chlorine-36 in the ice have been correlated to known Miyake events. At a radiocarbon conference in Zürich last year, physicists Andrew Smith and David Fink of the Australian Nuclear Science and Technology Organisation reported isotopic spikes at 774–75 C.E. and 993–94 C.E. in Antarctic ice cores. Continued analysis of the cores could reveal other, previously unknown Miyake events that dendrochronologists could follow up on.

Miyake, too, is continuing her quest to find more of the events that bear her name. She has joined forces with Arizona's Laboratory of Tree Ring Research, whose 700,000 tree ring samples–the largest collection in the world–offer an unparalleled record of past climate and cosmic events. There, Miyake and her Arizona colleagues, along with colleagues around the world, are aiming to patch together from recently deceased and long-dead trees a complete radiocarbon record stretching back 12,000 years. Filling in the gaps between Miyake events is also important, says Pearson, as they contain solar patterns that can be matched with tree-ring and ice-core records to improve radiocarbon dating. Pearson and Wacker are preparing to submit another 1100 years' worth of data from California's bristlecone pines, which can live for more than 4000 years, to an international radiocarbon calibration repository. Miyake is looking to samples of other long-lived trees from Russia, Finland, and Japan, and has already spotted several as-yet unpublished Miyake events, she says.

Miyake's pioneering analysis of the cedar continues to ripple through studies of the past, colleagues say. "She trusted her instincts and was so confident in what she had that she was able to convince us all that there was a real effect," Dee says. "And the rest, as they say, is history." ■

Dennis Normile contributed to this story.

such as star and planet alignments, to narrow dates further. But timelines remain blurry. Egypt's Old Kingdom era, for example, encompasses dozens of pharaohs who reigned from approximately 2700 to 2200 B.C.E. "The calendrical dates are not even really known to the century," Dee says.

Many native Egyptian tree species don't produce annual growth rings, complicating the quest for Miyake events there. But the Egyptians imported conifers—which builders often preferred for their straightness and strength—from Lebanon and Syria. The pandemic and laws restricting the export of Egyptian artifacts have slowed his team's work, Dee says. But he has his sights set on discovering a precise date for the construction of the Great Pyramid, which was built by the Old Kingdom Pharaoh Khufu. "There is an old Arab proverb which goes, 'Man fears time; time fears the Pyramids,'" Dee says. "If this structure, and his reign, could be precisely dated, that would solve a mystery that has endured since antiquity itself."

Another floating chronology Dee's team hopes to anchor is the Mesoamerican Long Count timekeeping system. Used by the Maya, Aztec, and other Mesoamerican civilizations, the system counts the days linearly from a supposed day of the world's creation. While there have been several attempts to correlate the Long Count with the Gregorian calendar, academics are still debating the precise creation date. Finding a Miyake event in wood from a Mesoamerican structure—such as a lintel in the Maya temple Tikal in Guatemala, whose construction is recorded in the Long Count—would settle the matter.

MIYAKE EVENTS ALSO PROMISE to date natural disasters that have altered the course of human history. "This technique can help answer questions about the rise and fall of civilizations," Pearson says.

In 2017, Clive Oppenheimer, a volcanologist at University of Cambridge, teamed up with Wacker and others to use a Miyake event to date a larch tree buried in ash by the so-called Millennium eruption of Mount Paektu, a volcano straddling the border of China and North Korea. Radiocarbon dates pegged the eruption—one of the biggest in the past 10,000 years—to sometime during the ninth to 11th centuries. Some historians hypothesized that the blast led to the fall in 926 C.E. of the Bohai Kingdom, which encompassed parts of the Korean Peninsula, Northeast China, and Russia's Far East. Oppenheimer and colleagues used the 774–75 C.E. Miyake event to date the larch's demise to 946 C.E., exonerating the eruption in the kingdom's collapse. "Thanks to the Miyake event, we could really nail it," Oppenheimer says. "With anything that kills trees—earthquakes, fires, biological pests—



A contemporary magazine illustration depicted the radiant aura of 1859.

ILLUSTRATION: *HARPER'S NEW MONTHLY MAGAZINE*

Past boon, future menace

Miyake events, blasts from the Sun or other stars that deposit spikes of radioactive carbon in tree rings, illuminate the past. But such a blast could seriously disrupt the present.

The largest solar storm to hit our planet in recorded history was the Carrington Event from 1–2 September 1859, named after a British astronomer who linked a massive solar flare others had spotted with the bizarre phenomena that soon followed. In the Northern Hemisphere, auroral displays were seen as far south as the Caribbean and were so bright in the northeastern United States that people could read newspapers at night. Telegraph stations caught fire as transmission wires sizzled under a barrage of electrons. Newspapers reported that birds began chirping in the night, and gold miners in California woke up in the wee hours to prepare breakfast, confusing the bright aurora for daybreak.

Yet compared with a Miyake event, those two scintillating days in 1859 are "barely a blip," says Charlotte Pearson, a dendrochronologist and radcarbon scientist at the University of Arizona. Carbon-14 (^{14}C) records in tree rings from 1859 show virtually no upticks—nothing approaching the massive spike caused by Miyake events.

A solar storm or gamma ray assault on the scale of the 774–75 C.E. event would similarly turn the night sky into an awe-provoking swirl of crackling, shimmering light. It would also do catastrophic damage to technologies and infrastructure. "We're massively vulnerable," Pearson says. The surge of high-energy particles would fry satellite electronics, and astronauts in orbit likely would absorb lethal doses of radiation; even airplane passengers might absorb dangerous levels, Miyake says. As Earth's magnetic field convulsed under the assault, surges of current would destroy power grids. Nearly anything running on electricity—from electric vehicles to cellphones to ventilators—would short out. At the very least, we'd experience widespread communications failures, she says.

The grave consequences have scientists trying to discern whether past Miyake events follow a pattern or cycle, or whether a future event would blindside us. "The goal is to try and understand the frequency and severity of these events with a view to actually do some kind of risk assessment," says physicist Andrew Smith of the Australian Nuclear Science and Technology Organisation. That includes working out how to better shield people and equipment on Earth and in space. "Obviously, a big [Miyake] event today would be a disaster," Smith says. "All of our tech would be in a smoking heap overnight."—*M. P.*

INSIGHTS

PERSPECTIVES



CLIMATE SCIENCE

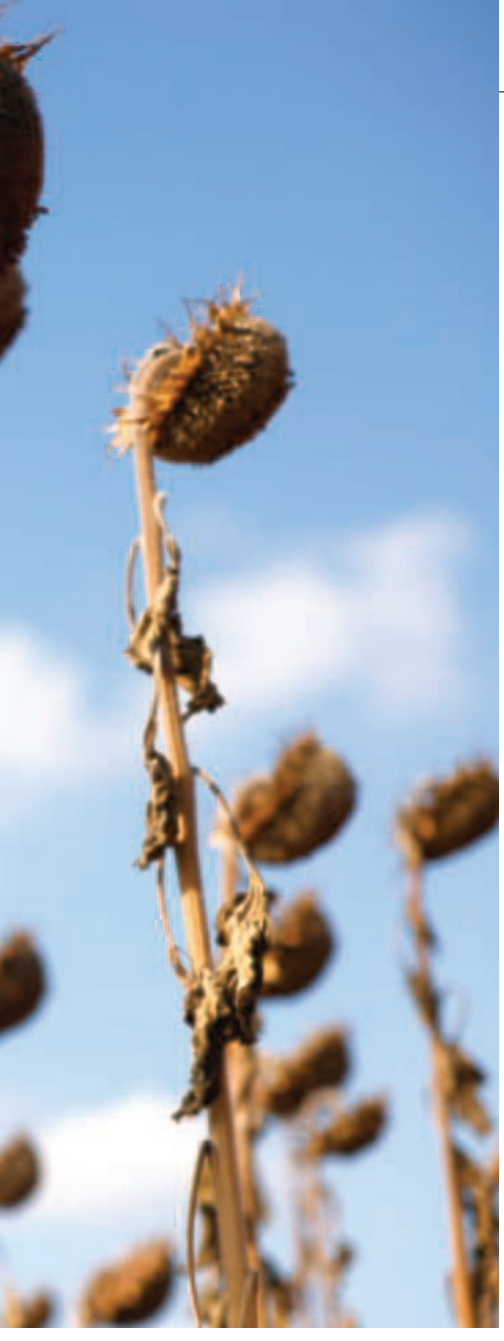
Droughts are coming on faster

Higher global temperatures are increasing the frequency of flash droughts

By **David W. Walker**[1] *and* **Anne F. Van Loon**[2]

Droughts are generally considered a long and slowly developing climate hazard. A persistent lack of rainfall dries out soil and vegetation, followed by declining water levels in rivers and lakes. Droughts may build and endure for years over large areas, with snowballing consequences for people and the environment. Across the Horn of Africa, for example, rains have failed for six consecutive rainy seasons, bringing widespread food insecurity and mass migration (*1*). Conversely, drought can arise rapidly, changing conditions from "normal" to problematically dry in a matter of weeks. These so-called "flash droughts" are akin to flash floods in that they appear suddenly with little warning. Consequently, impacts can be severe because there is little time to prepare for or deal with such a drought before it is too late. On page 187 of this issue, Yuan *et al.* (*2*) present a prognosis of a global transition to more frequent flash droughts.

As per most droughts, a flash drought is triggered by a period of anomalously low rainfall, though it is also associated with excessively high evaporation (prompted by higher temperatures). Therefore, it could be expected that as the planet warms, flash droughts will become a more common threat. Yuan *et al.* determined that over



Parts of Europe weathered a major drought in 2022 that devastated crops, such as this sunflower field in eastern France.

the past 70 years, the occurrence of flash droughts has increased as usually slower-onset droughts are coming on faster. The authors attribute this trend to anthropogenic climate change. Hence, flash drought frequency is likely to further increase in the future because droughts almost everywhere in the world will have faster onset.

The term "flash drought" originated in the United States in the early 21st century, yet flash droughts exist around the world with varying characteristics and are sufficiently well known to local populations that they receive local nomenclature, such as "veranicos" or "veranillos" ("little summers") in Brazil and Central America, respectively (*3*, *4*). Yuan *et al.* report that flash droughts occur more in humid regions. In these areas, sufficient water is available for evapotranspiration (transfer of water from land surfaces, water bodies, and plants to the atmosphere), and when conditions are right (low precipitation, high radiation, high wind speeds, and low resistance in soil or vegetation), a large amount of water can evaporate quickly.

Flash droughts materialize in soil moisture, quickly affecting soil moisture–dependent sectors such as agriculture and ecosystems. The minimum duration of a flash drought that is required to create impacts depends on various factors specific to each location. Some soils desiccate more rapidly than others, some areas have abundant water resources that could provide supplemental irrigation, some crops are more resistant to water and heat stress, and the proportion of populations reliant on agriculture for their livelihoods varies greatly between regions. It also depends on whether the flash drought recovers as quickly as it started or whether it develops into a normal long drought. Thus, the effects of a flash drought are clearly not only dependent on the severity of the drought itself. The hot spot regions for flash droughts determined by Yuan *et al.* are mostly found in low- and middle-income countries, and these regions generally have more vulnerable populations and lower financial resources for coping mechanisms. This is especially the case in arid and semi-arid regions. Although Yuan *et al.* found that flash droughts are more likely in the humid tropics, flash drought risk could be highest in more arid parts of the tropics.

A flash drought at certain times of year will have a much greater impact than at other times, according to preceding conditions and the particular crop growth stage. For example, a flash drought lasting mere days can substantially affect crop yields when it occurs during flowering or pollination stages, whereas later in the growing season, flash droughts can be beneficial for plant growth and fruit ripening (*5*). It may also be the case that a flash drought toward the end of the rainy season is less harmful because mitigating water resources have had time to accumulate. Flash drought timing is thus critically related to agricultural growing seasons, which vary around the world. Yuan *et al.* used a simplified version of these growing seasons by using April-to-September for the Northern Hemisphere and October-to-March for the Southern Hemisphere. However, growing seasons in the tropics may not correspond to these time periods because they are more dependent on seasonality in water availability than in temperature and radiation. It would be interesting to investigate flash drought trends in realistic growing seasons globally.

To plan adaptation, forecasting, and mitigation strategies, it is necessary to estimate the minimum drought duration that generates impacts and the associated critical timings, in addition to appreciating what other consequences arise from flash drought. To date, research has dominantly, and logically, focused on agricultural impacts (*6*). A starting point would be to monitor how society at large is affected by drought, which is increasingly recognized as a missing link in drought management. The lack of such monitoring is true for all drought types and is being tackled in various ways, with impact databases that store information from official reports, government-led impact monitoring, and crowdsourcing of impact information (*7*, *8*). The rapid onset of flash droughts and the urgency of impact identification and action suggest that real-time monitoring of social media traffic should be explored more (*8*).

An increasingly popular first step in drought management around the world is establishment of a drought monitor that shows the current state of drought in a region. The associated drought maps are mostly issued monthly, and even then after a delay that is needed for analysis and computation. Yet flash droughts may build and trigger outcomes in just weeks. Clearly, as noted by Yuan *et al.*, methods of drought detection are needed that operate on shorter time scales, and drought maps must be updated more frequently. Drought is most commonly identified when there are deviations in rainfall or when anomalies in vegetation health are detected through satellite remote sensing. The latter is especially useful in regions of scarce on-the-ground meteorological monitoring. However, rainfall alone is insufficient to identify flash drought, and declining vegetation health may only become apparent on satellite imagery once it is already too late to act. Closely monitoring soil moisture and evapotranspiration at the daily (or ideally subdaily) scale successfully identified flash droughts weeks earlier than the US Drought Monitor system (*9*).

Unfortunately, the crucial data for flash drought identification—soil moisture and actual evapotranspiration—are hard to obtain. Most observation methods measure

[1]Water Resources Management Group, Wageningen University, Wageningen, Netherlands. [2]Institute for Environmental Studies, Vrije Universiteit Amsterdam, Amsterdam, Netherlands. Email: david.walker@wur.nl; anne.van.loon@vu.nl

at a single point in space that is not representative for a larger area. Satellite soil moisture estimates provide global-scale data, although satellite sensors only reach the upper few centimeters of the soil, not the deeper soil layers also important for the development of flash droughts (*10*). Global land surface models provide data on soil moisture for the past (reanalysis data) and the future. For deeper soil layers, Yuan *et al*. used a model that simulates soil moisture up to a depth of 1 meter. This model type performs relatively well compared with observations (*11*), although studies also show that in these models, evapotranspiration processes draw water mostly from shallow soil and canopy storage and not from deeper soil moisture. This results in lower evapotranspiration values, especially in mid-latitude regions (*12*). This may indicate that the flash drought frequency results of Yuan *et al*. could be an underestimate for the mid-latitudes.

Therefore, the warning of Yuan *et al*. that the warming planet faces an increase in frequency of these extreme events should be taken seriously because in both the tropics and the mid-latitudes, the threat may be even greater than they suggest. With less time to prepare for flash droughts, proactive adaptation is required rather than more costly reactive drought management. ■

REFERENCES AND NOTES

1. United Nations High Commissioner for Refugees (UNHCR), "As the Horn of Africa drought enters a sixth failed rainy season, UNHCR calls for urgent assistance, 28 February 2023" (UNHACR, 2023); https://www.unhcr.org/news/briefing/2023/2/63fdbcee4/horn-africa-drought-enters-sixth-failed-rainy-season-unhcr-calls-urgent.html.
2. X. Yuan *et al*., *Science* **380**, 187 (2023).
3. A. R. Magalhães, M. H. Glantz, "Socioeconomic impacts of climate variations and policy responses in Brazil" (United Nations Environment Program, Secretariat for Planning and Coordination State of Ceara, Esquel Brasil Foundation, 1992).
4. V. Magaña, J. A. Amador, S. Medina, *J. Clim.* **12**, 1577 (1999).
5. R. Hamed, A. F. Van Loon, J. Aerts, D. Coumou, *Earth Syst. Dyn.* **12**, 1371 (2021).
6. J. A. Otkin *et al*., *Bull. Am. Meteorol. Soc.* **103**, E2188 (2022).
7. K. Stahl *et al*., *Nat. Hazards Earth Syst. Sci.* **16**, 801 (2016).
8. K. H. Smith, A. J. Tyre, Z. Tang, M. J. Hayes, F. A. Akyuz, *Bull. Am. Meteorol. Soc.* **101**, E1801 (2020).
9. T. W. Ford, D. B. McRoberts, S. M. Quiring, R. E. Hall, *Geophys. Res. Lett.* **42**, 9790 (2015).
10. A. J. Teuling *et al*., *Nat. Geosci.* **3**, 722 (2010).
11. M. Mujumdar *et al*., *J. Hydrol.* (Amst.) **597**, 126102 (2021).
12. J. Dong, F. Lei, W. T. Crow, *Nat. Commun.* **13**, 336 (2022).

ACKNOWLEDGMENTS

The authors thank R. Hamed for providing useful comments to a draft version of this Perspective. We are grateful for the International Association of Hydrological Sciences Panta Rhei Drought in the Anthropocene working group for discussions on this topic over the past years.

10.1126/science.adh3097

MEDICINE

Designer bugs as cancer drugs?

Skin microbiota can be engineered to induce antitumor T cells

By **Sairaj M. Sajjath, Anita Gola, Elaine Fuchs**

Human health relies on carefully tailored communication between the immune system and commensal microbiota, a diverse community of microorganisms that naturally reside in tissues. Several commensal bacteria trigger antigen-specific T cell responses, which help control tissue homeostasis (*1*). Could such interactions be harnessed therapeutically, by engineering symbionts to prime immunity against designated targets? On page 203 of this issue, Chen *et al*. (*2*) test this hypothesis by customizing the skin commensal *Staphylococcus epidermidis* to express a foreign antigen borne by melanoma cells. Upon colonization of mouse skin, these bacteria substantially bolstered organism-wide, T cell–driven attack of resulting tumors. The findings highlight the broader potential of microbiota-induced immunity toward host defense and open new doors to customized commensal bacteria therapies against a myriad of human diseases.

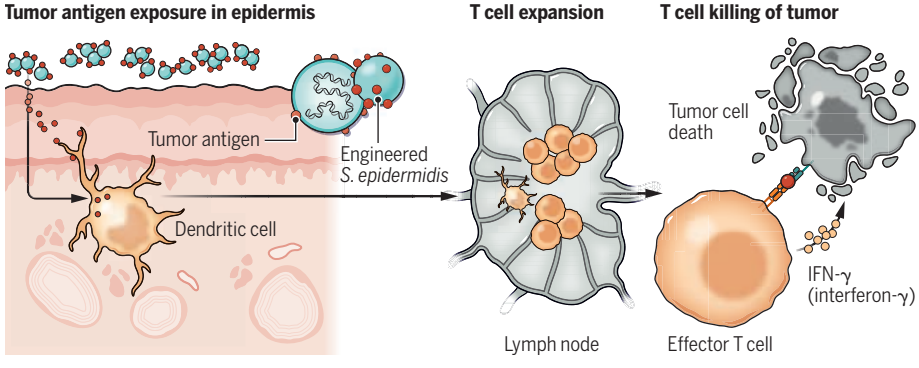
With a growing understanding of the microbiota's central role in optimal tissue fitness, strategies have emerged to therapeutically exploit this partnership. These primarily involve supplementing hosts with beneficial commensal bacteria or their metabolic products (*3*). Although still nascent, such approaches are promising: Food supplements that promote a healthy gut microbiota can rescue key health outcomes in malnourished children, while microbes associated with human pathology and treatment responsiveness have now been discovered for disorders including cancers, obesity, and allergies (*4*–*8*).

Concomitantly, it has become appreciated that individual symbionts can induce antigen-specific immune responses that dominantly influence tissue states, promoting such varied processes as inflammation, tissue regeneration, and immune tolerance, depending on the context (*1*). Indeed, when applied to mouse skin, human isolates of *S. epidermidis* drive locally increased numbers of $CD4^+$ (helper) and $CD8^+$ (cytotoxic) T lymphocytes, which preempt infections in colonized tissue through homeostatic production of inflammatory cytokines (*9*). Whether such antigen-specific responses can be systemically mobilized for targeted protective immunity throughout an organism had been unexplored until now.

Chen *et al*. genetically manipulated *S. epidermidis* to encode antigens (either test-case ovalbumin peptides or known tumor antigens) present on mouse melanoma cells. In contrast to nonengineered *S. epidermidis*, these modified commensals elicited potent T cell responses against both subcutaneous melanoma grafts and lung metastases, strikingly without causing inflammation in the skin colonized by the antigen-expressing bacteria. These "designer" commensals even impeded melanoma progression when applied

Antigen-engineered commensal bacteria

By genetically engineering the skin commensal *Staphylococcus epidermidis* to express tumor antigens, a robust effector T cell response can be induced without causing inflammation at the site of colonization in mice. When tumor cells expressing the same antigens are encountered, these T cells can be mobilized into circulation to systemically confer long-term host protection.



at a single point in space that is not representative for a larger area. Satellite soil moisture estimates provide global-scale data, although satellite sensors only reach the upper few centimeters of the soil, not the deeper soil layers also important for the development of flash droughts (*10*). Global land surface models provide data on soil moisture for the past (reanalysis data) and the future. For deeper soil layers, Yuan *et al.* used a model that simulates soil moisture up to a depth of 1 meter. This model type performs relatively well compared with observations (*11*), although studies also show that in these models, evapotranspiration processes draw water mostly from shallow soil and canopy storage and not from deeper soil moisture. This results in lower evapotranspiration values, especially in mid-latitude regions (*12*). This may indicate that the flash drought frequency results of Yuan *et al.* could be an underestimate for the mid-latitudes.

Therefore, the warning of Yuan *et al.* that the warming planet faces an increase in frequency of these extreme events should be taken seriously because in both the tropics and the mid-latitudes, the threat may be even greater than they suggest. With less time to prepare for flash droughts, proactive adaptation is required rather than more costly reactive drought management. ■

REFERENCES AND NOTES

1. United Nations High Commissioner for Refugees (UNHCR), "As the Horn of Africa drought enters a sixth failed rainy season, UNHCR calls for urgent assistance, 28 February 2023" (UNHACR, 2023); https://www.unhcr.org/news/briefing/2023/2/63fdbcee4/horn-africa-drought-enters-sixth-failed-rainy-season-unhcr-calls-urgent.html.
2. X. Yuan *et al.*, *Science* **380**, 187 (2023).
3. A. R. Magalhães, M. H. Glantz, "Socioeconomic impacts of climate variations and policy responses in Brazil" (United Nations Environment Program, Secretariat for Planning and Coordination State of Ceara, Esquel Brasil Foundation, 1992).
4. V. Magaña, J. A. Amador, S. Medina, *J. Clim.* **12**, 1577 (1999).
5. R. Hamed, A. F. Van Loon, J. Aerts, D. Coumou, *Earth Syst. Dyn.* **12**, 1371 (2021).
6. J. A. Otkin *et al.*, *Bull. Am. Meteorol. Soc.* **103**, E2188 (2022).
7. K. Stahl *et al.*, *Nat. Hazards Earth Syst. Sci.* **16**, 801 (2016).
8. K. H. Smith, A. J. Tyre, Z. Tang, M. J. Hayes, F. A. Akyuz, *Bull. Am. Meteorol. Soc.* **101**, E1801 (2020).
9. T. W. Ford, D. B. McRoberts, S. M. Quiring, R. E. Hall, *Geophys. Res. Lett.* **42**, 9790 (2015).
10. A. J. Teuling *et al.*, *Nat. Geosci.* **3**, 722 (2010).
11. M. Mujumdar *et al.*, *J. Hydrol.* (Amst.) **597**, 126102 (2021).
12. J. Dong, F. Lei, W. T. Crow, *Nat. Commun.* **13**, 336 (2022).

ACKNOWLEDGMENTS

The authors thank R. Hamed for providing useful comments to a draft version of this Perspective. We are grateful for the International Association of Hydrological Sciences Panta Rhei Drought in the Anthropocene working group for discussions on this topic over the past years.

10.1126/science.adh3097

MEDICINE

Designer bugs as cancer drugs?

Skin microbiota can be engineered to induce antitumor T cells

By **Sairaj M. Sajjath, Anita Gola, Elaine Fuchs**

Human health relies on carefully tailored communication between the immune system and commensal microbiota, a diverse community of microorganisms that naturally reside in tissues. Several commensal bacteria trigger antigen-specific T cell responses, which help control tissue homeostasis (*1*). Could such interactions be harnessed therapeutically, by engineering symbionts to prime immunity against designated targets? On page 203 of this issue, Chen *et al.* (*2*) test this hypothesis by customizing the skin commensal *Staphylococcus epidermidis* to express a foreign antigen borne by melanoma cells. Upon colonization of mouse skin, these bacteria substantially bolstered organism-wide, T cell–driven attack of resulting tumors. The findings highlight the broader potential of microbiota-induced immunity toward host defense and open new doors to customized commensal bacteria therapies against a myriad of human diseases.

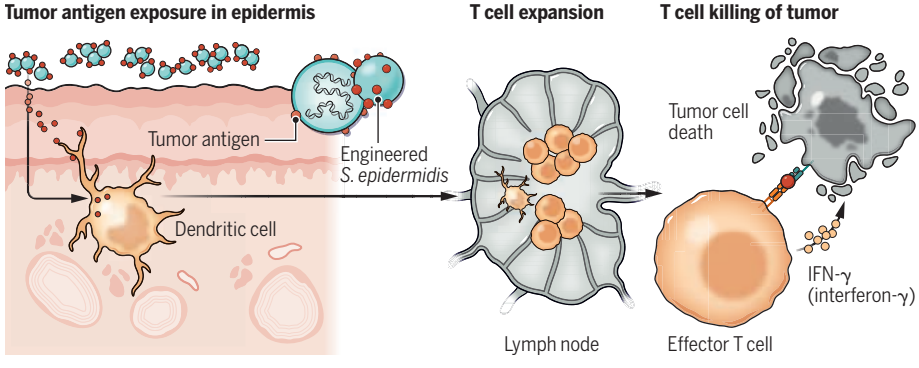
With a growing understanding of the microbiota's central role in optimal tissue fitness, strategies have emerged to therapeutically exploit this partnership. These primarily involve supplementing hosts with beneficial commensal bacteria or their metabolic products (*3*). Although still nascent, such approaches are promising: Food supplements that promote a healthy gut microbiota can rescue key health outcomes in malnourished children, while microbes associated with human pathology and treatment responsiveness have now been discovered for disorders including cancers, obesity, and allergies (*4*–*8*).

Concomitantly, it has become appreciated that individual symbionts can induce antigen-specific immune responses that dominantly influence tissue states, promoting such varied processes as inflammation, tissue regeneration, and immune tolerance, depending on the context (*1*). Indeed, when applied to mouse skin, human isolates of *S. epidermidis* drive locally increased numbers of $CD4^+$ (helper) and $CD8^+$ (cytotoxic) T lymphocytes, which preempt infections in colonized tissue through homeostatic production of inflammatory cytokines (*9*). Whether such antigen-specific responses can be systemically mobilized for targeted protective immunity throughout an organism had been unexplored until now.

Chen *et al.* genetically manipulated *S. epidermidis* to encode antigens (either test-case ovalbumin peptides or known tumor antigens) present on mouse melanoma cells. In contrast to nonengineered *S. epidermidis*, these modified commensals elicited potent T cell responses against both subcutaneous melanoma grafts and lung metastases, strikingly without causing inflammation in the skin colonized by the antigen-expressing bacteria. These "designer" commensals even impeded melanoma progression when applied

Antigen-engineered commensal bacteria

By genetically engineering the skin commensal *Staphylococcus epidermidis* to express tumor antigens, a robust effector T cell response can be induced without causing inflammation at the site of colonization in mice. When tumor cells expressing the same antigens are encountered, these T cells can be mobilized into circulation to systemically confer long-term host protection.



therapeutically (i.e., after tumor engraftment) and synergized strongly with immune checkpoint blockade to eliminate tumors and provide lasting protection against rechallenge. Thus, commensal-induced immunity might safely and systemically protect hosts against targets of choice (see the figure).

What dictates an effective commensal-guided response? At the heart of *S. epidermidis*–mediated protection was robust T cell activation. Melanoma cell killing required priming of both helper and cytotoxic T cells, each of which produced the proinflammatory cytokine interferon-γ (IFN-γ) in response to tumors. How other commensals, antigens, and sites of colonization may alter these responses is open for future study, where the blueprint established by Chen *et al.* will be a helpful guide.

Intriguingly, the antitumor T cells that had been primed by engineered *S. epidermidis* further displayed a shift from lymph-restricted "central memory" to tissue-homing "effector memory" phenotypes. Still unknown are the spatial and temporal dynamics of how these T cells can be activated, trafficked, and differentiated toward distinct memory states. Wild-type *S. epidermidis* was previously shown to generate skin-resident memory T cells during homeostasis, which were largely restricted to sites of colonization (*10*). Are these same programs co-opted to drive antitumor responses, and if so, how do they become endowed with migratory capacity? As T cell functionality is closely linked to the strength of antigen recognition, work will also be needed to determine whether host-derived tumor epitopes with weaker immunogenicity than those used by Chen *et al.* can still drive efficient cytotoxicity. Likewise, immune evasion by tumors—for example, through down-regulation of antigen processing and presentation machinery—will be critical to address for effective deployment of engineered commensal bacteria (*11*).

The findings of Chen *et al.* also raise the question of how immunity to the microbiota is compartmentalized in hosts. How do engineered commensals prime cytotoxicity to target sites, such as tumors, without disrupting homeostasis in colonized tissue? The observed combination of safety and efficacy suggests a key role for an intact, non-inflamed tissue barrier, which likely affects the means and extent to which microbes influence their host (*12*). Potentially central in this regard are epithelial cells, which are the primary interface between a host and its environment, and dendritic cells, which directly sample antigens in barrier sites to activate cognate T cells. Although both carry a battery of sensors to detect the microbiota, how each may regulate cytotoxicity to associated antigens is still unclear.

Inflammatory signals can also confer lasting epigenetic "memories" to multiple barrier constituents, including immune cells and long-lived barrier epithelial stem cells (*12*). How might such adaptations be holistically orchestrated to shape long-term immunity to engineered commensals? Elucidating these dynamics will offer important insight into the normal function of microbiota-directed immunity, as well as how this may go awry in settings of dysbiosis such as chronic inflammation and cancer (*1*, *3*, *13*).

Lastly, although Chen *et al.* highlight a new route to generate robust and specific T cell cytotoxicity, it is possible that engineered commensals might also be used to guide tolerogenic responses, especially given the intractability of certain autoimmune, allergic, and chronic inflammatory diseases. Indeed, many gut commensal bacteria can induce immune-suppressive regulatory T cells (*14*, *15*), but whether these or other symbionts might be repurposed toward targeted immune tolerance remains to be explored.

Looking forward, a toolkit of immune-modulatory commensal bacteria, expressing antigens, cytokines, and/or metabolites, could be coordinately leveraged to treat human pathologies. The findings reported by Chen *et al.* present an early but encouraging strategy to fine-tune immunity against cancer by antigen-engineering commensals, and highlight a need to continue investigating several key aspects of host–microbiota relationships that together could pave a way to safer, more effective, and widely applicable therapeutics. ■

REFERENCES AND NOTES

1. K. Honda, D. Littman, *Nature* **535**, 75 (2016).
2. Y. E. Chen *et al.*, *Science* **380**, 203 (2023).
3. M. T. Sorbara, E. G. Pamer, *Nat. Rev. Microbiol.* **20**, 365 (2022).
4. R. Y. Chen *et al.*, *N. Engl. J. Med.* **384**, 1517 (2021).
5. B. Routy *et al.*, *Science* **359**, 91 (2018).
6. V. Matson *et al.*, *Science* **359**, 104 (2018).
7. C. Depommier *et al.*, *Nat. Med.* **25**, 1096 (2019).
8. T. Feehley *et al.*, *Nat. Med.* **25**, 448 (2019).
9. S. Naik *et al.*, *Nature* **520**, 104 (2015).
10. O. J. Harrison *et al.*, *Science* **363**, aat6280 (2019).
11. S. Jhunjhunwala, C. Hammer, L. Delamarre, *Nat. Rev. Cancer* **21**, 298 (2021).
12. S. Naik, E. Fuchs, *Nature* **607**, 249 (2022).
13. G. D. Sepich-Poore *et al.*, *Science* **371**, abc4552 (2021).
14. J. Ochoa-Repáraz *et al.*, *J. Immunol.* **185**, 4101 (2010).
15. K. Atarashi *et al.*, *Nature* **500**, 232 (2013).

ACKNOWLEDGMENTS

S.M.S. is funded by the National Institutes of Health (NIH) Clinical and Translational Science Award (CTSA) through The Rockefeller University. A.G. is a Damon Runyon Cancer Research Foundation National Mah Jongg League Fellow (DRG 2409-20). The authors receive support for their work from NIH grants to E.F. (AR27883, AR050452, AR31737).

10.1126/science.adh3884

Howard Hughes Medical Institute, Robin Chemers Neustein Laboratory of Mammalian Cell Biology and Development, The Rockefeller University, New York, NY, USA. Email: fuchslb@rockefeller.edu

MEDICINE

Sleep like a bear

Reduced expression of a platelet protein protects against thrombosis during chronic immobilization

By **Mirta Schattner**

Deep vein thrombosis (DVT) is a common feature of short-term immobilization, which occurs, for example, for patients with acute paralysis. However, in chronic immobilization—e.g., owing to spinal cord injury (SCI)—the incidence of DVT equals that in the general population (*1*). The mechanisms that protect against DVT during chronic immobilization in SCI patients are unknown. On page 178 of this issue, Thienel *et al.* (*2*) report an analysis of the thromboinflammatory machinery in hibernating brown bears (*Ursus arctos*), which rarely suffer from DVT despite 6 months of immobilization during hibernation (*3*). They found that platelets from hibernating bears exhibit an antithrombotic signature characterized by decreased expression of heat shock protein 47 (HSP47) and overall reduced biomarkers of thromboinflammation compared with those of active brown bears. Further, this signature is conserved across different animal species, including mice, pigs, and humans, suggesting that this could be a therapeutic target to reduce DVT in immobilized patients.

Venous thromboembolism (VTE) is a cardiovascular disorder that encompasses two related conditions that are part of the same spectrum, namely DVT and pulmonary embolism. The spectrum of disease ranges from clinically unsuspected to clinically insignificant to massive embolism resulting in death. With an estimated incidence rate of 1 to 2 per 1000 persons per year, it is the third leading vascular disease after myocardial infarction and stroke and is a substantial contributor to the global burden of disease (*4*).

The pathophysiology of VTE is thought to be triggered by several factors, most notably the aberrant interplay between the endothelium, coagulation cascade, platelets, and the immune system. This thromboinflammation process can occur in sterile or infectious inflammatory states (*5*). Platelet activation was

Laboratorio de Trombosis Experimental e Inmunobiología de la Inflamación, Instituto de Medicina Experimental (IMEX), CONICET–Academia Nacional de Medicina, Buenos Aires, Argentina. Email: mschattner@hotmail.com

therapeutically (i.e., after tumor engraftment) and synergized strongly with immune checkpoint blockade to eliminate tumors and provide lasting protection against rechallenge. Thus, commensal-induced immunity might safely and systemically protect hosts against targets of choice (see the figure).

What dictates an effective commensal-guided response? At the heart of *S. epidermidis*–mediated protection was robust T cell activation. Melanoma cell killing required priming of both helper and cytotoxic T cells, each of which produced the proinflammatory cytokine interferon-γ (IFN-γ) in response to tumors. How other commensals, antigens, and sites of colonization may alter these responses is open for future study, where the blueprint established by Chen *et al.* will be a helpful guide.

Intriguingly, the antitumor T cells that had been primed by engineered *S. epidermidis* further displayed a shift from lymph-restricted "central memory" to tissue-homing "effector memory" phenotypes. Still unknown are the spatial and temporal dynamics of how these T cells can be activated, trafficked, and differentiated toward distinct memory states. Wild-type *S. epidermidis* was previously shown to generate skin-resident memory T cells during homeostasis, which were largely restricted to sites of colonization (*10*). Are these same programs co-opted to drive antitumor responses, and if so, how do they become endowed with migratory capacity? As T cell functionality is closely linked to the strength of antigen recognition, work will also be needed to determine whether host-derived tumor epitopes with weaker immunogenicity than those used by Chen *et al.* can still drive efficient cytotoxicity. Likewise, immune evasion by tumors—for example, through down-regulation of antigen processing and presentation machinery—will be critical to address for effective deployment of engineered commensal bacteria (*11*).

The findings of Chen *et al.* also raise the question of how immunity to the microbiota is compartmentalized in hosts. How do engineered commensals prime cytotoxicity to target sites, such as tumors, without disrupting homeostasis in colonized tissue? The observed combination of safety and efficacy suggests a key role for an intact, non-inflamed tissue barrier, which likely affects the means and extent to which microbes influence their host (*12*). Potentially central in this regard are epithelial cells, which are the primary interface between a host and its environment, and dendritic cells, which directly sample antigens in barrier sites to activate cognate T cells. Although both carry a battery of sensors to detect the microbiota, how each may regulate cytotoxicity to associated antigens is still unclear.

Inflammatory signals can also confer lasting epigenetic "memories" to multiple barrier constituents, including immune cells and long-lived barrier epithelial stem cells (*12*). How might such adaptations be holistically orchestrated to shape long-term immunity to engineered commensals? Elucidating these dynamics will offer important insight into the normal function of microbiota-directed immunity, as well as how this may go awry in settings of dysbiosis such as chronic inflammation and cancer (*1*, *3*, *13*).

Lastly, although Chen *et al.* highlight a new route to generate robust and specific T cell cytotoxicity, it is possible that engineered commensals might also be used to guide tolerogenic responses, especially given the intractability of certain autoimmune, allergic, and chronic inflammatory diseases. Indeed, many gut commensal bacteria can induce immune-suppressive regulatory T cells (*14*, *15*), but whether these or other symbionts might be repurposed toward targeted immune tolerance remains to be explored.

Looking forward, a toolkit of immune-modulatory commensal bacteria, expressing antigens, cytokines, and/or metabolites, could be coordinately leveraged to treat human pathologies. The findings reported by Chen *et al.* present an early but encouraging strategy to fine-tune immunity against cancer by antigen-engineering commensals, and highlight a need to continue investigating several key aspects of host–microbiota relationships that together could pave a way to safer, more effective, and widely applicable therapeutics. ■

REFERENCES AND NOTES

1. K. Honda, D. Littman, *Nature* **535**, 75 (2016).
2. Y. E. Chen *et al.*, *Science* **380**, 203 (2023).
3. M. T. Sorbara, E. G. Pamer, *Nat. Rev. Microbiol.* **20**, 365 (2022).
4. R. Y. Chen *et al.*, *N. Engl. J. Med.* **384**, 1517 (2021).
5. B. Routy *et al.*, *Science* **359**, 91 (2018).
6. V. Matson *et al.*, *Science* **359**, 104 (2018).
7. C. Depommier *et al.*, *Nat. Med.* **25**, 1096 (2019).
8. T. Feehley *et al.*, *Nat. Med.* **25**, 448 (2019).
9. S. Naik *et al.*, *Nature* **520**, 104 (2015).
10. O. J. Harrison *et al.*, *Science* **363**, aat6280 (2019).
11. S. Jhunjhunwala, C. Hammer, L. Delamarre, *Nat. Rev. Cancer* **21**, 298 (2021).
12. S. Naik, E. Fuchs, *Nature* **607**, 249 (2022).
13. G. D. Sepich-Poore *et al.*, *Science* **371**, abc4552 (2021).
14. J. Ochoa-Repáraz *et al.*, *J. Immunol.* **185**, 4101 (2010).
15. K. Atarashi *et al.*, *Nature* **500**, 232 (2013).

ACKNOWLEDGMENTS

S.M.S. is funded by the National Institutes of Health (NIH) Clinical and Translational Science Award (CTSA) through The Rockefeller University. A.G. is a Damon Runyon Cancer Research Foundation National Mah Jongg League Fellow (DRG 2409-20). The authors receive support for their work from NIH grants to E.F. (AR27883, AR050452, AR31737).

10.1126/science.adh3884

Howard Hughes Medical Institute, Robin Chemers Neustein Laboratory of Mammalian Cell Biology and Development, The Rockefeller University, New York, NY, USA. Email: fuchslb@rockefeller.edu

MEDICINE

Sleep like a bear

Reduced expression of a platelet protein protects against thrombosis during chronic immobilization

By **Mirta Schattner**

Deep vein thrombosis (DVT) is a common feature of short-term immobilization, which occurs, for example, for patients with acute paralysis. However, in chronic immobilization—e.g., owing to spinal cord injury (SCI)—the incidence of DVT equals that in the general population (*1*). The mechanisms that protect against DVT during chronic immobilization in SCI patients are unknown. On page 178 of this issue, Thienel *et al.* (*2*) report an analysis of the thromboinflammatory machinery in hibernating brown bears (*Ursus arctos*), which rarely suffer from DVT despite 6 months of immobilization during hibernation (*3*). They found that platelets from hibernating bears exhibit an antithrombotic signature characterized by decreased expression of heat shock protein 47 (HSP47) and overall reduced biomarkers of thromboinflammation compared with those of active brown bears. Further, this signature is conserved across different animal species, including mice, pigs, and humans, suggesting that this could be a therapeutic target to reduce DVT in immobilized patients.

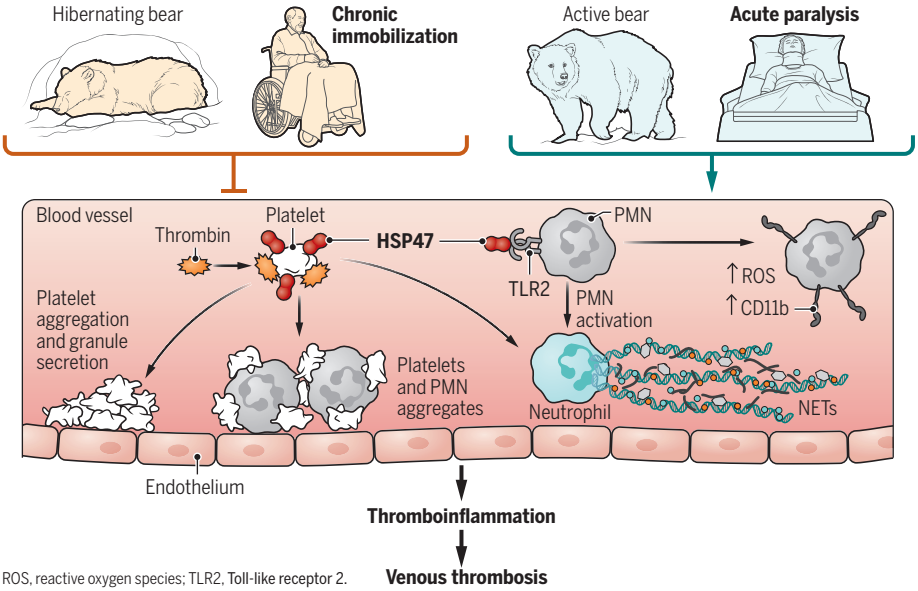
Venous thromboembolism (VTE) is a cardiovascular disorder that encompasses two related conditions that are part of the same spectrum, namely DVT and pulmonary embolism. The spectrum of disease ranges from clinically unsuspected to clinically insignificant to massive embolism resulting in death. With an estimated incidence rate of 1 to 2 per 1000 persons per year, it is the third leading vascular disease after myocardial infarction and stroke and is a substantial contributor to the global burden of disease (*4*).

The pathophysiology of VTE is thought to be triggered by several factors, most notably the aberrant interplay between the endothelium, coagulation cascade, platelets, and the immune system. This thromboinflammation process can occur in sterile or infectious inflammatory states (*5*). Platelet activation was

Laboratorio de Trombosis Experimental e Inmunobiología de la Inflamación, Instituto de Medicina Experimental (IMEX), CONICET–Academia Nacional de Medicina, Buenos Aires, Argentina. Email: mschattner@hotmail.com

Protection from venous blood clots

Platelet heat shock protein 47 (HSP47) contributes to blood clot formation by recruiting thrombin to platelets and triggering their aggregation and by activating polymorphonuclear leukocytes (PMNs). The latter generates neutrophil extracellular traps (NETs) comprising DNA and proteins that provide a scaffold for red blood cells, platelets, and procoagulant molecules. The expression of HSP47 is down-regulated during chronic immobilization, which protects against thromboinflammation and clot formation.



ROS, reactive oxygen species; TLR2, Toll-like receptor 2.

considered critical for arterial thrombosis, whereas hypercoagulation—not platelets—was thought to be the main trigger for the development of VTE. However, with the discovery of the thromboinflammation pathway over the past two decades, this vascular paradigm has changed, and platelets, particularly their interaction with inflammatory neutrophils and monocytes, are increasingly recognized as key factors in VTE (*4*). Among the various risk factors involved in the development of VTE, short-term immobility-related conditions predominate. VTE associated with immobility is triggered by restriction of venous blood flow, which leads to hypoxia of endothelial cells, and by recruitment of platelets and neutrophils, which generate extracellular DNA traps [neutrophil extracellular traps (NETs)] composed of DNA, histones, and antimicrobial proteins. NETs provide a scaffold for red blood cells, platelets, and procoagulant molecules that play key roles in the development of thromboinflammation (*5*).

HSPs are highly conserved molecular chaperones that support cellular functions during homeostasis and under environmental stress conditions. In addition, HSPs are secreted into the extracellular space and blood and can regulate inflammation and immune responses. HSP47 is mainly found in collagen-producing cells and contributes to the stabilization of the triple helix structures of procollagen. Although platelets do not produce collagen, they interact with subendothelial matrix collagen, which is exposed upon vessel wall injury and facilitates thrombus formation. Surface-bound HSP47 was previously identified in platelets as a key mediator of hemostasis that stabilizes collagen-induced platelet adhesion and promotes aggregation and thrombus formation (*6*). Thienel *et al.* extended these findings by showing that human platelet HSP47 facilitates the binding of thrombin, which triggers platelet aggregation and granule secretion, as well as the formation of platelet-neutrophil aggregates and the formation of NETs. Moreover, they showed that HSP47 binds to neutrophil Toll-like receptor 2 (TLR2), leading to the production of intracellular reactive oxygen species, increased expression of the integrin CD11b, and formation of NETs in a platelet-independent manner (see the figure).

Platelets are the main circulating cells that contribute to the formation of NETs. Platelet-neutrophil close contact (mediated by membrane cell adhesion molecules as well as the release of soluble factors in the synapses of both cells) is required for NET generation (*7*). Thienel *et al.* showed that the interplay between platelets and neutrophils is one of the main causes of thrombogenic differences in acute and chronic immobilization. They found that significantly reduced expression of HSP47 in hibernating bears and SCI patients correlates with a decreased effector response between platelets and neutrophils, resulting in the absence of a prothrombotic state despite long-term immobilization. Moreover, platelet HSP47 expression in healthy individuals decreased substantially after 27 days of voluntary bed rest. This suggests that the HSP47-TLR2 axis is a promising target to prevent thromboinflammation during immobilization.

However, it is important to consider that this study was conducted in immobilized, otherwise healthy chronic SCI patients and healthy volunteers. Acute immobilization–associated VTE usually occurs in patients who have other predisposing cardiovascular risk factors or acute illness, such as trauma or infection. Do these prothrombotic conditions overcome the effects of reduced HSP47 expression? Moreover, thromboinflammation and VTE may occur not only in immobilization but also in other clinical conditions, such as diabetes, cancer, and autoimmune diseases. Could HSP47 be a biomarker for the occurrence of VTE in these diseases? Strenuous physical exercise has been linked to the appearance of DVT in the upper limbs (*8*) as well as in the lower limbs of marathon athletes (*9*). If HSP47 expression is down-regulated under immobile conditions, how does intense exercise affect its expression? Another feature of thromboinflammation is the decreased antithrombotic and anti-inflammatory activity of the endothelium, mainly resulting from its interaction with activated platelets and neutrophils (*10*). Because there is virtually no therapy to prevent this endothelial dysfunction, it would be interesting to investigate whether platelet HSP47 can modulate endothelial cell activation.

Notably, much more is known about the risk factors and molecules that trigger thrombosis than about the protective mechanisms. Identifying signaling pathways that down-regulate prothrombotic and/or up-regulate thromboprotective molecules will not only enable us to complement existing therapies for the prevention of DVT in the various thromboinflammatory diseases but also to select safer targets with less host impact. The study by Thienel *et al.* shows that looking at nature can be a good way to learn about human biology. ■

REFERENCES AND NOTES

1. L. N. Godat, L. Kobayashi, D. C. Chang, R. Coimbra, *J. Trauma Acute Care Surg.* **78**, 475 (2015).
2. M. Thienel *et al.*, *Science* **380**, 178 (2023).
3. O. Fröbert, A. M. Frøbert, J. Kindberg, J. M. Arnemo, M. T. Overgaard, *J. Intern. Med.* **287**, 263 (2020).
4. A. M. Wendelboe, G. E. Raskob, *Circ. Res.* **118**, 1340 (2016).
5. K. Stark, S. Massberg, *Nat. Rev. Cardiol.* **18**, 666 (2021).
6. P. Sasikumar *et al.*, *J. Thromb. Haemost.* **16**, 946 (2018).
7. A. Carestia, T. Kaufman, M. Schattner, *Front. Immunol.* **7**, 271 (2016).
8. R. E. Keller, D. P. Croswell, G. I. S. Medina, T. T. W. Cheng, L. S. Oh, *J. Shoulder Elbow Surg.* **29**, 2417 (2020).
9. C. M. Hull, C. L. Hopkins, N. J. Purdy, R. C. Lloyd, J. A. Harris, *Scand. J. Med. Sci. Sports* **25**, 699 (2015).
10. H. Zhang *et al.*, *Clin. Transl. Med.* **13**, e1170 (2023).

10.1126/science.adh3276

MATERIALS SCIENCE

A mechanically one-way material

A material with asymmetric mechanical responses offers diverse potential applications

By **Bohan Sun**[1,2] *and* **Sung Hoon Kang**[1,2]

Mechanical movement or energy, in the form of wind (*1*), waves (*2*), and vibration (*3*), is abundant. However, most of this energy is scarcely used for reasons that include difficulty of collection or manipulation. Furthermore, unwanted propagation of mechanical energy can result in detrimental effects, including sound pollution, injury to humans, and damage to vehicles and infrastructures. But what if mechanical energy could be controlled or used through materials that allow preferential transport of mechanical energy in a particular direction? On page 192 of this issue, Wang *et al.* (*4*) report such a material that shows different mechanical responses based on the loading direction ("nonreciprocity"). This could open doors for applications in guiding, controlling, and damping mechanical energies, allowing, for example, guided sound propagation, directional mass transport or concentration of mechanical energy using random mechanical movement, and the ability to steer away from impact, earthquake, or vibration.

Recently, there has been keen interest in nonreciprocal properties in various fields, including optics (*5*), acoustics (*6*), quantum systems (*7*), and mechanics (*8*), because they can provide distinct capabilities such as one-way light propagation, unidirectional acoustic isolation, single-directional wave amplification, and one-way mechanical wave transmission. Although nonreciprocal mechanical behaviors have been demonstrated as a structure or a system, using a fishbone structure (*8*) or a "robotic" unit consisting of beams, sensor, motor, and microcontroller (*9*), it has been challenging to realize such behaviors in a material owing to the intricacy of current approaches that are required to achieve mechanical nonreciprocity. Realizing nonreciprocal materials without a complex structure or using a system approach can provide good designability so that parts can be made in any shape and can be assembled as needed. Wang *et al.* realized nonreciprocal materials by embedding aligned thin sheets of graphene oxide in a water-based gel (hydrogel). The oriented graphene oxide thin sheets switch between two different microstructures [unbuckled versus buckled state (*10*)] with corresponding different mechanical properties, based on the loading direction.

To understand why the mechanical properties change depending on the loading direction, imagine a sheet of paper. If you hold the paper with two hands and push along the plane of the paper, it easily bends out of plane, that is, it buckles under the compressive force. However, if you try to stretch the paper, it is difficult to do and it stays in the unbuckled state under tensile force. Likewise, when a force is applied to the nonreciprocal gel from right to left, the compressive force is applied to the inclined graphene oxide sheets, leading to buckling of the sheets and easy deformation along the force direction. But, if the force is applied from the opposite direction, the graphene oxide sheets need to be stretched, thus it is difficult to deform along the opposite direction.

"This could open doors for applications in guiding, controlling, and damping mechanical energies..."

Nonreciprocal properties endow materials with the ability to transform symmetric mechanical stimuli into an asymmetric response. Thus, nonreciprocal properties can open new opportunities where mechanical motion or energy exists—including mass transport, energy harvesting, and protection—because they can facilitate the steering of mechanical energy. Wang *et al.* demonstrated intriguing applications of harnessing nonreciprocity for generating mass transport in a desired direction such as horizontal, rotational, and even vertical directions against gravity by using symmetric vibrations. When symmetric vibrations were applied, the resulting deformation switched between large and small depending on the loading direction. As a result, there was a net directional movement, showing an asymmetric mechanical response. Because the nonreciprocal material is made of a hydrogel, it is soft and flexible. Such characteristics can be useful for wearable devices and in aquatic environments. Additionally, Wang *et al.* showed that when *Caenorhabditis elegans* (a nematode that has been studied as a model of development and behavior) were placed on the nonreciprocal gel, they exhibited directional migration without any gradient stimuli, showing potential to control their behavior.

Beyond the demonstration by Wang *et al.*, nonreciprocal mechanical properties could be useful for protecting objects—including human tissues, sensitive equipment, or buildings—by minimizing impact or vibration energy transfer. Such materials can also find applications in focusing mechanical energy by guiding it to the desired direction to amplify the mechanical energy or harvest it as another form of energy, for example, electricity using piezoelectric materials that convert mechanical energy into electrical energy (*11*).

The future research opportunities for nonreciprocal materials include scalability, programmability, material palette, and tunability. To use nonreciprocal materials in a wider range of applications, a more scalable synthesis approach would be beneficial. In the study of Wang *et al.*, nonreciprocity is realized for limited types of simple loading conditions such as shear or symmetric vibration along a specific direction. Realizing nonreciprocal materials for complex loading conditions, such as from any random direction, would expand the potential applications. In addition, the study of Wang *et al.* was based on materials with limited load-bearing capability, which can fail easily and lose their properties owing to evaporation of water. Realizing nonreciprocal materials with good load-bearing capability and in a solvent-free system could be useful for structural applications in the ambient environment. Finally, the ability to switch between reciprocal and nonreciprocal properties could further expand the capability to control mechanical behaviors because such materials could pass or block mechanical energy on demand. Nonreciprocal materials can lead us in exciting new directions that could only be imagined before. ■

[1]Department of Mechanical Engineering, Johns Hopkins University, Baltimore, MD, USA. [2]Hopkins Extreme Materials Institute, Johns Hopkins University, Baltimore, MD, USA. Email: shkang@jhu.edu

REFERENCES AND NOTES

1. S. Orrego *et al.*, *Appl. Energy* **194**, 212 (2017).
2. J. Cruz, *Ocean Wave Energy: Current Status and Future Perspectives* (Springer, 2007).
3. S. P. Beeby *et al.*, *J. Micromech. Microeng.* **17**, 1257 (2007).
4. X. Wang *et al.*, *Science* **380**, 192 (2023).
5. L. Feng *et al.*, *Science* **333**, 729 (2011).
6. R. Fleury *et al.*, *Science* **343**, 516 (2014).
7. V. Peano *et al.*, *Phys. Rev. X* **6**, 041026 (2016).
8. C. Coulais *et al.*, *Nature* **542**, 461 (2017).
9. M. Brandenbourger *et al.*, *Nat. Commun.* **10**, 4608 (2019).
10. S. Shan *et al.*, *Adv. Mater.* **27**, 4296 (2015).
11. S. Lee *et al.*, *Nano Energy* **103**, 107846 (2022).

ACKNOWLEDGMENTS

The authors acknowledge support from the Air Force Office of Scientific Research (award number FA9550-21-1-0368, program manager B.-L. Lee), the Johns Hopkins University Catalyst Award, and the Hanwha Non-Tenured Faculty Award.

10.1126/science.adh3098

POLICY FORUM

ARTIFICIAL INTELLIGENCE

Rethink reporting of evaluation results in AI

Aggregate metrics and lack of access to results limit understanding

By **Ryan Burnell[1], Wout Schellaert[2], John Burden[1,3], Tomer D. Ullman[4], Fernando Martinez-Plumed[2], Joshua B. Tenenbaum[5], Danaja Rutar[1], Lucy G. Cheke[1,6], Jascha Sohl-Dickstein[7], Melanie Mitchell[8], Douwe Kiela[9], Murray Shanahan[10,11], Ellen M. Voorhees[12], Anthony G. Cohn[13,14,15,16], Joel Z. Leibo[10], Jose Hernandez-Orallo[1,2,3]**

Artificial intelligence (AI) systems have begun to be deployed in high-stakes contexts, including autonomous driving and medical diagnosis. In contexts such as these, the consequences of system failures can be devastating. It is therefore vital that researchers and policymakers have a full understanding of the capabilities and weaknesses of AI systems so that they can make informed decisions about where these systems are safe to use and how they might be improved. Unfortunately, current approaches to AI evaluation make it exceedingly difficult to build such an understanding, for two key reasons. First, aggregate metrics make it hard to predict how a system will perform in a particular situation. Second, the instance-by-instance evaluation results that could be used to unpack these aggregate metrics are rarely made available (*1*). Here, we propose a path forward in which results are presented in more nuanced ways and instance-by-instance evaluation results are made publicly available.

Across most areas of AI, system evaluations follow a similar structure. A system is first built or trained to perform a particular set of functions. Then, the performance of the system is tested on a set of tasks relevant to the desired functionality of the system. In many areas of AI, evaluations use standardized sets of tasks known as "benchmarks." For each task, the system will be tested on a number of example "instances" of the task. The system would then be given a score for each instance based on its performance, e.g., 1 if it classified an image correctly, or 0 if it was incorrect. For other systems, the score for each instance might be based on how quickly the system completed its task, the quality of its outputs, or the total reward it obtained. Finally, performance across the various instances and tasks is usually aggregated to a small number of metrics that summarize how well the system performed, such as percentage accuracy.

But aggregate metrics limit our insight into performance in particular situations, making it harder to find system failure points and robustly evaluate system safety. This problem is also worsening as the increasingly broad capabilities of state-of-the-art systems necessitate ever more diverse benchmarks to cover the range of their capabilities. This problem is further exacerbated by a lack of access to the instance-by-instance results underlying the aggregate metrics, making it difficult for researchers and policy-makers to further scrutinize system behavior.

AGGREGATE METRICS

Use of aggregate metrics is understandable. They provide information about system performance "at a glance" and allow for simple comparisons across systems. But aggregate performance metrics obfuscate key information about where systems tend to succeed or fail (*2*). Take, for example, a system that was trained to classify faces as male or female that achieved classification accuracy of 90% (*3*). Based on this metric, the system appears highly competent. However, a subsequent breakdown of performance revealed that the system misclassified females with darker skin types a staggering 34.5% of the time, while erring only 0.8% of the time for males with lighter skin types. This example demonstrates how aggregation can make it difficult for policymakers to determine the fairness and safety of AI systems.

Compounding this problem, many benchmarks include disparate tasks that are ultimately aggregated together. For example, the Beyond the Imitation Game Benchmark (BIG-bench) for language models includes over 200 tasks that evaluate everything from language understanding to causal reasoning (*4*). Aggregating across these disparate tasks—as the BIG-bench leaderboard does—reduces the rich information in the benchmark to an overall score that is hard to interpret.

It is also easy for aggregation to introduce unwarranted assumptions into the evaluation process. For example, a simple average across tasks implicitly treats every task as equally important—in the case of BIG-bench, a sports understanding task has as much bearing on the overall score as a causal reasoning task. These aggregation decisions have huge implications for the conclusions that are drawn about system capabilities, yet are seldom considered carefully or explained.

Aggregate metrics depend not only on the capability of the system but also on the characteristics of the instances used for evaluation. If the gender classification system above were reevaluated by using entirely light-skinned faces, accuracy would skyrocket, even though the system's ability to classify faces has not changed. Aggregate metrics can easily give false impressions about capabilities when a benchmark is not well constructed.

Problems and trade-offs that arise when considering aggregate versus granular data and metrics are not specific to AI, but they are exacerbated by the challenges inherent in AI research and the research practices of the field. For example, machine learning evaluations usually involve randomly splitting data into training, validation, and test sets. An enormous amount of data is required to train state-of-the-art systems, so these datasets are often poorly curated and lack the detailed annotation necessary to conduct granular analyses. In addition, the research culture in AI is centered around outdoing the current state-of-the-art performance, as evidenced by the many lea-

[1]Leverhulme Centre for the Future of Intelligence, University of Cambridge, Cambridge, UK. [2]Valencian Research Institute for Artificial Intelligence, Universitat Politècnica de Valencia, València, Spain. [3]Centre for the Study of Existential Risk, University of Cambridge, Cambridge, UK. [4]Department of Psychology, Harvard University, Cambridge, MA, USA. [5]Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA. [6]Department of Psychology, University of Cambridge, Cambridge, UK. [7]Brain team, Google, Mountainview, CA, USA. [8]Santa Fe Institute, Santa Fe, NM, USA. [9]Stanford University, Stanford, CA, USA. [10]DeepMind, London, UK. [11]Department of Computing, Imperial College London, London, UK. [12]National Institute of Standards and Technology (Retired), Gaithersburg, MD, USA. [13]School of Computing, University of Leeds, Leeds, UK. [14]Alan Turing Institute, London, UK. [15]Tongji University, Shanghai, China. [16]Shandong University, Jinan, China. Email: rb967@cam.ac.uk

derboards, competitions, and challenges that offer prestige or monetary prizes (*5*). This research culture emphasizes aggregate metrics and incentivizes immediate publication of new findings at the expense of robust evaluation practices. In addition, the strict space restrictions and fast turnaround times enforced by high-impact AI conferences disincentivize researchers from reporting results in a granular way. Finally, the primary focus of most publications in AI is not the experimental results themselves but the new algorithms or techniques being evaluated. As a result, less attention has been paid in AI to issues around experimental design and reporting than in other fields such as psychology or physics.

INSTANCE-BY-INSTANCE EVALUATION

There are many situations in which the community might want to conduct analyses that go beyond those reported in a paper. For example, researchers often seek to investigate the extent to which AI systems are biased against minority or disadvantaged populations. It is also frequently useful to scrutinize patterns of performance to debug systems or to determine their safety in a particular deployment context. Moreover, in areas such as robotics and reinforcement learning, examining the trajectory of a system or its sequence of actions can help researchers better understand a system's strategy.

These supplemental evaluations often require access to the instance-by-instance evaluation results (i.e., the outputs and scores of the systems for each instance). But these results are rarely made available—one recent analysis found that only 4% of papers in top AI venues fully report the evaluation results (*1*). As systems and benchmarks continue to grow in size and complexity, it is becoming increasingly costly for researchers to recreate results by conducting their own evaluations. The result is that researchers and policy-makers are increasingly forced to take reported results at face value or to incur substantial and unnecessary costs just to recreate them.

A PATH FORWARD

To address these problems, we propose a broad set of solutions (see the box).

Moving beyond aggregate metrics

It is important that in-depth performance breakdowns are presented instead of, or alongside, aggregate metrics. Breakdowns can be created by identifying features of the problem space that might be relevant to performance and using those features to analyze, visualize, and predict performance (*6*). These kinds of granular analyses are not yet widely employed. However, researchers focused on system robustness and fairness have begun to demonstrate how valuable they can be.

For example, granular analyses can help researchers explore the concepts that a system has learned. Researchers examined the patterns of errors made by systems on a spatial reasoning benchmark designed to vary systematically on important features of the problem space (*7*). The researchers found that the systems performed much worse on problems involving the concept of "boundary" than on problems involving the concepts of "top" or "bottom," suggesting that the abstract concept of boundary is more difficult for systems to learn.

Instance-level analyses are essential for identifying "Clever Hans" phenomena in which a system can perform well by relying on unintended patterns in the dataset. For example, a computer vision system that was excellent at classifying images into categories such as "ship" or "horse" (*8*) was ultimately shown to have not really learned to identify ships or horses. Instead, it had learned to distinguish categories based on the surrounding background or watermarks naming the source of the image—features that the system could not rely on in the real world.

Granular analyses can also allow for more meaningful comparisons between AI systems and humans. For example, though an AI system was better overall at breast cancer screening than six human radiologists (*9*), an in-depth error breakdown showed that the AI system failed to detect various cancers that were detected by all six radiologists. These errors could not be easily explained by the researchers, suggesting that further investigation is needed to understand how the system detects cancers and why it failed in these cases. These findings demonstrate the complementary value of human and AI screenings in a way that aggregate metrics could not.

Of course, granular analysis approaches are not without their challenges. Annotating the features of each instance can be labor intensive. Granular reporting usually needs more space in publications, although detailed breakdowns can be provided in supplementary materials or online. Granular analyses also often involve slicing the data into smaller pieces, so care must be taken to ensure that findings are not simply artifacts of the data. These issues can be avoided by including a large and diverse range of instances, controlling for multiple comparisons, and specifying important features a priori where possible. Finally, deciding which features to include in performance breakdowns takes time, thought, and expertise. Rigorous theoretical and empirical work may be needed to build an understanding of the problem space. Data-driven approaches can help with this process—regression analyses or deep-learning models can be used to identify features that are predictive of performance.

Ultimately, the best way of presenting evaluation results will depend on the context. The costs of performing granular analyses must be weighed against the potential consequences of system failures, and there is no one-size-fits-all solution. But we think that a shift toward more granular reporting would be a win-win situation for developers and the wider

Guidelines for robust evaluation practices

We recommend that researchers and organizations adopt these guidelines to make it easier for the community to understand system capabilities and conduct follow-up analyses.

1. Wherever possible, reporting of system performance should be granular, with breakdowns across features of the problem space that have been either hypothesized or empirically shown to affect performance. Aggregation decisions should be clearly explained, and analyses conducted to explore system performance should be described.
2. Benchmarks should be designed to test specific capabilities and to systematically vary on important features of the problem space. Benchmark instances should be annotated to allow for granular analyses to be conducted.
3. All recorded evaluation results (e.g., success or failure, response time, partial or full trace, cumulative reward) for each system on each instance should be made available. These data can be reported in supplementary materials or uploaded to a public repository. In cases of cross validation or hyper-parameter optimization, results should ideally be reported for each run and validation split separately.
4. To enable researchers to conduct follow-up analyses, information about each test instance used in an evaluation should be made available, including data labels and all annotated features of those instances.

community in most subdisciplines of AI. Granular analyses would help developers pinpoint system weaknesses to guide improvements and avoid potentially catastrophic failures. Granular reporting would ease the evaluation burden that is currently placed on external groups such as algorithmic justice organizations, which often lack the resources and expertise to evaluate systems in detail.

These changes to reporting must go hand in hand with changes to how benchmark tasks are constructed. How well a system performs across different situations cannot be evaluated unless the benchmark comprehensively covers the problem space. The commonly used approach of collating a large dataset and randomly splitting it to create a test set does not assure this coverage, so alternative approaches to benchmark construction must be considered. For example, one can design tasks that test for specific concepts or cognitive abilities and ensure that task instances systematically vary on important features of the problem space (*7, 10, 11*). In this endeavor, techniques such as procedural or adversarial generation of task instances might be useful.

Bringing about these changes in research culture will require the participation and support of the entire community. Within academia, we recommend that publications report granular analyses wherever possible, and that reviewers and editors ask for performance breakdowns when they are not provided. It might also be valuable to alter the space limits in conference publications to enable in-depth reporting of evaluation results. More broadly, we think the field needs to reckon with the potentially detrimental effects of leaderboards and competitions on evaluation practices.

We also recommend that private organizations consider incorporating guidelines for granularity and aggregation into their wider evaluation and reporting practices (*12*). We are encouraged to see growing support for "model cards" that explain a system's architecture and justify evaluation decisions (*13*). Policy-makers should bear in mind the need for granular analyses when creating guidelines or safety standards for specific applications. For instance, the recently proposed Minimum Information for Medical AI Reporting (MINIMAR) standard could be modified to ask for explanations of aggregation decisions and performance breakdowns across features of the problem space.

Ensuring the availability of instance-by-instance evaluation results

A growing open-science movement has laid the groundwork for moving toward making instance-by-instance evaluation more common. Drawing on lessons from multiple disciplines, many have argued that a lack of transparency and reproducibility in AI research could stifle progress and lead to dangerous misestimations of AI capabilities (*5, 14, 15*). In response, various initiatives have been set up to promote code and data sharing, such as the Hugging Face Hub, the Machine Learning Open Source Software section of the *Journal of Machine Learning Research*, GitHub, OpenML, Papers With Code, and the Open Science Framework. However, instance-by-instance evaluation results are rarely included on these platforms. There have been few incentives to put in the extra work needed to clean, document, release, and maintain these results. Furthermore, many researchers fear that subsequent analyses might discover flaws or biases in their systems, or that they will be "scooped" by other researchers (*5*).

"...a shift toward more granular reporting would be a win-win situation for developers and the wider community in most subdisciplines of AI."

It is vital to incentivize the release of instance-by-instance results. In other disciplines, various kinds of requirements, incentives, and nudges have been implemented for similar purposes. For example, changes to journal and conference reporting guidelines to include instance-by-instance evaluation results could help encourage the sharing of these results. It should also be possible to ensure that researchers are credited for subsequent uses of their results—perhaps by giving the results a unique identifier that other researchers can cite.

Given that many AI systems are developed or deployed by nonacademic organizations, it is important to consider how policy-makers and industry organizations can encourage the sharing of results. For example, funding agencies could require the release of instance-by-instance results as a condition of funding, and private organizations could be encouraged to share instance-by-instance results whenever they publish preprints or press releases involving system evaluations. These solutions would complement wider efforts of groups such as the European Centre for Algorithmic Transparency to encourage transparency in AI. We recognize that there are some situations in which instance-by-instance results cannot be released (e.g., owing to privacy concerns or practical constraints), but in most cases it should be possible to do so. Even if no features were annotated, releasing instance-by-instance results would still allow other researchers to extract features themselves and perform additional analyses as long as the benchmark or test data are obtainable.

Some successful examples of results-sharing in AI give us confidence that these changes in broader research culture are possible. For example, researchers who developed the Holistic Evaluation of Language Models (HELM) benchmark made instance-by-instance results available for a variety of models across the entire benchmark. If other fields such as psychology and medicine can make progress on these issues even in the face of considerable data privacy challenges, AI should be able to do the same. ■

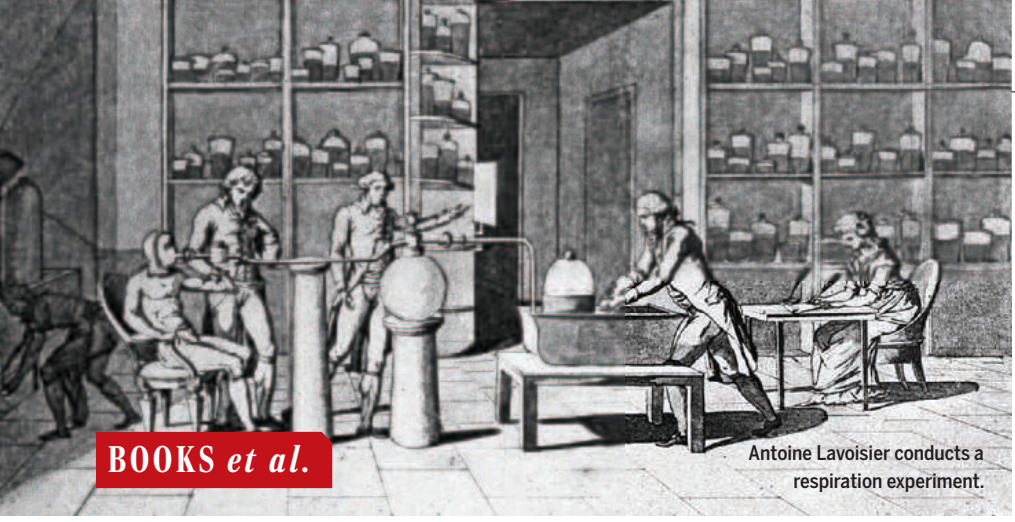
REFERENCES AND NOTES

1. O. E. Gundersen, S. Kjensmo, *Proceedings of the AAAI Conference on Artificial Intelligence* (Association for the Advancement of Artificial Intelligence, 2018), vol. 32, pp. 1644–1651.
2. B. Nushi, E. Kamar, E. Horvitz, *Proceedings of the Sixth AAAI Conference on Human Computation and Crowdsourcing* (Association for the Advancement of Artificial Intelligence, 2018), vol. 6, pp. 126–135.
3. J. Buolamwini, T. Gebru, *Proceedings of the 1st Conference on Fairness, Accountability and Transparency* (PMLR, 2018), vol. 81, pp. 77–91.
4. A. Srivastava *et al*., arXiv:2206.04615 (2022); http://arxiv.org/abs/2206.04615.
5. J. Pineau *et al*., *J. Mach. Learn. Res*. **22**, 1 (2021).
6. R. Burnell *et al*., in *Proceedings of the Thirty-First International Joint Conference on Artificial Intelligence*, Vienna, 23 to 29 July 2022 (International Joint Conferences on Artificial Intelligence), pp. 2827–2835.
7. V. V. Odouard, M. Mitchell, arXiv:2206.14187 (2022); https://arxiv.org/abs/2206.14187.
8. S. Lapuschkin *et al*., *Nat. Commun*. **10**, 1096 (2019).
9. S. M. McKinney *et al*., *Nature* **577**, 89 (2020).
10. D. Kiela *et al*., in *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (Association for Computational Linguistics, 2021), pp. 4110–4124.
11. J. Z. Leibo *et al*., in *Proceedings of the 38th International Conference on Machine Learning*, 18 to 24 July 2021 (PMLR), vol. 139, pp. 6187–6199.
12. I. D. Raji *et al*., arXiv:2001.00973 (2020); https://arxiv.org/abs/2001.00973.
13. M. Mitchell *et al*., in *Proceedings of the Conference on Fairness, Accountability, and Transparency* (Association for Computing Machinery, 2019), pp. 220–229.
14. B. Haibe-Kains *et al*., *Nature* **586**, E14 (2020).
15. M. Hutson, *Science* **359**, 725 (2018).

ACKNOWLEDGMENTS

We are grateful to J. Pineau and S. Russell for their feedback on the manuscript and to M. Brundage and L. Ahmad for helpful discussions about the need for instance-by-instance results for foundation models. Funding support was provided by the Future of Life Institute under grant RFP2-152 (J.H.O. and J.B.); US Defense Advanced Research Projects Agency HR00112120007 (RECoG-AI) (J.H.O., L.C., R.B., J.B., and D.K.); grant PID2021-1228300B-C42 (SFERA) funded by MCIN/AEI/10.13039/501100011033 and "ERDF A way of making Europe" (J.H.O., F.M.P., and W.S.); INNEST/2021/317 [Project cofunded by the European Union with the "Programa Operativo del Fondo Europeo de Desarrollo Regional (FEDER) de la Comunitat Valenciana 2014-2020"] (J.H.O. and W.S.); and UPV (Vicerrectorado de Investigación) grant PAI-10-21 (J.H.O. and W.S.).

