

# nature

## GENE GENIE

AI deciphers the evolutionary properties of DNA sequences

**Variety unpacked**  
How the pangenome project aims to capture human diversity

**Genomic imperative**  
100,000 sequences needed to preserve Africa's biodiversity

**Artificial biosynthesis**  
Constructed enzyme cascade generates cancer drug candidate

Vol. 598 | Number 7412

# Nature.2022.03.19

[ Sat, 19 Mar 2022]

- [This Week](#)
- [News in Focus](#)
- [Books & Arts](#)
- [Opinion](#)
- [Work](#)
- [Research](#)
- [Amendments & Corrections](#)

# This Week

- **[Are there limits to economic growth? It's time to call time on a 50-year argument](#)** [ 16 March 2022]  
Editorial • Researchers must try to resolve a dispute on the best way to use and care for Earth's resources.
- **[Gender equality will enhance research around the world](#)** [ 15 March 2022]  
Editorial • If researchers want to have maximum impact, women must be at the table.
- **[Open science, done wrong, will compound inequities](#)** [ 14 March 2022]  
World View • Research-reform advocates must beware unintended consequences.
- **[The geopolitics of fossil fuels and renewables reshape the world](#)** [ 11 March 2022]  
World View • To navigate the long road to net zero, energy researchers must grapple with the lessons of history.
- **[Turning up the heat on fat cells offers hope for treating obesity](#)** [ 08 March 2022]  
Research Highlight • Obese mice exposed to short bouts of heat therapy shed significant amounts of weight.
- **[The secrets of shark sleep](#)** [ 09 March 2022]  
Research Highlight • Look to a shark's posture, not its eyes, to gauge whether it's asleep.
- **[Paper strip holds high-accuracy, low-cost test for dreaded viruses](#)** [ 07 March 2022]  
Research Highlight • Screening method detects Zika and chikungunya viruses using materials that fit onto easily transported paper test cards.
- **[Microbial enzyme helps to churn out scarce COVID drug at full tilt](#)** [ 11 March 2022]  
Research Highlight • Synthesis of the COVID-19 antiviral molnupiravir is easier and more efficient when a bacterial enzyme is on the job.
- **[The world's rivers exhale a massive amount of carbon](#)** [ 07 March 2022]

Research Highlight • Nearly 6,000 measurements from rivers big and small allow scientists to estimate carbon emissions from Earth's waterways.

- **The hardy Hawaiian corals that could thrive in warming seas** [ 14 March 2022]

Research Highlight • Two species of tropical coral kept growing even as the water in their tanks became hotter and more acidic.

- **Madagascar's biggest mine achieves striking conservation success** [ 07 March 2022]

Research Highlight • Efforts to protect at-risk forests are projected to make up for the forest lost to an enormous mine.

- EDITORIAL
- 16 March 2022

# Are there limits to economic growth? It's time to call time on a 50-year argument

Researchers must try to resolve a dispute on the best way to use and care for Earth's resources.



Lead author Donella Meadows wrote that the book *The Limits to Growth* "was written not to predict doom but to challenge people to find ways of living that are consistent with the laws of the planet". Credit: Alamy

Fifty years ago this month, the System Dynamics group at the Massachusetts Institute of Technology in Cambridge had a stark message for the world: continued economic and population growth would deplete Earth's resources and lead to global economic collapse by 2070. This finding was from their 200-page book *The Limits to Growth*, one of the first modelling studies to forecast the environmental and social impacts of industrialization.

For its time, this was a shocking forecast, and it did not go down well. *Nature* called the study "another whiff of doomsday" (see [Nature 236, 47–49; 1972](#)). It was near-heresy, even in research circles, to suggest that some of the foundations of industrial civilization — mining coal, making steel, drilling for oil and spraying crops with fertilizers — might cause lasting damage. Research leaders accepted that industry pollutes air and water, but considered such damage reversible. Those trained in a pre-computing age were also sceptical of modelling, and advocated that technology would come to the planet's rescue. Zoologist Solly Zuckerman, a former chief scientific adviser to the UK government, said: "Whatever computers may say about the future, there is nothing in the past which gives any credence whatever to the view that human ingenuity cannot in time circumvent material human difficulties."

But the study's lead author, Donella Meadows, and her colleagues stood firm, pointing out that ecological and economic stability would be possible if action were taken early. *Limits* was instrumental to the creation of the United Nations Environment Programme, also in 1972. Overall, more than 30 million copies of the book have been sold.

But the debates haven't stopped. Although there's now a consensus that human activities have irreversible environmental effects, researchers disagree on the solutions — especially if that involves curbing economic growth. That disagreement is impeding action. It's time for researchers to end their debate. The world needs them to focus on the greater goals of stopping catastrophic environmental destruction and improving well-being.

Researchers such as Johan Rockström at the Potsdam Institute for Climate Impact Research in Germany advocate that economies can grow without making the planet unliveable. They point to evidence, notably from the Nordic nations, that economies can continue to grow even as carbon

emissions start to come down. This shows that what's needed is much faster adoption of technology — such as renewable energy. A parallel research movement, known as ‘post-growth’ or ‘degrowth’, says that the world needs to abandon the idea that economies must keep growing — because growth itself is harmful. Its proponents include Kate Raworth, an economist at the University of Oxford, UK, and author of the 2017 book *Doughnut Economics*, which has inspired its own global movement.

Economic growth is typically measured by gross domestic product (GDP). This composite index uses consumer spending, as well as business and government investment, to arrive at a figure for a country's economic output. Governments have entire departments devoted to ensuring that GDP always points upwards. And that is a problem, say post-growth researchers: when faced with a choice between two policies (one more green than the other), governments are likely to opt for whichever is the quicker in boosting growth to bolster GDP, and that might often be the option that causes more pollution.

A report published last week by the World Health Organization (see [go.nature.com/3j9xcpi](https://go.nature.com/3j9xcpi)) says that if policymakers didn't have a “pathological obsession with GDP”, they would spend more on making health care affordable for every citizen. Health spending does not contribute to GDP in the same way that, for example, military spending does, say the authors, led by economist Mariana Mazzucato at University College London.

Both communities must do more to talk to each other, instead of at each other. It won't be easy, but appreciation for the same literature could be a starting point. After all, *Limits* inspired both the green-growth and post-growth communities, and both were similarly influenced by the first study on planetary boundaries ([J. Rockström et al. \*Nature\* \*\*461\*\*, 472–475; 2009](https://doi.org/10.1038/nature08914)), which attempted to define limits for the biophysical processes that determine Earth's capacity for self-regulation.

Opportunities for cooperation are imminent. At the end of January, the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services announced a big study into the causes of biodiversity loss, including the role of economic systems. More than 100 authors from 40 countries and different fields will spend two years assessing the literature.

They will recommend “transformative change to the systems leading us to catastrophe”, says study co-chair, political scientist Arun Agrawal at the University of Michigan in Ann Arbor.

Another opportunity is an upcoming revision of the rules for what is measured in GDP. These will be agreed by countries’ chief statisticians and organized through the UN, and are due to be finalized in 2025. For the first time, the statisticians are asking how sustainability and well-being could be more closely aligned to GDP. Both post-growth and green-growth advocates have valuable perspectives.

Research can be territorial — new communities emerge sometimes because of disagreements in fields. But green-growth and post-growth scientists need to see the bigger picture. Right now, both are articulating different visions to policymakers, and there is a risk this will delay action. In 1972, there was still time to debate, and less urgency to act. Now, the world is running out of time.

*Nature* **603**, 361 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00723-1>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00723-1>

| [Section menu](#) | [Main menu](#) |

- EDITORIAL
- 15 March 2022

# Gender equality will enhance research around the world

If researchers want to have maximum impact, women must be at the table.



Efforts to improve life for women benefit both science and society. Credit: Getty

Last week, a group of inspiring scientists from all over the world met for a two-day online conference. The event, Breaking Barriers for Gender Equity Through Research, was organized and hosted by editors from across Springer Nature.

Delegates heard from women in disciplines such as mathematics and physics, in which the under-representation of women is particularly acute. There were discussions about gender equity in sustainable development, and

how women's lives are affected by crises such as wars, climate change and the current pandemic. There were conversations about mentorship, including the mentoring of men by women; about men championing women; and about the importance of teaching and learning listening skills.

The COVID-19 pandemic has been particularly hard on female researchers. Even before SARS-CoV-2 emerged, women, on average, published less than men, and faced additional and potentially long-lasting career setbacks. Last week's event, which coincided with this year's International Women's Day, with its theme of 'break the bias', was an opportunity to do something different: to reflect on the events of the past two years and to showcase ongoing efforts to advance gender equality in research and education, especially in low- and middle-income countries (LMICs) in Africa, Asia and the Middle East.

## **Biases can build up**

What speakers called compounded discrimination — also known as intersectionality — was a recurring theme. Someone experiencing discrimination on the basis of their gender might face further discrimination if they belong to an under-represented group or minority community. In many countries, for example, women face entrenched social biases about what they can and can't do, because of their gender — but if they are from a family or community in which incomes are low, they probably also have less access to finance, fewer opportunities for education and less chance of getting a skilled job.

Intersectionality isn't widely recognized in public policy in many countries, and this might be hindering the United Nations' Sustainable Development Goals (SDGs), which are designed to end poverty and protect the environment. Goal 5 is dedicated to achieving gender equality and empowering all women and girls. Only recently has there been widespread acceptance that achieving gender equality will help to achieve the other SDGs. Among them are the goals to end hunger, boost employment and achieve universal education. This underscores a need to accelerate a shift from addressing gender equality on its own to tackling it alongside other causes of inequality.

Addressing diversity in a range of contexts benefits everyone in tangible ways, and scientists' perspectives can help to address complex, real-world problems. For example, a better understanding of sex-based differences in immunity is helping to advance cancer immunotherapies ([Y. Ye et al. \*Nature Commun.\* 11, 1779; 2020](#)). And including data on gender in climate models provides a better understanding of different people's vulnerability to climate risks. Moreover, the safety of products and medical procedures can be compromised when differences between women and men aren't taken into account. In communities around the world, gender equality brings with it more resilient economies, reduced child mortality and better nutrition.

## Action needed now

The SDGs are designed to drive change in the longer term, but people who face discrimination need action now, and several initiatives are making progress on this front in research. The European Union's science funding programmes are embedding requirements on sex and gender equality into the grant application stage — and such rules have a ripple effect. If researchers elsewhere want to partner with European institutions, for example, they need to press their institutions to put a gender equality plan in place.

Many research institutions are also actively working to improve gender equality. One conference speaker, Segenet Kelemu, the director-general of the International Centre of Insect Physiology and Ecology, based in Nairobi, described a project to train 100,000 young people in Ethiopia — more than half of whom are women — in the science and business of bee-keeping and silkworm farming.

Similarly, international development agencies have been funding research on how public transport in LMICs can be made safer and more accessible to women ([see \[go.nature.com/3khoeda\]\(#\)](#)).

The results of these projects are clearly crucial to the communities involved, but they are not always widely shared. This needs to change so that others can benefit from the knowledge, and use it or build on it in their own local contexts. Institutions and funding bodies could make a big difference by

considering how local projects might be scaled up, providing a more meaningful impact.

The need to achieve gender equality across various spheres of society, including research, is more than a response to pressure brought to bear by equality targets or to the equality setbacks caused by the pandemic. It's an acknowledgement that being inclusive pays off across the board. Employees at gender-diverse institutions tend to be more satisfied and less likely to leave, and demonstrate better productivity, innovation and decision-making. So, if the global research community wants to have maximum impact, women must be at the table.

*Nature* **603**, 362 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00722-2>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00722-2>

| [Section menu](#) | [Main menu](#) |

- WORLD VIEW
- 14 March 2022

# Open science, done wrong, will compound inequities



Research-reform advocates must beware unintended consequences.

- [Tony Ross-Hellauer](#) 0

Ten years ago, as a new PhD graduate looking for my next position, I found myself in the academic cold. Nothing says “you are an outsider” more than a paywall asking US\$38 for one article. That fuelled my advocacy of open science and, ultimately, drove me to research its implementation.

Now, open science is mainstream, increasingly embedded in policies and expected in practice. But the ways in which it is being implemented can have unintended consequences, and these must not be ignored.

Since 2019, I've led ON-MERRIT, a project funded by the European Commission that uses a mixture of computational and qualitative methods to investigate how open science affects the research system. Many in the movement declare equity as a goal, but reality is not always on track for that. Indeed, I fear that without more critical thought, open science could become just the extension of privilege. Our recommendations for what to consider are out this week ([see go.nature.com/3kypbj8](https://go.nature.com/3kypbj8)).

Open science is a vague mix of ideals. Overall, advocates aim to increase transparency, accountability, equity and collaboration in knowledge production by increasing access to research results, articles, methods and tools. This means that data and protocols should be freely shared in high-quality repositories and research articles should be available without subscriptions or reading fees.

Making all that happen is expensive. Wealthy institutions and regions can afford this better than can poorer ones. At my university, in a high-income nation, I know I am privileged. In a collaboration to introduce open science at Ukrainian universities (including those displaced by conflict post-2014), I've been privy to difficult conversations about how to pay publication fees that are three times a professor's monthly salary, and how to meet data-sharing requirements to be eligible for funding when institutional support is lacking. And privilege comes in many forms. For instance, the fact that career-advancement criteria don't reward open practices puts early-career adherents at a disadvantage.

Failing to address structural inequalities directly means that the advantages of those who are already privileged will grow, especially given that they have the most influence over how open science is implemented.

A particularly pressing issue is open access (OA) publication fees, in which the benefit of free readership is being offset by new barriers to authorship. To support OA publishing, journals commonly charge authors, and charges are rising as the practice expands. My group and others have found that article-processing charges are creating a two-tier system, in which richer research teams publish more OA articles in the most prestigious journals. One analysis of 37,000 articles in hybrid 'parent' journals and their fully OA 'mirrors' (with the same editorial board and acceptance standards)

found that the geographic diversity of authors was much greater for non-OA articles than for OA articles ([A. C. Smith et al. \*Quant. Sci. Stud.\* 2, 1123–1143; 2022](#)). Another analysis found that authors of OA articles were more likely to be male, senior, federally funded and working at prestigious universities ([A. J. Olejniczak and M. J. Wilson \*Quant. Sci. Stud.\* 1, 1429–1450; 2020](#)). Worse still, citation advantages linked to OA mean that the academically rich will get even richer.

That open science can increase inequity should alarm science reformers. At the very least, they should commit to monitoring how researchers are affected.

To be sure, equity is not the sole priority for many advocates. When my team first announced its project, some critics objected. They argued that the key aims of open science are to improve research integrity by making processes and products more amenable to inspection, and boosting efficiency by making others' work reusable. Still, as our work has shown, equity is often cited as a key aim ([T. Ross-Hellauer et al. \*R. Soc. Open Sci.\* 9, 211032; 2022](#)).

Even those rooting for equity often argue that we should first enable access and then consider unintended side effects, such as marginalization of authors from low-income countries. But how change is implemented will have long-lasting consequences. Once new forms of inequity are in place, it will be too late to fix the system efficiently.

How can we ensure that open knowledge creation becomes fairer than it is now? First, we need more shared understanding and global dialogue. Open science is an umbrella term for a coalition of diverse practices with sometimes conflicting aims of transparency, participation and equity. We desperately need to untangle these.

Second, reform should encompass the research system as a whole, rather than country- or region-based policies that target specific practices. The UNESCO Recommendation on Open Science is an example of how this can work. Our recommendations include more focus on shared infrastructure, as well as on who participates and how. That means discussing ways to have

open access without publishing fees, plus making open practices easier, cheaper and more valued by promotion and grant evaluators.

I do think that open science can bring much good. Like many of its advocates, I aim to make research more accessible and collaborative and to establish a system that rewards current merit, not past success or current privilege. Any potential for open science to compound inequalities must be vigilantly monitored by the academic community — otherwise we idealists risk scoring an own goal.

*Nature* **603**, 363 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00724-0>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00724-0>

| [Section menu](#) | [Main menu](#) |

- WORLD VIEW
- 11 March 2022

# The geopolitics of fossil fuels and renewables reshape the world



To navigate the long road to net zero, energy researchers must grapple with the lessons of history.

- [Helen Thompson](#) 0

With 84% of our energy still coming from oil, coal and gas, much of the transition to renewable energy sources lies ahead. Just because a different future will arrive, it doesn't mean that the present will simply cede the stage. Renewables do not change the centrality of energy to geopolitics. Nor, given that the energy transition will be a long one, will it quickly end the geopolitics of fossil fuels.