10.1126/science.adf6369



BOOKS *et al.*

Antoine Lavoisier conducts a respiration experiment.

Elixir: A Parisian Perfume House and the Quest for the Secret of Life
Theresa Levitt
Harvard University Press, 2023. 320 pp.



HISTORY OF CHEMISTRY

The perfumer before Pasteur

A historian recounts the odorous origins of the first evidence for molecular structure

By **Michael D. Gordin**

For more than 3 years, loss of the sense of smell has been a telltale harbinger of infection for a considerable fraction of those who contract COVID-19, sometimes with lasting aftereffects. University of Mississippi historian Theresa Levitt found hers "strangely altered" in March 2020, while she was writing *Elixir*, a book that traces the intense connections between smell, health, science, and social upheaval. Even without this present context, the book would be almost impossible to put down.

Elixir concludes with Louis Pasteur's discovery that many organic molecules synthesized in a laboratory, although chemically identical with respect to composition, differ in structure from their counterparts isolated from living sources. Nature is left-handed; the laboratory is ambidextrous. Most of the book's excitement comes from the improbable path to reach this end point, which offered the first strong support for molecular structure and thus the atomic hypothesis. Written with the propulsive flow of a novel, it unfolds in two interconnected but sequential stories, each following a scientific hero. Neither hero is Pasteur.

Levitt first traces the fate of Édouard Laugier (1807–1869), scion of the respectable perfume business his grandfather Blaise had established on rue Bourg-l'Abbé in Paris in the 1770s. Blaise hailed from Grasse, in the south of France, renowned for its connection with his trade and boasting traditions of extracting scents from all sorts of botanical substances. The family prospered and developed higher professional aspirations. Édouard caught the chemistry bug and during the 1830s used his perfume shop as a site for scientific investigation as well as adornment. It was not easy, which is why *Elixir* is so much fun to read.

If you recall your European history, it will come as no surprise that the Laugiers were swept up in the tumult of the French Revolution. If you do not recall your European history, so much the better: You will receive a whirlwind tour from the point of view of pomades, perfumes, and eau de cologne.

Two teasers will not spoil the fun: Chemist Antoine Lavoisier's efforts to save the government's gunpowder supply were stymied by the demand for soap, as both these items required alkalis from the ashes of the barilla plant, whose supply Spain had cut off. More than a decade later, we see Napoleon bathing in and drinking gallons of perfume. (Perfumers did not replace the ethyl alcohol with poisonous methyl alcohol until 1855 in Britain and 1872 in France, when the switch was made for tax reasons.)

Halfway through the book, we encounter a new protagonist: a cantankerous small-town chemist Laugier rescued from his impoverished garret and brought into the shop's laboratory. This is Auguste Laurent, and the fact that you probably have never heard of him is an injustice of history that Levitt remedies.

Earlier than anyone, Laurent proposed theories of molecular substructure to understand organic chemistry. He also had a genius for alienating people and antagonizing everyone who was anyone in the Parisian scientific world, leading to his intellectual exile to Sèvres, to Bordeaux, and to Laugier's shop. (An exception was the equally crotchety physicist Jean-Baptiste Biot, an especially lively figure in *Elixir*.)

We watch Laurent turn every attempt to aid his career into a new opportunity to cause offense. He died in penury in 1853. But at a crucial moment, he collaborated with another young chemist from the provinces on his theories of chemical structure—Pasteur—who subsequently removed references to Laurent from his pathbreaking dissertation.

Laurent's story, rich in chemical history, is much leaner with respect to the politics and culture that beset Laugier's family, a consequence of the rapid professionalization of scientists, especially chemists. Levitt begins with an astounding tour of the guilds of glovers, spicers, distillers, perfumers, vinegar-makers, and apothecaries who were enmeshed in the upheavals of urban revolution. The academic laboratory artificially shut those economic and social pressures out such that the politics remaining, as Laurent's sad story reveals, tended to be of the petty academic variety.

Levitt's detailed research broaches one of the key questions in the history of science: How do ideas that are beyond the scope of their time ever find takers? In Laurent's case, it required a detour through an alternative space—Laugier's perfumery—to build up slowly (alas, also posthumously) enough evidence to change people's minds. ■

10.1126/science.adg9670

The reviewer is at the Department of History, Princeton University, Princeton, NJ 08544, USA.
Email: mgordin@princeton.edu

SCIENCE, SEX, AND GENDER



PODCAST

The study of sex and gender has long affected scientific thought, and in 2023, *Science* will explore some of these influences in a limited podcast series. This week on the *Science* podcast, the series' host, Angela Saini, and scholarly adviser Anne Fausto-Sterling discuss the impetus for the project and give listeners a sneak peek at the titles and authors who will be featured. https://bit.ly/3LUwdqY

10.1126/science.adh8326

HEALTH AND MEDICINE

Understanding the menstrual cycle

A polemical history shines a scientific spotlight on a remarkable and long-overlooked process

By **Lara Freidenfelds**

At one of the first professional conferences I attended as a history of science PhD student, I told a retired physician that I was writing my dissertation on the history of people's experiences of menstruation. "That is not interesting!" he shouted, unwilling to hear another word. I am therefore tickled by anthropologist Kate Clancy's tongue-in-cheek description of her life's work as a feminist "spite project" to demonstrate just how interesting and meaningful the menstrual cycle actually is. And demonstrate it she does.

In *Period*, Clancy wears her unapologetically progressive politics on her sleeve—the book is explicitly feminist and anti-colonialist. These interpretive lenses drive Clancy's commitment to her research and inspire the questions she asks.

In energetic and funny prose, Clancy castigates Western societies, especially scientists and physicians, for menstrual stigmas both ancient and modern. Throughout the book, Clancy points out how the language and assumptions of white, Western patriarchy have shaped what we know about menstruation and how we talk about it. And she flips the script, to great effect. For example, getting rid of "bad blood" and the "sloughing" of the endometrial lining become, in Clancy's retelling, an amazing process of "tissue remodeling."

Bathed in menstrual blood containing inflammation- and repair-managing cells, the uterus rebuilds its lining in a remarkably short time, without the scarring that typically occurs during tissue repair. When sperm enter the picture, and the uterus allows or facilitates access to a prepared egg, Clancy describes the ensuing process as "gamete fusion," connoting equality of agency, rather than "fertilization," avoiding the trope of an active male fertilizing a passive female.

Clancy's discussion of debates over the evolutionary purpose of menstruation is a treat. Does menstruation prevent infection from sperm? Does it save energy? Does it reduce preeclampsia risk? She gracefully demonstrates how a scientist reasons from her sources, concluding that a single explanation does not need to "win."

The book is more persuasive in its approach to science than to history. Clancy tends to reduce history to an analysis of oppressors exercising power over the oppressed and knowledge production as an exercise of power. She homes in on the ways in which 19th- and 20th-century eugenics theories shaped the science of menstrual cycles, for example, claiming that a eugenic focus on categorizing people according to normal curves led to our concern with "normal" cycles, one that leaves far too many people feeling that their bodies are "abnormal" and therefore bad. Although this critique is useful as far as it goes, interest in the regularity of menstrual cycles extends back millennia.

The term "menstruation" comes from "moon," and cultures all over the world have long connected menstruation to lunar cycles. Traditional names for menstruation such as "period" and "monthlies" in English and "les regles" ("the rules") in French connote regularity. Historical medical treatises and women's personal recipe books alike routinely contained recipes for "emmenagogues"—drugs to promote regular menstrual flow—and considered regular menstruation central to women's health and fertility.

Reading the history of science and medicine as most saliently about white patriarchal power misses an opportunity to see women's agency in this history. Women sought menstrual regularity, and cared for themselves accordingly, during the many centuries when an irregular cycle might signal an illness, infertility, or a mistimed pregnancy. Today, with modern birth control and home pregnancy tests, menstrual signals are no longer central to self-care, and menstruators can more easily embrace Clancy's enthusiasm for flexible and contingent cycling.

Clancy's historical narratives are also polemical to the point of hyperbole at times. People were not put to death in the Middle Ages for breaking menstrual taboos, as she claims, for example. The secondary source she cites here quotes a 17th-century Puritan cleric who interpreted the Biblical book Leviticus's ban on menstrual sex to mean that God wanted penalty of death, not just exile as indicated in the plain language of the biblical text. One preacher's extreme interpretation of Leviticus is a far cry from actual law and practice.

Overall, the book conveys a consistently positive view of menstrual blood, the menstrual cycle, and the bodies of people who menstruate. Looking for agency, optimism, power, and healing in menstruating bodies, Clancy demonstrates that, indeed, these qualities are there for those whose eyes are open to them.

Refreshingly, though, she resists the kind of essentialist menstrual activism that insists that empowerment means we menstruators must love our cycles. She makes room for those who relish their periods, those who tolerate them, and those who suppress them; for those for whom they may be central to gender identity, or anathema.

A main theme throughout the book is that there is a diversity of menstrual experiences, and that is a good way for bodies (and people) to be. In the end, Clancy simply wants periods to be understood and respected for the remarkable biological process they are. ■

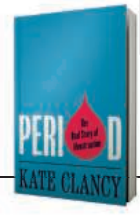
10.1126/science.adh0534



Volunteers sort donated menstrual products in Alexandria, Virginia.

PHOTO: ALLISON SHELLEY/THE WASHINGTON POST/GETTY IMAGES

Period: The Real Story of Menstruation
Kate Clancy
Princeton University Press, 2023. 264 pp.



The reviewer is a historian and the author of *The Modern Period: Menstruation in Twentieth-Century America* (Johns Hopkins Univ. Press, 2009) and *The Myth of the Perfect Pregnancy: A History of Miscarriage in America* (Oxford Univ. Press, 2020). Email: lara@post.harvard.edu



In the past, China's Zhejiang Province coastline supported abundant oyster populations.

Edited by **Jennifer Sills**

Protect oyster reefs in China's coastal zone

Oyster reefs contribute to water filtration, shoreline stabilization, carbon sequestration, and biodiversity enhancement (*1*). In the past 200 years, global oyster reefs have declined by 85% (*2*). Abundant oyster reefs with more than 30 oyster species used to thrive in China's coastal intertidal and subtidal zones (*3*). However, overfishing, water pollution, coastal development, biological disease, and habitat degradation have damaged oyster reefs in China's coastal zone, and most of them have become functionally extinct in the past few decades (*4*). China has taken steps to conserve and restore oyster reefs since 2000 (*5*), but more urgent action is required.

As of 2020, known oyster reefs cover about 1906 ha in China's waters (*5*). China has established several special protected areas to protect oysters, and four marine protected areas (about 43,959 ha in total) include oyster reefs within their borders (*5*). Oyster reef restoration projects have been implemented in estuaries in Tianjin, Jiangsu, Zhejiang, and Hong Kong (*6*). In addition, the Ministry of Agriculture and Rural Affairs listed oyster reefs as important to national marine ranching, a strategy that combines sustainable use of marine fishery resources with the protection of the ecological environment (*7*). The goal of this recognition is to conserve and restore marine biological resources in areas such as the waters off the shore of Caofeidian, Laizhou Bay, and the Dashentang sea area (*8*). However, these efforts have not reversed the degradation of China's oyster reefs.

To identify the full distribution and ecological status of oyster reefs, China should conduct a national survey and monitor any reefs that are found. Known reefs would benefit from stronger management of protected areas and surrounding waters and better legislation to limit illegal fishing and overharvesting. In addition, China should raise public awareness about the importance of oyster reefs, their ecological functions, and their role in coastal habitats alongside mangroves, coral reefs, and seagrass meadows (*5*). Together, government action and the mobilization of social forces could increase the effectiveness of oyster reef protection and restoration.

Wei Xu, Aifeng Tao*, Jinhai Zheng
Key Laboratory of Ministry of Education for Coastal Disaster and Protection, Hohai University, Nanjing 210098, China and College of Harbor, Coastal and Offshore Engineering, Hohai University, Nanjing 210098, China.
*Corresponding author. Email: aftao@hhu.edu.cn

REFERENCES AND NOTES

1. H. Yang, M. Ma, J. R. Thompson, R. J. Flower, *Science* **371**, 790 (2021).
2. M. W. Beck *et al.*, *Bioscience* **61**, 107 (2011).
3. X. Guo, C. Li, H. Wang, Z. Xu, *J. Shellfish Res.* **37**, 755 (2018).
4. D. A. Hemraj *et al.*, *Sci. Adv.* **8**, eabp8747 (2022).
5. The Nature Conservancy, "Research Report on Oyster Reef Habitat Conservation and Restoration in China" (2022); https://www.tnc.org.cn/edm/China_Oyster_Reef_Research_Report.pdf [in Chinese].
6. P. Yang *et al.*, "Present situation and protection restoration suggestions on the natural oyster reefs in China" *Geology in China* (2022). https://kns.cnki.net/kcms/detail/11.1167.P.20220822.1514.022.html [in Chinese].
7. M. Y. Jiao *et al.*, *J. Environ. Manage.* **335**, 117515 (2023).
8. Ministry of Agriculture and Rural Affairs, "National Construction Plan of Marine Ranch Demonstration Areas (2017-2025)" (2017); http://www.gov.cn/gongbao/content/2018/content_5277757.htm [in Chinese].

10.1126/science.adh1965

Restoring health priorities in Brazil

The Brazilian unified health system (SUS) was renowned for reducing health inequities (*1*). However, austerity measures imposed at the end of 2016 (*2*), a dismantling of the Ministry of Health's technical capacities (*3*), and the disastrous COVID-19 pandemic response by the Bolsonaro administration (*4*) have contributed to a rapid deterioration of health indicators. President Lula da Silva, who was inaugurated last January, faces the daunting task of addressing the setbacks and reconstructing the health sector.

Over the past 7 years, SUS has rapidly declined. About 33 million Brazilians face hunger (*5*), and hospitalizations due to child malnutrition caused by hunger have increased (*6*). Vaccination coverage is below the minimum necessary to prevent outbreaks, maternal mortality almost doubled in 2 years, and routine medical tests have declined, delaying treatment of chronic diseases (*3*, *7*). Thousands of pesticides were approved, many of which had been banned in other countries and associated with infant mortality and cancer (*8*, *9*), and delays in insecticide procurement disrupted vector control (*10*). Changes in legislation increased weapon ownership, directly contributing to an increase in homicides in Brazil (*11*). Environmental destruction resulted in a humanitarian crisis among the Yanomami people (*12*).

Given the many health-related problems Brazil faces, the Lula government must prioritize its response. First, the government should fortify the country's institutions, which will make other challenges

PHOTO: CHINA PHOTOS/STRINGER/GETTY IMAGES

0414_Letters_16933872.indd 142 4/10/23 4:17 PM

easier to address. Lula should begin by strengthening the financing and governance of the SUS, improving primary care and health networks (including telehealth), and promoting an equitable geographical allocation of human resources.

Once the health infrastructure is secure, the government should focus on reducing the unacceptably high numbers of preventable morbidity and mortality. This goal should include increasing vaccine coverage, reducing the waiting list for specialized care, and reorganizing healthcare delivery to Indigenous peoples, children and adolescents, and women. Improving health surveillance will also improve the health system's response to emergencies.

To make sure all voices are heard, it is imperative to build bridges between policymakers, academia, and civil society throughout these efforts. Reconstructing the health sector in Brazil is a social and humanitarian need as well as critical for global health security.

Marcia C. Castro[1]*, Nísia Trindade[2], Arthur Chioro[3], José Gomes Temporão[4], Humberto Costa[5], Alexandre Padilha[6], Adriano Massuda[7]

[1]Department of Global Health and Population, Harvard TH Chan School of Public Health, Boston, MA, USA. [2]Ministério da Saúde, Governo Federal do Brasil, Brasília, DF 70058-900, Brazil. [3]Departamento de Medicina Preventiva, Escola Paulista de Medicina, Universidade Federal de São Paulo, São Paulo, SP 04023-062, Brazil. [4]Centro de Estudos Estratégicos, Fundação Oswaldo Cruz, Rio de Janeiro, RJ 21040-361, Brazil. [5]Senado Federal, República Federativa do Brasil, Brasília, DF 70165-900, Brazil. [6]Secretaria de Relações Institucionais da Presidência da República, Governo Federal do Brasil, Brasília, DF 70150-900, Brazil. [7]Departamento de Administração, Centro de Pesquisa e Planejamento em Saúde, Fundação Getúlio Vargas, Escola de Administração do Estado de São Paulo, São Paulo, SP 01313-902, Brazil.
*Corresponding author.
Email: mcastro@hsph.harvard.edu

REFERENCES AND NOTES

1. M. C. Castro *et al.*, *Lancet* **394**, 345 (2019).
2. R. T. S. Lima, *Health Promot. Int.* **34** (Supplement_1), i20 (2019).
3. A. Massuda, R. Dall'Alba, A. Chioro, J. Gomes Temporão, M. C. Castro, *Lancet* **401**, 886 (2023).
4. L. G. Barberia, E. J. Gómez, *Lancet* **396**, 367 (2020).
5. "Rate of hospitalized babies due to malnutrition reaches its worst level in the past 13 years," FIOCRUZ (2022).
6. Rede Brasileira de Pesquisa em Soberania e Segurança Alimentar e Nutricional (Rede PENSSAN), "VIGISAN: Inquérito nacional sobre insegurança alimentar no contexto da pandemia da Covid-19 no Brasil" (Instituto Vox Populi, 2021); http://olheparaafome.com.br/VIGISAN_Inseguranca_alimentar.pdf [in Portuguese].
7. A. Bigoni *et al.*, *Lancet Reg. Health Am.* **10**, 100222 (2022).
8. C. Panis *et al.*, *Environ. Int.* **165**, 107321 (2022).
9. M. Dias, R. Rocha, R. R. Soares, *Rev. Econ. Stud.* 10.1093/restud/rdad011 (2023).
10. F. Borre, J. I. Borri, Y. Z. Cohen, M. Gasparoto, T. B. Gurung, *Epidemiologia* **3**, 97 (2022).
11. Fórum Brasileiro de Segurança Pública, "Armas de fogo e homicídios no Brasil," (Fórum Brasileiro de Segurança Pública, 2022); https://forumseguranca.org.br/publicacoes_posts/armas-de-fogo-e-homicidios-no-brasil [in Portuguese].
12. J. Watts, *Lancet* **401**, 631 (2023).

10.1126/science.adh1254

PHOTO: MURAD AHMED/MAKE SENSE



The authors represented postdocs at Denmark's annual democracy festival.

OUTSIDE THE TOWER

Denmark's big tent includes postdocs

As postdocs, we usually spend our time in the lab processing samples or at the computer analyzing data. Yet on one sunny day in June, we found ourselves in a tent on the Danish island of Bornholm, preparing to debate leaders of Denmark's scientific community, including a university dean and funding agency representatives. At Folkemødet, Denmark's annual democracy festival, people from all walks of life gather to discuss society's challenges. We hoped that by sharing our perspective, we could influence the decision-makers who shared the stage as well as educate the nonscientists in attendance about our work.

Call for submissions
Outside the Tower is an occasional feature highlighting scientists' advocacy experiences. Submit your advocacy story at http://cts.sciencemag.org.

Our time at the Danish Diabetes Academy Postdoc Summit had prepared us to advocate for early-career scientists. The unusual meeting focuses on breaking down the barriers between scientists and sharing the challenges faced by young researchers with the public through editorials (*1*), podcasts (*2*), and animated videos (*3*). Our outreach efforts at the summit led to our invitation to the Folkemøde debate, where we spoke on behalf of 100 science trainees.

The questions began: How do we avoid biases against applicants based on gender, race, and career stage? The funders reminded us of their duty to identify excellence, but we jumped to correct them. "This focus on excellency prevents new and exciting ideas from getting a chance to take root and grow," we explained. We advocated a lottery-based funding system instead, in which funds are allocated by random drawing after an initial screening phase, increasing the likelihood of funding for younger scientists and riskier ideas (*4*).

Although we were nervous as the debate started, our comfort grew as we passionately explained our position. We noticed that the other participants seemed interested in our ideas, and audience members nodded as we spoke. One of the funders replied that they are taking steps to try lottery-based funding systems.

We hope our experience can inspire other researchers to collaborate and build bridges with society. Young scientists' voices deserve to be heard.

Morten Dall[1] and Maria Hauge Pedersen[2]
[1]Novo Nordisk Foundation Center for Basic Metabolic Research, University of Copenhagen, DK-2200 Copenhagen, Denmark. [2]New York Genome Center, New York, NY 10013, USA. Email: dall@sund.ku.dk; mariahaugep@gmail.com

REFERENCES AND NOTES

1. Attendees of the Danish Diabetes Academy Winter School 2019, *Acta Physiol.* **229**, e13455 (2020).
2. Danish Diabetes and Endocrine Academy, "Series 1: Postdocs Talking - from Research to Society" (2021); https://ddeacademy.dk/podcasts.
3. Danish Diabetes and Endocrine Academy, "Animation 2: Work/life balance" (2022); https://twitter.com/DDEA_Denmark/status/1601124590192828416.
4. D. Adam, *Nature* **575**, 574 (2019).

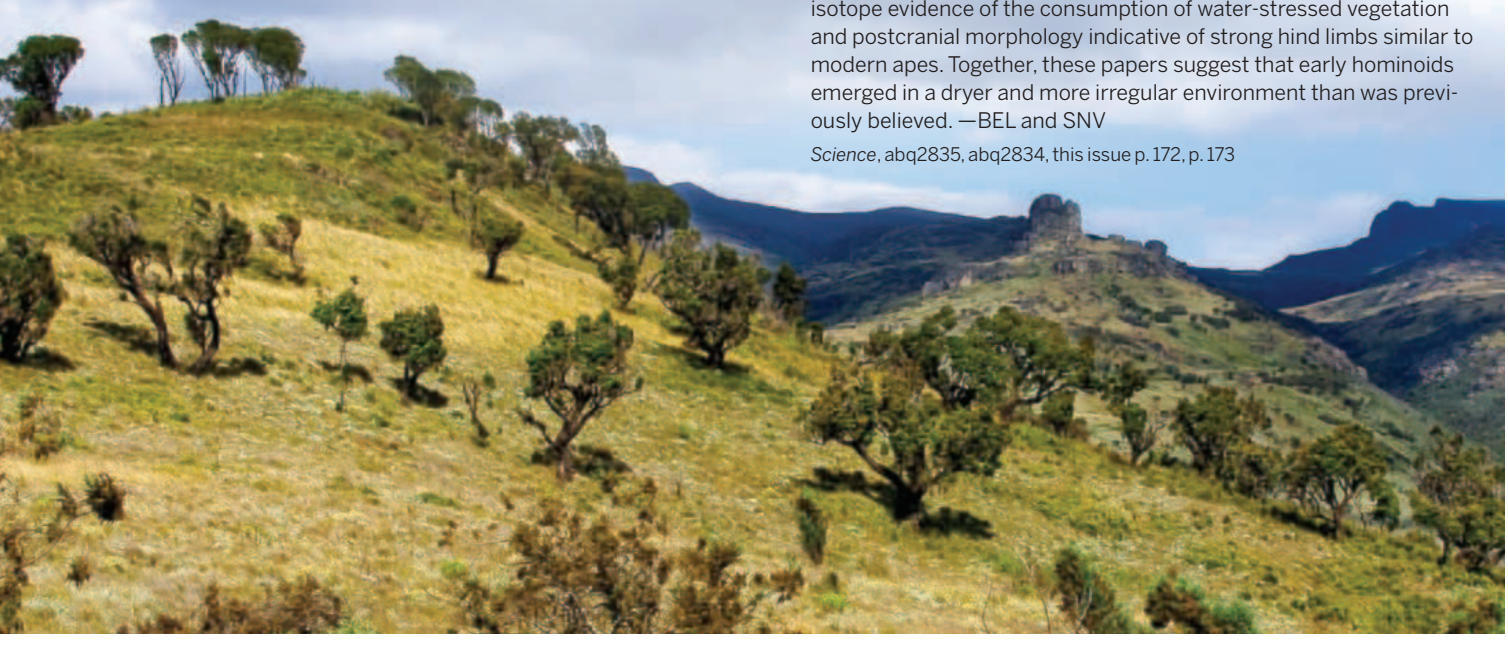
10.1126/science.adh4261

0414_Letters_16933872.indd 143 4/10/23 4:17 PM

RESEARCH

IN *SCIENCE* JOURNALS

Edited by **Michael Funk**



PALEOECOLOGY

A new habitat for hominoid emergence?

The hominoid lineage underwent a major morphological change in the Miocene, acquiring strong hind legs and a more upright posture. The prevailing hypothesis pertaining to these changes has been that they were adaptive for foraging on fruit in the terminal branches of tropical forest trees. A pair of papers now argue that, instead, such changes may have been driven by adaptation to feeding on leaves in seasonally dry and open forests. Peppe *et al*. used new data from fossil mammal study sites and found that the expansion of grassy biomes dominated by grasses with the C4 photosynthetic pathway in eastern Africa likely occurred more than 10 million years earlier than prior estimates. MacLatchy *et al*. looked at fossils of the earliest ape in this region at this time, *Morotopithecus*, and found isotope evidence of the consumption of water-stressed vegetation and postcranial morphology indicative of strong hind limbs similar to modern apes. Together, these papers suggest that early hominoids emerged in a dryer and more irregular environment than was previously believed. —BEL and SNV

Science, abq2835, abq2834, this issue p. 172, p. 173

ARCHAEOLOGY

Leatherworking 39,000 years ago

Humans have worked animal hides since deep antiquity, but the appearance of tailored clothing and the tools required to fashion them continues to be debated. Doyon *et al.* describe a bone artifact with 28 punctures arranged in two separate lines on its flat surface, which was found at the Early Upper Paleolithic site of Canyars, Spain. To understand how the punctures were made, they conducted experiments on similar bones and found that piercing a soft, pliable material by striking a pointed stone tool with a percussor left marks similar to those on the artifact. The authors inferred that the bone served as a punch board to work leather into tailored clothing in the absence of eyed needles in the region. —MSA

Sci. Adv. 10.1126/sciadv.adg0834 (2023).

PAIN

Releasing the brake on opioids

Opioids are effective for acute pain management, but increasing tolerance can lead to greater use and adverse side effects. Gabriel and Streicher found that opioids activate a self-limiting pathway that can be disabled to enhance their efficacy. In the mouse spinal cord, opioids activated the kinases ERK and AMPK. Whereas ERK promoted opioid efficacy against painful stimuli in mice, AMPK reduced it by inhibiting ERK. Blocking AMPK or the chaperone protein HSP90 (which stabilized AMPK) reduced pain sensitivity in opioid-treated mice. —LKF

Sci. Signal. **16**, eade2438 (2023).

CANCER IMMUNOLOGY

Engineered bacteria as melanoma vaccine

The skin microbiome usually lives in harmony with our tissues without inducing inflammation or mounting an infection. However, certain bacterial colonists, including the skin bacterium *Staphylococcus epidermidis*, can induce a highly specific adaptive immune response, the function of which is unclear. Chen *et al*. engineered an *S. epidermidis* strain to express melanoma tumor antigens and then tested their ability to drive antitumor immune responses (see the Perspective by Sajjath *et al*.). Engineered *S. epidermidis* generated tumor-specific T cells that infiltrated and reduced the growth of localized and metastatic melanoma. In combination with immune checkpoint inhibitors, these engineered skin bacteria caused mice to reject established tumors. These findings suggest that immune responses from engineered commensals may have therapeutic potential against other tumor antigens of interest. —PNK and STS

Science, abp9563, this issue p. 203; see also adh3884, p. 132



Heterogeneous grassland habitats, such as this site in modern Kenya, were present by at least 17 million years ago and may have influenced the evolution of hominoid locomotor versatility.

PHOTO: SINGLECELL ANIMATION LLC/SCIENCE SOURCE

EXOPLANETS

Astrometry guides imaging detection

Direct imaging of exoplanets is challenging because of the glare of the much brighter host star. The technique has detected only small numbers of gas giant exoplanets on wide orbits. Astrometry, the precise measurements of stellar positions, could potentially identify stars being pulled by the gravity of a nearby planet. Currie *et al*. used astrometry to identify a candidate host star, which they then studied with direct imaging. They found an object with the right properties to explain the astrometry signal and interpreted it as an exoplanet of about 14 Jupiter masses. The detection shows how astrometry can help to detect and characterize directly imaged exoplanets. —KTS

Science, abo6192, this issue p. 198

DROUGHT

Dry in a flash

Are flash droughts, those that develop unusually rapidly unlike those that develop more slowly and typically have been considered the archetype, becoming the new normal? Yuan *et al*. show that droughts have begun to intensify more rapidly since the 1950s and that flash droughts have become more common over much of the world (see the Perspective by Walker and Van Loon). This trend, which makes drought monitoring and forecasting more difficult, is associated with greater evapotranspiration and precipitation deficits caused by anthropogenic climate change and is projected to expand to all land areas in the future. —HJS

Science, abn6301, this issue p. 187; see also adh3097, p. 130

METAMATERIALS

Asymmetric responses in uniform materials

Typically, if a material is sheared in either direction, the response will be similar. Wang *et al*. designed hydrogels loaded with aligned and tilted graphene oxide sheets (see the Perspective by Sun and Kang). The shear modulus of this composite differs substantially as a function of the shear direction. When shearing causes the graphene oxide sheets to compress, they buckle; however, when they are sheared in the other direction, they are stretched. This asymmetry can convert symmetric vibrations into asymmetric motion. These materials open opportunities to guide, damp, and control mechanical forces using metamaterials of various shapes and sizes. —MSL

Science, adf1206, this issue p. 192; see also adh3098, p. 135

IN OTHER JOURNALS

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



KIDNEY

Kidneys RIG-ged for damage

Acute kidney injury can be caused by a variety of insults, including infections, drugs, and toxins. Studies have shown that supplementation of nicotinamide adenine dinucleotide (NAD^+) or related compounds can promote recovery from kidney injury, but the underlying mechanism for this was unclear. Using metabolomic analysis in injured human kidneys and healthy controls, as well as in mouse models, Doke *et al.* characterized alterations in NAD^+ metabolism and confirmed the benefits of NAD^+ supplementation. They also showed that a deficiency of NAD^+ in the setting of kidney injury impairs mitochondrial metabolism, leading to damage of the organelle and leakage of mitochondrial RNA. This RNA then activates RIG-I, a cytosolic RNA sensor pathway that normally responds to viral infection but in this case exacerbates kidney damage. —YN

Nat. Metab. (2023) 10.1038/s42255-023-00761-7.

CELL BIOLOGY

Endosome exocytosis in endothelial repair

The plasma membrane of vascular endothelial cells often suffers damage caused by blood circulation–induced shear and harmful interactions with blood cells. The source of membrane required for resealing endothelial cells after injury remains unclear. Raj *et al.* used live-cell imaging and correlative light and electron microscopy to study membrane repair in cultured human endothelial cells. Early endosomes are a crucial source of membrane for wound resealing. The authors found that localized early endosome exocytosis is initiated by calcium entering through the plasma membrane wound, which promotes membrane repair

ALSO IN *SCIENCE* JOURNALS

Edited by **Michael Funk**

HUMAN GENOMICS

Dynamic changes in human genomes

Postzygotic mutations accrue throughout a person's lifetime, beginning in utero. Despite their pervasiveness, the impact of these mutations on health is not yet well understood outside of selected mutations implicated in cancer development. Rockweiler *et al*. analyzed postzygotic mutations in a variety of organ sites from 948 human donors in the Genotype-Tissue Expression (GTEx) project to create an atlas of these mutations across tissues, and then validated their findings using several published datasets. The authors also reconstructed the evolution of postzygotic mutations in different organs over the course of a life span and identified mutations with particularly deleterious impacts on human health. —YN

Science, abn7113, this issue p. 171

COAGULATION

The bear necessities for anticoagulation

When humans experience acute immobility caused by illness or injury, they are at increased risk for potentially deadly venous thromboembolism. By contrast, hibernating bears remain immobile for months every year without experiencing any such complications. Similarly, chronically immobilized patients with spinal cord injury do not have an increased risk of thrombosis. By examining blood from humans experiencing chronic immobilization, healthy volunteers on extended bed rest, and free-ranging brown bears during their periods of activity and hibernation, Thienel *et al*. identified specific proteins that are down-regulated during periods of prolonged immobility and protect from thrombosis (see the Perspective by Schattner). —YN

Science, abo5044, this issue p. 178; see also adh3276, p. 133

FIBROSIS

Fibrogenic fiend

Macrophages are key players in the response to tissue injury, and specific macrophage subsets can drive fibrosis progression or resolution. Fabre *et al*. showed that *CD9*$^{+}$*TREM2*$^{+}$ macrophages, which are further defined by the expression of a key set of proteins, can mediate fibrosis. Analysis of human and murine hepatic and pulmonary fibrosis samples revealed accumulation of these macrophages near scar sites and proximal to activated mesenchymal cells. In vitro analysis revealed that type 3 cytokines can promote the differentiation of these macrophages, which are unable to degrade collagen I and contribute to collagen I–associated fibrosis. Blockade of these cytokines in mice was sufficient to limit fibrosis. These findings highlight a critical role for fibrogenic macrophages in this pathological process. —CNF

Sci. Immunol. **8**, eadd8945 (2023).

OPTOELECTRONICS

Surface field for carrier extraction

The efficiency of carrier extraction in photodetection devices can be substantially hindered by surface recombination. In the past, surface electric field has been shown to be effective in reducing surface recombination. However, knowledge regarding the interplay between surface electric field and carrier extraction is still limited. Zhang *et al.* conducted an experiment similar to the Haynes-Shockley scheme to reveal the role of the surface field in carrier extraction. They also introduced work-function-tunable indium tin oxide at the interface to optimize the surface field to achieve near-unity quantum efficiency. These findings could potentially be used to develop highly efficient thin-film optoelectronic devices. —FX

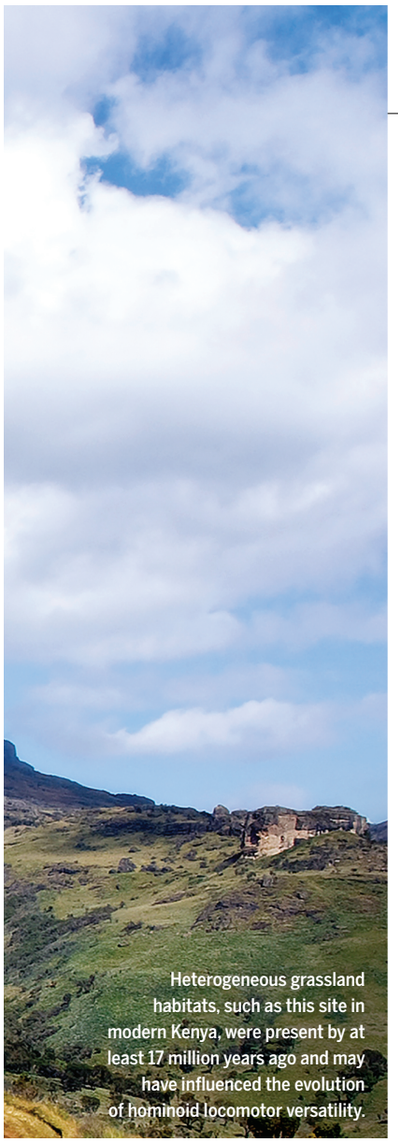
Sci. Adv. 10.1126/sciadv.adg2454 (2023).

VIRAL INFECTIONS

Antiviral advance

The Venezuelan and eastern equine encephalitis viruses (VEEV and EEEV, respectively) are mosquito-borne alphaviruses that can cause severe neurological and systemic disease in humans and equids. There is no approved vaccine or therapeutic available to treat human alphavirus infections. Cao *et al*. now describe the identification and optimization of a quinazolinone, BDGR-49, which had potent antiviral activity against VEEV and EEEV in vitro and showed outstanding prophylactic and therapeutic efficacy in mouse models. BDGR-49 was also associated with promising pharmacokinetic profiles and evidence of brain penetration. These findings highlight BDGR-49 as a promising compound for treating encephalitic alphavirus infections. —CNF

Sci. Transl. Med. **15**, eabl9344 (2023).



Heterogeneous grassland habitats, such as this site in modern Kenya, were present by at least 17 million years ago and may have influenced the evolution of hominoid locomotor versatility.

PHOTO: SINGLECELL ANIMATION LLC/SCIENCE SOURCE

EXOPLANETS

Astrometry guides imaging detection

Direct imaging of exoplanets is challenging because of the glare of the much brighter host star. The technique has detected only small numbers of gas giant exoplanets on wide orbits. Astrometry, the precise measurements of stellar positions, could potentially identify stars being pulled by the gravity of a nearby planet. Currie *et al*. used astrometry to identify a candidate host star, which they then studied with direct imaging. They found an object with the right properties to explain the astrometry signal and interpreted it as an exoplanet of about 14 Jupiter masses. The detection shows how astrometry can help to detect and characterize directly imaged exoplanets. —KTS

Science, abo6192, this issue p. 198

DROUGHT

Dry in a flash

Are flash droughts, those that develop unusually rapidly unlike those that develop more slowly and typically have been considered the archetype, becoming the new normal? Yuan *et al*. show that droughts have begun to intensify more rapidly since the 1950s and that flash droughts have become more common over much of the world (see the Perspective by Walker and Van Loon). This trend, which makes drought monitoring and forecasting more difficult, is associated with greater evapotranspiration and precipitation deficits caused by anthropogenic climate change and is projected to expand to all land areas in the future. —HJS

Science, abn6301, this issue p. 187; see also adh3097, p. 130

METAMATERIALS

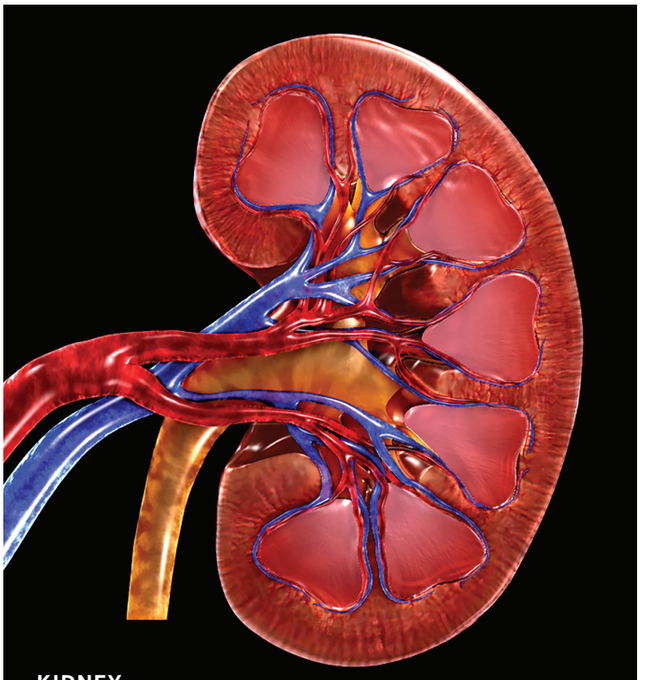
Asymmetric responses in uniform materials

Typically, if a material is sheared in either direction, the response will be similar. Wang *et al*. designed hydrogels loaded with aligned and tilted graphene oxide sheets (see the Perspective by Sun and Kang). The shear modulus of this composite differs substantially as a function of the shear direction. When shearing causes the graphene oxide sheets to compress, they buckle; however, when they are sheared in the other direction, they are stretched. This asymmetry can convert symmetric vibrations into asymmetric motion. These materials open opportunities to guide, damp, and control mechanical forces using metamaterials of various shapes and sizes. —MSL

Science, adf1206, this issue p. 192; see also adh3098, p. 135

IN OTHER JOURNALS

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



KIDNEY

Kidneys RIG-ged for damage

Acute kidney injury can be caused by a variety of insults, including infections, drugs, and toxins. Studies have shown that supplementation of nicotinamide adenine dinucleotide (NAD^+) or related compounds can promote recovery from kidney injury, but the underlying mechanism for this was unclear. Using metabolomic analysis in injured human kidneys and healthy controls, as well as in mouse models, Doke *et al.* characterized alterations in NAD^+ metabolism and confirmed the benefits of NAD^+ supplementation. They also showed that a deficiency of NAD^+ in the setting of kidney injury impairs mitochondrial metabolism, leading to damage of the organelle and leakage of mitochondrial RNA. This RNA then activates RIG-I, a cytosolic RNA sensor pathway that normally responds to viral infection but in this case exacerbates kidney damage. —YN

Nat. Metab. (2023) 10.1038/s42255-023-00761-7.

CELL BIOLOGY

Endosome exocytosis in endothelial repair

The plasma membrane of vascular endothelial cells often suffers damage caused by blood circulation–induced shear and harmful interactions with blood cells. The source of membrane required for resealing endothelial cells after injury remains unclear. Raj *et al.* used live-cell imaging and correlative light and electron microscopy to study membrane repair in cultured human endothelial cells. Early endosomes are a crucial source of membrane for wound resealing. The authors found that localized early endosome exocytosis is initiated by calcium entering through the plasma membrane wound, which promotes membrane repair

0414_ISIO_16947241.indd 169 4/10/23 4:02 PM



BEHAVIOR

Early birds might fail the grade

Adequate sleep and regular class attendance are essential for academic success. The influence of class start time on sleep and academic performance is unclear. Yeo *et al.* used digital data to monitor class attendance, grades, and sleep in more than 20,000 students attending early (8:00 a.m.) or later classes. Early classes were associated with lower attendance, reduced sleep, and poorer grades. Although cause-and-effect relationships cannot be made, the data suggest that reducing the number of early classes could result in increased attendance and better grades. —MMa

Nat. Hum. Behav. (2023)
10.1038/s41562-023-01531-x.

mediated by the early endosomal VAMP2 membrane fusion protein. Thus, early endosomes undergo a spatially constrained and calcium-induced exocytosis as a first line of defence to repair plasma membrane wounds and protect endothelial cells from damage. —SMH

Adv. Sci. (2023)
10.1002/advs.202300244.

REMOTE SENSING

Detecting drought

Predicting how the terrestrial carbon and water cycles are affected by ongoing climate change requires an understanding of vegetation's responses to extreme events such as droughts. Using satellite data, Zhang *et al.* developed a method to detect drought stress in plants from space. Applying machine learning to observations from the Orbiting Carbon Observatory 3, the authors were able to monitor changes in fluorescence efficiency (a proxy of vegetation stress) in response to a 2020 drought across the southwestern United States. Measuring physiological responses from satellite data also has the potential for distinguishing differential responses between vegetation types. —BEL

Glob. Chang. Biol. (2023)
10.1111/gcb.16683.

QUANTUM SCIENCE

Probing 2D spin dynamics

Nitrogen-vacancy (NV) centers, defects in diamond with favorable optical properties, have been used as sensors of magnetic field, pressure, and temperature. The creation of these defects typically results in a three-dimensional (3D) distribution within the sample. Davis *et al.* engineered a thin layer of two types of defects, NV centers and substitutional nitrogen impurities called P1. The more dilute NV centers served to monitor the many-body dynamics of the denser P1 spin ensemble. By combining theoretical predictions with experimental probes of NV center decoherence, the researchers were able to demonstrate the 2D nature of their system and measure the correlation time of the P1 ensemble. —JS

Nat. Phys. (2023)
10.1038/s41567-023-01944-5.

ASTEROIDS

Forming Dimorphos from spin and a ring

In September 2022, the Double Asteroid Redirection Test (DART) spacecraft deliberately collided with the asteroid moon Dimorphos, changing its orbit around the larger asteroid Didymos. Madeira *et al.* calculated how the system could have formed, and believe that asymmetrical heating by sunlight caused Didymos to rotate so fast that it ejected material near its equator. The ejecta then spread into a ring, most of which eventually fell back onto Didymos, but some coalesced to form Dimorphos. The results are consistent with the masses, spin rates, and mutual orbit of the two bodies and with their morphologies in the images taken by DART. The system will be visited by another spacecraft in 2026. —KTS

Icarus (2023)
10.1016/j.icarus.2023.115428.

BIOSYNTHESIS

Peptide hormones on demand

Peptides and proteins are an important and growing fraction of drugs, but they typically require complex, expensive production processes. DeWinter *et al.* created a process for the production of small-peptide hormones through cell-free protein synthesis using a fusion protein tag to boost expression and stability. On-bead purification and proteolysis enable the generation of peptides corresponding to native hormones for some of the targets attempted. In theory, such in vitro production of complex molecules may be useful in small-scale biologics production for point-of-care applications. —MAF

ACS Synth. Biol. (2023)
10.1021/acssynbio.2c00680.

RESEARCH ARTICLE SUMMARY

HUMAN GENOMICS

The origins and functional effects of postzygotic mutations throughout the human life span

Nicole B. Rockweiler*, Avinash Ramu, Liina Nagirnaja, Wing H. Wong, Michiel J. Noordam, Casey W. Drubin, Ni Huang, Brian Miller, Ellen Z. Todres, Katinka A. Vigh-Conrad, Antonino Zito, Kerrin S. Small, Kristin G. Ardlie, Barak A. Cohen, Donald F. Conrad*

INTRODUCTION: Mutation lays the foundation for genetics, evolution, and our very existence and demise. Historically, genetics has focused on inherited variants and has only recently begun to examine the genetic changes that occur after fertilization, known as postzygotic mutations (PZMs). This bias is partially because of technological limitations and the simplifying assumption that all cells in a multicellular organism share the same genome.

RATIONALE: Most PZM research has been single-tissue studies. An exciting next generation of PZM studies now examine PZMs across multiple tissues within an individual. However, the relatively small number of individuals and tissue types examined thus far have limited the ability to ascribe sources of mutation variation among individuals or to provide detailed descriptions of embryonic mutations that occur after the first few cell divisions.

RESULTS: To expand the field's knowledge of the origins and functional consequences of PZMs, we sought to answer four key questions: (i) Are PZMs detectable? (ii) Where do PZMs occur? (iii) When do PZMs occur? (iv) When do PZMs contribute to phenotypic variation?

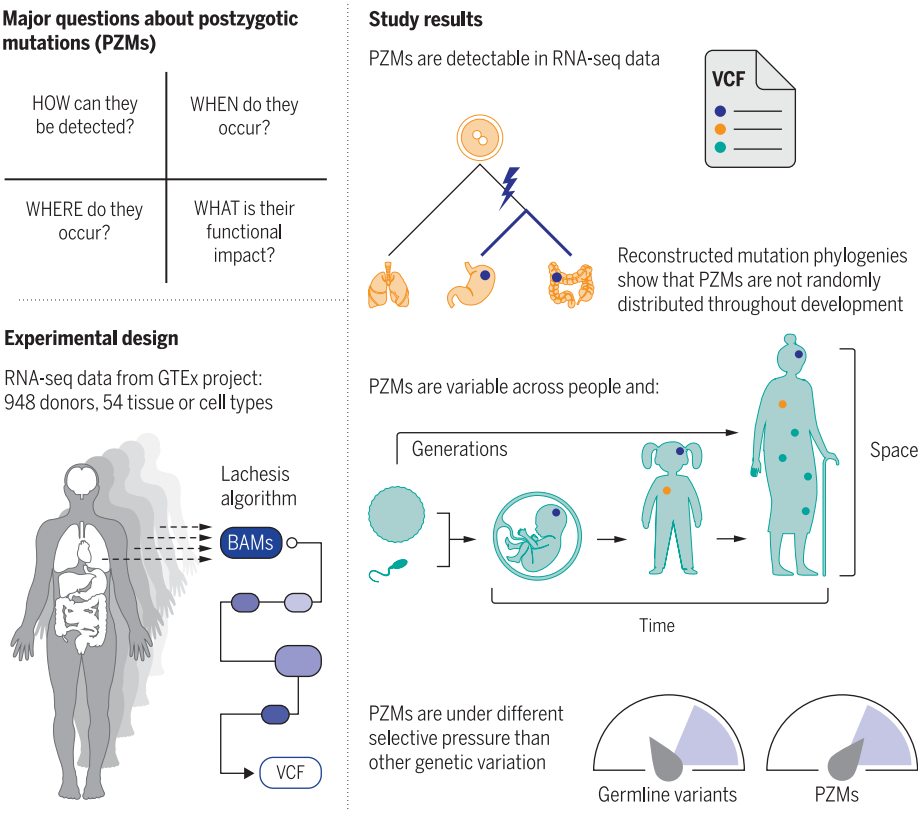
We developed a suite of methods called Lachesis to detect single-nucleotide DNA PZMs from bulk RNA sequencing (RNA-seq) data. We applied these methods to the final major release of the NIH Genotype-Tissue Expres (GTEx) project—a catalog of 17,382 sam derived from 948 donors across 54 diverse tissues and cell types—to generate one of the largest and most diverse catalogs of PZMs in normal individuals.

PZMs were pervasive and highly variable among donors and tissues. Nearly half of the variation in mutation burden among tissue samples was explained by technical and biological effects, such as age and tissue type. We also found that 9% of this variation was attributed to donor-specific effects. This means that there may be systematic differences among individuals in the number of mutations that they carry due to genetic and/or environmental effects, even after controlling for age. The types of mutations, i.e., mutation spectra, were also variable across tissues, which suggests that mutational mechanisms may be different across tissues.

To estimate when PZMs occur during development, we first identified putative prenatal PZMs in the catalog and then mapped them to a developmental tree. Mutation burden and spectra varied throughout prenatal development, with early embryogenesis being the most mutagenic.

Finally, to investigate the functional consequences of PZMs, we compared the predicted deleteriousness and selection strength on PZMs across space and time. We found that the predicted functional impact of PZMs varies during prenatal development and across tissues, and we identified a class of low-frequency prenatal mutations apparently more deleterious than all other forms of human genetic variation considered. The deleteriousness of germline mutations decreased through the lifecycle: testicular germ cells carried more deleterious mutations than ejaculated sperm and sperm resulting in viable offspring.

CONCLUSION: In this work, we present methods for detecting PZMs and a comprehensive and diverse atlas of PZMs in normal development and aging. Akin to how expansive surveys of normal germline variation are immensely beneficial for human and medical genetics, this catalog contributes to our understanding of normal postzygotic variation so that abnormal variation can be identified and interpreted. Uncovering the effects of these PZMs on human health and disease is an exciting and valuable endeavor. ■



Postzygotic DNA mutations accumulate in the human body, are copied into RNA, and are detected by RNA-seq. A developmental map of these mutations resolves when and where in the body they occur and finds that the selection pressure on PZMs is different from inherited germline variants. BAMs, binary alignment maps; VCF, variant call format.

The list of author affiliations is available in the full article online.
*Corresponding author. Email: nrockweiler@wustl.edu (N.B.R.); conradon@ohsu.edu (D.F.C.)
Cite this article as N. B. Rockweiler *et al.*, *Science* **380**, eabn7113 (2023). DOI: 10.1126/science.abn7113

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abn7113

RESEARCH ARTICLE

HUMAN GENOMICS

The origins and functional effects of postzygotic mutations throughout the human life span

Nicole B. Rockweiler[1]*†, Avinash Ramu[1], Liina Nagirnaja[2], Wing H. Wong[3]‡, Michiel J. Noordam[1], Casey W. Drubin[1], Ni Huang[1]§, Brian Miller[2], Ellen Z. Todres[4], Katinka A. Vigh-Conrad[2], Antonino Zito[5]¶, Kerrin S. Small[5], Kristin G. Ardlie[4], Barak A. Cohen[1], Donald F. Conrad[1,2,6]*

Postzygotic mutations (PZMs) begin to accrue in the human genome immediately after fertilization, but how and when PZMs affect development and lifetime health remain unclear. To study the origins and functional consequences of PZMs, we generated a multitissue atlas of PZMs spanning 54 tissue and cell types from 948 donors. Nearly half the variation in mutation burden among tissue samples can be explained by measured technical and biological effects, and 9% can be attributed to donor-specific effects. Through phylogenetic reconstruction of PZMs, we found that their type and predicted functional impact vary during prenatal development, across tissues, and through the germ cell life cycle. Thus, methods for interpreting effects across the body and the life span are needed to fully understand the consequences of genetic variants.

The effects of age ravage all tissues of the body, but the pace and consequences of age-related decay vary among tissues and people. The accumulation of DNA damage is thought to be a primary agent of age-related disease (*1*), and surveys of postzygotic mutations (PZMs) in normal tissues [for example, blood (*2*–*4*), brain (*5*), and skin (*6*, *7*)] and across the body (*8*–*10*) have found PZMs to be pervasive across the genome and individuals. However, beyond cancer, there are few conditions where PZMs are known to have a causal role. Because of the high cost and technological challenges of PZM studies, a general understanding of how and when mutation affects the function of specific cell and tissue types is essential for defining research priorities. One way to prioritize hypotheses about mutation and disease is to systematically characterize the consequences of PZMs on cellular fitness across a broad range of tissues. Surveys of normal tissues have found that PZMs appear to accrue neutrally (*10*, *11*), but positive and negative selection do occur in specific genes and cellular contexts, which suggests that PZMs affect cellular function.

Another fundamental question is how the timing of mutation modulates risk for diseases. As clearly demonstrated in oncology, it is possible to detect disease-causing PZMs and augment clinical care years before clinical disease is recognized (*12*, *13*). If PZMs that confer risk for disease accrue across the life span, the PZM profile in a healthy individual could contain actionable prognostic information. Although the relative contributions of prenatal and postnatal PZMs to disease risk are unclear, because of the massive cell proliferation during development, prenatal PZMs have the potential to affect many cells and, thus, to play an important role in disease.

Most PZM research has been single-tissue studies largely focused on tissues that are easily accessible, such as blood, liver, skin, and colon. An exciting next generation of PZM studies now examines PZMs across multiple tissues within an individual (*8*–*10*, *14*). However, the relatively small numbers of individuals and tissue types used in such studies have limited the ability to ascribe sources of mutation variation among individuals or to provide detailed descriptions of embryonic mutations that occur after the first few cell divisions. To expand our knowledge of PZMs in normal tissues, we developed a suite of methods called Lachesis to identify single-nucleotide PZMs from bulk RNA sequencing (RNA-seq) data and predict when the mutations occurred during development and aging (figs. S1 and S2). We ran the algorithm on the final major release of the Genotype Tissue Expression project (GTEx)—a collection of RNA-seq data from 17,382 samples derived from 948 donors across 54 diverse tissues and cell types—to generate one of the most comprehensive databases of PZMs in normal tissues (*15*, *16*) (tables S1 to S8). We used this atlas and the rich metadata on GTEx donors to characterize sources of variation in PZM burden among individuals and unveil the spatial, temporal, and functional variation of PZMs in normal development and aging.

DNA PZMs are accurately detected in bulk tissue RNA-seq

We evaluated the accuracy of the algorithm using several in silico and experimental methods (figs. S3 and S4 and tables S3 to S6). For experimental validation, we obtained four independent DNA- and RNA-based validation datasets generated from the same tissue samples as the primary data covering 296 specific genomic sites across 95 samples. The original PZM variant allele frequency (VAF) estimates from RNA-seq were well correlated with the VAFs from DNA-seq (Spearman's $\rho = 0.82$, P value $= 2.3 \times 10^{-25}$), which suggests that RNA-seq–based VAFs are representative of true mutant cell frequencies. PZMs with VAFs as low as 0.16% and PZMs found in multiple tissues and multiple donors were validated. The average false discovery rate (FDR) across all validation datasets was 27% and was lower than with published methods for detecting PZMs from RNA-seq (34 to 82%) (*8*, *9*, *17*) (fig. S1E and table S3). Because mutations may fail to validate as a result of spatial variation in mosaicism, the FDRs may be overestimated. A small subset of samples (~5%) had an extraordinarily high number of detected PZMs; validation data from these samples produced an average FDR estimate of 98% (table S3). We conclude that these outliers were likely technical artifacts and not hypermutated tissues.

We used power simulations to estimate the algorithm's sensitivity. As expected, simulated PZMs with larger VAFs and higher coverage had higher PZM detection power. At the middle quintile of coverage [(673, 1395) fold coverage), PZMs with VAFs as low as 0.66% could be detected in at least 90% of simulations, which suggests that the method has reasonable sensitivity (fig. S1F).

PZMs are pervasive and highly variable among donors and tissues

After sample and PZM quality control, 56,585 PZMs were detected with VAFs as low as 0.04% and a median VAF of 0.5% (table S7). These mutations are not a random sample of PZMs from the genome but a critically important subset located in the so-called allowable transcriptome—a filtered set of transcribed positions based on GENCODE 26 gene models (*16*) (table S1). In total, 100% of the donors and 77% of the tissue samples had detectable

[1]Department of Genetics, Washington University School of Medicine, St. Louis, MO 63110, USA. [2]Division of Genetics, Oregon National Primate Research Center, Oregon Health & Science University, Beaverton, OR 97006, USA. [3]Department of Pediatrics, Division of Hematology and Oncology, Washington University School of Medicine, St. Louis, MO 63110, USA. [4]Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA. [5]Department of Twin Research and Genetic Epidemiology, King's College London, London SE1 7EH, UK. [6]Center for Embryonic Cell and Gene Therapy, Oregon Health & Science University, Portland, OR 97239, USA.
***Corresponding author. Email: nrockweiler@wustl.edu (N.B.R.); conradon@ohsu.edu (D.F.C.)**
†Present address: Program in Medical and Population Genetics and Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA 02142, USA.
‡Present address: Departments of Genetics and Medicine, Stanford University, CA 94305, USA.
§Present address: T-Therapeutics Ltd., Cambridge CB21 6AD, UK.
¶Present address: Department of Molecular Biology, Massachusetts General Hospital, Boston, MA 02114, USA.

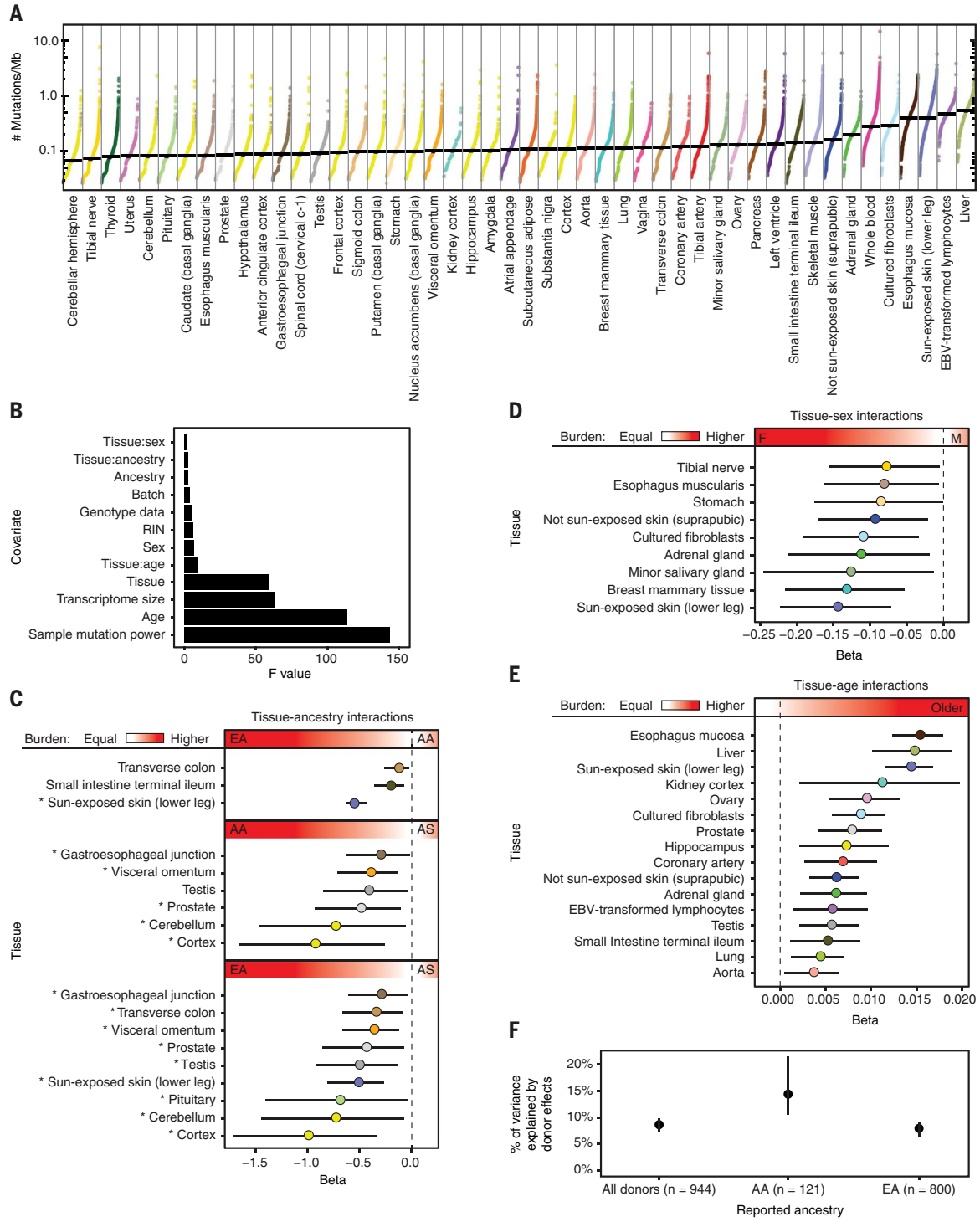


Fig. 1. PZM burden is correlated with biological and technical variables. (**A**) Each data point represents a single tissue sample and is colored by tissue. Median normalized PZM burden in a tissue denoted by the horizontal black line. Tissues are sorted by increasing median normalized PZM burden. A pseudocount of one mutation was added to each sample before normalization and log transformation for visualization. (**B**) We fit a regression model for single-tissue PZM burden using 12 covariates and 48 tissues. Shown here are the type II ANOVA F statistics for each covariate in the model. Larger F statistics correspond to greater explanatory power of the covariate. (**C** and **D**) Regression coefficients of tissue-ancestry interactions (C) and tissue-sex interactions (D) indicate strong effects of ancestry and sex on PZM burden. AA, African American; AS, Asian American; EA, European American. The asterisks in (C) denote differences in mutation burden among ancestry groups that are consistent with cancer incidence trends (*18*). (**E**) Significant positive tissue-age interaction effects were detected for 16/48 (33%) tissues. In (C) to (E), the red gradient and text labels within indicate the meaning of the regression coefficients' sign and magnitude. (**F**) Variance component estimates of donor-specific random effects on PZM burden indicate that 8 to 15% of variation among tissues can be ascribed to donor effects, which could be genetic and environmental. Dashed vertical lines at beta = 0 in interaction plots denote no association between mutation burden and interaction. [(C), (D), (E), and (F)] Error bars represent 95% CIs. [(A), (C), (D), and (E)] Tissues are colored using the GTEx coloring convention (see table S8 for a complete legend).

mosaicism (table S2). We defined the mutation burden of a sample as the number of PZMs detected in a sample and the normalized mutation burden of a sample as the mutation burden normalized by the size of the sample's transcriptome (the number of megabases with at least 20× total coverage). The median normalized mutation burden in a tissue ranged from 0.03 PZMs per expressed megabase in the cerebellar hemisphere to 0.47 PZMs per expressed megabase in the liver (Fig. 1A and table S8). The observed normalized mutation burden was more variable within a tissue than between tissues [mean median absolute deviation (MAD) within a tissue = 0.07 PZMs per expressed megabase; MAD across tissues = 0.02 PZMs per expressed megabase]. This observation suggests that processes generating detectable PZMs may be more variable across donors than across tissue types.

To build further support for the validity of our per-tissue estimates of mutation burden, we compared our data with a recent multitissue survey of PZMs based on DNA-seq of three donors (*10*, *14*). Encouragingly, when comparing 12 tissues assessed by both studies, we found reasonably high correlation in estimated PZM burden (*16*) (fig. S5). The Pearson correlation for the average burden was 0.8 ($P = 0.0018$, Pearson's correlation test).

PZM burden is correlated with biological and technical variables

To partition and quantify potential sources of single-tissue PZM burden, we fit linear models relating technical and biological metadata to single-tissue PZM burdens and selected the best fitting model identified from detailed model comparisons (*16*). The final model contained 12 covariates and explained 48% of the variation in mutation burden. All covariates yielded F test $P < 0.05$ in a type II analysis of variance (ANOVA) and included both biological (age; tissue; and interactions of tissue with age, sex, and self-reported ancestry) and technical (for example, mutation detection power and RNA extraction batch) sources of variation (Fig. 1B and table S9). In total, 20.8% (10/48) of tissues showed significant (Wald test Q value < 0.05) associations with self-reported ancestry, including, as expected, a much lower burden of mutation in sun-exposed skin in African Americans and Asian Americans compared with European Americans (*8*). The incidence rates of cancer types affecting these tissues have ancestry associations that are consistent with (in the same direction as) the mutation burden associations in 83% (15/18) of comparisons (*18*), which suggests that variation in PZM burden in normal tissues may contribute to differences in cancer risk among ancestries (Fig. 1C and table S10). Unexpectedly, males had lower burden in all three skin-related sample types compared with females (Fig. 1D). This result was essentially unchanged when removing genes inferred to have sex-biased expression (fig. S6). Age was positively associated with 33% (16/48) of tissues and was the strongest for esophagus mucosa, liver, and sun-exposed skin (Fig. 1E). We note that power may have been too low to detect some associations; for example, there were few young GTEx brain donors.

Extending this model to include a random donor effect, we estimated that 8.8% of variation in PZM burden can be attributed to systematic properties of donors that extend across some or all tissues of a donor, even after controlling for metadata such as age and sex. This donor variance component estimate was larger in African Americans [14.1%; 95% confidence interval (CI): 10.5 to 21.5%] than in European Americans (8%; 95% CI: 6.5 to 9.1%) (Fig. 1F). These unexplained donor-specific effects could have both genetic and environmental bases. Notably, a recent study estimated that 5.2% of variance in germline mutation rate could be attributed to family-specific effects (*19*). In total, our results indicate that variation in PZM rate among individuals is less constrained than variation in germline mutation rate and that there is considerable scope for heritable variation in observable PZM burden. The inability of the models to explain all variation implies that there are additional factors associated with detectable mutation burden and/or that stochasticity plays a major role in mosaicism (*20*, *21*). A reanalysis of the data that incorporated information on the apparent clonality of mutations produced models with similar biological conclusions and less explanatory power (*16*) (fig. S7 and table S11).

Mutation spectra are variable across tissues and reflect known biological processes

Diverse processes mutate the human genome with characteristic mutational signatures (*22*). Thus, the observed mutation spectra can provide insight on the types and relative activities of the unobserved mutation processes that occurred. We estimated the contribution of canonical mutation signatures for each tissue. Because of the relatively low number of detected mutations, mutation spectra were reliably deconstructed for only four tissues or cell types (*16*) (fig. S8). Consistent with expectations and previous studies (*3*, *6*, *7*), the mutations were resolved into mutational signatures associated with age in all tissues and ultraviolet light exposure in skin-related tissues (fig. S9).

For a higher powered, but coarser-grained analysis of mutation spectra, we assessed the frequency of the six base substitutions across all tissues (fig. S8). Mutation spectra were highly variable across tissues, which suggests that mutational mechanisms and their relative activity may vary across the human body. C > T was the most common mutation type across tissues, whereas C > G and T > A were the least common. Hierarchical clustering of the mutation types revealed two significant large clusters ($P < 1 \times 10^{-3}$, bootstrap resampling). We denoted these as cluster A (marked by depleted T > G) and cluster B (marked by elevated T > G). Cluster membership was associated with mutation burden, which suggests that the underlying mutation mechanisms may be coupled to the frequency of mutagenic events ($P = 3.8 \times 10^{-2}$, Mann-Whitney U test). Additionally, cluster B was enriched with neural ectoderm tissues compared with cluster A ($P = 7.7 \times 10^{-6}$, Fisher's exact test). These clusters could not be attributed to differences in sample processing (*16*) (fig. S10). We speculated that these clusters may reflect differences in the relative contributions of mutations acquired during prenatal development and mutations that accrue during age-related tissue renewal. To further study the properties of prenatal and postnatal PZMs, we developed methods to define the developmental origin of each PZM.

The developmental origins of prenatal PZMs

Multitissue PZMs exhibit prenatal properties

We defined a multitissue PZM as a PZM that was detected in at least two tissues from the same donor. Because the PZM burden was relatively low across tissues (Fig. 1) and PZMs are predominantly under neutral selection (*11*), we hypothesized that a multitissue PZM was the result of a single PZM that occurred in a common ancestor of the mutated tissues. Because the common ancestors of any set of GTEx tissues (excluding cell lines) occurred before the end of organogenesis, multitissue PZMs may have occurred prenatally. Consistent with this hypothesis, we found several lines of evidence suggesting that the multitissue PZMs occurred prenatally (*16*) (figs. S11 and S12). We found a significant positive correlation between VAF and the fraction of the donor's tissues that had the multitissue mutation detected (Spearman's $\rho = 0.34$; $P = 9.7 \times 10^{-56}$, Spearman's rank correlation test; fig. S11A). Controlling for technical and biological confounders, age was not significantly associated with multitissue mutation burden for most tissues but was significantly associated with single-tissue mutation burden for a large number of tissues (fig. S11, B and C). Additionally, the multitissue age regression coefficients were significantly smaller than the single-tissue age regression coefficients ($P = 0.016$, Wilcoxon signed-rank test) (fig. S11D). We denoted these multitissue mutations as prenatal PZMs, and all other mutations were called postnatal PZMs. We note that there may be an error rate associated with this classification because some mutations labeled as postnatal may have been prenatal mutations

lost in some tissues or may have been undetected in some donors because of limited samples.

PZM burden and spectra vary throughout prenatal development with most mutations occurring during early embryogenesis

To determine when and where PZMs occur in prenatal development, we developed a method called LachesisMap to map the origin of 1864 prenatal mutation events (*16*) (Fig. 2A, figs. S13 to S16, and table S12). Briefly, the method takes as input a directed rooted tree representing the developmental relationships among the tissues and a list of multitissue PZMs and maps the PZMs to the tree while accounting for differential mutation detection power across the genome, human body, and developmental tree. The algorithm outputs a list of edge weights that represent the estimated fraction of PZMs that occurred in that spatiotemporal window of development.

The mutation burdens across developmental time and space were highly variable, with edge weights ranging from 0.04 to 23%, and appeared compatible with an exponential distribution (P = 0.56, Kruskal-Wallis test). The ensemble of observed edge weights was significantly different from random ($P = 2.2 \times 10^{-308}$, multinomial goodness-of-fit test), and the majority of individual edge weights (56%, 14/25) were significantly different from random after Benjamini-Hochberg correction (permutation tests) (fig. S17). The top two edge weights, representing 41% of prenatal mutation events, were the zygote-to-gastrula transition and the ectoderm-to–neural ectoderm transition, which suggests that most detectable prenatal mutations occur during early

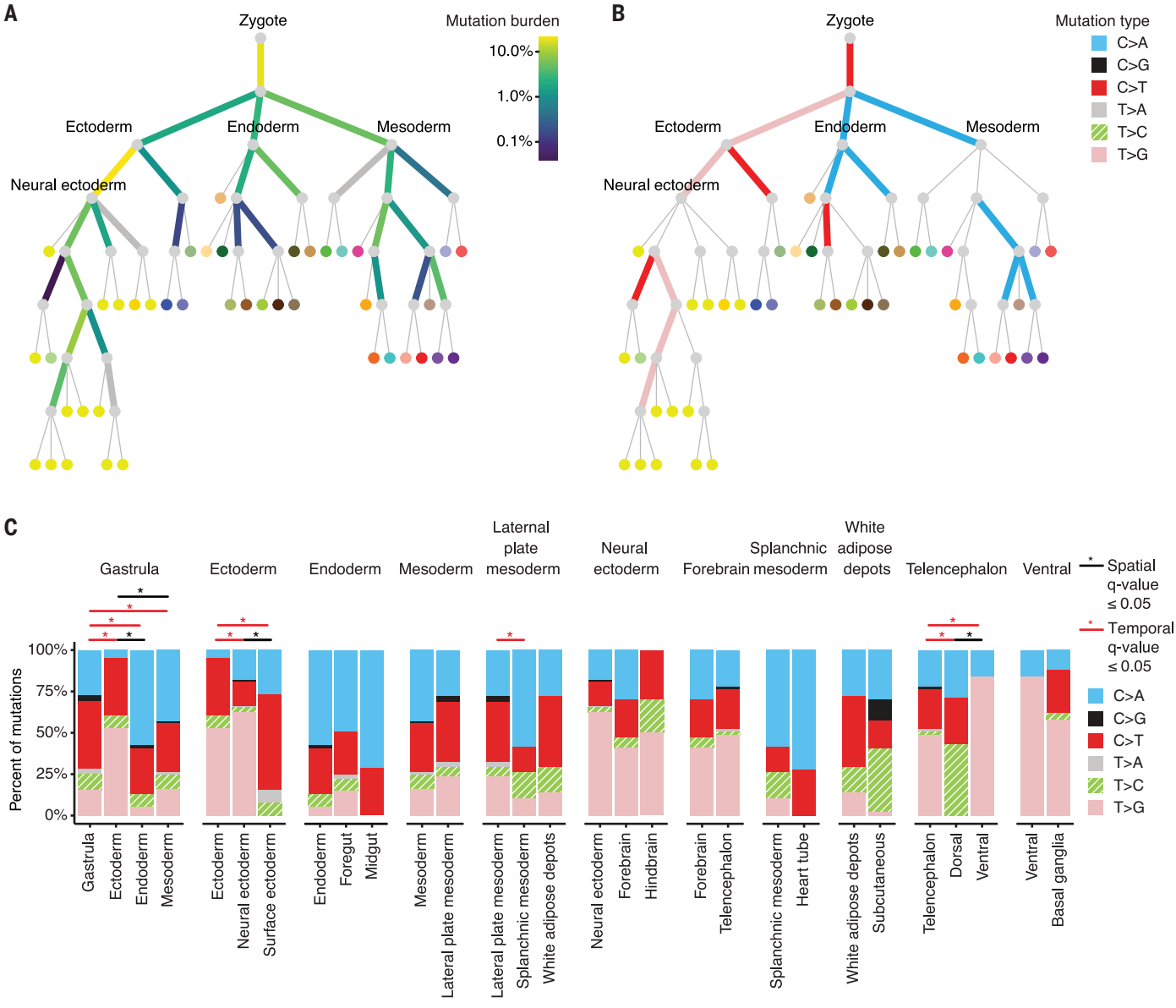


Fig. 2. Mutation burden and spectra of prenatal PZMs across time and space. (**A**) Prenatal PZM mutation burden. Edge color represents the percent of prenatal PZMs mapped to that period in development. Thick gray edges are edges with limited mutation detection power. (**B**) Edge color represents the predominant mutation type of mutations mapped to that edge, as established by binomial testing. Thin gray edges are edges with no predominant mutation type. See fig. S13A for the full set of vertex labels. Adult tissues (leaves of tree) are colored using the GTEx coloring convention (see table S8 for a complete legend). (**C**) Local variation in mutation spectra across developmental space and time. Each facet represents the mutation spectra observed in a parent edge (leftmost bar plot) and its children's edges. Statistically significant differences in mutation spectra are annotated with asterisks.

embryogenesis (*14*, *23*). Notably, the edge mutation burdens were not explained by differential edge mapping power across the developmental tissue tree (*16*) (fig. S17). It is also important to note that these are estimates for mutations that are detectable in adulthood—the data do not allow for extrapolating to all developmentally acquired mutations because some fraction is likely lost through cell death, revertant mosaicism, etc.

We next investigated whether the mutational processes, as proxied by their mutation spectra, varied over development, using binomial tests to establish the predominant mutation type on each edge. There was a strong dichotomy between ectoderm lineages, which tended to have T > G mutations, and endoderm and mesoderm lineages, which tended to have C > A mutations (Fig. 2B). These observations could not be attributed to differences in sample processing (*16*) (fig. S10).

In addition to global changes in mutation across the tree, we also examined local changes by comparing mutation spectra between sibling edges (local spatial differences) and parent-child edges (local temporal differences) (Fig. 2C). Significant spatial and temporal variation was detected during gastrulation and in ectodermal lineages ($Q < 0.05$, multinomial goodness-of-fit test). Differences in mutation spectra across developmental space [$n = 4/8$ (50%) sibling edge comparisons] occurred at similar rates as differences along developmental time [$n = 8/18$ (44%) parent-child comparisons] ($P = 1.00$, Fisher's exact test).

Together, these results suggest that the mutational mechanisms that operate during development may vary across space and time. Although published data are limited, others have also detected variations in mutation spectra in fetal stem cells in humans (*24*) and during early embryogenesis and gametogenesis in mice (*25*). We repeated these analyses using a simplified germ layer tree and observed similar results as from the full developmental tissue tree (fig. S18), which suggests that the development tree definition does not substantially affect the results.

The functional consequences of PZMs across the human life span

The GTEx PZM atlas provides a great opportunity to compare the quality and fitness consequences of mutations that arise at different stages of the human life cycle. First, we annotated the PZM atlas with Combined Annotation Dependent Depletion (CADD), a widely used machine learning classifier of genetic variation (*26*). The CADD score of a genetic variant is a quantitative prediction of deleteriousness, measured on an evolutionary time scale. A mutation was defined as deleterious if the PHRED-scaled CADD score was ≥20. We performed a series of systematic comparisons of PZM CADD scores to identify differences across mutation VAF, developmental time, developmental location, and tissue type.

When comparing prenatal and postnatal PZMs, we found a major effect of VAF on the distribution of CADD scores (Fig. 3A and fig. S19). For prenatal PZMs, low-VAF PZMs were much more deleterious than high-VAF PZMs (odds ratio = 1.9, $P = 2.6 \times 10^{-7}$, Fisher's exact test), whereas no such difference was observed for postnatal PZMs ($P = 0.24$, Wald test). Furthermore, we found that for low-VAF PZMs, deleteriousness decreased over time (odds ratio = 0.58, $P = 1.4 \times 10^{-9}$) but remained constant for high-VAF PZMs ($P = 0.15$). These results suggest that mutations that appear deleterious on an evolutionary time scale may be benign or even beneficial to a growing fetus so long as the mutation remains in a small fraction of cells.

Next, we explored whether deleteriousness varied across the adult human body by comparing postnatal PZMs in each adult tissue. PZM deleteriousness was similar across tissues; however, there were a few exceptions (Fig. 3B). PZMs in 6/48 (13%) tissues were significantly less deleterious than the average tissue, and PZMs in 3/48 (6%) tissues were more deleterious ($Q < 0.05$, Wald test). When analyzed together, the PZMs from all brain regions were also more deleterious than average ($P = 0.02$, Fisher's exact test).

Finally, to provide context for our results, we compared the deleteriousness of GTEx PZMs with other classes of single-nucleotide genetic variation: (i) random mutations (simulated from two different models of neutral evolution), (ii) standing germline variation [from gnomAD, a comprehensive database of germline genetic variation (*27*)], (iii) inherited de novo mutations from cases of disease and controls [from denovo-db, a curated database of de novo mutations (*28*)], and (iv) somatic mutations observed in cancer (from TCGA, a comprehensive database of cancer somatic mutations) (*29*).

The low-VAF prenatal PZMs were the most deleterious class of genetic variation investigated (Fig. 3C and fig. S19). Using the simulated random mutations as a reference, we found postnatal PZMs, de novo mutations in cases, and somatic cancer mutations to be significantly enriched for deleterious mutations ($Q < 0.05$, Fisher's exact test). De novo mutations in controls and high-VAF prenatal PZMs were not statistically different from simulated random mutations. Inherited germline variants were depleted of deleterious mutations, with the extent of depletion increasing with population frequency. These observations were recapitulated in three validation datasets that used a variety of nucleic acid sources and variant-calling methods (*16*) (figs. S20 to S23 and table S13).

The selective constraint on the transcribed exome varies throughout the human life span

The deleteriousness results suggest that selection pressure may be different across classes of genetic variation. We investigated this hypothesis by estimating the selection pressure on PZMs and other classes of genetic variation using dN/dS, a normalized rate of nonsynonymous to synonymous mutations (*30*). dN/dS values >1 were interpreted as evidence for positive selection, whereas negative selection can lead to dN/dS values <1. Using dNdScv, a method for the study of somatic evolution (*11*), we assessed dN/dS across VAF, developmental time, developmental location, and tissue type and contextualized the results by comparing selection pressures on PZMs with other classes of genetic variation as before (*16*) (Fig. 3, D and E, and figs. S24 to S26).

For most tissues of the body, single-tissue dN/dS was not significantly different from 1, consistent with previous work (*11*). However, for postnatal missense mutations, dN/dS was higher for high-VAF PZMs compared with low-VAF PZMs for all tissues en masse and for three tissues or cell types individually [whole blood, Epstein-Barr virus (EBV)–transformed lymphocytes, and adrenal gland] (fig. S24). Additionally, dN/dS estimates for high-VAF postnatal PZMs were higher in cancer driver genes than noncancer driver genes for all tissues en masse, sun-exposed skin, and esophagus mucosa—tissues where the action of adaptive evolution has already been documented (*7*, *31*) (Fig. 3D). These observations are consistent with the expectation that positive selection on a mutation may result in clonal growth, and we detected mutations associated with clonal hematopoiesis of indeterminant potential (CHIP) in the blood of individuals without apparent hematological malignancies (*16*) (fig. S27). Six separate CHIP mutations were detected in seven samples (table S14). Two of the mutations (IDH2 R140Q and MYD88 L273P) are in the 99.99th percentile of recurrent mutations in hematopoietic and lymphoid cancers and have been shown to have gain-of-function properties (*32*, *33*). In total, 0.1% (1/746) of whole-blood donors and 3.5% (6/174) of EBV-transformed lymphocyte donors had a CHIP mutation. Notably, none of the CHIP-positive donors had a history of cancer. The observed CHIP prevalence in GTEx is similar to what we would expect given the age demographics of the cohort and published prevalence rates (*2*).

dN/dS for the low-VAF prenatal PZM class was nominally greater than 1 (missense dN/dS = 1.25, $P = 0.047$) (Fig. 3E). The high-VAF postnatal nonsense mutations showed dN/dS values much less than 1, which can be attributed to sampling bias against transcripts carrying premature stop codons, because of nonsense-mediated decay (*34*). Altogether, the deleteriousness and selection results suggest a dichotomy

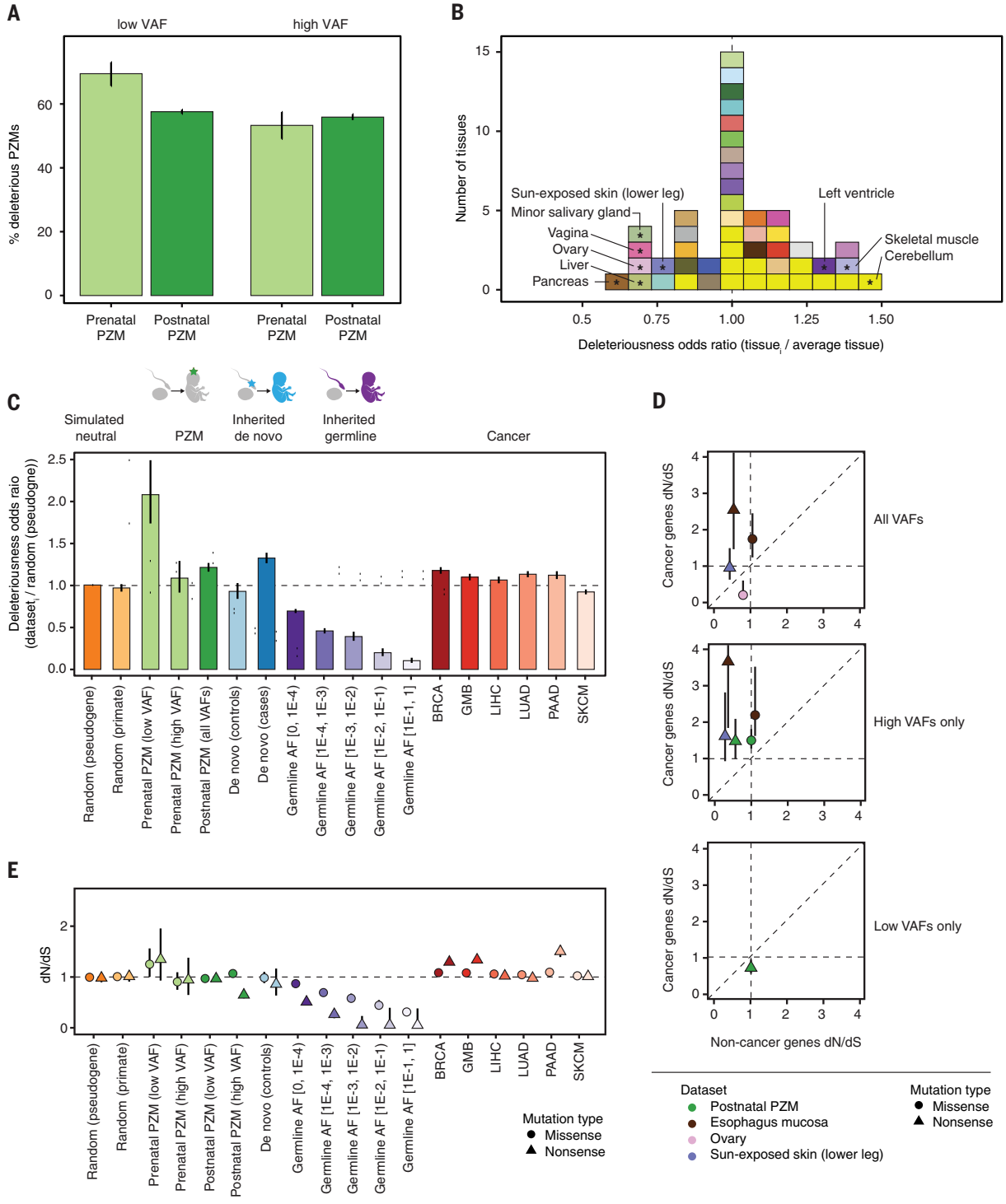


Fig. 3. Deleteriousness and selective pressure changes as a function of VAF, space, time, and classes of genetic variation. (**A**) Relative odds of detecting deleterious mutations across developmental time (gray bars) and VAF bins (green bars). (**B**) Histogram of the odds of detecting deleterious postnatal PZMs in each tissue compared with the average tissue. Tissues are colored using the GTEx coloring convention (see table S8 for a complete legend). Tissues with significant odds ratios (at $Q \leq 0.05$) are marked with asterisks and labeled with their names. The vertical dashed line at odds ratio = 1 indicates no difference in odds. (**C**) Relative odds of detecting deleterious PZM mutations compared with different classes of genetic variation. Dashed line at odds ratio = 1 indicates no difference in odds of detecting deleterious mutations compared with reference group. Error bars represent 95% CIs. (**D**) Comparison of postnatal PZM selection pressure in cancer and noncancer genes. For clarity, only PZM datasets that had different selection pressure between cancer and noncancer genes are shown. (Top) PZM datasets that had variable selection when using all mutations. (Middle) High-VAF mutations. (Bottom) Low-VAF mutations. Error bars represent 95% CIs. Some CIs are smaller than the data point so are not directly visible. (**E**) dN/dS values for classes of genetic variation, as in (C). CIs are plotted behind each data point and are sometimes smaller than the data point size. dN/dS = 1 indicates neutral expectation. AF, allele frequency; BRCA, breast invasive carcinoma; GBM, glioblastoma multiforme; LIHC, liver hepatocellular carcinoma; PAAD, pancreatic adenocarcinoma; SKCM, skin cutaneous melanoma.

between growth within an individual versus growth within a population: Mutations that are selected for within parts of an individual may be detrimental when considered across the entire life span.

Characterization of germ cell PZMs

Construction of a catalog of germ cell PZMs throughout the germ cell life cycle

Although a great deal is known about germline variation (*27*) and de novo mutations (*25*, *35*–*39*), much less is known about the PZMs that seed these forms of inherited genetic variation. To better understand PZMs in germ cells, we characterized and contrasted the mutation burden, spectra, and deleteriousness of germ cell PZMs across the germ cell life cycle.

Because of cell composition differences between male and female gonads, PZMs in testes samples could be confidently mapped to germ cells, but PZMs in ovary samples could not (*16*) (fig. S28 and table S15). Therefore, only testicular germ cell PZMs were analyzed further. Germ cell PZMs were classified into gonosomal (present in somatic and germ cells) and germ cell–specific. In total, 571 germ cell PZMs were identified in the bulk testis from 281 testis donors, of which 12% were putative gonosomal PZMs and the remaining 88% were putative germ cell–specific PZMs. As expected, germ cell–specific PZM burden was positively associated with donor age ($P = 0.03$), but gonosomal mutation burden was not ($P = 0.28$). Additionally, as expected, germ cell–specific PZMs had lower VAFs than gonosomal PZMs ($P = 1.3 \times 10^{-14}$, Mann-Whitney U test; fig. S28E).

Testicular germ cell PZMs represent the full reservoir of mutations that can be passed on to progeny. We hypothesized that the selection pressures on spermatogenesis, fertilization, and prenatal development may alter the types of mutations that pass through each of these bottlenecks of life. To examine germ cell PZMs that passed the spermatogenesis bottleneck, we generated whole-exome sequencing data on small, 200-cell pools of ejaculated sperm and identified and validated 83 PZMs in the same genomic regions that we assessed in the GTEx RNA-seq samples (defined as the allowable transcriptome) (*16*) (tables S1 and S16 and fig. S29). To examine germ cell PZMs that completed prenatal development, we used ~17,000 de novo mutations in the allowable transcriptome from denovo-db (*28*).

The mutation spectra for each germ cell mutation dataset were statistically different from the others (Fig. 4A and table S17; chi-square test). Although C > T was the most common mutation type in all datasets, C > A was the most variable. Hierarchical clustering of the spectra nested the classes in developmental order, indicating that the mutation spectra shift during development (Fig. 4A, inset). Given the complex ascertainment of these diverse mutation call-sets, we cannot exclude the possibility that some of the apparent structure is attributable to differences in mutation detection among sources, as a result of either bioinformatic or experimental effects.

Deleterious mutations are likely purged during the germ cell life cycle

Consistent with the action of purifying selection on male germ cells, we found that mutation deleteriousness decreased over the germ cell life cycle when comparing testicular germ cell PZMs and de novo mutations in controls (Fig. 4B). By contrast, de novo mutations from cases of disease were just as likely to be deleterious as testis PZMs. To replicate these observations, we performed a similar analysis using only DNA-based measurements from published datasets (*10*, *16*, *40*) (fig. S30). Both the fraction of coding mutations and the odds of detecting a deleterious mutation decreased over the germ cell life cycle in the independent datasets (fig. S30). Donor age was not associated with PZM deleteriousness in each dataset.

The mutation rate during male gametogenesis is dynamic

We estimated the mutation rate (the number of mutations in the transcriptome per cell division) for each of three major stages during male gametogenesis (*16*). Consistent with previous work (*37*), the observed mutation rate was higher in prenatal time points than the postnatal time point (Fig. 4C). The observed lower mutation rate during adulthood may be a strategy to limit the number of deleterious mutations that are passed to the next generation. Unlike Rahbari *et al.* (*37*) and other studies that have used transmitted de novo mutations to measure mutation rates (*35*, *36*, *38*, *39*), these estimates reflect mutation rates in germ cells in the testis and thus offer insight on germ cell mutagenesis.

Blood is a poor surrogate for measuring mosaicism of gonosomal PZMs

Motivated by the fact that only a small subset of tissue types is easily and ethically accessible in antemortem human subjects research, we hypothesized that more-accessible tissues may be useful surrogates for examining prenatal PZMs in less-accessible tissues. The results of such analyses may shed light on the cellular dynamics of human development and implications for preconception genetic counseling and de novo mutation discovery.

We fit a mixed-effects model to predict whether a gonosomal PZM was detected in a somatic tissue while controlling for technical effects (*16*). Unexpectedly, 88% (38/43) of tissues had significantly higher odds of detecting gonosomal PZMs than in blood (Fig. 4D), which suggests that blood is a poor surrogate for detecting gonosomal PZMs. Additionally, 76% (32/42) of somatic tissues had a significant linear correlation between the somatic VAF and the germ cell VAF (Fig. 4, E and F; $Q < 0.05$, Pearson's correlation test), which suggests that somatic tissues may offer a faithful representation of gonosomal PZMs in germ cells. These observations were not an artificial result of germline variant filtering or our cross-sample mutation-calling strategy (*16*) (fig. S31). Although 82% of GTEx donors were genotyped using blood, for 6% of GTEx donors, a nonblood tissue was used for genotyping, and for 12% of donors, no genotyping data were available. There was no detectable difference among these three groups in the probability of detecting a prenatal PZM in blood while controlling for other confounders; this suggests that poor detection of gonosomal PZMs in blood is not simply the result of aggressive germline filtering using genotype calls from blood-derived DNA (*16*).

Discussion

In this work, we present one of the most comprehensive and diverse surveys of PZM variation in normal individuals, which should prove a valuable resource for understanding the causes and consequences of PZMs across the body. By linking these mutation calls to the vast data and tissue resources of the GTEx project, there are a number of analyses that could be attempted. First, if there is a heritable component to PZM burden, variants modulating this burden may be detectable using genome-wide association studies (GWASs) (*41*, *42*). Second, the impact of PZMs on gene expression traits, both in cis and trans, can be directly assessed (*9*, *43*). Third, the spatial and cell-type distribution of the mutations reported here could be mapped in banked tissue samples from the GTEx donors (*7*, *44*), and the mutation type and burden of each sample could be associated with histology images collected by the GTEx project. We performed extensive validation of our PZM call-set, and these validation data will be helpful in training algorithms for PZM detection.

We observed a number of notable features regarding the developmental origins of mutations that deserve follow up. Most intriguing is a class of low-VAF prenatal mutations that appear to have the highest fraction of deleterious mutations across the human life span, even considering disease states. This observation, based on a definition of deleteriousness on an evolutionary time scale, suggests that the functional consequences of mutation can have opposite fitness effects at different stages of the life cycle of genomes and in different cellular contexts. One well-established example of marked differences in fitness effects between somatic and germline cells is the RAS-MAPK pathway, in which gain-of-function mutations provide a

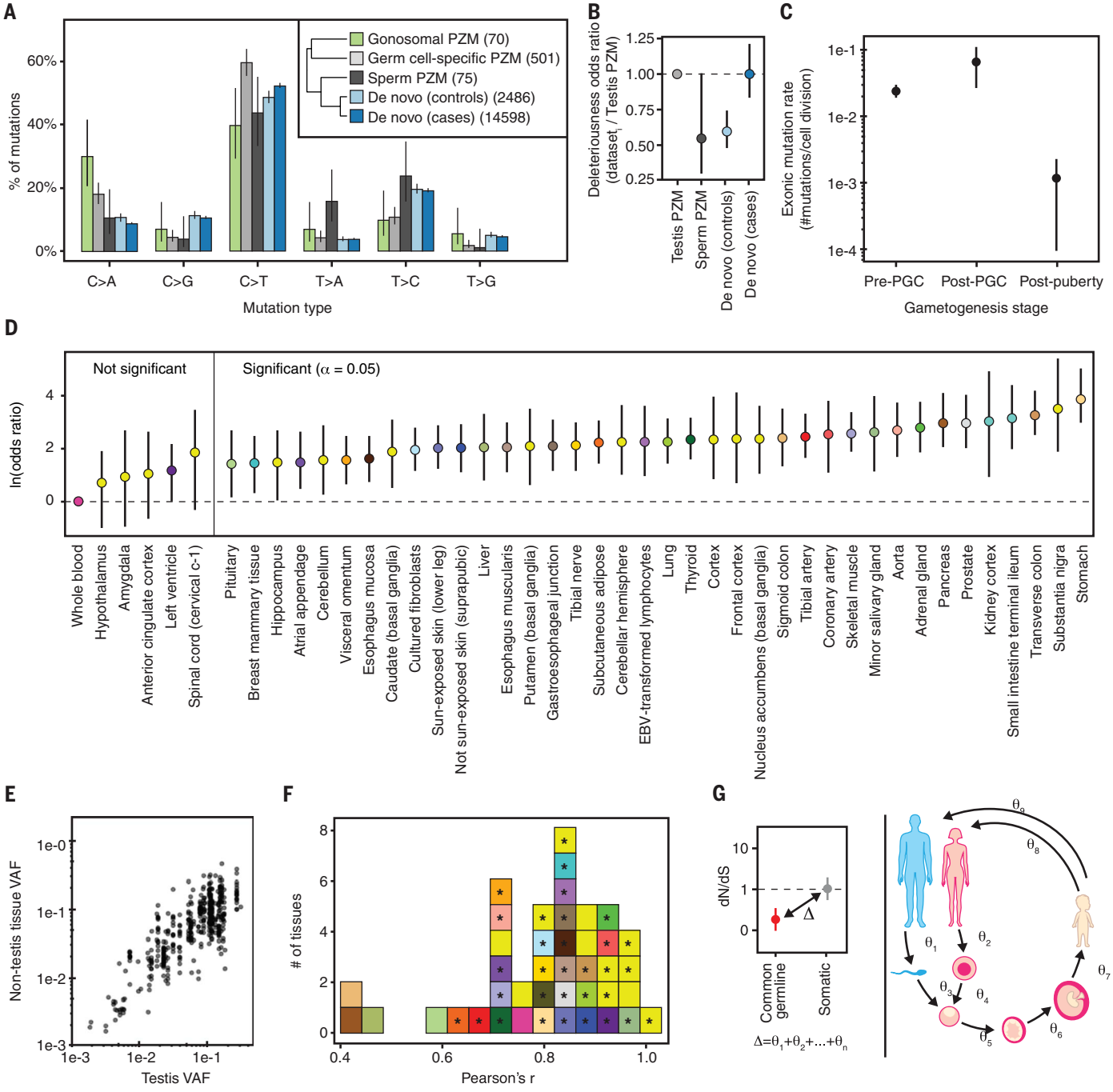


Fig. 4. Germ cell PZM characteristics. (**A**) Mutation spectra of different germ cell mutation classes. Number of mutations used in each dataset is listed in the inset. (Inset) Hierarchical clustering of germ cell mutation spectra. (**B**) Relative odds of detecting deleterious mutations across germ cell datasets compared with testis PZMs. Bars are colored by dataset. Horizontal black line at odds ratio = 1 denotes no difference in odds. (**C**) Germ cell mutation rate varies during gametogenesis in males. (**D**) Most somatic tissues have a higher odds of detecting a gonosomal PZM than blood. Natural log odds ratio for detecting a gonosomal PZM in each somatic tissue compared with blood. Dashed line at Y = 0 denotes no difference in odds. (**E**) Comparison of gonosomal PZM VAF in nontestis tissues versus testis tissue. (**F**) Distribution of tissue-specific Pearson correlations of $\log_{10}$-transformed gonosomal PZM VAFs in each somatic tissue and testis. Significant correlations at $Q \leq 0.05$ are marked with asterisks. (**G**) Schematic of the difference in selective constraint between germline and somatic genetic variation partitioned into discrete stages of the life cycle. [(A), (B), (C), and (D)] Error bars denote 95% CIs.

transmission advantage to male germ cells but are often reproductively lethal for the resulting conceptus (*45*, *46*). Although some parallels have been noted between molecular mechanisms of carcinogenesis and normal embryogenesis (*47*, *48*), there are essentially no data on the potential adaptive effects of PZMs on embryonic or fetal development in healthy individuals.

We advise caution in the interpretation of the dN/dS values for multitissue PZMs. Although we have evaluated obvious sources of technical error, such as the multitissue ascertainment (fig. S25) and small sample size (fig. S26), there may be other complexities influencing this rather general statistic, including recurrent mutation

and changes in mutation processes throughout development. Clearly it will be important to continue research into appropriate statistical methods for assessing fitness consequences of PZMs from multitissue datasets.

We found that blood-derived RNA appeared to be a poor proxy for detection of gonosomal mutations. On the basis of these results, for trio studies, we recommend that sperm (a direct readout of germ cells) should be profiled in males and skin (which is predicted to be more than five times as likely as blood to contain a gonosomal PZM) should be profiled in females. It should be noted that these findings on gonosomal PZMs were based on analysis of data exclusively from male tissues. We are optimistic that this conclusion will hold for female gonosomal mutations. In humans, male and female germ cells are both formed from a common progenitor cell type, primordial germ cells (PGCs). Early embryonic development, up to and including the formation of PGCs, is the time frame in which gonosomal mutations occur and is thought to occur identically in males and females (*49*). The developmental phylogeny that relates PGCs and the three germ layers is unclear, and the patterns of gonosomal mutations observed across human tissues may yield important insight into the matter. Some studies have indicated that PGCs may be most related to the mesoderm: Incipient mesoderm or mesendoderm cells can be induced to form PGC-like cells in vitro (*50*, *51*), and PGCs may share expression markers with mesoderm or primitive streak (*52*). However, loss of BLIMP1, a key driver of germline identity, from germline-competent cells leads to activation of a default neuronal differentiation program (*50*). When mapping gonosomal mutations frequencies across the body, we found that brain tissues were most similar to the testis (Fig. 4F). This might be an indication that PGCs and the ectoderm share a closer developmental origin.

We reported a large difference in deleteriousness and dN/dS inferred from PZMs and inherited germline variants, consistent with strong purifying selection reducing the transmission of deleterious mutation across generations. An important future direction is to dissect and quantify the physiological basis of this purifying selection (Fig. 4G). With careful thought and experimental design, it should be possible to model the steps of the human life cycle where purifying selection can occur, to estimate the strength of selection at each step, and to translate these data into life stage–specific measures of selective constraint for each gene in the genome. This would be of great benefit to human geneticists, who rely heavily on selective-constraint measures aggregated across the life cycle (such as CADD) for interpretation of genetic variants in the context of disease (*53*, *54*). Stage-specific constraint metrics could augment current methods for variant interpretation to be more relevant to the tissue and developmental time affected by a disease.

Materials and methods summary

GTEx data

We detected PZMs in the GTEx v8 dataset. To achieve a high-quality dataset, we removed RNA-seq samples that had an RNA integrity number (RIN) < 6, were derived from tissues with overall poor quality (*8*), or had an extremely high PZM mutation burden (table S2). We also confirmed that none of the analyzed samples were from transplanted tissue. After our quality control, there were 14,672 samples from 944 donors from 48 diverse tissue and cell types. Library preparation, sequencing, alignment, and GTEx quality control are described in detail in Aguet *et al.* (*15*).

Algorithms for detecting PZMs

LachesisDetect contains four basic steps. First, alignment files are filtered for extremely high-quality alignments. Next, the algorithm leverages cohort-wide information by simultaneously analyzing all samples to estimate position-specific error models for >115 Mb of the transcriptome. LachesisDetect uses these models to detect putative PZMs with single-sample calling. Third, the method removes sources of false-positive PZMs, such as RNA editing and allele-specific expression of germline variants, using >15 filters based on theoretical and experimental validation metrics. In the last step, the method leverages donor information by jointly analyzing all samples in a donor to detect mutations with low power and estimate empirical false-positive rates (fig. S1).

PZM validation

We performed several orthogonal validation experiments to quantify the FDR of the mutation-calling algorithm. These efforts included both in silico and experimental approaches and involved analysis of both DNA and RNA from the tissues used for mutation detection. A summary of the validation results is provided in table S3. We analyzed independent genomics datasets generated by the ENCODE project on four GTEx donors, encompassing 245 DNA assays and 67 RNA assays generated from four GTEx donors. Finally, we generated our own validation data by performing targeted DNA-seq of >1650 putative PZMs using DNA from GTEx donors.

Mutation burden modeling

To evaluate biological and technical sources of variation in mutation burden, we used linear mixed-effect models. We explored a large variety of model choices to arrive at our final modeling framework, comparing modeling choices using deviance, stability of model fitting, and other diagnostics. We used type II ANOVA to summarize the relative contributions of covariates.

Algorithm for mapping PZMs to a developmental tree

We manually derived two developmental tissue trees that represent the phylogenetic relationships among GTEx tissues during human development using information from the literature—the full tree and the simplified germ layer tree. We then developed an algorithm, LachesisMap, to reconstruct the phylogenetic history of multitissue PZMs. The algorithm jointly analyzes all multitissue PZMs and accounts for missing data as well as differential PZM detection power resulting from differences in VAF, expression level, and tissue profiling in the dataset.

Sperm sequencing experiments

Ejaculated sperm and venous blood were collected from a European American. Sperm samples had normal sperm density, sperm motility, and morphology. Fresh ejaculates were stained using the LIVE/DEAD Sperm Viability Kit (Invitrogen) and propidium iodide (PI). Sperm samples were then selectively sorted using fluorescence-activated cell sorting (FACS) into 96 well plates (~200 sperm cells per well) and 5-ml Falcon tubes based on their staining. We used MALBAC amplification (*55*) to prepare up to 1.5 μg of DNA from each pool of sperm using a kit, and six pools were selected for sequencing. Exome library preparation was performed according to the manufacturer's protocol using 50 ng of preamplified MALBAC reactions or DNA extracted from blood.

REFERENCES AND NOTES

1. D. Melzer, L. C. Pilling, L. Ferrucci, The genetics of human ageing. *Nat. Rev. Genet.* **21**, 88–101 (2020). doi: 10.1038/s41576-019-0183-6; pmid: 31690828
2. G. Genovese *et al.*, Clonal hematopoiesis and blood-cancer risk inferred from blood DNA sequence. *N. Engl. J. Med.* **371**, 2477–2487 (2014). doi: 10.1056/NEJMoa1409405; pmid: 25426838
3. M. Xie *et al.*, Age-related mutations associated with clonal hematopoietic expansion and malignancies. *Nat. Med.* **20**, 1472–1478 (2014). doi: 10.1038/nm.3733; pmid: 25326804
4. A. L. Young, G. A. Challen, B. M. Birmann, T. E. Druley, Clonal haematopoiesis harbouring AML-associated mutations is ubiquitous in healthy adults. *Nat. Commun.* **7**, 12484 (2016). doi: 10.1038/ncomms12484; pmid: 27546487
5. M. A. Lodato *et al.*, Somatic mutation in single human neurons tracks developmental and transcriptional history. *Science* **350**, 94–98 (2015). doi: 10.1126/science.aab1785; pmid: 26430121
6. A. Abyzov *et al.*, One thousand somatic SNVs per skin fibroblast cell set baseline of mosaic mutational load with patterns that suggest proliferative origin. *Genome Res.* **27**, 512–523 (2017). doi: 10.1101/gr.215517.116; pmid: 28235832
7. I. Martincorena *et al.*, High burden and pervasive positive selection of somatic mutations in normal human skin. *Science* **348**, 880–886 (2015). doi: 10.1126/science.aaa6806; pmid: 25999502
8. K. Yizhak *et al.*, RNA sequence analysis reveals macroscopic somatic clonal expansion across normal tissues. *Science* **364**, eaaw0726 (2019). doi: 10.1126/science.aaw0726; pmid: 31171663
9. P. E. García-Nieto, A. J. Morrison, H. B. Fraser, The somatic mutation landscape of the human body. *Genome Biol.* **20**, 298 (2019). doi: 10.1186/s13059-019-1919-5; pmid: 31874648

10. L. Moore *et al*., The mutational landscape of human somatic and germline cells. *Nature* **597**, 381–386 (2021). doi: 10.1038/s41586-021-03822-7; pmid: 34433962
11. I. Martincorena *et al*., Universal patterns of selection in cancer and somatic tissues. *Cell* **171**, 1029–1041.e21 (2017). doi: 10.1016/j.cell.2018.06.001; pmid: 29906452
12. T. N. Wong *et al*., Role of *TP53* mutations in the origin and evolution of therapy-related acute myeloid leukaemia. *Nature* **518**, 552–555 (2015). doi: 10.1038/nature13968; pmid: 25487151
13. X. Chen *et al*., Non-invasive early detection of cancer four years before conventional diagnosis using a blood test. *Nat. Commun.* **11**, 3475 (2020). doi: 10.1038/s41467-020-17316-z; pmid: 32694610
14. T. H. H. Coorens *et al*., Extensive phylogenies of human development inferred from somatic mutations. *Nature* **597**, 387–392 (2021). doi: 10.1038/s41586-021-03790-y; pmid: 34433963
15. F. Aguet *et al*., The GTEx Consortium atlas of genetic regulatory effects across human tissues. bioRxiv 787903 [Preprint] (2019). https://doi.org/10.1101/787903
16. See the supplementary materials.
17. M. Enge *et al*., Single-cell analysis of human pancreas reveals transcriptional signatures of aging and somatic mutation patterns. *Cell* **171**, 321–330.e14 (2017). doi: 10.1016/j.cell.2017.09.004; pmid: 28965763
18. US Cancer Statistics Working Group, US Cancer Statistics Data Visualizations Tool (Centers for Disease Control and Prevention, 2021); https://www.cdc.gov/cancer/dataviz.
19. J. M. Goldmann *et al*., Differences in the number of de novo mutations between individuals are due to small family-specific effects and stochasticity. *Genome Res.* **31**, 1513–1518 (2021). doi: 10.1101/gr.271809.120; pmid: 34301630
20. S. A. Frank, Somatic evolutionary genomics: Mutations during development cause highly variable genetic mosaicism with risk of cancer and neurodegeneration. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 1725–1730 (2010). doi: 10.1073/pnas.0909343106; pmid: 19805033
21. M. D. Lynch *et al*., Spatial constraints govern competition of mutant clones in human epidermis. *Nat. Commun.* **8**, 1119 (2017). doi: 10.1038/s41467-017-00993-8; pmid: 29066762
22. L. B. Alexandrov *et al*., The repertoire of mutational signatures in human cancer. *Nature* **578**, 94–101 (2020). doi: 10.1038/s41586-020-1943-3; pmid: 32025018
23. J.-J. Gao *et al*., Highly variable recessive lethal or nearly lethal mutation rates during germ-line development of male *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 15914–15919 (2011). doi: 10.1073/pnas.1100233108; pmid: 21890796
24. E. Kuijk *et al*., Early divergence of mutational processes in human fetal tissues. *Sci. Adv.* **5**, eaaw1271 (2019). doi: 10.1126/sciadv.aaw1271; pmid: 31149636
25. S. J. Lindsay, R. Rahbari, J. Kaplanis, T. Keane, M. E. Hurles, Similarities and differences in patterns of germline mutation between mice and humans. *Nat. Commun.* **10**, 4053 (2019). doi: 10.1038/s41467-019-12023-w; pmid: 31492841
26. M. Kircher *et al*., A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* **46**, 310–315 (2014). doi: 10.1038/ng.2892; pmid: 24487276
27. K. J. Karczewski *et al*., The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434–443 (2020). doi: 10.1038/s41586-020-2308-7; pmid: 32461654
28. T. N. Turner *et al*., denovo-db: A compendium of human *de novo* variants. *Nucleic Acids Res.* **45**, D804–D811 (2017). doi: 10.1093/nar/gkw865; pmid: 27907889
29. ICGC/TCGA Pan-Cancer Analysis of Whole Genomes Consortium, Pan-cancer analysis of whole genomes. *Nature* **578**, 82–93 (2020). doi: 10.1038/s41586-020-1969-6; pmid: 32025007
30. N. Goldman, Z. Yang, A codon-based model of nucleotide substitution for protein-coding DNA sequences. *Mol. Biol. Evol.* **11**, 725–736 (1994). doi: 10.1093/oxfordjournals.molbev.a040153; pmid: 7968486
31. I. Martincorena *et al*., Somatic mutant clones colonize the human esophagus with age. *Science* **362**, 911–917 (2018). doi: 10.1126/science.aau3879; pmid: 30337457
32. P. S. Ward *et al*., The common feature of leukemia-associated IDH1 and IDH2 mutations is a neomorphic enzyme activity converting α-ketoglutarate to 2-hydroxyglutarate. *Cancer Cell* **17**, 225–234 (2010). doi: 10.1016/j.ccr.2010.01.020; pmid: 20171147
33. V. N. Ngo *et al*., Oncogenically active *MYD88* mutations in human lymphoma. *Nature* **470**, 115–119 (2011). doi: 10.1038/nature09671; pmid: 21179087
34. M. A. Rivas *et al*., Effect of predicted protein-truncating genetic variants on the human transcriptome. *Science* **348**, 666–669 (2015). doi: 10.1126/science.1261877; pmid: 25954003
35. 1000 Genomes Project, Variation in genome-wide mutation rates within and between human families. *Nat. Genet.* **43**, 712–714 (2011). doi: 10.1038/ng.862; pmid: 21666693
36. J. C. Roach *et al*., Analysis of genetic inheritance in a family quartet by whole-genome sequencing. *Science* **328**, 636–639 (2010). doi: 10.1126/science.1186802; pmid: 20220176
37. R. Rahbari *et al*., Timing, rates and spectra of human germline mutation. *Nat. Genet.* **48**, 126–133 (2016). doi: 10.1038/ng.3469; pmid: 26656846
38. A. Kong *et al*., Rate of de novo mutations and the importance of father's age to disease risk. *Nature* **488**, 471–475 (2012). doi: 10.1038/nature11396; pmid: 22914163
39. J. J. Michaelson *et al*., Whole-genome sequencing in autism identifies hot spots for de novo germline mutation. *Cell* **151**, 1431–1442 (2012). doi: 10.1016/j.cell.2012.11.019; pmid: 23260136
40. X. Yang *et al*., Developmental and temporal characteristics of clonal sperm mosaicism. *Cell* **184**, 4772–4783.e15 (2021). doi: 10.1016/j.cell.2021.07.024; pmid: 34388390
41. A. Hodgkinson *et al*., High-resolution genomic analysis of human mitochondrial RNA sequence variation. *Science* **344**, 413–415 (2014). doi: 10.1126/science.1251110; pmid: 24763589
42. P.-R. Loh *et al*., Insights into clonal haematopoiesis from 8,342 mosaic chromosomal alterations. *Nature* **559**, 350–355 (2018). doi: 10.1038/s41586-018-0321-x; pmid: 29995854
43. J. Ding *et al*., Systematic analysis of somatic mutations impacting gene expression in 12 tumour types. *Nat. Commun.* **6**, 8554 (2015). doi: 10.1038/ncomms9554; pmid: 26436532
44. A. Erickson *et al*., Spatially resolved clonal copy number alterations in benign and malignant tissue. *Nature* **608**, 360–367 (2022). doi: 10.1038/s41586-022-05023-2; pmid: 35948708
45. G. J. Maher *et al*., Selfish mutations dysregulating RAS-MAPK signaling are pervasive in aged human testes. *Genome Res.* **28**, 1779–1790 (2018). doi: 10.1101/gr.239186.118; pmid: 30355600
46. A. Goriely, A. O. M. Wilkie, Paternal age effect mutations and selfish spermatogonial selection: Causes and consequences for human disease. *Am. J. Hum. Genet.* **90**, 175–200 (2012). doi: 10.1016/j.ajhg.2011.12.017; pmid: 22325359
47. N. M. Aiello, B. Z. Stanger, Echoes of the embryo: Using the developmental biology toolkit to study cancer. *Dis. Model. Mech.* **9**, 105–114 (2016). doi: 10.1242/dmm.023184; pmid: 26839398
48. T. Gong *et al*., A time-resolved multi-omic atlas of the developing mouse liver. *Genome Res.* **30**, 263–275 (2020). doi: 10.1101/gr.253328.119; pmid: 32051188
49. Y.-T. Lin, B. Capel, Cell fate commitment during mammalian sex determination. *Curr. Opin. Genet. Dev.* **32**, 144–152 (2015). doi: 10.1016/j.gde.2015.03.003; pmid: 25841206
50. K. Sasaki *et al*., Robust in vitro induction of human germ cell fate from pluripotent stem cells. *Cell Stem Cell* **17**, 178–194 (2015). doi: 10.1016/j.stem.2015.06.014; pmid: 26189426
51. T. Kobayashi *et al*., Principles of early human development and germ cell program from conserved model systems. *Nature* **546**, 416–420 (2017). doi: 10.1038/nature22812; pmid: 28607482
52. R. C. V. Tyser *et al*., Single-cell transcriptomic characterization of a gastrulating human embryo. *Nature* **600**, 285–289 (2021). doi: 10.1038/s41586-021-04158-y; pmid: 34789876
53. S. Petrovski, Q. Wang, E. L. Heinzen, A. S. Allen, D. B. Goldstein, Genic intolerance to functional variation and the interpretation of personal genomes. *PLOS Genet.* **9**, e1003709 (2013). doi: 10.1371/journal.pgen.1003709; pmid: 23990802
54. K. Eilbeck, A. Quinlan, M. Yandell, Settling the score: Variant prioritization and Mendelian disease. *Nat. Rev. Genet.* **18**, 599–612 (2017). doi: 10.1038/nrg.2017.52; pmid: 28804138
55. P. G. Engström *et al*., Systematic evaluation of spliced alignment programs for RNA-seq data. *Nat. Methods* **10**, 1185–1191 (2013). doi: 10.1038/nmeth.2722; pmid: 24185836
56. N. Rockweiler, D. Contrad, conradlab/RockweilerEtAl: v0.1.0-alpha, version 0.1.0-alpha, Zenodo (2022); https://doi.org/10.5281/zenodo.7378922.

ACKNOWLEDGMENTS

We thank the GTEx donors and families for their generous and selfless donation of tissues and organs for the advancement of science. We thank T. Lappalainen (New York Genome Center and KTH Royal Institute of Technology), H. Leitch (London Institute of Medical Sciences), G. Coop (UC Davis), I. Martincorena (Sanger Institute), the M. Griffiths laboratory (Washington University), the S. Montgomery laboratory (Stanford University), and the Conrad and Cohen laboratories (Washington University) for data sharing and helpful discussions. Sequencing for the validation experiments was performed by the Genome Technology Access Center (GTAC) in the Department of Genetics at Washington University School of Medicine. We appreciate obtaining access to de novo mutations on SFARI base (https://base.sfari.org). We thank the following funders for their support of the TwinsUK resource: Wellcome Trust, Medical Research Council, European Union, Chronic Disease Research Foundation (CDRF), the National Institute for Health Research (NIHR)–funded BioResource Clinical Research Facility, and the Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London. We also acknowledge that this manuscript includes several analyses that focus on GTEx donors with European and African American ancestry and less so on other ancestries in GTEx. This decision was made so that we would have higher statistical power to detect potentially rare genetic signals that may vary across populations. We hope that in future studies, larger numbers of Black, Indigenous, and People of Color are included so that we can learn about all populations in an equitable manner. **Funding:** This study was supported by National Institutes of Health grants R01MH101810 (to D.F.C.), R01HG007178 (to D.F.C.), and R01HD078641 (to D.F.C.); National Human Genome Research Institute grant T32HG000045 (to B.A.C.); and Medical Research Council grants MR/M004422/1 (to K.S.S.) and MR/R023131/1 (to K.S.S.). **Author contributions:** D.F.C. conceived the study. N.B.R. designed and performed research and analyzed data. W.H.W. and E.Z.T. assisted with experimental design and sample procurement and generated sequencing libraries for validation experiments. M.J.N. and L.N. performed and analyzed data from the sperm experiment. B.M. and A.Z. contributed to analyses. A.R., C.W.D., and N.H. contributed to the study design. K.A.V.-C. contributed to data visualization. D.F.C., K.G.A., and B.A.C. supervised the project. N.B.R. and D.F.C. wrote the manuscript in consultation with all authors. **Competing interests:** D.F.C. is an advisor to Paterna Biosciences. All other authors declare that they have no competing interests. **Data and materials availability:** All GTEx protected data are available through the database of Genotypes and Phenotypes (dbGaP) (accession no. phs000424.v8). Access to the raw sequence data is provided through the AnVIL platform (https://gtexportal.org/home/protectedDataAccess). The cancer results are based on data generated by the TCGA Research Network: https://www.cancer.gov/ccg/research/genome-sequencing/tcga. De novo mutations were obtained from denovo-db (https://denovo-db.gs.washington.edu/denovo-db/). Data generated on GTEx tissues by the ENCODE project are available from https://www.encodeproject.org. Source code used in this study is available from https://github.com/conradlab/RockweilerEtAl. Archived source code at the time of publication is available from Zenodo (*56*). **License information:** Copyright © 2023 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

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abn7113
Materials and Methods
Figs. S1 to S31
Tables S1 to S17
References (*57*–*123*)
MDAR Reproducibility Checklist

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

Submitted 16 December 2021; accepted 17 March 2023
10.1126/science.abn7113

RESEARCH ARTICLE SUMMARY

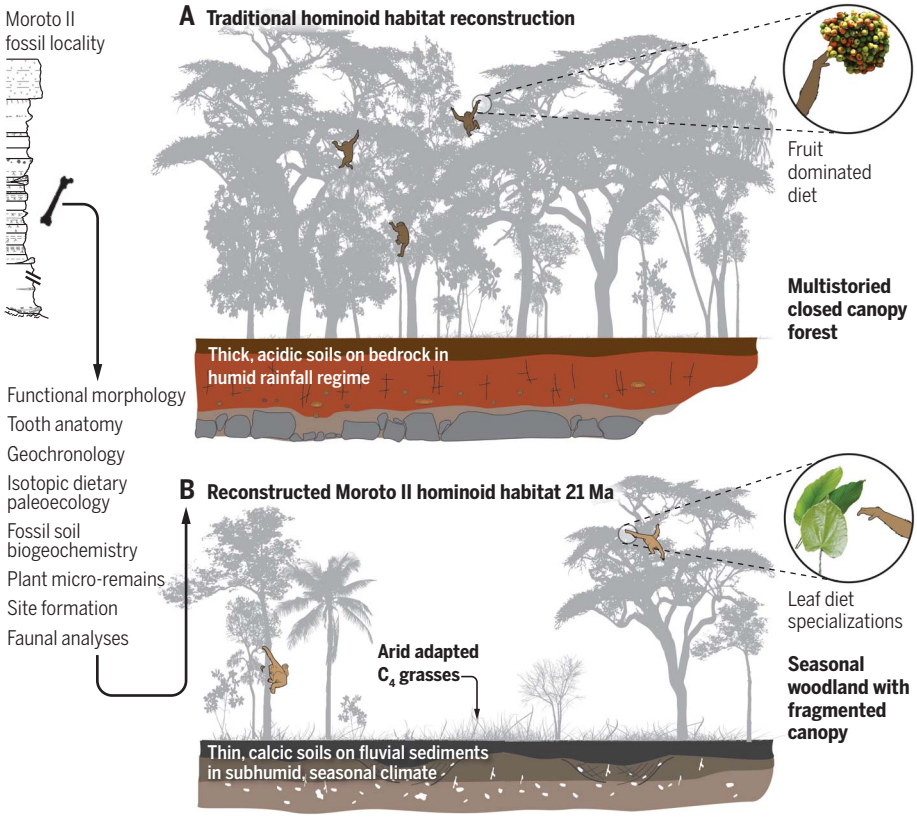
PALEOECOLOGY

The evolution of hominoid locomotor versatility: Evidence from Moroto, a 21 Ma site in Uganda

Laura M. MacLatchy*, Susanne M. Cote, Alan L. Deino, Robert M. Kityo, Amon A. T. Mugume, James B. Rossie, William J. Sanders, Miranda N. Cosman, Steven G. Driese, David L. Fox, April J. Freeman, Rutger J. W. Jansma, Kirsten E. H. Jenkins, Rahab N. Kinyanjui, William E. Lukens, Kieran P. McNulty, Alice Novello, Daniel J. Peppe, Caroline A. E. Strömberg, Kevin T. Uno, Alisa J. Winkler, John D. Kingston

INTRODUCTION: Inherent in traditional views of ape origins is the idea that, like living apes, early large-bodied apes lived in tropical forests. In response to constraints related to locomoting in forest canopies, it has been proposed that early apes evolved their quintessential upright torsos and acrobatic climbing and suspensory abilities, enhancing their locomotor versatility, to distribute their weight among small supports and thus reach ripe fruit in the terminal branches. This feeding and locomotor transition from a quadruped with a horizontal torso is thought to have occurred in the Middle Miocene due to an increasingly seasonal climate and feeding competition from evolving monkeys. Although ecological and behavioral comparisons among living apes and monkeys provide evidence for versions of terminal branch forest frugivory hypotheses, corroboration from the early ape fossil record has been lacking, as have detailed reconstructions of the habitats where the first apes evolved.

RATIONALE: The Early Miocene fossil site of Moroto II in Uganda provides a unique opportunity to test the predictions of terminal branch forest frugivory hypotheses. Moroto II documents the oldest [21 million years ago (Ma)] well-established paleontological record of ape teeth and postcranial bones from a single locality and preserves paleoecological pro… to reconstruct the environment. The follo… ing lines of evidence from Moroto II were analyzed: (i) the functional anatomy of femora and a vertebra attributed to the ape *Morotopithecus*; (ii) dental traits, including molar shape and isotopic profiles of *Morotopithecus* enamel; (iii) isotopic dietary paleoecology of associated fossil mammals; (iv) biogeochemical signals from paleosols (ancient soils) that reflect local relative proportions of C_3 (trees and shrubs) and C_4 (tropical grasses and sedges that can endure water stress) vegetation as well as rainfall; and (v) assemblages of phytoliths, microscopic plant-derived silica bodies that reflect past plant communities.



Hominoid habitat comparisons. Shown are reconstructions of a traditionally conceived hominoid habitat (**A**) and the 21 Ma Moroto II, Uganda, habitat (**B**).

RESULTS: A short, strong femur biomechanically favorable to vertical climbing and a vertebra indicating a dorsostable lower back confirm that ape fossils from Moroto II shared locomotor traits with living apes. Both *Morotopithecus* and a smaller ape from the site have elongated molars with well-developed crests for shearing leaves. Carbon isotopic signatures of the enamel of these apes and of other fossil mammals indicate that some mammals consistently fed on water-stressed C_3 plants, and possibly also C_4 vegetation, in a woodland setting. Carbon isotope values of pedogenic carbonates, paleosol organic matter, and plant waxes all point to substantial C_4 grass biomass on the landscape. Analysis of paleosols also indicates subhumid, strongly seasonal rainfall, and phytolith assemblages include forms from both arid-adapted C_4 grasses and forest-indicator plants.

CONCLUSION: The ancient co-occurrence of dental specializations for leaf eating, rather than ripe fruit consumption, along with apelike locomotor abilities counters the predictions of the terminal branch forest frugivory hypotheses. The combined paleoecological evidence situates *Morotopithecus* in a woodland with a broken canopy and substantial grass understory including C_4 species. These findings call for a new paradigm for the evolutionary origins of early apes. We propose that seasonal, wooded environments may have exerted previously unrecognized selective pressures in the evolution of arboreal apes. For example, some apes may have needed to access leaves in the higher canopy in times of low fruit availability and to be adept at ascending and descending from trees that lacked a continuous canopy. ■

The list of author affiliations is available in the full article online.
*Corresponding author. Email: maclatch@umich.edu
Cite this article as L. M. MacLatchy *et al.*, *Science* **380**, eabq2835 (2023). DOI: 10.1126/science.abq2835

READ THE FULL ARTICLE AT
https://doi.org/10.1126/science.abq2835

RESEARCH ARTICLE

PALEOECOLOGY

The evolution of hominoid locomotor versatility: Evidence from Moroto, a 21 Ma site in Uganda

Laura M. MacLatchy[1,2]*, Susanne M. Cote[3], Alan L. Deino[4], Robert M. Kityo[5], Amon A. T. Mugume[5,6], James B. Rossie[7], William J. Sanders[1,2], Miranda N. Cosman[1], Steven G. Driese[8], David L. Fox[9], April J. Freeman[8], Rutger J. W. Jansma[10], Kirsten E. H. Jenkins[11], Rahab N. Kinyanjui[12,13,14], William E. Lukens[15], Kieran P. McNulty[16], Alice Novello[17,18], Daniel J. Peppe[8], Caroline A. E. Strömberg[18], Kevin T. Uno[19], Alisa J. Winkler[20,21], John D. Kingston[1]

Living hominoids are distinguished by upright torsos and versatile locomotion. It is hypothesized that these features evolved for feeding on fruit from terminal branches in forests. To investigate the evolutionary context of hominoid adaptive origins, we analyzed multiple paleoenvironmental proxies in conjunction with hominoid fossils from the Moroto II site in Uganda. The data indicate seasonally dry woodlands with the earliest evidence of abundant C_4 grasses in Africa based on a confirmed age of 21 million years ago (Ma). We demonstrate that the leaf-eating hominoid *Morotopithecus* consumed water-stressed vegetation, and postcrania from the site indicate ape-like locomotor adaptations. These findings suggest that the origin of hominoid locomotor versatility is associated with foraging on leaves in heterogeneous, open woodlands rather than forests.

Living nonhuman hominoids are distinguished from most other primates by postcranial adaptations that facilitate arboreal positional behaviors in which the torso is upright (orthograde) and the fore- and hindlimbs often have wide excursions and are oriented in different directions. Most hypotheses seeking to explain the adaptive significance of these versatile behaviors propose that they help medium- and large-bodied primates forage for ripe fruit in the terminal branches or the upper canopy of forests, where small, flexible branches must be navigated [e.g., (*1*–*4*)]. Such terminal branch forest frugivory hypotheses predict that versatile positional behaviors (including both postures and locomotion) would have first appeared in hominoids that were committed frugivores and were foraging in tropical forest habitats.

The Moroto II Miocene fossil site in Uganda (Fig. 1) provides a unique opportunity to test these predictions because it documents the oldest well-established paleontological record of hominoid teeth and postcranial bones from a single locality. Here, we report on new radiometric dates, data on fossil hominoid diet and locomotion, and multiple paleoecological proxies from Moroto II that allow us to elucidate the evolutionary context and refine the nature and timing of key aspects of early hominoid adaptive origins.

Background

Terminal branch forest frugivory hypotheses

Hominoid arboreal positional behavior is particularly versatile because it involves transitions among a range of postural and locomotor modes and across a range of substrates. This versatile behavior fundamentally stems from having a dorsostable, orthograde torso, which in turn facilitates differential limb use and the ability to have limbs aligned differently while serving a weight-bearing function. Although this use of the term "versatile" is not established in the literature, which has an accepted positional behavioral terminology that that both consolidates and simplifies nomenclature (*4*–*7*), we use it as shorthand to refer to behaviors such as vertical climbing on substrates oriented ≥45°, arboreal travel during which multiple limbs are propulsive but oriented differently, and suspension (below-branch behaviors) (movies S1 to S3). Versatile behavior thus represents an alternative to quadrupedalism, in which the torso is roughly horizontal (pronograde), limbs are generally below the body, and joint excursions are more modest (*4*). Versatile behaviors are thought to be especially important when moving in the trees with a large body size and have been linked to hominoid-specific anatomical specializations, including dorsoventrally shallow but mediolaterally broad torsos; dorsally placed scapulae; relatively short, dorsostable lumbar regions; relatively long forelimbs paired with relatively short hindlimbs; mobile limb joints; manual and pedal specializations for grasping; and tail loss (*8*, *9*).

In attempting to explain the evolutionary origins of the unique hominoid body plan, researchers have focused on particular versatile behaviors, including suspension (*10*, *11*), quadrumanous climbing (*12*), and vertical climbing (*13*, *14*), all of which normally incorporate the use of limbs outside of the parasagittal plane. These capabilities also rely on an upright, stable torso, so it has been further suggested that it is the capacity for orthogrady that may be largely responsible for hominoid positional behavioral distinctiveness (*6*, *15*).

Several influential hypotheses have been advanced to explain the adaptive significance of the acquisition of versatile positional behaviors in hominoids and are here grouped under a single umbrella as the terminal branch forest frugivory hypotheses. The earlier forms of these hypotheses focused on substrate. It was proposed that versatile behaviors contribute to the fitness of medium- and large-bodied apes by allowing them to use multiple supports to distribute their weight and thereby successfully forage either in the terminal branches generally or in the middle and upper canopy of forests [e.g., (*1*, *2*, *16*, *17*)]. Enhanced competition with cercopithecoid monkeys and/or increasing seasonality were invoked (*2*, *10*, *18*) to explain the increase in body size. Such models of hominoid-cercopithecoid divergence also typically situate cercopithecoids in more open habitats [e.g., (*3*, *18*, *19*)].

Forests figure prominently in these models of hominoid divergence. They are generally defined as distinct from woodlands in having greater structural complexity and taxonomic richness, usually including taller trees with multistoried, intermingling crowns, with the highest level of foliage (the canopy) more closed (*20*, *21*). By contrast, woodlands in tropical and subtropical Africa tend to have a single-story broken canopy, extensive ground

[1]Department of Anthropology, University of Michigan, Ann Arbor, MI 48109, USA. [2]Museum of Paleontology, University of Michigan, Ann Arbor, MI 48109, USA. [3]Department of Anthropology and Archaeology, University of Calgary, Calgary, Alberta T2N 1N4, Canada. [4]Berkeley Geochronology Center, Berkeley, CA 94709, USA. [5]Department of Zoology, Entomology and Fisheries Sciences, Makerere University, Kampala, Uganda. [6]Uganda National Museum, Department of Museums and Monuments, Ministry of Tourism, Wildlife and Antiquities, Kampala, Uganda. [7]Department of Anthropology, Stony Brook University, Stony Brook, NY 11794, USA. [8]Department of Geosciences, Baylor University, Waco, TX 76798, USA. [9]Department of Earth & Environmental Sciences, University of Minnesota, Minneapolis, MN 55455, USA. [10]Institute of Human Origins, School of Human Evolution and Social Change, Arizona State University, Tempe, AZ 85281, USA. [11]Department of Social Sciences, Tacoma Community College, Tacoma, WA 98466, USA. [12]Earth Sciences Department, National Museums of Kenya, Nairobi, Kenya. [13]Max Planck Institute for Geoanthropology, Jena D-07743, Germany. [14]Human Origins Program, National Museum of Natural History, Smithsonian Institution, Washington, DC 20013, USA. [15]Department of Geology & Environmental Science, James Madison University, Harrisonburg, VA 22807, USA. [16]Department of Anthropology, University of Minnesota, Minneapolis, MN 55455, USA. [17]CEREGE, Aix-Marseille Université, CNRS, IRD, Collège de France, INRAE, Aix en Provence, France. [18]Department of Biology & Burke Museum of Natural History and Culture, University of Washington, Seattle, WA 98195, USA. [19]Division of Biology and Paleo Environment, Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY 10964, USA. [20]Roy M. Huffington Department of Earth Sciences, Southern Methodist University, Dallas, TX 75275, USA. [21]Section of Anatomy, Department of Surgery, University of Texas Southwestern Medical Center, Dallas, TX 75390, USA.
*Corresponding author. Email: maclatch@umich.edu

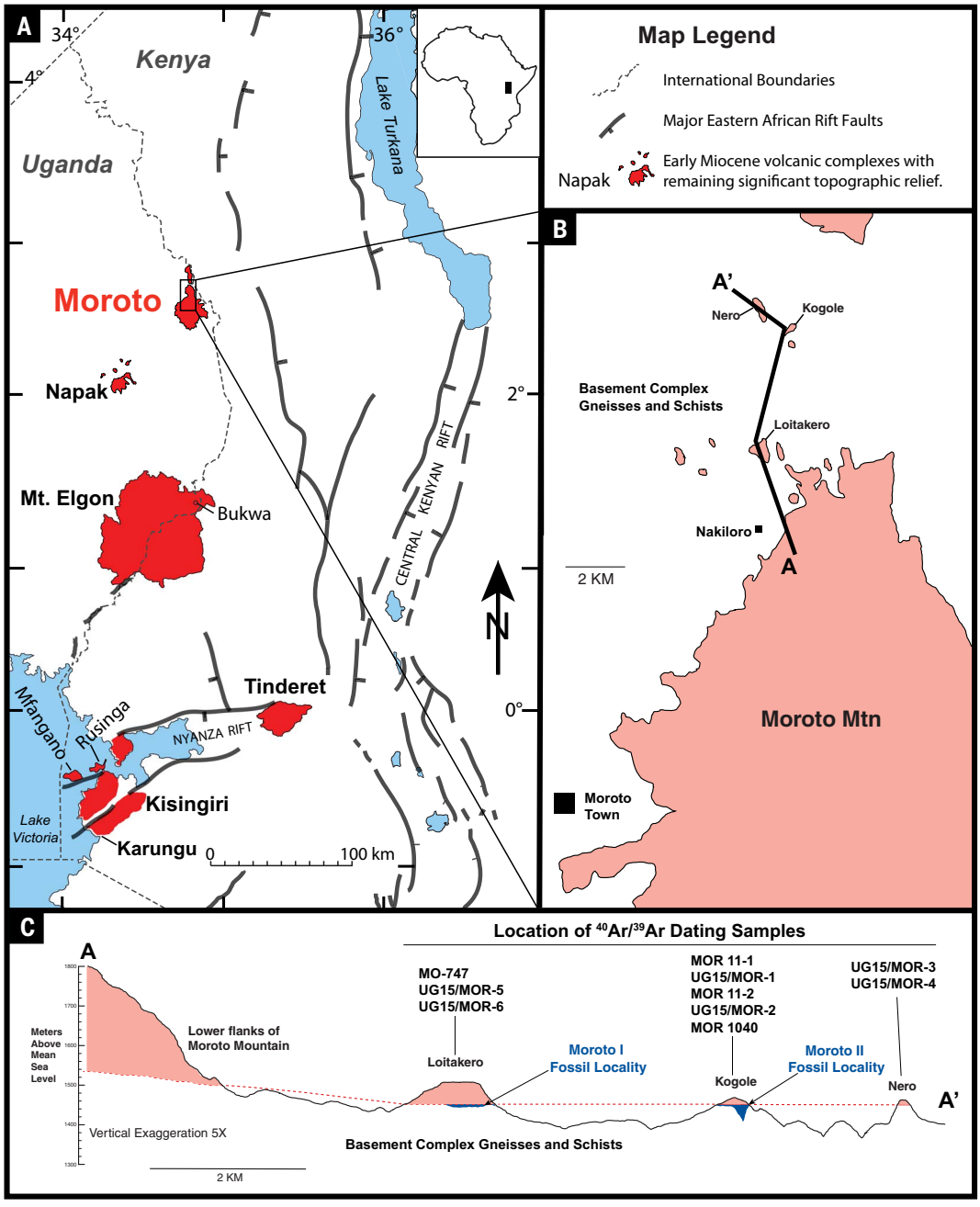


Fig. 1. Location maps and cross sections of the Moroto region. (**A**) Regional map of the Uganda-Kenya rift system depicting major Early Miocene fossil site complexes (in red). (**B**) Inset of current Moroto volcanic edifice and environs, including the location of Moroto II at Kogole and Moroto I at Loitakero (volcanics are shown in pink and the metamorphic basement in white). (**C**) Topographic cross-sectional profile [depicted in (B)] showing the intersection with the hills of Loitakero, Kogole, and Nero overlying the fossil sites (in blue). Locations of $^{40}Ar/^{39}Ar$ dating samples from lavas capping three topographic highs are shown.

cover including grasses, and distinctive dry-wet seasons. Consequently, forests have greater connectivity between adjacent trees, as well as more shade and moisture. Ripley (*18*) proposed that terminal branch feeding on fruit would have occurred in evergreen forests. Temerin and Cant (*3*) also postulated that hominoid divergence occurred within a forest environment. These primatologists, along with paleontologists [e.g., (*19*)], speculated that the development of nonforested habitats occurred in the Middle Miocene (~15 Ma), with ancestral cercopithecoids exploiting the woodlands or savannas and hominoids remaining in tropical forests (*19*, *22*).

Frugivory was emphasized in most detail by Temerin and Cant's (*3*) insightful model of hominoid origins, which held that versatility evolved to increase speed and decrease the path length of arboreal travel in forest habitats under conditions of decreased fruit availability. These authors postulated that suspension- and forelimb-dominated climbing allowed apes to outcompete monkeys (as well as arboreal squirrels, bats, and birds) for rare, small, and/or dispersed high-quality fruit patches, whereas monkeys, which are restricted to longer, above-branch pathways and must frequently return to the core of a tree, exploited abundant ripe and nonripe fruit and leaves. Temerin and Cant

(3) proposed that a reduction in fruit availability, driven by a cooler and drier climate, was the selection pressure leading to the ecological divergence of hominoids and cercopithecoids. A more recent synthesis by Hunt (4) also invoked competition with cercopithecoids, in which small size, bilophodonty, and greater tolerance for the chemical defenses of plants allowed them to more successfully exploit unripe or small patches of fruit and leaves in the center of tree crowns. Apes, in turn, evolved postcranial specializations to feed on fruit in the terminal branches (4).

Although most reconstructions of hominoid foraging ecology focus on fruit, Sarmiento (23) proposed that the slow movements found in larger forms of select mammals, including hominoids, lorisines, colobines, the fossil lemur *Paleopropithecus*, and sloths, are correlated with folivory, although no causal links to an ecological context were proposed. The importance of terminal branch resources has been strengthened by research documenting that at least some fruits are larger, more abundant, denser, and most nutritionally valuable in the upper canopy (24); and leaves in the upper canopy may have higher protein (25), with young leaves having relatively higher protein than mature leaves (26). The higher quality and/or quantity of food in the upper canopies is thus a potential driver of versatile positional behaviors among hominoids.

There remains a prevailing and persistent perception that early hominoid evolution played out in a forested context and that frugivory was an important component. This relates to a number of long-held observations and interpretations, including: (i) all modern hominoids (and several fossil taxa) exhibit and are defined by arboreal adaptations for which behavioral versatility (as we define it) is key; (ii) the preferred habitats of modern hominoids are forests, so forest dwelling could be construed as a synapomorphy of the hominoid clade; (iii) the preferred diet of modern hominoids is largely focused on fruit; and (iv) the Early Miocene is viewed as consisting of a pan-equatorial African forest [see (27), which presents an expansive, regional interpretation of the vegetation of the Early Miocene]. The terminal branch forest frugivory hypotheses represent a formal distillation of these concepts, and we invoke them as representative of broad perceptions that persist today.

The Moroto II locality

As described by Bishop and Whyte (28), fossils in Moroto District were first found just north of Moroto Mountain (Moroto I) in 1959 by J. G. Wilson. The site of Moroto II was subsequently discovered in 1962 (28). The sites were prospected by Bishop and colleagues between 1961 and 1965 (28–30), resulting in the William Bishop fossil collection. Ongoing research at Moroto II by our team has produced fossil collections from the years 1994 to 2017.

Initial K/Ar dates for the capping basaltic lavas at the fossil sites in the 1960s yielded ages of 12.5 and 14.3 Ma (31), but more recent $^{40}Ar/^{39}Ar$ incremental dating techniques resulted in an age of 20 to 20.6 Ma (32). Although this date has been questioned and an age of 17.5 Ma advocated on the basis of qualitative biostratigraphic interpretations (33–35), the more comprehensive geochronologic strategy described below unequivocally supports an age of 21 Ma for the Moroto fossil localities.

The hominoids of Moroto II

Morotopithecus bishopi is one of at least two large-bodied primates from the Early Miocene fossil site of Moroto II in Uganda (32, 36). Its status as one of the oldest fossil hominoids, coupled with derived postcranial traits, including a dorsostable lower back, has placed it in a unique position among its Miocene contemporaries (37–39), because other Early Miocene (e.g., *Proconsul* and *Ekembo*) and even Middle Miocene (e.g., *Equatorius* and *Kenyapithecus*) African taxa retain ancestral, monkey-like body plans (37–39). Although the distinctiveness of *M. bishopi* from *Afropithecus turkanensis* (40, 41) and the composition of the *Morotopithecus* hypodigm (42, 43) have been questioned, recent (36) and current analyses support our assignments of functionally informative dental and postcranial specimens recovered at Moroto II to *Morotopithecus*.

The distinctiveness of the holotype maxilla and partial face UMP 62-11 described by Allbrook and Bishop (29), as well as the fragmentary mandibular remains UMP 62-10 and UMP 66-01 described by Pilbeam (44), were recently reassessed by MacLatchy *et al.* (36) after the discovery of another partial mandible (UMP MORII 03'551) and an isolated *Proconsul*-like tooth (UMP MORII 03′559) at Moroto II. Comparisons of taxonomically, developmentally, and functionally diagnostic characteristics, including tooth crown morphology and tooth row proportions, relative enamel thickness, enamel-dentine junction morphology, long-period line periodicity, and dental wear patterns, clearly discriminate among *Afropithecus*, *Morotopithecus*, and the *Proconsul* clade and support genus-level distinctiveness for all three (36). In addition, it was concluded, given the similarity in the tooth row proportions between the holotype UMP 62-11 and the mandible UMP MORII 03'551 and the complementary morphology between mandibles MORII 03'551, UMP 66-01, and UMP 62-10 (including having a long tooth row and vertical planum alveolare), that all of these specimens were best assigned to *M. bishopi* (36).

The functional assessments of these dental specimens provide evidence for specialized leaf eating (36) rather than the more generalized fruit-based diet reconstructed for primitive catarrhines (45, 46). Both the upper and lower molars are mesiodistally elongated, and an unworn M_2 has conspicuous, well-developed shearing crests (36), a characteristic long accepted as being reflective of folivory (45, 46). The presence of features for specialized leaf eating in both the upper and lower dentition further supports the attribution of all of these specimens to *Morotopithecus* (36). Although incipient signals of folivory have been documented in some smaller-bodied Early Miocene catarrhines (45), this combination of features was previously unknown for any great ape–sized (>30 kg) catarrhine before the Late Miocene. A folivorous diet is corroborated by thin occlusal enamel (23, 36), which, although not necessarily diagnostic for leaf eating, rules out the hard-object feeding inferred for other fossil taxa, including *Afropithecus* (47).

Two hominoid postcranial remains from Moroto II attributed to *M. bishopi* represent the oldest clear evidence of derived locomotor adaptations in hominoids: a lumbar vertebra [UMP 67-28; originally described by Walker and Rose (48)] indicative of a dorsoventrally stable lumbar spine consistent with orthogrady (37, 49, 50) and associated partial femora [UMP MORII 94′80; first described by Gebo *et al.* (32)] reflective of vertical climbing and slow forms of orthograde or pronograde clamber (15, 50, 51). These are core derived behaviors of the hominoid clade emphasized in the terminal branch forest frugivory hypotheses. The right femur is now complete and new functional implications are described below.

The precise location within Moroto II where the hominoid vertebra UMP 67-28 and some dental remains (UMP 66-01, UMP 62-10, and the holotype UMP 62-11) were found ~50 years ago is not documented, so there will always be ambiguity about their association. However, the mandible bearing the lightly worn M_2 with well-developed shearing crests (UMP MORII 03′551) and the right and left femora (UMP MORII 94′80) were found by our team within a single stratigraphic level (Fig. 2A) (36) from a restricted spatial distribution (see geologic setting below) and are the only hominoid fossils known from this level. As noted above, we believe that the holotype UMP 62-11 and mandible UMP MORII 03′551 belong in the same taxon because of their complementary morphology and specialized features for folivory. Likewise, the vertebra UMP 67-28 and the femora UMP MORII 94′80 share adaptive characteristics (see below). The co-occurrence of hominoid folivory and ape-like climbing abilities (regardless of taxonomic attribution) at a single, constrained stratigraphic level counters the predictions of the terminal branch forest frugivory hypotheses and makes reconstructing the environment and age of Moroto II highly relevant for developing adaptive explanations for the early stages of the hominoid lineage.

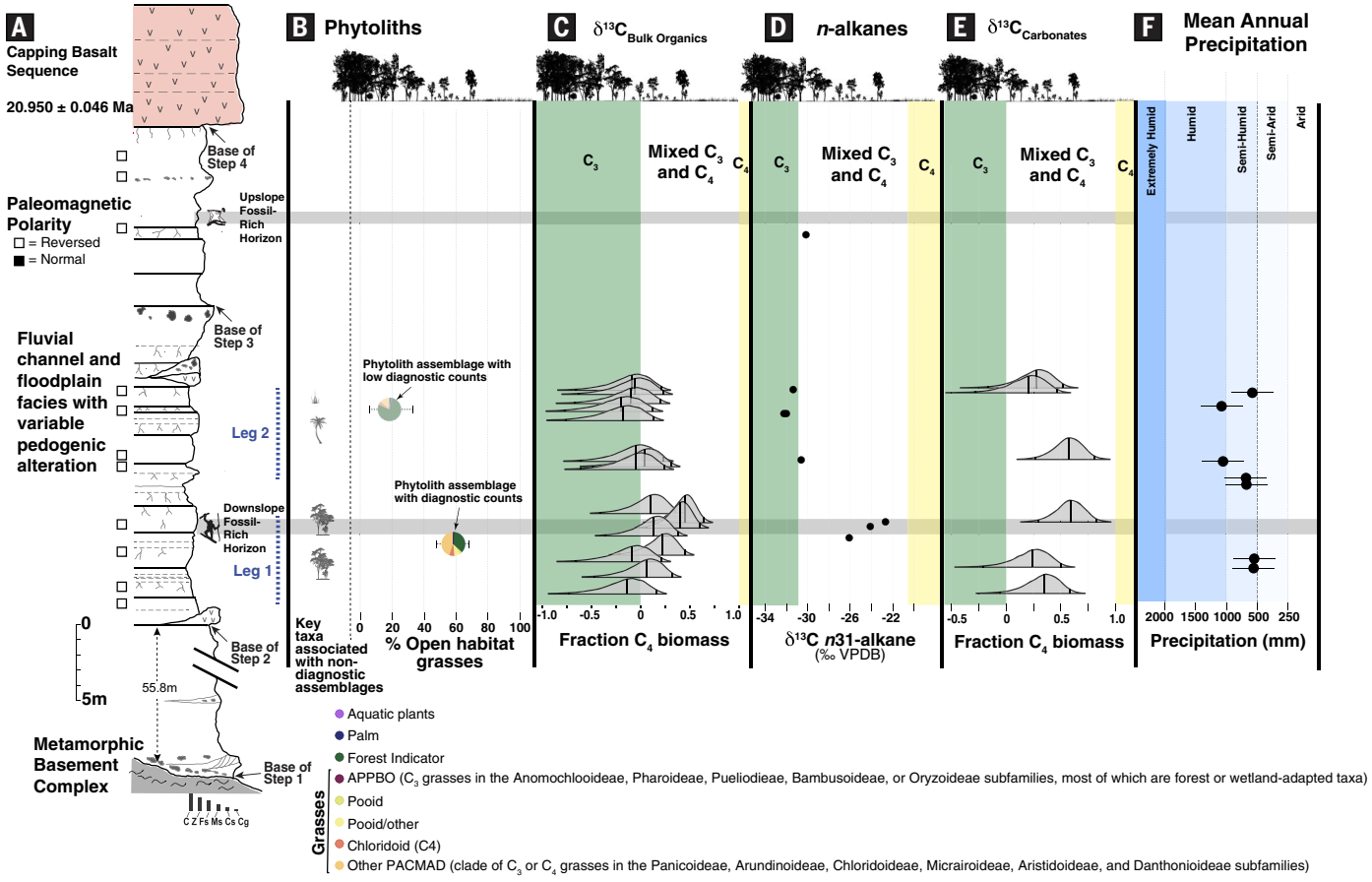


Fig. 2. Moroto II paleoenvironmental proxies in stratigraphic context. (**A**) Composite stratigraphic section at Moroto II depicting generalized sedimentary lithofacies, relationship to overlying basalt sequence, magnetostratigraphy, and fossil-rich horizons. (**B**) Percentage of open habitat grasses on the basis of phytolith assemblages, with the dark pie chart representing diagnostic assemblage data; the faded pie chart assemblage representing low diagnostic counts; and the grass, forest, and palm images representing key taxa associated with nondiagnostic assemblages. (**C**) Fraction of C_4 biomass on the basis of $\delta^{13}C_{Bulk\ Organics}$ ($\delta^{13}C_{OM}$). Fractions of <0 and >1 result from Monte Carlo resamples less than and greater than median plant $\delta^{13}C$ reference values, respectively (*52*). (**D**) C_{31} *n*-alkane homolog. (**E**) Fraction of C_4 biomass on the basis of $\delta^{13}C_{Carbonates}$ ($\delta^{13}C_{PC}$) components (*52*). (**F**) MAP estimates on the basis of the random forest (RF)–MAP machine learning model (*89*) with uncertainty ranges.