For nearly 200 years, fossil-fuel energy has been central to geopolitics. The relationship between western Europe and China changed decisively in

1839, when Britain deployed coal-fired steam ships in the First Opium War. This move opened up China to a succession of imperial powers. The turn to oil in the twentieth century made the United States the world's dominant power and began the decline of Europe's great powers. For the past decade, the United States and Russia have competed with each other to sell gas to Europe, as they did oil at the start of the past century.

Energy makes for dramatic geopolitical conflicts with after-effects that last decades. Take the Suez Crisis in 1956. US president Dwight Eisenhower used his country's financial might to stop Anglo-French military action against Egypt that was designed to protect western European energy interests in the Middle East. The United States had encouraged these interests, wanting to protect supply from the Western Hemisphere for itself. Aghast that their supposed NATO ally could betray them, several European countries began their turn towards what was then Soviet, and is now Russian, oil. In the 1970s, this Soviet-European energy relationship was extended to gas.

Ever since Russia's President Vladimir Putin first made it clear in 2008, in Georgia, that he does not accept the borders created by the dissolution of the Soviet Union, this dependency has constrained the European Union's policy towards Russia. Complementary fossil-fuel interests have also turned China and Russia into tacit allies.

Energy research — from materials science to emissions modelling to carbon pricing — that does not account for such realities can only ever give a partial picture.

As early as the 1990s, it was evident that addressing climate change would be constrained by geopolitics, and that choices about which new energy sources to develop would have geopolitical consequences. The United States refused to ratify the 1997 Kyoto Protocol on greenhouse-gas emissions because almost the entire Senate thought that an agreement that imposed obligations on the United States but not on China — classed as a developing country — would disadvantage the US economy. Meanwhile, the 1998–2005 coalition government in Berlin moved to renewable energy and began to phase out nuclear power, deepening German dependency on

Russian gas. At the same time, Putin began a two-decade strategic effort to remove Ukraine from Russia's gas -transport system.

Climate change does create acute incentives for cooperation between geopolitical rivals, especially the world's two largest carbon emitters: China and the United States. Despite the deterioration in Sino-US relations from around 2010, president Barack Obama struck an emissions agreement with Chinese President Xi Jinping in November 2014, which was the essential prelude to the Paris climate accord the following year. Yet even this moment of US–Chinese cooperation could not transcend geopolitics. In the same year, Xi also reached an agreement with Putin to build the Power of Siberia gas pipeline. This opened in 2019 and is the first to take gas east to Asia rather than west to Europe. For China, that is at least as important as an accommodation with Washington DC over climate.

States are competing to manufacture green-energy infrastructure, such as solar panels and wind turbines, and to mass-produce electric vehicles. In May 2015, the Chinese Communist Party announced a plan — Made in China 2025 — to turn the country into a superpower of high-tech manufacturing, including of electric vehicles, and ensure that it produces 70% of the core resources necessary. Former US president Donald Trump's trade and technology war with Beijing was mainly a response to this Chinese ambition, and gained cross-party support.

There is a discernible fear in Washington DC that an age of green energy will be the age of China. Renewables infrastructure depends heavily on rare-earth minerals, whose production China almost entirely dominates. Deng Xiaoping, a former leader of the Chinese Communist Party, once quipped: "The Middle East has oil and China has rare earths." Over the past decade, China has also been willing to use this control as a geopolitical weapon, imposing an export ban on all rare earths to Japan in 2010 after a conflict about a fishing trawler in the East China Sea. For the United States, playing catch-up on creating a domestic industry around the extraction of 'tech metals' has become a national imperative.

Quite simply, there is no way that governments — or the scholars who seek to advise them — can be serious about the energy transition without having a realistic strategy for the problems that history tells us will arise as the

geopolitics of old and new energy sources and technologies combine. Unless these predicaments are faced — by citizens as energy consumers, by scientists and social scientists, as well as by governments — they will become ever harder.

*Nature* **603**, 364 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00713-3>

---

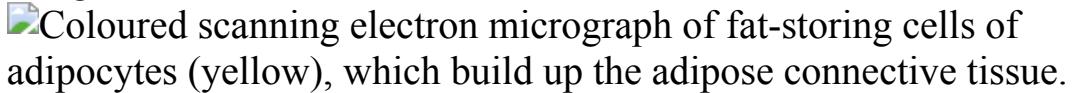
This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00713-3>

| [Section menu](#) | [Main menu](#) |

- RESEARCH HIGHLIGHT
- 08 March 2022
- Correction [11 March 2022](#)

# Turning up the heat on fat cells offers hope for treating obesity

Obese mice exposed to short bouts of heat therapy shed significant amounts of weight.



Coloured scanning electron micrograph of fat-storing cells of adipocytes (yellow), which build up the adipose connective tissue.

Fat cells (yellow; artificially coloured). Low levels of external heating can cause such cells, in both humans and mice, to burn their stored energy.  
Credit: Prof. P. Motta/Dept. of Anatomy/University "La Sapienza"  
Rome/Science Photo Library

Warming fat tissue in both humans and mice stimulates fat cells to produce heat, and might be an effective treatment to counter obesity<sup>1</sup>.

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

[Buy](#)

All prices are NET prices.

### **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

*Nature* **603**, 366 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00646-x>

### **Updates & Corrections**

- **Correction 11 March 2022:** The original version of this story listed the wrong reference. This has now been corrected.

### **References**

1. Li, Y. *et al.* *Cell* <https://doi.org/10.1016/j.cell.2022.02.004> (2022).

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00646-x>

| [Section menu](#) | [Main menu](#) |

- RESEARCH HIGHLIGHT
- 09 March 2022

# The secrets of shark sleep

Look to a shark's posture, not its eyes, to gauge whether it's asleep.

 Carpet shark (*Cephaloscyllium isabellum*) at the sea floor of Doubtful Sound, Fiordland National Park, New Zealand.

Are you sleeping? That's one way for a draughtsboard shark (above) to save energy. Credit: Nature Picture Library/Alamy

All animals sleep, but not every species sleeps like we do.

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.  
VAT will be added later in the checkout.  
Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

[Buy](#)

All prices are NET prices.

### **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

*Nature* **603**, 366 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00645-y>

## **References**

1. Kelly, M. L. *et al.* *Biol. Lett.* <https://doi.org/10.1098/rsbl.2021.0259> (2022).

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00645-y>.

- RESEARCH HIGHLIGHT
- 07 March 2022

# Paper strip holds high-accuracy, low-cost test for dreaded viruses

Screening method detects Zika and chikungunya viruses using materials that fit onto easily transported paper test cards.

 Coloured transmission electron micrograph (TEM) of chikungunya viruses (red) infecting a cell.

Infection by chikungunya virus (red; artificially coloured) can be detected with a newly developed test that is both accurate and inexpensive. Credit: O. Schwartz, M. Sourisseau, MC. Prevost, Institute Pasteur/Science Photo Library

A cheap, easy way to diagnose Zika and chikungunya infections could streamline testing for these pathogens in areas lacking expensive laboratory equipment, field experiments show<sup>1</sup>.

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

[Buy](#)

All prices are NET prices.

### **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

*Nature* **603**, 366 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00644-z>

## **References**

1. Karlikow, M. *et al.* *Nature Biomed. Eng.* <https://doi.org/10.1038/s41551-022-00850-0> (2022).
2. Pardee, K. *et al.* *Cell.* **165**, 1255-1266 (2016).

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00644-z>

| [Section menu](#) | [Main menu](#) |

- RESEARCH HIGHLIGHT
- 11 March 2022

# Microbial enzyme helps to churn out scarce COVID drug at full tilt

Synthesis of the COVID-19 antiviral molnupiravir is easier and more efficient when a bacterial enzyme is on the job.



Technicians wearing protective suits produce molnupiravir, the first tablet drug used to treat coronavirus disease (COVID-19).

The antiviral pill molnupiravir is produced using a cumbersome process that can be streamlined with help from a bacterial enzyme. Credit: Ziad Ahmed/NurPhoto via Getty

An engineered bacterial enzyme turbocharges the manufacture of molnupiravir, an antiviral drug used to treat COVID-19<sup>1</sup>.

Molnupiravir is one of the few approved drugs that can prevent severe COVID-19. But current methods for preparing it require multiple steps and yield relatively low amounts of the drug.

A more efficient route involves a precursor molecule called *N*-hydroxycytidine: just one chemical reaction converts it to molnupiravir. Anthony Green at the University of Manchester, UK, and his colleagues observed that, under certain conditions, an enzyme from the bacterium *Escherichia coli* could generate small amounts of the precursor from the common biomolecule cytidine.

The team mutated the DNA encoding the enzyme, improving the enzyme's efficiency, and used *E. coli* cells to make the upgraded enzyme. Finally, the researchers optimized the synthetic process by lowering the reaction's

temperature. That prompted *N*-hydroxy-cytidine to crystallize in high purity from solution as soon as it formed.

In a scaled-up demonstration, the researchers prepared 137 grams of *N*-hydroxy-cytidine, converting 71% of reactants to the precursor molecule in under three hours. During that time, each enzyme molecule facilitated more than 85,000 chemical transformations.

*Nature* **603**, 366 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00675-6>

## References

1. Burke, A. J. *et al.* *J. Am. Chem. Soc.* **144**, 3761–3765 (2022).

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00675-6>

| [Section menu](#) | [Main menu](#) |

- RESEARCH HIGHLIGHT
- 07 March 2022

# The world's rivers exhale a massive amount of carbon

Nearly 6,000 measurements from rivers big and small allow scientists to estimate carbon emissions from Earth's waterways.

Aerial view of landscape with a river coloured by glacial melt in Kirkjubæjarklaustur, Sudurland, Iceland.

A river in Iceland. The carbon emissions of rivers in northern regions vary more from month to month than do those of tropical rivers. Credit: Daniel Bosma/Getty

Scientists now have a global view of the amount of carbon dioxide released into Earth's atmosphere by rivers and streams throughout the year<sup>1</sup>.

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

[Buy](#)

All prices are NET prices.

### **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

*Nature* **603**, 367 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00674-7>

## **References**

1. Liu, S. *et al.* *Proc. Natl Acad. Sci. USA* **119**, e2106322119 (2022).

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00674-7>

- RESEARCH HIGHLIGHT
- 14 March 2022

# The hardy Hawaiian corals that could thrive in warming seas

Two species of tropical coral kept growing even as the water in their tanks became hotter and more acidic.



A flame angelfish (*Centropyge loricula*) shelters in finger coral (*Porites compressa*), which might be resilient to rising water temperatures and acidity. Credit: Thomas Kline/Design Pics/Getty

Some species of coral might be able to adapt to a world altered by climate change, at least if countries curb their greenhouse-gas emissions<sup>1</sup>.

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

[Buy](#)

All prices are NET prices.

### **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

*Nature* **603**, 367 (2022)

*doi:* [\*https://doi.org/10.1038/d41586-022-00719-x\*](https://doi.org/10.1038/d41586-022-00719-x)

## **References**

1. McLachlan, R. H. *et al. Sci. Rep.* **12**, 3712 (2022).

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00719-x>

- RESEARCH HIGHLIGHT
- 07 March 2022

# Madagascar's biggest mine achieves striking conservation success

Efforts to protect at-risk forests are projected to make up for the forest lost to an enormous mine.

 Indri-Indri is the largest living species of lemurs and are found around a giant nickel-cobalt mine in Madagascar.

The indri (*Indri indri*), the world's largest lemur, lives in the forest around an immense mine in eastern Madagascar. Credit: Sébastien Desbureaux

An ambitious goal to prevent an enormous nickel mine from causing net forest loss is on track for success, according to a statistical analysis<sup>1</sup>.

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

[Buy](#)

All prices are NET prices.

### **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

*Nature* **603**, 367 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00643-0>

## **References**

1. Devenish, K., Desbureaux, S., Willcock, S. & Jones, J. P. G. *Nature Sustain.* <https://doi.org/10.1038/s41893-022-00850-7> (2022).

| [Section menu](#) | [Main menu](#) |

# News in Focus

- **[Moon junk, kids' COVID antibodies and 'first stars' debate](#)** [ 16 March 2022]  
News Round-Up • The latest science news, in brief.
- **['I lost two years of my life': US scientist falsely accused of hiding ties to China speaks out](#)** [ 07 March 2022]  
News • Acquitted nanotechnology researcher Anming Hu returns to his lab after two years — and is still grappling with the aftershocks of his ordeal.
- **[Major CRISPR patent decision won't end tangled dispute](#)** [ 09 March 2022]  
News • Fights over who invented the gene-editing technology are becoming more complex, and could carry on for years.
- **[Police rely on radiocarbon dating to identify forged paintings](#)** [ 09 March 2022]  
News • Advances in the technique have bolstered its reputation as a tool for investigating faked artwork.
- **[ResearchGate dealt a blow in copyright lawsuit](#)** [ 04 March 2022]  
News • Court ruling says that the academic-networking website is responsible for papers uploaded by its users — but ResearchGate plans to appeal the decision.
- **[This US Supreme Court decision could derail Biden's climate plan](#)** [ 03 March 2022]  
News Explainer • Controversial lawsuit has put the US government's ability to slash carbon emissions on the line.
- **[A more-inclusive genome project aims to capture all of human diversity](#)** [ 16 March 2022]  
News Feature • Researchers are building a human 'pangenome' that would represent the entirety of human genetic variation. But not everyone is ready to sign on.

- NEWS ROUND-UP
- 16 March 2022

# Moon junk, kids' COVID antibodies and 'first stars' debate

The latest science news, in brief.



Some children produce no antibodies after having COVID-19.Credit: Varuth Pongsapipatt/SOPA/LightRocket via Getty

## Kids have low levels of COVID antibodies

Children infected with SARS-CoV-2 are about half as likely as adults to produce antibodies against the virus, despite having similar symptoms and

levels of virus in their bodies, according to a small study in Australia ([Z.Q. Toh et al. \*JAMA Netw. Open\* 5, e221313; 2022](#)).

The team looked at 57 children with a median age of 4 and 51 adults with a median age of 37, who tested positive for SARS-CoV-2 in late 2020. They found that children and adults had similar viral loads, but [only 37% of the children produced SARS-CoV-2 antibodies](#), compared with 76% of the adults. The findings were published on 9 March.

Researchers say that children could be producing fewer antibodies because they have a more robust innate immune response than adults. This is the first line of defence against pathogens, and is non-specific.

Children could also be better at responding to infections where they enter the body. This means that the body clears the virus quickly and it doesn't "hang around" to trigger antibody production, says Donna Farber, an immunologist at Columbia University in New York City.

But because antibodies are likely to be important against reinfection, the findings raise questions about how well protected children might be against future infections.

## **Space junk's Moon crash adds to 60+ years of lunar debris**

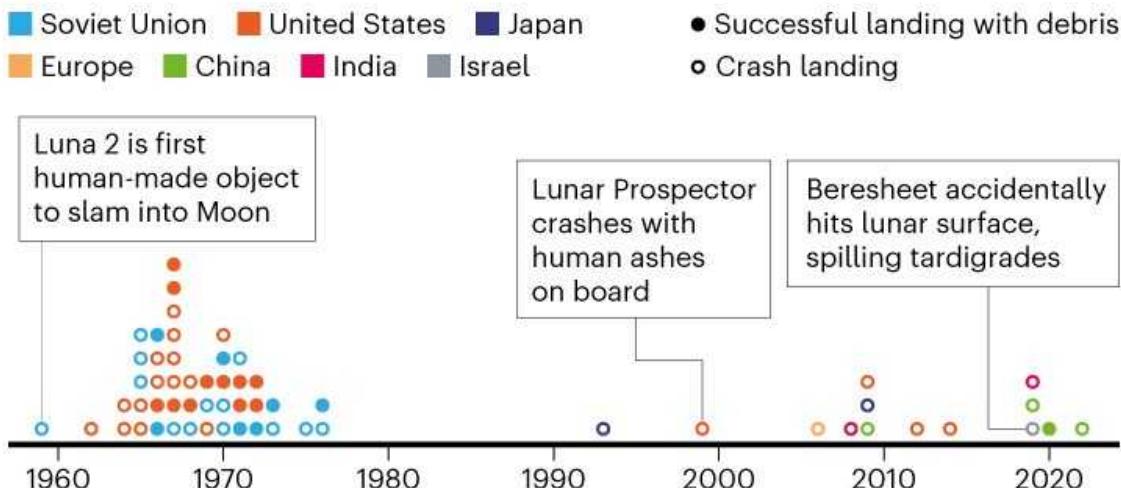
On 4 March, humanity set a record for littering, when an old rocket booster smashed into the far side of the Moon. It is the first time a piece of human-made space debris has hit a celestial body other than Earth without being aimed there.

The booster is probably part of a rocket that launched a small Chinese spacecraft, called Chang'e 5-T1, towards the Moon in 2014. Although Chang'e 5-T1 returned to Earth successfully, the booster is thought to have been zipping around chaotically in space ever since. Lunar gravity pulled it into a fatal collision with the far side of the Moon.

## MOON CRASHES

A spent rocket booster will hit the Moon on 4 March — the first time this has happened for space junk not guided there. But it will be far from the first human-made object that has ever crashed into the lunar surface.

Shown are planetary missions that have smashed into the Moon, either as a full spacecraft or as parts\*.

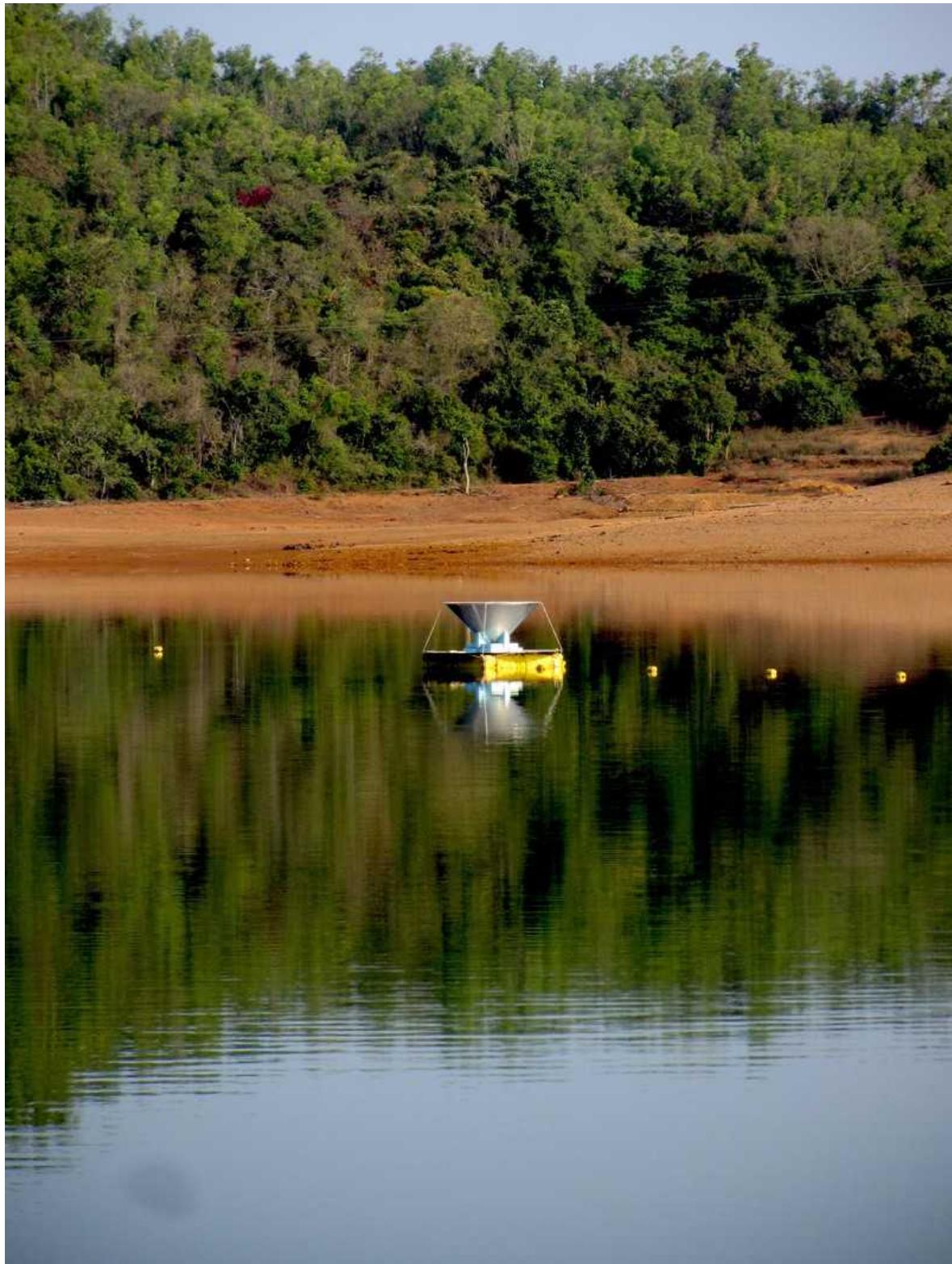


©nature

Source: Data from Jonathan McDowell

The incident posed no immediate danger to humans or other spacecraft, but with at least half a dozen craft scheduled to reach the Moon this year, concern is growing about the lunar surface becoming an [unintentional dumping ground](#).

Plenty of other spacecraft — and spacecraft bits — have hit the Moon (see ‘Moon crashes’). The first was the Soviet Union’s Luna 2 in 1959, which became the first human-made object to make contact with another celestial body when it crashed a little north of the lunar equator. The most recent was China’s Chang’e 5 lander (a different spacecraft from Chang’e 5-T1), which dropped an ascent vehicle onto the Moon in 2020 as it flew lunar samples back to Earth.



The SARAS 3 antenna provided results suggesting that the EDGES signal was caused by instrument error.Credit: Ravi Subrahmanyam

# Study challenges detection of first stars

The [first major attempt to replicate evidence of the ‘cosmic dawn’](#) — the appearance of the Universe’s first stars 180 million years after the Big Bang — has muddled the picture.

Four years after radioastronomers reported finding a signature of the cosmic dawn, radioastronomer Ravi Subrahmanyam and his collaborators describe how they floated an antenna on a reservoir along the Sharavati River, in the Indian state of Karnataka, in search of that signal. “When we looked for it, we did not find it,” says Subrahmanyam, who led the effort at the Raman Research Institute in Bengaluru, India. The results appeared on 28 February ([S. Singh \*et al.\* \*Nature Astron.\* <https://doi.org/hkpj>; 2022](#)).

The findings are “a very important landmark in the field”, says Anastasia Fialkov, a theoretical physicist at the University of Cambridge, UK. She and others had been unconvinced that the cosmic-dawn signals were real. The Raman team’s results are the first to put the claim to a serious test, she says — but she thinks that they don’t yet have the power to completely rule it out.

Light from the most ancient stars in the observable Universe has had to travel so far that it is too faint to view directly with ordinary telescopes. But radioastronomers have been looking for an indirect effect, using the spectrum of radio waves. In 2018, astronomers reported seeing a dip in the primordial radio spectrum, centred at a frequency of about 78 megahertz — which the team took to be evidence of the cosmic dawn ([J. D. Bowman \*et al.\* \*Nature\* \*\*555\*\*, 67–70; 2018](#)).

*Nature* **603**, 369 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00725-z>

---

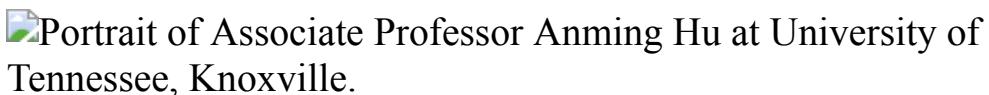
This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00725-z>

- NEWS
- 07 March 2022

# ‘I lost two years of my life’: US scientist falsely accused of hiding ties to China speaks out

Acquitted nanotechnology researcher Anming Hu returns to his lab after two years — and is still grappling with the aftershocks of his ordeal.

- [Natasha Gilbert](#)



Anming Hu is back at work but fears his time away from the lab will be ‘catastrophic’ for his career. Credit: Jessica Tezak for *Nature*

Anming Hu walked back into his laboratory at the University of Tennessee (UT), Knoxville, on 1 February for the first time in about two years to find it stripped of research equipment. Gone were the delicate lasers, lenses and voltage metres — worth tens of thousands of dollars — that he and his students had used to conduct their studies.

The nanotechnology researcher spent much of his first few weeks back at work searching for his prized tools. Some he found in colleagues’ offices and labs, other pieces — some of them broken — he discovered packed into storage cupboards, he says. Others are still missing.

“I want to get them back because those [are] my treasures,” he says.

Hu’s tepid return to the university follows a turbulent few years during which he was accused by the US government of hiding ties with China, put

under house arrest and, eventually, acquitted of all charges. This string of events occurred around the same time as the launch of the China Initiative — a US government effort to counter economic espionage that [frequently targeted academic researchers](#) for failing to disclose funds from China or partnerships with Chinese institutions.

Hu, a Canadian citizen of Chinese descent, had been working in the United States for more than four years when the FBI first interviewed him in 2018, the year the initiative was launched. Almost two years later, in February 2020, he was indicted for wire fraud and for making false statements about his affiliation with a Chinese university on research grant applications submitted to NASA.

“That is the day I lost everything. I worked hard for years, and it happened in a few minutes,” he says.

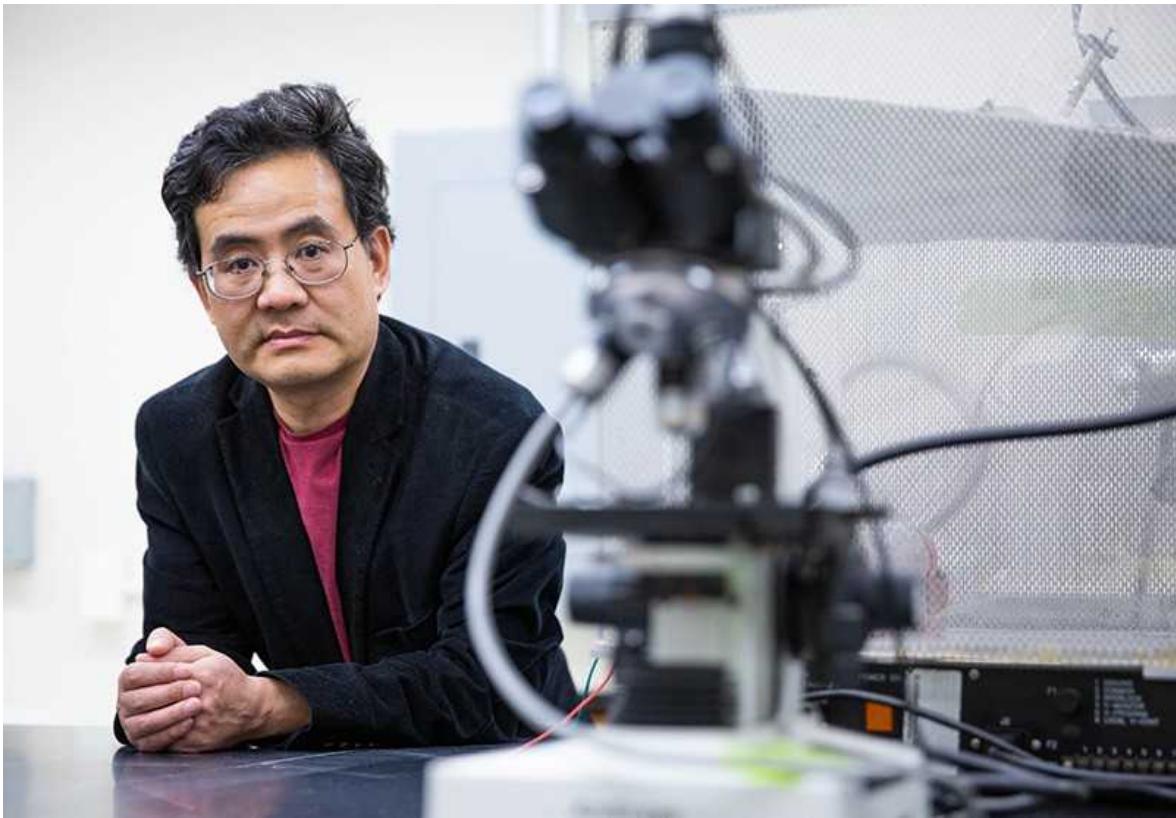
## Accountability needed

The US Department of Justice (DoJ) [announced on 23 February](#) that it was effectively ending the China Initiative, and was replacing it with a broader strategy covering China and several other countries, including Russia and North Korea. Academic researchers and civil-liberties groups had been calling for the initiative to be shut down, arguing that it was racially profiling scientists.

When announcing the changes to the initiative, US assistant attorney-general Matthew Olsen said that his review of the original programme did not find evidence that the DoJ’s actions had been racially prejudiced. But he did acknowledge that the programme, as well as several controversial cases that targeted people from China, gave the appearance of bias.

Hu welcomes the end of the initiative, but he would like to see the government held to account for its actions.

“I lost two years of my life,” he says. “Who is taking the consequences of that?”



Much of Anming Hu's equipment disappeared from his lab during his enforced absence. Credit: Jessica Tezak for *Nature*

After Hu's indictment, UT Knoxville suspended him and then stopped his pay. Eight months later, the university fired him after he lost his right to work in the United States. Hu's work visa expired in June 2020 while he was awaiting trial, and the US Citizenship and Immigration Services closed his application for permanent residence two months later. During his time in court, the jury was unable to reach a verdict, resulting in a mistrial. In September 2021, a judge acquitted Hu on all counts, finding no evidence that the nanotechnology researcher had committed the crimes he was accused of.

"It is very painful in my heart. The memories are very hard," Hu says.

## Under house arrest

Hu was under house arrest for more than a year while awaiting trial. During this time, he relied heavily on his church community to bring him groceries

— and even take out his rubbish — because he was not allowed to go outside, he says. Although UT Knoxville had cut ties with Hu, some of his colleagues quietly tagged along with the church group to do the shopping or wave to him from the street. With his wife and two of his children in Canada, he felt isolated and missed his family.

“It made me want to cry every night,” Hu says.

Once he had finished building his legal defence, Hu dived into unpaid work to help to distract him while he awaited trial, he says. He reviewed around 400 research papers and wrote 6 research papers of his own, including some on nanojoining — a technology used to link nanometre-sized building blocks to manufacture nanodevices and systems. He also finished writing a book on laser manufacturing and 3D printing. But he fears that the two-year gap in his laboratory research will be “catastrophic” for his career.

Others accused under the China Initiative also worry about the personal and professional toll. Gang Chen, a prominent mechanical engineer at the Massachusetts Institute of Technology (MIT) in Cambridge, [told \*Nature\* in an e-mail](#) that the US Congress should hold the FBI and DoJ to account for their treatment of researchers. Chen was arrested in January 2021 for allegedly failing to disclose on grant applications that he had ties to China and had received funding from Chinese institutions — charges the DoJ later dropped.

Chen [wrote in an editorial in \*Science\*](#) that MIT openly supported him after he was arrested, and says that other universities and funding agencies should also stand up for faculty members who are wrongfully prosecuted. This strikes a chord with Hu, who did not feel supported by his university or his funding agency.

## On trial

According to court documents from Hu’s trial, UT Knoxville knew of evidence from 2014 onwards that showed Hu was not hiding his affiliations with Chinese institutions. For instance, Hu submitted a letter to the university in 2016 disclosing his collaboration with a researcher at the

National Synchrotron Radiation Laboratory in Hefei, China, that he planned to include in his application materials for a NASA-funded grant proposal. Hu worked with this researcher on nanostructure characterization.

A university employee requested that Hu remove the letter from the application because it would violate an assurance document the university was submitting stating that the proposal had no restricted ties to China, the court documents say. A 2011 US law prohibits NASA from funding collaborations with China or Chinese businesses.

The university also gave Hu inconsistent advice about what links to China were allowed when applying for NASA funding; for example, on one occasion the same university employee wrote to Hu in an e-mail that, as UT Knoxville understood it, the NASA restrictions didn't apply to its faculty members, staff members or students, the court documents show.

Hu also submitted a letter outlining his collaboration with the Beijing University of Technology (BJUT) to a researcher at NASA's Jet Propulsion Laboratory (JPL) in Pasadena, California, who was a subcontractor for NASA. Hu collaborated with a scientist at BJUT on nanomaterial development. However, the JPL researcher told Hu that he would not include the letter in the application package, because the collaboration was not permitted under the NASA restrictions, court documents say.

When asked for comment by *Nature*, the NASA Office of Inspector General replied: "It is our policy not to comment on or generally acknowledge investigative activities conducted by our office."

That the university did not openly defend him during his case particularly stung, Hu says. "They cannot throw the professor under the bus. They should have taken their responsibility."

Tyra Haag, a spokesperson for UT Knoxville, told *Nature* in a statement that the university "followed its policies and processes" throughout the Hu case, and supported his efforts to resolve his immigration status once the case was over. Now UT Knoxville "is providing support for him to restart his career and his research" and is "grateful and pleased to welcome him back to the faculty", the statement says.