Results

Geologic and taphonomic setting

Fossiliferous sediments exposed at the site of Moroto II represent a fluvial system with deposition in a deeply incised channel eroded directly in the underlying Precambrian metamorphic basement complex and are directly overlain by a basaltic lava sequence. Deposits exposed at Moroto II consist almost exclusively of fluvial facies with minor fine-grained lacustrine or ponding facies, perhaps within the paleochannel(s) or floodplains. Sediments have been overprinted by pedogenic processes to varying degrees, suggesting that the channel and associated floodplains were intermittently inactive or meandered laterally. The site is currently ~11 km north of Mount Moroto, a volcanic complex that formed along the tectonically active west margin of the eastern branch of the East African Rift System during the Early Miocene, in conjunction with a series of eruptive centers extending north to the Turkana Basin and south into western Kenya (Fig. 1A). A lack of volcanoclastic debris in the sedimentary sequence indicates deposition on the basement complex that was independent of any input from the Moroto volcano. These observations suggest that although the site was formed adjacent to a developing eruptive center, the fossil assemblage more likely sampled ecosystems associated with the basement peneplain rather than the flanks of a volcano.

Fossils were recovered from multiple fluvial channels in fine- to coarse-grained sediments. Occasionally, elements from multiple taxa have been recovered together in small pockets within floodplain or overbank deposits, but most of the specimens derive mainly from two concentrations within the succession (Fig. 2A). The spatial proximity of the specimens recovered at each of these two stratigraphic levels (confined to an ~10 × 10 m area) indicates that fossils are generally associated. Taphonomic analysis (*52*) indicates that fossil assemblages likely represent time-averaged streamside accumulations in an open environment without a dominant biogenic accumulating process. Weathering patterns reveal that remains were often exposed to sunlight or frequent wetting and drying, which may suggest a more open environment (*53*). Attritional death assemblages such as Moroto II may be faithful indicators of broader ancient living assemblages (*54*, *55*), with the exception of small taxa that are underrepresented in our collections.

Chronology

The basaltic sequence overlying the fossiliferous sediments represents the initial input of volcanic material in the vicinity of the fossil localities. This overlying lava flow is in direct superposition with a paleosol surface in the Moroto II succession, indicating relatively rapid emplacement of the lavas within and over the active channel system without a protracted temporal hiatus (Figs. 1C and 2A). There are currently a number of topographic highs immediately northwest of the main Moroto volcanic edifice capped by erosional

remnants of the same sequence of lavas (Fig. 1C). Ten basalt samples were dated by $^{40}Ar/^{39}Ar$ laser incremental heating [Fig. 1C, figs. S1 to S3, and tables S1 to S3 (*52*)] from outcrops at multiple levels within the lava sequence at Moroto II (Kogole Hill and the nearby hills of Loitakero, which overlies the Moroto I fossil site) and Nero (*52*) (Fig. 1 and figs. S1 to S3). Ages are consistent within error, yielding a mean of 20.950 ± 0.046 Ma, corroborating previous radiometric ages of 20.6 Ma for the Loitakero basalt and ~20 Ma for the Kogole basalt (*32*).

On the basis of the reversed polarity of all samples at Moroto II, coupled with a representative age for the Moroto capping lavas of 20.950 ± 0.046 Ma and the stratigraphic relationship of the lava flows to the Moroto II sediments, our preferred correlation to the Geomagnetic Polarity Time Scale (GPTS) is to C6Ar, which occurred from 21.204 to 21.130 Ma (± 0.037 Ma) (*56*). Given that all samples are reversed, if the section sampled the entire chron, then the maximum duration of deposition of Moroto II was 365 ± 36.9 thousand years. On the basis of the sedimentological and paleosol features of the deposits, the site was likely deposited over a much shorter duration (*52*) (figs. S4 and S5 and tables S4 and S5).

This geochronology is entirely consistent with biostratigraphic data from the mammalian assemblage at Moroto II. There is a long sequence of Early Miocene fossil sites in Uganda and western Kenya that document faunal change in eastern Africa [e.g., (*57*, *58*]. A major aspect of this faunal change is the gradual transition from the afrotherian-dominated assemblages characteristic of the Paleogene to assemblages that are dominated by laurasiatherian immigrants by the onset of the Middle Miocene. Our detailed comparisons of the identifiable mammal fossils (n = 146) (table S6) provide strong support for the radiometric date, because some common laurasiatherian taxa are completely absent. By 20 Ma, and documented at sites such as Songhor and Rusinga, eastern African faunas incorporate additional laurasiatherian immigrants such as suids, tragulids, or perissodactyls and additional rodent lineages [e.g., (*57*, *59*–*61*)]. By contrast, at Moroto II, there are no suids or tragulids, no perissodactyls, and hyracoids are relatively abundant (table S6), suggesting that this assemblage is older than 20 Ma.

The radiometric age is further supported by the proboscideans, which are the best-known mammals in the Moroto II assemblage. *Eozygodon*, in particular, has a limited stratigraphic range in Africa and is only known from sites ~20 to 19 Ma and older (*62*). In eastern Africa, *Eozygodon* is only found at Moroto II and Meswa Bridge [22.5 Ma (*63*)]. The Moroto II material displays more derived dental features (table S6), supporting a biostratigraphic framework in which Meswa Bridge is older than Moroto. In addition, at younger sites in eastern Africa, such as Rusinga, *Eozygodon* is replaced by the more derived *Zygolophodon*, and a second proboscidean genus at Moroto II, *Progomphotherium*, is succeeded by *Archaeobelodon* (*62*). Therefore, these two proboscidean taxa also support an age of >20 Ma. Overall, biostratigraphy provides strong support that Moroto II is older than the well-sampled ~20 Ma site complexes of Napak and Tinderet, but younger than the ~28- to 23 Ma Late Oligocene faunas from eastern Africa (*62*).

Hominoid and other catarrhine fossils

Femur

New catarrhine fossils from Moroto II include additional fragments of the right femur UMP MORII 94′80 attributed to *Morotopithecus* (*32*, *50*), which complete the previously partial shaft for a total length of 240.7 mm (data S1). Very few complete femora from large-bodied Miocene hominoids exist. Relative to the midshaft, femoral head, and bicondylar breadths, UMP MORII 94′80 is very short (*52*) (Fig. 3), like the femora of extant great apes and unlike those of large cercopithecoids and the large-bodied Miocene hominoid *Ekembo nyanzae* (*51*), which is considered to be a generalized pronograde quadruped (*38*) (Figs. 3 and 4A). The cortical bone of UMP MORII 94′80 is thick, as in living great apes (*50*, *51*), with a ratio of cortical area divided by subperiosteal area (CA/PA) of 0.83 at about midshaft (*52*) (fig. S6). This is greater than the CA/PA ratio of the Early Miocene ape *Ekembo heseloni* [0.67 for KNM-RU 2036 (*50*)]. Both CA (282 mm^2) and strength (J) (18117.49 mm^4) at about midshaft are high relative to shaft length (fig. S7) and resemble proportions found in *Pongo* (*52*). This heavy reinforcement of cortical bone in *Pongo* has been interpreted to reflect an emphasis on axial loading, and may be associated with static postures (*64*) and slow movement with prolonged muscular recruitment (*50*). The finding that the hominoid represented by UMP MORII 94′80 had very short but strong legs is also consistent with the biomechanical expectation for large-bodied vertical climbers, in which the ability to remain on vertical supports (especially large ones) requires differentiated limb use involving (i) sustained muscular contractions, (ii) leaning away from the substrate using relatively long forelimbs, and (iii) having relatively short hindlimbs that flex at the hip and knee (*65*) (Fig. 4B). In combination, the short legs and large body size of *Morotopithecus* also suggest that its vertical climbing kinematics may have been more similar to those of great apes than to those of hylobatids. Isler (*14*) has shown that compared with the smaller gibbons (who have longer hindlimbs for their body size; Fig. 3), great apes position the body closer to the substrate, climb more slowly, and use shorter strides with a more extended thigh and abducted hip.

As noted previously (*50*), UMP MORII 94′80 has a broad bicondylar breadth; a well-developed popliteal groove; a wide, shallow patella groove; and a medial condyle that is broader than the lateral condyle (*50*). These features are correlated with ape-like capacities in the knee for rotation and hindlimb abduction (*50*, *66*, *67*). Proximally, the femoral head is somewhat small compared with extant hominoids, as reflected in the smaller body mass estimates that were calculated by Ruff (*51*) on the basis of femoral head breadth (29.5 kg) versus bicondylar breadth (37.3 kg) using a combined catarrhine sample. However, the relative femoral head surface area falls between those of African apes and monkeys, implying at least moderate hip mobility, and is broadly consistent with locomotion that included vertical climbing and with locomotor inferences stemming from the femoral shaft and distal femur (*68*).

Vertebra

Lumbar vertebra UMP 67-28 from Moroto II shares anatomical features with hominoids that are absent in cercopithecoid monkeys (*37*, *48*, *49*). For example, transverse processes arise from the base of robust pedicles and are oriented dorsally (*49*), increasing the moment arm of the muscles iliocostalis and longissimus dorsi to resist ventroflexion (*69*). Also, the vertebra lacks anapophyses (tailed catarrhines have anapophyses as insertions for caudal vertebral extensors), has reduced ventral keeling, and exhibits a caudally inclined spinous process (which is associated with reduced dorsoventral mobility) (*37*, *49*). New cleaning and preparation of this vertebra in 2016 revealed the presence of salient ligamentous pits medial to the postzygapophyses; these had been filled with sediment and had not previously been obvious (data S2). The pits represent insertions for the ligamentum flavum, intervertebral ligaments positioned between laminae of adjacent vertebrae in mammals. Their primary function in the lumbar column is to limit ventroflexion and stabilize lumbar vertebrae on one another. The depth of these ligamentous pits in UMP 67-28 suggests ligaments of substantial strength. With the features mentioned above, these ligaments can be interpreted to help maintain orthogrady and lumbar stability, as opposed to increasing spinal mobility. These features differ from those found in the lumbar vertebrae of *Ekembo*, which were interpreted as being associated with monkey-like flexion and extension of the spine during quadrupedal locomotion (*37*).

Dental specimens

In addition to *Morotopithecus*, at least three smaller, non-cercopithecoid catarrhine taxa are represented in the 1994–2017 fossils collected

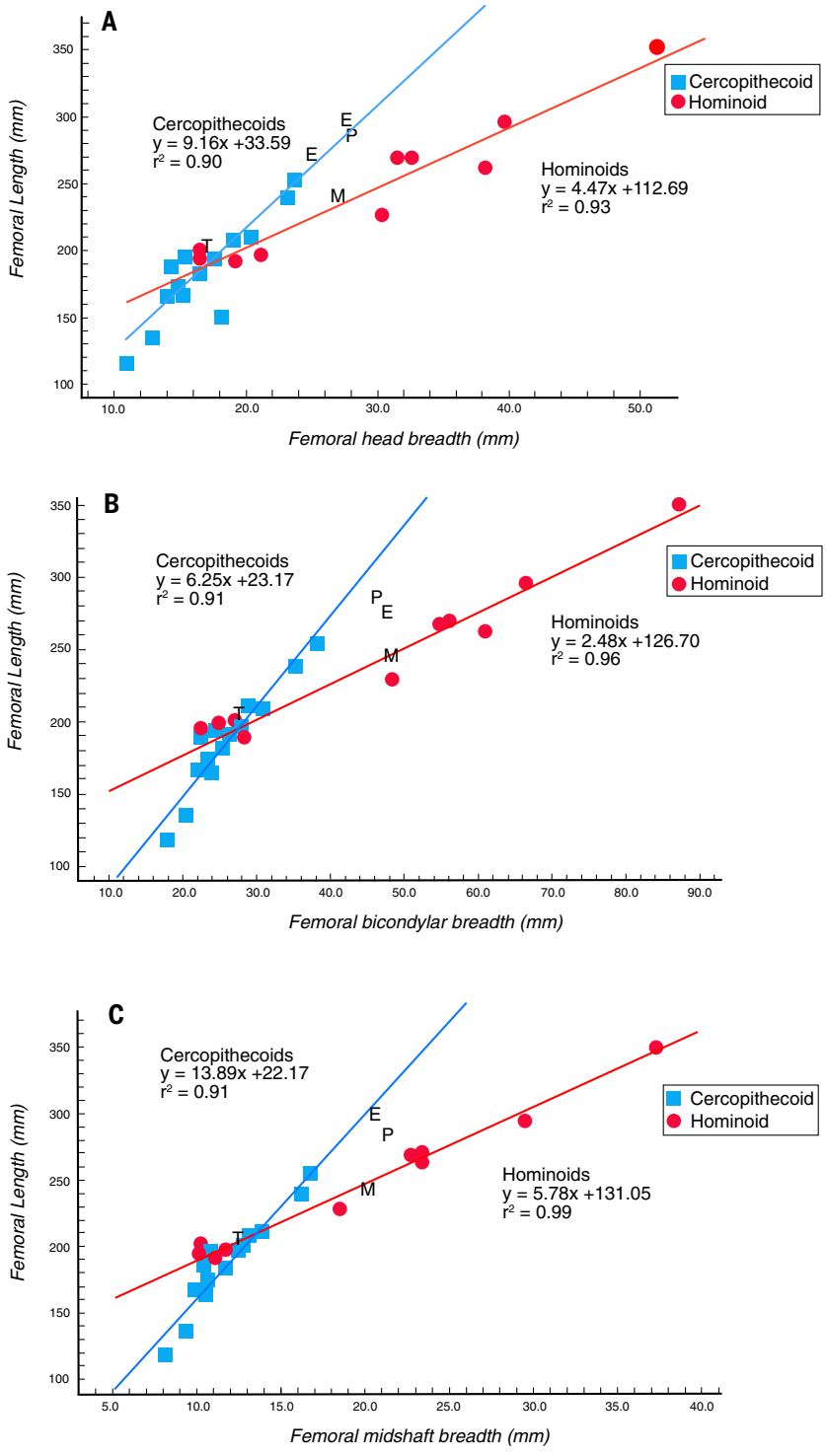


Fig. 3. Catarrhine femoral proportions. (**A** to **C**) Bivariate plots of mean femur length for extant hominoids and cercopithecoids regressed on mean femur bicondylar breadth (A), mean femur head superoinferior breadth (B), and mean femur midshaft breadth (C). Extant values are from (*51*) [see also (*52*)]. Fossil taxa are plotted and are from the following specimens: *Morotopithecus* (M) is from UMP MORII 94′80; *Paracolobus* (P) is from BC3; *Ekembo* (E) are from KNM-RU 5527 [(A) and (B)] and KNM-MW 13142 [(B) and (C)]; and *Turkanapithecus* (T) is from KNM WT 16950. Note that hylobatids and cercopithecoids cannot be distinguished on the basis of femoral proportions, and that femur length is only a diagnostic trait at larger body size.

by our team (table S6): a proconsulid (UMP MORII 03′559), aff. *Rangwapithecus* (UMP MOR II 11′1938; figs. S8 and S9 and table S7), and a small-bodied non-cercopithecoid catarrhine (UMP MORII 02′352) that is represented only by a lower canine that does not fit in the hypodigm of any known small catarrhine genus (fig. S10).

The lower third molar (UMP MOR II 11′1938) attributable to aff. *Rangwapithecus* is the first record of the genus outside of Kenya, and the oldest. We describe it in brief because *Rangwapithecus* has very well developed shearing crests compared with other Early Miocene catarrhines (*45*) and microwear denotive of folivory (*46*), indicating that another, smaller catarrhine with this dietary specialization was present at Moroto II. The specimen compares favorably to *R. gordoni* from Songhor and shares with it several traits, including a large and elongate crown, peripherally positioned conical cusps linked by rounded crests, and delicate secondary enamel crenulations (*42*, *70*). However, there are some features suggesting that it is more primitive and/or unique, including a narrower mesial buccal notch and less discrete buccal cusp outlines. Traits such as the well-developed mesostylid and lack of a crest between the hypoconulid and entoconid are reminiscent of traits found in the older nyanzapithecine from the Late Oligocene, *Rukwapithecus* (*71*). On balance, we find the specimen to have strong affinities to *Rangwapithecus*, with the differences likely reflecting the ancestral nyanzapithecine condition, which is compatible with its greater antiquity.

Stable isotopes of enamel

Dietary carbon isotopic signatures of molars from *Morotopithecus*, aff. *Rangwapithecus*, and the proconsulid reveal that all three taxa had ^{13}C-enriched diets relative to modern hominoids (Fig. 5B and table S8), indicating dietary niches and habitat preferences distinct from those of living apes (*52*). Specifically, the carbon isotopic signals of the enamel ($\delta^{13}C_{enamel}$) of the Moroto hominoids are consistent with water-stressed C_3 dietary vegetation, encroaching on the isotopic range of modern catarrhines that forage in more open habitats, such as *Papio* or some of the earliest hominins (Fig. 5B). Although the $\delta^{18}O_{enamel}$ value of *Morotopithecus* is within the isotopic range of associated large-bodied terrestrial fossil herbivores at Moroto II (e.g., proboscideans and anthracotheres), it is ^{18}O enriched relative to the other catarrhines (figs. S11B and S12 and table S8). As a taxon with derived orthograde and climbing capabilities, this suggests that *Morotopithecus* may have (despite its large size) more easily foraged in peripheral or higher parts of the canopy, where evapotranspiration and irradiance were higher and vegetation was more water stressed (*72*).

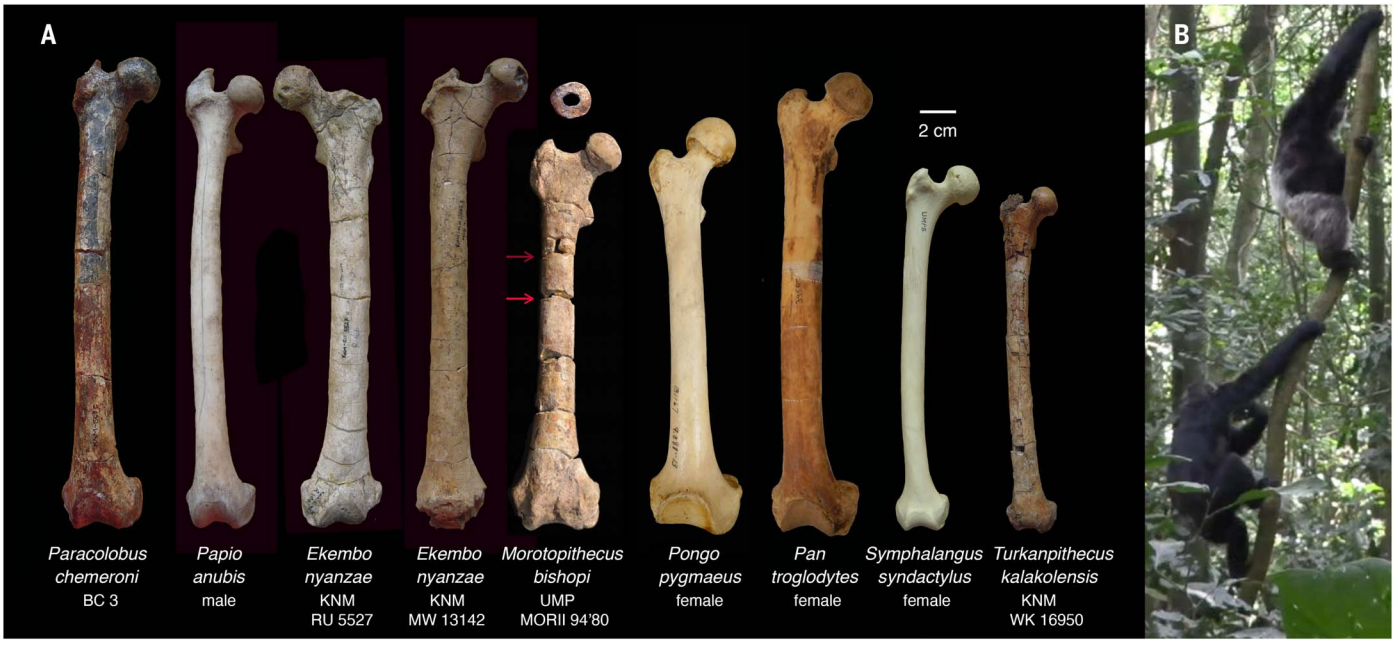


Fig. 4. Catarrhine femora. (**A**) Catarrhine femora, anterior view. The Pliocene cercopithecoid *Paracolobus* [mass ~36 kg based on regressions in (*51*)] and *Papio* [male mass ~25 kg (*103*)] have proportionately longer shafts than extant hominoids. *Ekembo* (~36 kg) and *Morotopithecus* (~35 kg) have similar femoral midshaft and joint dimensions and inferred body sizes (*51*), but the *Morotopithecus* femur is shorter. *Turkanapithecus* [~10 kg based on regressions in (*51*)] is Early Miocene in age and is the size of extant *Symphalangus* [~12 kg (*51*)]. The larger fossil taxa are about the same size as female *Pongo* [mass ~36 kg (*103*)], but smaller than female *Pan* [mass ~46 kg (*103*)]. *Pan*, *Pongo*, and *Morotopithecus* have shorter femora for their size than do cercopithecoids and *Ekembo*. The top red arrow adjacent to the *Morotopithecus* femur denotes where cortical bone was exposed and analyzed (*50*) when this fragment was hypothesized to be midshaft. The bottom arrow indicates the break exposing the cortical bone analyzed in this study, with its cross section pictured above. (**B**) Short femora and hindlimbs are biomechanically advantageous for vertical climbing. Shown are Jackson (adult male, top) and Sarah (adult female), chimpanzees from Kibale Forest in Uganda, vertically climbing a liana.

Paleoecology

Paleoecological reconstructions of fossil catarrhine and hominoid sites in eastern Africa have suggested a primitive phyloecological context for hominoid origins characterized by a complex canopy structure(s) associated with forested habitats [e.g., (*73*–*76*)]. However, evidence and interpretations from fossil sites in eastern Africa have also yielded empirical data implicating more heterogeneous habitats that include woodland and even wooded grassland components [e.g., (*27*, *77*, *78*)]. Overall, the fossil record leading up to the Early Miocene in eastern Africa has revealed variable ecosystems, although the general distribution and relevance of any particular habitat(s) for hominoid evolution remains unknown. Given this difficulty, it is important to situate the evolution of hominoids and their vertebrate communities specifically within local and regional paleoecological frameworks.

To assess the paleoecological context of the Moroto II fossil assemblage, several independent approaches were used, including (i) isotopic dietary ecology of fossil herbivore guilds; (ii) biogeochemical analyses of paleosol bulk organics, *n*-alkanes, and pedogenic carbonate; (iii) elemental characterization of paleosols to estimate mean annual precipitation (MAP); and (iv) analyses of phytolith assemblages. These approaches provide a composite perspective on the physiognomy and physiology of the vegetation but also reflect climatic factors such as seasonality and precipitation. Cumulatively, these data have relevance for constraining aspects of the canopy architecture and resource distribution that are critical for determining the adaptive significance of early hominoid locomotor and foraging behaviors.

Stable carbon isotopes of fossil enamel

Isotopic paleodietary ecology has figured prominently in discussions of the evolution of eastern African terrestrial communities, particularly in using $\delta^{13}C$ proxies to establish the relative significance of C_4 grasses in past ecosystems [e.g., (*79*, *80*)]. However, $\delta^{13}C_{enamel}$ datasets are comparatively rare from sites older than ~9 Ma. This is in part because Early and Middle Miocene sites were presumed to predate significant C_4 biomass in eastern Africa. Thus, enamel samples were not analyzed because the paleoecological reconstructive potential of isotopes based only on C_3 plant variation is limited. This study provides one of the first applications of this technique to the Early Miocene of eastern Africa, thereby filling a crucial data gap.

The application of enamel isotopic analyses for differentiating dietary guilds within C_3-dominated biomes or among C_3 browsers and grazers currently remains coarse. Although there is no strict relationship between canopy cover or architecture and $\delta^{13}C_{enamel}$ values, general trends related to water stress, irradiance, and canopy density have been established [e.g., (*81*, *82*)], providing guidelines for reconstructing general aspects of past landscapes and habitats. A framework generated for interpreting the isotopic signatures of fossil enamel at Moroto II [Fig. 5 (*52*)] is based on analysis of a compilation of $\delta^{13}C$ values of modern and fossil vegetation and herbivore enamel in both C_3-dominated habitats and mixed C_3/C_4 ecosystems.

$\delta^{13}C_{enamel}$ values of fossil enamel from 25 fossil herbivores at Moroto II, representing nine genera from seven families, range from –6.8‰ to –15.3‰, with ~85% of the values falling between –12‰ and –8‰ (Fig. 5B and table S8). This dietary range is most consistent with foraging in mesic to arid woodland ecosystems, as depicted in Fig. 5, with limited evidence of closed canopy forest or open woodland and

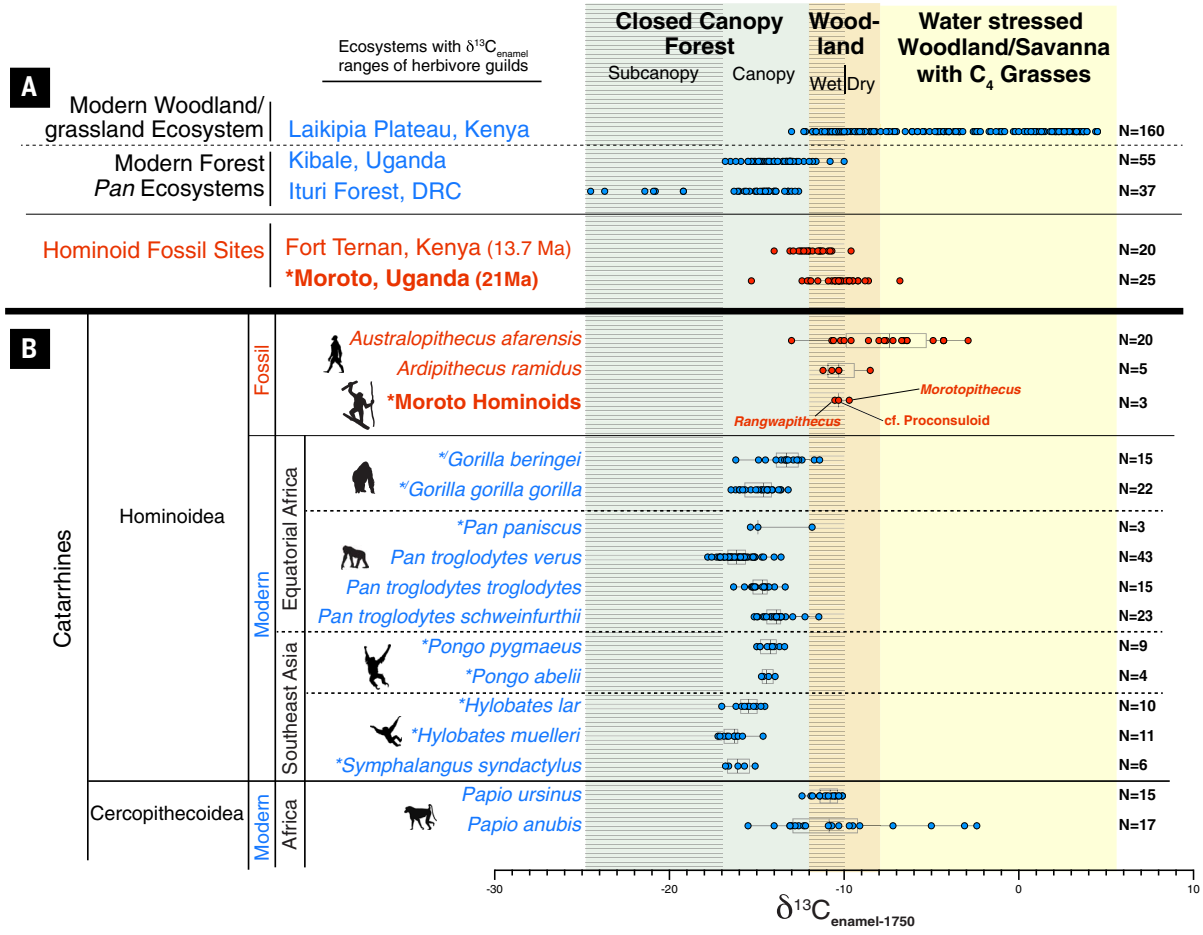


Fig. 5. Dietary $\delta^{13}C_{enamel}$ values from modern and fossil ecosystems and taxa. (**A**) An isotopic ecological framework (*52*) relating foraging strategy to habitat type applied to $\delta^{13}C_{enamel}$ dietary signatures from a range of herbivores collected from three modern ecosystems in equatorial Africa (*79*, *104*, *105*). Isotopic dietary ranges of fossil herbivores analyzed from Fort Ternan (*106*) and the Moroto sites I and II (this study) are shown. Extant taxa are shown in blue and fossil taxa in red. (**B**) Comparative isotopic dietary signals of fossil hominins (*107*, *108*), three Moroto hominoids (this study), and extant catarrhines used to constrain foraging habitats. Modern catarrhine values include gorillas [(*109*); this study, n = 21], chimpanzees [(*110*) and references therein; this study, n = 3)], orangutans (this study, n = 13), gibbons (this study, n = 27), and baboons (*109*). *, Data from this study; */, data from this study and previous research; all other data are from previous studies.

savanna microhabitats. The more positive dietary values (~>−9‰) indicate feeding on ^{13}C-enriched C_3 vegetation (typically associated with water-stressed modern habitats) and possibly including minor C_4 (or CAM) dietary components (i.e., largely arid-adapted grasses and/or succulents in the tropics). The Moroto II isotopic fossil dietary range differs from modern herbivores feeding in tropical forest and woodland and grassland ecosystems in equatorial Africa (Fig. 5A), with no clear modern analogs yet established isotopically. Isotopic variation of fossil herbivores at Moroto II is also distinguished from the more closed, forested profile established for the Middle Miocene fossil site of Fort Ternan (Fig. 5A).

Paleosol carbon isotopes

Multiple environmental indicators were also assessed on the basis of analyses of paleosols throughout the sequence (*52*), including the interval associated with the *Morotopithecus* mandible and femora (Fig. 2A). These indicators are biogeochemical proxies, primarily reflecting local relative proportions of C_3 and C_4 vegetation. The carbon isotope values of bulk organic matter ($\delta^{13}C_{OM}$; n = 16) from Moroto II range from −29.9‰ to −20.1‰ [relative to the Vienna PeeDee Belemnite (VPDB)] (Fig. 2C), and pedogenic carbonate values ($\delta^{13}C_{PC}$; n = 6) range from −7.7‰ to −2.5‰ (VPDB) (Fig. 2E). Although some paleosol samples at Moroto II are consistent with entirely C_3 vegetation, paleosols stratigraphically adjacent to the *Morotopithecus* fossil–bearing interval had the largest proportion of reconstructed C_4 vegetation, with up to 59% C_4 [95% confidence interval (CI) = 29 to 83%; $\delta^{13}C_{PC}$ = −2.5‰] based on micritic pedogenic carbonates and up to 45% (95% CI = 13 to 64%; $\delta^{13}C_{OM}$ = −20.1‰) based on paleosol organic matter (Fig. 2C and tables S9 and S10). Overall, the $\delta^{13}C_{PC}$ and $\delta^{13}C_{OM}$ datasets suggest that C_3 vegetation was interspersed with either water-stressed C_3 plants or C_4 grasses throughout the stratigraphic succession.

Paleosol n-alkanes

n-Alkanes are plant leaf wax–derived biomarkers found in paleosols that can be used to reconstruct vegetation. Biomarker concentrations and molecular distributions indicate excellent preservation of *n*-alkanes in the Moroto II paleosols and support the interpretation that isotope data reflect the Early Miocene plant community (table S11). The carbon isotope values of C_{31}*n*-alkanes range from −31.4‰ to −22.7‰ (Fig. 2D), which represent a range of C_3- to C_4-dominated ecosystems, respectively. Carbon isotope values from other plant-derived *n*-alkanes are presented in table S12. The $\delta^{13}C$ values of the C_{35}*n*-alkanes for three samples range from −24.1‰ to −20.7‰, which indicate

high abundances of C_4 vegetation on the landscape. The C_{35} homolog has been shown to be a highly sensitive indicator of C_4 vegetation in other Miocene samples from Africa (*83*, *84*) and from the Siwalik sequence in Pakistan and India (*85*). Several studies have shown that C_4 grasses, especially from Africa, produce much higher abundances of long-chain *n*-alkanes (i.e., C_{33} and C_{35}) than do C_3 plants (*86*, *87*). Samples with $\delta^{13}C$ values associated with C_4 vegetation from Moroto II also have relatively high concentrations of C_{33} and C_{35} homologs and together provide compelling isotopic and molecular evidence for C_4 grasses at Moroto II.

Seven samples from Moroto II have paired carbon isotope data for both bulk organic matter and *n*-alkanes. The $\delta^{13}C$ values of bulk organic matter for these samples are strongly and significantly correlated with the $\delta^{13}C$ values of C_{31} alkanes (r = 0.98, $P << 0.001$) and C_{35} alkanes (r = 0.94, P = 0.002), corroborating the reliability of the bulk organic carbon isotope results. Only two sampling levels at Moroto II included analyses of both pedogenic carbonate and *n*-alkanes, because the sample sizes from other levels were too small for meaningful comparisons.

Paleosols and hydroclimate

Paleosols at Moroto II typically formed on the finer-grained fluvial deposits and include gleyed Vertisols, calcic and noncalcic Inceptisols, and Alfisols (figs. S12 and S13). Qualitatively, these features are consistent with a seasonally varying hydroclimate that requires water deficit and surplus for Vertisol development and seasonal dewatering for pedogenic carbonate formation with a potential maximum MAP of ~1000 mm given the rarity of modern calcic soils above this threshold (*88*). Paleo-MAP was estimated using three proxies that use the bulk elemental composition of the uppermost subsoil (B) horizon as input data (*52*, *89*). The samples from the Moroto II sequence (Fig. 2F, figs. S13 and S14, and tables S13 to S15) yielded MAP values of 626 to 1313 mm/year, indicative of a subhumid paleoclimate within the range of modern woodlands in Africa (*90*, *91*). Further, none of the results from Moroto II supports MAP values in excess of 1500 to 2000 mm regardless of wet season length or fire regime, conditions that currently characterize closed canopy forests in equatorial Africa [e.g., (*91*)].

Phytoliths

Phytoliths are microscopic silica bodies precipitated in living plant cells and cell walls that remain in soils or sediments when plants die and decay. Preserved phytolith assemblages allow reconstruction of vegetation structure and type on a relatively local scale, especially in tropical and subtropical ecosystems (*92*–*94*). The six variably preserved phytolith assemblages from Moroto II confirm inferences from isotopic data of habitat heterogeneity, either spatially or temporally (*52*) (Fig. 2B and tables S16 to S18). The best-preserved assemblage, sampled from the *Morotopithecus* interval, is dominated by grass phytoliths (63.1% of diagnostic morphotypes) that are mainly from PACMAD grasses (C_3 or C_4 taxa in the Panicoideae, Arundinoideae, Chloridoideae, Micrairoideae, Aristidoideae, or Danthonioideae subfamilies), and includes phytoliths of the arid-adapted, C_4 Chloridoideae PACMAD subfamily. This assemblage also contains forest-indicator phytoliths, including rare evidence of palms, indicating spatial heterogeneity in the vegetation. Compositionally, this assemblage has affinities with modern phytolith assemblages from wooded grasslands, but it is also consistent with open forests containing a substantial grass understory (e.g., riparian forest in a savanna landscape, dry forest, or forest openings) (*27*).

Discussion

Context of the hominoid remains from Moroto

Moroto II records not only the earliest evidence of clearly derived hominoid postcrania and the first record of folivory for a large-bodied hominoid, but also the oldest known African ecosystem with C_4 grasses. Compilation of lithostratigraphic, pedogenic, biogeochemical, phytolith assemblage, and taphonomic data from Moroto II indicate a seasonal, mixed C_3/C_4 environment with broken canopy forest or woodland habitats. Evidence of water-stressed C_3 woodland–based diets for the catarrhines and associated fauna, as well as pedogenic proxies documenting a major C_4 component, are congruent with habitats characterized by a fragmentary canopy cover. These reconstructions provide empirical evidence supporting the concept that Early Miocene apes may have been broadly associated with heterogeneous, woodland habitats (*27*, *77*).

These data call into question previous assumptions regarding hominoid dietary and locomotor coevolution because, in combination, they fail to support prevailing terminal branch forest frugivory hypotheses that orthograde locomotor versatility evolved to facilitate feeding on fruit in a tropical forest. Instead, the earliest known appearance of enhanced orthogrady and shortened hindlimbs occurred in a seasonal woodland and wooded grassland context, where soft, ripe fruit would be periodically limited. Indeed, the morphology of dental remains reveals that leaves were an important component of the diet of at least two Moroto hominoids, *Morotopithecus* (*36*) and the nyanzapithecine aff. *Rangwapithecus*.

Positional behavioral versatility and its anatomical correlates may have evolved multiple times in large-bodied catarrhines. There is no evidence to indicate, for example, that *Morotopithecus* is closely related to the Eurasian taxa *Pierolapithecus* (~12 Ma) or *Hispanopithecus* (~10 Ma), the next-oldest hominoids inferred to have had dorsostable lumbar spines (*95*, *96*). These hominoids have been associated with forested environments (*97*, *98*), but because they are so much younger, their adaptive and phylogenetic significance reflects more the evolution of crown hominoids (i.e., that are closely related to the living hominoids) (*99*) than hominoid origins in Africa.

In the absence of additional fossils such as hand bones, it is not possible to judge whether suspension (*10*, *11*), quadrumanous climbing (*12*), and vertical climbing (*13*, *14*) were more or less associated with hominoid adaptive origins (*39*). However, the record from Moroto II supports the hypothesis that orthogrady is key to hominoid distinctiveness (*6*, *15*) because it is among the first identifiable steps in the evolution of anatomical features associated with the modern-type hominoid adaptive niche.

Furthermore, in this first recorded instance, locomotor and postural innovation seem to have evolved in hominoids for effective arboreal feeding on leaves and fruit situated in higher and/or peripheral aspects of tree crowns in seasonal woodland habitats characterized by gaps in the canopy. Such environments may exert hitherto unrecognized selective pressures on the evolution of arboreal apes, such as the need, in times of seasonal low fruit availability, to access young leaves with relatively high protein (*26*) in the terminal or higher aspects of the canopy. Although some breaks in the canopy could be bridged through enhanced climbing adaptations, larger gaps would have necessitated repeated descent and ascent of tree trunks and/or lianas, which is consistent with the locomotor behavior indicated by the femoral and vertebral morphology of the hominoid fossils of Moroto II.

Hominoid adaptive evolution in eastern Africa

Documenting the early stages of hominoid evolution as the lineage diverged from other catarrhines is critical for interpreting the adaptive significance of traits that ultimately define modern hominoids, including humans. Although high catarrhine taxonomic and phylogenetic diversity in the Early Miocene have been recognized for some time (*42*), a schema is now emerging of high adaptive diversity among these primates (*100*), as well as high environmental variability (*27*, *73*, *78*), a perspective reinforced by reconstructions of hominoid behavior and ecology at Moroto II. These findings suggest that early hominoid morphologies and environments do not match classic ideas about the origins of the group, revealing a complexity yet to be studied and understood. The traditional paradigm for interpreting the origin of the hominoid lineage has involved forested habitats, but the data described herein

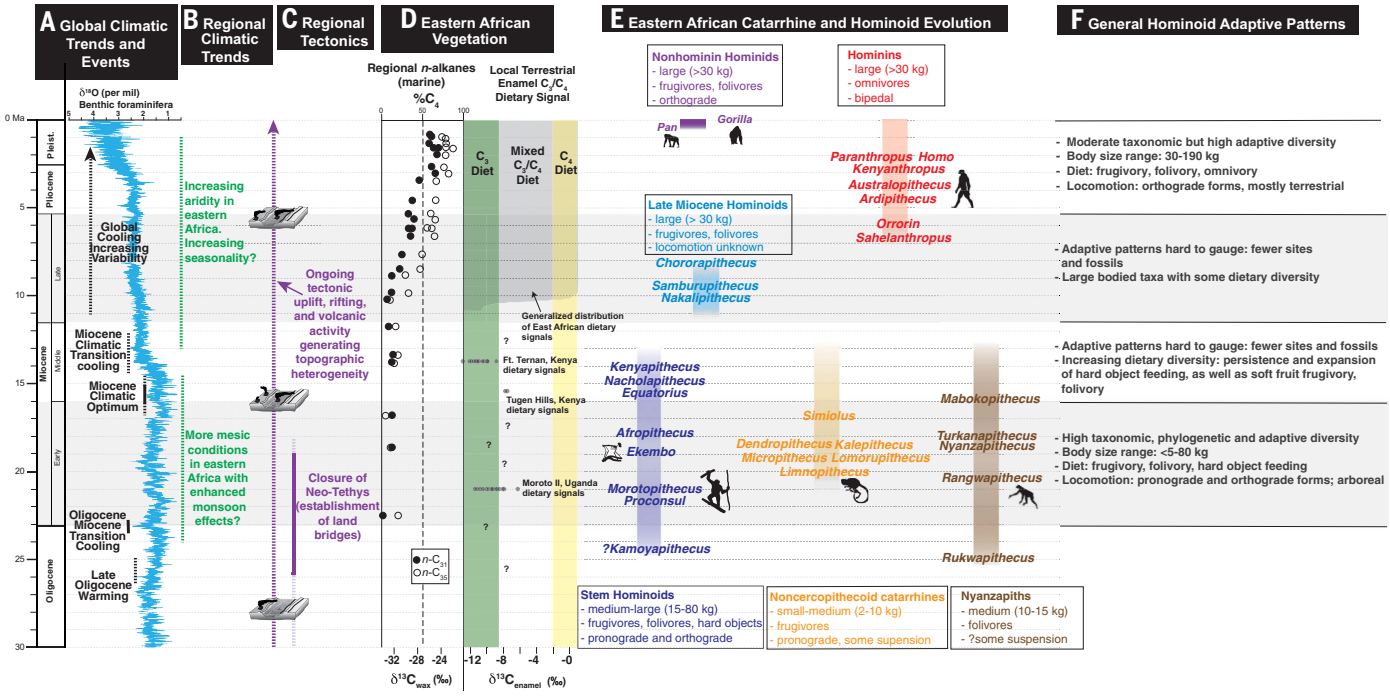


Fig. 6. Environments and catarrhine adaptive evolution. (**A**) Global climatic trends and events (*111*–*115*). Benthic foraminifera $\delta^{18}O$ are shown in blue (*115*). (**B**) Regional climatic trends (seasonality and variability) [e.g., (*116*, *117*)]. (**C**) Regional tectonic events and trends in eastern Africa (*118*, *119*). (**D**) Regional *n*-alkane record of terrestrial vegetation (*83*, *120*) and eastern Africa fossil dietary record (this study) (*121*, *122*). The gray area depicts reported herbivore dietary isotopic signals that include a range of C_3 and C_4 dietary components and an apparent increase in C_4 dietary biomass after ~10 Ma ago (*84*, *100*, *120*). (**E**) African catarrhines grouped using biogeography, chronology, and adaptive inferences, not strict phylogeny. The group "stem hominoids" is well supported (*99*), but including *Kamoyapithecus*, although plausible (*123*), is not typical [e.g., (*71*)] and is informed by size and inferred adaptive similarity (*8*). Nyanzapithecinae (e.g., *Rangwapithecus* and *Rukwapithecus*) are folivorous and considered a monophyletic group (*124*). "Noncercopithecoid catarrhines" are poorly known and [with possible exceptions (*125*)] unlikely to be hominoids (*42*). Late Miocene hominoids, related by biogeography and age, remain enigmatic (*42*, *126*–*129*). *Pan* and *Gorilla* share adaptive similarity (*130*), but *Pan* is more related to hominins. Hominin adaptations, including canine reduction (*131*) and bipedality (*108*, *132*), are often linked to habitat heterogeneity. (**F**) Chronologically informed adaptive patterns.

reveal that in the Early Miocene, hominoid adaptive diversity may have been influenced by environmental heterogeneity (Fig. 6), as has already been documented in the Late Miocene/Pliocene and invoked in scenarios of hominin origins (*101*, *102*). This requires more careful consideration of the nature of habitat heterogeneity in the past and a reevaluation of how it may be linked to features (e.g., bipedality and dental and foraging specializations) associated with early hominin expansion into more open microhabitats within diverse ecosystems.

Materials and methods summary

Material and methods for all techniques are available in the supplementary materials. This includes radiometric, paleomagnetic, and biostratigraphic dating methods; morphological, taphonomic, and faunal analyses; paleosol biogeochemical approaches (carbon isotopic analysis from both bulk organic and carbonate samples, *n*-alkane analysis, and paleosol-based hydroclimate analysis); and paleoecological analyses including isotopic dietary paleoecology and phytolith analyses.

REFERENCES AND NOTES

1. V. Avis, Brachiation: The crucial issue for man's ancestry. *Southwest. J. Anthropol.* **18**, 119–148 (1962). doi: 10.1086/soutjanth.18.2.3629013
2. J. R. Napier, Evolutionary aspects of primate locomotion. *Am. J. Phys. Anthropol.* **27**, 333–341 (1967). doi: 10.1002/ajpa.1330270306; pmid: 4968059
3. L. A. Temerin, J. G. Cant, The evolutionary divergence of Old World monkeys and apes. *Am. Nat.* **122**, 335–351 (1983). doi: 10.1086/284139
4. K. D. Hunt, Why are there apes? Evidence for the co-evolution of ape and monkey ecomorphology. *J. Anat.* **228**, 630–685 (2016). doi: 10.1111/joa.12454; pmid: 27004976
5. K. D. Hunt *et al.*, Standardized descriptions of primate locomotor and postural modes. *Primates* **37**, 363–387 (1996). doi: 10.1007/BF02381373
6. S. K. S. Thorpe, R. H. Crompton, Orangutan positional behavior and the nature of arboreal locomotion in Hominoidea. *Am. J. Phys. Anthropol.* **131**, 384–401 (2006). doi: 10.1002/ajpa.20422; pmid: 16617429
7. L. A. Sarringhaus, L. M. MacLatchy, J. C. Mitani, Locomotor and postural development of wild chimpanzees. *J. Hum. Evol.* **66**, 29–38 (2014). doi: 10.1016/j.jhevol.2013.09.006; pmid: 24238359
8. L. M. MacLatchy, W. J. Sanders, C. L. Wuthrich, Hominoid origins. *Nat. Educ. Knowl.* **6**, 1–13 (2015).
9. T. Harrison, "The implications of *Oreopithecus bambolii* for the origins of bipedalism," in *Origine (s) de la bipédie chez les hominidés*, Y. Coppens, B. Senut, Eds. (Editions du Centre National de la Recherche Scientifique, 1991), vol. 235, pp. 235–244.
10. S. Cachel, A paleoecological model for the origin of higher primates. *J. Hum. Evol.* **8**, 351–359 (1979). doi: 10.1016/0047-2484(79)90058-7
11. J. G. H. Cant, Positional behavior of female bornean orangutans (*Pongo pygmaeus*). *Am. J. Primatol.* **12**, 71–90 (1987). doi: 10.1002/ajp.1350120104; pmid: 31973515
12. J. G. Fleagle, Locomotion and posture of the Malayan siamang and implications for hominoid evolution. *Folia Primatol. (Basel)* **26**, 245–269 (1976). doi: 10.1159/000155756; pmid: 1010498
13. J. T. Stern, R. L. Susman, Electromyography of the gluteal muscles in *Hylobates*, *Pongo*, and *Pan*: Implications for the evolution of hominid bipedality. *Am. J. Phys. Anthropol.* **55**, 153–166 (1981). doi: 10.1002/ajpa.1330550203
14. K. Isler, 3D-kinematics of vertical climbing in hominoids. *Am. J. Phys. Anthropol.* **126**, 66–81 (2005). doi: 10.1002/ajpa.10419; pmid: 15386239
15. M. G. Nowak, U. H. Reichard, "Locomotion and posture in ancestral hominoids prior to the split of hylobatids," in *Evolution of Gibbons and Siamang*, U. H. Reichard, H. Hirai, C. Barelli, Eds. (Springer, 2016), pp. 55–89.
16. M. Cartmill, K. Milton, The lorisiform wrist joint and the evolution of "brachiating" adaptations in the hominoidea. *Am. J. Phys. Anthropol.* **47**, 249–272 (1977). doi: 10.1002/ajpa.1330470206; pmid: 410307
17. R. Tuttle, "Parallelism, brachiation, and hominoid phylogeny" in *Phylogeny of the Primates: A Multidisciplinary Approach*, W. P. Luckett, F. S. Szalay, Eds. (Springer, 1975), pp. 447–480. doi: 10.1007/978-1-4684-2166-8_17
18. S. Ripley, "Environmental grain, niche diversification, and positional behavior in Neogene primates : An evolutionary hypothesis" in *Environment, Behavior, and Morphology: Dynamic Interactions in Primates*, M. E. Morbeck, H. Preuschoft, N. Gomberg, Eds. (Gustav Fischer, 1979), pp. 371–412; ; https://ci.nii.ac.jp/naid/10016066823/.

19. P. Andrews, L. Aiello, "An evolutionary model for feeding and positional behaviour," in *Food Acquisition and Processing in Primates*, D. J. Chivers, B. A. Wood, A. Bilsborough, Eds. (Springer, 1984), pp. 429–466; http://link.springer.com/10.1007/978-1-4757-5244-1_20.
20. F. E. Putz, K. H. Redford, The importance of defining 'forest': Tropical forest degradation, deforestation, long-term phase shifts, and further transitions. *Biotropica* **42**, 10–20 (2010). doi: 10.1111/j.1744-7429.2009.00567.x
21. C. McElhinny, P. Gibbons, C. Brack, J. Bauhus, Forest and woodland stand structural complexity: Its definition and measurement. *For. Ecol. Manage.* **218**, 1–24 (2005). doi: 10.1016/j.foreco.2005.08.034
22. P. Andrews, Evolution and environment in the Hominoidea. *Nature* **360**, 641–646 (1992). doi: 10.1038/360641a0; pmid: 1465127
23. E. E. Sarmiento, Cautious climbing and folivory: A model of hominoid differentiation. *Hum. Evol.* **10**, 289–321 (1995). doi: 10.1007/BF02438967
24. A. Houle, C. A. Chapman, W. L. Vickery, Intratree variation in fruit production and implications for primate foraging. *Int. J. Primatol.* **28**, 1197–1217 (2007). doi: 10.1007/s10764-007-9214-9
25. N. J. Dominy, P. W. Lucas, S. J. Wright, Mechanics and chemistry of rain forest leaves: Canopy and understorey compared. *J. Exp. Bot.* **54**, 2007–2014 (2003). doi: 10.1093/jxb/erg224; pmid: 12867549
26. C. A. Chapman, L. J. Chapman, L. Naughton-Treves, M. J. Lawes, L. R. McDowell, Predicting folivorous primate abundance: Validation of a nutritional model. *Am. J. Primatol.* **62**, 55–69 (2004). doi: 10.1002/ajp.20006; pmid: 14983464
27. D. J. Peppe *et al*., Earliest evidence for C4 grass-dominated habitats and vegetation heterogeneity in eastern Africa. *Science* **380**, XXXX (2023). doi: 10.1126/science.abq2835
28. W. W. Bishop, F. Whyte, Tertiary mammalian faunas and sediments in Karamoja and Kavirondo, East Africa. *Nature* **196**, 1283–1287 (1962). doi: 10.1038/1961283a0
29. D. Allbrook, W. W. Bishop, New fossil hominoid material from Uganda. *Nature* **197**, 1187–1190 (1963). doi: 10.1038/1971187a0
30. W. W. Bishop, More fossil primates and other miocene mammals from north-east Uganda. *Nature* **203**, 1327–1331 (1964). doi: 10.1038/2031327a0
31. W. W. Bishop, J. A. Miller, F. J. Fitch, New potassium-argon age determinations relevant to Miocene fossil mammal sequence in East Africa. *Am. J. Sci.* **267**, 669–699 (1969). doi: 10.2475/ajs.267.6.669
32. D. L. Gebo *et al*., A hominoid genus from the early Miocene of Uganda. *Science* **276**, 401–404 (1997). doi: 10.1126/science.276.5311.401; pmid: 9103195
33. M. Pickford, B. Senut, D. Gommery, E. Musiime, Distinctiveness of *Ugandapithecus* from *Proconsul*. *Estud. Geol. (Madr.)* **65**, 183–241 (2009). doi: 10.3989/egeol.39926.071
34. M. Pickford, P. Mein, Early Middle Miocene mammals from Moroto II, Uganda. *Beitrage Zur Palaontologie* **30**, 361–386 (2006).
35. M. Pickford, B. Senut, D. Gommery, E. Musiime, New Catarrhine fossils from Moroto II, Early Middle Miocene (ca 17.5 Ma) Uganda. *C. R. Palevol* **2**, 649–662 (2003). doi: 10.1016/j.crpv.2003.10.001
36. L. MacLatchy, J. Rossie, A. Houssaye, A. J. Olejniczak, T. M. Smith, New hominoid fossils from Moroto II, Uganda and their bearing on the taxonomic and adaptive status of *Morotopithecus bishopi*. *J. Hum. Evol.* **132**, 227–246 (2019). doi: 10.1016/j.jhevol.2019.03.008; pmid: 31203849
37. C. V. Ward, Torso morphology and locomotion in *Proconsul nyanzae*. *Am. J. Phys. Anthropol.* **92**, 291–328 (1993). doi: 10.1002/ajpa.1330920306; pmid: 8291620
38. C. V. Ward, "Postcranial and Locomotor Adaptations of Hominoids" in *Handbook of Paleoanthropology*, W. Henke, I. Tattersall, Eds. (Springer, 2015), pp. 1363–1386; http://link.springer.com/10.1007/978-3-642-39979-4_34.
39. L. MacLatchy, The oldest ape. *Evol. Anthropol.* **13**, 90–103 (2004). doi: 10.1002/evan.10133
40. M. Pickford, New reconstruction of the Moroto hominoid snout and a reassessment of its affinities to *Afropithecus turkanensis*. *Hum. Evol.* **17**, 1–19 (2002). doi: 10.1007/BF02436425
41. B. A. Patel, A. Grossman, Dental metric comparisons of *Morotopithecus* and *Afropithecus*: Implications for the validity of the genus *Morotopithecus*. *J. Hum. Evol.* **51**, 506–512 (2006). doi: 10.1016/j.jhevol.2006.07.002; pmid: 16914180
42. T. Harrison, "Dendropithecoidea, Proconsuloidea, and Hominoidea" in *Cenozoic Mammals of Africa*, L. Werdelin, W. J. Sanders, Eds. (Univ. of California Press, 2010), pp. 429–470.
43. M. Pickford, B. Senut, D. Gommery, S. Musalizi, E. Musiime, Revision of the Miocene Hominoidea from Moroto I and II, Uganda. *Geo-Pal Uganda* (2017).
44. D. R. Pilbeam, Tertiary Pongidae of East Africa: Evolutionary relationships and taxonomy. *Bull. Peabody Mus. Nat. Hist.* **31**, 1–185 (1969).
45. R. F. Kay, P. S. Ungar, "Dental evidence for diet in some Miocene Catarrhines with comments on the effects of phylogeny on the interpretation of adaptation," in *Function, Phylogeny, and Fossils*, D. R. Begun, C. V. Ward, M. D. Rose, Eds. (Springer, 1997), pp. 131–151; ; http://link.springer.com/10.1007/978-1-4899-0075-3_7.
46. P. Ungar, M. Teaford, R. Kay, Molar microwear and shearing crest development in Miocene catarrhines. *Anthropologie* **42**, 21–35 (1962).
47. M. Leakey, A. Walker, "Afropithecus" in *Function, Phylogeny, and Fossils: Miocene Hominoid Evolution and Adaptations*, D. R. Begun, C. V. Ward, M. D. Rose, Eds. (Springer, 1997), *Advances in Primatology* series, pp. 225–239; doi: 10.1007/978-1-4899-0075-3_11
48. A. Walker, M. D. Rose, Fossil hominoid vertebra from the Miocene of Uganda. *Nature* **217**, 980–981 (1968). doi: 10.1038/217980a0; pmid: 5642859
49. W. J. Sanders, B. E. Bodenbender, Morphometric analysis of lumbar vertebra UMP-67-28: Implications for spinal function and phylogeny of the Miocene Moroto hominoid. *J. Hum. Evol.* **26**, 203–237 (1994). doi: 10.1006/jhev.1994.1012
50. L. MacLatchy, D. Gebo, R. Kityo, D. Pilbeam, Postcranial functional morphology of *Morotopithecus bishopi*, with implications for the evolution of modern ape locomotion. *J. Hum. Evol.* **39**, 159–183 (2000). doi: 10.1006/jhev.2000.0407; pmid: 10968927
51. C. B. Ruff, Long bone articular and diaphyseal structure in Old World monkeys and apes. II: Estimation of body mass. *Am. J. Phys. Anthropol.* **120**, 16–37 (2003). doi: 10.1002/ajpa.10118; pmid: 12489135
52. Materials and methods are available as supplementary materials.
53. A. K. Behrensmeyer, Taphonomic and ecologic information from bone weathering. *Paleobiology* **4**, 150–162 (1978). doi: 10.1017/S0094837300005820
54. A. Aslan, A. K. Behrensmeyer, Taphonomy and time resolution of bone assemblages in a contemporary fluvial system: The East Fork River, Wyoming. *Palaios* **11**, 411 (1996). doi: 10.2307/3515209
55. D. Western, A. K. Behrensmeyer, Bone assemblages track animal community structure over 40 years in an African savanna ecosystem. *Science* **324**, 1061–1064 (2009). doi: 10.1126/science.1171155; pmid: 19461002
56. J. G. Ogg, "Geomagnetic polarity time scale," in *Geologic Time Scale 2020*, F. M. Gradstein, J. G. Ogg, M. D. Schmitz, G. M. Ogg, Eds. (Elsevier, 2020), pp. 159–192; https://www.sciencedirect.com/science/article/pii/B978012824360200005X.
57. M. Pickford, Preliminary Miocene mammalian biostratigraphy for Western Kenya. *J. Hum. Evol.* **10**, 73–97 (1981). doi: 10.1016/S0047-2484(81)80026-7
58. S. Cote *et al*., Evidence for rapid faunal change in the early Miocene of East Africa based on revised biostratigraphic and radiometric dating of Bukwa, Uganda. *J. Hum. Evol.* **116**, 95–107 (2018). doi: 10.1016/j.jhevol.2017.12.001; pmid: 29477184
59. S. M. Cote, thesis, Harvard University (2008).
60. M. Pickford, B. Senut, D. Hadoto, J. Musisi, C. Kariira, Recent discoveries In the Miocene sites At Moroto, Northeast Uganda-biostratigraphical and paleoecological implications. *Comptes Rendus Acad. Sci. Ser. II* **302**, 681 (1986).
61. R. E. Drake, J. A. Van Couvering, M. H. Pickford, G. H. Curtis, J. A. Harris, New chronology for the Early Miocene mammalian faunas of Kisingiri, Western Kenya. *J. Geol. Soc. London* **145**, 479–491 (1988). doi: 10.1144/gsjgs.145.3.0479
62. W. Sanders, E. Gheerbrant, J. Harris, H. Saegusa, C. Delmer, "Proboscidea," in *Cenozoic Mammals of Africa*, L. Werdelin, W. J. Sanders, Eds. (Univ. of California Press, 2010), pp. 161–252.
63. L. Werdelin, "Chronology of Neogene mammal localities," in *Cenozoic Mammals of Africa*, L. Werdelin, W. J. Sanders, Eds. (Univ. of California Press, 2010), pp. 27–43; http://urn.kb.se/resolve?urn=urn:nbn:se:nrm:diva-89.
64. C. B. Ruff, J. A. Runestad, Primate limb bone structural adaptations. *Annu. Rev. Anthropol.* **21**, 407–433 (1992). doi: 10.1146/annurev.an.21.100192.002203
65. W. L. Jungers, Hindlimb and pelvic adaptations to vertical climbing and clinging in *Megaladapis*, a giant subfossil prosimian from Madagascar. *Yearb. Phys. Anthropol.* **20**, 508–524 (1977).
66. J. G. Fleagle, Locomotor behavior and skeletal anatomy of sympatric Malaysian leaf-monkeys (*Presbytis obscura* and *Presbytis melalophos*). *Yearb. Phys. Anthropol.* **20**, 440–453 (1977).
67. C. Tardieu, L'articulation du genou des Primates catarhiniens et hominidés fossiles. Implications phylogénétique et taxinomique. *Bull. Mem. Soc. Anthropol. Paris* **10**, 355–372 (1983). doi: 10.3406/bmsap.1983.3908
68. C. B. Ruff, Long bone articular and diaphyseal structure in old world monkeys and apes. I: Locomotor effects. *Am. J. Phys. Anthropol.* **119**, 305–342 (2002). doi: 10.1002/ajpa.10117; pmid: 12448016
69. L. J. Shapiro, "Functional morphology of the vertebral column in primates," in *Postcranial Adaptation in Non-Human Primates*, D. Gebo, Ed. (Northern Illinois Univ. Press, 1993), pp. 121–149.
70. S. Cote, N. Malit, I. Nengo, Additional mandibles of *Rangwapithecus gordoni*, an early Miocene catarrhine from the Tinderet localities of Western Kenya. *Am. J. Phys. Anthropol.* **153**, 341–352 (2014). doi: 10.1002/ajpa.22433; pmid: 24288165
71. N. J. Stevens *et al*., Palaeontological evidence for an Oligocene divergence between Old World monkeys and apes. *Nature* **497**, 611–614 (2013). doi: 10.1038/nature12161; pmid: 23676680
72. J. Krigbaum, M. H. Berger, D. J. Daegling, W. S. McGraw, Stable isotope canopy effects for sympatric monkeys at Tai Forest, Cote d'Ivoire. *Biol. Lett.* **9**, 20130466 (2013). doi: 10.1098/rsbl.2013.0466; pmid: 23843220
73. L. A. Michel *et al*., Remnants of an ancient forest provide ecological context for Early Miocene fossil apes. *Nat. Commun.* **5**, 3236 (2014). doi: 10.1038/ncomms4236; pmid: 24549336
74. P. Andrews, J. A. Van Couvering, "Palaeoenvironments in the East African Miocene," in *Approaches to Primate Paleobiology*, F. S. Szalay, Ed. (Karger, 1975), vol. 5, pp. 62–103.
75. E. N. Evans, J. A. Van Couvering, P. Andrews, Palaeoecology of Miocene sites in western Kenya. *J. Hum. Evol.* **10**, 99–116 (1981). doi: 10.1016/S0047-2484(81)80027-9
76. A. Walker, Fossil mammal locality on Mount Elgon, Eastern Uganda: Lower Miocene fossils from Mount Elgon Uganda. *Nature* **223**, 591–593 (1969). doi: 10.1038/223591a0
77. P. Andrews, *An Ape's View of Human Evolution* (Cambridge Univ. Press, 2015).
78. W. E. Lukens *et al*., The Early Miocene critical zone at Karungu, Western Kenya: An equatorial, open habitat with few primate remains. *Front. Earth Sci. (Lausanne)* **5**, 87 (2017). doi: 10.3389/feart.2017.00087
79. T. E. Cerling *et al*., Dietary changes of large herbivores in the Turkana Basin, Kenya from 4 to 1 Ma. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 11467–11472 (2015). doi: 10.1073/pnas.1513075112; pmid: 26240344
80. J. D. Kingston, T. Harrison, Isotopic dietary reconstructions of Pliocene herbivores at Laetoli: Implications for early hominin paleoecology. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **243**, 272–306 (2007). doi: 10.1016/j.palaeo.2006.08.002
81. H. V. Graham *et al*., Isotopic characteristics of canopies in simulated leaf assemblages. *Geochim. Cosmochim. Acta* **144**, 82–95 (2014). doi: 10.1016/j.gca.2014.08.032
82. G. D. Farquhar, J. R. Ehleringer, K. T. Hubick, Carbon isotope discrimination and photosynthesis. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **40**, 503–537 (1989). doi: 10.1146/annurev.pp.40.060189.002443
83. P. J. Polissar, C. Rose, K. T. Uno, S. R. Phelps, P. deMenocal, Synchronous rise of African C4 ecosystems 10 million years ago in the absence of aridification. *Nat. Geosci.* **12**, 657–660 (2019). doi: 10.1038/s41561-019-0399-2
84. K. T. Uno, P. J. Polissar, K. E. Jackson, P. B. deMenocal, Neogene biomarker record of vegetation change in eastern Africa. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 6355–6363 (2016). doi: 10.1073/pnas.1521267113; pmid: 27274042
85. P. J. Polissar *et al*., Hydrologic changes drove the Late Miocene expansion of C4 grasslands on the Northern Indian Subcontinent. *Paleoceanogr. Paleoclimatol.* **36**, e2020PA004108 (2021). doi: 10.1029/2020PA004108

86. R. T. Bush, F. A. McInerney, Leaf wax *n*-alkane distributions in and across modern plants: Implications for paleoecology and chemotaxonomy. *Geochim. Cosmochim. Acta* **117**, 161–179 (2013). doi: 10.1016/j.gca.2013.04.016
87. Y. Garcin *et al*., Reconstructing C3 and C4 vegetation cover using n-alkane carbon isotope ratios in recent lake sediments from Cameroon, Western Central Africa. *Geochim. Cosmochim. Acta* **142**, 482–500 (2014). doi: 10.1016/j.gca.2014.07.004
88. K. Zamanian, K. Pustovoytov, Y. Kuzyakov, Pedogenic carbonates: Forms and formation processes. *Earth Sci. Rev.* **157**, 1–17 (2016). doi: 10.1016/j.earscirev.2016.03.003
89. W. E. Lukens *et al*., Recursive partitioning improves paleosol proxies for rainfall. *Am. J. Sci.* **319**, 819–845 (2019). doi: 10.2475/10.2019.01
90. F. White, Long distance dispersal, overland migration and extinction in the shaping of tropical African floras. *Bothalia* **14**, 395–403 (1983). doi: 10.4102/abc.v14i2.1184
91. K. Guan *et al*., Terrestrial hydrological controls on land surface phenology of African savannas and woodlands. *J. Geophys. Res. Biogeosci.* **119**, 1652–1669 (2014). doi: 10.1002/2013JG002572
92. D. Barboni, L. Bremond, R. Bonnefille, Comparative study of modern phytolith assemblages from inter-tropical Africa. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **246**, 454–470 (2007). doi: 10.1016/j.palaeo.2006.10.012
93. A. Novello, thesis, University of Poitiers (2012).
94. K. Neumann, A. G. Fahmy, N. Müller-Scheeßel, M. Schmidt, Taxonomic, ecological and palaeoecological significance of leaf phytoliths in West African grasses. *Quat. Int.* **434**, 15–32 (2017). doi: 10.1016/j.quaint.2015.11.039
95. S. Moyà-Solà, M. Köhler, A *Dryopithecus* skeleton and the origins of great-ape locomotion. *Nature* **379**, 156–159 (1996). doi: 10.1038/379156a0; pmid: 8538764
96. S. Moyà-Solà, M. Köhler, D. M. Alba, I. Casanovas-Vilar, J. Galindo, *Pierolapithecus catalaunicus*, a new Middle Miocene great ape from Spain. *Science* **306**, 1339–1344 (2004). doi: 10.1126/science.1103094; pmid: 15550663
97. I. Casanovas-Vilar *et al*., Biochronological, taphonomical, and paleoenvironmental background of the fossil great ape *Pierolapithecus catalaunicus* (Primates, Hominidae). *J. Hum. Evol.* **55**, 589–603 (2008). doi: 10.1016/j.jhevol.2008.05.004; pmid: 18691737
98. J. Marmi, I. Casanovas-Vilar, J. M. Robles, S. Moyà-Solà, D. M. Alba, The paleoenvironment of *Hispanopithecus laietanus* as revealed by paleobotanical evidence from the Late Miocene of Can Llobateres 1 (Catalonia, Spain). *J. Hum. Evol.* **62**, 412–423 (2012). doi: 10.1016/j.jhevol.2011.12.003; pmid: 22265456
99. S. Almécija *et al*., Fossil apes and human evolution. *Science* **372**, eabb4363 (2021). doi: 10.1126/science.abb4363; pmid: 33958446
100. C. Wuthrich, L. M. MacLatchy, I. O. Nengo, Wrist morphology reveals substantial locomotor diversity among early catarrhines: An analysis of capitates from the early Miocene of Tinderet (Kenya). *Sci. Rep.* **9**, 3728 (2019). doi: 10.1038/s41598-019-39800-3; pmid: 30842461
101. J. D. Kingston, Shifting adaptive landscapes: Progress and challenges in reconstructing early hominid environments. *Am. J. Phys. Anthropol.* **134** (Suppl 45), 20–58 (2007). doi: 10.1002/ajpa.20733; pmid: 18046753
102. T. E. Cerling *et al*., Woody cover and hominin environments in the past 6 million years. *Nature* **476**, 51–56 (2011). doi: 10.1038/nature10306; pmid: 21814275
103. R. J. Smith, W. L. Jungers, Body mass in comparative primatology. *J. Hum. Evol.* **32**, 523–559 (1997). doi: 10.1006/jhev.1996.0122; pmid: 9210017
104. S. V. Nelson, Chimpanzee fauna isotopes provide new interpretations of fossil ape and hominin ecologies. *Proc. Biol. Sci.* **280**, 20132324 (2013). doi: 10.1098/rspb.2013.2324; pmid: 24197413
105. T. E. Cerling, J. A. Hart, T. B. Hart, Stable isotope ecology in the Ituri Forest. *Oecologia* **138**, 5–12 (2004). doi: 10.1007/s00442-003-1375-4; pmid: 14530961
106. T. E. Cerling, J. M. Harris, S. H. Ambrose, M. G. Leakey, N. Solounias, Dietary and environmental reconstruction with stable isotope analyses of herbivore tooth enamel from the Miocene locality of Fort Ternan, Kenya. *J. Hum. Evol.* **33**, 635–650 (1997). doi: 10.1006/jhev.1997.0151; pmid: 9467773
107. T. D. White *et al*., *Ardipithecus ramidus* and the paleobiology of early hominids. *Science* **326**, 75–86 (2009). doi: 10.1126/science.1175802; pmid: 19810190
108. J. G. Wynn *et al*., Diet of *Australopithecus afarensis* from the Pliocene Hadar Formation, Ethiopia. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 10495–10500 (2013). doi: 10.1073/pnas.1222559110; pmid: 23733965
109. N. E. Levin, Y. Haile-Selassie, S. R. Frost, B. Z. Saylor, Dietary change among hominins and cercopithecids in Ethiopia during the early Pliocene. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 12304–12309 (2015). doi: 10.1073/pnas.1424982112; pmid: 26371308
110. M. A. Malone, L. M. MacLatchy, J. C. Mitani, R. Kityo, J. D. Kingston, A chimpanzee enamel-diet $\delta^{13}C$ enrichment factor and a refined enamel sampling strategy: Implications for dietary reconstructions. *J. Hum. Evol.* **159**, 103062 (2021). doi: 10.1016/j.jhevol.2021.103062; pmid: 34536662
111. H. M. Beddow, D. Liebrand, A. Sluijs, B. S. Wade, L. J. Lourens, Global change across the Oligocene-Miocene transition: High-resolution stable isotope records from IODP Site U1334 (equatorial Pacific Ocean). *Paleoceanography* **31**, 81–97 (2016). doi: 10.1002/2015PA002820
112. K. Hyeong *et al*., Southward shift of the Intertropical Convergence Zone due to Northern Hemisphere cooling at the Oligocene-Miocene boundary. *Geology* **42**, 667–670 (2014). doi: 10.1130/G35664.1
113. M. Harzhauser *et al*., Biogeographic responses to geodynamics: A key study all around the Oligo–Miocene Tethyan Seaway. *Zool. Anz.* **246**, 241–256 (2007). doi: 10.1016/j.jcz.2007.05.001
114. J. R. Super *et al*., North Atlantic temperature and pCO2 coupling in the early-middle Miocene. *Geology* **46**, 519–522 (2018). doi: 10.1130/G40228.1
115. D. De Vleeschouwer, M. Vahlenkamp, M. Crucifix, H. Pälike, Alternating Southern and Northern Hemisphere climate response to astronomical forcing during the past 35 m.y. *Geology* **45**, 375–378 (2017). doi: 10.1130/G38663.1
116. T. L. P. Couvreur *et al*., Tectonics, climate and the diversification of the tropical African terrestrial flora and fauna. *Biol. Rev. Camb. Philos. Soc.* **96**, 16–51 (2021). doi: 10.1111/brv.12644; pmid: 32924323
117. B. Jacobs, A. Pan, C. Scotese, "A review of the Cenozoic vegetation history of Africa," in *Cenozoic Mammals of Africa*, L. Werdelin, W. Sanders, Eds. (Univ. of California Press, 2010), pp. 57–72.
118. D. Macgregor, History of the development of the East African Rift System: A series of interpreted maps through time. *J. Afr. Earth Sci.* **101**, 232–252 (2015). doi: 10.1016/j.jafrearsci.2014.09.016
119. T. O. Rooney, The Cenozoic magmatism of East-Africa: Part I — Flood basalts and pulsed magmatism. *Lithos* **286–287**, 264–301 (2017). doi: 10.1016/j.lithos.2017.05.014
120. K. T. Uno *et al*., Late Miocene to Pliocene carbon isotope record of differential diet change among East African herbivores. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 6509–6514 (2011). doi: 10.1073/pnas.1018435108; pmid: 21464327
121. M. E. Morgan, J. D. Kingston, B. D. Marino, Carbon isotopic evidence for the emergence of C4 plants in the Neogene from Pakistan and Kenya. *Nature* **367**, 162–165 (1994). doi: 10.1038/367162a0
122. T. Cerling *et al*., Global vegetation change through the Miocene/Pliocene boundary. *Nature* **389**, 153–158 (1997). doi: 10.1038/38229
123. M. G. Leakey, P. S. Ungar, A. Walker, A new genus of large primate from the Late Oligocene of Lothidok, Turkana District, Kenya. *J. Hum. Evol.* **28**, 519–531 (1995). doi: 10.1006/jhev.1995.1040
124. J. B. Rossie, L. MacLatchy, A new pliopithecoid genus from the early Miocene of Uganda. *J. Hum. Evol.* **50**, 568–586 (2006). doi: 10.1016/j.jhevol.2005.12.007; pmid: 16472841
125. J. B. Rossie, A. Hill, A new species of *Simiolus* from the middle Miocene of the Tugen Hills, Kenya. *J. Hum. Evol.* **125**, 50–58 (2018). doi: 10.1016/j.jhevol.2018.09.002; pmid: 30502897
126. G. Suwa, R. T. Kono, S. Katoh, B. Asfaw, Y. Beyene, A new species of great ape from the late Miocene epoch in Ethiopia. *Nature* **448**, 921–924 (2007). doi: 10.1038/nature06113; pmid: 17713533
127. Y. Kunimatsu *et al*., A new Late Miocene great ape from Kenya and its implications for the origins of African great apes and humans. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 19220–19225 (2007). doi: 10.1073/pnas.0706190104; pmid: 18024593
128. H. Ishida, M. Pickford, A new Late Miocene hominoid from Kenya: *Samburupithecus kiptalami* gen. et sp. nov. *Comptes Rendus Académie Sci. - Ser. IIA - Earth. Planet. Sci.* **325**, 823–829 (1997).
129. S. Katoh *et al*., New geological and palaeontological age constraint for the gorilla-human lineage split. *Nature* **530**, 215–218 (2016). doi: 10.1038/nature16510; pmid: 26863981
130. D. P. Watts, "The apes: taxonomy, biogeography, life histories, and behavioral ecology," in *The Evolution of Primate Societies*, J. B. Silk, J. C. Mitani, J. Call, P. M. Kapppeler, R. A. Palombit, Eds. (Univ. of Chicago Press, 2012), pp. 113–142.
131. M. Brunet *et al*., A new hominid from the Upper Miocene of Chad, Central Africa. *Nature* **418**, 145–151 (2002). doi: 10.1038/nature00879; pmid: 12110880
132. M. Pickford, B. Senut, D. Gommery, J. Treil, Bipedalism in *Orrorin tugenensis* revealed by its femora. *C. R. Palevol* **1**, 191–203 (2002). doi: 10.1016/S1631-0683(02)00028-3

ACKNOWLEDGMENTS

We thank the Uganda National Council of Science and Technology for permission to conduct research, the Uganda National Museum for issuing an excavation license, and the former Ministry of Tourism, Trade and Industry and the current Ministry of Tourism, Wildlife and Antiquities for assisting with permissions. We are very grateful for the research support of Commissioner for Antiquities and Museums R. Mwanja, as well as for expert assistance from staff from the Uganda National Museum, particularly E. Musiime and S. Musalizi. E. Mbua, F. Manthi, T. Mukhuyu, and J. Kibii graciously provided access to comparative collections at the National Museums of Kenya. We thank B. Acayo, M. Aleper, N. Abiti, C. Ajambo, I. Arney, J. DeSilva, H. Kasozi, J. Kiyemba, S. Lapeyre-Montrose, M. Mafabi, M. Malone, M. S. Wampamba, J. Nabbale, V. Nakimuli, B. Nalikka, D. Ongwen, and E. Toloa for invaluable assistance in the field. We also thank K. Yezzi-Woodley for assistance with taphonomic data collection and R. Rogers and M. Tappen for advice in developing taphonomic sampling methods. Preparation of fossils from Moroto was done in the Vertebrate Fossil Preparation Laboratory at the Museum of Paleontology, University of Michigan, as well as by W. Downs. We thank A. Rountrey for conducting the photogrammetry on the postcranial specimens and J. Klausmeyer for his work on the summary figure. L. Sarringhaus and S. Namaganda generously assisted with the videotaping of extant chimpanzees in the Ngogo community, Kibale Forest, Uganda. J. Mitani provided feedback on an earlier version of this paper. We also thank D. Pilbeam, C. Chapman, and D. Gebo for the critical roles they played in the origin of this project. The comments from one anonymous reviewer and reviewer S. Almécija were greatly appreciated. This manuscript is publication no. 11 supporting Research on Eastern African Catarrhine and Hominoid Evolution (REACHE). **Funding:** This work was supported by the National Science Foundation (grants BCS 0215877, BCS 0456589, BCS 1208369, and BCS 1241811 to L.M.M.; grant BCS 1241807 to K.P.M. and D.L.F.; grant BCS 1241812 to D.J.P. and S.G.D.; grant BCS 1241817 to J.B.R.; and grant BCS 1241918 to A.L.D.); the LSB Leakey Foundation (L.M.M. and R.J.W.J.); a Wenner-Gren Foundation Workshop Grant (S.M.C., K.P.M., L.M.M., F.K.M.); the University of Michigan (L.M.M. and J.D.K.); the Natural Sciences and Engineering Research Council of Canada (S.M.C.); the University of Calgary (S.M.C.); and the University of Minnesota (R.J.W.J.). **Author contributions:** Conceptualization: L.M.M., J.D.K., J.B.R., S.M.C., K.P.M.; Funding acquisition: L.M.M., S.M.C., A.L.D., K.P.M., D.J.P., J.B.R.; Investigation: L.M.M., J.D.K., S.M.C., M.N.C., A.L.D., S.G.D., D.L.F., A.J.F., R.J.W.J., K.E.H.J., R.M.K., R.N.K., W.E.L., K.P.M., A.A.T.M., A.N., D.J.P., W.J.S., C.A.E.S., K.T.U., A.J.W.; Methodology: S.M.C., M.N.C., A.L.D., S.G.D., D.L.F., A.J.F., R.J.W.J., K.E.H.J., J.D.K., R.M.K., R.N.K., W.E.L., L.M.M., K.P.M., A.A.T.M., A.N., D.J.P., J.B.R., W.J.S., C.A.E.S., K.T.U., A.J.W. Project administration: L.M.M., R.M.K., A.A.T.M.; Supervision: L.M.M., J.D.K., S.M.C., R.M.K., A.A.T.M., D.J.P., K.P.M.; Visualization: J.D.K., A.L.D., S.G.D., R.J.J., L.M.M., D.J.P., C.A.E.S.; Writing – original draft: L.M.M., J.D.K., S.M.C., A.L.D., S.G.D., D.L.F., K.E.H.J., R.J.W.J., W.E.L., D.J.P., J.B.R., W.J.S., C.A.E.S., K.T.U.; Writing – review and editing: L.M.M., J.D.K., S.M.C., D.L.F., R.M.K., K.P.M., A.A.T.M., J.B.R., C.A.E.S. **Competing interests:** The authors declare no competing interests. **Data and materials availability:** All data are available in the main text or the supplementary materials. All Ugandan fossils are accessioned at the Uganda National Museum, Kampala, Uganda. The accession numbers for fossils can be found in table S6. **License information:** Copyright © 2023 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

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abq2835
Materials and Methods
Figs. S1 to S14
Tables S1 to S18
References (*133*–*197*)
Movies S1 to S3
Data S1 and S2 (3D Models)
MDAR Reproducibility Checklist

Submitted 8 April 2022; accepted 1 March 2023
10.1126/science.abq2835

RESEARCH ARTICLE

PALEOECOLOGY

Oldest evidence of abundant C_4 grasses and habitat heterogeneity in eastern Africa

Daniel J. Peppe[1]†, Susanne M. Cote[2]†, Alan L. Deino[3], David L. Fox[4]†, John D. Kingston[5]†, Rahab N. Kinyanjui[6,7,8], William E. Lukens[9]†, Laura M. MacLatchy[5,10]†, Alice Novello[11,12]‡, Caroline A. E. Strömberg[12]†, Steven G. Driese[1], Nicole D. Garrett[13], Kayla R. Hillis[14], Bonnie F. Jacobs[15], Kirsten E. H. Jenkins[16], Robert M. Kityo[17], Thomas Lehmann[18], Fredrick K. Manthi[6], Emma N. Mbua[6], Lauren A. Michel[14], Ellen R. Miller[19], Amon A. T. Mugume[17,20], Samuel N. Muteti[6,13], Isaiah O. Nengo[21,22]§, Kennedy O. Oginga[1]¶, Samuel R. Phelps[23]#, Pratigya Polissar[24], James B. Rossie[22], Nancy J. Stevens[25], Kevin T. Uno[26], Kieran P. McNulty[13]*†

The assembly of Africa's iconic C_4 grassland ecosystems is central to evolutionary interpretations of many mammal lineages, including hominins. C_4 grasses are thought to have become ecologically dominant in Africa only after 10 million years ago (Ma). However, paleobotanical records older than 10 Ma are sparse, limiting assessment of the timing and nature of C_4 biomass expansion. This study uses a multiproxy design to document vegetation structure from nine Early Miocene mammal site complexes across eastern Africa. Results demonstrate that between ~21 and 16 Ma, C_4 grasses were locally abundant, contributing to heterogeneous habitats ranging from forests to wooded grasslands. These data push back the oldest evidence of C_4 grass–dominated habitats in Africa—and globally—by more than 10 million years, calling for revised paleoecological interpretations of mammalian evolution.