## Rebuilding his reputation

After Hu's acquittal, faculty members at UT Knoxville [campaigned for the university to rehire him](#). The university helped Hu to apply for a temporary work permit, reinstated him as a tenured professor and gave him US\$300,000 to help to restart his research, some of which he will spend on repairing and replacing his equipment. But Hu says that the university has not yet apologized to him.

He also says that NASA has told him he can reapply for funding, but that he doesn't have any immediate plans to do so. From now on, he will be careful who he collaborates with and where he seeks funding from, he says.

Hu ultimately decided to return to UT Knoxville because he felt it would help to show that he wasn't guilty and that he didn't deserve to be penalized for the mistakes of others, he says.

"I need to get back my reputation. I need to let other people see I didn't do anything wrong," he says.

For the time being, Hu wants to focus on his research, including new plans for projects on nanorobots. He also hopes to use his position at the university to get the word out about how situations like his can arise — and to prevent it from happening to others. In addition to speaking to the media about his experiences, Hu plans to speak at events and to advocacy groups. He doesn't want others to suffer in the same way he has, he says. He warns other researchers that they "need to be very careful".

The China Initiative might have ended, but Hu says he will be looking out for changes in the way researchers of Chinese descent are treated, and whether law-enforcement agencies, including the FBI, are held accountable for their actions.

"I would like to see that the DoJ respects academic freedom and encourages normal international collaboration and exchange," Hu says. "We should be alert that this does not mean that the prosecution goes underground."

*doi: <https://doi.org/10.1038/d41586-022-00528-2>*

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00528-2>

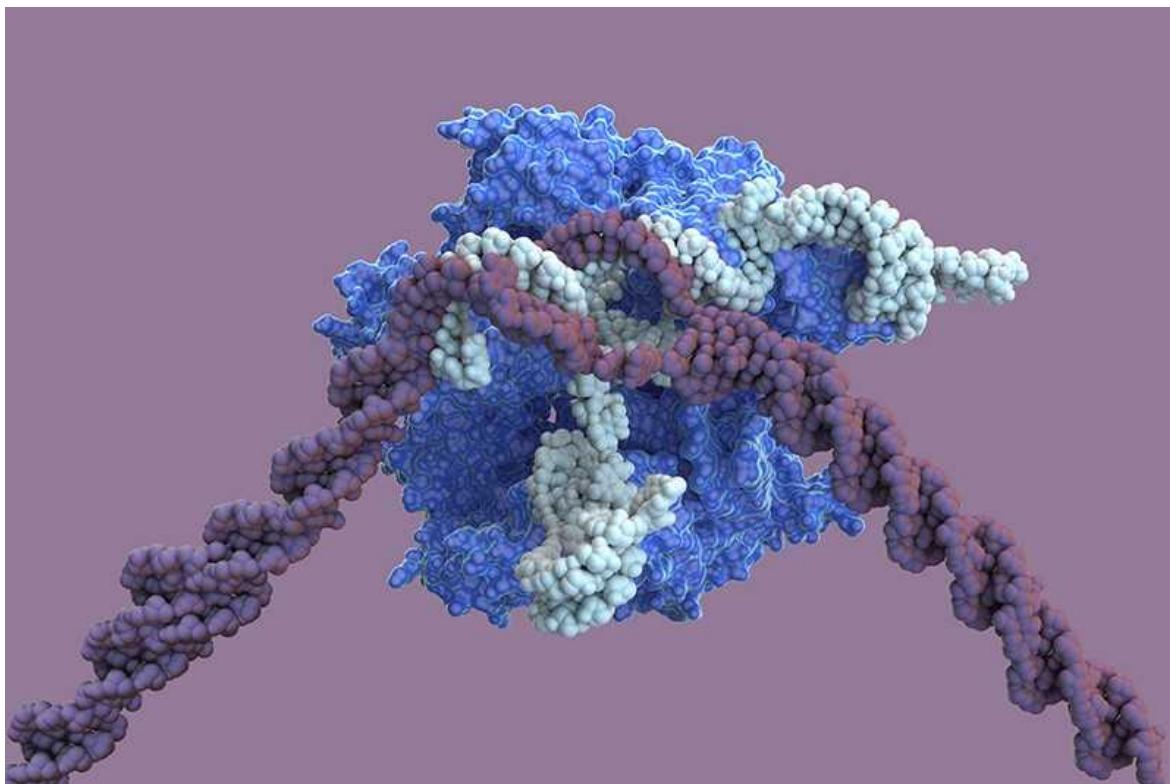
| [Section menu](#) | [Main menu](#) |

- NEWS
- 09 March 2022

# Major CRISPR patent decision won't end tangled dispute

Fights over who invented the gene-editing technology are becoming more complex, and could carry on for years.

- [Heidi Ledford](#)



The CRISPR–Cas9 complex (blue) can cut DNA (purple). Disputes about who invented the gene-editing system, and who should benefit from key patents, are ongoing. Credit: Ella Maru Studio/Science Photo Library

A long-running dispute between two groups that claim to have invented the revolutionary CRISPR–Cas9 gene-editing tool is likely to remain unresolved for years to come, lawyers say — despite the US patent office’s latest decision to award key patent rights to one of the teams.

The ultimate outcome of the patent row — which began in 2016 — could mean millions of dollars in royalties for the victor, if and when CRISPR-based therapies make it to market.

But the dispute shows little sign of ending, and the intellectual property around CRISPR is growing more complex. Patent offices in other countries have reached different decisions about who invented what. Additional parties have entered the battle, contesting ownership of pivotal, early CRISPR–Cas9 patents. And patents on CRISPR-related technology have proliferated as new techniques are developed.

“This could be going on for years,” laughs Catherine Coombes, a patent attorney in the York, UK, office of the intellectual property law firm Murgitroyd. “Many years.”

## Discovery dispute

The patent case in question has pitted two high-profile research teams against each other. One, led by molecular biologist Feng Zhang at the Broad Institute of MIT and Harvard in Cambridge, Massachusetts, has won several key decisions by the US Patent and Trademark Office (USPTO). The other team includes biochemist Jennifer Doudna at the University of California, Berkeley, and microbiologist Emmanuelle Charpentier, who began her work on CRISPR at the University of Vienna. In 2020, Doudna and Charpentier [shared the Nobel Prize in Chemistry](#) for discovering CRISPR–Cas9 gene editing — but patents and Nobel prizes are not necessarily judged by the same criteria.

The Berkeley team (abbreviated as CVC in the latest USPTO decision) filed for its original patent in 2012, a few months earlier than the Broad Institute. But at that time, the USPTO awarded patents on the basis of who was first to invent a technology, rather than who was first to file the patent, and the two

teams have been arguing for years about who initially developed CRISPR–Cas9 gene editing. (The USPTO changed its procedures in 2013 and the United States, like much of the rest of the world, now awards patents according to who files the application first.)

There are now more than 11,000 families of patents on CRISPR-related technologies, according to the business-intelligence firm Centredoc in Neuchâtel, Switzerland. But the initial patents that the CVC and Broad teams filed are considered to be broad and foundational: many companies that wish to sell products such as therapies or crops made with CRISPR–Cas9 gene editing might be required to obtain licences whichever team wins these fights.

“It makes sense why there’s been so much focus on these patents: the technology is so significant,” says Daniel Lim, a patent litigator at the law firm Kirkland & Ellis International in London. “But there’s so much more than this one, isolated and really unique feature of the old US legal system — there’s a whole world out there.”

Over the years, the USPTO has repeatedly ruled in favour of the Broad, but the CVC team prompted another evaluation of patents covering what will probably be the most lucrative application of gene editing: modifying genomes in eukaryotes, a group of organisms that includes humans and crops. Both CVC and the Broad claimed to be the first to have adapted CRISPR–Cas9 gene-editing for use in eukaryotes, and the case involved poring over lab notebooks and e-mails to work out when each team achieved success. On 28 February, the USPTO determined that the Broad got there first, potentially by a matter of weeks.

## Experimental treatments

Even if the decision stands, it could be some time before the Broad stands to earn much in royalties. No human therapy based on CRISPR has yet been approved, although several are working their way through the pipeline. On 28 February, Intellia Therapeutics in Cambridge, Massachusetts, announced that [its experimental CRISPR–Cas9 treatment](#) for a rare condition called transthyretin amyloidosis reduced production of an errant protein by up to

93%, with effects lingering for at least a year. And by the end of 2022, a team of two companies — CRISPR Therapeutics, of Cambridge and Zug, Switzerland, and Vertex Pharmaceuticals of Boston, Massachusetts — plans to file for approval from the US Food and Drug Administration for its experimental sickle-cell disease treatment based on CRISPR–Cas9.

Both of these teams licensed patents from CVC, rather than the Broad. That means they might eventually need to come to an agreement with the latter team, which could be entitled to a portion of any profits from the treatments.

That need not delay the progress of the therapies, says analyst Geulah Livshits at Chardan, a health-care investment bank in New York City. Two companies have claimed patent rights to the lipid nanoparticle that Moderna Therapeutics in Cambridge, Massachusetts, used to encase its mRNA-based COVID-19 vaccine. But that ongoing dispute did not keep the vaccine from reaching the market, Livshits notes.

The CRISPR–Cas9 patent landscape remains unclear, she adds. CVC might appeal against the USPTO’s decision in a federal patent court, but the case is unlikely to go all the way to the US Supreme Court, says Kevin Noonan, chair of the biotechnology and pharmaceuticals group at the law firm McDonnell Boehnen Hulbert & Berghoff in Chicago, Illinois. “There’s no big policy question here for the court to address,” he says.

It’s also possible that neither CVC nor the Broad will reap the rewards of the foundational CRISPR–Cas9 patents. Both are facing challenges on these patents from two other companies: ToolGen in Seoul and Sigma-Aldrich, now owned by Merck in Darmstadt, Germany.

And in the European Union, the key CRISPR–Cas9 patents in the Broad’s portfolio have been tossed out altogether owing to missing paperwork. In the course of finalizing its patents, the Broad team decided to drop one of its inventors from the filings — but neglected to get written approval from him, a requirement in the EU system. As a result, CVC has the upper hand in Europe. “Europe’s gone in a completely different direction to the US,” says Coombes, “which makes things interesting from a licensing perspective.”

## Gene-editing alternatives

Companies now also have the option of avoiding these patents altogether by using different CRISPR systems. Such systems occur naturally in many bacteria and archaea, and can have various properties. Over the past two years, says Fabien Palazzoli, a senior patent analyst at Centredoc, there has been a dramatic increase in the number of patent applications claiming new diagnostic tests for viruses and bacteria, possibly spurred by the COVID-19 pandemic.

Relatively few of these use CRISPR–Cas9, he says; instead, they use alternative enzymes such as Cas13 or Cas14, which is remarkably small and easy to transport into human cells. Labs have also engineered new CRISPR-associated enzymes, such as base editors, that are better able to make specific edits. Patent filings on base editors are doubling every year, Palazzoli says, and now exceed 730.

Against the backdrop of so much activity, it will be years until it becomes clear how much the original CVC and Broad patents are worth, says Coombes. “I don’t think CRISPR–Cas9 is the be all and end all,” she says. “There’s still a lot up for grabs.”

*Nature* **603**, 373–374 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00629-y>

---

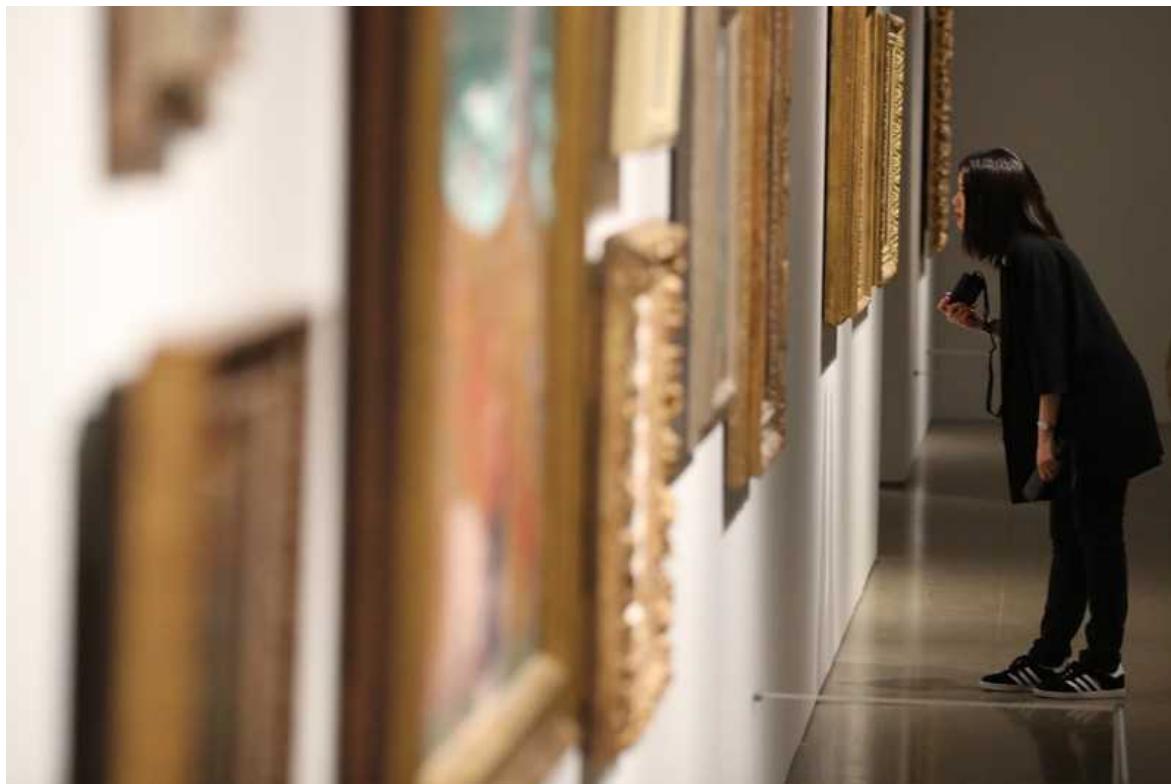
This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00629-y>

- NEWS
- 09 March 2022
- Clarification [09 March 2022](#)

# Police rely on radiocarbon dating to identify forged paintings

Advances in the technique have bolstered its reputation as a tool for investigating faked artwork.

- [Carolyn Wilke](#)



Researchers have used radiocarbon dating to identify two works of art as forgeries. Credit: Steve Russell/Toronto Star/Getty

Radiocarbon dating has unmasked two forged paintings in France — probably the first time the technique has been used in a police investigation. The paintings were supposedly impressionist and pointillist works from around the early twentieth century. But a team led by heritage scientist Lucile Beck at the University of Paris-Saclay used radiocarbon levels in the fibres of their canvases to date them to sometime within the past 70 years. The researchers concluded that the paintings are modern forgeries in a 4 February report published in *Forensic Science International*<sup>1</sup>.

The use of radiocarbon dating is gaining steam in the forensic analysis of artwork, thanks to advances that require smaller samples than ever before. Removing tinier samples from artwork is more palatable to auction houses, museums and owners of paintings. If there is a chance a painting is genuine — and therefore valuable — they don't want the collection of larger samples to damage it, says art historian Anna Tummers at Leiden University in the Netherlands, who was not part of the new research.

The technique's success might persuade more of the art world to seek radiocarbon dates, which can more definitively pinpoint when a painting was made, Tummers says. Researchers typically use imaging and chemical analysis to sniff out art forgeries. These methods can peer beneath brushstrokes to see how a painting's materials have aged, but can't conclusively nail down a painting's date.

The consequences of faked artwork extend beyond forgers lining their pockets in the global art market, which moves tens of billions of dollars every year. Forgeries pollute people's understanding of artworks' meaning, Tummers says. "If we don't weed them out carefully, it might really distort our understanding of our own heritage and our own history."

## Suspected forgeries

The two paintings were part of a trove of artwork that French investigators uncovered in a restorer's workshop in 2019. Of some 600 paintings, dozens appeared to be mid-level masterpieces from the late nineteenth and early twentieth century. But experts questioned the works' authenticity because the paint seemed relatively fresh.



Beck's team removed a fibre from the canvas of an impressionist painting, to test whether the piece was a forgery.Credit: L. Beck

To investigate the potential forgeries, the French government's Central Office for the Fight against Illicit Trafficking in Cultural Property (OCBC) called on Beck. When she walked into the OCBC office, she was amused to see what would normally be a drab office with paintings displayed across its floor and tables.

Beck's team selected a few works, including an impressionist garden scene and a pointillist port landscape, to test. In the office, the researchers used scalpels to remove samples including a small bit of fibre from the canvases.

All living things take in carbon, including radioactive carbon-14, from the atmosphere and from food. When a plant — such as linen or hemp, commonly used to make canvas — dies, the carbon-14 that it incorporated continues to decay. Radiocarbon dating measures what's left to estimate the time that's passed, says Mariaelena Fedi, a physicist at the National Institute for Nuclear Physics in Florence, Italy. The technique gives an absolute

earliest date of an artwork, because there could be years between harvesting linen for a canvas and making the painting.

Atomic-bomb testing, which began in the 1940s and took off in the 1950s, bumped up the amount of carbon-14 in the atmosphere beyond naturally produced levels. Carbon-14 peaked around 1964 and tailed off after a partial ban on nuclear tests. Researchers can easily identify materials containing modern bomb-derived radiocarbon because their carbon-14 concentrations are higher than pre-1950s levels.

Beck's team tested its samples to see whether they bore the signature of that bomb-derived radioactive carbon-14. In the laboratory, the researchers cleaned and dried the material, whittling down several milligrams into around one milligram of carbon that was pressed into a graphite puck to be measured with accelerator mass spectrometry.

The canvas fibres from both the impressionist and pointillist paintings clearly contained carbon from either the mid-1950s or after the year 2000, the researchers reported. (This is because the carbon-14 concentrations they measured could correspond to either side of the atomic bomb peak.) Another fibre, plucked from the varnish of the pointillist painting and perhaps coming from a brush, also dated to after 1950. Beck acknowledges that, ideally, the team would do further chemical analysis to support its findings, but the researchers were limited by the tight time frame of the investigation.

## Advances in the field

Although this seems to be the first report of using radiocarbon dating to identify faked artwork in a police investigation, researchers have been laying the groundwork for the past decade.

Fedi and her colleagues made the first report of radiocarbon dating to detect a forged painting in the Peggy Guggenheim Collection in Venice, Italy, in 2014. The team collected snippets of the canvas and dated them to after the presumed artist's death — concluding that the painting had been forged<sup>2</sup>.

In 2019, Laura Hendriks and colleagues used a known forgery of a village scene to test a radiocarbon dating method that used a much smaller sample size than previous techniques. Hendriks, a chemist at the University of Applied Sciences and Arts of Western Switzerland in Fribourg, converted a sample of the scene's oil paint to carbon-dioxide gas before introducing it to the mass spectrometer. The team was able to date the forgery using only micrograms collected from the painting<sup>3</sup>. The tiny sample was “just a few crumbs of dust, basically”, Hendriks says.

Such advances are good news for the field. There's a huge need for objective tools to suss out fakes, Fedi says, and radiocarbon dating is excellent when combined with other methods and the expertise of art scholars who can help to interpret the history of these complex objects.

*Nature* **603**, 374-375 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00582-w>

## Updates & Corrections

- **Clarification 09 March 2022:** An earlier version of this story mentioned potential contamination in the analysis. This would be for the analysis of paint, not canvas. The text has been updated to eliminate confusion.

## References

1. Beck, L. *et al. Forensic Sci. Int.* **333**, 111214 (2022).
2. Caforio, L. *et al. Eur. Phys. J. Plus* **129**, 6 (2014).
3. Hendriks, L. *et al. Proc. Natl Acad. Sci. USA* **116**, 27 (2019).

---

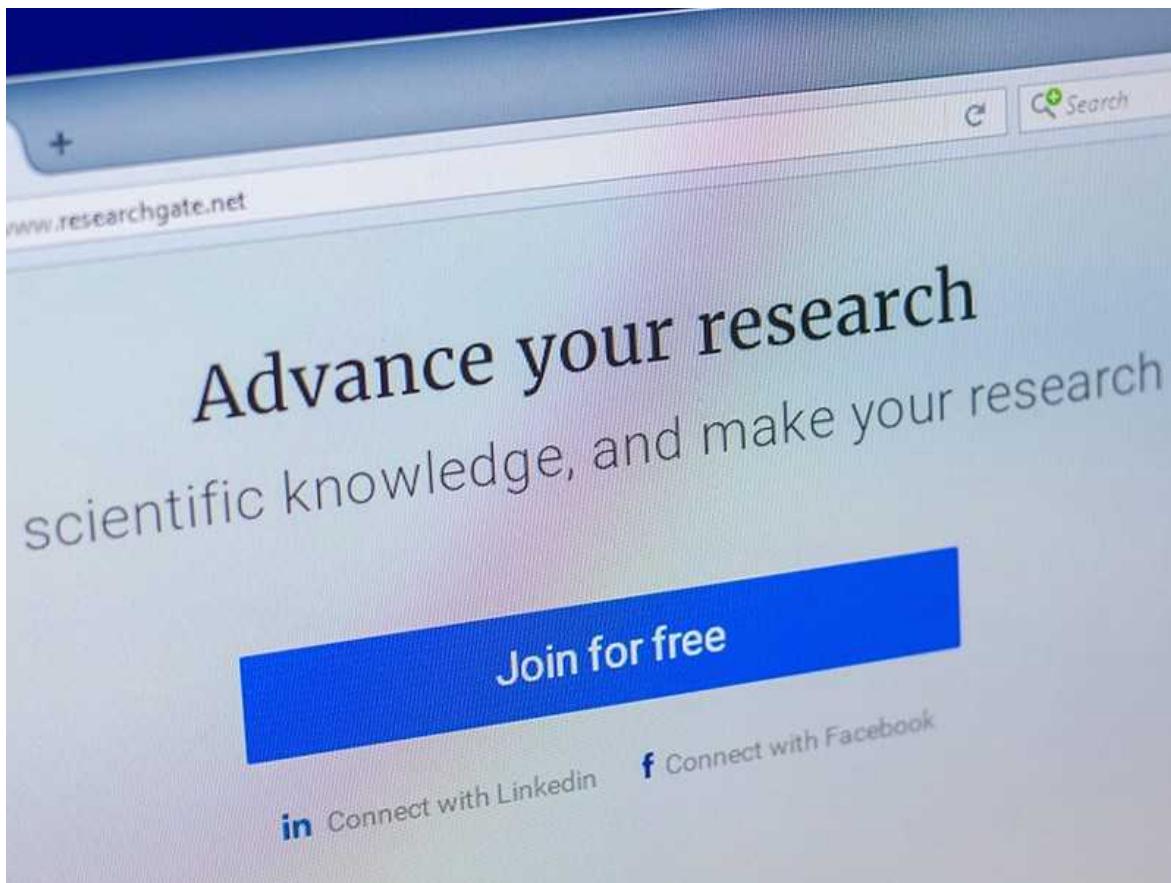
This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00582-w>

- NEWS
- 04 March 2022

# ResearchGate dealt a blow in copyright lawsuit

Court ruling says that the academic-networking website is responsible for papers uploaded by its users — but ResearchGate plans to appeal the decision.

- [Diana Kwon](#)



ResearchGate was founded in 2008 and has 20 million users from more than 190 countries. Credit: Alamy

A landmark court case in which two major academic publishers sued the popular website ResearchGate for hosting 50 of their copyrighted papers has come to a close — although both sides say that they will appeal. The court in Munich, Germany, has not only prohibited ResearchGate from hosting the papers, but also ruled that it is responsible for copyright-infringing content uploaded on its platform. The decision has the potential to set a precedent for further restrictions on the site, which has 20 million users worldwide.

Neither side emerged a clear winner in this case, says Nancy Sims, a librarian at the University of Minnesota in Minneapolis who specializes in copyright issues. “Each party got some pieces that were very favourable to them and some pieces that were less favourable to their claims.”

Elsevier, which is based in Amsterdam, and the American Chemical Society (ACS), in Washington DC, [filed the lawsuit](#) against ResearchGate in a regional court in Munich in 2017, alleging that the site had made copyrighted material freely available. ResearchGate said that it could not be held responsible for the content, which was uploaded by authors.

“We are pleased with the verdict,” says a spokesperson for the Coalition for Responsible Sharing, a group of publishers — including Elsevier and the ACS — that was formed in 2017. “The clear aim of the legal action was to clarify the responsibilities of ResearchGate for the content that it illicitly distributes on its site, which it does for its own commercial gain.”

“We believe that the outputs of scientific research, the majority of which is funded by public money, should be shared as openly as possible and we’ll continue to support researchers in sharing their work easily and legally,” says a spokesperson for ResearchGate. They add that it is up to publishers to work with platforms to remove uploaded content, and that in the years since the case was first brought they had introduced software that could help prevent the sharing of copyrighted material. “We’ve built a fully compliant solution, and are pleased that Elsevier and ACS are using it.” In a [public statement](#) posted after the Munich court’s 31 January ruling, ResearchGate said that it will appeal parts of the decision.

## **Wins and losses**

ResearchGate has had a complicated relationship with academic publishers since its foundation in 2008. Many researchers use the platform to upload and share documents, including their published research. As well as initiating legal action, publishers have been sending take-down notices to the site for years demanding it remove paywalled articles.

The Munich lawsuit focused on 50 articles on the site, all of which have since been removed. The court ruled that the uploaded articles breached the publishers' copyrights, and that ResearchGate was responsible for the content on its platform. "Only a sample of 50 articles was included, but the verdict speaks to the responsibility of the site in general," says the spokesperson for the Coalition for Responsible Sharing.

However, the direct implications for any article other than the 50 named in the lawsuit are unclear, says Guido Westkamp, who studies intellectual property and comparative law at Queen Mary University of London. If the reasoning applied in the court's judgment remains unaltered after an appeal, the precedent will also apply to any future lawsuits against ResearchGate, he says, in Germany at least. But the ruling is "far from a blocking order", he adds. "In principle, any other content is subject to a new lawsuit."

The court rejected the publishers' request to be paid damages, noting that they did not have sufficient evidence to show that standard copyright licensing agreements, which are typically signed by only a paper's corresponding author, prove that all authors agree to transfer ownership to the publisher. The publishers have said that they [plan to appeal this part of the ruling](#).

## **A shifting tide**

In the years since the lawsuit was first filed, the academic publishing world has undergone a sea change.

A growing open-access movement, driven in part by mandates from funders, has led many publishers to make more of their articles freely available. And

while Elsevier and the ACS have been locked in a legal battle with ResearchGate, several other big publishers, such as Wiley and Springer Nature, have instead partnered with the platform to enable it to share published papers. (*Nature*'s news team is editorially independent of its publisher, Springer Nature.)

In the European Union, there has also been a shift in copyright law. One article of sweeping copyright legislation that came into effect last year makes content-sharing platforms responsible for the content uploaded by their users. That could make ResearchGate liable for damages in the future, Westkamp says. “However, publishers will still need to demonstrate that the rights had been assigned, including, as the case may be, by all co-authors.”

ResearchGate has [stated](#) that it believes it is not subject to these rules because of the “nature of [its] business”. Nevertheless, the platform has developed a software tool known as Jarvis, which matches rights information from publishers to papers during the uploading process to prevent researchers from posting content they are not permitted to share.

The EU legislation “essentially says that ResearchGate has to do what ResearchGate has just been ruled it has to do”, says Lisa Hinchliffe, who studies and coordinates the information-literacy services at the University of Illinois at Urbana-Champaign. “Not to say there’s no value in a ruling, but the world has indeed changed in the five years since this lawsuit was filed.”

## **The saga continues**

The legal battle between ResearchGate and the publishers is a long way from over. ResearchGate’s appeal to the Munich court’s ruling — which applies only in Germany — is unresolved. And the site is embroiled in a [similar lawsuit](#) with the same two publishers in the United States.

ResearchGate potentially has a stronger case in the US court, because decisions could hinge on the issue of fair use rather than copyright law alone, Westkamp says. “The public-interest argument would be way stronger in the States than it is in Germany,” he says, and ResearchGate could argue

that copyrighted papers should be made available on its platform for the sake of freedom of access to knowledge.

In the long term, he says, it is up to publishers to decide how to deal with ResearchGate. “From a purely commercial point of view, you don’t want competition,” he says. But attempts to “eradicate” the platform could anger the authors and members of the scientific community on whom publishers rely. “It is not in the interest of authors not to have platforms like ResearchGate.”

*Nature* **603**, 375-376 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00513-9>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00513-9>

| [Section menu](#) | [Main menu](#) |

- NEWS EXPLAINER
- 03 March 2022

# This US Supreme Court decision could derail Biden's climate plan

Controversial lawsuit has put the US government's ability to slash carbon emissions on the line.

- [Jeff Tollefson](#)



Climate-change activists rallied outside the Supreme Court on 28 February as the justices heard arguments in the case *West Virginia vs Environmental Protection Agency*. Credit: Bryan Olin Dozier/NurPhoto/Shutterstock

The US Supreme Court heard oral arguments this week in a controversial lawsuit that could deal yet another blow to President Joe Biden's climate agenda. Depending on how the court rules, the lawsuit has the power not only to prevent the US Environmental Protection Agency (EPA) from regulating future greenhouse-gas emissions, but also to potentially reshape other US agencies' regulatory powers.

The unusual case hinges on a years-long legal tussle over two EPA policies crafted under former presidents Barack Obama and Donald Trump that sought to regulate power-plant emissions in opposing ways. Neither policy ever took effect, and it's that fact that sets this case apart: normally, the Supreme Court would not agree to hear regulatory cases in which there is no regulation to debate. The group of Republican-led states and coal companies suing the EPA, however, are raising the spectre of future regulations that could hamper a crucial sector of the US economy — the electricity industry.

“This is not about stopping climate-change efforts,” West Virginia attorney-general Patrick Morrisey, a Republican, said in a statement released after the arguments this week. “The future of our nation is at stake. This case will determine who decides the major issues of the day.”

The Biden EPA and its allies — including environmentalists and public-health advocates — see it as a brazen attempt to restrict the government’s power over industry and pollution control. “The arguments being raised here really are breathtaking in terms of trying to limit the government’s ability to protect public health and welfare,” says Sean Donahue, a lawyer with Donahue, Goldberg & Littleton in Washington DC, who represents the group weighing in on behalf of the EPA.

Here *Nature* explains the case, and examines what’s at stake.

## What is the history of the lawsuit?

In 2015, during the Obama administration, the EPA finalized a flagship climate rule, dubbed the Clean Power Plan, which sought to curtail emissions from the electricity sector to at least 30% below 2005 levels by 2030. The plan would have set reductions targets for US states; to meet

them, coal- and gas-burning power plants could have upgraded their technology to boost efficiency and decrease emissions, but the bulk of the reductions would have needed to come from electric utilities shifting towards more renewable energy sources, such as wind and solar. The Obama EPA said that this ‘generation shifting’ approach was consistent with the Clean Air Act, a law that requires the agency to consider the best-available technologies when crafting regulations to curb air pollution.



Under fire in the Supreme Court case is the EPA’s authority to regulate emissions from coal-fired power plants such as this one, in Adamsville, Alabama. Credit: Andrew Carballeiro Reynolds/AFP/Getty

The more industry-friendly Trump administration repealed the Clean Power Plan in 2018 and [replaced it with a weaker version](#) dubbed the Affordable Clean Energy plan, which more narrowly interpreted the Clean Air Act. It also limited pollution controls to technologies that could be installed at individual power plants. Critics said it would do little, if anything, to encourage a broader shift towards clean energy.

The situation came to a head on Trump's final day in office in early 2021, when a federal appeals court in Washington DC dismissed the Trump plan and rejected its repeal of the original Clean Power Plan. The new Supreme Court case, *West Virginia vs Environmental Protection Agency*, hinges on the fact that the appeals court expressly rejected the Trump administration's arguments that the Clean Air Act does not authorize the EPA to require generation shifting across the electricity industry.

Under Biden, the EPA has declined to revive the Obama administration's Clean Power Plan and is instead crafting its own plan for power plants. This means there are no actual regulations to challenge in this Supreme Court case, says GianCarlo Canaparo, a lawyer with the Heritage Foundation, a conservative think tank in Washington DC. But given the appeals court's ruling, Canaparo adds, the plaintiffs rightfully fear that the Biden EPA will craft its new plan by interpreting the Clean Air Act as the Obama EPA did.

To bring their case to the Supreme Court in the absence of a standing regulation, the plaintiffs have invoked something called the major questions doctrine, which argues that courts must prevent agencies from going beyond what Congress intends when it passes legislation of vast economic significance — for instance, using the Clean Air Act to reshape the electricity industry. If the high court follows this logic, it could adopt the Trump administration's narrow view for regulating emissions, or go even further and limit the EPA's power to craft regulations without express consent from Congress. The latter, critics fear, could have implications for other agencies.