Grasses using the C_4 photosynthetic pathway are ubiquitous across Earth's low to mid-latitudes, dominating modern tropical lowland grassland and savannah ecosystems. C_4 grassy biomes play an important role in regulating global climate and have been linked to key adaptations and diversification in mammalian faunas (*1*–*3*). The C_4 photosynthetic pathway is physiologically advantageous under conditions of aridity, higher temperatures and irradiance, seasonal climates, and low atmospheric partial pressure of carbon dioxide (pCO_2) (*1*). Phylogenetic analyses show that, beginning in the Paleocene, C_4 photosynthesis evolved independently at least 21 times from C_3 ancestors within the generally warm-adapted and shade-tolerant PACMAD (Panicoideae, Arundinoideae, Chloridoideae, Micrairoideae, Aristidoideae, Danthonioideae) grass clade (*2*). However, such ancient origins are not corroborated by the geologic record. Likewise, abundant C_4 biomass, for which we use a working definition of ≳30% C_4 biomass, is not documented before ~10 million years ago (Ma), except in isolated cases (*4*–*6*). Instead, major ecological shifts from C_3 vegetation to C_4 grasslands are well documented regionally and globally between 10 and 1 Ma (*2*, *3*, *7*, *8*). This apparent contradiction between phylogenetic estimates and geological evidence of C_4 emergence times has confounded attempts to explain the origins and global spread of C_4 grasslands (*2*, *3*, *9*).

Within eastern Africa, the expansion of C_4 grasslands has been intensely studied because of their relevance for interpreting the evolution of numerous mammalian lineages, including the hominin clade. Carbon isotope data from multiple substrates suggest that ecologically important C_4 biomass appeared in the region only after 10 Ma (*8*, *10*–*13*). However, older vegetation records are sparse (*14*, *15*). This paucity of data from the Paleogene and early Neogene has led to inferences that equatorial Africa was largely forested, with C_4 grasses and open habitats making up only minor elements of the landscape until the Late Miocene (*4*, *7*, *11*).

Here, we present a multiproxy study of Early Miocene fossil mammal site complexes in Kenya and Uganda (Fig. 1) that combines carbon isotope analyses of bulk soil organic matter, plant wax biomarkers, and pedogenic carbonates with inferences from phytolith (microscopic plant silica body) assemblages, forming a comprehensive approach to paleoecological reconstruction (*16*). Specifically, we assess the role of grasses, especially C_4 grasses, in early Neogene habitats of eastern Africa.

Early Miocene habitat heterogeneity, open habitats, and locally abundant C_4 vegetation

Results from carbon isotope analyses of ancient soil organic matter at these sites demonstrate a surprisingly high degree of habitat heterogeneity, with most datasets yielding $\delta^{13}C$ values that reflect vegetation types ranging from closed canopy to wooded grassland, and no clear trends toward increasing or decreasing openness through time in the Early Miocene or spatially across eastern Africa (Fig. 2A and data S1) (*16*). Every site yielded some paleosol samples with $\delta^{13}C$ values indicating that standing biomass was a mixture of C_3 plants and water-stressed C_3 and/or C_4 biomass (Fig. 2A) (*16*, *17*). Soil organic matter $\delta^{13}C$ values at Moroto, Tinderet, Bukwa, Rusinga, and Karungu reflect more open values that extend into the range of modern woodlands, bushlands, and shrublands, with Moroto, Bukwa, and Rusinga indicating a proportion of C_4 biomass equivalent to those characterizing extant wooded grasslands (Fig. 2A). Taken together, we interpret these results to reflect regionally abundant C_4 grasses that were heterogeneously present on the landscape.

Interpreting ancient soil organic matter $\delta^{13}C$ values is complicated by potential preferential degradation of isotopically light, labile components and by higher rates of decomposition of C_4-derived carbon observed in mixed C_3-C_4 systems (*16*, *18*). Furthermore, although the carbon in soil organic matter predominantly

[1]Department of Geosciences, Baylor University, Waco, TX 76798, USA. [2]Department of Anthropology and Archaeology, University of Calgary, Calgary, AB T2N 1N4, Canada. [3]Berkeley Geochronology Center, Berkeley, CA 94709, USA. [4]Department of Earth and Environmental Sciences, University of Minnesota, Minneapolis, MN 55455, USA. [5]Department of Anthropology, University of Michigan, Ann Arbor, MI 48109, USA. [6]Department of Earth Sciences, National Museums of Kenya, Nairobi 00100, Kenya. [7]Max Planck Institute for Geoanthropology, D-07743 Jena, Germany. [8]Human Origins Program, National Museum of Natural History, Smithsonian Institution, Washington, DC 20013, USA. [9]Department of Geology & Environmental Science, James Madison University, Harrisonburg, VA 22807, USA. [10]Museum of Paleontology, University of Michigan, Ann Arbor, MI 48109, USA. [11]CEREGE, Aix-Marseille Université, CNRS, IRD, Collège de France, INRAE, 13545 Aix en Provence, France. [12]Department of Biology, Burke Museum of Natural History and Culture, University of Washington, Seattle, WA 98195, USA. [13]Department of Anthropology, University of Minnesota, Minneapolis, MN 55455, USA. [14]Department of Earth Sciences, Tennessee Tech University, Cookeville, TN 38505, USA. [15]Roy M. Huffington Department of Earth Sciences, Southern Methodist University, Dallas, TX 75275, USA. [16]Department of Social Sciences, Tacoma Community College, Tacoma, WA 98466, USA. [17]Department of Zoology Entomology and Fisheries Sciences, Makerere University, Kampala, Uganda. [18]Department Messel Research and Mammalogy, Senckenberg Research Institute and Natural History Museum, 60325 Frankfurt, Germany. [19]Department of Anthropology, Wake Forest University, Winston-Salem, NC 27109, USA. [20]Uganda National Museum, Department of Museums and Monuments, Ministry of Tourism, Wildlife and Antiquities, Kampala, Uganda. [21]Turkana Basin Institute, Stony Brook University, Stony Brook, NY 11794, USA. [22]Department of Anthropology, Stony Brook University, Stony Brook, NY 11794, USA. [23]Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA 02138, USA. [24]Ocean Sciences Department, University of California, Santa Cruz, Santa Cruz, CA 95064, USA. [25]Department of Biomedical Sciences, Heritage College of Osteopathic Medicine, and Ohio Center for Ecological and Evolutionary Studies, Ohio University, Athens, OH 45701, USA. [26]Division of Biology and Paleo Environment, Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY 10964, USA.
*Corresponding author. Email: kmcnulty@umn.edu
†These authors contributed equally to this work.
‡Present address: Aix-Marseille Université, Direction de la Recherche et de la Valorisation, 13001 Marseille, France.
§Deceased. ¶Present address: Terra Guidance, Englewood, CO 80110, USA. #Present address: CIM Group, New York, NY 10022, USA.

comes from bulk plant material, it can also come from soil microbiota, mycorrhizae, or fungi. However, long-chain *n*-alkanes are a single, plant-derived molecule, and our $\delta^{13}C$ values from leaf wax *n*-alkanes extracted from the same samples as soil organic matter yielded parallel results: substantial variation within stratigraphic sections and abundant C_4 biomass at Moroto and Rusinga (Fig. 2B, figs. S1 and S2, and data S2 and S3) (*16*). The C_{31} *n*-alkane includes input from both woody and grassy vegetation, so moderate $\delta^{13}C$ values for this homolog cannot unequivocally distinguish low concentrations of C_4 biomass from water-stressed C_3 biomass (*8*, *16*). However, the C_{35} *n*-alkane is preferentially produced by C_4 grasses (*19*), therefore samples at Moroto with high $\delta^{13}C$ values in both C_{35} and C_{31} *n*-alkanes unambiguously indicate considerable C_4 biomass (Fig. 2B and fig. S1) (*20*). In addition, the high relative abundance of long-chain *n*-alkanes (C_{33} and C_{35}) from Moroto is similar to patterns found in modern C_4 grasslands (*19*) and in sediments from C_4-dominated environments (fig. S1) (*21*). Interestingly, *n*-alkane $\delta^{13}C$ values from marine cores off of western and eastern African coasts do not provide evidence of regionally abundant Early Miocene C_4 vegetation (*8*, *10*, *11*). This apparent discrepancy likely reflects differences in scale and resolution between paleosols sampled in this study, which represent local vegetation, and sediments deposited at deep-sea marine core sites, which integrate plant wax carbon isotope records over thousands of square kilometers.

Pedogenic carbonate isotopic values confirm variation within and among sites. Pedogenic carbonate was not found at all sites, nor uniformly throughout sections where it was present (*16*), demonstrating spatial and temporal variation in climate (i.e., mesic conditions) and/or pedogenic processes (e.g., acidic pH) that inhibited carbonate precipitation. Carbonate $\delta^{13}C$ distributions are shifted even more toward open habitats relative to organic matter (Fig. 2C and data S4) (*16*), with a few $\delta^{13}C$ values indicating C_4-dominated vegetation for some paleosols at multiple sites (Moroto, Napak, Karungu, and Buluk). Bulk soil organic matter

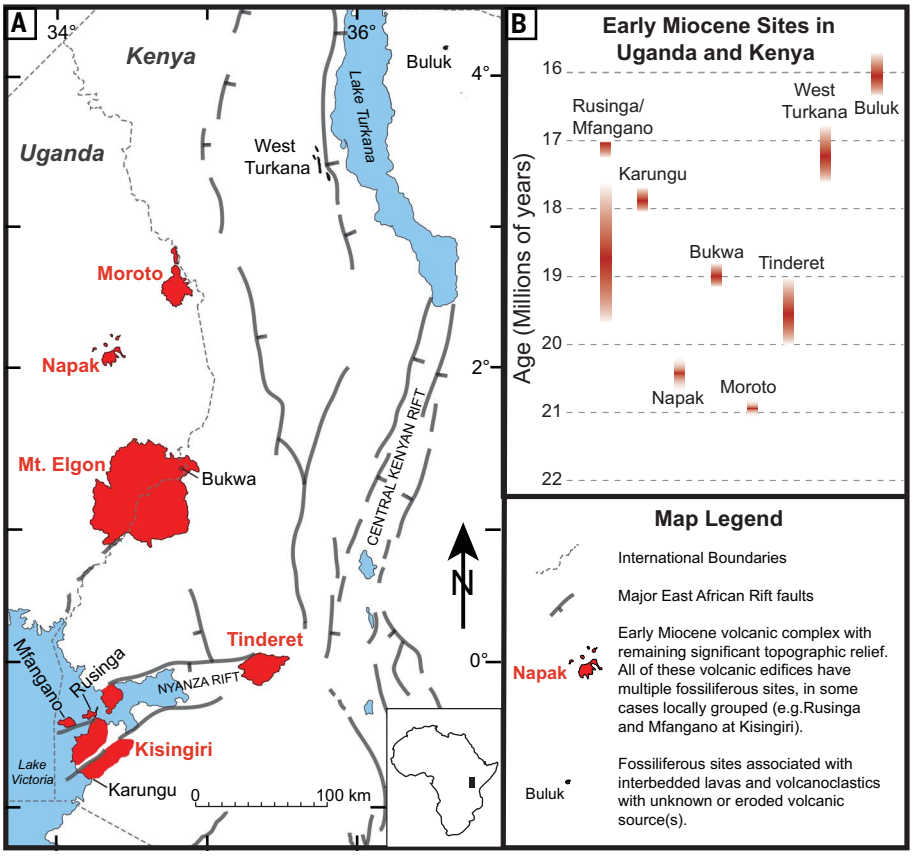


Fig. 1. Early Miocene fossil site complexes in eastern Africa. (**A**) Map of sites. Areas in red represent current exposures of primarily volcanic and associated sedimentary rock. (**B**) Chronology of fossil sites sampled. Further site information is provided in the supplementary materials (*16*).

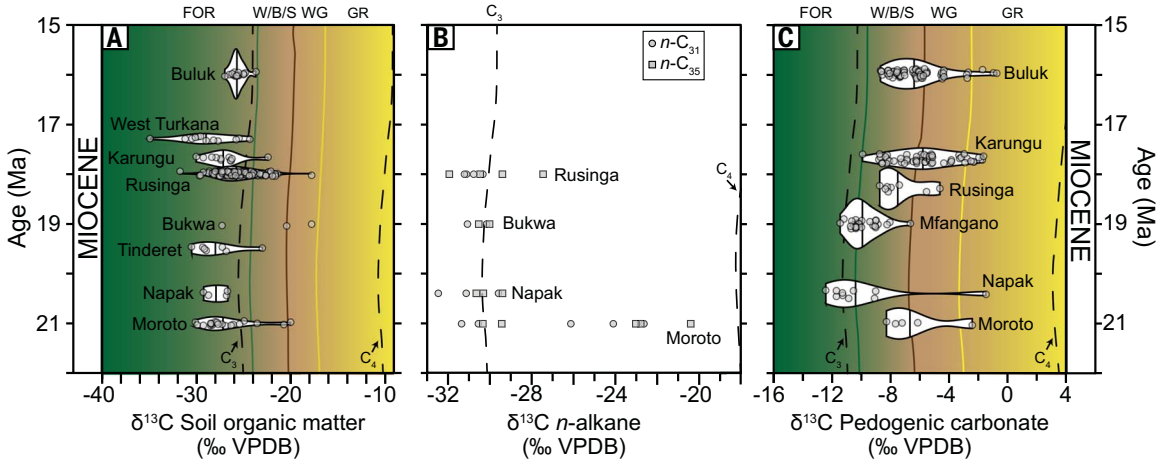


Fig. 2. Stable carbon isotope ratios ($\delta^{13}C$) for three proxies. (**A**) Bulk soil organic matter, (**B**) C_{31} and C_{35} *n*-alkane homologs, and (**C**) pedogenic carbonates. Data distributions in (A) and (C) shown as violin plots (mirrored kernel density estimates), with violin areas normalized in each panel and trimmed to data ranges in each group; points are jittered for clarity, and vertical crossbars indicate the median of the group. Dashed lines represent mean C_3 and C_4 endmember $\delta^{13}C$ values for soil organic matter and pedogenic carbonates (*34*) and *n*-alkanes (*11*) based on data from modern soils; shaded regions with colored boundaries represent UNESCO biome classifications after (*35*). Biome boundaries for *n*-alkanes have not been sufficiently characterized and thus are not shown. Endmembers and biome boundaries have been corrected for Miocene atmospheric CO_2 $\delta^{13}C$ value after (*36*). Detailed methodology and data are provided in (*16*) and tables S1 and S2 and data S1, S2, and S4. VPDB, Vienna Pee Dee belemnite; FOR, forest; W/B/S, woodland/bushland/shrubland; WG, woody grassland; GR, grassland.

and carbonate $\delta^{13}C$ values from the same stratigraphic levels are not significantly correlated (fig. S3B), and the mean offset between them [$\Delta^{13}C_{CO3\text{-}OM}$ = 20.0 per mil (‰)] is greater than theoretical expectations (~14 to 17‰) (*22*), possibly indicating that pedogenic carbonates or organic matter (or both) do not preserve original environmental signatures. This interpretation is unlikely for several reasons. First, samples with matched soil organic matter and *n*-alkane $\delta^{13}C$ values show strong, significant correlations (Fig. 2, A and B, and fig. S3), supporting the fidelity of the isotopic signals from our larger bulk organic dataset. Second, our field sampling strategy involved collection of pedogenic carbonates well within the subsoil, typically at least 0.5 m in depth, on the basis of the description of paleosol horizonation and macroscopic pedogenic morphologies (*16*) (data S4). Third, the carbonates used for paleovegetation reconstruction were prepared by drilling micritic crystal textures on cut and polished rock samples after petrographic analysis, including cathodoluminescence petrography, whenever possible (*16*). We also purposefully sampled sparry carbonate in addition to pedogenic micrite to assess for diagenetic contamination (fig. S4). However, we found no correspondence between crystal texture and either $\delta^{13}C$ or $\delta^{18}O$ values (fig. S5) (*16*), from which we infer that diagenesis cannot account for $\delta^{13}C$ values above a C_3 endmember threshold. Fourth, paleosol morphologies and reconstructed hydroclimate from independent proxies at multiple sites (*6*, *20*, *23*) are consistent with high soil respiration rates, minimizing the possibility that infiltration of atmospheric CO_2 into soil profiles drove the elevated carbonate $\delta^{13}C$ values that we measured.

We propose instead that both pedogenic carbonates and soil organic matter preserve environmental signals that capture different components of floral compositions at different time resolutions. Many paleosols that we sampled were deposited and formed in dynamic, aggradational landscapes associated with rifting and volcanism (*20*, *23*), which would have caused short-term local changes in the composition of vegetation as the plant community responded to frequent disturbances. Owing to faster humus turnover times relative to the pace of carbonate accumulation (*24*), soil organic matter would retain the last phase of vegetation for any one profile, yet carbonates would capture a longer-term, but seasonally biased, signal of respired CO_2, reflecting vegetation that was most productive during warm and dry seasons (i.e., C_4 plants). Consequently, carbonates would be more sensitive to recording C_4 plant presence compared with bulk organic matter from the same paleosols, especially if net primary productivity were higher in wetter seasons with more C_3 biomass.

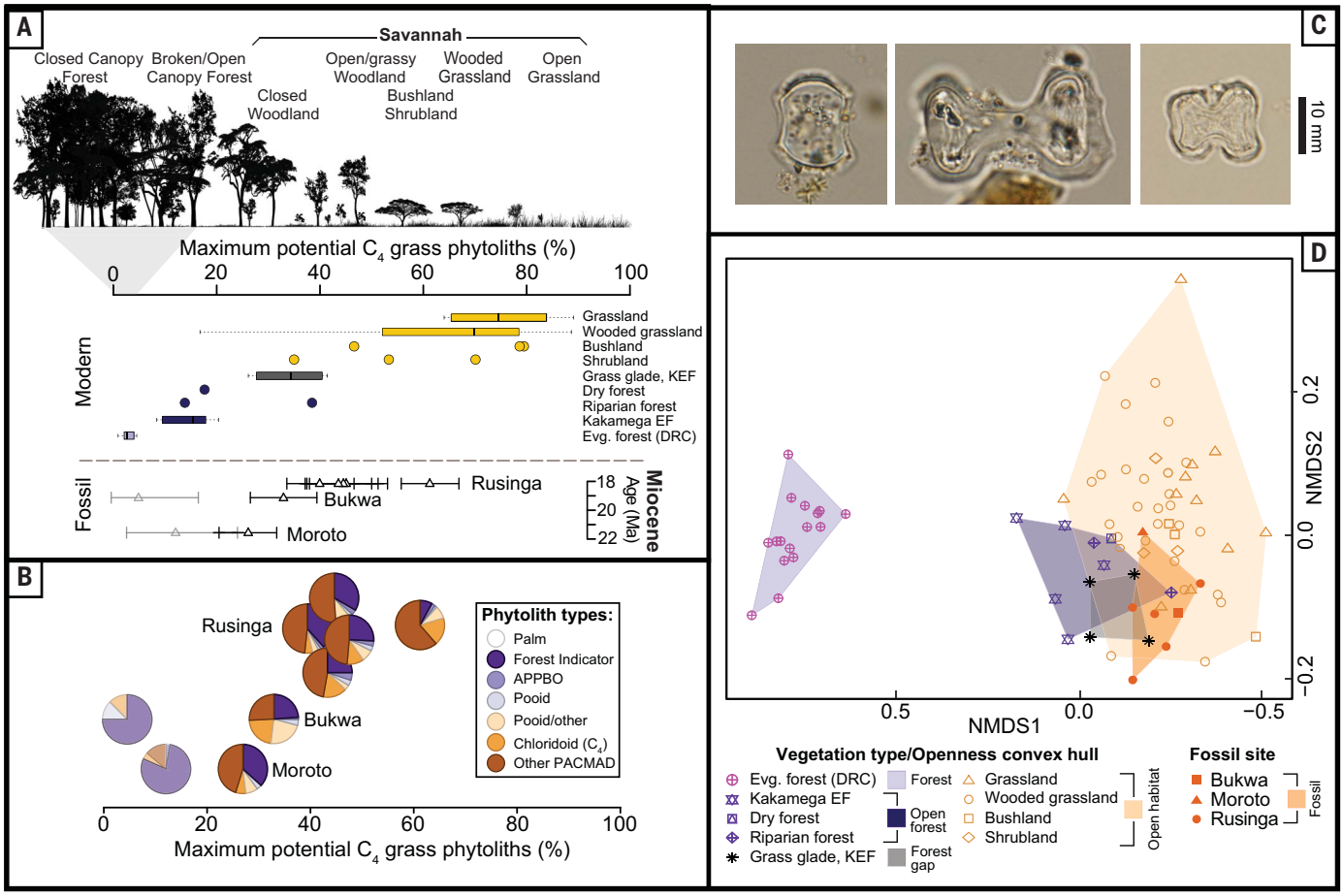


Fig. 3. Comparison of phytolith assemblages from fossil sites and modern vegetation types in eastern and central Africa. (**A**) Maximum potential C_4 (i.e., PACMAD) grass phytoliths of diagnostic counts from modern vegetation types (top; showing median and quartiles for types with more than four data points) and fossil sites (bottom; showing bootstrapped 95% confidence intervals around estimates). Light-gray fossil ranges represent assemblages with low diagnostic counts. (**B**) Pie charts representing diagnostic assemblage data (excluding nondiagnostic grass phytoliths). Sites are arranged stratigraphically, and ages for sites are shown in (A) and Fig. 1. Faded pie charts as in (A). (**C**) Representative PACMAD grass phytolith types from left to right: Saddle (C_4 Chloridoideae), Bilobate (PACMAD), and Cross (PACMAD). (**D**) Ordination [nonmetric multidimensional scaling (NMDS); *k* = 2] of modern and fossil assemblage data. Details and data are provided in (*16*) and data S5 to S11. APPBO, closed-habitat grasses, which are diagnostic of early diverging grasses in the families Anomochlooideae, Pharoideae, Puelioideae, Bambusoideae, and Oryzoideae; DRC, Democratic Republic of Congo; Evg. forest or EF, evergreen forest; KEF, Kakamega evergreen forest.

The ecological significance of C_4 grasses and habitat heterogeneity is further corroborated by phytolith assemblages documenting vegetation on these Early Miocene landscapes (*16*). Every site that yielded phytoliths provided evidence of grasses, and all except Napak and Moruorot had phytoliths from PACMAD grasses (i.e., potential C_4 grasses) (data S5 and S6). Well-preserved assemblages from Moroto, Bukwa, and Rusinga could be analyzed quantitatively, and most are dominated by grass phytoliths (up to 91.9% of diagnostic phytolith counts; Fig. 3A, fig. S6, and data S5) (*16*). As many as 90.8% of grass phytoliths (excluding unknown and damaged forms) are diagnostic of PACMADs, and up to 29.0% of grass phytoliths are typical of the strictly-C_4 Chloridoideae PACMAD subclade (Fig. 3B). Nonchloridoid PACMAD phytoliths could represent C_3 or C_4 grasses, but strong C_4 isotopic signals from the same sites suggest that many derived from C_4 species. Nevertheless, to acknowledge this inherent uncertainty, we refer to the sum of PACMAD phytoliths as "maximum potential C_4 phytoliths" (Figs. 3A and 4). Consistent with isotopic evidence for habitat heterogeneity, the relative abundance of forest indicator phytoliths varied within and across sites as well, ranging from 8.1 to 81.8% (Fig. 3A and data S5).

Proportions of potential C_4 grasses in fossil phytolith assemblages most closely resemble modern eastern and central African habitats with substantial grass components: wooded grasslands, shrubland, bushland, riparian forests, and grassy glades within forests (Fig. 3A). Early Miocene grass communities were diverse, including phytoliths typical of closed-habitat grasses, various PACMAD grasses (Fig. 3, B and C), and potentially Pooideae grasses (strictly C_3) (table S5). Nonmetric multidimensional scaling distinguishes closed habitats, represented by the Guineo-Congolian lowland rainforest, from open habitats in the form of (wooded) grasslands, shrublands, and bushlands (Fig. 3D, fig. S7, and table S1) (*16*). Assemblages from forests with substantial grassy glades (i.e., Kakamega) or that are surrounded by open, grass-dominated vegetation (i.e., riparian forests in savannah landscapes) partly overlap with open, grassy vegetation types, but most Early Miocene assemblages fall outside these compositional fields, instead showing greater similarity with C_4 grass–dominated habitats.

Together, our data from multiple chronologically constrained paleoecological proxies reveal that: (i) C_4 grasses were present in the Early Miocene of eastern Africa, and at some sites were a dominant part of the vegetation; (ii) open habitats (e.g., woodland, bushland, shrubland, and wooded grassland biomes) were a substantial and important component of Early Miocene ecosystems, contributing to landscape heterogeneity at individual sites and across the region both spatially and temporally (Figs. 3 and 4, and fig. S8); and (iii) the range of habitats and $\delta^{13}C$ values (Fig. 4, A to D, and fig. S8) and the C_4 grass abundances (Fig. 3, A and B) during the Early Miocene were at times similar to grassy savannah and woodland ecosystems reconstructed for the Late Miocene and Plio-Pleistocene (Fig. 4, A to D, and fig. S8) (*25*), only lacking evidence of true grasslands or rainforests [but see (*26*)]. Although it remains a challenge to differentiate spatial and temporal heterogeneity at a local level because our paleosol and paleoecological proxies are time-averaged to various degrees, temporally restricted horizons at Moroto and Rusinga document penecontemporaneous forested to grassy woodland microhabitats (*20*, *23*), suggesting that landscape-scale spatial heterogeneity contributed to the variable paleovegetation signals in our records.

Implications for Neogene ecosystems of eastern Africa

This study unequivocally extends the record of C_4 grasses in Africa, roughly coeval with

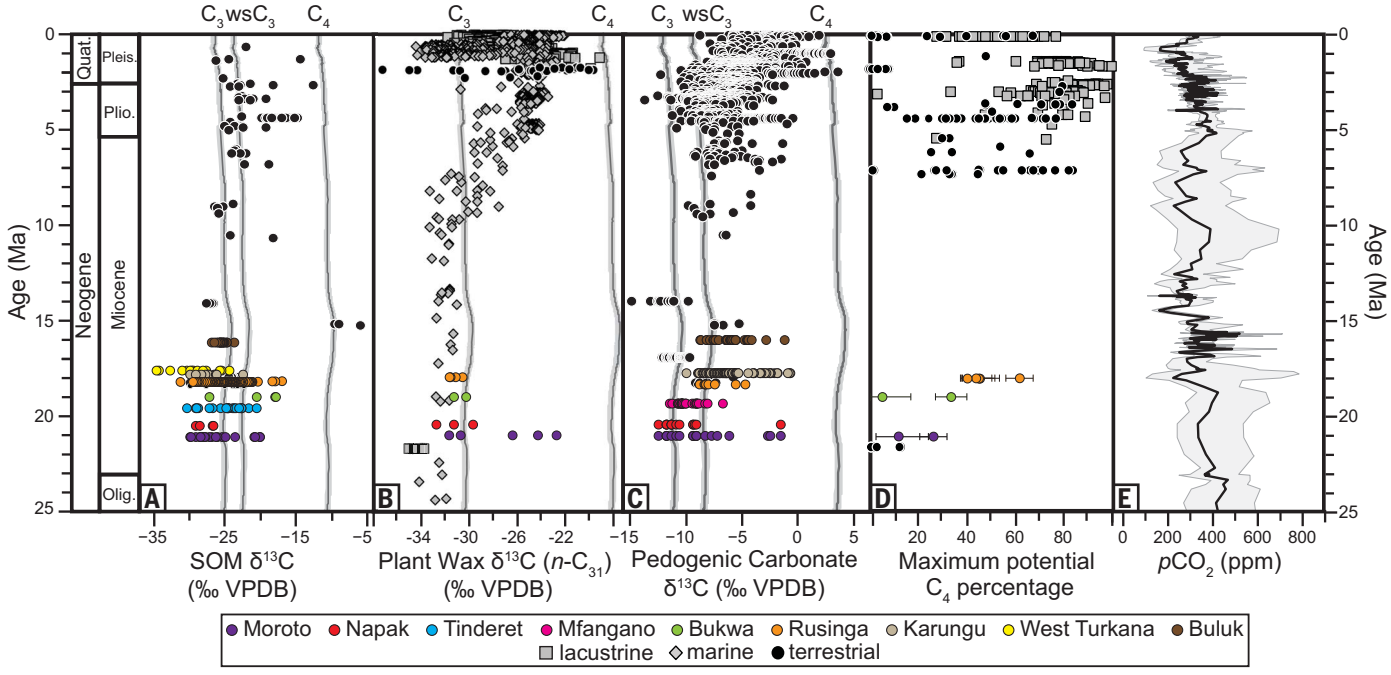


Fig. 4. Long-term record of isotopes, phytoliths, and CO_2 from the late Oligocene–Holocene for eastern and central Africa. (**A**) Soil organic matter (SOM) $\delta^{13}C$ values from terrestrial samples. (**B**) $\delta^{13}C$ values for C_{31} *n*-alkane homolog ($\delta^{13}C_{wax}$) from terrestrial, marine, and lacustrine samples. (**C**) $\delta^{13}C$ from pedogenic carbonates sampled from paleosols. (**D**) Maximum potential C_4 (i.e., PACMAD) grass phytoliths of diagnostic counts with 95% confidence bars. (**E**) CO_2 compilation for the past 25 million years from (*29*). ppm, parts per million. Solid line depicts grand mean of all CO_2 proxy data, and gray outline indicates 95% confidence interval. [(A) to (C)] Black lines represent mean C_3, mean water-stressed (wsC_3), and C_4 endmember $\delta^{13}C$ values calculated and corrected for atmospheric CO_2 $\delta^{13}C$ as for Fig. 2. Data, data sources, and methods are provided in (*16*), table S1, and data S1, S2, S4, S5, S12, and S13. Quat., Quaternary; Pleis., Pleistocene; Plio., Pliocene; Olig., Oligocene. Holocene not shown owing to the scale of figure.

the earliest records of chloridoid phytoliths from North America (*3*). Furthermore, our data document the occurrence of ecologically abundant and, in several cases, locally dominant C_4 grasses, predating by more than 10 million years the near-global spread of C_4 grasslands in the Late Miocene–Early Pleistocene (Fig. 4) (*2*, *3*). This finding contradicts the longstanding view that C_4 grasses did not gain ecological importance in eastern Africa until the late Neogene (*2*, *7*, *8*, *11*). Expansion of C_4 grasses has been linked to a range of local to global factors, including decreasing atmospheric pCO_2, shifts in disturbance regime (e.g., fire, herbivory, volcanism), and changes in regional climate (e.g., aridity, seasonality) (*2*, *8*, *9*, *22*, *27*, *28*). However, changes in pCO_2 are not likely to have played a role in structuring Early Miocene vegetation, as CO_2 levels remained relatively high during this time (Fig. 4) (*29*). Changes in the fire regime and fauna–vegetation interaction may have influenced plant communities, but current evidence from eastern Africa is insufficient to evaluate these hypotheses in the Early Miocene. We hypothesize that regional tectonic factors (e.g., recurrent volcanic eruptions and edifice building, extension-induced formation of escarpments and grabens) expressed variably at the scale of individual sites were the primary controls on heterogeneity and regional C_4 grass dominance in eastern Africa. Orographic climate effects related to doming and uplift associated with incipient rifting would have disrupted both local and regional seasonality and drainage patterns, further enhancing ecological variability. The presence of topographic barriers may also explain the lack of agreement between in situ signals of C_4 vegetation reported in this study and regionally integrated records from marine cores off of the eastern and western coasts of Africa, which first record C_4 signals at 10 Ma (*8*).

Our discovery that heterogeneously distributed C_4 vegetation preceded the rise to dominance of C_4 grassland habitats in eastern Africa stands in stark contrast to data from North America, South America, Australia, and eastern Asia, which indicate that C_4 grasslands replaced C_3-dominated grassy mosaics or other open, semiarid to arid vegetation (*3*, *9*, *30*). Thus, the C_3–C_4 transition in eastern Africa may have been more protracted and complex than elsewhere. Additionally, habitat heterogeneity demonstrated by our data counters hypotheses of predominantly forested habitats in the Early Miocene of eastern Africa [(*2, 31, 32*), but see (*33*)], instead supporting a more nuanced interpretation of ecosystem evolution (*15*, *20*, *23*). At local scales—those at which organisms interact with their environments—the Early Miocene in eastern Africa was characterized by a variety of habitats, including open vegetation with C_4 grasses. The evident importance of open habitats in the early Neogene compels a deeper time perspective when considering adaptive hypotheses linking mammalian community evolution with grassland expansion in the fossil record.

REFERENCES AND NOTES

1. R. F. Sage, *New Phytol.* **161**, 341–370 (2004).
2. E. J. Edwards *et al*., *Science* **328**, 587–591 (2010).
3. C. A. E. Strömberg, *Annu. Rev. Earth Planet. Sci.* **39**, 517–544 (2011).
4. B. Jacobs, J. Kingston, L. Jacobs, *Ann. Mo. Bot. Gard.* **86**, 590–643 (1999).
5. J. D. Kingston, in *Late Cenozoic Environments and Hominid Evolution: a tribute to Bill Bishop*, P. Andrews, P. Banham, Eds. (The Geological Society, 1999), pp. 69–84.
6. W. E. Lukens *et al*., *Front. Earth Sci. (Lausanne)* **5**, 87 (2017).
7. T. Cerling *et al*., *Nature* **389**, 153–158 (1997).
8. P. J. Polissar, C. Rose, K. T. Uno, S. R. Phelps, P. deMenocal, *Nat. Geosci.* **12**, 657–660 (2019).
9. A. T. Karp, A. K. Behrensmeyer, K. H. Freeman, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 12130–12135 (2018).
10. S. J. Feakins *et al*., *Geology* **41**, 295–298 (2013).
11. K. T. Uno, P. J. Polissar, K. E. Jackson, P. B. deMenocal, *Proc. Natl. Acad. Sci. U.S.A.* **113**, 6355–6363 (2016).
12. K. T. Uno *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **108**, 6509–6514 (2011).
13. N. E. Levin, *Annu. Rev. Earth Planet. Sci.* **43**, 405–429 (2015).
14. B. F. Jacobs, A. D. Pan, C. R. Scotese, in *Cenozoic Mammals of Africa*, L. Werdelin, W. J. Sanders, Eds. (Univ. of California Press, 2010), pp. 57–72.
15. T. L. P. Couvreur *et al*., *Biol. Rev. Camb. Philos. Soc.* **96**, 16–51 (2021).
16. Materials and methods are available as supplementary materials.
17. T. E. Cerling, J. M. Harris, M. G. Leakey, N. Mudida, in *Lothagam: The Dawn of Humanity in Eastern Africa*, M. G. Leakey, J. M. Harris, Eds. (Columbia Univ. Press, 2003), pp. 583–604.
18. J. G. Wynn, *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **251**, 437–448 (2007).
19. R. T. Bush, F. A. McInerney, *Geochim. Cosmochim. Acta* **117**, 161–179 (2013).
20. L. M. MacLatchy *et al*., *Science* **380**, eabq2835 (2023).
21. Y. Garcin *et al*., *Geochim. Cosmochim. Acta* **142**, 482–500 (2014).
22. T. E. Cerling, *Global Biogeochem. Cycles* **6**, 307–314 (1992).
23. L. A. Michel *et al*., *Sedimentology* **67**, 3567–3594 (2020).
24. M. I. Bird, A. R. Chivas, J. Head, *Nature* **381**, 143–146 (1996).
25. A. Novello *et al*., *J. Hum. Evol.* **106**, 66–83 (2017).
26. E. D. Currano *et al*., *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **539**, 109425 (2020).
27. S. Hoetzel, L. Dupont, E. Schefuß, F. Rommerskirchen, G. Wefer, *Nat. Geosci.* **6**, 1027–1030 (2013).
28. P. Sepulchre *et al*., *Science* **313**, 1419–1423 (2006).
29. G. L. Foster, D. L. Royer, D. J. Lunt, *Nat. Commun.* **8**, 14845 (2017).
30. C. A. E. Strömberg, R. E. Dunn, C. Crifò, E. B. Harris, in *Methods in Paleoecology: Reconstructing Cenozoic Terrestrial Environments and Ecological Communities*, D. A. Croft, D. F. Su, S. W. Simpson, Eds. (Springer Cham, 2018), pp. 235–287.
31. R. J. Morley, K. Richards, *Rev. Palaeobot. Palynol.* **77**, 119–127 (1993).
32. S. J. Feakins, P. deMenocal, in *Cenozoic Mammals of Africa*, L. Werdelin, W. J. Sanders, Eds. (Univ. of California Press, 2010), pp. 45–55.
33. P. Andrews, *An Ape's View of Human Evolution* (Cambridge Univ. Press, 2015).
34. B. H. Passey *et al*., *J. Geol.* **110**, 123–140 (2002).
35. T. E. Cerling *et al*., *Nature* **476**, 51–56 (2011).
36. B. J. Tipple, S. R. Meyers, M. Pagani, *Paleoceanography* **25**, PA3202 (2010).
37. D. Peppe *et al*., Supplementary Data for Oldest evidence of significant C4 grasses and habitat heterogeneity in eastern Africa, version 1, Texas Data Repository (2023); https://doi.org/10.18738/T8/TCAMKJ.

ACKNOWLEDGMENTS

This work was conducted under the authority of, and with permits from, the National Commission for Science, Technology, and Innovation (Kenya) and the Department of Museums and Monuments, Ministry of Tourism, Wildlife and Antiquities (Uganda). We thank the staff of the National Museums of Kenya, Turkana Basin Institute, Uganda Museum, and Makerere University for their ongoing research support, with particular thanks to R. Mwanja and J. Kibii. Our gratitude extends to the many collaborators who provided assistance in the field and in museums: D. Aketch, I. Arney, K. Cheng, J. Erus Edung, S. Ekitela, A. Ekuwom, S. Longoria Eyanae, M. Fuchs, J. Greschbach, C. Hemm, R. Henning, H. Kasozi, J. Kiseu, U. Lächele, M. Macharwas, M. Malone, U. Menz, R. Moru, T. Mukhuyu, B. Nalikka, C. Ochieng, S. Owuor Odhiambo, B. Onyango, C. Ouma, T. Rumpf, J. Shaduma, J. Siembo, J. Tausch, and J. Torgeson. T. Cerling graciously provided modern soil samples from eastern Africa from which phytoliths were extracted. This is publication #12 supporting Research on Eastern African Catarrhine and Hominoid Evolution (REACHE). **Funding:** This work was funded by National Science Foundation grant BCS 1241807 (K.P.M. and D.L.F.), National Science Foundation grant BCS 0852609 (K.P.M.), National Science Foundation grant BCS 1241811 (L.M.M.), National Science Foundation grant BCS 1241812 (D.J.P. and S.G.D.), National Science Foundation grant BCS 1241817 (J.B.R.), National Science Foundation grant BCS 1241918 (A.L.D.), National Science Foundation grant EAR 1253713 (C.A.E.S.), National Science Foundation grant EAR 1053549 (B.F.J.), National Science Foundation grant BCS 1638796 (N.J.S.), National Science Foundation grant DGE 16-44869 (S.R.P.), Leakey Foundation grant (K.P.M.), Leakey Foundation grant (N.D.G.), Leakey Foundation grant (K.E.H.J.), Leakey Foundation grant (D.J.P. and K.P.M.), Leakey Foundation grant (L.A.M.), Leakey Foundation grant (E.R.M. and I.O.N.), Leakey Foundation Baldwin Fellowship (K.O.O.), McKnight Land-Grant Fellowship (K.P.M.), National Geographic Society grant (E.R.M. and I.O.N.), Wenner-Gren Foundation Workshop grant (S.M.C., K.P.M., L.M.M., and F.K.M.), Leverhulme Trust Fellowship (K.P.M.), EU Horizon 2020 for Research and Innovation MSCA IOF#659596 (A.N.), Karl und Marie Schack-Stiftung fund (T.L.), Vereinigung von Freunden und Förderern der Goethe-Universität Frankfurt (T.L.), Natural Sciences and Engineering Research Council of Canada (S.M.C.), Tennessee Tech University URECA Program (L.A.M.), Tennessee Tech University Department of Earth Sciences (L.A.M.), Center for Climate and Life, Columbia University (K.T.U.), and Vetlesen Foundation (K.T.U.). **Author contributions:** Conceptualization: S.M.C., A.L.D., S.G.D., D.L.F., J.D.K., L.M.M., F.K.M., E.N.M., K.P.M., A.N., D.J.P., J.B.R., N.J.S., and C.A.E.S. Methodology: S.M.C., S.G.D., D.L.F., J.D.K., R.N.K., K.E.H.J., W.E.L., L.M.M., F.K.M., K.P.M., L.A.M., E.R.M., I.O.N., A.N., K.O.O., D.J.P., S.R.P., P.P. C.A.E.S., and K.T.U. Data analysis: S.G.D., D.L.F., N.D.G., K.R.H., J.D.K., R.N.K., W.E.L., L.A.M., A.N., K.O.O., D.J.P., S.R.P., P.P., C.A.E.S., and K.T.U. Investigation: S.M.C., A.L.D., S.G.D., D.L.F., N.D.G., K.E.H.J., J.D.K., R.N.K., R.M.K., T.L., W.E.L., L.M.M., K.P.M., L.A.M., E.R.M., A.A.T.M., S.N.M., I.O.N., A.N., K.O.O., D.J.P., J.B.R., and C.A.E.S. Funding acquisition: S.M.C., A.L.D., S.G.D., D.L.F., N.D.G., B.F.J., K.E.H.J., J.D.K., T.L., L.M.M., F.K.M., K.P.M., L.A.M., E.R.M., I.O.N., A.N., K.O.O., D.J.P., J.B.R., and C.A.E.S. Project administration: S.M.C., L.M.M., K.P.M., E.R.M., I.O.N., D.J.P., and J.B.R. Writing – original draft: S.M.C., D.L.F., J.D.K., L.M.M., K.P.M., A.N., D.J.P., and C.A.E.S. Writing – review & editing: S.M.C., A.L.D., S.G.D., D.L.F., N.D.G., B.F.J., K.E.H.J., J.D.K., R.N.K., R.M.K., T.L., W.E.L., L.M.M., F.K.M., E.N.M., K.P.M., L.A.M., E.R.M., A.A.T.M., S.N.M., I.O.N., A.N., D.J.P., S.R.P., P.P., J.B.R., N.J.S., C.A.E.S., and K.T.U. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data are available in the supplementary materials and in the Texas Data Repository online archive (*37*).

License information: Copyright © 2023 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

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abq2834
Materials and Methods
Figs. S1 to S8
Tables S1 to S3
References (*38–144*)
MDAR Reproducibility Checklist
Data S1 to S13

Submitted 8 April 2022; accepted 10 March 2023
10.1126/science.abq2834

COAGULATION

Immobility-associated thromboprotection is conserved across mammalian species from bear to human

Manuela Thienel[1,2]*, Johannes B. Müller-Reif[3,4]*, Zhe Zhang[1]†, Vincent Ehreiser[1,2], Judith Huth[1,2], Khrystyna Shchurovska[1,2], Badr Kilani[1,2], Lisa Schweizer[3], Philipp E. Geyer[3,4], Maximilian Zwiebel[3], Julia Novotny[1], Enzo Lüsebrink[1], Gemma Little[5], Martin Orban[1], Leo Nicolai[1,2], Shaza El Nemr[1,2], Anna Titova[1,2], Michael Spannagl[6], Jonas Kindberg[7,8], Alina L. Evans[9], Orpheus Mach[10], Matthias Vogel[10], Steffen Tiedt[11], Steffen Ormanns[12], Barbara Kessler[13], Anne Dueck[2,14], Andrea Friebe[7,8], Peter Godsk Jørgensen[15], Monir Majzoub-Altweck[16], Andreas Blutke[16], Amin Polzin[17], Konstantin Stark[1,2], Stefan Kääb[1,2], Doris Maier[10], Jonathan M. Gibbins[5], Ulrich Limper[18,19], Ole Frobert[20,21,22,23]*, Matthias Mann[3]*, Steffen Massberg[1,2]*, Tobias Petzold[1,2]*

Venous thromboembolism (VTE) comprising deep venous thrombosis and pulmonary embolism is a major cause of morbidity and mortality. Short-term immobility-related conditions are a major risk factor for the development of VTE. Paradoxically, long-term immobilized free-ranging hibernating brown bears and paralyzed spinal cord injury (SCI) patients are protected from VTE. We aimed to identify mechanisms of immobility-associated VTE protection in a cross-species approach. Mass spectrometry–based proteomics revealed an antithrombotic signature in platelets of hibernating brown bears with heat shock protein 47 (HSP47) as the most substantially reduced protein. HSP47 down-regulation or ablation attenuated immune cell activation and neutrophil extracellular trap formation, contributing to thromboprotection in bears, SCI patients, and mice. This cross-species conserved platelet signature may give rise to antithrombotic therapeutics and prognostic markers beyond immobility-associated VTE.

Immobilization is a major risk factor for the development of venous thromboembolism (VTE) generating tremendous socioeconomic costs (*1*). Immobility-associated VTE is initiated by flow restriction in veins causing endothelial cell hypoxia (*2*). Recruitment of platelets to the activated endothelium leads to local activation of the plasmatic coagulation system and innate immune cells (for example neutrophils and monocytes), orchestrating deleterious thromboinflammation, the central mechanism of VTE (*3*, *4*). Despite long-term immobilization, chronically paralyzed patients with spinal cord injuries exhibit a VTE risk similar to that of the general population (*5*–*7*). Hibernating brown bears *(Ursus arctos)* are another example of an organism that experiences a period of long-term immobilization, during winter. Although hibernating bears enter a physically immobile state that is accompanied by considerably reduced cardiac output, there is no obvious evidence for VTE (*8*–*11*). In our present study, we identified a platelet-derived antithrombotic molecular signature that is conserved across different species and protects against immobility-induced VTE.

Results

VTE in nonhibernating brown bears

Immobilization is a major risk factor linked to increased rates of VTE even in the absence of additional healthcare-related factors, such as obesity, acute trauma, or medical interventions (*1*, *12*). In some animal species, hibernation leads to immobilization for several weeks or months. During their annual 4- to 7-month-long hibernation episodes, brown bears show notable plasticity by establishing an energy saving metabolism to adapt to hostile environmental changes during winter. However, this has not been linked to increased VTE-associated complications. A case in point are Swedish brown bears with a mortality rate during hibernation of only 0.4% compa with 58.4% during activity (fig. S1A). In t bears we found no echocardiographic evidence of pulmonary embolism during hibernation (right ventricle diameter, fig. S1B) and D-dimer levels were reduced compared with active bears (Fig. 1A). This indicates that immobilization during hibernation is not linked to an obvious excess in VTE rates. Nevertheless, nonhibernating brown bears may develop VTE (table S1). Analysis of autopsy samples revealed characteristic histological pulmonary thrombus structure in some bears, including intravascular neutrophil extracellular traps (NETs), prothrombogenic externalized nuclear DNA-histone strands that foster intravascular thrombosis (*13*, *14*) (Fig. 1, B and C). By spatial proteomics of thrombus samples, we identified key components of the coagulation system (e.g., factor X, factor V), platelets [e.g., glycoprotein (GP)V, GPIbα] and neutrophils (e.g., neutrophil elastase) in both brown bears and patients with VTE (Fig. 1, D and E, and table S2). Thus, brown bears possess a thromboinflammatory machinery similar to humans enabling them to develop VTE but appear to be protected from immobility-associated VTE during hibernation.

Antithrombotic platelet signatures during hibernation

To identify factors potentially contributing to a naturally occurring antithrombotic mechanism during hibernation, we serially captured thirteen subadult brown bears for blood sampling during hibernation and the following activity period in spring in middle Sweden (Fig. 2A). First, we performed hematological characterization of the Scandinavian brown bear cohort. Hemoglobin and hematocrit values were elevated during winter reflecting intravascular dehydration (*15*) (fig. S2A). Whole blood thromboelastographic analysis showed that clot formation upon activation of the intrinsic—but not extrinsic—coagulation pathway is delayed during inactivity (Fig. 2B), even when we adjusted for seasonal differences in body temperature (winter 33°C, summer 37°C). Thrombus mass (clot firmness) was mildly reduced during hibernation (Fig. 2B).

[1]Department of Cardiology, University Hospital, LMU Munich, 81377 Munich, Germany. [2]DZHK (German Centre for Cardiovascular Research), partner site Munich Heart Alliance, 80802 Munich, Germany. [3]Department of Proteomics and Signal Transduction, Max Planck Institute of Biochemistry, 82152 Martinsried, Germany. [4]Omicera Diagnostics, 82152 Martinsried, Germany. [5]Institute for Cardiovascular and Metabolic Research, School of Biological Sciences, Health and Life Sciences Building, University of Reading, RG6 6UR, UK. [6]Anesthesiology and Transfusion Medicine, Cell Therapeutics and Hemostaseology, Ludwig-Maximilians-University Munich, 80336 Munich, Germany. [7]Norwegian Institute for Nature Research, 7034 Trondheim, Norway. [8]Scandinavian Brown Bear Research Project, Tackåsen 2, SE-79498 Orsa, Sweden. [9]Department of Forestry and Wildlife Management, Faculty of Applied Ecology and Agricultural Sciences, Inland Norway University of Applied Sciences, 2480 Koppang, Norway. [10]Zentrum für Rückenmarkverletzte mit Neuro-Urologie, BG Unfallklinik Murnau, 82418 Murnau am Staffelsee, Germany. [11]Neurologische Klinik und Poliklinik, Klinikum der Universität München, Ludwig-Maximilians- University Munich, 81377 Munich, Germany. [12]Pathologisches Institut, Klinikum der Universität München, Ludwig-Maximilians- University Munich, 81377 Munich, Germany. [13]Gene Center, Ludwig-Maximilians-University Munich, 81377 Munich, Germany. [14]Institute of Pharmacology and Toxicology, Technical University of Munich, 80802 Munich, Germany. [15]Herlev and Gentofte University Hospital, Borgmester Ib Juuls Vej 1, DK-2730, Herlev, Copenhagen, Denmark. [16]Institute of Veterinary Pathology, Center for Clinical Veterinary Medicine, LMU Munich, 80539 Munich, Germany. [17]Division of Cardiology, Pulmonology, and Vascular Medicine, Heinrich Heine University Medical Center Dusseldorf, 40225 Dusseldorf, Germany. [18]Institute of Aerospace Medicine, German Aerospace Center (DLR), 51147 Cologne, Germany. [19]Department of Anesthesiology and Intensive Care Medicine, Merheim Medical Center, Hospitals of Cologne, University of Witten/Herdecke, 51109 Cologne, Germany. [20]Faculty of Health, Department of Cardiology, Örebro University, 701 85 Örebro, Sweden. [21]Department of Clinical Medicine, Faculty of Health, Aarhus University, 8000 Aarhus, Denmark. [22]Department of Clinical Pharmacology, Aarhus University Hospital, 8000 Aarhus, Denmark. [23]Steno Diabetes Center Aarhus, Aarhus University Hospital, 8000 Aarhus, Denmark.

*Corresponding author. Email: tobias.petzold@med.uni-muenchen.de (T.P.); steffen.massberg@med.uni-muenchen.de (S.M.); mmann@biochem.mpg.de (M.M.); ole.frobert@regionorebrolan.se (O.F.)

†These authors contributed equally to this work.

Population-based studies have associated plasma coagulation factor activities with VTE risk in humans (*16*, *17*). When characterizing alterations within the intrinsic coagulation pathway of the bears in more detail, we found a prolonged plasma clotting time but with values that were still within the human reference range (*18*) (fig. S2B). Further, single coagulation factor (factor VIII, factor IX, factor XI, factor XII) function was also only mildly reduced during winter (fig. S2C). Mass spectrometry (MS)-based plasma proteome profiling (*19*) showed no major differences in coagulation-related plasma proteins between active state and hibernation. Instead, plasma proteome differences were dominated by enzymes that reflected a seasonal switch in metabolism, dominating intra-individual differences (fig. S2, D to F). Seasonal alterations in coagulation factors therefore are not the main feature responsible for VTE protection during hibernation. Instead, thromboelastography revealed that the seasonal differences in clot formation depend on platelets and disappear after platelet inhibition (FIBTEM, Fig. 2B). Of note, plasma levels of the physiological endothelial cell–derived platelet inhibitor prostacyclin were reduced during winter (fig. S2G). This prompted us to focus on the characterization of the platelet population as the central blood component regulating physiological hemostasis and pathological thrombosis.

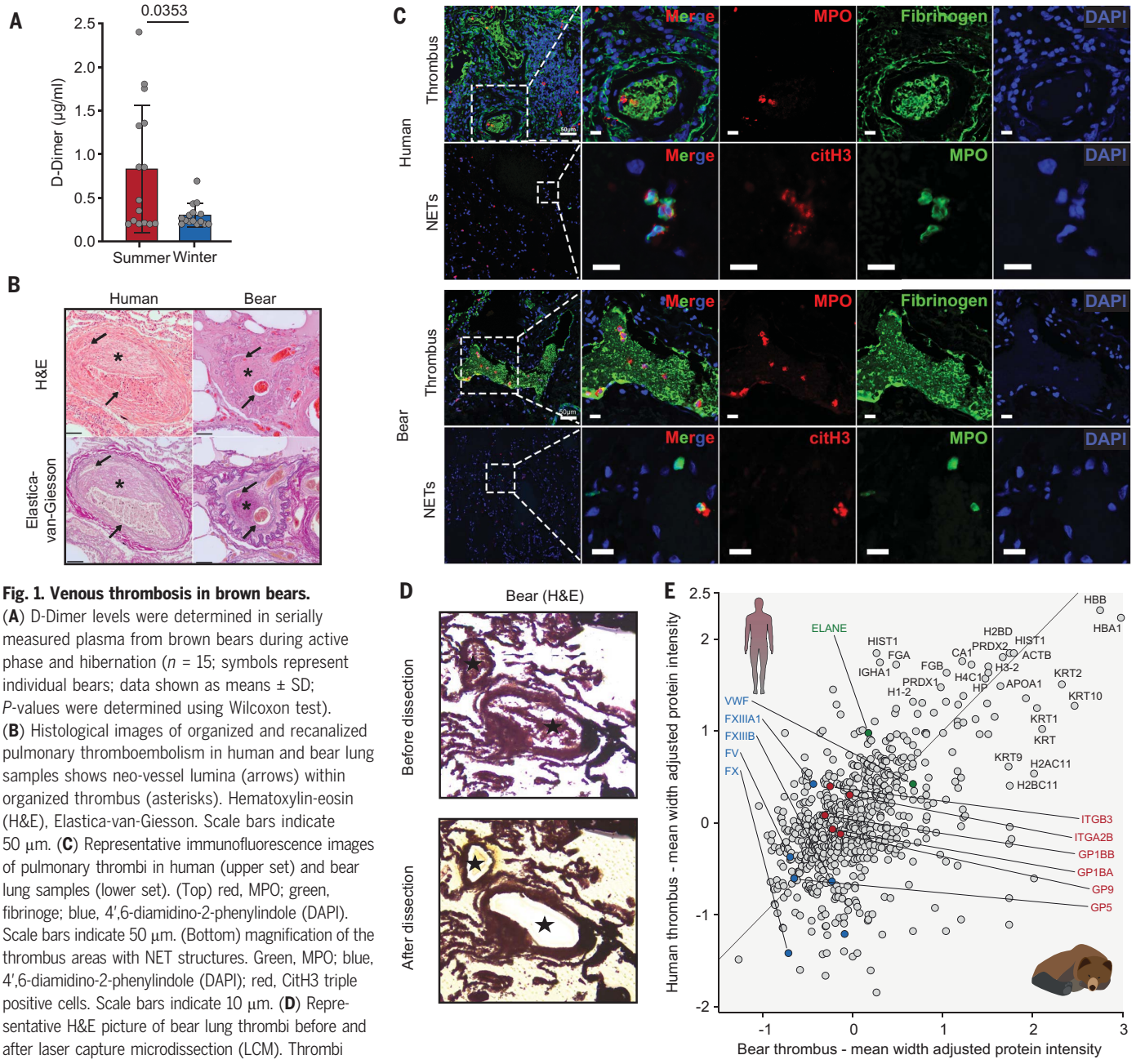
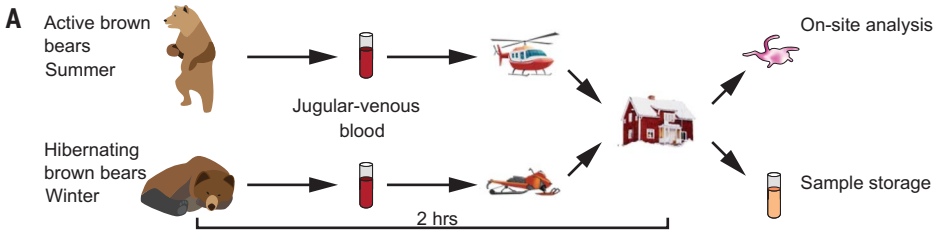
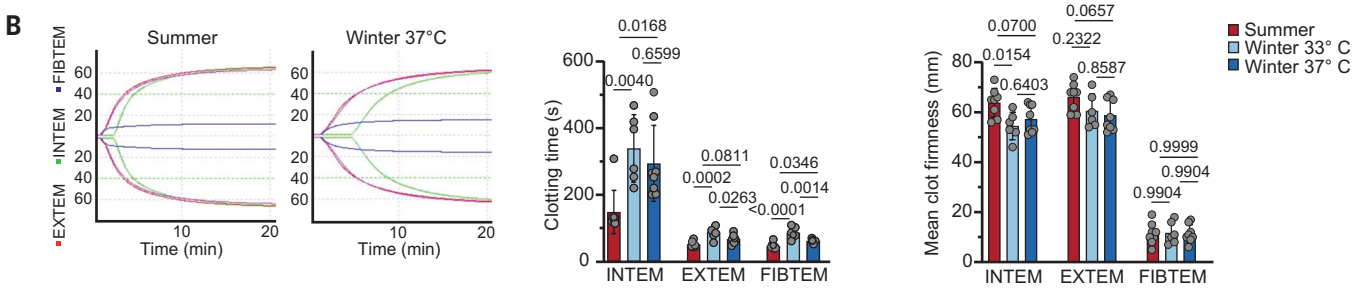
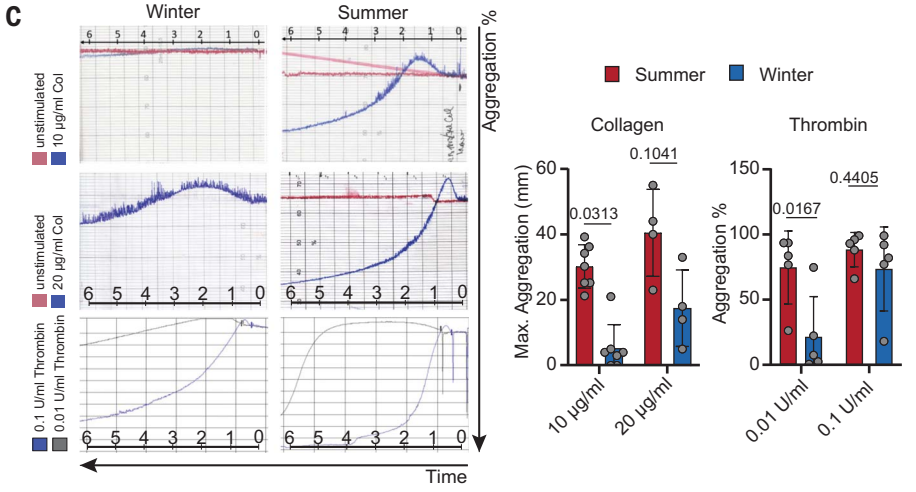
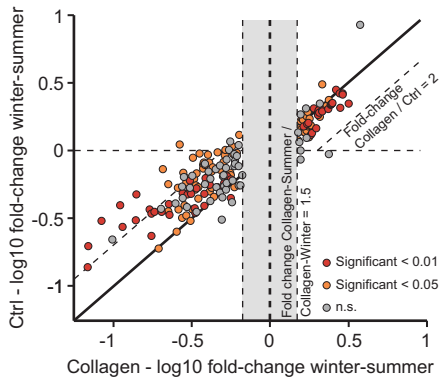


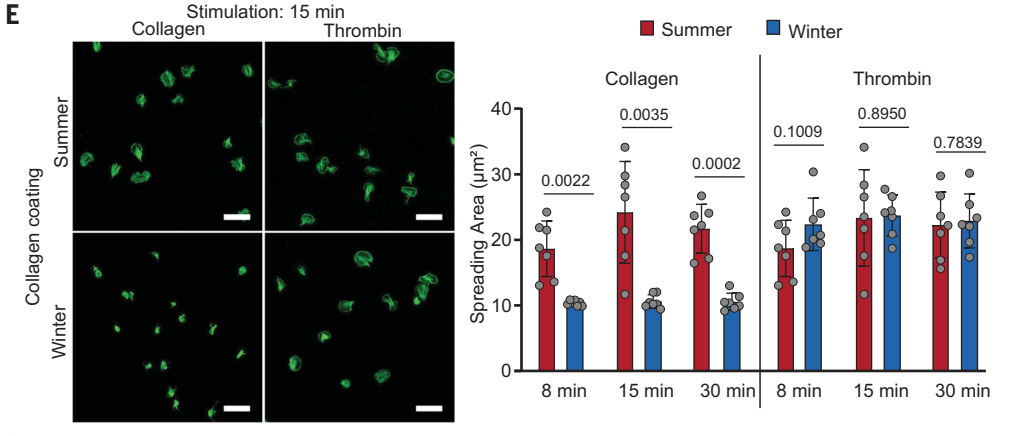
Fig. 1. Venous thrombosis in brown bears. (**A**) D-Dimer levels were determined in serially measured plasma from brown bears during active phase and hibernation (n = 15; symbols represent individual bears; data shown as means ± SD; P-values were determined using Wilcoxon test). (**B**) Histological images of organized and recanalized pulmonary thromboembolism in human and bear lung samples shows neo-vessel lumina (arrows) within organized thrombus (asterisks). Hematoxylin-eosin (H&E), Elastica-van-Giesson. Scale bars indicate 50 µm. (**C**) Representative immunofluorescence images of pulmonary thrombi in human (upper set) and bear lung samples (lower set). (Top) red, MPO; green, fibrinoge; blue, 4′,6-diamidino-2-phenylindole (DAPI). Scale bars indicate 50 µm. (Bottom) magnification of the thrombus areas with NET structures. Green, MPO; blue, 4′,6-diamidino-2-phenylindole (DAPI); red, CitH3 triple positive cells. Scale bars indicate 10 µm. (**D**) Representative H&E picture of bear lung thrombi before and after laser capture microdissection (LCM). Thrombi within vessels (asterisks) were cut out and pooled per individual for MS-based proteomics analysis. (**E**) Comparison of average, mean width adjusted protein intensities from bear and human lung thrombi (n = 3 per organism) cut out from H&E stained FFPE tissue slides with LCM. The proteins with highest intensities (gray), proteins from the coagulation cascade (blue), platelet markers (red), and neutrophile specific elastase (green) are highlighted with gene names.

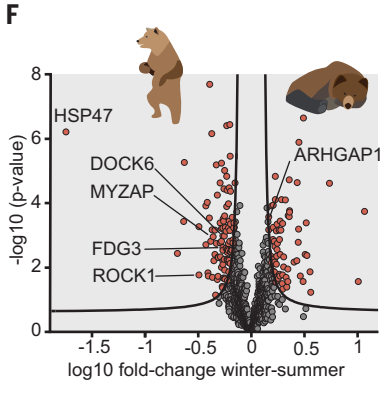

A
Active brown bears Summer
Jugular-venous blood
Hibernating brown bears Winter
2 hrs
On-site analysis
Sample storage

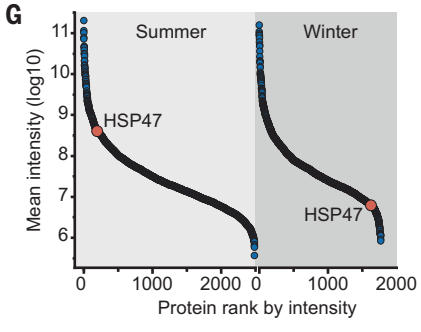

B
Summer
Winter 37°C
FIBTEM
INTEM
EXTEM
Time (min)
Clotting time (s)
Mean clot firmness (mm)
INTEM
EXTEM
FIBTEM
Summer
Winter 33° C
Winter 37° C

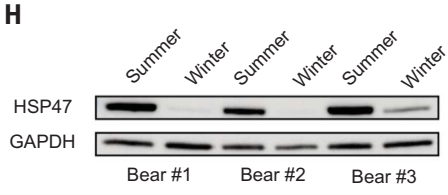

C
Winter
Summer
unstimulated
10 µg/ml Col
20 µg/ml Col
0.1 U/ml Thrombin
0.01 U/ml Thrombin
Aggregation %
Time
Summer
Winter
Collagen
Thrombin
Max. Aggregation (mm)
Aggregation %
10 µg/ml
20 µg/ml
0.01 U/ml
0.1 U/ml


D
Ctrl - log10 fold-change winter-summer
Collagen - log10 fold-change winter-summer
Fold-change Collagen / Ctrl = 2
Fold change Collagen-Summer / Collagen-Winter = 1.5
Significant < 0.01
Significant < 0.05
n.s.