## **So coal companies and Republican-led states are challenging EPA's authority. What about the electricity industry?**

Many of the largest utility companies that provide electricity to consumers have lined up on the side of the EPA in this case. One reason is that when people and groups sue utilities to seek compensation for climate change-induced damages, the companies have been able to defend themselves by pointing out that greenhouse gases are regulated by the federal government.

If the EPA loses that authority, that defence evaporates, potentially opening the door to an avalanche of lawsuits.

Utilities also fear a world in which agencies such as the EPA have been stripped of their powers, leaving the US Congress — notoriously slow to act and frequently deadlocked by partisanship — responsible for the details of agency-crafted regulations. “You’re talking about utter gridlock,” says Thomas Lorenzen, general counsel for the Edison Electric Institute based in Washington DC, utility companies’ largest trade association. “We want the court to understand fully these potential consequences” before it goes down the road advocated by West Virginia and the coal companies, Lorenzen says.

## **Which way does it look like the Supreme Court will go?**

Nobody knows. In 2007, the Supreme Court ruled 5–4 that the EPA had the authority to regulate greenhouse gases from vehicles, and by extension other sources. But in recent years, Trump appointed three justices to the court, [making it more conservative](#). Last month, Biden nominated Ketanji Brown Jackson to the court, but if she is confirmed by the Senate, she would replace liberal justice Stephen Breyer and so would not significantly alter the balance of the court’s power. She would also arrive too late for this particular case.

*West Virginia vs. EPA* will be a major test of how aggressively this new court is going to be reshaping legal doctrines, says Cara Horowitz, co-executive director of the Emmett Institute on Climate Change and the Environment at the University of California, Los Angeles.

Going by the justices’ lines of questioning during nearly two hours of oral arguments on 28 February, Horowitz thinks it unlikely that the court will dismiss the case outright. Instead, she expects it will either declare that the EPA has no authority to regulate power-plant emissions, or sharply limit the agency’s authority, in line with the Trump administration’s Affordable Clean Energy plan.

The Supreme Court arguments came on the same day that the United Nations' Intergovernmental Panel on Climate Change [released its latest report](#), which documents the accelerating impacts of climate change on people and natural ecosystems. “It makes clear that we don’t have time to waste squabbling over legal authorities,” Horowitz says. “But it’s a good bet that the court’s decision in this case will make that work harder, not easier.”

A decision on the case is expected as early as June.

*Nature* **603**, 376-377 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00618-1>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00618-1>

| [Section menu](#) | [Main menu](#) |

- NEWS FEATURE
- 16 March 2022

# A more-inclusive genome project aims to capture all of human diversity

Researchers are building a human ‘pangenome’ that would represent the entirety of human genetic variation. But not everyone is ready to sign on.

- [Roxanne Khamsi](#) 0

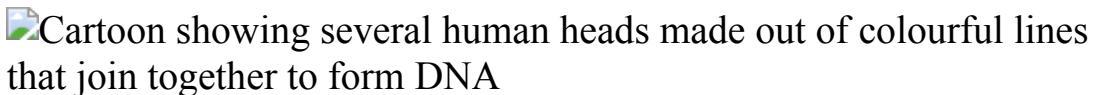


Illustration by Ana Kova

Several years ago, after an exhaustive search for uncharted variation in the human genome, Evan Eichler stumbled on something extraordinary. Eichler, a geneticist at the University of Washington in Seattle, and his colleagues struck on a massive stretch of DNA, about 400,000 letters long, that contained extra copies of genes — probably passed on from an ancient hominin group known as the Denisovans<sup>1</sup>. It appeared in about 80% of people living in Papua New Guinea, but practically nowhere else.

“We were shocked by the size,” Eichler says. “We always knew there would be archaic segments in our genome.” But the segment’s length and its absence in much of the world, he says, “transformed our thinking”.

This and other unexpected discoveries have made Eichler and other geneticists increasingly dissatisfied with the breadth and depth of the available maps of the human genome. The first draft genome from the US\$2.7-billion Human Genome Project, released in 2001, was [meant to](#)

become a reference point for future genetic research. But 93% of its sequence came from just 11 individuals, many of whom were recruited through a newspaper advertisement in Buffalo, New York; a whopping 70% of the DNA comes from just one man.

By 2003, that reference genome, known as GRCh38, would be deemed technically complete, but it still had hundreds of gaps and sections containing copious errors. These shortcomings came with consequences. Eichler worked with clinical geneticists at his university's medical centre and found that the reference genome lacks a region that has variants associated with Baratela-Scott syndrome, which can cause cognitive delays and skeletal malformations in children. Because that portion was missing, there was no quick way for the physicians to check for DNA errors there.

Genome maps have improved, but still don't adequately capture humanity's vast diversity. For example, in 2018, one group of researchers sequenced 910 individuals of African descent and discovered a sequence consisting of 300 million DNA letters, or bases, that was unfamiliar<sup>2</sup>. That's roughly 10% of the entire genome.

To create a reference that is more complete and more representative, Eichler has joined forces with a number of high-profile scientists, mostly in the United States. Their goal is to capture almost all human genetic variability — the dizzying number of genetic remixes in the human species, including additions, deletions and other types of mutation.

Rather than depicting the genome as a linear readout from a single individual, it would contain multiple paths branching in and out like the tangle of train lines on the map for the London Underground. These would represent the varieties of sequence that can be found in different populations, such as the long stretch of DNA found in many people from Papua New Guinea.

In 2019, Eichler and his colleagues started the Human Pangenome Project, a \$30-million effort funded by the US National Human Genome Research Institute (NHGRI) in Bethesda, Maryland. The initial goal is to do detailed, reference-quality genome sequencing of about 350 people from different backgrounds and to share those data as freely as possible.

The effort will pose a significant technical challenge, but the scientists behind it, including Karen Miga at the University of California, Santa Cruz, and Ting Wang at the Washington University School of Medicine in St. Louis, argue it is worth it. They see it as crucial to making genomic medicine more equitable<sup>3</sup>. “To account for diversity is to better serve humanity,” Wang says. “It is about both equity and equality. It is about building a more inclusive genomic resource for humankind.”

The researchers in the pangenome effort are aware of the history of past missions to capture human genetic diversity, some of which were seen as ‘vampire’ projects that took data from marginalized populations and failed to respect their needs and wishes. In response to this, the pangenome effort engages bioethicists throughout the project, not just at periodic junctures, as was done by initiatives in the past. “They are not a separate entity working in silo, they are involved in every step of the project, including all the technical decisions,” Wang says.

Nevertheless, some geneticists focused on the needs of Indigenous communities are wary of the initiative. They aren’t calling for an end to the Human Pangenome Project per se, but they say that marginalized groups deserve control of their genetic data, and of the sequencers, too. “As we position ourselves to be in control of these technologies, we’re empowering our communities,” explains Keolu Fox, a geneticist at the University of California, San Diego, who is Native Hawaiian. “Nothing is as real deal as we are. We’re from our communities.”

## Panning out

The concept of a pangenome traces back to the study of a bacterium known as *Streptococcus agalactiae*, or group B streptococcus, which can cause deadly infections in newborns. Scientists analysing six strains of the bacterium published a paper in 2005 trying to capture all of the microbe’s genetic nuances<sup>4</sup>. What they produced was a core genome shared by all six strains and a “dispensable” genome of partially shared and strain-specific genes.

It was a tricky task, because bacteria swap and share bits of DNA, even with other species, mostly through a process known as horizontal gene transfer. “There’s a lot of things that can happen in bacteria,” says Candice Hirsch, a plant geneticist at the University of Minnesota in Saint Paul. As a result, biologists are continually updating the bacterial reference genomes. Humans, by contrast, do not add new variation as easily. That makes characterizing a human pangenome more feasible, Hirsch says.



Geneticist Krystal Tsosie.Credit: Tomás Karmelo Amaya/New York Times/Redux/eyevine

But what it lacks in dynamics, the human genome makes up for in length and repetition. Chromosome 1, for example, the largest of the 24 different human chromosomes, stretches over about 250 million base pairs. That's more than 100 times the length of *S. agalactiae*. And it is riddled with long stretches of simple, repeated sequences and duplications of other, more complex segments. Until the past decade, scientists' main option for sequencing DNA involved breaking it into fragments and reading it in small chunks. This allows them to detect single-letter changes in DNA relatively easily. But the short reads make it hard to recognize when a long stretch of

DNA contains more than one copy of a gene. Eichler, who has specialized in identifying structural variants such as gene duplications and deletions, has opted for a newer approach, called ‘long-read sequencing’, which analyses bigger stretches of DNA at a time. This is what enabled him to find the previously unnoticed variant in people from Papua New Guinea.

In 2018, Eichler and other scientists gathered at the NHGRI to discuss a human pangenome effort. There, Eichler reconnected with a fellow scientist who shared his passion for long-read technology, Erich Jarvis, a neuroscientist and molecular biologist at Rockefeller University in New York City.

“We kept raising our hands and saying, ‘You’re not going to be able to do that unless you have high-quality reference genomes,’” Jarvis recalls. But long-read sequencing would require more money, and not everyone was keen to deploy it. Jarvis recalls feeling frustrated by some of the debates. “I even chipped a little bit of my front tooth on a fork at a restaurant. I was biting on it so hard,” he says. Ultimately, he and others pushing for the long-read approach won.

Miga, who brings to the project a reputation for completing difficult-to-read sections of DNA, was already using long-read technology. She, along with Jarvis, Eichler and others, published the first-ever completely sequenced human genome, capturing all 3 billion letters, including the messy, highly repetitive sections that cap the ends of chromosomes — known as telomeres<sup>5</sup>. This first telomere-to-telomere genome sequence corrected numerous errors from previous references and uncovered around 100 unnoticed genes that probably code for proteins.

It was no simple feat, however. Typically, human cells contain two sets of 23 chromosomes — one from an egg and one from a sperm cell. But duplicated sequences and other structural DNA variations get jumbled up when machines try to read both sets at the same time. To circumvent this, the scientists analysed the DNA of a cell line derived from what’s known as a molar pregnancy, in which a sperm fertilizes an egg with no nucleus. The DNA contained only one set of chromosomes.

The 350 genomes for the Human Pangenome Project, by contrast, will come from diploid cell lines, that is, cells that contain copies from both parents, so scientists will have to use complex computational tools to tease the genomes apart and make sure they capture the structural variation accurately.

The pangenome effort has already completed around 70 detailed genomes. It aims to finish telomere-to-telomere versions of all 350 by the end of the grant, in mid-2024.

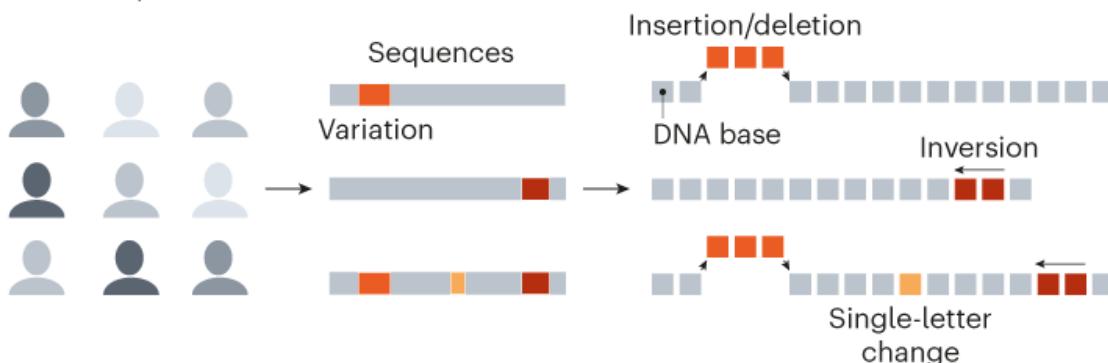
And scientists are already working on ways to visualize the diversity and showcase the variations. Up until now, including for the GRCh38 reference genome, the convention has been to have a simple linear representation and a companion database with variations listed for different positions in the sequence, such as single-letter changes. “The community has used this convenient fiction of linear reference sequence for 20 years,” says Benedict Paten, a computational biologist at the University of California, Santa Cruz. Paten, whose office is next to Miga’s, is collaborating with a group to improve the sophistication of the pangenome visualization. In this new visualization, coloured lines represent distinct variants. More-frequent variations are indicated with thicker lines. “Making human variation intuitive and easy to understand is part of our mission in integrating the pangenome,” Paten says (see ‘Visualizing a pangenome’).

# VISUALIZING A PANGENOME

The Human Pangenome Project aims to capture all of the variability in the human genome around the world. By analysing this variation and creating innovative ways to display it, the effort counters the assumption that there is a consensus of what a human genome looks like.

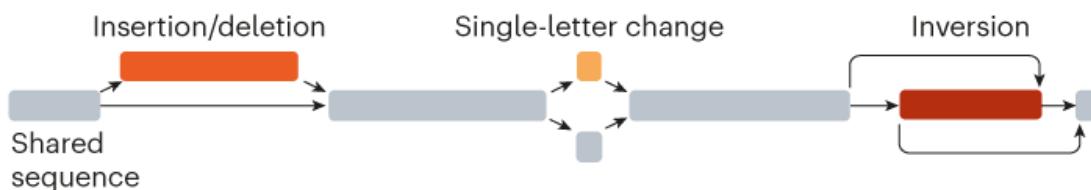
## Gathering samples

Researchers will have to produce high-quality sequences for hundreds of individuals and catalogue the variants, including single-letter changes, insertions, deletions and inversions.



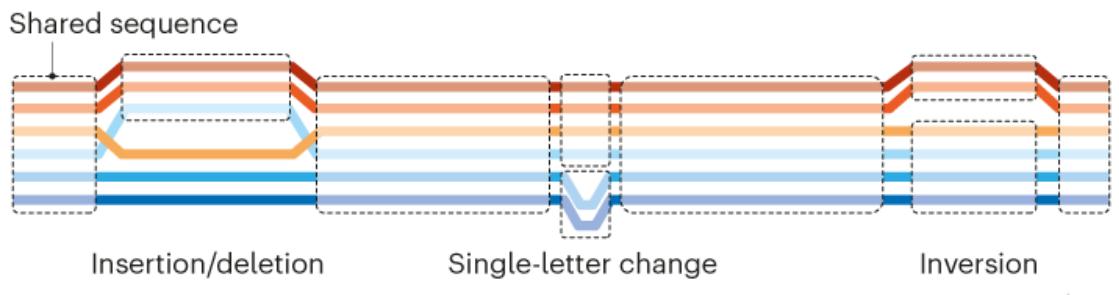
## Visualizing variation

Graphical models can present variation data in a way that doesn't assume a standard, or default reference genome.



## Exploring the pangenome

Representations that look like subway maps allow researchers to compare the variations in a population at a sequence level.



©nature

## Missteps and departures

Many of the 350 people whose genomes will be analysed in the Human Pangenome Project participated in the 1000 Genomes Project, an effort launched in 2008 to catalogue common and rare variants from 26 diverse populations. The DNA samples that were collected as part of that effort will be retrieved from cold storage and repurposed for the more detailed long-reads of the pangenome sequencing project. The consent forms that those individuals signed years ago also cover the use of their DNA data for the new project. But the Human Pangenome Project is taking further measures to ensure ethical collection and use of genetic data. In contrast to other major genetic sequencing efforts, in which scientists made decisions and then only had them vetted by an Institutional Review Board, for example, the Human Pangenome Project has social ethicists who are “embedded” in the decision-making process and continuously vetting the project, Eichler says.

As Wang puts it: “It’s really about how to guide the nerdy scientists who may not think about social issues to do their science in the most appropriate manner.”

In many ways, the leaders of the pangenome project are trying to overcome the ethically thorny legacy of past endeavours. The Human Genome Diversity Project, for example, launched in 1991 as an effort to collect DNA information from people around the globe, engendered staunch opposition from several communities. Indigenous groups, among others, felt they were being treated as living fossils, headed towards extinction<sup>6</sup>.

“Scientists were collecting Indigenous peoples’ genomic data largely for the benefit of other, non-Indigenous peoples, which, when done without regards to Indigenous data sovereignty, is a means of continued data extraction,” says Krystal Tsosie, a geneticist and bioethicist at Vanderbilt University in Nashville, Tennessee, and a member of the Navajo Nation.

The next decade brought even more concern over ethical transgressions in genetic studies of under-represented groups, notably when the Havasupai Tribe filed a lawsuit against the Arizona Board of Regents and Arizona State University researchers in 2004. Members of the tribe had donated their DNA for genetic studies on type 2 diabetes, but discovered that it had been used without their consent for studies on schizophrenia and migration<sup>7</sup>.

The researchers had also used stigmatizing words such as ‘inbreeding’ to explain genetic phenomena that were actually the consequence of population bottlenecks related to genocidal events, says Tsosie. She adds that, in the past, geneticists doing sequencing projects have [often used racial language](#) and failed to properly acknowledge the lasting legacy of colonialism in science, and the threat it poses to Indigenous people.

For several years, Fox and others have been calling for a massive departure from this approach. They say that Indigenous groups should have greater agency when it comes to the collection of their genetic data. Fox, who was a graduate student in Eichler’s lab, says that he’s not convinced that the pangenome project and others like it are involving the diverse groups they seek to sample in a way that truly empowers them. “I love Evan, man. When I have problems, I call him for advice,” he says. “Despite that, you know, we don’t agree on everything.”



Members of the Native BioData Consortium including Joseph Yracheta (far right) using a genome sequencing machine. Credit: Joseph Yracheta

Fox advocates for an approach that puts sequencing power in the hands of the people. He and Tsosie are involved in the Native BioData Consortium, a non-profit research institute led by Indigenous scientists and tribal members in the United States that has been working to help Indigenous groups to acquire and run DNA sequencers on their own territory. The first sequencer was delivered to the Cheyenne River Sioux reservation in December 2020 says consortium co-founder Joseph Yracheta, a public-health geneticist at the Johns Hopkins Bloomberg School of Public Health in Baltimore. In February, Yracheta joined a Human PanGenome Project working group focused on ethical, legal and social implications of the project.

Fox is currently focused on genetic complexity found in the Pacific islands. He and his team mates are taking a holistic approach to sequence the genomes of agricultural species and other organisms in the environment in tandem, and are building a genomics institute to serve the community. Fox notes that the latest technologies, such as a ‘distributed ledger’ computer system that securely ties a person to their genetic data, can give people greater autonomy about whom they allow to access and use their information. “There are so many advancements in the data sciences right now that really allow for a new level of agency for participants,” Fox says.

Eichler is supportive of Fox’s path. “I applaud his efforts to engage Indigenous scientists into genomics research — we need more of it,” Eichler says. “It is not an either-or scenario, however, in my opinion.” He adds that the Human PanGenome Project is encouraging Indigenous scientists to generate their own reference genomes. In those scenarios, “we will work together to make it happen by providing expertise and tools as needed”.

## No mutation without representation

Tsosie says that Indigenous groups might collaborate with big diversity projects in the future, but that it would have to happen in a way that would ensure that such communities can do their own sequencing. Moreover, although these major genome projects are often open-data efforts, Tsosie says it would be wise for there to be protections added for Indigenous people’s deposited DNA sequences such that they be available only through

access requests [to avoid exploitation](#). “If it’s going to happen, it needs to happen in the best way that represents Indigenous people,” she says.

It’s not just advocates from Indigenous communities in the United States who have voiced concerns about representation and data ownership. Others have argued that the pangenome project hasn’t adequately involved researchers from regions outside the United States, according to Jarvis, who is on the project’s sampling committee. He recognizes that some see the initiative as a largely US effort, but says that he and his collaborators are working to broaden it and involve scientists and participants from different parts of the world. For example, they have reached out to leaders of the Human Heredity and Health in Africa (H3Africa) programme to involve scientists in Africa who can do sequencing in countries there. (No sequencing effort seems immune from ethical challenges, however — even the H3Africa programme has had to straddle different countries’ rules and norms governing the use of participant data, for example.)

Jarvis says he wants the Human Pangenome Project to achieve a better representation of human genetic diversity. “I’m a person of colour. I grew up as an African American. I grew up as an under-represented minority in the sciences,” he says. “My diversity is not represented. So I have a personal motivation and a societal one to make sure that this pangenome really represents populations.”

As they push forwards, the scientists also acknowledge that 350 genomes will not represent all human diversity. Ultimately, the true number of genomes needed to do this is difficult to pin down, and genetics often teaches us that rare differences can be important. “I don’t think there is any magic number,” says Adam Phillippy, head of the Genome Informatics Section at the NHGRI, and an investigator on the pangenome project.

Juggling the massive scientific undertaking while trying to avoid ethical pitfalls is something that weighs heavily on the pangenome researchers. “I’m sure there will be things that we will do that people will criticize five or ten years from now. I’m almost 100% sure of it,” Eichler says. “But if we can go in with a clear conscience and say, we tried to do everything we possibly could to do it right, I feel that that’s something.”

*Nature* **603**, 378–381 (2022)

doi: <https://doi.org/10.1038/d41586-022-00726-y>

## References

1. Hsieh, P. *et al.* *Science* **366**, eaax2083 (2019).
2. Sherman, R. M. *et al.* *Nature Genet.* **51**, 30–35 (2019).
3. Miga, K. H. & Wang, T. *et al.* *Annu. Rev. Genom. Hum. Genet.* **22**, 81–102 (2021).
4. Tettelin, H. *et al.* *Proc. Natl Acad. Sci. USA* **102**, 13950–13955 (2005).
5. Nurk, S. *et al.* Preprint at bioRxiv  
<https://doi.org/10.1101/2021.05.26.445798> (2021).
6. Dodson, M. & Williamson, R. *J. Med. Ethics* **25**, 204–208 (1999).
7. Garrison, N. A. *et al.* *Sci. Technol. Hum. Values* **38**, 201–223 (2013).

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00726-y>.

| [Section menu](#) | [Main menu](#) |

## Books & Arts

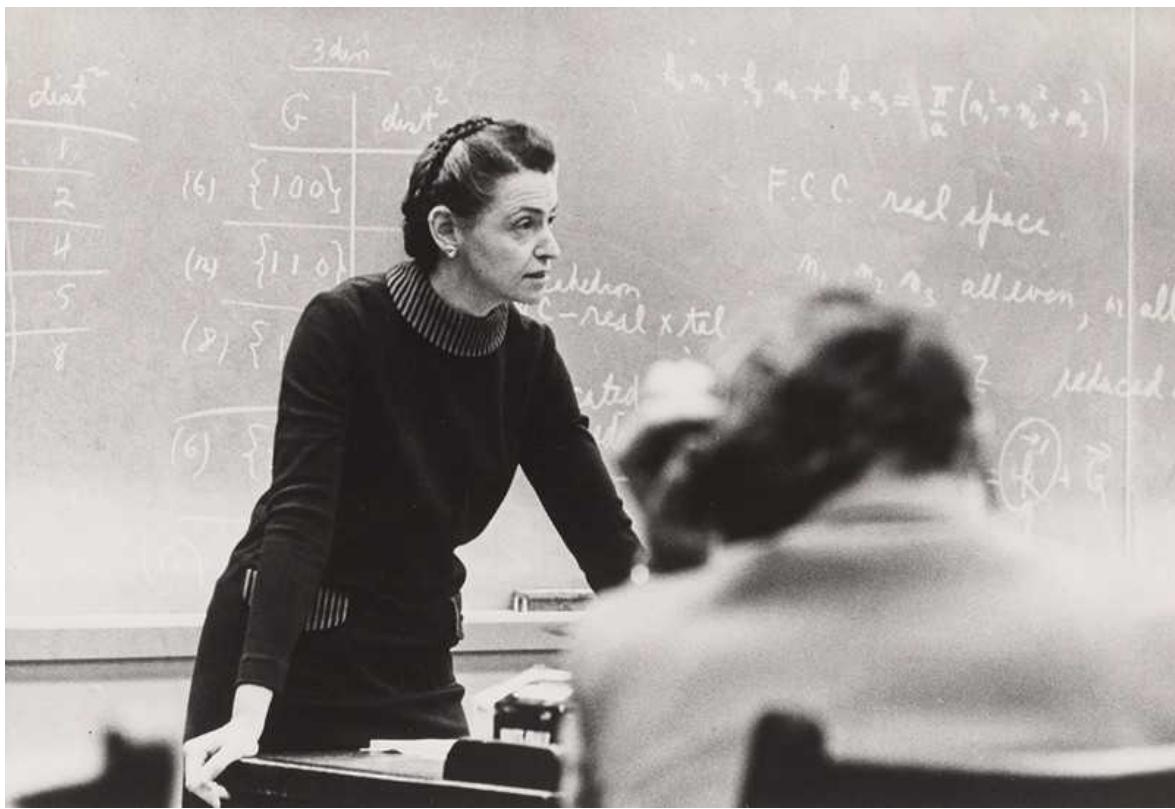
- **[Queen of carbon, champion of women in science](#)** [ 15 March 2022]  
Book Review • From graphite to graphene and government, Mildred Dresselhaus fought for equity all the way.
- **[An open-access history: the world according to Smits](#)** [ 14 March 2022]  
Book Review • The Plan S architect, scourge of paywalls, reveals how the policy sausage got made.

- BOOK REVIEW
- 15 March 2022

# Queen of carbon, champion of women in science

From graphite to graphene and government, Mildred Dresselhaus fought for equity all the way.

- [Ariana Remmel](#) 0



Mildred Dresselhaus lecturing at the Massachusetts Institute of Technology in Cambridge. Credit: MIT Museum

## **Carbon Queen: The Remarkable Life of Nanoscience Pioneer Mildred Dresselhaus** Maia Weinstock MIT Press (2022)

Carbon nanostructures are all around us, from sports gear to microelectronics and reinforced concrete. Among the people we have to thank for that is Mildred Dresselhaus. Her trailblazing research into the fundamental physics of materials such as graphite and carbon nanotubes in the second half of the twentieth century, and her advocacy of equality in science, saw her dubbed queen of carbon.

Among the stars of her generation, Dresselhaus overcame hardship and discrimination to win just about every honour short of a Nobel prize. Even before she began hobnobbing with scientific celebrities — taking daily walks with physicist Enrico Fermi, for example — Dresselhaus’s childhood talents as a violinist earned her an audience with first lady Eleanor Roosevelt. By the end of her illustrious career, a 2014 visit to Washington DC to receive a US Presidential Medal of Freedom from Barack Obama was almost a footnote. Now, science writer Maia Weinstock chronicles Dresselhaus’s extraordinary life in the spirited biography *Carbon Queen* — bringing to it the same exuberance with which she once created a custom Lego figure to celebrate her heroine.

## **Promising properties**

Born in 1930 to Polish Jewish immigrants in Depression-era New York City, Dresselhaus was a determined student. Food was often so scarce that she began working odd jobs at the age of eight to support her family, Weinstock relates — even while competing for the limited educational opportunities afforded to girls. After postgraduate physics studies at Harvard University in Cambridge, Massachusetts, she earned her doctorate from the University of Chicago in Illinois in 1958. The same year, she married physicist Gene Dresselhaus, who would become a collaborator.

In the 1960s, at the Massachusetts Institute of Technology (MIT) Lincoln Laboratory in Lexington, Dresselhaus fell for the physical properties of carbon — an intellectual challenge but a research backwater at the time. “The little bit I learned made me wonder why no one was interested,” she

told *The University of Chicago Magazine* in 2015. There was little pressure to churn out findings at the breakneck pace of more fashionable topics. This gave Dresselhaus a buffer against academic expectations as she raised four children with her husband, Weinstock explains.

Dresselhaus wanted to understand how electrons flow through graphite, which is made of stacks of single-atom-thick sheets of carbon bonded in a network of hexagons like a honeycomb. Graphite has some of the same properties as a conventional semiconductor but its electronic structure is unique, and Dresselhaus recognized that the unassuming material might be hiding some fascinating physics, Weinstock writes.

One of her earliest moments of acclaim came from a 1968 paper upending the prevailing view of graphite's electronic structure ([P. R. Schroeder et al. Phys. Rev. Lett. 20, 1292; 1968](#)). It was so controversial that a reviewer revealed himself to Dresselhaus to warn her against publishing. She and her colleagues went ahead knowing that they risked ruining their careers, as she recounted in a 1987 lecture. The work unleashed a flood of publications confirming the findings, and jump-started a broader interest in carbon-based materials.



The custom Mildred Dresselhaus Lego figurine created by author Maia Weinstock.Credit: Maia Weinstock

Dresselhaus made a habit of bucking trends, following her experimental evidence even when the conclusions ran counter to accepted knowledge. In the 1980s and 1990s, she made discoveries that presaged the existence of buckminsterfullerene, or ‘buckyballs’, and the possibility of elongating these carbon spheres to form ‘buckytubes’. She predicted that the properties of a nanotube, made of a rolled-up sheet of carbon a single atom thick, would depend on the orientation of the hexagons. This was validated years later.

In the final chapter of her research career, Dresselhaus worked on the fundamental properties of graphene, a single sheet of carbon atoms that looks like chicken wire. When Andre Geim and Konstantin Novoselov won the 2010 Nobel Prize in Physics for “groundbreaking experiments regarding the two-dimensional material graphene”, both acknowledged Dresselhaus’s contributions in their award speeches.

## **Stalwart advocate**

The queen of carbon was also a champion for women in science, as Weinstock shows. Early in her career, Dresselhaus was often the only woman in a research group or institution. One adviser told her that fellowships and grant money were wasted on women. Such blatant discrimination made it difficult for her to envision a long-term future in research — she expected that her presence in a male-dominated laboratory would be tolerated at best — but she persevered. Weinstock's biography takes pains to present Dresselhaus's ideals to a modern audience that considers tolerance of scientists from historically excluded groups to be a bare minimum for equity. Although Dresselhaus did not dwell on these incidents, Weinstock highlights how she decided to pay forward the support she got from mentors such as Nobel-prizewinning physicist Rosalyn Yalow.

As a tenured professor at MIT, Dresselhaus advocated for admissions reforms and created support systems that cultivated the inclusion of more women at the institution and beyond. Rising through the ranks of academic administration, in 2000 she became director of the US Department of Energy Office of Science, where she managed national research laboratories and a budget of US\$2.8 billion, and continued to champion young scientists.

Weinstock navigates the complexities of theoretical physics and research bureaucracy deftly. She describes forms of carbon — from diamond to graphite — and their properties with sleek diagrams and colourful analogies that unpack basic principles and broader implications. And she situates Dresselhaus in the cultural context of her time in research. But the rosy rags to riches telling often flattens Dresselhaus as if she were a character from a fairytale, fated to change the world. Nevertheless, as a much more recently trained chemist, I found the story an arresting reminder of the shoulders on which we now stand.

*Nature* **603**, 383-384 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00728-w>

| [Section menu](#) | [Main menu](#) |

- BOOK REVIEW
- 14 March 2022

# An open-access history: the world according to Smits

The Plan S architect, scourge of paywalls, reveals how the policy sausage got made.

- [Richard Van Noorden](#)



Robert-Jan Smits at the launch of the European Union's Horizon 2020 research-funding programme. Credit: Luigi Mistrulli/Sipa/Shutterstock

## **Plan S for Shock: Science. Shock. Solution. Speed.** *Robert-Jan Smits & Rachael Pells* Ubiquity (2022)

In 2018, a group of influential research funders struck a blow in the decades-long fight to end paywalls in science. Peer-reviewed papers from research they supported must be made open access immediately on publication, they declared: free to read, download and redistribute.

This radical pledge, called Plan S, began to take effect last year. It is now supported by around two dozen funders, most in Europe. (One estimate of their influence is that around 12% of articles in the most highly cited journals acknowledge these funders.) Big names include the European Commission, national funders in France and the United Kingdom, the London-based charity Wellcome, the Bill & Melinda Gates Foundation in Seattle, Washington, and the Howard Hughes Medical Institute in Chevy Chase, Maryland. This has nudged many subscription journals into offering open-access (OA) publishing routes. But there is still a long way to go to eliminate paywalls, as Robert-Jan Smits, architect of Plan S, reflects in an insider account of his frenetic year putting the policy together.

In *Plan S for Shock*, Smits — now president of the executive board of Eindhoven University of Technology in the Netherlands — and journalist co-author Rachael Pells emphasize that nothing less than a transformation of academia's reward systems is needed to achieve full OA. The authors also introduce dozens of other viewpoints, from researchers, publishers (including Springer Nature, which publishes *Nature*) and non-academics frustrated that they can't easily access the knowledge they need. The book is itself free to download: it is published by the OA Ubiquity Press in London, and Pells was paid for her time by OA publisher Frontiers, based in Lausanne, Switzerland. (Frontiers had a key role in encouraging Smits to introduce Plan S.)

Few disagree with the compelling case, laid out in the book's first section, that it will benefit science and society if anyone can access peer-reviewed scholarly work for free. Scientists' career incentives, however, are tied not to open publishing, but to getting their work into the most prestigious journals possible. And as Smits noted when he became the European Commission's OA envoy in 2018 (after eight years in charge of the commission's research

directorate), the majority of journals, especially the highly selective and prestigious ones, run on subscriptions. “The academic community had become entangled in a cobweb it couldn’t get out of,” Smits says.