E
Stimulation: 15 min
Collagen
Thrombin
Summer
Winter
Collagen coating
Spreading Area (µm²)
8 min
15 min
30 min


F
HSP47
DOCK6
MYZAP
FDG3
ROCK1
ARHGAP1
-log10 (p-value)
log10 fold-change winter-summer


G
Summer
Winter
HSP47
HSP47
Mean intensity (log10)
Protein rank by intensity


H
Summer
Winter
Summer
Winter
Summer
Winter
HSP47
GAPDH
Bear #1
Bear #2
Bear #3

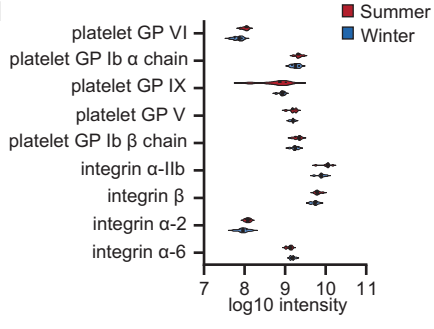

I
Summer
Winter
platelet GP VI
platelet GP Ib α chain
platelet GP IX
platelet GP V
platelet GP Ib β chain
integrin α-IIb
integrin β
integrin α-2
integrin α-6
log10 intensity

Fig. 2. Platelet function in hibernating brown bears. (**A**) Field study scheme is shown. (**B**) Quantification of clotting time and mean clot firmness using thromboelastography after INTEM, EXTEM, and FIBTEM activation. Differences between blood from summer (n = 8), cool blood from winter (33°C; n = 6), and warmed blood from winter (37°C; n = 8) were analyzed. Symbols represent individual bears; data shown as means ± SD; P-values were determined with one-way analysis of variance (ANOVA) test. (**C**) Light transmission aggregometry of washed platelets from winter and summer bears was performed after stimulation with the indicated agonists. Maximum aggregation after 6 minutes was assessed and representative aggregation curves are shown (n = 5 for thrombin 0.01 U/ml and thrombin 0.1 U/ml, n = 4 for collagen 20 μg/ml, n = 7 for collagen 10 μg/ml; symbols represent individual bears; data shown as means ± SD; P-values were determined with paired t-test for collagen 20 μg/ml and thrombin 0.1 U/ml; Wilcoxon test for collagen 10 μg/ml and unpaired t-test for thrombin 0.01 U/ml). (**D**) Secreted proteins from bear platelets (winter and summer) were analyzed upon collagen stimulation or control treatment with MS-based proteomics. Proteins changing significantly (FDR 5%) in one of the two group comparisons are highlighted in colors and only proteins with a fold change >1.5 times for the collagen stimulation are shown. Dashed diagonal lines mark the 2 times fold change threshold between x and y axis comparison. (**E**) Spreading size of individual platelets stimulated with the indicated agonists on soluble collagen I–coated surfaces was analyzed after 8, 15, and 30 minutes (n = 7; symbols represent individual bears; data shown as means ± SD; P-values were determined with paired t-test). Representative images show spreading after 15 minutes; scale bar represents 20 μm. (**F**) Statistical analysis of the platelet proteome of hibernating and active bear platelets reveals HSP47 as the most changed protein (fold change of 55.26). Changed proteins involved in ROCK1 signaling are marked with gene names. Proteins significantly changing (FDR < 0.05) are highlighted in red in an unpaired analysis of eight individuals' platelets during hibernation phase and activity. (**G**) Proteomes of bear platelets ranked by mean protein intensity of the eight individuals for summer and winter. The rank of HSP47 changes from 184 (summer) to 1624 (winter). (**H**) Representative immunoblots of three individual bears during winter and summer for HSP47 in platelets are shown. (**I**) Group expression (MS-based proteomics) of a panel of platelet surface activation markers is shown as violin plots, which reveal an absence of regulation between the summer and winter samples.

Platelet and neutrophil counts were reduced (by 37.9 and 20.9%, respectively) in winter compared with summer (fig. S2A). The changes in platelets and neutrophils were mild compared with other well-studied hibernating animals such as hamsters and other rodents in which a substantial decrease in body temperature below 10°C leads to profound sequestering of platelets in the liver, neutrophil margination to the vessel wall, and lymphocyte storage in secondary lymphoid organs (*20*–*22*). To analyze seasonal platelet function on a single cell type level adjusted for the differences in platelet counts, we freshly isolated washed platelets. Compared with summer, in winter platelet aggregation was delayed and reduced upon stimulation with fibrillar collagen and low, but not high, doses of thrombin (Fig. 2C) whereas adenosine diphosphate (ADP), Thromboxane A analog U46619, and thrombin receptor activating peptide (TRAP) induced platelet-shape change but failed to trigger platelet aggregation (fig. S2H). Platelet aggregation is accompanied by the release of the platelet granule compartment, fostering aggregation by providing auto- and paracrine-acting platelet agonists and fibrinogen as cross-linking substrates (*23*, *24*). Using MS-based proteomics we found attenuated platelet protein release in response to collagen during winter compared with summer (Fig. 2D and fig. S2I). Spreading of individual collagen-activated platelets on soluble collagen and fibronectin, the substrates of α2β1 and α5β1 integrin adhesion receptors, was reduced in winter compared with summer. By contrast, we did not detect substantial seasonal differences in platelet spreading in response to thrombin or on fibrinogen, the substrate of the most abundantly expressed platelet integrin $\alpha IIb\beta_3$ (Fig. 2E and fig. S3, A to C). These data show that collagen- and thrombin-induced platelet signaling during hibernation is attenuated, affecting distinct platelet responses.

To elucidate the molecular basis of reduced platelet reactivity toward collagen and thrombin during inactivity, we again chose MS-based proteomics as it only requires that the species has a sequenced genome; no active transcription and translation or available antibodies are necessary as in affinity-based protein identification (*25*, *26*). This technology has successfully been applied to biomarker discovery on body fluids and derived cell types, including platelets (*27*, *28*). Despite poor annotation of the bear genome, we identified and quantified more than 2678 proteins. Platelet proteomes are markedly different between hibernating and active brown bears, with the abundance of 151 proteins (5.6%) significantly [False Discovery Rate (FDR) < 5%] changing with an absolute average of 2.1-fold. Of these, 80 were expressed at lower levels and 71 at higher levels during inactivity (Fig. 2F). The gene ontology (GO) terms associated with catabolism such as proteolysis, peptidase activity, and proteasome complex were enriched in the more highly expressed platelet proteins during inactivity, whereas no specific term enrichment was seen in the less-expressed proteins (FDR < 0.02). Focusing on proteins involved in platelet activation, we found decreased amounts of ROCK1 (Rho-associated coiled-coil-containing protein kinase 1) during inactivity. ROCK1 is regulated by the small Rho-GTPase RhoA, and two of its activating guanidine exchange factors, FYVE and DOCK6, were also down-regulated, whereas ARHGAP1, a deactivator, was up-regulated. These findings already indicated a seasonal functional conversion of the platelet proteome toward lower activity. However, the platelet protein with the largest difference between hibernating and active bears was the gene product of serine proteinase inhibitor H1 (SERPINH1), also known as heat shock protein 47 (HSP47). On average, HSP47 was 55-fold down-regulated in the platelets of hibernating bears, 10 times more than the next most down-regulated protein (RRBP1, fivefold change) (Fig. 2, F to I).

Platelet HSP47 regulates thromboinflammation

HSP47 belongs to the SERPIN superfamily but possesses no serine protease inhibitory activity. In a fibroblast context it acts as an endoplasmic reticulum resident chaperone, specifically facilitating collagen folding and secretion into the extracellular space. In different fibrotic diseases and skeletal malformation pathologies, HSP47 acts as a key disease driver (*29*). Recently, it has been detected on the platelet membrane, promoting collagen signaling together with GPVI and integrin α2β1, the two major collagen receptors in experimental arterial thrombosis and stroke (*30*–*33*). Given its strong regulation and cross-species conservation, we decided to functionally investigate the potential protective effect of HSP47 down-regulation in the context of VTE.

To define the functional role of platelet HSP47 in VTE in vivo, we generated *PF4-Cre(+)/HSP47*$^{-/-}$ or *PF4-Cre(−)/HSP47*$^{fl/fl}$ bone marrow chimeric mice and subjected them to venous flow restriction (Fig. 3A and fig. S4, A and B). HSP47 is abundantly expressed by platelets but not leukocytes. *PF4-Cre* mediated recombination in chimeric mice therefore leads to specific deletion of HSP47 in the platelet compartment (Fig. 3B and fig. S4C). Thrombus frequency and sizes in *PF4-Cre(+)/HSP47*$^{-/-}$ chimeric mice were drastically reduced compared with controls (Fig. 3C). At the same time, NETs were reduced in *PF4-Cre(+)/HSP47*$^{-/-}$ compared with control mice (Fig. 3D). As patients with VTE have increased plasma levels of NET markers, we also quantified circulating citH3/DNA complexes (Fig. 3E) as well as circulating DNA

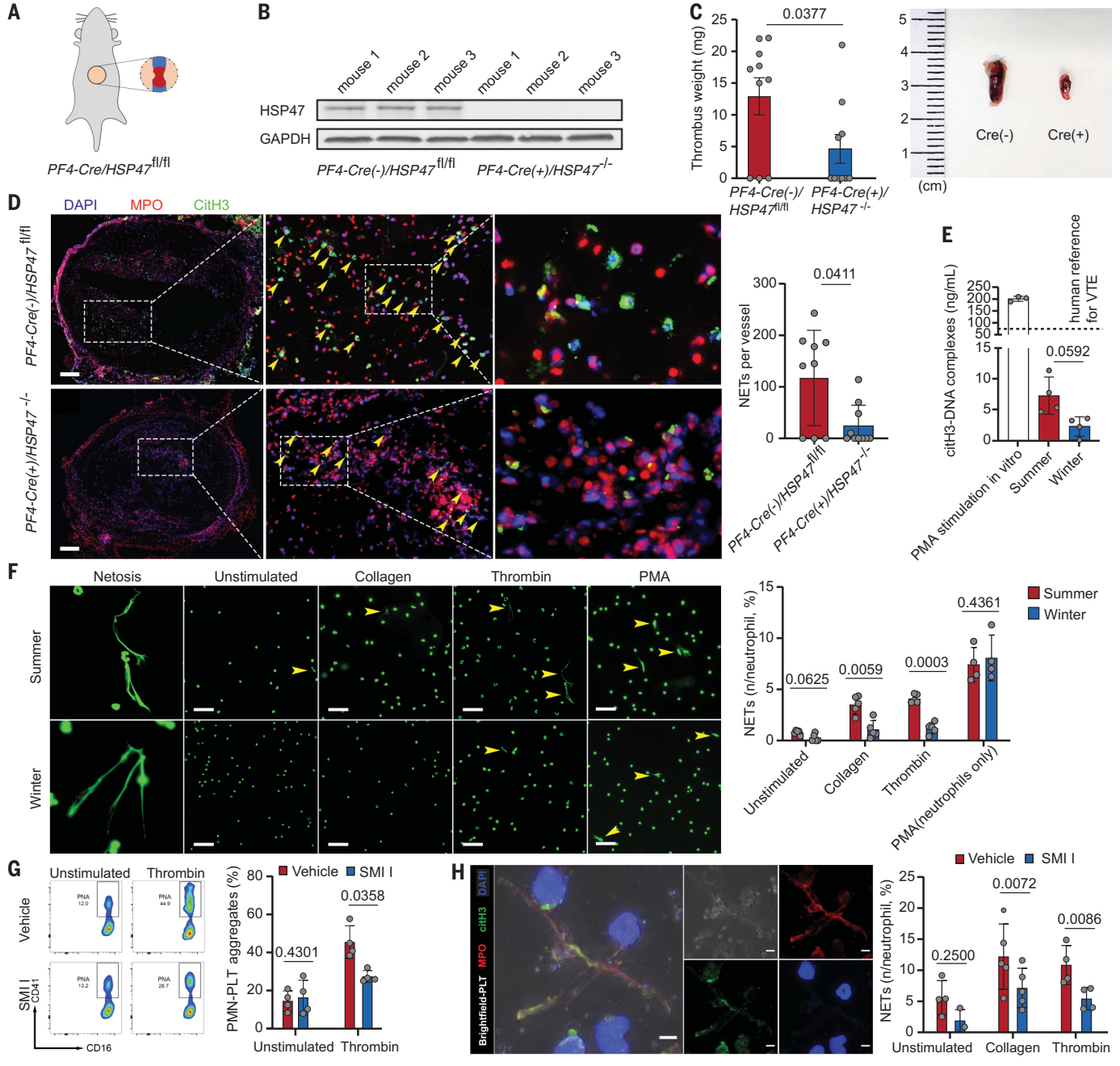


Fig. 3. Antithrombotic mechanism under HSP47 inhibition. (**A**) Scheme of a vena cava thrombosis model. (**B**) Immunoblots for HSP47 in platelets of individual HSP47 knockout mice and littermate controls are shown. (**C**) Thrombus weight was analyzed after 48 hours (n = 10; symbols represent individual mice; data shown as means ± SD; P-values were determined using unpaired t-test). (**D**) NETs in explanted thrombi of HSP47-deficient mice and control animals were analyzed with confocal microscopy. Scale bar represents 200 μm. Magnifications of thrombus areas with NET structures [MPO- 4′,6-diamidino-2-phenylindole (DAPI) -CitH3 triple positive cells] are shown on the right. Quantitative comparison of NETs per vessel in the HSP47-deficient mice (*PF4-Cre(+)/HSP47*$^{-/-}$] and littermate controls [*PF4-Cre(−)/HSP47*$^{fl/fl}$] is shown as a bar plot on the right (n = 10; data shown as means ± SD; P-values were determined with Mann-Whitney test). (**E**) Levels of citH3-DNA complexes in paired hibernating and active bears. Dashed line indicates human reference in healthy controls. Supernatant from PMA-stimulated bear neutrophils was used as positive control (n = 4 for winter and summer bears, n = 3 for PMA-stimulated neutrophils; symbols represent individual bears; data shown as means ± SD; P-values were determined with paired t-test). (**F**) Quantification of NETs per neutrophil in bears after stimulation of platelets with the indicated agonists (n = 5 per group; data shown as means ± SD; P-values were determined with paired t-test for thrombin, collagen, and PMA; Wilcoxon test was used for unstimulated). Scale bar represents 50 μm; magnification of NET structure is shown on the left. (**G**) In vitro formation of human PMN-platelet aggregates in the presence or absence of HSP47 inhibitor Col003 (SMI I) after platelet stimulation with the indicated agonists. Representative flow-cytometric contour plots are shown (n = 4; symbols represent independent experiments; data shown as means ± SD; P-values were determined using paired t-test). (**H**) NET formation was analyzed using confocal microscopy following in vitro co-culture of human neutrophils and platelets after stimulation with the indicated agonists following HSP47 inhibition (SM I; symbols represent independent experiments; data shown as means ± SD; P-values were determined using Wilcoxon test for unstimulated and paired t-test for collagen and thrombin). Scale bar represents 5 μm.

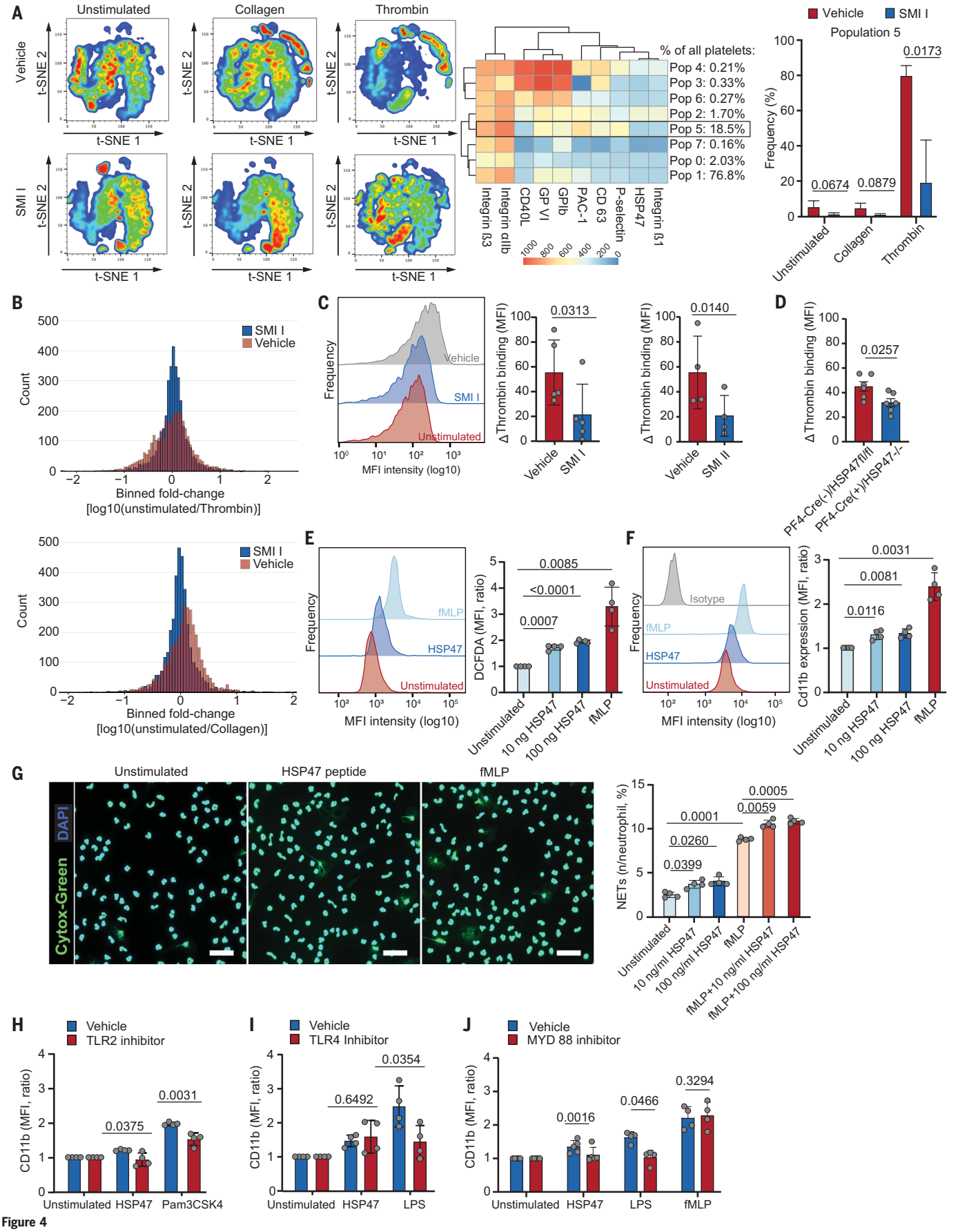

A
Unstimulated
Collagen
Thrombin
Vehicle
SMI I
t-SNE 1
t-SNE 2
% of all platelets:
Pop 4: 0.21%
Pop 3: 0.33%
Pop 6: 0.27%
Pop 2: 1.70%
Pop 5: 18.5%
Pop 7: 0.16%
Pop 0: 2.03%
Pop 1: 76.8%
Integrin β3
Integrin αIIb
CD40L
GP VI
GPIb
PAC-1
CD 63
P-selectin
HSP47
Integrin β1
Population 5
Frequency (%)
0.0674
0.0879
0.0173
B
Count
Binned fold-change
[log10(unstimulated/Thrombin)]
[log10(unstimulated/Collagen)]
C
Frequency
MFI intensity (log10)
Δ Thrombin binding (MFI)
0.0313
0.0140
D
0.0257
PF4-Cre(-)/HSP47fl/fl
PF4-Cre(+)/HSP47-/-
E
fMLP
HSP47
DCFDA (MFI, ratio)
0.0085
<0.0001
0.0007
Unstimulated
10 ng HSP47
100 ng HSP47
fMLP
F
Isotype
Cd11b expression (MFI, ratio)
0.0031
0.0081
0.0116
G
DAPI
Cytox-Green
HSP47 peptide
NETs (n/neutrophil, %)
0.0005
0.0059
0.0001
0.0260
0.0399
10 ng/ml HSP47
100 ng/ml HSP47
fMLP+10 ng/ml HSP47
fMLP+100 ng/ml HSP47
H
TLR2 inhibitor
CD11b (MFI, ratio)
0.0375
0.0031
Pam3CSK4
I
TLR4 Inhibitor
0.6492
0.0354
LPS
J
MYD 88 inhibitor
0.0016
0.0466
0.3294

Figure 4

Fig. 4. HSP47 regulates thromboinflammation. (**A**) Multicolor flow cytometry on platelets following activation with indicated stimuli in the presence or absence of HSP47 inhibitor (SMI I) is shown. T-distributed stochastic neighbor embedding (*t*-SNE) of platelet heterogeneity is displayed by Flow-SOM plots. Heatmap of the mean fluorescence intensity (MFI) for each platelet subcluster relative to the maximum MFI of the surface marker is shown. Frequency of population 5 following SMI I or vehicle treatment is shown (n = 4; bars show means ± SD; P-values were determined using paired t-test). (**B**) Changing peptides from phospho-proteome analysis of activated platelets under SMI I treatment compared with vehicle treatment after collagen/thrombin stimulation. The phospho-proteome of six individuals was compared by independent t-test and average fold changes of all peptides are displayed in the histograms. (**C**) Thrombin binding to the surface of platelets from healthy control patients was analyzed using flow cytometry after addition of two different HSP47 small molecule inhibitors [n = 5 for SMI I (Col003), n = 4 for SMI II (RH00007SC)]. Symbols represent individuals; data shown as means ± SD; P-values were determined using paired t-test. (**D**) Thrombin binding to the surface of platelets from HSP47-deficient mice (*PF4-Cre(+)/HSP47*$^{-/-}$) and littermate controls [*PF4-Cre(−)/HSP47*$^{fl/fl}$] was examined. Symbols represent individuals; data shown as means ± SD; P-values were determined using unpaired t-test. (**E**) ROS production and (**F**) CD11b surface expression on neutrophils upon stimulation with purified HSP47 peptide and fMLP. Bars show means ± SD; symbols represent individual experiments, n = 4; P-values were determined using paired t-test. (**G**) Percentage of NET-forming neutrophils was analyzed after stimulation with HSP47 peptide or co-stimulation with fMLP and HSP47 peptide (n = 4; bars show means ± SD; P-values were determined using paired t-test). Scale bar represents 30 μm. (**H**) Expression of CD11b (MFI ratio to unstimulated) in neutrophils upon stimulation with the indicated agonists in presence or absence of TLR2 inhibitor. Bars show means ± SD; symbols represent individual experiments, n = 4; P-values were determined using paired t-test. (**I**) Expression of CD11b (MFI ratio to unstimulated) in neutrophils upon stimulation with the indicated agonists in presence or absence of TLR4 inhibitor. Bars show means ± SD; symbols represent individual experiments, n = 4; P-values were determined using multiple t-test. (**J**) CD11b surface expression (MFI ratio to unstimulated) in neutrophils upon stimulation with the indicated agonists in presence or absence of MyD88 inhibitor. Bars show means ± SD; symbols represent individual experiments, n = 5 for unstimulated and HSP47; n = 4 for LPS and fMLP; P-values were determined using paired t-test.

(fig. S4D) in bears (*34*–*36*). Along with Hsp47 down-regulation and D-dimer levels (Fig. 1A), both markers were reduced during hibernation. Consistent with these findings we observed reduced prothrombotic NET formation in an in vitro co-culture model of neutrophils and platelets isolated from hibernating compared with active brown bears whereas NET formation of PMA-stimulated neutrophils was unaffected (Fig. 3F). When we pretreated human platelets with two HSP47-specific small molecule inhibitors (SMI; Col003, SMI II; RH00007SC) (*30*, *37*), platelet-neutrophil aggregates were reduced upon stimulation with thrombin (Fig. 3G and fig. S4E) and prothrombotic NET formation was attenuated (Fig. 3H and fig. S4F).

Activated platelets fuel NET formation by release of granule content and direct physical interactions with neutrophils (*38*, *39*). To dissect the functional role of HSP47 in both processes, we first quantified platelet activation and granule release by multicolor flow cytometry. SMI I treatment reduced the fraction of activated platelets that express high levels of granule secretion markers (CD62P^{high}, CD63^{high}) as well as active integrin αIIbβ3 (PAC-1^{high}) and attenuated collagen platelet aggregation (t-SNE cluster population 5; Fig. 4A and fig. S5, A and C). In phospho-proteome analysis of aggregation assays, phosphorylation patterns were sensitive to HSP47 inhibition with SMI I (Fig. 4B and fig. S5, D to G) as well as SMI II (fig. S5H). Pharmacological inhibition of HSP47 with SMIs almost completely abolished binding of thrombin to the platelet surface (Fig. 4C). We also found reduced thrombin binding to HSP47-deficient mouse platelets (Fig. 4D), suggesting that HSP47 regulates thrombin recruitment to platelet surfaces.

Next, we investigated whether the interaction of HSP47 with neutrophils affects their function. Flow cytometric analysis of neutrophils treated with purified HSP47 revealed increased intracellular reactive oxygen species production and increased leukocyte integrin CD11b expression, only somewhat less efficient than N-Formylmethionine-leucyl-phenylalanine (fMLP), an established highly potent neutrophil activator (Fig. 4, E and F). Direct or co-stimulation with HSP47 augmented NET formation even in the absence of platelets (Fig. 4G). To further elucidate the molecular mechanism by which HSP47 activates neutrophils, we analyzed toll-like receptor (TLR) signaling, which is known to be induced by other heat shock proteins. Indeed, inhibition of TLR2 but not TLR4 reduced CD11b expression in response to HSP47 peptide (Fig. 4, H and I). In line with this, we found reduced CD11b expression after inhibition of MyD88 by Pepinh-MYD peptide, a canonical adaptor downstream of TLR2 and -4 signaling pathway, after stimulation with recombinant HSP47 (Fig. 4J). In summary, we identified a mechanism by which HSP47 facilitates thrombin binding to platelet surfaces as well as neutrophil activation through the TLR2-MyD88 signaling axis.

HSP47 down-regulation in immobile patients

We addressed the relevance of thromboprotective HSP47 down-regulation in scenarios beyond hibernation. Chronically paralyzed patients with spinal cord injury (SCI) exhibit a VTE risk similar to the general population for unknown reasons despite long term immobilization (*5*–*7*). We therefore recruited chronically immobilized patients who were either bedridden or wheelchair-bound as a result of SCI and matched healthy controls (Fig. 5A and table S3). HSP47 levels before SCI were unknown, but platelet HSP47 expression was reduced in SCI compared with matched healthy controls, parallel to the hibernating bear population. The expression of other key platelet receptors (GPVI and GPIbα) remained unchanged, indicating a specific response (Fig. 5B and fig. S6A). Immunoblotting and confocal fluorescence microscopy confirmed reduced protein expression whereas HSP47 cellular distribution did not differ (Fig. 5, C and D). Coagulation and whole-blood thrombus formation was unchanged between patients and controls (fig. S6, B and C). Circulating DNA and citrullinated histone 3 levels—biomarkers of NET formation—were also similar, indicating absence of a prothrombotic state in chronically immobilized patients (fig. S6D). Consistent with our findings regarding human platelets treated with HSP47 SMI, thrombin binding to platelets was reduced in SCI patients (Fig. 5E). Quantifying agonist-induced platelet activation indicated normal aggregation responses (fig. S6E) whereas platelet granule release and integrin activation (i.e., CD62P and PAC-1) were attenuated in chronic immobilization (Fig. 5F). The reduced HSP47 levels on SCI platelets were associated with attenuated neutrophil platelet interactions and NETosis compared with healthy controls (Fig. 5, G to I). A merged analysis of both groups revealed a direct correlation between HSP47 expression and platelet neutrophil aggregates or NETs (Fig. 5J), indicating that HSP47 expression correlates with thrombogenicity in immobilized patients. When we analyzed *HSP47* mRNA in platelets, we found reduced *HSP47* transcripts in SCI patients compared with healthy controls, indicating that HSP47 regulation occurs at the megakaryocyte (MK) level (Fig. 5K). In patients acutely immobilized with intracranial bleeding (table S4), the kinetics of platelet HSP47 down-regulation—but not GPVI or GPIbα

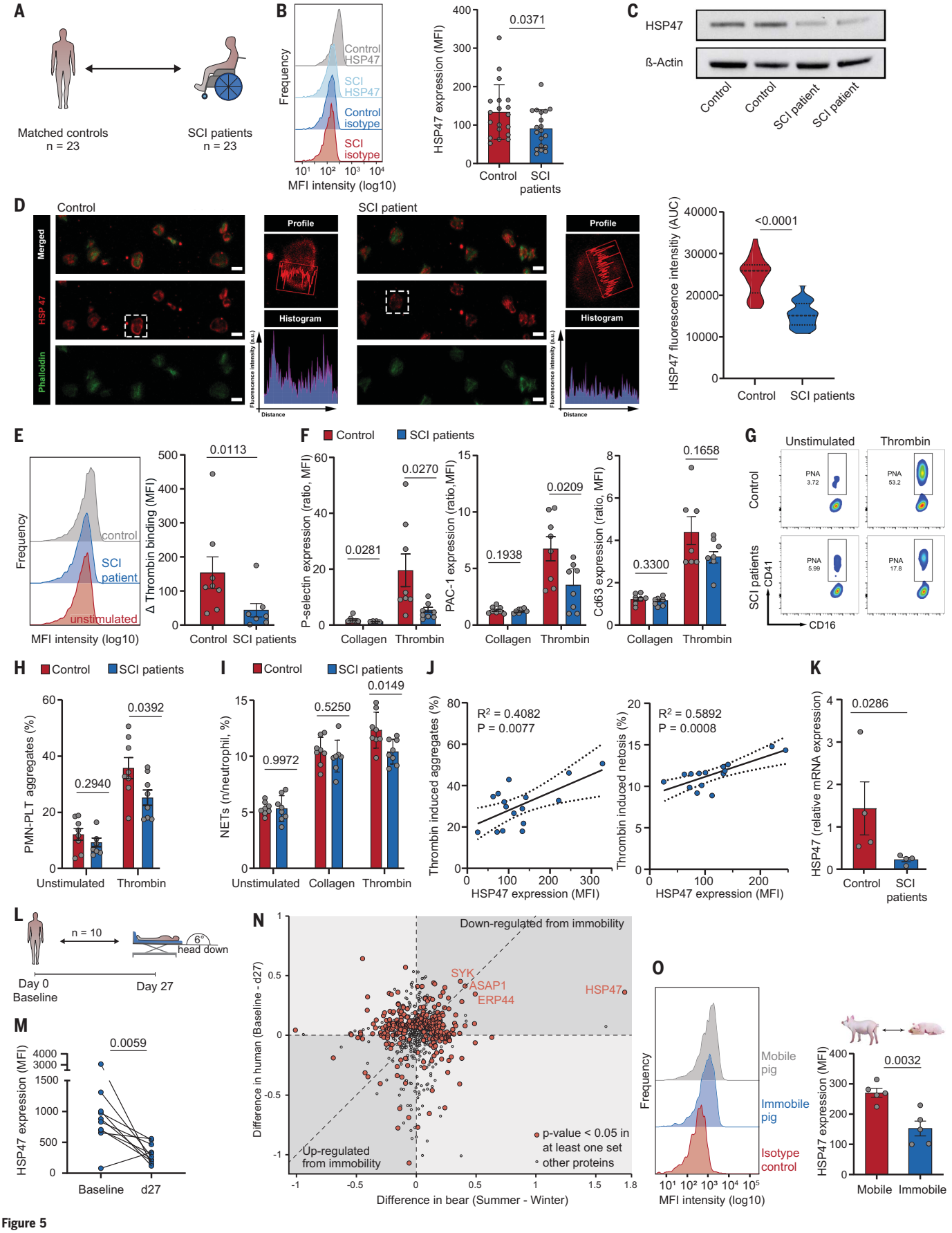

A
Matched controls
n = 23
SCI patients
n = 23
B
Frequency
Control HSP47
SCI HSP47
Control isotype
SCI isotype
MFI intensity (log10)
HSP47 expression (MFI)
0.0371
Control
SCI patients
C
HSP47
ß-Actin
Control
Control
SCI patient
SCI patient
D
Control
SCI patient
Merged
HSP 47
Phalloidin
Profile
Histogram
Fluorescence intensity (a.u.)
Distance
HSP47 fluorescence intensity (AUC)
<0.0001
Control
SCI patients
E
Frequency
control
SCI patient
unstimulated
MFI intensity (log10)
Δ Thrombin binding (MFI)
0.0113
Control
SCI patients
F
Control
SCI patients
P-selectin expression (ratio, MFI)
0.0281
0.0270
PAC-1 expression (ratio,MFI)
0.1938
0.0209
Cd63 expression (ratio, MFI)
0.3300
0.1658
Collagen
Thrombin
G
Unstimulated
Thrombin
Control
SCI patients
CD41
CD16
PNA
3.72
53.2
5.99
17.8
H
Control
SCI patients
PMN-PLT aggregates (%)
0.2940
0.0392
Unstimulated
Thrombin
I
Control
SCI patients
NETs (n/neutrophil, %)
0.9972
0.5250
0.0149
Unstimulated
Collagen
Thrombin
J
Thrombin induced aggregates (%)
R² = 0.4082
P = 0.0077
HSP47 expression (MFI)
Thrombin induced netosis (%)
R² = 0.5892
P = 0.0008
HSP47 expression (MFI)
K
HSP47 (relative mRNA expression)
0.0286
Control
SCI patients
L
n = 10
6°
head down
Day 0
Baseline
Day 27
M
HSP47 expression (MFI)
0.0059
Baseline
d27
N
Down-regulated from immobility
SYK
ASAP1
ERP44
HSP47
Up-regulated from immobility
p-value < 0.05 in at least one set
other proteins
Difference in human (Baseline - d27)
Difference in bear (Summer - Winter)
O
Frequency
Mobile pig
Immobile pig
Isotype control
MFI intensity (log10)
HSP47 expression (MFI)
0.0032
Mobile
Immobile

Figure 5

Fig. 5. Thrombus protection by HSP47 down-regulation. (**A**) Schematic view of the recruited SCI cohort and healthy controls. (**B**) Platelet expression of HSP47 (MFI) in healthy controls and patients with SCI. Representative flow cytometry histograms are shown (n = 19; data shown as means ± SD; P-values were determined using Mann-Whitney test). (**C**) Representative immunoblot for HSP47 in human platelets from immobile patients and matched healthy controls. (**D**) Immunofluorescence microscopy of HSP47 (red) and actin (green) was performed in platelets from patients with SCI and healthy controls. HSP47 fluorescence intensity was quantified [violin plot shows median (thick dotted line) and interquartile range (thin dotted lines); cells from three independent experiments were analyzed]. Representative cells are shown in both profile and histogram (MFI). Scale bar indicates 10 μm; P-values were determined using Mann-Whitney test. (**E**) Thrombin binding was analyzed using flow cytometry in patients with SCI (n = 8 for patients with SCI and healthy controls. Symbols represent individual humans; data shown as means ± SD; P-values were determined using Mann-Whitney test. Representative flow cytometry is shown). (**F**) Quantification of platelet activation marker expression (MFI ratios to unstimulated) by flow cytometry following stimulation with the indicated agonists (bars represent means ± SD; symbols represent single SCI patients or healthy controls; n = 8 per group; P-values were determined using unpaired t-test for P-selectin stimulation with thrombin, PAC-1, CD63; Mann-Whitney was used for P-selectin stimulation with collagen). (**G**) Quantification of platelet-PMN aggregates following platelet stimulation with thrombin. Representative flow-cytometric contour plots from eight independent experiments are shown. (**H**) Quantification of human PMN-platelet aggregates formation after platelet stimulation with thrombin in vitro (symbols represent individual SCI patients and healthy controls n = 8 per group; bars show mean ± SD: P-values were determined using unpaired t-test). (**I**) Quantification of NETs per neutrophil after stimulation of platelets with indicated agonists (bars represent means ± SD; symbols represent single patients or healthy controls; n = 8 per group; P-values were determined using unpaired t-test). (**J**) Pooled (patients and healthy controls; n = 16) correlation analysis of HSP47 expression with thrombin induced aggregates or thrombin-induced NETs, given as r^2-values; analysis was done using Pearson correlation. (**K**) *HSP47* mRNA was quantified in platelets from patients with SCI and healthy controls (n = 4). Bars represent means ± SD; symbols represent single patients or healthy controls; P-values were determined using Mann-Whitney test. (**L**) Schematic view of the bed rest cohort analyzed. (**M**) Serial analysis of HSP47 expression at indicated time points is shown (symbols represent single individuals; n = 10 per group; P-values were determined using Wilcoxon test). (**N**) Correlation of protein fold changes from proteomics analysis of the hibernating versus active bears and human individuals before and after 27 days of bedrest. Proteins with a t-test P-value < 0.05 in one of the two sets are highlighted in red. Proteins in the upper right quadrant are co–down-regulated in bears and humans upon immobility (e.g., HSP47, SYK, ASAP1), whereas proteins in the bottom left quadrant are co–up-regulated upon immobility. (**O**) Platelet expression of HSP47 (MFI) in immobile and mobile pigs (n = 5; symbols represent individual pigs; data shown as means ± SD; P-values were determined using unpaired t-test). Representative flow cytometry histograms are shown.

levels—followed platelet turnover, consistent with a regulation of HSP47 expression at the MK level (fig. S6F).

Finally, to study the effect of immobilization in a physiological setting, we analyzed a longitudinal cohort of healthy individuals undergoing voluntary bed rest immobilization (see materials and methods and Fig. 5L). After 27 days of bed rest, platelet HSP47 was drastically reduced whereas GPVI or GPIbα surface expression did not change (Fig. 5M and fig. S6, G to H). In a metaproteome analysis, we performed a bioinformatic matching for homologous proteins between the human and bear proteome (Fig. 5N and table S5). This enabled a cross-species comparison of immobility-induced proteome changes, where we focused on proteins involved in platelet activation. In addition to HSP47 we observed down-regulation of SYK, ASAP1, and Erp44 upon immobility in bears and humans. All of these proteins are known to be involved in platelet activation (*40*–*42*). We conclude that immobility is a trigger for functional platelet proteome conversion in an antithrombotic manner. To explore immobility-induced HSP47 down-regulation in platelets across a wider range of species, we added pigs as another mammalian model. After giving birth to their piglets, lactating pigs were limited in mobility for 21 to 28 days. Compared with free-ranging indoor pigs, platelet HSP47 expression was reduced as detected using flow-cytometry (Fig. 5O). However, mobile pigs were neither lactating nor gave birth within the preceeding weeks, which is a limiting factor that might affect HSP47 expression.

Discussion

Combining functional analyses in free-ranging brown bears, mice, pigs, and humans, we uncovered a cross-species conserved physiological process upon immobilization. Through HSP47 down-regulation, chronic immobilization conveys protection from VTE by reversing the pro-thromboinflammatory state that normally follows immobilization. We therefore suggest that HSP47 is a promising target to prevent VTE in patients exposed to immobilization, that might be evolutionarily conserved across mammals. Our findings provide mechanistic insight into the paradox that the VTE risk of long-term immobilized SCI patients resembles the risk of the general population. Further investigation is needed regarding the molecular basis of mechanosensation during hemostasis and thrombosis in various contexts, including microgravity in space, which can lead to venous thrombosis (*43*, *44*). However, our data also establish a function of HSP47 driving thromboinflammation in the tightly controlled interplay of innate immune cells and platelets in general. Toward clinical application, our findings thus suggest a concept for treating thromboinflammatory cardiovascular diseases by making use of naturally occurring thromboprotective mechanisms.

Immobility is not the single cause of VTE and we did not investigate the role of HSP47 in other thromboembolic environments such as cancer or coagulation disorders. Nevertheless, patients at risk of VTE beyond immobilization might benefit from treatments mimicking the here-described thromboprotective mechanism.

REFERENCES AND NOTES

1. F. Khan, T. Tritschler, S. R. Kahn, M. A. Rodger, *Lancet* **398**, 64–77 (2021).
2. K. Stark, S. Massberg, *Nat. Rev. Cardiol.* **18**, 666–682 (2021).
3. M. L. von Brühl *et al.*, *J. Exp. Med.* **209**, 819–835 (2012).
4. B. Engelmann, S. Massberg, *Nat. Rev. Immunol.* **13**, 34–45 (2013).
5. L. N. Godat, L. Kobayashi, D. C. Chang, R. Coimbra, *J. Trauma Acute Care Surg.* **78**, 475–481 (2015).
6. M. Giorgi Pierfranceschi *et al.*, *Thromb. Haemost.* **109**, 34–38 (2013).
7. R. A. Maxwell *et al.*, *J. Trauma* **52**, 902–906 (2002).
8. O. Fröbert, A. M. Frøbert, J. Kindberg, J. M. Arnemo, M. T. Overgaard, *J. Intern. Med.* **287**, 263–270 (2020).
9. P. G. Jørgensen *et al.*, *Cardiovasc. Ultrasound* **12**, 36 (2014).
10. A. U. Friedrich *et al.*, *Thromb. Res.* **158**, 16–18 (2017).
11. T. L. Iles, T. G. Laske, D. L. Garshelis, P. A. Iaizzo, *J. Exp. Biol.* **220**, 455–459 (2017).
12. M. J. Engbers, J. W. Blom, M. Cushman, F. R. Rosendaal, A. van Hylckama Vlieg, *J. Thromb. Haemost.* **12**, 290–296 (2014).
13. K. Martinod, D. D. Wagner, *Blood* **123**, 2768–2776 (2014).
14. A. Brill *et al.*, *J. Thromb. Haemost.* **10**, 136–144 (2012).
15. K. G. Welinder *et al.*, *J. Biol. Chem.* **291**, 22509–22523 (2016).
16. J. C. Meijers, W. L. Tekelenburg, B. N. Bouma, R. M. Bertina, F. R. Rosendaal, *N. Engl. J. Med.* **342**, 696–701 (2000).
17. M. Cushman, E. S. O'Meara, A. R. Folsom, S. R. Heckbert, *Blood* **114**, 2878–2883 (2009).
18. S. M. Bates, J. I. Weitz, *Circulation* **112**, e53–e60 (2005).
19. P. E. Geyer *et al.*, *Cell Syst.* **2**, 185–195 (2016).
20. B. Sahdo *et al.*, *Int. J. Med. Sci.* **10**, 508–514 (2013).
21. H. R. Bouma *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 2052–2057 (2011).
22. H. R. Bouma *et al.*, *J. Leukoc. Biol.* **94**, 431–437 (2013).
23. J. A. Coppinger *et al.*, *Blood* **103**, 2096–2104 (2004).
24. S. R. Piersma *et al.*, *J. Proteomics* **72**, 91–109 (2009).
25. J. B. Müller *et al.*, *Nature* **582**, 592–596 (2020).
26. M. Heck, B. A. Neely, *J. Proteome Res.* **19**, 3595–3606 (2020).
27. M. Zeiler, M. Moser, M. Mann, *Mol. Cell. Proteomics* **13**, 3435–3445 (2014).
28. P. E. Geyer, L. M. Holdt, D. Teupser, M. Mann, *Mol. Syst. Biol.* **13**, 942 (2017).
29. S. Ito, K. Nagata, *J. Biol. Chem.* **294**, 2133–2141 (2019).
30. P. Sasikumar *et al.*, *J. Thromb. Haemost.* **16**, 946–959 (2018).
31. W. J. Kaiser, L. M. Holbrook, K. L. Tucker, R. G. Stanley, J. M. Gibbins, *J. Proteome Res.* **8**, 2903–2914 (2009).

32. T. Petzold *et al*., *Blood* **122**, 2723–2731 (2013).
33. S. Wu *et al*., *Front. Pharmacol.* **12**, 792263 (2022).
34. M. Jiménez-Alcázar, N. Kim, T. A. Fuchs, *Semin. Thromb. Hemost.* **43**, 553–561 (2017).
35. P. Smith *et al*., *PLOS ONE* **17**, e0270865 (2022).
36. J. A. Diaz *et al*., *J. Vasc. Surg. Venous Lymphat. Disord.* **1**, 341–348.e1 (2013).
37. S. Ito *et al*., *J. Biol. Chem.* **292**, 20076–20085 (2017).
38. A. Carestia *et al*., *J. Leukoc. Biol.* **99**, 153–162 (2016).
39. N. Maugeri *et al*., *J. Thromb. Haemost.* **12**, 2074–2088 (2014).
40. P. Andre *et al*., *Blood* **118**, 5000–5010 (2011).
41. L. M. Holbrook *et al*., *Br. J. Haematol.* **148**, 627–637 (2010).
42. P. A. Randazzo *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **97**, 4011–4016 (2000).

ACKNOWLEDGMENTS

We thank all members of the Proteomics and Signal Transduction Group and the Clinical Proteomics Group for help and discussions, and I. Paron and M. Oroshi in particular for technical assistance. We thank A. R. Graesli, M. Lorenz, I. Saleh, and E. Raatz for technical assistance. Bear captures were led by S. Brunberg and D. Alquist. **Funding:** This work was supported by the following: German Research Foundation (PE2704/2-1 and PE2704/3-1 (to T.P.); SFB 1123-project B06 (to S.M.); TRR267 – ID 403584255 project B06 (to A.D.); LMU Munich's Institutional Strategy LMU excellent within the framework of the German Excellence Initiative (806 32 006) (to T.P.); the DZHK (German Center for Cardiovascular Research) (to T.P.) (External collaboration no. 100378833); by the LMU Munich's Clinician scientist program in vascular medicine (PRIME) (to M.T.); and by the LMU Munich's institutional Förderprogramm für Forschung und Lehre (FöFoLe) (to M.T.) This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant agreement 833440 to S.M.) The work was also partially supported by the Max Planck Society for the Advancement of Science awarded to M.M. The SANS bedrest study was funded by DLR and NASA (contract 80JSC019P0010 and 80JSC018P0078). DLR and NASA designed the bed rest study, but had no role in the design, data analysis, manuscript draft, or decision to publish the part of the study described in this manuscript. U.L. received funding from the internal grant program (project IFF 2020-26) of the Faculty of Health at Witten/Herdecke University, Germany, during preparation of this work. J.M.G. and G.L. received funding from the British Heart Foundation FS/17/31/32848. The Scandinavian brown bear research project is funded by the Norwegian Environment Agency and the Swedish Environmental Protection Agency.
Author contributions: Conceptualization: O.F. and T.P. Methodology: M.T., J.M.R., Z.Z., V.E., P.G., M.S., J.K., A.E., O.M., M.V., S.T., A.D., A.F., P.J., M.M.A., A.B., M.Z., A.P., K.S., S.K., J.G., D.M., U.L., M.M., and S.M. Investigation: M.T., J.M.R., Z.Z., V.E., J.H., K.S., B.K., L.S., J.N., E.L., G.L., M.O., L.N., S.E.N., A.P., B.K., and U.L. Visualization: Z.Z., V.E., J.M.R., M.T., K.S., A.T., and S.O. Supervision: T.P., S.M., M.M., and O.F. Writing - original draft: M.T., J.M.R., Z.Z., M.M., and T.P. Writing - review and editing: J.M.R., M.M., S.M., M.T., and T.P. **Competing interests:** The authors declare that they have no competing interests. **Data and material availability:** All data are available in the main text, supplementary material or online storage: The MS-based proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository and are available via ProteomeXchange with identifiers PXD030465, PXD031521, and PXD039087. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.sciencemag.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abo5044
Materials and Methods
Figs. S1 to S6
Tables S1 to S6
References (*43*–*57*)
MDAR Reproducibility Checklist

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

Submitted 9 February 2022; resubmitted 27 December 2022
Accepted 10 March 2023
10.1126/science.abo5044

DROUGHT

A global transition to flash droughts under climate change

Xing Yuan[1,2]*, Yumiao Wang[1,2], Peng Ji[1,2], Peili Wu[3], Justin Sheffield[4], Jason A. Otkin[5]

Flash droughts have occurred frequently worldwide, with a rapid onset that challenges drought monitoring and forecasting capabilities. However, there is no consensus on whether flash droughts have become the new normal because slow droughts may also increase. In this study, we show that drought intensification rates have sped up over subseasonal time scales and that there has been a transition toward more flash droughts over 74% of the global regions identified by the Intergovernmental Panel on Climate Change Special Report on Extreme Events during the past 64 years. The transition is associated with amplified anomalies of evapotranspiration and precipitation deficit caused by anthropogenic climate change. In the future, the transition is projected to expand to most land areas, with larger increases under higher-emission scenarios. These findings underscore the urgency for adapting to faster-onset droughts in a warmer future.

Droughts are periods of time with a persistent water deficit (*1*, *2*), which can cause devastating impacts on regional economies and environments (*3*–*5*), as well as on human health (*6*). Droughts mainly originate from large-scale internal climate variability, in which ocean-atmosphere teleconnections associated with phenomena such as the El Niño–Southern Oscillation, Pacific Decadal Variability, and Atlantic Multidecadal Variability play critical roles in drought formation and persistence over interannual to decadal time scales (*7*, *8*). For droughts that occur over shorter seasonal time scales, the dominant drivers can also include local or remote land-atmosphere feedbacks (*9*, *10*). The multiscale interactions among these different parts of the climate system raise challenges for drought forecasting and impact mitigation. Droughts are also influenced by anthropogenic forcings such as climate change (*2*, *11*), land use or land cover change, and human water consumption and management (*12*, *13*). As global warming accelerates the terrestrial water cycle (*14*, *15*), agricultural and hydrological droughts have increased substantially in many regions (*11*, *16*, *17*) and are projected to become more frequent, longer, and more severe in a warmer future (*2*, *11*, *18*). Such statements are based on analysis of droughts at seasonal, annual, or decadal time scales. However, recent studies have shown that droughts also occur frequently at subseasonal time scales worldwide (*4*, *5*, *19*–*24*) and can develop into severe droughts within a few weeks. These rapid-onset droughts are termed "flash droughts" in contrast with conventional droughts that evolve slowly. In addition to large precipitation deficits, flash droughts are also caused by abnormally high evapotranspiration that depletes soil water quickly (*25*–*29*), which challenges current drought monitoring and forecasting capabilities (*30*–*34*) that were developed to detect slowly evolving droughts.

The concept of flash droughts was proposed at the beginning of the 21st century but did not receive wide attention until the occurrence of the severe US drought in the summer of 2012 (*5*, *28*, *30*, *34*). This drought was regarded as one of the most severe US droughts since the 1930s Dust Bowl and caused more than US$30 billion of economic losses (*35*). One of the distinctive features of this drought was its extremely rapid onset, with many locations going from drought-free to extreme drought conditions within a month. This rapid intensification was unexpected, and no operational prediction models captured its onset (*30*). In this regard, some flash droughts can be considered as the onset stage of a long-term drought, the impacts of which are amplified by a subsequent persistent period of severe drought conditions (*23*, *30*, *36*). Moreover, even without a transition to seasonal drought, these rapidly evolving subseasonal droughts have substantial impacts on vegetation growth (*37*) and can trigger compound extreme events such as heat waves or wildfires. Previous studies have focused on the evolution and changing characteristics of flash droughts (*5*, *20*–*29*) and found that human-induced climate change has increased the frequency of flash droughts throughout southern Africa (*20*) and China (*21*). A recent study presents a 36-year climatology report of global flash droughts and shows substantial increase[s] in flash droughts throughout several key regions (*38*). However, no consensus has been reached on whether there has been a transition from slow to flash droughts at the global scale, because the frequency of slower-developing droughts at subseasonal time scales may also increase. There is currently no robust evidence that drought intensification rates have increased globally, although several studies have speculated such increases by relating drought onset with global warming (*16*, *21*).

In this study, we investigated changes in the speed of global drought onset and the partitioning between flash and slow droughts. We divided subseasonal droughts into flash droughts (*21*, *28*) and slow droughts by onset speed measured by the declining rate of soil moisture and present their global distributions during the local growing season over the past 64 years. We then estimated the global trend of the ratio of the number of flash droughts to total subseasonal droughts and the global trend of the onset speed of subseasonal droughts and attributed these trends to anthropogenic climate change on the basis of the sixth Coupled Model Intercomparison Project (CMIP6) (*39*) climate model simulations (table S1). We also showed how these trends vary over different IPCC SREX (Intergovernmental Panel on Climate Change Special Report on EXtreme events) regions (*40*).

Global distributions of flash and slow droughts

On the basis of estimates of soil moisture from three global reanalyses from 1951 to 2014, subseasonal drought events are identified as pentad-mean soil moisture declines from above the 40th percentile to below the 20th percentile and then increase to above the 20th percentile again [supplementary materials (SM), materials and methods]. The minimum duration for subseasonal droughts is 20 days to exclude dry spells that are too short to cause substantial impacts. We then divided the subseasonal droughts into flash and slow droughts depending on the rate of the reduction in soil moisture (*21*) during the onset stage (fig. S1). We used the ratio of flash drought events to the total number of subseasonal drought events, and the subseasonal drought onset speed (SM, materials and methods), to quantify the transition to flash droughts by determining whether there are significant trends in these two indices. Flash droughts tend to occur more often than slow droughts over humid regions with lower aridity (Fig. 1A and fig. S2), where flash-drought frequency is two to three times greater than other regions (fig. S3A). By contrast, slow drought occurrence has smaller spatial variability (fig. S3B). Flash droughts usually last for 30 to 45 days, whereas slow droughts usually last for 40 to 60 days (fig. S4). The uncertainty across

[1]School of Hydrology and Water Resources, Nanjing University of Information Science and Technology, Nanjing 210044, Jiangsu, China. [2]Key Laboratory of Hydrometeorological Disaster Mechanism and Warning of Ministry of Water Resources/Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters, Nanjing University of Information Science and Technology, Nanjing 210044, Jiangsu, China. [3]Met Office Hadley Centre, Exeter EX1 3PB, UK. [4]Geography and Environmental Science, University of Southampton, Southampton SO17 1BJ, UK. [5]Cooperative Institute for Meteorological Satellite Studies, Space Science and Engineering Center, University of Wisconsin–Madison, Madison, WI 53706, USA.
*Corresponding author. Email: xyuan@nuist.edu.cn

the three reanalyses is low over most humid and semihumid regions but high over arid regions (fig. S5).

The regions with a higher flash drought ratio also have faster drought onset speeds (Fig. 1B), which are associated with large precipitation deficits and/or increases in evapotranspiration. Compared with slow droughts, larger precipitation deficits occur during the onset stage of flash droughts over most global land areas (fig. S6A). In addition to the precipitation deficit, the increase in evapotranspiration (fig. S6B) accelerates the drawdown of soil moisture, which results in a higher likelihood of flash drought over humid regions, such as Europe, North Asia, southern China, eastern and northwestern parts of North America, and the Amazon. Evapotranspiration over these regions is energy limited, and the enhanced radiation because of fewer clouds drives the increase in evapotranspiration and speeds up drought onset. Over regions with higher aridity (such as northern China, western India, and parts of Africa), evapotranspiration is water limited (*41*), and the decrease in evapotranspiration during the onset stage suggests that precipitation deficit is the main driver of flash droughts (fig. S6).

Detection and attribution of changes in global droughts

Given that global land evapotranspiration is increasing in a warming climate (*14*, *42*), it was hypothesized that drought onset may speed up globally (*16*). In this study, we provide robust estimations that there are upward trends in the global mean flash drought ratio ($P < 0.1$) and subseasonal drought onset speed ($P < 0.1$) from 1951 to 2014 (Fig. 2, A and B), which means that subseasonal droughts have developed faster and shifted from slow to flash droughts at global scale. To assess whether the global trends are sensitive to the definition of flash droughts (*43*), we increased and decreased the soil moisture thresholds for drought starting and ending points as well as drought onset speed and found that the upward global trends remain significant ($P < 0.1$) (fig. S7).

The upward global trends are well captured by the state-of-the-art CMIP6/ALL multi-model ensemble simulations ($P < 0.1$) (Fig. 2, A and B), in which both the anthropogenic climate forcings (anthropogenic emission of, for example, greenhouse gases and aerosols) and natural climate forcings (solar and volcanic activities) are considered. The CMIP6/ALL ensemble simulations also roughly capture the spatial patterns of long-term climatology of flash drought ratio and subseasonal drought onset speed (fig. S8). However, the global trends are not captured by the CMIP6/NAT ensemble simulations that only consider natural climate forcings (Fig. 2, A and B). The best estimates of scaling factors (SM, materials and methods) show that only the ANT (ALL-NAT; anthropogenic forcings) signal is detectable, with contributions of 48% (10 to 86%) and 39% (13 to 70%) to the increases in flash drought ratio and subseasonal drought onset speed, respectively (Fig. 2, C and D). With CMIP5 models included, the detection and attribution results remain similar (fig. S9). We therefore conclude that the global transition to more frequent flash droughts during the past 64 years is influenced by anthropogenic climate change.

During the onset stage of subseasonal droughts, there has been a significant increase ($P < 0.1$) in strong anomalies of global evapotranspiration and a significant decrease ($P < 0.1$) in strong anomalies of precipitation surplus (precipitation minus evapotranspiration) during the past 64 years, in which anthropogenic contributions are detectable (fig. S10). The decrease in strong anomalies of precipitation surplus is dominated by the increase in strong anomalies of evapotranspiration because strong anomalies of precipitation show a small and insignificant decreasing trend ($P < 0.1$). Again, the results were similar after incorporating CMIP5 models (fig. S11). Therefore, anthropogenic climate change has significantly ($P < 0.1$) amplified the strong anomalies of global evapotranspiration and precipitation surplus and ultimately

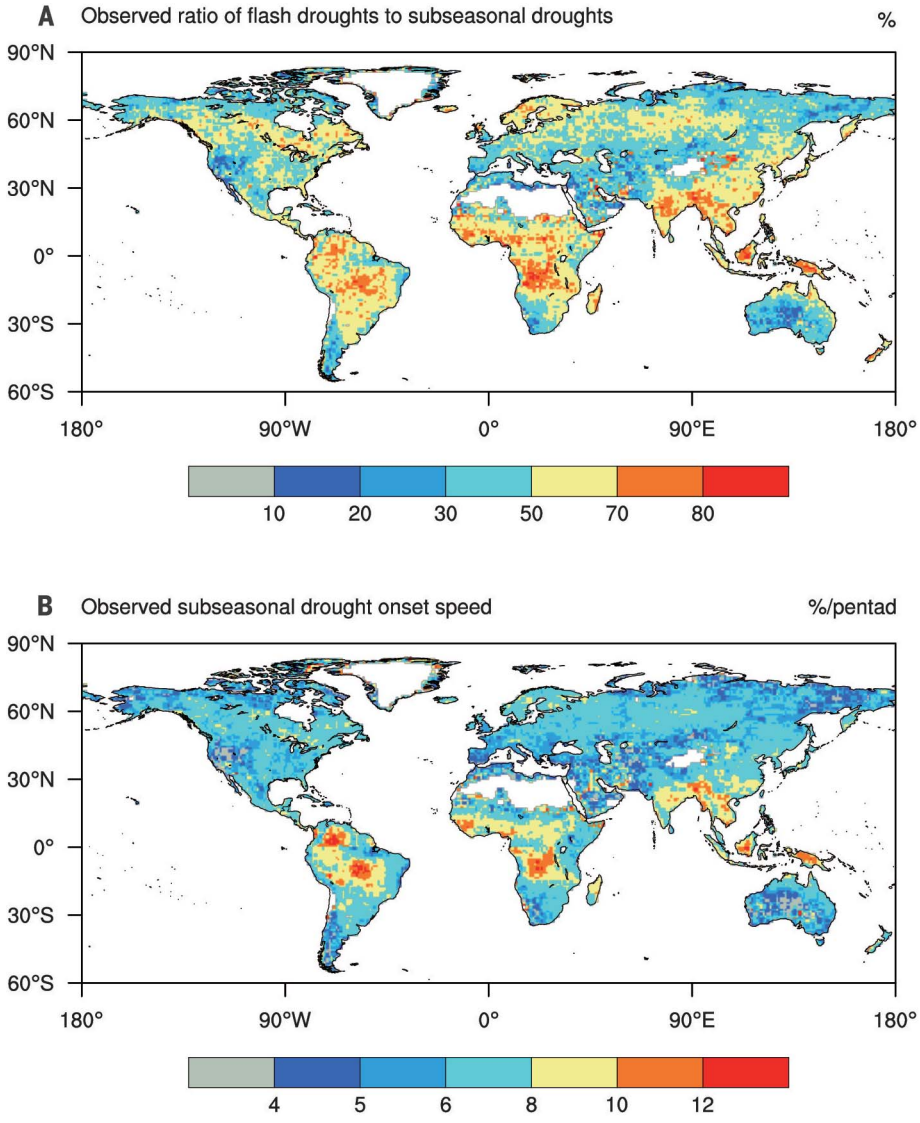


Fig. 1. Spatial distributions of the flash drought ratio and the onset speed of subseasonal droughts. (**A**) The ratio (%) of flash drought events to the sum of flash and slow drought events [subseasonal drought events (SM, materials and methods)]. (**B**) The mean onset speed (%/pentad) for both flash and slow droughts. All statistics are based on the average results from ERA5, GLDASv2.0/Noah, and GLDASv2.0/Catchment global reanalysis data during the growing seasons of 1951 to 2014 (April to September for the Northern Hemisphere and October to March for the Southern Hemisphere).

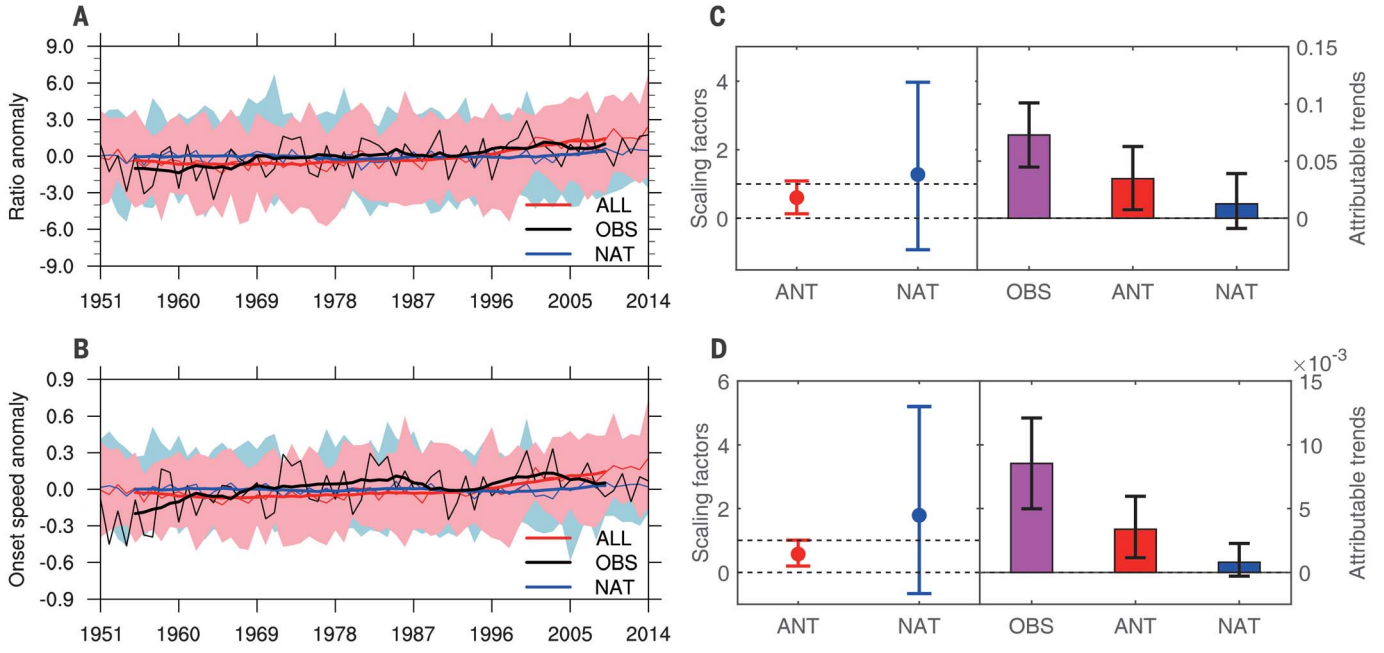


Fig. 2. Attribution for changes in global mean flash drought ratio and onset speed of subseasonal droughts. (**A**) Observed and simulated anomalies of the ratio (%) of flash drought events to subseasonal drought events averaged over the globe from 1951 to 2014. The black line indicates the results based on three global reanalysis datasets (OBS, mean of three reanalyses), and red and blue lines show the ensemble mean results based on CMIP6 climate model simulations with ALL and NAT forcings, respectively (table S1). The thick lines are 10-year running means, and the pink and cyan shadings display the 5 to 95% ranges of ALL and NAT ensemble simulations, respectively. (**C**) The best estimates of the scaling factors (left axis) and attributable increasing trends (%/year, right axis) from two-signal [ANT (ALL-NAT) and NAT] analysis of the changes in flash drought ratio for the period of 1951 to 2014. The time series used for detection and attribution are nonoverlapping 2-year averages (SM, materials and methods). Error bars indicate their corresponding 5 to 95% uncertainty ranges. (**B**) and (**D**) are the same as (A) and (C), except for the anomalies of onset speed of subseasonal droughts (%/pentad), scaling factors, and attributable trends (%/pentad/year) for the changes in onset speed from 1951 to 2014. Ratio and onset speed were identified at each grid cell and then averaged over the globe (excluding Antarctic, Greenland, and deserts) with consideration of the weights of grid areas. All of the statistics were calculated during the growing seasons (April to September for the Northern Hemisphere and October to March for the Southern Hemisphere).

has sped up drought onset and enhanced the global transition to more frequent flash droughts.

Regional drought changes in the past and projected future

A significant global transition to flash droughts is driven by regional increases in flash drought ratio over 74% of the IPCC SREX regions, notably for the significant increases ($P < 0.1$) over East and North Asia, Europe, Sahara, and the west coast of South America (Fig. 3A). Moreover, the onset speed of subseasonal droughts has increased over most regions, with significant increases ($P < 0.1$) over North Asia, Australia, Europe, Sahara, and the west coast of South America (Fig. 3B). These regions' significant increases in flash drought ratio and subseasonal drought onset speed (Fig. 3, A and B) are largely because of the increases in the frequency and onset speed of flash droughts (fig. S12). The regions with increasing onset speed but decreasing flash drought ratio suggest that the transition from slow to flash droughts might not be stable (Fig. 3, A and B). For example, East Africa, Northeast Brazil, and western North America show a historical decline in the flash drought ratio (Fig. 3A), but the frequency increases for both flash and slow droughts (fig. S12, A and C). These regions may eventually switch to a more stable transition once the onset speed increases to a certain level in the future. There are also regions with decreased frequency for both flash and slow droughts (such as eastern North America, southern South America, North Australia, and Southeast Asia), but the drought onset speed has increased (fig. S12). Almost all regions—except the Amazon and West Africa—show increasing trends in flash drought ratio and/or subseasonal drought onset speed (Fig. 3, A and B). For the Amazon, there is no evidence of a transition to flash droughts because drought onset speed decreases and flash-drought frequency decreases, whereas slow-drought frequency increases (fig. S12). For West Africa, both flash and slow droughts increase, whereas flash droughts occur faster and slow droughts occur slower, which suggests a more extreme drought condition even without an obvious transition signal (fig. S12). The results are similar for those with different drought thresholds (figs. S13 to S15).

Because of the regional differences in the responses to global warming, projecting drought changes at the regional scale is more challenging than that at the global scale (*44*–*46*). The CMIP6 climate model ensemble simulations roughly capture the historical changes in flash drought ratio and subseasonal drought onset speed, and 67 and 81% of the IPCC SREX regions show the same trends in ratio and speed between climate models and observations (Fig. 3, A and B, and fig. S16). Under a moderate emission scenario (SSP245) from 2015 to 2100, future projections show significant increasing trends ($P < 0.1$) in the flash drought ratio and subseasonal drought onset speed over almost all IPCC SREX regions (Fig. 3, C and D). Under a higher emission scenario (SSP585), the increasing trends become stronger over most regions (fig. S17). The projection results are similar for different drought thresholds (figs. S18 and S19) and different sets of climate models (fig. S20). Although flash droughts would only increase across 59% of the regions, and slow droughts would decrease over most regions, onset speeds for both flash and slow droughts would increase over most regions (figs. S21 and S22). Therefore, when droughts do occur in the future, they are more likely to be rapid-onset droughts. Although there are

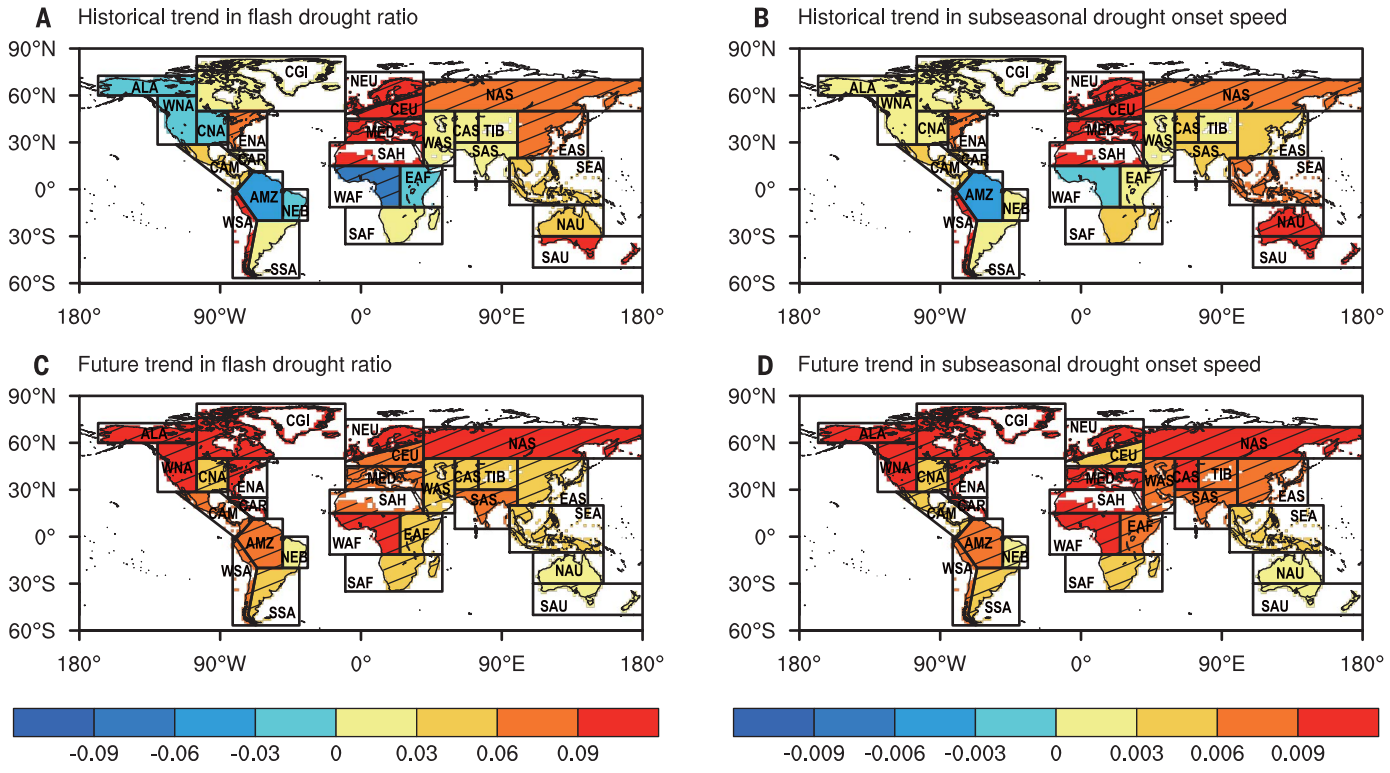


Fig. 3. Historical and future trends in flash drought ratio and onset speed of subseasonal droughts averaged over the IPCC SREX regions. (**A**) Observed trends (%/year) in regional mean ratio of flash drought events to subseasonal drought events from 1951 to 2014 based on the mean time series of three global reanalyses. (**B**) The same as (A), except for the trends (%/pentad/year) in regional mean subseasonal drought onset speed. (**C**) Projected future trends (%/year) in regional mean flash drought ratio from 2015 to 2100 based on CMIP6 climate model ensemble mean simulations under SSP245 scenario. (**D**) The same as (C), except for the trends (%/pentad/year) in regional mean subseasonal drought onset speed. Ratio and onset speed were identified at each grid cell for each model and were then averaged over the IPCC SREX regions with consideration of the weights of grid areas. All of the statistics were calculated during the growing seasons (April to September for the Northern Hemisphere and October to March for the Southern Hemisphere). Hatching represents a significant trend with $P < 0.1$ based on the nonparametric Mann-Kendall test.

uncertainties in the climate model projections, the results suggest that the transition to flash droughts is more stable and rapid in a warmer future, and the higher-emission scenario would lead to a greater risk of flash droughts with quicker onset, which poses a substantial challenge for climate adaptation.

Implications for climate adaptation

The transition toward more frequent flash droughts presents challenges for unraveling the anthropogenic influence on compound extremes (*44*), broadening our understanding of drought impacts (*5*, *37*) across time scales and improving drought prediction capability (*33*) for timely early warning. The increasing drought onset speed primarily comes from intensifying rainfall deficit and increasing evapotranspiration caused by anthropogenic climate change (fig. S10), which dries the soil quickly and creates ideal conditions for heat waves. Because the cooccurrence of heat waves and droughts is increasing globally according to the latest IPCC Sixth Assessment Report (*44*), the anthropogenic-enhanced transition to flash droughts suggests the need to understand flash drought–heat wave interactions both locally and remotely under climate change.

The transition to flash droughts may have irreversible impacts on terrestrial ecosystems (*5*, *37*). The impacts of extreme droughts on vegetation productivity are expected to increase in a warming future (*47*–*49*), but the findings are for long-term droughts with slow evolution. Because flash droughts develop more rapidly with higher temperatures (*27*), ecosystems may not have enough time to adapt to the sudden onset of large water deficits and heat extremes, resulting in a rapid reduction in ecosystem productivity (*37*). In addition, possible future increases in the durations of flash and slow droughts (fig. S23), as well as the increase in total subseasonal drought days caused by the increase in flash drought days (fig. S24), also suggest exacerbated impacts on ecosystems. Assessing such exacerbated impacts will comprehensively broaden our understanding of drought-vegetation interactions at time scales from yearly down to subseasonal.

The acceleration of drought onset also raises substantial challenges for drought monitoring and prediction (*33*). Effective monitoring of subseasonal droughts needs careful and objective selection of drought indices because various types of droughts (such as meteorological, agricultural, and hydrological droughts) also have different implications at the subseasonal time scale. The temporal resolutions of most current monitoring approaches are generally too coarse to capture the onset of flash droughts, and more frequent updates with drought indices suitable for shorter time scales are needed (*33*). For predictions, current approaches are aimed at predicting droughts at seasonal to decadal time scales, depending on oceanic and terrestrial sources of drought predictability (*7*–*10*). For subseasonal drought prediction, the Madden-Julian Oscillation, Southern and Northern Annular Modes, and Indian Ocean Dipole may provide relevant sources of predictability (*33*, *50*), but these large-scale signals should be connected with local synoptic anomalies [through Rossby wave train (*50*)] because most flash droughts do not have a wide spatial coverage. The linkage of these teleconnections with local or remote land-atmospheric coupling (*10*) could

provide a source of predictability for flash droughts.

Anthropogenic climate change is driving the transition to flash droughts, which has a wide range of implications for our understanding of climate change and its impacts, as well as how we can adapt to these changes. Improved understanding is needed for the adaptive capacity of natural ecosystems and human-managed environments that may be more susceptible to flash droughts and associated compound extreme events. Early warning of flash drought onset on time scales of a few weeks can be hugely beneficial for mitigating their impacts and managing the risk of this new normal.

REFERENCES AND NOTES

1. A. K. Mishra, V. P. Singh, *J. Hydrol. (Amst.)* **391**, 202–216 (2010).
2. T. R. Ault, *Science* **368**, 256–260 (2020).
3. A. B. Smith, J. L. Matthews, *Nat. Hazards* **77**, 1829–1851 (2015).
4. X. Yuan, Z. Ma, M. Pan, C. Shi, *Geophys. Res. Lett.* **42**, 4394–4401 (2015).
5. J. A. Otkin *et al*., *Agric. For. Meteorol.* **218-219**, 230–242 (2016).
6. B. I. Cook *et al*., *Wiley Interdiscip. Rev. Clim. Change* **7**, 411–432 (2016).
7. M. Hoerling, A. Kumar, *Science* **299**, 691–694 (2003).
8. G. J. McCabe, M. A. Palecki, J. L. Betancourt, *Proc. Natl. Acad. Sci. U.S.A.* **101**, 4136–4141 (2004).
9. J. K. Roundy, C. R. Ferguson, E. F. Wood, *J. Hydrometeorol.* **14**, 622–635 (2013).
10. R. D. Koster, Y. Chang, H. Wang, S. Schubert, *J. Clim.* **29**, 7345–7364 (2016).
11. A. Dai, *Nat. Clim. Chang.* **3**, 52–58 (2013).
12. A. F. Van Loon *et al*., *Nat. Geosci.* **9**, 89–91 (2016).
13. A. AghaKouchak *et al*, *Rev. Geophys.* **59**, e2019RG000683 (2021).
14. Y. Zhang *et al*., *Remote Sens. Environ.* **222**, 165–182 (2019).
15. X. Zhang *et al*., *Nature* **448**, 461–465 (2007).
16. K. E. Trenberth *et al*., *Nat. Clim. Chang.* **4**, 17–22 (2014).
17. K. Marvel *et al*., *Nature* **569**, 59–65 (2019).
18. Y. Pokhrel *et al*., *Nat. Clim. Chang.* **11**, 226–233 (2021).
19. T. W. Ford, D. B. McRoberts, S. M. Quiring, R. E. Hall, *Geophys. Res. Lett.* **42**, 9790–9798 (2015).
20. X. Yuan, L. Y. Wang, E. F. Wood, *Bull. Am. Meteorol. Soc.* **99**, S86–S90 (2018).
21. X. Yuan *et al*., *Nat. Commun.* **10**, 4661 (2019).
22. S. S. Mahto, V. Mishra, *Environ. Res. Lett.* **15**, 104061 (2020).
23. T. Parker, A. Gallant, M. Hobbins, D. Hoffmann, *Environ. Res. Lett.* **16**, 064033 (2021).
24. Y. Wang, X. Yuan, *Geophys. Res. Lett.* **48**, e2020GL091901 (2021)
25. K. C. Mo, D. P. Lettenmaier, *J. Hydrometeorol.* **17**, 1169–1184 (2016).
26. L. Wang, X. Yuan, Z. Xie, P. Wu, Y. Li, *Sci. Rep.* **6**, 30571 (2016).
27. T. W. Ford, C. F. Labosier, *Agric. For. Meteorol.* **247**, 414–423 (2017).
28. J. A. Otkin *et al*., *Bull. Am. Meteorol. Soc.* **99**, 911–919 (2018).
29. J. I. Christian *et al*., *J. Hydrometeorol.* **20**, 833–846 (2019).
30. M. P. Hoerling *et al*., *Bull. Am. Meteorol. Soc.* **95**, 269–282 (2014).
31. D. J. Lorenz, J. A. Otkin, M. Svoboda, C. R. Hain, Y. Zhong, *J. Geophys. Res. Atmos.* **123**, 8365–8373 (2018).
32. L. G. Chen *et al*., *Atmosphere (Basel)* **10**, 498 (2019).
33. A. G. Pendergrass *et al*., *Nat. Clim. Chang.* **10**, 191–199 (2020).
34. M. Liang, X. Yuan, *Front. Earth Sci. (Lausanne)* **9**, 615969 (2021).
35. NOAA National Centers for Environmental Information (NCEI), "U.S. Billion-Dollar Weather and Climate Disasters 1980-2022" (NCEI, 2022); https://www.ncei.noaa.gov/access/billions/events.pdf.
36. J. A. Otkin *et al*., *Atmosphere (Basel)* **12**, 741 (2021).
37. M. Zhang, X. Yuan, *Hydrol. Earth Syst. Sci.* **24**, 5579–5593 (2020).
38. J. I. Christian *et al*., *Nat. Commun.* **12**, 6330 (2021).
39. V. Eyring *et al*., *Geosci. Model Dev.* **9**, 1937–1958 (2016).
40. S. Seneviratne *et al*., *Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation*, C. B. Field *et al*., Eds. (Cambridge Univ. Press, 2012), pp. 109–230.
41. R. D. Koster, S. D. Schubert, H. Wang, S. P. Mahanama, A. M. DeAngelis, *J. Hydrometeorol.* **20**, 1241–1258 (2019).
42. H. Douville, A. Ribes, B. Decharme, R. Alkama, J. Sheffield, *Nat. Clim. Chang.* **3**, 59–62 (2013).
43. M. Osman *et al*., *Hydrol. Earth Syst. Sci.* **25**, 565–581 (2021).
44. IPCC, Summary for Policymakers in *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, V. Masson-Delmotte *et al*., Eds. (Cambridge Univ. Press, 2021).
45. A. M. Ukkola *et al*., *J. Hydrometeorol.* **19**, 969–988 (2018).
46. D. Hoffmann, A. J. E. Gallant, M. Hobbins, *J. Hydrometeorol.* **22**, 1439–1454 (2021).
47. C. Xu *et al*., *Nat. Clim. Chang.* **9**, 948–953 (2019).
48. W. Yuan *et al*., *Sci. Adv.* **5**, eaax1396 (2019).
49. X. Li *et al*., *Nat. Ecol. Evol.* **4**, 1075–1083 (2020).
50. S. D. Schubert, Y. Chang, A. M. DeAngelis, H. Wang, R. D. Koster, *J. Clim.* **34**, 1701–1723 (2021).

ACKNOWLEDGMENTS

Funding: Funding was provided by the National Natural Science Foundation of China 41875105, the National Key R&D Program of China 2018YFA0606002 and 2022YFC3002803, the Natural Science Foundation of Jiangsu Province for Distinguished Young Scholars BK20211540, and the UK-China Research & Innovation Partnership Fund through the Met Office Climate Science for Service Partnership (CSSP) China as part of the Newton Fund. **Author contributions:** Conceptualization: X.Y.; Methodology: X.Y., Y.W., and P.J.; Critical insights: P.W., J.S., and J.A.O.; Writing – original draft: X.Y.; Writing – review and editing: X.Y., Y.W., P.J., P.W., J.S., and J.A.O. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** The soil moisture, precipitation, and evapotranspiration datasets from the GLDAS2.0 land surface reanalysis are available at the GES DISC website for the Catchment land surface model (https://disc.gsfc.nasa.gov/datasets/GLDAS_CLSM10_3H_2.0/summary?keywords=GLDAS) and the Noah land surface model (https://disc.gsfc.nasa.gov/datasets/GLDAS_NOAH10_3H_2.0/summary?keywords=GLDAS). The soil moisture, precipitation and evapotranspiration datasets from ERA5 reanalysis are publicly available at the CDS website for 1951 to 1978 (https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels-preliminary-back-extension?tab=form) and for 1979 to 2014 (https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels?tab=form). The monthly potential evapotranspiration data produced by GLDAS2.0 are publicly available at https://disc.gsfc.nasa.gov/datasets/GLDAS_NOAH10_M_2.0/summary?keywords=GLDAS. The monthly potential evaporation data produced by ERA-5 are publicly available at the CDS website for 1951 to 1978 (https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels-monthly-means-preliminary-back-extension?tab=form) and for 1979 to 2014 (https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-single-levels-monthly-means?tab=form). The daily soil moisture, precipitation, and evapotranspiration (converted from latent heat flux) data from CMIP5 and CMIP6 are available at the WCRP website (https://esgf-node.llnl.gov/search/cmip5 and https://esgf-node.llnl.gov/projects/cmip6/), and the daily soil moisture data from the CESM1 Large Ensemble are available at the UCAR website (https://www.cesm.ucar.edu/projects/community-projects/LENS/data-sets.html). Statistical methods are noted in the text and figure captions. The computer codes for analyzing data and drawing plots are developed in Fortran or NCAR Command Language (NCL) scripts. The codes for flash and slow droughts are available at https://github.com/Hydroclimate2023/global-flash-drought. **License information:** Copyright © 2023 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

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abn6301
Materials and Methods
Figs. S1 to S26
Tables S1 to S3
References (*51*–*56*)

Submitted 9 December 2021; accepted 17 February 2023
10.1126/science.abn6301

METAMATERIALS

Mechanical nonreciprocity in a uniform composite material

Xiang Wang[1]*, Zhihao Li[1,2], Shuxu Wang[1,2], Koki Sano[1]†, Zhifang Sun[1], Zhenhua Shao[3], Asuka Takeishi[3,4], Seishiro Matsubara[5], Dai Okumura[5], Nobuyuki Sakai[6], Takayoshi Sasaki[6], Takuzo Aida[1,2], Yasuhiro Ishida[1]*

Mechanical nonreciprocity, or the asymmetric transmission of mechanical quantities between two points in space, is crucial for developing systems that can guide, damp, and control mechanical energy. We report a uniform composite hydrogel that displays substantial mechanical nonreciprocity, owing to direction-dependent buckling of embedded nanofillers. This material exhibits an elastic modulus more than 60 times higher when sheared in one direction compared with the opposite direction. Consequently, it can transform symmetric vibrations into asymmetric ones that are applicable for mass transport and energy harvest. Furthermore, it exhibits an asymmetric deformation when subjected to local interactions, which can induce directional motion of various objects, including macroscopic objects and even small living creatures. This material could promote the development of nonreciprocal systems for practical applications such as energy conversion and biological manipulation.

In many physical systems, the law of reciprocity ensures that signal propagation through the system always occurs in a symmetric manner. In simpler words, if a system responds to an input in one way, then it also does the same when the input is applied in the opposite direction. Breaking this symmetry induces nonreciprocity, a feature that has been increasingly pursued over the past several years in various fields, including optics (*1*–*4*), acoustics (*5*–*7*), quantum systems (*8*, *9*), and mechanics (*10*, *11*). Systems that exhibit nonreciprocity to mechanical quantities have recently emerged because of their potential to guide, damp, and control mechanical signals and energies in ways that have not been imagined for reciprocal systems (*12*–*14*). To realize such systems, breaking of either the time-reversal symmetry (dynamic systems) or the material deformation symmetry (static systems) is required. One simple example is the realization of nonreciprocal transmission of the displacement field through the fabrication of silicon rubber into a fishbone-structured metamaterial (*10*). However, mechanical nonreciprocity has so far mainly been realized through the complicated design of active robotics (*15*, *16*) or metamaterial frameworks (*17*, *18*). Even the most advanced systems still rely on the shaping and connection of reciprocal materials, which limits the design freedom and practical applications of such systems. As an alternative strategy, we propose the pursuit of mechanical nonreciprocity in a uniform material as its intrinsic property. Such a material is expected to show nonreciprocal response in any condition regardless of shape and size and, in principle, can manipulate any objects with which it interacts.

Mechanical reciprocity of common materials can be readily demonstrated by the following experiment: When a 10-mm cube of conventional polyacrylamide hydrogel was fixed at the bottom and sheared left and right at the top, with the force gradually increased to ±0.8 N, it showed the same extent of deformation (Fig. 1A and movie S1). Through this deformation process, the displacement changed linearly with complete symmetry (Fig. 1Aii). Such symmetry is observed not only for materials with a linear response but also for those with a nonlinear response (fig. S1).

To create a uniform material that shows a nonreciprocal response to shear forces, some symmetry breaking should be introduced to the internal nanostructure. Inspiration can be gained from some of the asymmetric designs of metamaterial frameworks (*10*). These designs are too complicated to be directly used, so we propose a simpler model that takes advantage of a phenomenon called buckling (*19*, *20*). Suppose that an elastic beam, tilted clockwise relative to the height direction, connects a floor and ceiling (Fig. 1B, top), where only pivotal motion is allowed at the connection points. The floor is unmovable, whereas the ceiling is horizontally slidable and retains its height. When the ceiling slides to the left, the beam experiences a compressive force and eventually buckles, so that it no longer effectively restricts the sliding motion. By contrast, when it slides to the right, the beam experiences a tensile force and does not buckle at all, so that it continues restricting the sliding motion before breaking. Overall, such direction-dependent buckling of the tilted beam induces an asymmetric response of the entire system to shear forces. We envisioned that this mechanism would be useful for creating a uniform material with mechanical nonreciprocity, in which the floor and ceiling are replaced with an elastic matrix and the tilted beam is replaced with nanofillers (Fig. 1B, bottom).

Asymmetric shear response induced by directional-dependent buckling of graphene oxide

To confirm our hypothesis, we developed a composite hydrogel consisting of a polyacrylamide network as the matrix and graphene oxide (GO) nanosheets as the nanofillers, in which the nanosheets were unidirectionally oriented in a tilted manner (Fig. 1C; see fig. S2 for material synthesis), just like the tilted beam in Fig. 1B. GO nanosheets were chosen because they easily buckle when subjected to in-plane compression, owing to their single-atom thickness (*21*, *22*). A plane indicator **P** is defined to help illustrate the orientation of the GO plane, and the angle between the height direction and **P** is designated as the tilt angle θ (details in fig. S3). Figure 1D shows the deformation of a 10-mm cube of such a composite hydrogel with $\theta = 45°$ when it was sheared to the right (left) with an applied force $F_0 = 0.8$ N ($-F_0 = -0.8$ N). Surprisingly, the hydrogel exhibited an almost 10-times larger displacement when it was sheared to the left than when it was sheared to the right (movie S1). The force-displacement relationship also depicted the highly asymmetric response of the hydrogel, which is strong evidence of mechanical nonreciprocity (Fig. 1Dii). We also confirmed that the hydrogel showed a similar asymmetric behavior even when the shear force was applied at a single point (fig. S4). This hydrogel is referred to as the nonreciprocal (NR) gel, and the directions exhibiting larger and smaller shear deformation are referred to as the softer and harder directions, respectively.

For a more quantitative analysis of the asymmetric response to shear, the stress-strain relationship of the NR gel was measured within the strain range of ±20% (Fig. 2A). Upon changing the left-sided strain up to −20%, the elastic modulus continuously decreased, resulting in a sluggish change in the stress (green in Fig. 2A). By contrast, upon changing the right-sided strain up to +20%, the elastic modulus was substantially enhanced with an increase in the strain, so that the strain-stress relationship exhibited a typical J-shaped curve (red in Fig. 2A).

[1]RIKEN Center for Emergent Matter Science, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan. [2]Department of Chemistry and Biotechnology, School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan. [3]RIKEN Center for Brain Science, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan. [4]RIKEN Cluster for Pioneering Research, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan. [5]Department of Mechanical Systems Engineering, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan. [6]International Center for Materials Nanoarchitectonics (WPI-MANA), National Institute for Materials Science (NIMS), 1-1 Namiki, Tsukuba, Ibaraki 305-0044, Japan.
***Corresponding author. Email: xiang.wang@riken.jp (X.W.); y-ishida@riken.jp (Y.I.)**
†Present address: Department of Chemistry and Materials, Faculty of Textile Science and Technology, Shinshu University, 3-15-1 Tokida, Ueda, Nagano 386-8567, Japan.

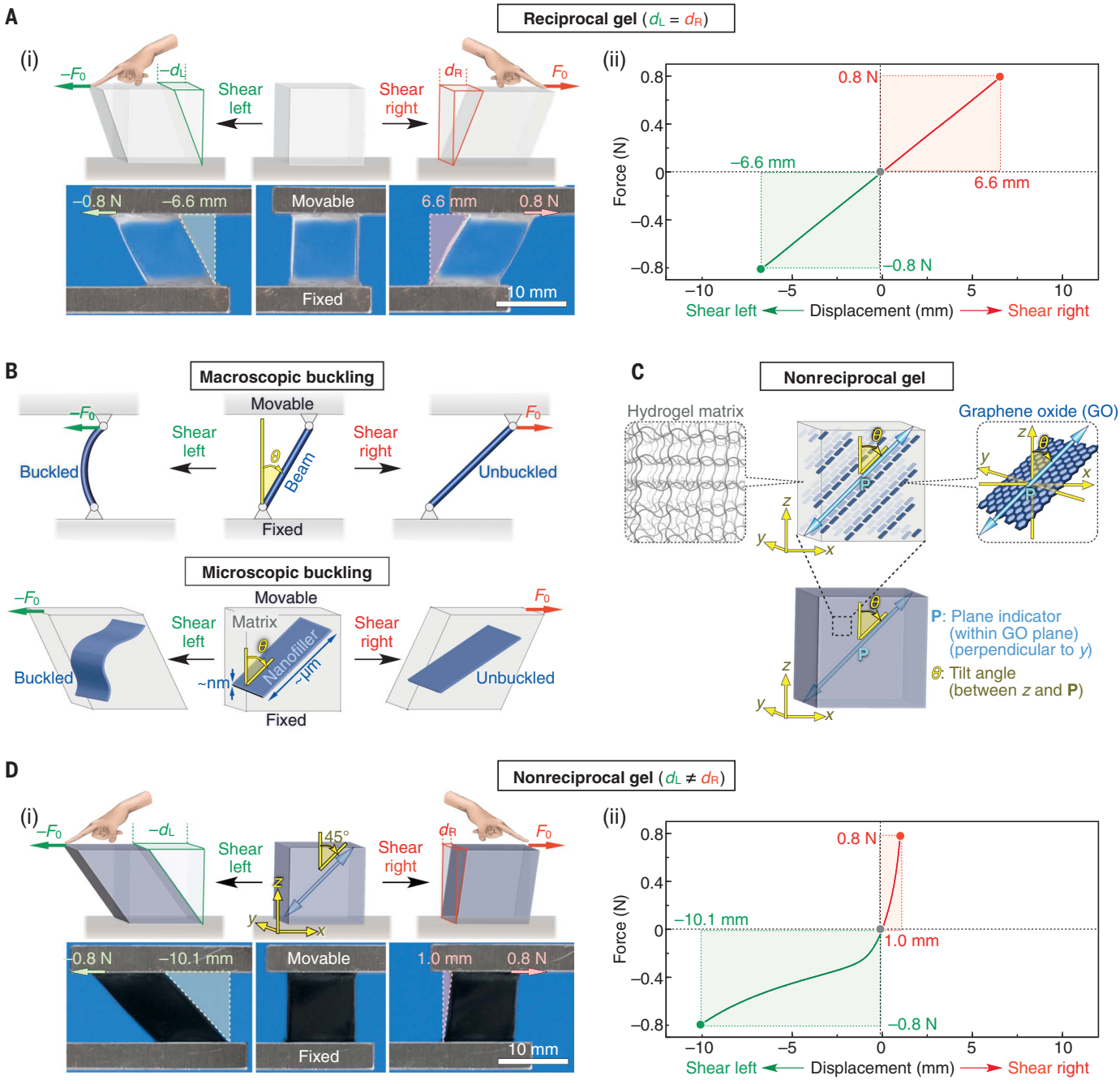


Fig. 1. Design and properties of the nonreciprocal gel with an asymmetric response to shear forces. (**A**) Symmetric deformation of a 10-mm cube of the reciprocal gel (polyacrylamide hydrogel) in response to shear forces applied at the top of the gel: (i) optical images and (ii) relationship between the displacement and applied force. d_L and d_R are the left and right displacements, respectively. (**B**) Macroscopic buckling of a tilted beam connecting a fixed floor and a slidable ceiling (top) and microscopic buckling of a tilted nanofiller embedded in a matrix (bottom). (**C**) Design of the nonreciprocal (NR) gel, which is composed of hydrogel as a matrix and unidirectionally oriented nanosheets of GO as nanofillers. The planes of the GO nanosheets include the *y* axis of the *x*-*y*-*z* orthonormal coordinate, where the *z* axis is parallel to gravity. The vector within the GO plane that is perpendicular to the *y* axis is designated as the plane indicator **P**. The angle between the *z* axis and **P** is designated as the tilt angle θ. For details, see fig. S3. (**D**) Asymmetric deformation of a 10-mm cube of the NR gel in response to shear forces applied at the top of the gel: (i) optical images and (ii) relationship between the displacement and applied force. The light blue arrow in (C) and (D) is the plane indicator **P** in the NR gel as defined in (C).

An extremely large difference existed between the elastic modulus at the left-shear strain of −20% (E_L) and the elastic modulus at the right-shear strain of +20% (E_R); E_R and E_L were estimated to be 380 and 5.7 kPa, respectively.

To confirm whether this asymmetric response really originates from the direction-dependent buckling of nanofillers (Fig. 2B), we monitored the changes in the alignment of the GO nanosheets during the shear deformation by investigating the light transmittance of the NR gel along the *y* axis (Fig. 2C and movie S2). Light transmittance is an effective indicator of the GO orientation because its light-absorbing efficiency highly depends on the angle between the GO plane and light path. The original NR gel showed considerable light

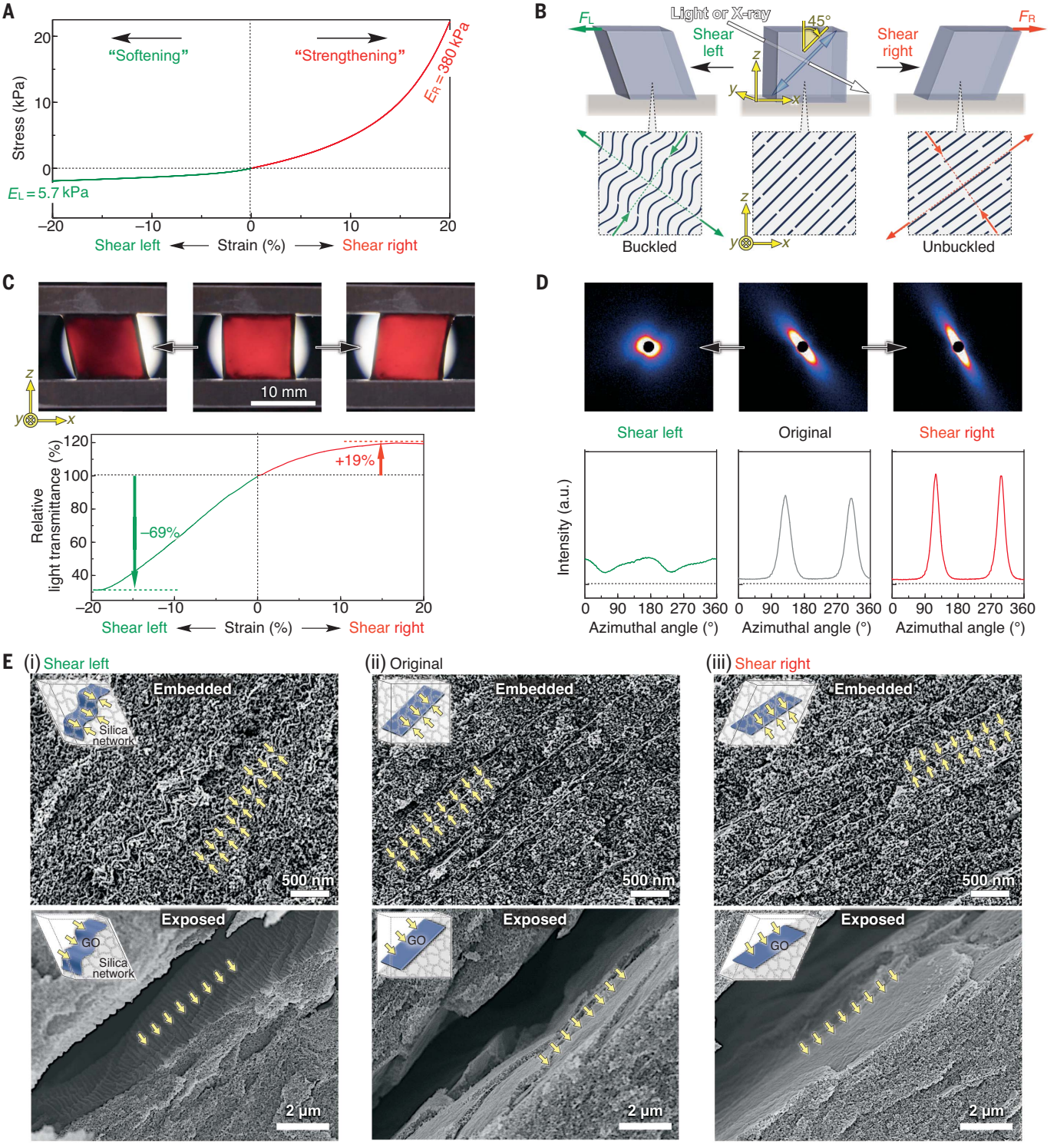


Fig. 2. Direction-dependent buckling of GO nanosheets in the NR gel that causes the asymmetric response to shear forces. (**A**) Stress-strain curve of the NR gel within the strain range of −20 to +20%, with right defined as positive. (**B**) Schematic of the buckling of the GO nanosheets upon shear deformation. The light blue arrow is the plane indicator **P** in the NR gel. F_L and F_R are the left and right forces, respectively. (**C**) Changes in the appearance of a 10-mm cube of the NR gel upon shear deformation (±20% strain) with irradiating unpolarized light parallel to the *y* axis (top) and the relationship between the strain and light transmittance (bottom). (**D**) Changes in the 2D SAXS image of a 1-mm-thick film of the NR gel upon shear deformation with irradiating x-ray parallel to the *y* axis (top) and azimuthal angle–intensity plot of the 2D SAXS image (bottom). a.u., arbitrary units. (**E**) Cross-sectional SEM images of the NR gel fixed by the silica network at the (i) left-sheared, (ii) original, and (iii) right-sheared states. The top row shows GO nanosheets embedded in the silica network, where the edges of the GO nanosheets appear as bright lines that are highlighted with yellow arrows and the silica network is observed as a porous skeleton of ~5-nm spheres. The bottom row shows GO nanosheets exposing their surfaces (smooth, as highlighted by yellow arrows) at the cracks of the silica network (porous), which were formed during cross-sectional cutting.

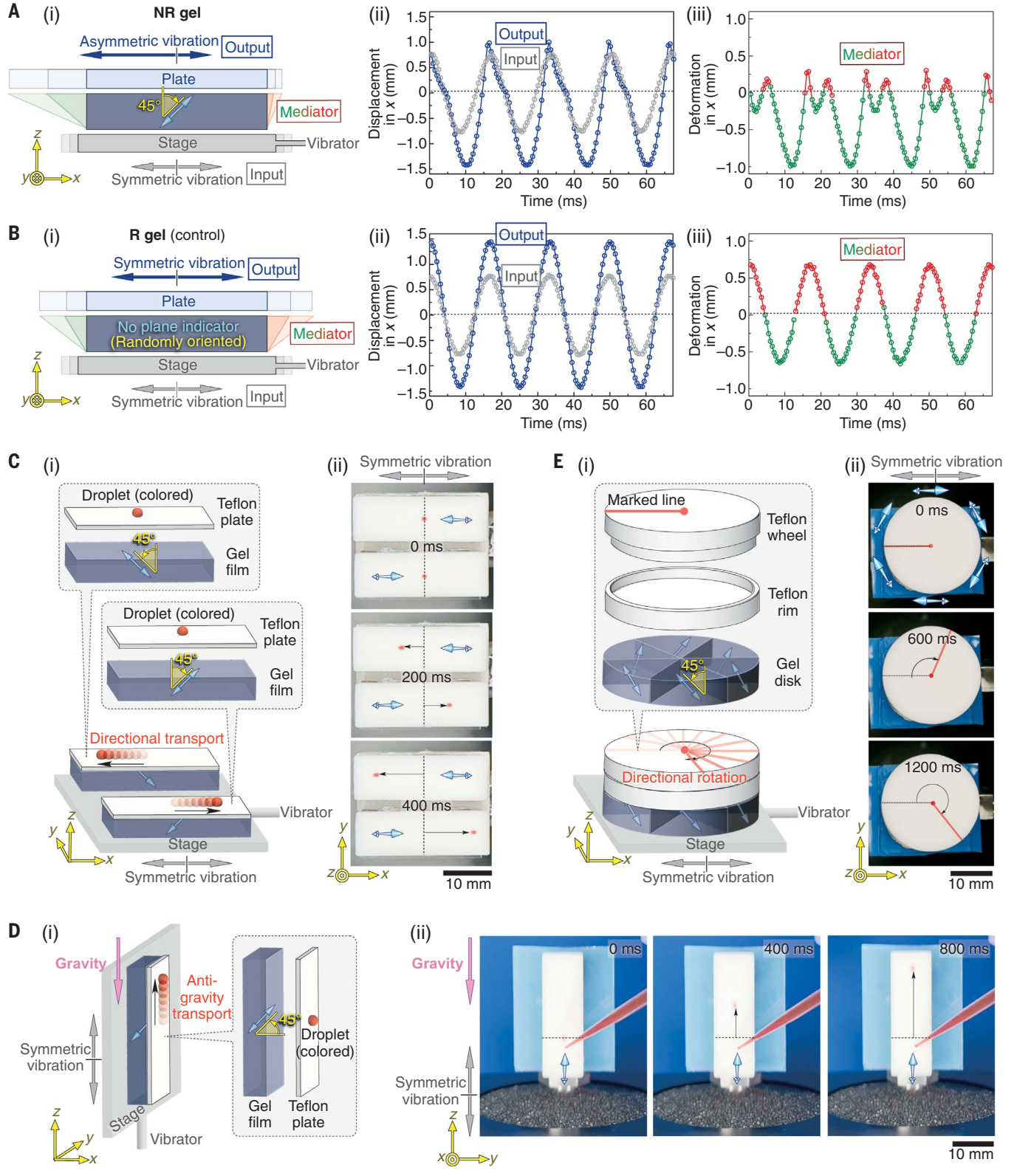


Fig. 3. Conversion of symmetric vibration into asymmetric vibration by the NR gel and its application to directional mass transport and rotation. (**A** and **B**) Vibration experiment of the NR gel (A) and the R gel that contains randomly oriented GO nanosheets as a control (B). Shown are (i) the schematic of the setup, (ii) time course changes in the *x* position of the bottom stage (input; gray) and that of the top plate (output; blue), and (iii) time-course changes in the shear deformation of the gel film estimated from (ii). (**C** to **E**) Directional transport of water droplets (C), antigravity transport of water droplets (D), and directional rotation of a wheel (E) using the NR gel. Shown are (i) the schematic of the setup and (ii) time-course pictures. For (A) to (E), 2-mm-thick gel films were used. In (A) and (C) to (E), the light blue arrow is the plane indicator **P** in the NR gel.

transmittance because all of the GO nanosheets exposed their edges to the light source (Fig. 2C, top center). Upon shearing to the left (softer) up to −20% strain, the light transmittance decreased by 69% (Fig. 2C, top left), which is indicative of the disordering of the GO orientation because of buckling. By contrast, upon shearing to the right (harder) up to +20% strain, the light transmittance increased (Fig. 2C, top right). This was because the alignment degree of the GO nanosheets was enhanced because they experienced a tensile force along the same direction to which they were aligned (shearing-induced alignment). Additional proof of the direction-dependent buckling is the results of the small-angle x-ray scattering (SAXS) analysis, which is often used to evaluate the ordering of nanostructures. When an x-ray beam was passed through to the NR gel along the y axis, a highly anisotropic scattering pattern was observed, which proved the high ordering of the GO nanosheets (Fig. 2D, top center) in the original NR gel. Upon shearing to the left (softer), the scattering pattern became isotropic (Fig. 2D, top left), indicating the complete disordering of the GO orientation. By contrast, upon shearing to the right (harder), the anisotropic scattering pattern became sharper (Fig. 2D, top right), which was consistent with the result of the light-transmittance measurement (Fig. 2C, top right). The transition between the ordered and disordered states was reversible and durable, as demonstrated by a cycling test of more than 100 cycles (fig. S5). The cycling test also revealed that the stress-strain curve showed little hysteresis (fig. S5A), indicating that the NR gel was highly elastic, which is favorable for transferring mechanical energies in an efficient manner.

By using scanning electron microscopy (SEM), we succeeded in visualizing the buckling of GO nanosheets in a more direct manner. For SEM analysis, samples of the NR gel in the original, left-sheared, and right-sheared states were rigidified by a silica network so that the configuration of the GO nanosheets at each state was fixed (details in the supplementary materials). In the cross-sectional images of the left-sheared (softer) sample (fig. S6), all of the GO nanosheets were buckled with a regular periodicity of ~300 nm along the direction in which they were compressed. Such buckling could be more easily visualized when the nanosheet surface was exposed because of cracks during cross-sectional cutting. By contrast, the SEM images of the original-state sample (fig. S7) and the right-sheared (harder) sample (fig. S8) did not show any sign of such periodic buckling. For a fair comparison between these three states, SEM images taken under similar magnifications and observation directions are summarized in Fig. 2E, which verifies that the buckling of the GO nanosheets is a shear direction–dependent phenomenon. Such a direction-dependent behavior causes a large difference in the reinforcement ability of the GO nanosheets under the left and right shear, thus giving rise to the nonreciprocity (details in the supplementary text, section 2.1).

We then investigated the influence of the tilt angle θ of the GO nanosheets. Upon increasing θ from 0° to 90°, E_R/E_L initially increased, reaching a maximum at θ = 45°, and then decreased (fig. S9). The structures at θ = 0° and 90° were symmetric, and so their mechanical response was also completely symmetric (fig. S10), which verifies the importance of structural asymmetry for nonreciprocal systems. We also investigated the influence of the size of the GO nanosheets, where the larger GO nanosheets had a higher E_R/E_L (details in fig. S11). Meanwhile, the influence of the concentration of GO nanosheets was not monotonic. Upon increasing GO concentration, E_R/E_L increased, reaching a maximum at 0.3 wt %, and then decreased (details in fig. S12). It is also worth noting that similar hydrogels with an asymmetric response to shear can also be fabricated by replacing the GO nanosheets with other 2D and even 1D nanofillers (fig. S13) or by replacing polyacrylamide with other polymers (fig. S14), confirming the generality of our strategy.

Guiding mechanical energies for mass transport and directional rotation

Nonreciprocal mechanical systems are expected to guide mechanical energies in a directional manner. As shown in Fig. 3A, a machine-generated sinusoidal vibration (the input vibration) was applied to a NR gel film from the bottom, and the vibration at the top of the gel, which was pasted with a plate as loading, was monitored (the output vibration). When a symmetric vibration (0.75-mm amplitude and 60-Hz frequency) was applied from the bottom (gray in Fig. 3Aii), a highly asymmetric vibration was afforded at the top (blue in Fig. 3Aii). This conversion was enabled by the vibrational shear deformation of the NR gel, which was highly asymmetric in terms of the frequency, amplitude, and peak shape (Fig. 3Aiii). By contrast, a reference hydrogel containing randomly oriented GO nanosheets, designated as the reciprocal gel (R gel; fig. S1), afforded a symmetric vibration under the same conditions (blue in Fig. 3Bii) because of its symmetric deformation (Fig. 3Biii).

Using the asymmetric vibration generated above, efficient vectoral transport of objects was realized, as exemplified by the following water droplet transport experiment (*23*). A NR gel film was placed on the vibration stage and covered with a Teflon plate, on which a 0.3-μl droplet of water was placed (Fig. 3Ci). To exclude the effects of artifacts such as the tilt of the stage, another gel film, with its harder direction opposite to the previous one, was also set on the same stage. When a symmetric vibration was applied from the bottom, the water droplets started drifting unidirectionally toward the harder directions of the NR gel films (fig. S15 and movie S3). This unidirectional transport operated stably when the water droplets were continuously loaded (movie S3). The droplets could even climb up against gravity smoothly when the vibration stage was vertically arranged (Fig. 3D and movie S4).

As a uniform material, one advantage of our NR gel is its good designability—it can be fabricated into pieces of any shape or size, and such pieces can be assembled on demand. Taking advantage of this, we developed a device by assembling six sector-shaped pieces of NR gels (details in fig. S16) that could convert vibrational noises into the rotational motion of a wheel loaded on the device (Fig. 3E and movie S5). The rotational device was quite stable and tolerant to external mechanical disturbances, and the direction was controllable (fig. S16 and movie S5). Importantly, for potential application, the response of this rotational device was so robust that it could generate a regular rotational motion from noisy vibrations with irregular frequencies and small amplitudes, including those from an oil pump and a wireless speaker (movie S6), implying its promise for use as an energy-harvesting system.

Manipulation of objects and creatures for directional movement

Apart from global deformation of the whole gel, another notable feature of the NR gel is its behavior when subjected to local deformation. For example, when a 2-mm-diameter metal cylinder laid parallel to the y axis was pushed vertically onto the surface of the NR gel, it showed a highly asymmetric deformation (Fig. 4Aii and movie S7), in contrast to the case of the R gel (Fig. 4Ai). At the left (softer) side of the cylinder, the gel bent sharply and contacted tightly with the cylinder. By contrast, at the right (harder) side, the gel bent gently, so that the deformation propagated over a wider distance (Fig. 4Aii). Such deformation can be attributed to the fact that the buckling occurred only at the left (softer) side, as proven by the light-transmittance experiment (fig. S17). The asymmetric deformation of the NR gel was consistent with the theoretical prediction that was based on the finite element analysis (Fig. 4Aiii and movie S8; details in the supplementary text). This theoretical study also indicated a highly asymmetric distribution of elastic strains: The softer side showed much higher strain than the harder side. As a result of such asymmetric deformation, our NR gel gains the potential to induce the directional motion of objects with which it interacts. For instance, when a

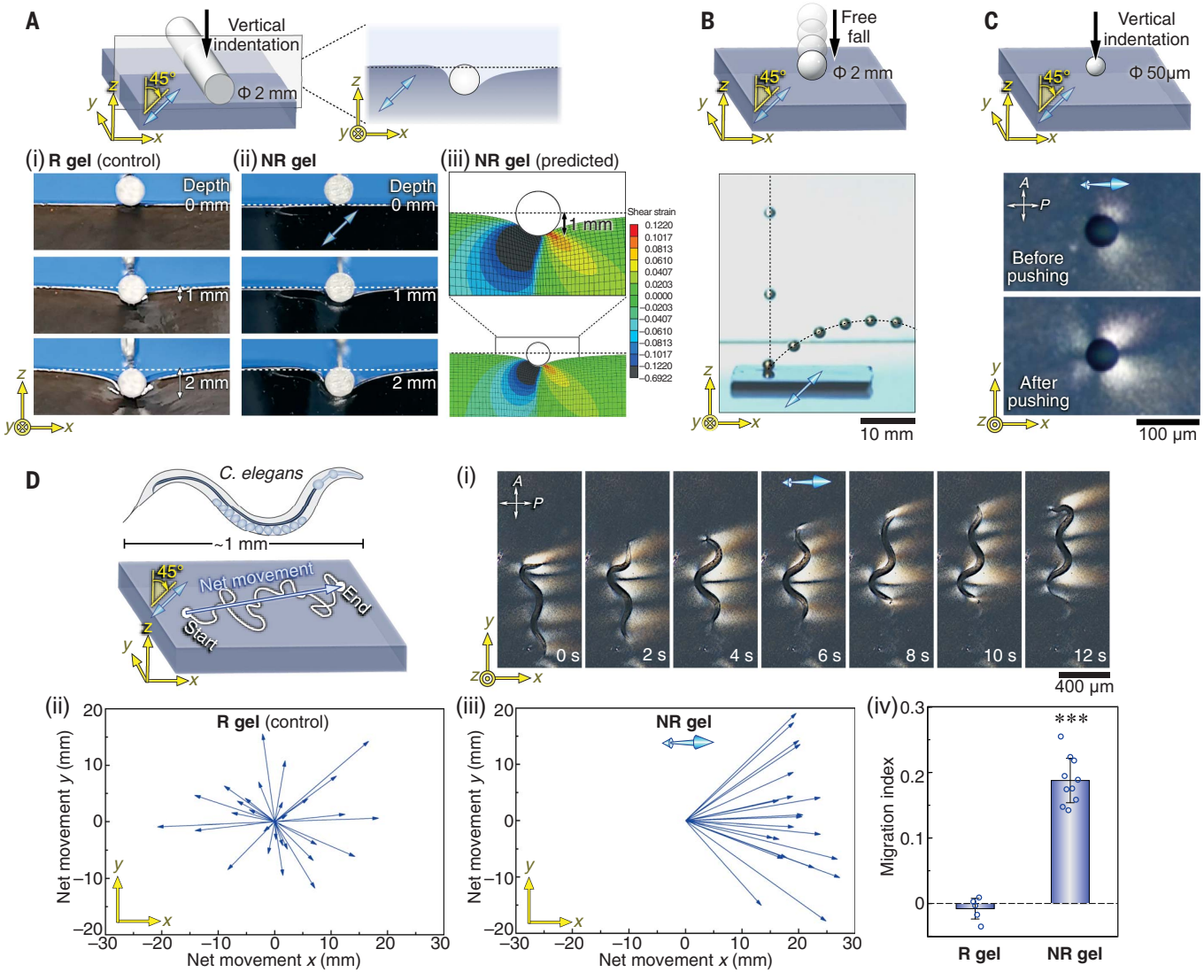


Fig. 4. Asymmetric deformation of the NR gel upon local compression and its application to directional movement of various objects. (**A**) Side-view photos of the R gel with randomly oriented GO nanosheets as a control (i) and the NR gel (ii) upon vertical indentation of a 2-mm-diameter cylinder laid parallel to the *y* axis. A theoretical prediction of the deformation and strain distribution in the NR gel based on the finite element analysis is shown in (iii). ϕ, diameter. (**B**) Time-lapse photo (bottom) showing the trajectory of a 2-mm-diameter metal ball free-dropped and bounced on the NR gel. (**C**) POM images (bottom) of the NR gel upon vertical indentation of a 50-μm-diameter microbead. To visualize the deformed regions, the polarizer (*P*) and analyzer (*A*) were set parallel and perpendicular, respectively, to the *x* axis. For enhancing optical contrast, the NR gel was prepared with titanium oxide nanosheets instead of GO nanosheets. (**D**) Directional migration of *C. elegans* on the NR gel. Shown in (i) are POM images of a worm crawling on the surface of the NR gel, using the same technique as was used in the microbead indentation experiment. Net movement of individual worms (arrows) in the representative assay on the R gel as a control (ii) and on the NR gel (iii). Shown in (iv) is the mean migration index of worms placed on the R gel as a control and the NR gel. The migration index was calculated as [(run distance to right) − (run distance to left)]/(total run distance). Each dot represents the migration index of a biologically independent assay performed with at least 30 worms. The error bars indicate standard error of the mean. ***$p < 0.001$ (one-sided Mann-Whitney U test). For (A), a 10-mm-thick gel slab was used. For (B) to (D), 2-mm-thick gel films were used. In (A) to (D), the light blue arrow is the plane indicator **P** in the NR gel.

2-mm-diameter metal ball was free-dropped onto a NR gel film, it always bounced to the harder side (Fig. 4B and movie S9).

Even micrometer-sized objects could also induce asymmetric deformation of the NR gel. To visualize such small deformations, polarized optical microscopy (POM) was used. When a 50-μm-diameter microbead was vertically pushed onto the gel, the bright area that emerged to the right (harder) side of the ball was much larger than that to the left (softer) side (Fig. 4C), indicating a highly asymmetric deformation. We then attempted to use our NR gel to manipulate small living creatures. As a proof-of-concept example, we chose *Caenorhabditis elegans* (Fig. 4D), a type of nematode that has long attracted much attention as a model organism for studying animal development and behavior (*24*, *25*). Indeed, when *C. elegans* crawled on the NR gel, an asymmetric deformation always occurred around the worm (Fig. 4Di and fig. S18), strongly implying that the migration of *C. elegans* might be biased by the NR gel.

To confirm this hypothesis, a NR gel film was placed on a horizontal stage, and a group of *C. elegans* containing ~30 worms was transferred to the middle of the film; the migration of the worms was then monitored for 50 min (Fig. 4D). Consequently, all the worms migrated

to the right (harder) side and eventually gathered at the right end of the gel film (movie S10). As shown in the net movement of all the worms, the migration was robustly biased to the right (Fig. 4Diii). To exclude the effects of other possible stimuli such as light, odor, and so on, worm migration was also conducted simultaneously on two NR gels placed side by side with their harder directions pointing opposite to each other. As expected, *C. elegans* always migrated toward the harder side of the NR gels (movie S10). By contrast, when the R gel that contained randomly oriented GO nanosheets was used instead, the worm migration was completely random (Fig. 4Dii and movie S10). For a more statistical study, the worm migration experiment was repeated 10 times for the NR gel as well as five times for control, and the same results were observed with excellent reproducibility (Fig. 4Div).

Conclusions

We have developed a uniform material that displays mechanical nonreciprocity and exhibits asymmetric deformation to various types of forces. This material can induce the directional transport of a wide range of objects, from macroscopic objects to small living creatures, suggesting potential applications in various fields such as mechanics, energy, and biology. This material can also broaden the scope of current metamaterials systems and inspire more interesting designs that cannot be realized by current existing reciprocal materials (for example, see fig. S19). In principle, the present design for nonreciprocal materials also applies to various kinds of elastomers other than hydrogels. One promising candidate is cross-linked liquid crystal elastomers, which are known to show large structural anisotropy. Our material may pave the way toward practical applications of nonreciprocal mechanical systems and inspire the development of nonreciprocal materials in other physical systems, such as optics and acoustics.

REFERENCES AND NOTES

1. R. J. Potton, *Rep. Prog. Phys.* **67**, 717–754 (2004).
2. D. L. Sounas, A. Alù, *Nat. Photonics* **11**, 774–783 (2017).
3. L. Feng *et al*., *Science* **333**, 729–733 (2011).
4. H. Xu, L. Jiang, A. A. Clerk, J. G. E. Harris, *Nature* **568**, 65–69 (2019).
5. R. Fleury, D. L. Sounas, C. F. Sieck, M. R. Haberman, A. Alù, *Science* **343**, 516–519 (2014).
6. L. Shao *et al*., *Nat. Electron.* **3**, 267–272 (2020).
7. C. Rasmussen, L. Quan, A. Alù, *J. Appl. Phys.* **129**, 210903 (2021).
8. V. Peano, M. Houde, F. Marquardt, A. A. Clerk, *Phys. Rev. X* **6**, 041026 (2016).
9. H. K. Lau, A. A. Clerk, *Nat. Commun.* **9**, 4320 (2018).
10. C. Coulais, D. Sounas, A. Alù, *Nature* **542**, 461–464 (2017).
11. H. Nassar *et al*., *Nat. Rev. Mater.* **5**, 667–685 (2020).
12. D. L. Sounas, J. Soric, A. Alù, *Nat. Electron.* **1**, 113–119 (2018).
13. Y. Wang *et al*., *Phys. Rev. Lett.* **121**, 194301 (2018).
14. L. M. Nash *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **112**, 14495–14500 (2015).
15. M. Brandenburger, X. Locsin, E. Lerner, C. Coulais, *Nat. Commun.* **10**, 4608 (2019).
16. G. Librandi, E. Tubaldi, K. Bertoldi, *Nat. Commun.* **12**, 3454 (2021).
17. M. Shaat, *Sci. Rep.* **10**, 21676 (2020).
18. M. Shaat, M. A. Moubarez, M. O. Khan, M. A. Khan, A. Alzo'ubi, *Phys. Rev. Appl.* **14**, 014005 (2020).
19. R. M. Jones, *Buckling of Bars, Plates, and Shells* (Bull Ridge Publishing, 2006).
20. N. Hu, R. Burgueño, *Smart Mater. Struct.* **24**, 063001 (2015).
21. A. V. Thomas *et al*., *Adv. Mater.* **27**, 3256–3265 (2015).
22. S. Deng, V. Berry, *Mater. Today* **19**, 197–212 (2016).
23. M. Morozov, O. Manor, *Curr. Opin. Colloid Interface Sci.* **36**, 37–45 (2018).
24. N. A. Croll, *J. Zool.* **176**, 159–176 (1975).
25. P. M. Meneely, C. L. Dahlberg, J. K. Rose, *Curr. Protoc. Essent. Lab. Tech.* **19**, e35 (2019).

ACKNOWLEDGMENTS

We thank F. Araoka (RIKEN) for helpful discussions.
Funding: This work was financially supported by JST CREST grant numbers JPMJCR17N1 and JPMJCR22B1, Japan. X.W. acknowledges support from the Incentive Research Project from RIKEN. S.W. acknowledges support from the Junior Research Associate (JRA) Program from RIKEN. **Author contributions:** X.W., Z.L., and Y.I. conceived the project. X.W. designed and conducted all experiments. Z.L., S.W., K.S., Z.Su., and T.A. helped with these experiments. S.M. and D.O. performed the theoretical analysis. Z.Sh. and A.T. guided the experiments that used *C. elegans* and analyzed the data. N.S. and T.S. synthesized the titanium oxide and niobium oxide nanosheets. X.W. and Y.I. cowrote the paper.
Competing interests: The authors declare no competing interests.
Data and materials availability: All data needed to evaluate the conclusions in the paper are present in the paper or the supplementary materials.

License information: Copyright © 2023 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

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adf1206
Supplementary Methods
Supplementary Text
Figs. S1 to S30
References (*26*–*44*)
MDAR Reproducibility Checklist
Movies S1 to S10

Submitted 30 September 2022; accepted 22 February 2023
10.1126/science.adf1206

EXOPLANETS

Direct imaging and astrometric detection of a gas giant planet orbiting an accelerating star

Thayne Currie[1,2,3]*, G. Mirek Brandt[4], Timothy D. Brandt[4], Brianna Lacy[5,6], Adam Burrows[5], Olivier Guyon[1,7,8], Motohide Tamura[7,9,10], Ranger Y. Liu[11]†, Sabina Sagynbayeva[12], Taylor Tobin[13]‡, Jeffrey Chilcote[13], Tyler Groff[14], Christian Marois[15,16], William Thompson[16], Simon J. Murphy[17,18], Masayuki Kuzuhara[7,9], Kellen Lawson[14], Julien Lozi[1], Vincent Deo[1], Sebastien Vievard[1], Nour Skaf[1], Taichi Uyama[9,19], Nemanja Jovanovic[20], Frantz Martinache[21], N. Jeremy Kasdin[22], Tomoyuki Kudo[1], Michael McElwain[14], Markus Janson[23], John Wisniewski[24], Klaus Hodapp[25], Jun Nishikawa[9,10], Krzysztof Hełminiak[26], Jungmi Kwon[10], Masahiko Hayashi[9]

Direct imaging of gas giant exoplanets provides information on their atmospheres and the architectures of planetary systems. However, few planets have been detected in blind surveys with direct imaging. Using astrometry from the Gaia and Hipparcos spacecraft, we identified dynamical evidence for a gas giant planet around the nearby star HIP 99770. We confirmed the detection of this planet with direct imaging using the Subaru Coronagraphic Extreme Adaptive Optics instrument. The planet, HIP 99770 b, orbits 17 astronomical units from its host star, receiving an amount of light similar to that reaching Jupiter. Its dynamical mass is 13.9 to 16.1 Jupiter masses. The planet-to-star mass ratio [(7 to 8) × 10^{-3}] is similar to that of other directly imaged planets. The planet's atmospheric spectrum indicates an older, less cloudy analog of the previously imaged exoplanets around HR 8799.

About 20 extrasolar gas giant planets have been detected with direct imaging by adaptive optics–assisted ground-based telescopes (*1*–*8*). These detections were made by blind (unbiased) surveys, in which targets were selected based on properties of the stellar system, including age and distance. However, the low yields of these blind surveys have shown that exoplanets detectable with direct imaging are rare (*9*).

Direct imaging provides constraints on an exoplanet's atmospheric properties, including their effective temperature, surface gravity, cloudiness, and chemical composition (*10*). However, direct imaging alone does not directly measure a planet's mass, which is instead inferred by using luminosity evolution models. These models have poor observational constraints and rely on often-uncertain host star ages (*11*). The typically wide orbits of directly imaged exoplanets, and short temporal coverage of their astrometry (precise positions and proper motions across the sky), provide poor constraints on orbital parameters derived from direct imaging data alone (*12*). The low number of direct imaging detections, and their poor mass and orbital constraints, impede our understanding of gas giant (jovian) exoplanet atmospheres and the architectures of their planetary systems.

Using an indirect detection method to guide direct imaging searches could, in principle, improve discovery yields and provide better constraints on the atmospheres, orbits, and masses of jovian exoplanets. A few previous surveys have adopted this approach, using long-term radial-velocity (RV) trends to select direct imaging targets. They resulted in direct imaging detections of stellar companions and intermediate- to high-mass brown dwarfs, but not planets (*13*) (supplementary text). An alternative approach is to use stars' astrometry to identify which stars are undergoing a proper motion acceleration, potentially caused by an undetected planetary-mass companion. Astrometric data can identify evidence for planetary-mass companions around young stars, which are unsuitable for precise RV

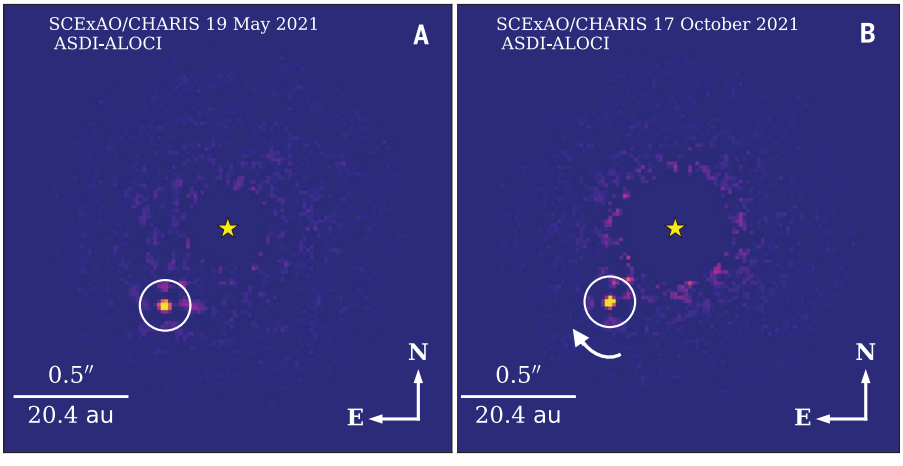


Fig. 1. Direct images of HIP 99770 b. We show the two highest-quality data sets, using SCExAO/CHARIS observations on (**A**) 19 May 2021 and (**B**) 17 October 2021. Data were processed using the adaptive, locally optimized combination of images (ALOCI) algorithm, combined with angular differential imaging (ADI) and spectral differential imaging SDI (ASDI) to remove the light of the host star (*32*–*34*). The star symbol indicates the host star's location. HIP 99770 b is within the white circle. The white arrow in (B) indicates the direction of the planet's orbital motion. The color intensity scalings are linear with a minimum of zero and maximum scaled to the mean signal within the core of the point spread function.

[1]Subaru Telescope, National Astronomical Observatory of Japan, Hilo, HI 96720, USA. [2]University of Texas-San Antonio, San Antonio, TX 78006, USA. [3]Eureka Scientific, Oakland, CA 94602, USA. [4]Department of Physics, University of California, Santa Barbara, Santa Barbara, CA 93106, USA. [5]Department of Astrophysical Sciences, Princeton University, Princeton, NJ 08544, USA. [6]Department of Astronomy, University of Texas-Austin, Austin, TX 78712, USA. [7]Astrobiology Center, Osawa, Mitaka, Tokyo 181-8588, Japan. [8]Steward Observatory, The University of Arizona, Tucson, AZ 85721, USA. [9]National Astronomical Observatory of Japan, Osawa, Mitaka, Tokyo 181-8588, Japan. [10]Department of Astronomy, Graduate School of Science, The University of Tokyo, Tokyo 113-0033, Japan. [11]Department of Astronomy, Columbia University, New York, NY 10027, USA. [12]Department of Physics and Astronomy, State University of New York–Stony Brook, Stony Brook, NY 11790, USA. [13]Department of Physics, University of Notre Dame, Notre Dame, IN 46556, USA. [14]NASA-Goddard Space Flight Center, Greenbelt, MD 20771, USA. [15]National Research Council-Herzberg, Victoria, BC V9E 2E7, Canada. [16]Department of Physics and Astronomy, University of Victoria, Victoria, BC V8W 2Y2, Canada. [17]Sydney Institute for Astronomy, School of Physics, University of Sydney, Sydney, Australia. [18]Centre for Astrophysics, University of Southern Queensland, Toowoomba, QLD 4350, Australia. [19]Infrared Processing and Analysis Center, California Institute of Technology, Pasadena, CA 91125, USA. [20]Department of Astronomy, California Institute of Technology, Pasadena, CA 91125, USA. [21]Universite Cote d'Azur, Observatoire de la Cote d'Azur, Laboratoire Lagrange, Nice 06000, France. [22]Department of Mechanical Engineering, Princeton University, Princeton, NJ 08544, USA. [23]Department of Astronomy, Stockholm University, Stockholm 114 19, Sweden. [24]Department of Physics and Astronomy, George Mason University, Fairfax, VA 22030, USA. [25]Institute for Astronomy, University of Hawai'i, Hilo, HI 96720, USA. [26]Nicolaus Copernicus Astronomical Center, Polish Academy of Sciences, Torun 87-100, Poland.

***Corresponding author. Email: currie@naoj.org**

†Present address: Parsons School of Design, New York, NY 10011, USA.

‡Present address: Department of Astronomy, University of Michigan, Ann Arbor, MI 48109, USA.

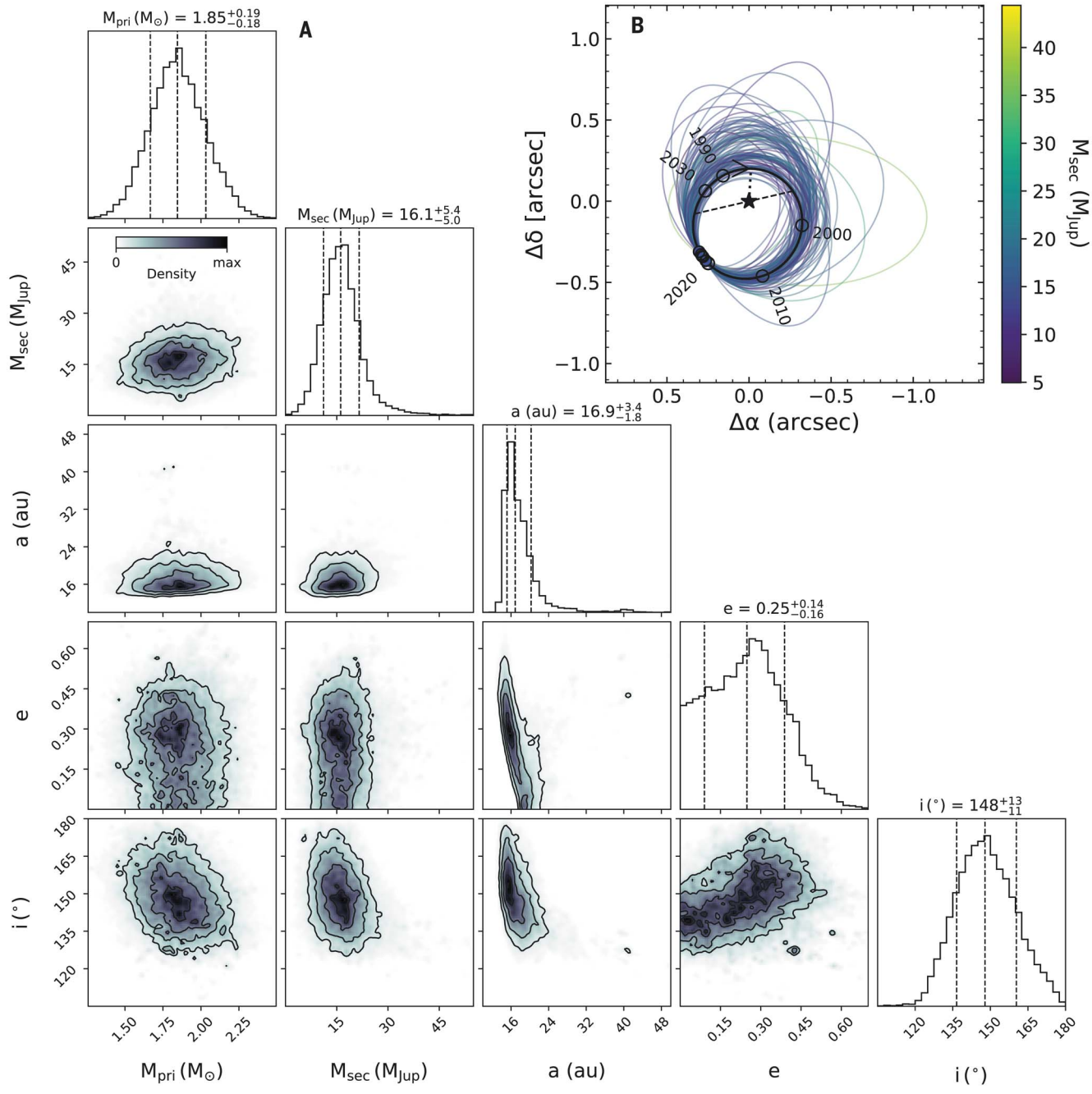


Fig. 2. Astrometric analysis of HIP 99770 b. (**A**) A corner plot showing the posterior probability distributions (gray shading) for each orbital parameter, the mass of HIP 99770 b, and the mass of the host star. Contours indicate the 1σ, 2σ, and 3σ levels of uncertainty. Histograms above each column indicate the marginalized probability distributions for single parameters, with vertical lines indicating the best-fitting values and 1σ uncertainties. (**B**) The best-fitting orbit (black), with circles indicating the predicted locations at labeled epochs. Also shown are 100 other orbits randomly drawn from the posterior probability distributions, color-coded by HIP 99770 b's mass.

measurements but are the most likely to host self-luminous gas giant planets (*14*). Combining the imaged planet's relative astrometry (measured with respect to the host star) with the host star's absolute astrometry can directly constrain planet masses and orbital properties (*14*). The astrometric precision of the Gaia mission, combined with measurements from the Hipparcos mission taken 25 years earlier, is sufficient to detect jovian planets orbiting nearby stars at separations similar to that of Jupiter or Neptune in the Solar System.

Astrometry of HIP 99770

HIP 99770 (also cataloged as HD 192640 or 29 Cygni) is a chemically peculiar star of spectral type A, located at a distance of $d \sim 40.74$ parsecs (pc) from the Sun. It has an effective temperature of ≈8000 K, a luminosity of ≈13.9 times the luminosity of the Sun ($L_\odot$), and a mass of ≈1.7 to 2.0 times the mass of the Sun ($M_\odot$) (*15*) (full stellar properties and references are listed in table S1). Different analysis methods—including kinematics, Hertzsprung-Russell diagram positions, and asteroseismology using the Transiting Exoplanet Space Satellite (TESS)—have assigned ages of either ~40 million years (Myr)

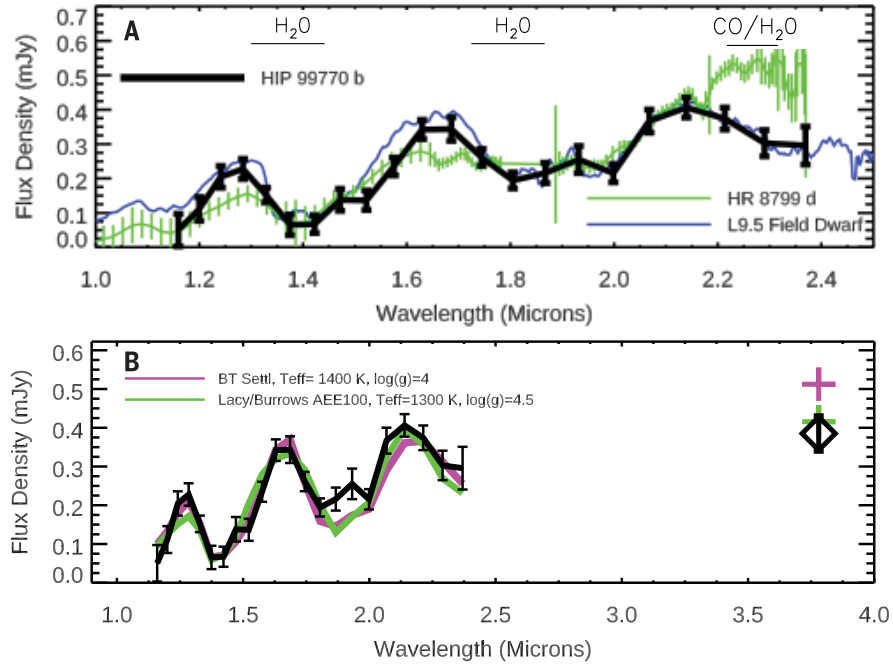


Fig. 3. Atmospheric characterization of HIP 99770 b. (**A**) The spectrum of HIP 99770 b (black) compared to spectra of HR 8799 d (green) (*35*, *36*) and the L9.5 field dwarf SIMPJ0956-1447 (blue) (*27*), all normalized to HIP 99770 b's spectrum at 2.1 μm. Horizontal bars indicate the wavelength ranges for absorption due to H_2O and CO. (**B**) HIP 99770 b's spectrum (black line) and L_p photometry (black diamond) compared to atmospheric models from the BT-Settl (magenta) and Lacy/Burrows (green) grids (χ^2_ν = 1.345 and 1.397, respectively) (*30*). Colored lines are model predicted spectra; plus signs are model predicted photometry. The best-fitting radii and implied masses are 0.92 R_J, 3.3 M_{Jup}, and 1.05 R_J, 13.5 M_{Jup}, respectively. Both models include clouds and atmospheric dust. The Lacy/Burrows models include nonequilibrium carbon chemistry and opacities. The best-fitting Lacy/Burrows model is designated AEE100, which has intermediate cloudiness and a modal size of cloud particles of 100 μm (*15*, *29*, *37*, *38*). Both panels are in flux density units of milli-jansky (mJy). All error bars indicate 1σ uncertainties.

or 115 to 414 Myr to HIP 99770. Data from the Wide-field Infrared Survey Explorer (WISE), Spitzer Space Telescope, and Herschel Space Observatory show that the star is surrounded by a cold debris disk detected at far-infrared (far-IR) wavelengths extending to >150 astronomical units (au) (*15*).

As part of a joint astrometry and direct imaging search for young exoplanets, we investigated evidence that HIP 99770 shows an astrometric acceleration due to an unseen companion, in the form of a deviation from linear motion across the sky. We used astrometry from the Hipparcos-Gaia Catalogue of Accelerations (HGCA) (*16*), a cross-calibration of the Hipparcos and Gaia missions. HIP 99770's average proper motion between the Hipparcos and Gaia missions differs from the proper motion measured around 2016 in both Gaia Data Release 2 (Gaia DR2) and Gaia Early Data Release 3 (Gaia eDR3) reductions (*17*). The HGCA using Gaia EDR3 data found a high χ^2 ~ 7.23 for a constant linear motion model, indicating a statistically significant acceleration (≈2.21 Gaussian sigma) at >97.31% confidence, equivalent to a false positive rate of <2.69%. The earlier DR2-based version of HGCA also found a similarly significant acceleration (χ^2 ~ 9.80). Other catalogs adopting independent analyses of Hipparcos and Gaia-eDR3 data also list HIP 99770 as having a statistically significant astrometric acceleration (*18*). The astrometric acceleration due to a secondary companion (which we designate with the subscript sec) depends on its mass M and projected separation R, in the form of a lower limit on M_{sec}/R^2_{sec}. For HIP 99770, this ratio is consistent with an object with ≈11 Jupiter masses (M_{Jup}) located at a fiducial separation of 0."5 (≈20 au), which is potentially detectable using direct imaging. Previous high-contrast imaging observations of this star using the Gemini Observatory have ruled out a stellar or substellar companion at wider separations (>1″) (*15*).

Direct imaging observations of HIP 99770

Motivated by this evidence of an astrometric acceleration, we observed HIP 99770 with the Subaru Coronagraphic Extreme Adaptive Optics (SCExAO) instrument, coupled to the Coronagraphic High-Resolution Imager and Spectrograph (CHARIS) (*19*, *20*) on the 8.2-m Subaru telescope. We used the low-resolution (broadband) mode, which provides 22 spectral channels covering the atmospheric near-IR windows simultaneously (wavelengths λ = 1.16 to 2.37 μm). We performed two shallow observing sequences in July and September 2020 and three deeper follow-up SCExAO/CHARIS observations between May and October 2021. We also obtained one complementary data set in the thermal IR (using the L_p filter centered on λ = 3."78 μm) using the Near Infra Red Camera 2 (NIRC2) instrument on the 10-m Keck II telescope.

The SCExAO/CHARIS images show a faint point source, which we designate HIP 99770 b, located at an angular separation ρ ~ 0.″43 to 0.″44 southeast of the host star (Fig. 1). In the highest-quality data, obtained by using point-spread function (PSF) subtraction methods to remove light from the star (*15*), HIP 99770 b is visible across the entire spectral range of the instrument. Every CHARIS data set detects the object (fig. S4), providing astrometry at epochs spanning 15 months (table S3). We also detect HIP 99770 b in the Keck–NIRC2 thermal IR data (fig. S4C).

We used the direct imaging astrometry to reject the possibility that HIP 99770 b is a stationary background object at the >15σ level (*15*). Between our first epoch (29 July 2020) and fourth epoch (13 July 2021), a background star would appear to move by ~65 milliarcseconds (mas) to both the west and south (table S2), owing to the star's proper motion. HIP 99770 b instead moved 23 ± 6 mas to the east and 29 ± 6 mas to the north—in the direction opposite to that expected for a stationary background star. To masquerade as an orbiting companion, a distant background object would require an even higher proper motion than HIP 99770 A. We reject the possibility that HIP 99770 b is a nonzero proper motion background star at the >5.7σ level from astrometry alone and at a >7.4σ level from astrometry and the planet's spectral features considered jointly (*15*).

Planet orbit and mass

To determine HIP 99770 b's orbital properties and mass, we simultaneously fit a model to its relative astrometry (from the imaging data) and the host star's proper motions and astrometric acceleration [from the Gaia and Hipparcos data (*16*)] using ORVARA, a Markov Chain Monte Carlo (MCMC) code (*16*, *21*). We used parallel tempering MCMC, in which progressively hotter chains are more accepting of poor fits to the data, enabling them to effectively explore multimodal posterior probability distributions (*22*). The coldest chain, the one that is least accepting of poor fits, is used for statistical inference. We used chains with 20 temperatures,

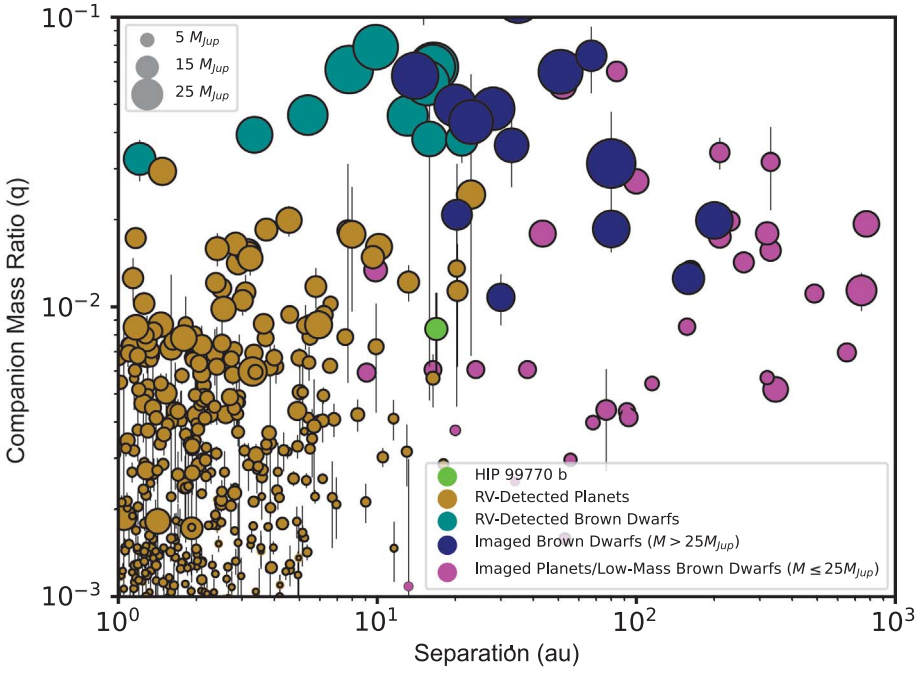


Fig. 4. Mass ratio and orbital separation for substellar companions. Data are from the NASA Exoplanet Archive (*39*) and the Extrasolar Planets Encyclopaedia (*40*) with small modifications (*15*). We show only objects detected by either direct imaging or radial-velocity techniques. Symbols sizes are proportional to the object mass, as indicated in the legend. HIP 99770 b (green point) lies in the region with RV-detected and directly imaged exoplanets.

each with 100 walkers taking 150,000 steps, to map the joint posterior probability distribution of the orbital parameters. We saved every 50th step and used only the second half of the coldest chain to obtain our probability distributions; the resulting distributions are then independent of how we initialized the chains. For our baseline model, we conservatively adopted a uniform prior on the companion mass, M_p, because this prior tends to yield higher companion masses than the alternative log-normal prior ($1/M_p$). For the host star mass, we adopted a Gaussian prior of $1.8 \pm 0.2\,M_\odot$. For all other parameters, we adopted standard uninformative priors (*23*). Adopting a $1/M_p$ companion prior or changing the prior for the primary mass resulted in a lower companion mass but otherwise had little effect on our results (*15*).

The posterior probability distributions output by ORVARA for HIP 99770 b constrain the orbital parameters (Fig. 2). For our baseline model, we derive a semimajor axis of $16.9^{+3.4}_{-1.9}$ au, similiar to that for the orbit of HR 8799 e, a gas giant planet previously detected with direct imaging (*1*). HIP 99770 b's orbital eccentricity ($e = 0.25^{+0.14}_{-0.16}$) is more consistent with values for directly imaged planets than for brown dwarfs (*12*). The model estimate of the host star mass is $1.85 \pm 0.1\,M_\odot$.

The dynamical mass of HIP 99770 b in the baseline model is $16.1^{+5.4}_{-5.0} M_{Jup}$, so the planet-to-star mass ratio is $q = \left(8.4^{+2.8}_{-2.6}\right) \times 10^{-3}$. Adopting a $1/M_p$ prior on the companion mass, instead of a uniform prior, yields a lower mass of $13.9^{+6.1}_{-5.1} M_{Jup}$ [so $q = \left(7.3^{+3.2}_{-2.7}\right) \times 10^{-3}$]. HIP 99770 b's mass is lower than the empirical mass separation between massive planets and low-mass brown dwarfs (supplementary text). HIP 99770 b's mass ratio is similar to those of some directly imaged exoplanets, including HR 8799 e ($q \approx 6 \times 10^{-3}$) and TYC 8998-760-1 b ($q \approx 0.01$ to 0.015), but is much lower than mass ratios measured for the brown dwarfs GJ 758 B and HD 33632 Ab (both $q > 0.04$) (*23*–*25*).

Planet atmospheric properties

We compare HIP 99770 b's CHARIS spectrum (data S1) to theoretical spectral templates (*26*) and the empirical Montreal Spectral Library (*27*) to estimate its spectral type and effective temperature. HIP 99770 b lies at the boundary between cloudy, methane-poor L-type substellar objects and (nearly) cloud-free, T-type objects showing methane absorption (the L/T transition). Among the spectral templates, the L7 template best matches the observed spectrum. In the spectral library, the closest match to HIP 99770 b is the L9.5 field brown dwarf SIMP J0956081−144706 (hereafter SIMPJ0956−1447). We therefore assign HIP 99770 b a spectral type of L7 to L9.5, corresponding to an effective temperature $T_{eff} \sim 1300$ to 1500 K (*28*).

HIP 99770 b's atmospheric spectrum appears to be intermediate between those of SIMPJ0956−1447 and the directly imaged, extremely cloudy L/T transition exoplanet HR 8799 d (Fig. 3A). Thick clouds in the atmospheres of L/T transition objects are expected to produce a planetary photosphere—the location where the optical depth τ is ≈1—which is more uniform with wavelength, resulting in a spectrum that is flatter, more blackbody-like, and redder [from J band (1.25 μm) to H band (1.65 μm) to K band (2.16 μm)] than for an object with thinner clouds (*29*). HIP 99770 b's spectrum is flatter and slightly redder than that of SIMPJ0956−1447. However, it is more peaked than HR 8799 d's spectrum at J and H by a factors of ~2 and ~1.5, and has a bluer J–K color, suggesting a cloud thickness intermediate between those of SIMPJ0956−1447 and HR 8799 d. At 2.2 to 2.4 μm, HR 8799 d lacks CO absorption owing to nonequilibrium atmospheric chemistry (*10*). As HIP 99770 b's spectrum is less luminous at those wavelengths, the planet's atmosphere is likely closer to being in chemical equilibrium.

We compared the spectrum and photometry of HIP 99770 b to two grids of atmospheric models, spanning a range of temperatures and surface gravities: the Lacy/Burrows grid (data S2) and the BT-Settl grid (*15*, *29*, *30*). These models adopt different prescriptions for clouds and atmospheric dust. We find that the atmospheric models require the presence of clouds to match HIP 99770 b's spectrum. The best-fitting models from the Lacy/Burrows grid are less cloudy and less dusty than for the HR 8799 planets (*29*). The best-fitting models have a surface gravity of $\log(g/\mathrm{cm\ s^{-2}}) = 4$ to 4.5 and temperatures of 1300 to 1600 K, with a preference for $T_{eff} = 1300$ to 1400 K (Fig. 3B) (*15*). Considering all models, the 2σ confidence intervals span $\log(\mathrm{g/cm\ s^{-2}}) = 4$ to 5 and temperatures of 1250 to 1600 K. The best-fitting luminosity of HIP 99770 b is $\log(L/L_\odot) = -4.53 \pm 0.02$, corresponding to a star-to-planet luminosity ratio of 2×10^{-6}. Masses inferred from luminosity evolution models are consistent with the planet's dynamical mass if the planet is ~80 to 200 Myr old (*15*).

Comparisons to other systems

We compare the architecture of the HIP 99770 system to that of the Solar System. HIP 99770 b orbits at ~16.9 au, intermediate between the distances of Saturn and Uranus from the Sun. The system's cold debris disk is likely >150 au from the star, about 3.5 times the typical distance from the Sun to Kuiper belt objects. However, the $1.85\,M_\odot$ star HIP 99770 is more luminous than the Sun. Therefore, the amount of light that HIP 99770 b receives (its insolation) is similar to that at ~4.5 au in the Solar System,

just interior to the orbit of Jupiter. Likewise, HIP 99770's debris disk, if located at 150 au, lies at a luminosity-scaled distance similar to that of the Kuiper belt from the Sun. Thus, like the HR 8799 system (*1*), the outer regions of the HIP 99770 system bear some characteristics of a scaled-up version of the outer Solar System, albeit with a single, far more massive planet.

The mass ratios (q) and orbital separations (a_p) of substellar companions provide a coarse diagnostic of formation processes: The companion mass function for substellar objects reaches a local minimum at $q \sim 0.025$, with smaller q values mostly being planets formed in a disk, whereas larger values are mostly brown dwarf companions formed by molecular cloud fragmentation (*31*). HIP 99770 b's mass ratio and separation ($q \sim 0.0084$, $a_p \sim 16.9$ au) are more consistent with those for planets below the deuterium-burning limit, detected by both direct imaging and RV (Fig. 4). This is similar to some RV-detected companions orbiting stars more massive than the Sun, which also have masses around (or slightly above) the deuterium-burning limit of ≈13 to 14 M_{Jup}. The deuterium-burning limit itself does not distinguish between planets and brown dwarfs: HIP 99770 b's mass is more consistent with those of other directly imaged planets than brown dwarfs (*15*) (supplementary text).

HIP 99770 b is an extrasolar planet jointly detected and characterized by using direct imaging and astrometry. HIP 99770 b joins β Pic b and HR 8799 e as directly imaged planets with measured atmospheric spectra and well-constrained dynamical masses (supplementary text).

REFERENCES AND NOTES

1. C. Marois, B. Zuckerman, Q. M. Konopacky, B. Macintosh, T. Barman, *Nature* **468**, 1080–1083 (2010).
2. A.-M. Lagrange *et al*., *Science* **329**, 57–59 (2010).
3. J. Carson *et al*., *Astrophys. J. Lett.* **763**, L32 (2013).
4. T. Currie *et al*., *Astrophys. J. Lett.* **780**, L30 (2014).
5. B. Macintosh *et al*., *Science* **350**, 64–67 (2015).
6. G. Chauvin *et al*., *Astron. Astrophys.* **605**, L9 (2017).
7. M. Keppler *et al*., *Astron. Astrophys.* **617**, A44 (2018).
8. T. Currie *et al*., *Nat. Astron.* **6**, 751–759 (2022).
9. E. L. Nielsen *et al*., *Astron. J.* **158**, 13 (2019).
10. Q. M. Konopacky, T. S. Barman, B. A. Macintosh, C. Marois, *Science* **339**, 1398–1401 (2013).
11. D. S. Spiegel, A. Burrows, *Astrophys. J.* **745**, 174 (2012).
12. B. P. Bowler, S. C. Blunt, E. L. Nielsen, *Astron. J.* **159**, 63 (2020).
13. J. R. Crepp *et al*., *Astrophys. J.* **831**, 136 (2016).
14. T. D. Brandt, T. J. Dupuy, B. P. Bowler, *Astron. J.* **158**, 140 (2019).
15. Materials and methods are available as supplementary materials.
16. T. D. Brandt, *Astrophys. J. Suppl. Ser.* **254** (suppl.), 42 (2021).
17. Gaia Collaboration *et al., Astron. Astrophys.* **649**, A1 (2021).
18. P. Kervella, F. Arenou, F. Thévenin, *Astron. Astrophys.* **657**, A7 (2022).
19. N. Jovanovic *et al*., *Publ. Astron. Soc. Pac.* **127**, 890–910 (2015).
20. T. D. Groff *et al*., in *Ground-based and Airborne Instrumentation for Astronomy VI*, C. J. Evans, L. Simard, H. Takami, eds. (2016), vol. 9908 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, p. 990800.
21. D. Foreman-Mackey, D. W. Hogg, D. Lang, J. Goodman, *Publ. Astron. Soc. Pac.* **125**, 306–312 (2013).
22. W. D. Vousden, W. M. Farr, I. Mandel, *Mon. Not. R. Astron. Soc.* **455**, 1919–1937 (2016).
23. G. M. Brandt *et al*., *Astron. J.* **162**, 301 (2021).
24. G. M. Brandt, T. D. Brandt, T. J. Dupuy, D. Michalik, G.-D. Marleau, *Astrophys. J. Lett.* **915**, L16 (2021).
25. A. J. Bohn *et al*., *Astrophys. J. Lett.* **898**, L16 (2020).
26. K. L. Cruz *et al*., *Astron. J.* **155**, 34 (2018).
27. J. Gagné, D. Lafrenière, R. Doyon, L. Malo, É. Artigau, *Astrophys. J.* **783**, 121 (2014).
28. D. C. Stephens *et al*., *Astrophys. J.* **702**, 154–170 (2009).
29. T. Currie *et al*., *Astrophys. J.* **729**, 128 (2011).
30. F. Allard, D. Homeier, B. Freytag, *Philos. Trans. R. Soc. A* **370**, 2765–2777 (2012).
31. M. Reggiani *et al*., *Astron. Astrophys.* **586**, A147 (2016).
32. T. Currie *et al*., *Astrophys. J. Lett.* **760**, L32 (2012).
33. C. Marois, D. Lafrenière, R. Doyon, B. Macintosh, D. Nadeau, *Astrophys. J.* **641**, 556–564 (2006).
34. W. B. Sparks, H. C. Ford, *Astrophys. J.* **578**, 543–564 (2002).
35. A. Z. Greenbaum *et al*., *Astron. J.* **155**, 226 (2018).
36. A. Zurlo *et al*., *Astron. Astrophys.* **587**, A57 (2016).
37. A. Burrows, D. Sudarsky, I. Hubeny, *Astrophys. J.* **640**, 1063–1077 (2006).
38. N. Madhusudhan, A. Burrows, T. Currie, *Astrophys. J.* **737**, 34 (2011).
39. NASA Exoplanet Archive, https://exoplanetarchive.ipac.caltech.edu/.
40. J. Schneider, C. Dedieu, P. Le Sidaner, R. Savalle, I. Zolotukhin, *Astron. Astrophys.* **532**, A79 (2011).

ACKNOWLEDGMENTS

We thank the three anonymous referees for their helpful comments and suggestions, which improved the quality of this paper. We thank the Subaru Time Allocation committee, NASA-Keck Time Allocation Committee, and recent Subaru Director Michitoshi Yoshida for their support of this program through Open Use and Director's Discretionary Time allocations. This research is based on data collected at Subaru Telescope, which is operated by the National Astronomical Observatory of Japan. The authors acknowledge the very important cultural role and reverence that the summit of Maunakea holds within the Hawaiian community. We are most fortunate to have the opportunity to conduct observations from this mountain. We thank J. Gagne, E. Mamajek, R. Gray, and J. Jones for extensive and helpful conversations regarding HIP 99770's properties. This research has made use of the NASA Exoplanet Archive, which is operated by the California Institute of Technology, under contract with the National Aeronautics and Space Administration under the Exoplanet Exploration Program. The development of SCExAO was supported by the Japan Society for the Promotion of Science (Grant-in-Aid for Research 23340051, 26220704, 23103002, 19H00703, and 19H00695), the Astrobiology Center of the National Institutes of Natural Sciences, Japan, the Mt Cuba Foundation, and the director's contingency fund at Subaru Telescope. **Funding:** T.C. was funded under NASA–XRP program NNX17AF88G and NASA-Keck Principal Investigator Award (Program 2020A_N027). B.L. is supported by a 51 Pegasi Postdoctoral Fellowship. C.M. acknowledges the support of the Natural Sciences and Engineering Research Council of Canada (NSERC). S.J.M. was supported by the Australian Research Council (ARC) through Future Fellowship FT210100485. V.D. acknowledges support from NASA grant 80NSSC19K0336 and Heising-Simons Foundation grant no. 2020-1823. M.T. is supported by JSPS KAKENHI grant 18H05442.
Author contributions: T.C. proposed and conducted the observations, reduced data, and led writing of the paper. G.M.B. and T.D.B. assisted with target selection and performed the joint direct imaging and astrometry analysis. B.L. and A.B. generated planet atmosphere models. R.Y.L. and S.S. performed separate data reductions; T.T. contributed to data reduction. O.G., J.L., V.D., S.V., N.S., J.C., T.G., and T.K. operated Subaru's facility AO system, SCExAO, and CHARIS during observations. C.M. and W.T. reduced and analyzed archival Gemini data. S.J.M. performed astroseismology analysis. M.K., T.U., M.M., M.J., J.W., K.H., J.N., K.H., and J.K. contributed to proposals and draft comments. N.J., F.M., N.J.K., and M.H. developed, commissioned, and maintained SCExAO and CHARIS and contributed to paper comments. M.T. oversaw group collaboration, proposals, and draft circulation.
Competing interests: The authors declare no competing interests.
Data and materials availability: The raw CHARIS data are available from the Subaru SMOKA archive at https://smoka.nao.ac.jp/. The Keck data are available from the Keck Observatory Archive at https://koa.ipac.caltech.edu/cgi-bin/KOA/nph-KOAlogin. The archival data we used from the WISE mission, Spitzer Space Telescope, and Herschel Space Observatory are available from the NASA/IPAC Infrared Science Archive https://irsa.ipac.caltech.edu/applications/Gator/; the Spitzer data we used are from the Spitzer Enhanced Imaging Products data set. The Gemini data are available from the Gemini Observatory archive at https://archive.gemini.edu/searchform. The TESS data are available from the Mikulski Archive for Space Telescopes at https://mast.stsci.edu/portal/Mashup/Clients/Mast/Portal.html. In all cases, data can be found by searching for "HIP 99770" or "HIP99770." The measured spectrum of HIP 99770 b is available as data S1. The Lacy/Burrows atmospheric models that we used to interpret the spectrum are available as data S2. The derived properties of the host star and the planet are listed in tables S1 and S4, respectively. The measured relative astrometric is reported in table S3. **License information:** Copyright © 2023 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. https://www.sciencemag.org/about/science-licenses-journal-article-reuse

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abo6192
Materials and Methods
Supplementary Text
Figs. S1 to S14
Tables S1 to S6
References (*41*–*144*)
Data S1 and S2

Submitted 18 February 2022; accepted 16 March 2023
10.1126/science.abo6192

CANCER IMMUNOLOGY

Engineered skin bacteria induce antitumor T cell responses against melanoma

Y. Erin Chen[1,2,3,4], Djenet Bousbaine[1,2,3], Alessandra Veinbachs[1,2,3], Katayoon Atabakhsh[1,2,3], Alex Dimas[1,2,3], Victor K. Yu[1,2,3], Aishan Zhao[1,2,3], Nora J. Enright[1,2,3], Kazuki Nagashima[1,2,3], Yasmine Belkaid[5,6], Michael A. Fischbach[1,2,3,7]*

Certain bacterial colonists induce a highly specific T cell response. A hallmark of this encounter is that adaptive immunity develops preemptively, in the absence of an infection. However, the functional properties of colonist-induced T cells are not well defined, limiting our ability to understand anticommensal immunity and harness it therapeutically. We addressed both challenges by engineering the skin bacterium *Staphylococcus epidermidis* to express tumor antigens anchored to secreted or cell-surface proteins. Upon colonization, engineered *S. epidermidis* elicits tumor-specific T cells that circulate, infiltrate local and metastatic lesions, and exert cytotoxic activity. Thus, the immune response to a skin colonist can promote cellular immunity at a distal site and can be redirected against a target of therapeutic interest by expressing a target-derived antigen in a commensal.

Certain members of the commensal microbiota elicit a potent T cell response upon colonization (*1*–*10*). This immune program has three defining features. First, the subtype of T cell elicited is determined by the bacterial colonist and its tissue context. For example, the gut commensal segmented filamentous bacterium induces T helper 17 (T_H17) cells (*1*), specific *Clostridium* species induce regulatory T cells (*11*), and certain strains of the skin colonist *Staphylococcus epidermidis* induce $CD8^+$ T cells (*2*). Second, colonist-induced T cells express a T cell receptor (TCR) specific for the bacterial inducer (*1*, *8*, *9*, *12*). This is a degree of specificity normally associated with an immune response to an infecting pathogen. Third, this process takes place across an intact, uninflamed epithelial barrier (*1*–*3*). Thus, physiologic colonization leads to adaptive immunity in the absence of an infection.

This work sought to answer two fundamental questions, one basic and one applied. First, we sought to understand the host's "goal" in responding to colonists by probing the functional properties of colonist-induced T cells. The $CD8^+$ T cells elicited by *S. epidermidis* are known to promote skin homeostasis and accelerate wound closure (*12*, *13*), which are unusual functions for a $CD8^+$ T cell. Are these colonist-induced immune cells capable of carrying out core functions of cellular immunity (migration and antigen-specific cytotoxicity) that are important for preventing sepsis and opportunistic infections?

Second, antigen-specific immune responses are of keen interest to researchers in oncology (*14*), infectious disease (*15*, *16*), and autoimmunity (*17*). Commensals induce antigen-specific T cells simply by colonizing, a process that is durable, safe, and avoids the undesirable side effects commonly associated with immune stimulation (*3*). Can anticommensal immunity, particularly its antigen specificity, be harnessed therapeutically by redirecting it against a target of interest?

We reasoned that both questions could be addressed by the same approach (Fig. 1A). By expressing tumor antigens in *S. epidermidis*, we could elicit T cells that were licensed by the commensal immune program but specific for a tumor. In so doing, we could probe whether these T cells were capable of leaving the colonized tissue, infiltrating a tumor, and engaging in cytotoxic activity.

Here, we show that strains of *S. epidermidis* engineered to express model antigens and tumor-derived neoantigens elicit tumor-specific $CD4^+$ and $CD8^+$ T cells upon colonization. These immune cells migrate to distal skin sites and the lungs and exhibit potent antitumor activity. *S. epidermidis* must be alive to elicit a response, suggesting that this program requires active colonization. Colonist-induced T cells synergize potently with immune checkpoint blockade. Aggressive B16-F10 melanomas are often rejected even when treatment is initiated after they are established. Thus, a commensal-directed immune program can function outside of tissue homeostasis, and engineered colonists are a simple but powerful means of generating antigen-specific T cells against a target of interest.

[1]Department of Bioengineering, Stanford University, Stanford, CA 94305, USA. [2]Department of Microbiology and Immunology, Stanford University School of Medicine, Stanford, CA 94305, USA. [3]ChEM-H Institute, Stanford University, Stanford, CA 94305, USA. [4]Dermatology Service, San Francisco Veterans Administration Health Care System, San Francisco, CA 94121, USA. [5]Metaorganism Immunity Section, Laboratory of Host Immunity and Microbiome, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892, USA. [6]NIAID Microbiome Program, National Institute of Allergy and Infectious Diseases, Bethesda, MD 20892, USA. [7]Chan Zuckerberg Biohub, Stanford, CA 94305, USA.
*Corresponding author. Email: fischbach@fischbachgroup.org

Engineering antigen expression into *S. epidermidis*

We chose to focus on *S. epidermidis* NIHLM087, an isolate from healthy human skin that elicits antigen-specific $CD4^+$ and $CD8^+$ T cells in mice and nonhuman primates (*2*, *3*, *12*). It does so in the context of physiologic skin colonization. There is no infection, the skin barrier is not breached, and there are no inflammatory processes (*2*, *3*, *13*, *18*).

Even though *S. epidermidis* is a ubiquitous human colonist, genetic manipulation of primary human isolates such as NIHLM087 is challenging because *S. epidermidis* has stringent restriction systems and phage-transduction susceptibility that differ substantially among strains (*19*–*22*), and as with other Gram-positive bacteria, electroporation is an inefficient means of introducing DNA (*23*). We developed a genetic system for *S. epidermidis,* which includes a previously described method of passaging plasmids through *Escherichia coli* DC10B (Δ*dcm*) and also incorporates heat shock and growth in hyperosmolar sorbitol (*24*) (Fig. 1B). This system enabled us to construct mutations in 13 of the 17 *S. epidermidis* strains that we tested, including primary human isolates from diverse phylogenetic groups (Fig. 1C).

We next sought to engineer *S. epidermidis* NIHLM087 (hereafter, *S. epidermidis*) to express non-native antigens. Notably, the process that determines which *S. epidermidis* antigens are selected for presentation from among thousands of distinct proteins in the bacterial cell has not been studied, so it was unclear whether a non-native peptide could compete against native *S. epidermidis* antigens for $CD8^+$ T cell recognition.

Our strategy accounted for this element of uncertainty (Fig. 1D). We began with ovalbumin (OVA) as a model antigen because it harbors well-characterized antigenic peptides that are recognized by OVA-specific $CD8^+$ or $CD4^+$ T cells (OT-I or OT-II, respectively). We were unsure whether a non-native antigen could be expressed in an undomesticated human isolate at a high enough level to be presented on the surface of an antigen-presenting cell, or whether the subcellular localization of the antigen in *S. epidermidis* would affect its processing and presentation. We therefore generated three sets of strains in which the nature of the antigen and its localization were varied. One strain for cytoplasmic expression consisted solely of full-length OVA (cOVA). Another three strains were designed for antigen displayed on the cell wall in which OVA, OT-I, or OT-II was spliced between the N-terminal sortase signal peptide and a C-terminal cell wall–spanning region from *S. aureus* protein A (*25*). This approach yielded *S. epi*-wOVA, *S. epi*-wOT-I, and *S. epi*-wOT-II, respectively ("w" denotes "wall-attached."). Three more strains were generated for antigen secretion,

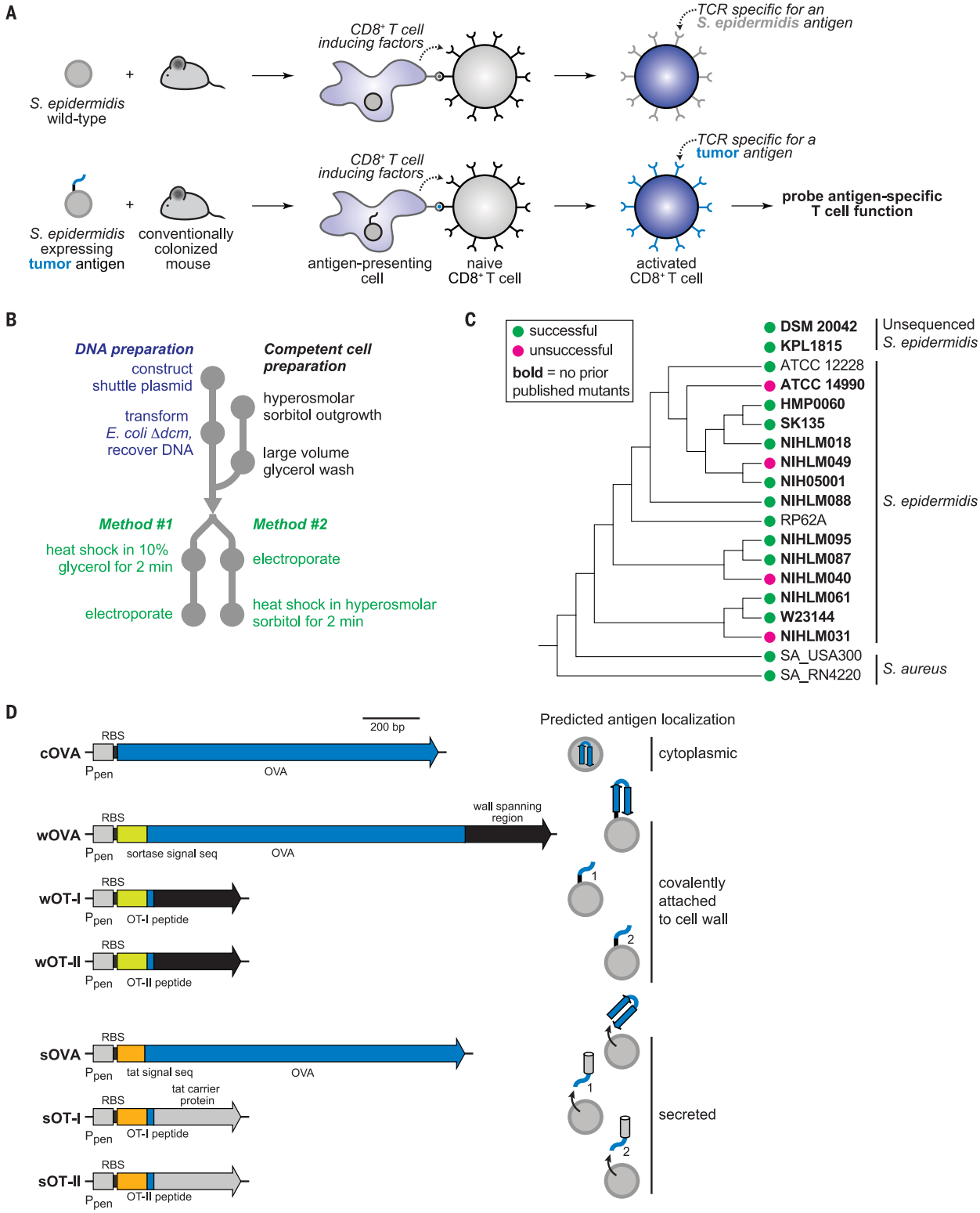


Fig. 1. Engineering *S. epidermidis* to express non-native antigens. (**A**) Experimental approach. By colonizing mice with an engineered strain of *S. epidermidis* that expresses a tumor antigen, we elicit T cells in vivo that are licensed by the commensal immune program but specific for a tumor. These T cells can then be subjected to functional assays in vivo, such as tumor killing. (**B**) Schematic of the genetic system that we developed for *S. epidermidis*. Advances include improved preparation of DNA and competent cells (materials and methods). *Staphylococcus* is subjected to heat shock before (Method #1) or after (Method #2) electroporation with the plasmid. (**C**) Phylogenetic tree of *S. epidermidis* and *S. aureus* strains. Circles indicate the strains for which our transformation method was successful (green) or unsuccessful (magenta). Strains in bold were not previously published as genetically tractable. (**D**) Design of the ovalbumin (OVA)–derived constructs expressed in *S. epidermidis* and their predicted localization. OVA is expressed either as a full-length protein (OVA), a class I major histocompatibility complex (MHC)–restricted antigenic peptide (OT-I, "1"), or a class II MHC-restricted antigenic peptide (OT-II, "2"). RBS, ribosome binding site; P_{pen}, promoter; bp, base pair.

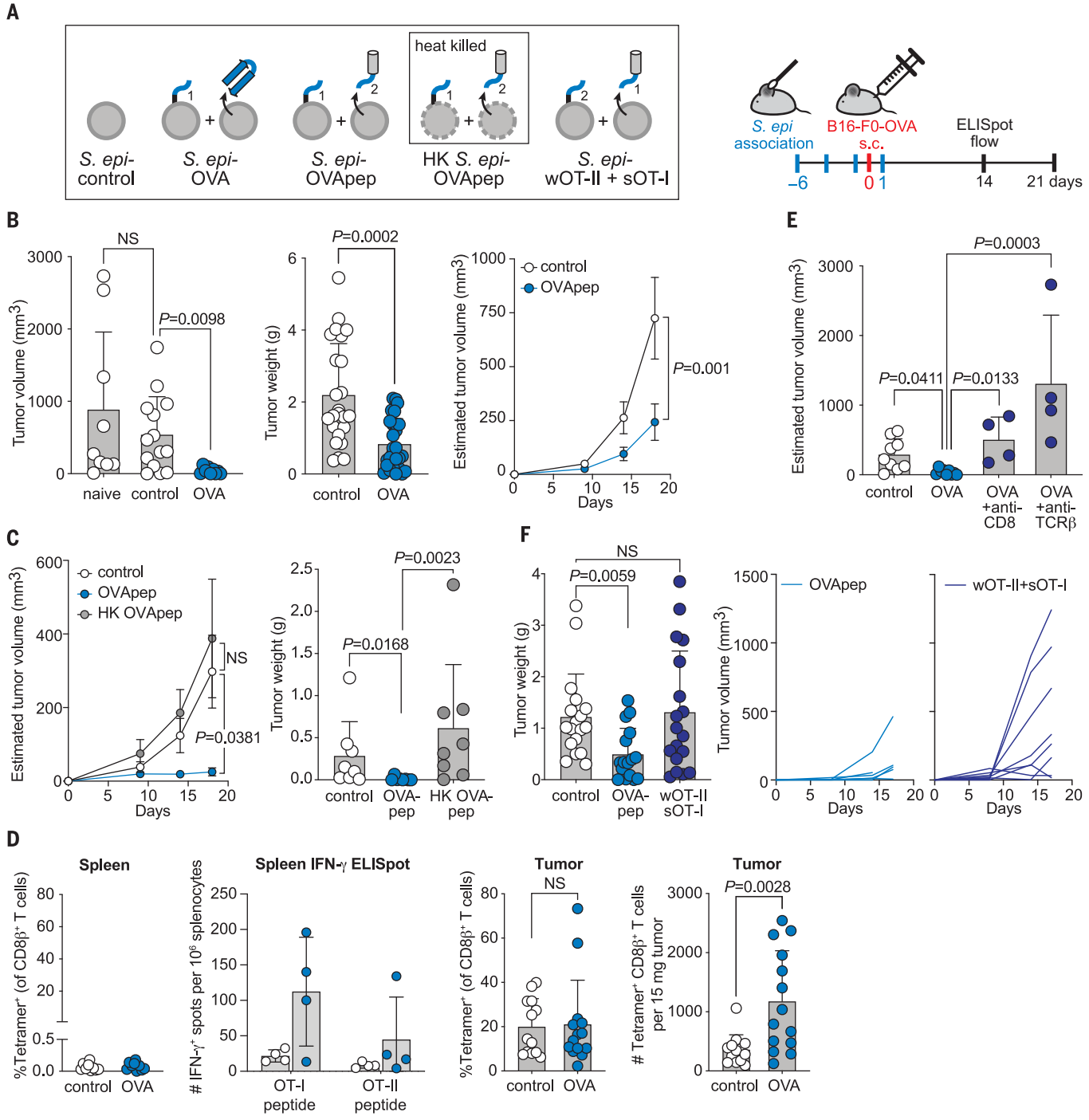


Fig. 2. Engineered *S. epidermidis* strains slow tumor progression and stimulate antigen-specific T cells in vivo. (**A**) Experimental schematic. Mice are naïve or colonized with live or heat-killed (HK; 95°C for 30 min) *S. epidermidis* strains (shown in box; OT-I, "1"; OT-II, "2") on days −6 to +1. When HK is used, mice are associated twice weekly throughout the entire experiment. On day 0, B16-F0-OVA melanoma cells are injected subcutaneously into the flank. Tissues are collected for analysis on days 14 or 21. (**B**) (Left) Day 19 blinded caliper measurements of subcutaneous B16-F0-OVA tumors (n = 5 to 10 mice per group). (Middle) Day 21 masses of dissected tumors (n = 24 to 28 mice per group). (Right) Blinded caliper measurements of tumors over time (n = 16 mice per group). (**C**) (Left) Blinded caliper measurements over time and (right) masses of subcutaneous B16-F0-OVA tumors on day 21 from mice associated with live or HK strains (gray dots). (**D**) Frequency of OT-I-specific T cells in the indicated organs as measured by H2-K^b-SIINFEKL tetramer staining or IFN-γ$^+$ ELISpot assay at day 14. Tetramer staining is gated on live CD90.2$^+$TCRβ$^+$CD8β$^+$ cells. (**E**) Endpoint caliper measurements of tumors from *S. epidermidis*–associated mice treated or not with anti-CD8α (2.43) or anti-TCRβ (H57-597) neutralizing antibodies (dark blue dots) (n = 4 to 9 mice per group). (**F**) Masses of subcutaneous B16-F0-OVA tumors on day 21 from mice colonized with *S. epidermidis* harboring control or various OVA constructs (n = 8 to 10 mice per group). For tumor burden bar graphs, nonparametric testing was used to generate P values (Mann-Whitney U test for two groups and Kruskal-Wallis H test for more than two groups). For flow cytometry data, parametric testing was used (unpaired Student's t test). For tumor-growth time courses, two-way analysis of variance (ANOVA, mixed-effects model) with multiple comparison testing was used. All experiments show unilateral tumors pooled from two [(B), left and right, (E), and (F)] or three [(B), middle] independent experiments.

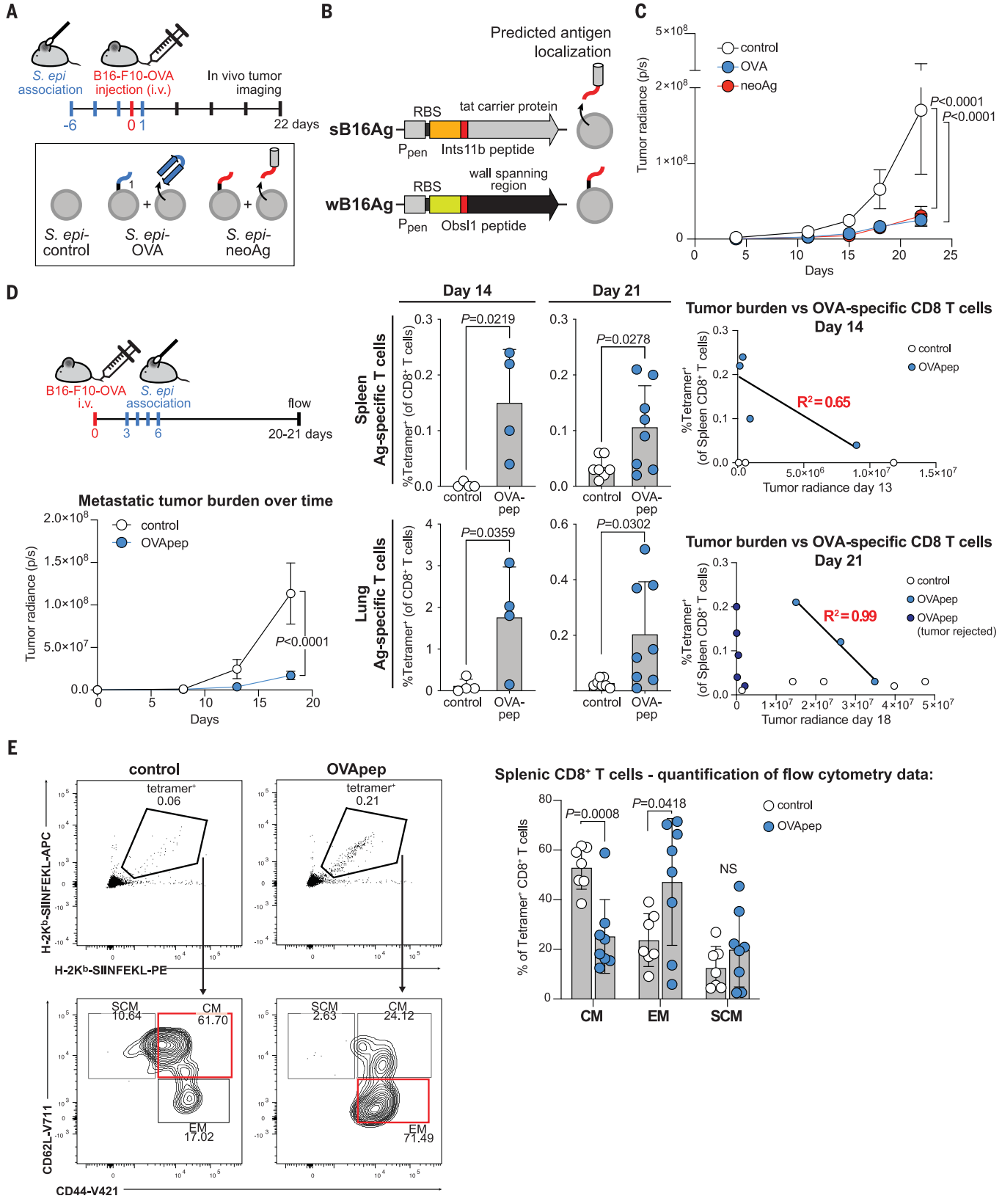


Fig. 3. Engineered *S. epidermidis* strains slow the progression of metastatic tumors. (**A**) Study design. Mice associated with *S. epidermidis* strains are injected intravenously (day 0) with B16-F10-OVA melanoma cells that constitutively express luciferase (OT-I, "1"). (**B**) Schematic of neoantigen expression constructs and their predicted subcellular localization within *S. epidermidis*. The neoantigen sequence (red box) encodes peptides centered around Obsl1(T1764M) for the wall-attached construct (wB16Ag) or Ints11(D314N) for the secreted construct (sB16Ag). (**C**) Average measurement of tumor bioluminescence over time (n = 20 mice per group, two experiments pooled). (**D**) Treatment of metastatic B16-F10-OVA melanoma with association of

S. epi-OVApep. (Left) Tumor burden over time as quantified by bioluminescence imaging. (Middle) Frequency of OT-I–specific T cells in the spleen at days 14 and 21 by H2-K^b-SIINFEKL tetramer staining. Cells are gated on live CD90.2$^+$ TCRβ$^+$ CD8β$^+$ cells. (Right) Frequency of tetramer$^+$ CD8$^+$ T cells versus tumor radiance. (**E**) Memory cell phenotypes of tetramer$^+$ CD8$^+$ T cells in the spleens of control or treated mice with metastatic melanoma. SCM, stem cell–like memory; CM, central memory; EM, effector memory. For tumor burden bar graphs, nonparametric testing was used to generate *P* values (Mann-Whitney *U* test for two groups and Kruskal-Wallis *H* test for more than two groups). For flow cytometry data, parametric testing was used (unpaired Student's *t* test). For tumor growth time courses, two-way ANOVA (mixed-effects model) with multiple comparison testing was used.

in which OT-I or OT-II was spliced into the secreted protein FepB (Tat pathway) (*26*, *27*) at the predicted signal sequence cleavage site or full-length OVA was fused to the N-terminal signal sequence (*S. epi*-sOT-I, *S. epi*-sOT-II, and *S. epi*-sOVA; "s" denotes "secreted."). We verified the production and localization of the OVA constructs by immunoblot and electron microscopy (fig. S1). Although *S. epidermidis* has no described Tat secretion system, the Tat signal peptide enabled efficient production and secretion of OVA.

Engineered strains of *S. epidermidis* stimulate antigen-specific T cells in vitro and in vivo

The engineered strains of *S. epidermidis* could stimulate antigen-specific T cells in vitro (fig. S2), so we tested whether they could prime T cells in vivo (fig. S3). Wild-type (WT) specific pathogen–free (SPF) C57BL/6 mice were topically colonized with *S. epi*-wOT-I or *S. epi*-sOT-II followed by the adoptive transfer of OT-I or OT-II cells. OT-I– or OT-II–expressing *S. epidermidis* could prime OVA-specific CD8$^+$ or CD4$^+$ T cells, respectively, without any barrier breach or skin inflammation (fig. S3A). Thus, under physiologic conditions of colonization, engineered *S. epidermidis* strains stimulate T cells specific for non-native epitopes.

S. epidermidis-OVA-induced T cells slow the progression of subcutaneous melanoma

We next sought to determine whether colonization with engineered *S. epidermidis* could elicit OVA-specific T cells from the native polyclonal repertoire, and if so, whether they would be functionally equipped to exert antitumor activity.

WT SPF C57BL/6 mice were injected subcutaneously with B16-F0-OVA, a C57BL/6-derived melanoma cell line that expresses ovalbumin. Mice were either unassociated (naïve) or gently swabbed with live *S. epidermidis* (Fig. 2A). Control mice were associated with a strain of *S. epidermidis* engineered to express an unrelated protein (*S. epi*-control). Treated mice were associated with a combination of *S. epi*-wOT-I and *S. epi*-sOVA (hereafter, "*S. epi*-OVA") or a combination of *S. epi*-wOT-I and *S. epi*-sOT-II (hereafter, "*S. epi*-OVApep") to elicit OVA-specific CD8$^+$ and CD4$^+$ T cells (fig. S3) because both T cell populations are required for optimal control of B16 melanoma (*28*). *S. epi*-OVA or *S. epi*-OVApep colonization markedly reduced tumor growth (Fig. 2B). *S. epi*-control colonization did not reduce tumor growth as compared with the absence of colonization (Fig. 2B). Thus, the antitumor effect is not due to an increased number of *S. epidermis*–specific CD8$^+$ T cells or a change in immune tone. This effect instead appears to be mediated by OVA-specific T cells because *S. epidermidis* exerts antitumor activity only when OVA is present.

Heat-killing *S. epi*-OVApep abolished the antitumor response (Fig. 2C), consistent with the observation that heat-killed *S. epi*-wOT-I failed to prime OT-I T cells in vivo (fig. S3B). Thus, the engineered bacterial colonist is not simply a source of antigen and adjuvant. Rather, bacterial viability, occupancy of specialized niches, and (potentially) prolonged antigen exposure may be required for the immune stimulatory response, even though no infection is mounted.

Next, we asked whether the antitumor effect was mediated by the OVA-specific T cells elicited by engineered *S. epidermidis*. First, compared with *S. epi*-control colonization, colonization with *S. epi*-OVApep caused increased infiltration of OVA-specific CD8$^+$ T cells into the tumor and an increase in OVA-specific interferon-γ (IFN-γ)$^+$ CD8$^+$ and CD4$^+$ T cells in the spleen (Fig. 2D and fig. S4A). The skin of tumor-bearing mice contained very limited numbers of OVA-specific CD8$^+$ T cells, which were biased toward the production of IFN-γ over interleukin-17A (fig. S4, B to D). Furthermore, there was no evidence of skin or systemic inflammation, in contrast to immunization with OVA-CFA (figs. S5 and S6). Thus, OVA-specific T cells elicited by *S. epidermidis* can migrate to and infiltrate a tumor without skin or systemic toxicity.

Second, we depleted CD8$^+$ T cells or all αβ T cells in tumor-bearing, *S. epi*-OVA–colonized mice. In both cases, the therapeutic effect of *S. epi*-OVA colonization was eliminated (Fig. 2E and fig. S4F). When mice were associated with either *S. epi*-wOT-I or *S. epi*-sOT-II rather than the combination of both, tumor weights were not significantly reduced, suggesting that OVA-specific CD4$^+$ and CD8$^+$ T cells were both required. When the subcellular localization of OT-I and OT-II in *S. epidermidis* was swapped by colonizing with *S. epi*-wOT-II plus *S. epi*-sOT-I, the antitumor effect was also lost, even though these strains are respectively capable of priming CD4$^+$ OT-II and CD8$^+$ OT-I cells (Fig. 2F and fig. S3B). Thus, the minimal requirements for antitumor efficacy are a wall-attached CD8$^+$ antigen and a CD4$^+$ antigen. Furthermore, in vivo priming of OVA-specific T cells does not completely predict their functional outcome. More broadly, our findings are consistent with a model in which *S. epidermidis*–induced T cells enter circulation, infiltrate the tumor, and are required for antitumor activity.

S. epidermidis–induced T cells are active in tissues distinct from the site of colonization

S. epi-OVA–induced T cells can limit subcutaneous tumor growth. Even though *S. epidermidis* is a skin colonist, we wondered whether *S. epidermidis*–induced T cells could traffic into other tissues. We reasoned that translocation of microbial colonists across barrier tissues can cause life-threatening sepsis, so one of the central goals of barrier immunity may be to limit systemic microbial dissemination.

To address this question, we used B16-F10, a more aggressive variant of B16, in a model of metastatic melanoma. Luciferase-expressing B16-F10-OVA cells were injected intravenously, resulting in lung metastases, which were quantified with in vivo luminescence imaging (Fig. 3A and fig. S7A). Topical colonization with *S. epi*-OVA before intravenous tumor cell injection slowed tumor progression substantially, and in some cases, promoted tumor rejection or near-rejection (Fig. 3C and fig. S7B). Thus, the antitumor effect of *S. epi*-OVA is not restricted to the skin and subcutaneous tissues.

The potency of the antitumor response induced by *S. epi*-OVA was unexpected. Even when used as monotherapy, colonist-induced T cells limited the growth of an aggressive tumor, which is notable because colonization is a mild treatment and does not induce any of the normal side effects of immune stimulation. This result is consistent with the view that *S. epidermidis* could be used as a tool for eliciting antigen-specific T cells in situ. Antigen-specific T cells generated by other methods (ex vivo engineering) are expensive (*29*), prone to exhaustion (*30*, *31*), and have had limited success in treating solid tumors (*32*). Thus, we tested whether *S. epi*–induced T cells have functional attributes that would enable them to be useful in a real-world setting, such as specificity for real-world antigens, activity against established tumors, and synergy with existing immunotherapies.

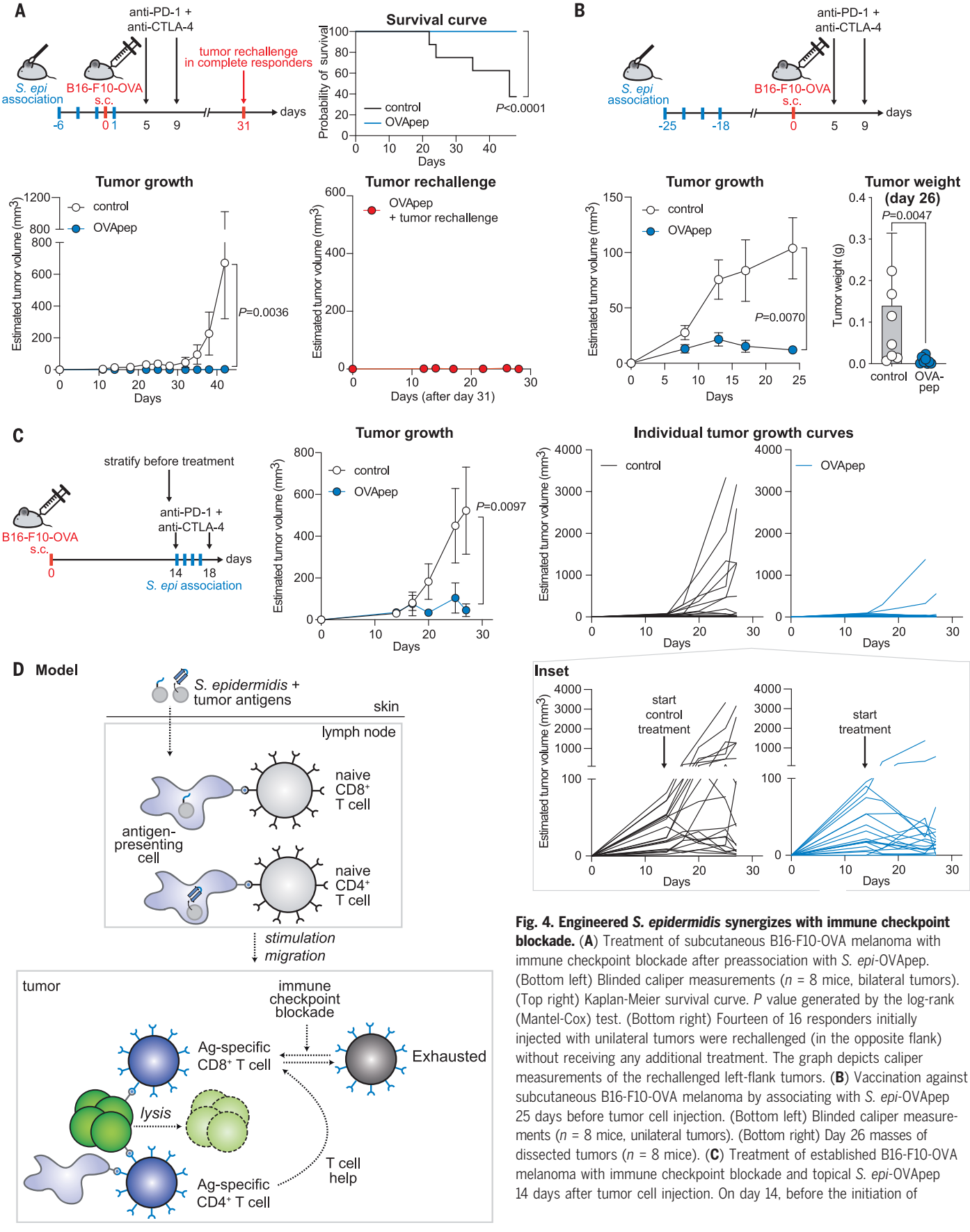


Fig. 4. Engineered *S. epidermidis* synergizes with immune checkpoint blockade. (**A**) Treatment of subcutaneous B16-F10-OVA melanoma with immune checkpoint blockade after preassociation with *S. epi*-OVApep. (Bottom left) Blinded caliper measurements (n = 8 mice, bilateral tumors). (Top right) Kaplan-Meier survival curve. P value generated by the log-rank (Mantel-Cox) test. (Bottom right) Fourteen of 16 responders initially injected with unilateral tumors were rechallenged (in the opposite flank) without receiving any additional treatment. The graph depicts caliper measurements of the rechallenged left-flank tumors. (**B**) Vaccination against subcutaneous B16-F10-OVA melanoma by associating with *S. epi*-OVApep 25 days before tumor cell injection. (Bottom left) Blinded caliper measurements (n = 8 mice, unilateral tumors). (Bottom right) Day 26 masses of dissected tumors (n = 8 mice). (**C**) Treatment of established B16-F10-OVA melanoma with immune checkpoint blockade and topical *S. epi*-OVApep 14 days after tumor cell injection. On day 14, before the initiation of

treatment, mice with no tumors or particularly large tumors (>100 mm^3) were excluded from the analysis. Blinded caliper measurements are shown (middle) averaged or (right) as growth curves of individual tumors. Inset with segmented *y* axis shows tumors initially growing and then regressing after treatment (n = 19 to 20 mice, unilateral tumors). For bar graphs, the Mann-Whitney *U* test was used to generate *P* values. For tumor growth time courses, two-way ANOVA with multiple comparison testing was used. (**D**) Model: Engineered strains of *S. epidermidis* colonize the skin and induce antigen-presenting cells to stimulate antigen-specific T cells, which traffic to the tumor and restrict tumor growth. Immune checkpoint blockade synergizes with engineered commensals. All data shown are representative of two independent experiments.

S. epidermidis producing a native neoantigen slows tumor progression

We asked whether *S. epidermidis* could elicit T cells specific for a real tumor antigen, and if so, whether their antitumor activity would be comparable to that of OVA-specific T cells. We engineered *S. epidermidis* to express two neoantigen-containing peptides (termed "*S. epi*-neoAg") naturally present in B16-F10 melanoma that were previously shown to generate $CD8^+$ and $CD4^+$ T cell responses (*33*) (Fig. 3B). Colonization with *S. epi*-neoAg before intravenous injection of B16-F10-OVA cells potently restricted metastatic tumor growth (Fig. 3, A to C, and fig. S7, A and B). A similar approach also reduced tumor progression in the TRAMP-C2 prostate carcinoma model (fig. S8) (*34*). Thus, commensal-induced T cells can be redirected against multiple host antigens and in different tumor contexts.

S. epi-OVApep colonization is effective after tumor cell injection

Next, we tested whether colonizing mice with engineered *S. epidermidis* after tumor cell injection would yield a therapeutic response. Using the intravenous B16-F10-OVA model of metastatic melanoma, we colonized with *S. epi*-OVApep starting 3 days after tumor inoculation (Fig. 3D). The reduction in tumor burden was comparable to the prophylaxis model and was accompanied by an increase in OVA-specific $CD8^+$ T cells in both the lung and spleen. The frequency of splenic OVA-specific $CD8^+$ T cells correlated well with the reduction in tumor burden in *S. epi*-OVApep–associated mice but not in control mice (Fig. 3D, right), supporting the hypothesis that *S. epi*-OVApep–induced $CD8^+$ T cells are important for metastatic tumor control.

Because control mice also had OVA-specific $CD8^+$ T cells circulating in the lung and spleen due to antigen exposure from the tumor, albeit a smaller number than in *S. epi*-OVApep–colonized mice, we asked whether the OVA-specific T cells differed functionally between control and treatment mice. We observed an unusual shift in the OVA-specific $CD8^+$ T cell pool from mostly central memory cells in control mice to effector or effector memory cells in *S. epi*-OVApep–associated mice (Fig. 3E and fig. S7D). Thus, *S. epi*-OVA–primed T cells are functionally distinct from tumor-primed T cells, even when they share the same antigen specificity. The effector phenotype of *S. epi*-OVA–primed T cells may explain their enhanced capacity to control tumors (*35*–*37*).

Engineered *S. epidermidis* synergizes with immune checkpoint blockade

The therapeutic effect of *S. epi*-OVA is likely mediated by effector and effector memory T cells that can become exhausted in the tumor environment (*38*, *39*), and a large proportion of *S. epi*-OVA–induced tumor-infiltrating lymphocytes (TILs) were positive for programmed cell death protein–1 (PD-1), a marker of exhaustion (fig. S9, A to C) (*40*). Therefore, we reasoned that adding immune checkpoint blockade would relieve the exhaustion of *S. epi*-OVA–induced TILs and augment antitumor activity.

We tested this hypothesis in the setting of subcutaneous B16-F10-OVA, which does not respond to immune checkpoint blockade alone (*41*, *42*). In contrast to *S. epi*-control colonization, *S. epi*-OVApep colonization combined with immune checkpoint blockade led to the rejection of 15 of 16 tumors (Fig. 4A). In a repeat experiment, we initially injected tumors in the right flank and 31 days later challenged the complete responders with left-flank tumors. All mice rejected the left-flank tumors, and 9 of 14 mice continued to be tumor free in the right flank 2 months after the initial tumor challenge (Fig. 4A). Furthermore, we could vaccinate mice against melanoma by associating them 25 days before tumor inoculation, by which time the effector T cell response to *S. epi*-OVApep would have contracted (Fig. 4B) (*3*). Thus, the memory T cell response elicited by earlier *S. epi*-OVApep association in combination with checkpoint blockade is sufficient to restrict tumor growth.

Next, we tested whether a combination of engineered *S. epidermidis* and checkpoint inhibition could treat established tumors. We waited either 5 or 14 days after B16-F10-OVA inoculation to colonize mice with *S. epi*-control or *S. epi*-OVApep and administer immune checkpoint blockade. The reduction in tumor burden was almost as pronounced as in the prophylaxis model, with regression of most tumors (Fig. 4C and fig. S9D). No inflammation, increase in ear thickness, or significant OVA-specific T cell infiltration of the skin was observed (fig. S4B and fig. S5, B to D). Thus, the combination of a tumor antigen-expressing commensal with checkpoint blockade yields a potent and durable antitumor response that can treat well-established tumors.

Discussion

Our approach, which uses a commensal microbe as the adjuvant and colonization as the mode of tumor antigen delivery, differs from previous strategies in important ways. In contrast to recent studies with native bacterial species from the gut microbiome, we find that the antitumor activity of *S. epidermidis* is not due to a change in the innate immune milieu or to an increase in colonist-specific $CD8^+$ T cells (*43*–*46*). Rather, it requires the tumor antigen to be present. Moreover, colonization with *S. epi*-OVA does not result in an infection, require intratumoral delivery, or generate any measurable inflammation, making it safer, simpler, and more specific than approaches that require inflammation or tissue infiltration for antitumor activity (*47*–*49*).

Our approach is more comparable to other means of generating tumor-specific T cells, such as ex vivo T cell engineering [to generate chimeric antigen receptor (CAR) T and TCR T cells] and tumor vaccines. Although much research still needs to be done, engineered *S. epidermidis* is safe and simple and yields a potent antitumor response in combination with checkpoint blockade. This is notable, given that B16-F10 melanoma is an unusually fast-growing and immunologically cold cancer. Other approaches that have yielded comparable treatment outcomes have been far more aggressive (*50*–*52*). Colonist-induced T cells appear to have different functional properties than antigen-specific T cells generated by other means. We found that *S. epi*-OVA–induced $CD8^+$ T cells were biased toward an effector memory (rather than central memory) phenotype, which may explain their potency within a tumor. In contrast to vaccination methods that are episodic, a colonist may provide continuous, durable stimuli over multiple weeks, leading to a prolonged supply of effector T cells.

The immune memory that we observed in combination with checkpoint blockade is notable, given that the plasmid was lost (fig. S10). This indicates that even a week of antigen exposure can induce a long-lasting response, even if this exposure occurs much earlier than the tumor challenge. In the future, a strain that expresses antigen stably may lead to more prolonged antigen exposure—the equivalent of a "prime" and a constant "boost"—and consequently, a more robust memory immune cell response. Furthermore, the observation of tumor

rejection in some mice in treatment mode, combined with the observation that *S. epidermidis* induces $CD8^+$ T cells in nonhuman primates (*12*), suggests this may be a credible strategy for translation.

More broadly, our results imply that it might be possible to engineer antigen expression into other commensal bacterial strains to elicit a wide range of antigen-specific immune cell responses. Understanding how to redirect each one may open the door to immunotherapies for other diseases and deepen our understanding of commensal–host interactions.

REFERENCES AND NOTES

1. Y. Yang *et al*., *Nature* **510**, 152–156 (2014).
2. S. Naik *et al*., *Science* **337**, 1115–1119 (2012).
3. S. Naik *et al*., *Nature* **520**, 104–108 (2015).
4. T. C. Scharschmidt *et al*., *Immunity* **43**, 1011–1021 (2015).
5. V. K. Ridaura *et al*., *J. Exp. Med.* **215**, 785–799 (2018).
6. H. Chung *et al*., *Cell* **149**, 1578–1593 (2012).
7. M. Xu *et al*., *Nature* **554**, 373–377 (2018).
8. E. Ansaldo *et al*., *Science* **364**, 1179–1184 (2019).
9. M. M. Wegorzewska *et al*., *Sci. Immunol.* **4**, eaau9079 (2019).
10. L. Cervantes-Barragan *et al*., *Science* **357**, 806–810 (2017).
11. K. Atarashi *et al*., *Science* **331**, 337–341 (2011).
12. J. L. Linehan *et al*., *Cell* **172**, 784–796.e18 (2018).
13. O. J. Harrison *et al*., *Science* **363**, eaat6280 (2019).
14. A. K. Palucka, L. M. Coussens, *Cell* **164**, 1233–1247 (2016).
15. N. N. Jarjour, D. Masopust, S. C. Jameson, *Immunity* **54**, 14–18 (2021).
16. T. Carvalho, F. Krammer, A. Iwasaki, *Nat. Rev. Immunol.* **21**, 245–256 (2021).
17. H. B. Streeter, D. C. Wraith, *Curr. Opin. Immunol.* **70**, 75–81 (2021).
18. S. Tamoutounour *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **116**, 23643–23652 (2019).
19. J. Y. H. Lee *et al*., *Microb. Genom.* **2**, e000077 (2016).
20. I. R. Monk, I. M. Shah, M. Xu, M.-W. Tan, T. J. Foster, *mBio* **3**, e00277-11 (2012).
21. I. R. Monk, J. J. Tree, B. P. Howden, T. P. Stinear, T. J. Foster, *mBio* **6**, e00315 (2015).
22. V. Winstel, P. Kühner, H. Rohde, A. Peschel, *Nat. Protoc.* **11**, 949–959 (2016).
23. J. L. Moran, N. N. Dingari, P. A. Garcia, C. R. Buie, *Bioelectrochemistry* **123**, 261–272 (2018).
24. S. K. Costa, N. P. Donegan, A.-R. Corvaglia, P. François, A. L. Cheung, *J. Bacteriol.* **199**, e00271-17 (2017).
25. M. Hansson *et al*., *J. Bacteriol.* **174**, 4239–4245 (1992).
26. L. Biswas *et al*., *J. Bacteriol.* **191**, 5921–5929 (2009).
27. E. Turlin, M. Débarbouillé, K. Augustyniak, A.-M. Gilles, C. Wandersman, *PLOS ONE* **8**, e56529 (2013).
28. J. Kline, L. Zhang, L. Battaglia, K. S. Cohen, T. F. Gajewski, *J. Immunol.* **188**, 2630–2642 (2012).
29. C. H. June, R. S. O'Connor, O. U. Kawalekar, S. Ghassemi, M. C. Milone, *Science* **359**, 1361–1365 (2018).
30. E. W. Weber *et al*., *Science* **372**, eaba1786 (2021).
31. R. C. Lynn *et al*., *Nature* **576**, 293–300 (2019).
32. L. Lindo, L. H. Wilkinson, K. A. Hay, *Front. Immunol.* **11**, 618387 (2021).
33. S. Kreiter *et al*., *Nature* **520**, 692–696 (2015).
34. B. A. Foster, J. R. Gingrich, E. D. Kwon, C. Madias, N. M. Greenberg, *Cancer Res.* **57**, 3325–3330 (1997).
35. S. van Duikeren *et al*., *J. Immunol.* **189**, 3397–3403 (2012).
36. M. Enamorado *et al*., *Nat. Commun.* **8**, 16073 (2017).
37. T. N. Gide *et al*., *Cancer Cell* **35**, 238–255.e6 (2019).
38. F. S. Hodi *et al*., *N. Engl. J. Med.* **363**, 711–723 (2010).
39. J. R. Brahmer *et al*., *N. Engl. J. Med.* **366**, 2455–2465 (2012).
40. K. Sakuishi *et al*., *J. Exp. Med.* **207**, 2187–2194 (2010).
41. D. Zamarin *et al*., *Sci. Transl. Med.* **6**, 226ra32 (2014).
42. C. Twyman-Saint Victor *et al*., *Nature* **520**, 373–377 (2015).
43. M. Sade-Feldman *et al*., *Cell* **175**, 998–1013.e20 (2018).
44. V. Gopalakrishnan *et al*., *Science* **359**, 97–103 (2018).
45. B. Routy *et al*., *Science* **359**, 91–97 (2018).
46. T. Tanoue *et al*., *Nature* **565**, 600–605 (2019).
47. Y. Paterson, P. D. Guirnalda, L. M. Wood, *Semin. Immunol.* **22**, 183–189 (2010).
48. C.-Z. Wang, R. A. Kazmierczak, A. Eisenstark, *Int. J. Microbiol.* **2016**, 5678702 (2016).
49. D. C. Binder *et al*., *Cancer Immunol. Res.* **1**, 123–133 (2013).
50. W. W. Overwijk *et al*., *J. Exp. Med.* **198**, 569–580 (2003).
51. W. W. Overwijk, N. P. Restifo, *Curr. Protoc. Immunol.* **39**, 20.1.1–20.1.29 (2001).
52. D. C. Palmer *et al*., *Proc. Natl. Acad. Sci. U.S.A.* **105**, 8061–8066 (2008).

ACKNOWLEDGMENTS

We are indebted to members of the Fischbach Group for helpful suggestions and comments on the manuscript. We are grateful to N. Collins, E. Enamorado, T. Yamamoto, R. Tussiwand, and S. Park for helpful discussions and to H. Pimentel and J. Rajniak for guidance on statistical analyses. We are indebted to the Stanford animal facility staff for help with animal husbandry. Cell sorting and flow cytometry analyses were performed on instruments in the Stanford Shared FACS Facility with help from M. Weglarz, L. Nichols, and R. Zermeno. Electron micrographs were generated by the Washington University Center for Cellular Imaging supported by Washington University School of Medicine in St. Louis, The Children's Discovery Institute of Washington University and St. Louis Children's Hospital (CDI-CORE-2015-505 and CDI-CORE-2019-813), and the Foundation for Barnes-Jewish Hospital (3770 and 4642). Electron microscopy and data generation were performed by G. Strout, S. Sviben, and J. Fitzpatrick. Phycoethrin-labeled and allophycocyanin-labeled H2-Kb-SIINFEKL tetramers were obtained through the NIH Tetramer Core Facility. *S. epidermidis* strain NIHLM087 was a gift from J. Segre (NIH). pMS182 (pLI50-Ppen-GFP-mut2) was a gift from S. Walker (Harvard University). B16-FO-OVA was a gift from N. Reticker-Flynn and E. Engleman (Stanford University). **Funding:** M.A.F. was supported by Open Philanthropy, the Howard Hughes Medical Institute (HHMI)–Simons Faculty Scholar Award, the Leona M. and Harry B. Helmsley Charitable Trust, the Chau Hoi Shuen Foundation, NIH grant DK110174, the Chan Zuckerberg Biohub, Stand Up to Cancer, the Stanford Microbiome Therapies Initiative, and MAC3 Impact Philanthropies. Y.E.C. was supported by an HHMI Hanna H. Gray Fellowship. D.B. was supported by an Early Postdoc Mobility Fellowship from the Swiss National Science Foundation. K.N. was supported by the Human Frontier Science Program LT000493/2018-L and a Fellowship of the Astellas Foundation for Research on Metabolic Disorders. Y.B. was supported by the Division of Intramural Research of NIAID (1ZIA-AI001115 and 1ZIA-AI001132). **Author contributions:** Conceptualization: Y.E.C., D.B., K.N., Y.B., and M.A.F. Methodology: Y.E.C., D.B., K.N., Y.B., and M.A.F. Investigation: Y.E.C., D.B., K.A., A.D., A.Z., N.E., V.K.Y., and A.V. Visualization: Y.E.C., D.B., and M.A.F. Funding acquisition: Y.E.C., D.B., K.N., Y.B., and M.A.F. Supervision: Y.E.C., D.B., Y.B., and M.A.F. Writing – original draft: Y.E.C. and M.A.F. Writing – review and editing: Y.E.C., D.B., A.V., Y.B., and M.A.F. **Competing interests:** M.A.F. is a cofounder and director of Federation Bio, a company developing microbiome-based therapeutics. Y.E.C. and K.N. were consultants for Federation Bio. Y.E.C., D.B., K.N., and M.A.F. are inventors on patent applications submitted by Stanford University and the Chan Zuckerberg Biohub that cover methods for using engineered bacteria to elicit antigen-specific immune cells (patent application no. US20220362358). **Data and materials availability:** All data are available in the main text or supplementary materials. **License information:** Copyright © 2023 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

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.abp9563
Materials and Methods
Figs. S1 to S10
Tables S1 and S2
References (*53*–*68*)
MDAR Reproducibility Checklist

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

Submitted 8 March 2022; accepted 28 February 2023
10.1126/science.abp9563

WORKING LIFE

By **Jorge Cortez Campaña**

Sand through the hourglass

Six months ago, I thought my academic career was finally taking off. Growing up, I was a curious child charmed by nature who dreamed of becoming a scientist. Because I was born into a poor family with little schooling, it was not an easy path. But I made it to university, the first in my family to do so, and chose to study mining engineering, which is highly valued in my home country of Chile. It seemed a prudent choice that would balance economic stability and intellectual satisfaction. By serving as a teaching assistant I also discovered I love teaching. Halfway through my degree, I had no doubts about the next step in my career: pursuing a Ph.D. so I could work at a university, conducting research and teaching. But I thought I could take my time.

Given my upbringing, concerns about economic stability soon overshadowed my passion for a research career. So, after finishing my bachelor's degree I found a part-time teaching position that offered a comfortable salary. Even so, it was easy to fall into a whirlpool of debt, buying a flat, a new car, holidays—each expense making it more difficult to start the Ph.D. I dreamed of. Time flew.

It took me 5 years and a pandemic to stabilize financially and go ahead with the doctorate. I found a Ph.D. group I really enjoy working in, and everything finally seemed to be on track. I was all set to write my first research proposal, conduct my first fieldwork, and do an internship abroad. My future prospects seemed limitless.



"I still don't know how long I'm going to live ... but I will keep going for as long as I can."

Then, I tore my calf muscle while playing with a kitten my partner and I had recently adopted. The doctor said it should heal on its own, but after a few weeks the pain had not gone away. The doctor sent me for scans to see if I might have blood clots. His worries were confirmed: I had clots not only in my right leg, but also in my lungs. I was in a panic; in just 3 weeks I was due to fly to England to start my internship. A blood thinner medication cleared up the clots, I felt good, and a new flight was booked. I thought this would just be a blip in my plans, one I would quickly forget.

But soon the pain in my leg returned, and in the back of my mind I knew I would not travel. The doctor found that, even though I was taking anticoagulants, I had developed new clots. Further examination found a 10-centimeter tumor in my liver: I had stage 4 cancer.

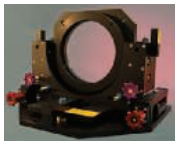
As a lover of numbers, I asked my oncologist what the survival rate was. One in 20 patients, he said. I heard the words but couldn't process what they meant. I asked again. Only 5% survive, he said. When I got home and did my own literature searching, I found the doctor's discouraging prognosis repeated in paper after paper. What had begun as the best year of my life had become perhaps the last year of my life. The worst part was not my own pain, but the sadness I was causing for the people who loved me—my partner, my sister, my parents.

I paused my work and studies to focus on my treatment, which includes chemotherapy and immunotherapy. My mother dropped everything to take care of me, for which I am immensely grateful.

My illness also spurred me to focus on what I really enjoy—and to realize I find even more satisfaction in my work than I had known. Although my Ph.D. is officially on hold, working on my research when I can helps me forget my disease for a few moments here and there, like a temporary escape door. And I'm proud to have made it to a Ph.D. program in the first place, even with the detours and delays. I don't regret the time it has taken—every step along the way has taught me important lessons and is a part of who I am. I am trying to approach my disease with the same perspective.

Thankfully I have responded well to my treatments and the tumor has shrunk by more than 50%. I still don't know how long I'm going to live—maybe a few years—but I feel like I have more time. I want to close circles in my life, including my Ph.D. I have so many ideas I want to work on and sometimes I don't know where I should start, with the sand running through the hourglass right behind me. But I will keep going for as long as I can. ■

Jorge Cortez Campaña is a researcher at the Pontifical Catholic University of Chile and the University of La Serena. Send your career story to SciCareerEditor@aaas.org.



Optical Mounts
Optical Surfaces offer a range of optical mounts designed to mechanically support parabolic, spherical, and flat mirrors from 100 mm to 600 mm (4 in to 24 in) in diameter. Employing high-precision optics in applications where optical stability is critical is pointless unless they are stress-free and positioned with absolute precision. All OS Series Mounts are constructed in black anodized aluminium and are proven in labs worldwide to provide long-running, trouble-free operation. Optical Surfaces is also experienced in designing and supplying custom optical mounts for securely holding and allowing precise positioning of larger-diameter, irregularly shaped, or lightweighted mirrors.
Optical Surfaces
For info: +44-208-668-6126
www.optisurf.com

Multimode Plate Reader
PerkinElmer launched the EnVision Nexus system, its fastest and most sensitive multimode plate reader to date, designed for demanding high-throughput screening (HTS) applications and to accelerate drug discovery efforts. The EnVision Nexus platform, with complementary microplates and optimized reagents from PerkinElmer's drug discovery reagents portfolio, including proprietary HTRF and AlphaLISA technologies and latest reagent kits, provides researchers increased assay flexibility. The EnVision Nexus system features high-throughput, exceptionally fast dual detectors, enabling researchers to screen millions of samples with increased accuracy, speed, and sensitivity. For walkaway convenience, the system can be equipped with a plate stacker for 20 or 50 plates and can be fully integrated for 24/7 workflow-driven automation. Easy to set up, use, and maintain, the new EnVision Nexus platform works across major detection technologies, including fluorescence intensity, fluorescence polarization, absorbance, luminescence, and time-resolved fluorescence (TRF).
PerkinElmer
For info: +1 781 663 6900
www.perkinelmer.com

Ruggetized Lenses for Space Projects
Resolve Optics is a leading designer and international supplier of ruggedized lenses and optical systems for spaceborne applications. The cost of launching payloads into space is considerable—incorporating only "space ready" ruggedized lenses or optical systems into your satellite or space observation system is therefore critical for achieving longer-term, high-performance operation. During launch into space, lenses in high-precision instruments and cameras are subject to vibration and shock. While slight vibration might cause malfunction in lenses, severe vibration can result in damage that is beyond repair. Shock waves are also highly detrimental to the performance and operation of high-precision optical devices. Over the past 20 years, Resolve Optics has developed the expertise to design ruggedized lenses and optical systems proven to withstand the vibration effects and shocks encountered during launch. To ensure all ruggedized lenses and optical systems produced can withstand the launch vibration and shock profile provided by the customer, Resolve Optics has invested in in-house testing equipment.
Resolve Optics
For info:+44-1494-777100
www.resolveoptics.com

Bst P DNA/RNA Polymerase (glycerol-free)
Bst P DNA/RNA Polymerase is an upgraded version of Bst DNA/RNA Polymerase through enzyme electronic restructure and evolution screening (in silico design & in vitro evolution), which is generally used for LAMP or RT-LAMP amplification of DNA or RNA. The whole Bst P DNA/RNA System includes hot-start aptamers, which ensures that the enzyme activity blocking efficiency is >95% at <30°C, and the enzyme activity is completely released within 1 min at >60°C. This characteristic facilitates the establishment of the reaction system at room temperature and greatly reduces the nonspecific amplification at low temperatures. The reaction temperature is further raised to 70°C, which greatly reduces the formation of primer dimer, improves the amplification specificity, and makes the nucleic acid release of crude samples more sufficient. Additionally, the whole portfolio contains Helicase, so Premium LAMP amplification (pLAMP) is allowed without using F3/B3 primers. At the same time, Helicase assists in strand unwinding, which further reduces the concentration of FIP/BIP primers. This will further reduce nonspecific amplification and greatly improve amplification homogeneity.
SBS Genetech
For info: tech@sbsbio.com
www.sbsgenetech.com

NOG Mice
FcResolv NOG models from Taconic Biosciences are the first and only super immunodeficient mouse models lacking residual expression of murine Fc gamma receptors (FcyRs), which are known to interact with antibody-based therapeutics and confound results. FcResolv models provide greater clarity, higher confidence, and improved translatability of study results by eliminating this known source of false positives and negatives. FcResolv models are ideal for investigating any indication which utilizes a therapy incorporating an Fc domain, including oncology, immuno-oncology, and autoimmune disease research. FcResolv NOG models are suitable for engraftment of tumor xenografts, human cells and tissue, and immune system humanization. FcResolv hIL-15 NOG models support human NK engraftment for studies on therapeutics with an ADCC mechanism of action.
Taconic Biosciences
For info: +1 (888) 822-6642
www.taconic.com/mouse-model/fcresolv-nog

Reactor Modules for Flow Chemistry
Uniqsis offers a range of reactor modules designed to enable labs with a limited budget to build flow chemistry systems without compromising performance. The Cold Coil IIFlow Reactor Module is designed to provide a flexible, entry-level solution for low-temperature flow chemistry applications. Used in conjunction with an external thermoregulation circulator, the Cold Coil II can maintain stable temperatures between -78°C and +150°C for extended periods of time. With an operating temperature range of ambient to +300°C, the HotCoil reactor heating module represents a cost-effective entry solution into flow chemistry. HotChip is a standalone heater module that provides precise temperature control of flow chemistry glass static mixer (GSM) chip reactors. The Uniqsis range of chemically resistant borosilicate glass chip reactor blocks can be used as static mixers to ensure effective mixing before a coil reactor or to perform fast exothermic reactions where reaction times are often less than 1 min.
Uniqsis
For info: +44-845-864-7747
www.uniqsis.com/paAccessories.aspx

Electronically submit your new product description or product literature information! Go to www.science.org/about/new-products-section for more information.

Newly offered instrumentation, apparatus, and laboratory materials of interest to researchers in all disciplines in academic, industrial, and governmental organizations are featured in this space. Emphasis is given to purpose, chief characteristics, and availability of products and materials. Endorsement by *Science* or AAAS of any products or materials mentioned is not implied. Additional information may be obtained from the manufacturer or supplier.