## Bold move

To break the bind, Smits corralled a coalition of the willing: funders prepared to insist that scientists publish outside paywalls. After publishers retorted that they would probably fold if their journals had to flip immediately to OA models, Plan S made temporary allowances for publishing in hybrids: subscription journals that offer an OA option. Researchers are also allowed to publish behind a paywall and simultaneously post peer-reviewed accepted manuscripts in online repositories, such as PubMed Central or their institution’s own repository. Some journals are trialling this model; others oppose it. (The book also mentions widely used pirate sites for accessing papers published behind paywalls, such as Sci-Hub.)

Plan S remains too bold for most funders. Many, including government agencies in the United States and China, have so far been reluctant to require immediate open access to research. Unfortunately, this book doesn’t dive deeply into the reasons why. And it sometimes seems out of date. It notes, for instance, that in 2018, at a conference in Berlin, China promised to support Plan S — but doesn’t say that nothing came of this promise.

It’s clear that a major problem is money. Journals often generate revenues for OA publishing through per-paper author fees called article processing charges (APCs), deterring scientists whose funders or universities won’t or can’t cover the costs. There are plenty of low or no-cost OA platforms, usually subsidized by governments or funders. But high-prestige, selective journals are generally expensive (*Nature*’s APC is US\$11,390; *Cell*’s is \$9,900). *Science* is experimenting with allowing those researchers with Plan S mandates — and no one else — to share their manuscripts immediately online under OA licences. Smits says he had hoped to regulate the APC market — his draft plans suggested that funders cap the OA fees they’d cover at €2,000 (US\$2,200) per paper. Plan S funders voted against this idea, which Smits regrets.

## Incentive shift

The longer-term view is that science's incentive system is faulty: funders and universities should not assess scientists on where they publish, only on what they publish. Quoting the Declaration on Research Assessment, a global initiative to change these incentives, Smits supports eliminating journal-based metrics, such as the Journal Impact Factor, from academic assessment. Many scientists are on board with this, but the move might not remove the lure of prestige. Plan S funders have said they'll value the "intrinsic merit" of work and not the publication channel, but for now, it's difficult to know whether grant assessors will really stop using publication in big-name journals as a marker of work worth paying attention to.

As the book goes on to discuss, some argue that the concept of journals that filter the literature before publishing it — the selectivity that builds prestige — needs to disappear. They say that all scientific papers should be published openly first, and that recommendation and summary services could then pick out the best or most interesting for a particular community. Today's journals might, perhaps, evolve into these products.

None of these ideas is new to Plan S cognoscenti. For them, the book's drama is in the details of how the policy sausage gets made. (A final reveal is that Plan S was itself named after Smits — it doesn't stand for shock, science, solution or speed, as he used to suggest). But aficionados will probably spot misstatements about the size of journal profits and taxpayer-funded research, along with a foreword that equates open access with author fees, a mistake the book itself avoids.

Perhaps the acid test for *Plan S for Shock* is whether it can reach a broad audience of scientists, rather than the already engaged experts. The book sometimes suffers from thickets of jargon that confuse many discussions of academic publishing — but its lively quotes could pull readers through.

This matters because one enduring puzzle of the OA movement is that it's never really managed to engage the majority of scientists. Many researchers still don't bother to archive their past peer-reviewed papers online, even if it's been years since the work appeared in journals and it could now be

shared — and even though studies suggest that openly shared articles attract more citations.

In a perceptive quote, Paul Ayrис, chief executive of UCL Press in London, notes that change in academic organizations is never really implemented by top-down enforcement. What's needed instead is to work with academics to show them the benefits they will reap from the change. Smits describes Plan S as a stone cast into water, sending ripples and waves through science publishing even though its ultimate goal is far from complete. Perhaps this account will convert a few more scientists to the cause.

*Nature* **603**, 384-385 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00717-z>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00717-z>

| [Section menu](#) | [Main menu](#) |

# Opinion

- **[C. Thomas Caskey \(1938–2022\)](#)** [ 11 March 2022]  
Obituary • Geneticist who demonstrated the universality of life's code.
- **[Paul Farmer \(1959–2022\)](#)** [ 10 March 2022]  
Obituary • Global public-health pioneer and equity advocate.
- **[Africa: sequence 100,000 species to safeguard biodiversity](#)**  
[ 15 March 2022]  
Comment • Build a major genomics resource on the continent to help breeders and conservationists.
- **[Ukraine: thousands sign plea for scientific sanctions against Russia](#)** [ 08 March 2022]  
Correspondence •
- **[Secure molybdenum isotope supplies for diagnostics](#)** [ 15 March 2022]  
Correspondence •
- **[Predatory journals and conferences — three drivers](#)** [ 15 March 2022]  
Correspondence •
- **[Forest protection: invest in professionals and their careers](#)**  
[ 15 March 2022]  
Correspondence •

- OBITUARY
- 11 March 2022

## C. Thomas Caskey (1938–2022)

Geneticist who demonstrated the universality of life's code.

- [Jan Witkowski](#) 0



Credit: H. Parsons

Charles Thomas (Tom) Caskey contributed to many advances in human genetics. He demonstrated the universality of the genetic code; developed one of the earliest applications of DNA fingerprinting in forensic science; was key to the early phases of the Human Genome Project (HGP); and helped to initiate personalized medicine. Enthusiastic, optimistic and

extraordinarily generous with his time and knowledge, Caskey mentored a generation of young scientists and physicians. He has died aged 83.

Caskey attended the University of South Carolina before entering Duke University Medical School in Durham, North Carolina. When he qualified in 1965, the Vietnam War was ramping up. As required by the ‘Doctor Draft’, he joined a select group of young physicians, known pejoratively (but later honourably) as the Yellow Berets, in the US Public Health Service. He joined the laboratory of future Nobel laureate Marshall Nirenberg at the National Institutes of Health (NIH) in Bethesda, Maryland.

Four years earlier, Nirenberg and Heinrich Matthaei had deciphered the first codon — the three-nucleotide ‘word’ that encodes an amino acid — of the genetic code. Caskey demonstrated the universality of the code by showing that transfer RNAs from bacteria, amphibians and mammals recognized the same codons. Nirenberg wrote: “These results had a profound philosophical impact on me because they indicate that all forms of life on this planet use essentially the same language.” Caskey also identified the signals that tell the cell’s protein-synthesis machinery when to stop translating a messenger RNA.

In 1971, Caskey moved to Houston, Texas, to establish what is now the Department of Molecular and Human Genetics at Baylor College of Medicine (BCM). His research shifted to the gene encoding the enzyme hypoxanthine phosphoribosyl transferase (HPRT), which is mutated in the rare neurological disorder Lesch–Nyhan syndrome. After the cloning of the first mammalian genes in the mid-1970s, he spent 1979 with Sydney Brenner at the Laboratory of Molecular Biology in Cambridge, UK, learning to use recombinant DNA techniques to characterize mutated genes. His many studies on the HPRT gene culminated in 1990 in the publication of its complete DNA sequence.

At BCM, Caskey created an environment that stimulated gene discovery and the development of diagnostic tools by establishing facilities for mapping, cloning and sequencing disease-related genes. The successes of the highly motivated and skilful researchers he recruited, mentored and trained included determining the genes and mutations underlying fragile X

syndrome, myotonic dystrophy, Charcot–Marie–Tooth disease and Rett syndrome.

In 1986, Caskey established the Robert J. Kleberg DNA Diagnostics Laboratory at BCM, one of the first DNA diagnostic labs in the United States. It initially offered prenatal diagnosis of sickle-cell disease (using allele-specific oligonucleotides) and Duchenne muscular dystrophy (using linkage analysis), adding more tests as more genes were discovered. Progress in human genetics has been extraordinary, and today Baylor Genetics, successor to the Kleberg lab, offers whole-genome sequencing, completed in five days, as a diagnostic tool.

Caskey participated in several advisory committees of both the Department of Energy and the NIH that evaluated the merits of an international HGP. He served as president of the Human Genome Organization and chair of Genome Canada. He and BCM led by example. The HPRT gene, which in 1990 was the longest human gene to have been sequenced, was the first to be read using an automated sequencer. The BCM Human Genome Sequencing Center, founded in 1996, ultimately contributed 10% of the human genome sequence completed by the HGP in 2003.

One of the department's most widely used discoveries came about as part of the research programme on gene mapping. In 1991, Caskey reported that short tandem repeat sequences found in non-coding regions of the genome, such as TATATATATA or GTCGTCGTCGTC, were highly variable from one individual to another and could be used for identity testing. The test he patented can be directly linked to hundreds of millions of forensic DNA tests performed since, and was used to identify people killed in the Gulf War, as well as victims of disasters. Atypical expansion of repeats causes fragile X syndrome and more than 60 other genetic disorders.

In 1994, Ed Scolnick, a former postdoc of Caskey's and at the time head of research at the pharmaceutical company Merck, headquartered in Kenilworth, New Jersey, invited him to join the company. The result was the Merck Genome Research Institute, which applied genomic approaches to drug and vaccine development. Caskey had a strong entrepreneurial streak: he also served as managing director of venture-capital company Cogene

Ventures in Houston, an investor in early-stage life-sciences enterprises based on genomic technology.

Tom was a wonderful participant in many meetings I organized at the Banbury Center at Cold Spring Harbor Laboratory in New York. He would ask probing, helpful questions in his instantly identifiable Southern drawl. An entertaining companion and a consummate story-teller, Tom, as he admitted, was never one to let the truth spoil a good tale. He was an inveterate traveller and a passionate sailor: eminent scientists occasionally found themselves at sea in severe storms, fearing their imminent demise.

In 1982, Tom closed a Banbury meeting with these words: “There is a new scientific world emerging in our study of the genetic character of man and his inherited flaws.” We all owe Tom an immense debt of gratitude for helping to bring that new world into being.

*Nature* **603**, 386 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00714-2>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00714-2>

| [Section menu](#) | [Main menu](#) |

- OBITUARY
- 10 March 2022

# Paul Farmer (1959–2022)

Global public-health pioneer and equity advocate.

- [Amy Maxmen](#)



Credit: Gary Coronado/Palm Beach Post/ZUMA/Alamy

“What happens to poor people is never divorced from the actions of the powerful,” Paul Farmer wrote in his 2005 book *Pathologies of Power*. A doctor, medical anthropologist and activist, Farmer devoted his life to advocating for health equity. As a co-founder of Partners In Health (PIH), a non-profit organization that provides free medical care in low-income

countries including Haiti, Peru and Rwanda, he used the group's results to change global guidelines on how to treat tuberculosis and HIV. During the COVID-19 pandemic, Farmer and his colleagues denounced monopolies on vaccines that help to account for why fewer than 10% of people have been fully vaccinated in low-income countries ([P. Erfani et al. BMJ 374, n1837; 2021](#)). Farmer treated patients up until his death in Rwanda, aged 62.

In 1990, he earned a PhD in anthropology alongside a medical degree from Harvard Medical School in Boston, Massachusetts, where he later taught global health and social medicine. His views were shaped by his own experiences as a child in the United States and as a young adult in Haiti. They were deepened by his knowledge of social theory, political theory and the Catholic philosophy of 'liberation theology'. This study focused his thinking on the systemic oppression of poor people.

Farmer's 12 books and more than 200 manuscripts reveal the principles that guided his actions. The money needed to save lives exists if lives are valued equally, Farmer argued. He criticized the public-health field for cost-effectiveness analyses used by governments and donors to calculate when medical technologies that they take for granted are worthwhile for those who cannot pay for them themselves. In an editorial in the *World Health Organization Bulletin* in 2003, he criticized those in the public-health community who pushed for HIV prevention rather than care in poor countries because it was cheaper ([P. Farmer Bull. World Health Organ. 81, 699; 2003](#)). At the time, HIV drugs were astronomically expensive — but did not need to be, as he pointed out. Two years later, policy changes allowed generics to enter the market and prices fell drastically.

To achieve his vision of a world in which everyone has access to health care, Farmer sometimes bent the rules. In the early days of PIH in the 1990s, he and co-founder Jim Yong Kim smuggled some US\$92,000-worth of second-line tuberculosis drugs out of Brigham and Women's Hospital in Boston, Massachusetts, where they both worked, to their patients in Peru, according to [a profile in The New Yorker](#). (A PIH donor later reimbursed the hospital for the drugs.) Years later, Farmer and Kim threw their support behind an ambitious World Health Organization initiative to treat three million people with HIV/AIDS by 2005. They understood that donors prefer to fund

programmes that focus narrowly on a single disease, rather than on health care broadly.

In private, however, they wanted to restructure systems, says Farmer's close colleague, Adia Benton, an anthropologist at Northwestern University in Evanston, Illinois. "They knew that to get this many people on HIV care, you'd have to change patent law, you'd have to change manufacturing, you'd have to build obstetrician–gynaecology clinics. He'd argue that you can't take care of maternal-to-child transmission without good prenatal care," she explains. "They were hustlers."

PIH stands apart from most other aid organizations in that it attempts not only to build clinics but also to ensure they remain sustainable by operating within government-run services and enlisting local staff at every level. "I am a living testament to that," says Bailor Barrie, the executive director of PIH Sierra Leone, a branch established during the Ebola outbreak in 2014.

Farmer and Barrie met when the latter was a global-health student in Farmer's class at Harvard Medical School. "He is my teacher, my mentor, my colleague, my friend. I call him Pa because he is like a father," Barrie says. "I am devastated."

Farmer encouraged scientists studying diseases to account for forces such as racism, sexism and poverty that hinder people's abilities to take advantage of the fruits of science. Although rooting out the sources of oppression is a tall order, Farmer described how medical programmes could work better if they strove for equity ([J. S. Mukherjee et al. \*Lancet Glob. Health\* 7, E410–E411; 2019](#)). In that paper, he wrote about how 100% of people who received small monthly stipends and food alongside free tuberculosis treatment from his team in Peru were cured of the disease, compared with only 56% of those given the drugs alone.

Benton suggests that Farmer's political orientation stemmed from his experience of growing up in relative poverty: one of six children, he lived in a bus, a boat and a tent. His medical care for a broken leg as a graduate student in 1988 cost about twice his mother's annual salary as a cashier, he once wrote in the *London Review of Books*. He reflected that most of his bill was covered by Harvard's medical insurance, whereas such health costs impoverish some 30 million households each year around the world — if

people get care at all. “Three of the Haitian founders of PIH, all in their twenties, had recently died stupid deaths,” he wrote, of the preventable and easily treated diseases of sepsis, malaria and typhoid fever.

Farmer spent his career trying to convince people that health care is a human right. For this, he became a celebrity in global-health spheres. He leaves a legacy of researchers committed to carrying his mission forward. “He wasn’t just a guy with a vision,” Benton says. “He was brilliant.”

*Nature* **603**, 387 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00715-1>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00715-1>

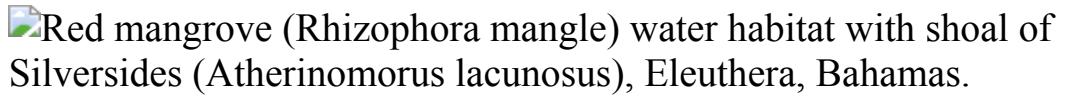
| [Section menu](#) | [Main menu](#) |

- COMMENT
- 15 March 2022

# Africa: sequence 100,000 species to safeguard biodiversity

Build a major genomics resource on the continent to help breeders and conservationists.

- [ThankGod Echezona Ebenezer](#)<sup>0</sup>,
- [Anne W. T. Muigai](#)<sup>1</sup>,
- [Simplice Nouala](#)<sup>2</sup>,
- [Bouabid Badaoui](#)<sup>3</sup>,
- [Mark Blaxter](#)<sup>4</sup>,
- [Alan G. Buddie](#)<sup>5</sup>,
- [Erich D. Jarvis](#)<sup>6</sup>,
- [Jonas Korlach](#)<sup>7</sup>,
- [Josiah O. Kuja](#)<sup>8</sup>,
- [Harris A. Lewin](#)<sup>9</sup>,
- [Roksana Majewska](#)<sup>10</sup>,
- [Ntanganedzeni Mapholi](#)<sup>11</sup>,
- [Suresh Maslamoney](#)<sup>12</sup>,
- [Michèle Mbo'o-Tchouawou](#)<sup>13</sup>,
- [Julian O. Osuji](#)<sup>14</sup>,
- [Ole Seehausen](#)<sup>15</sup>,
- [Oluwaseyi Shorinola](#)<sup>16</sup>,
- [Christian Keambou Tiambo](#)<sup>17</sup>,
- [Nicola Mulder](#)<sup>18</sup>,
- [Cathrine Ziyomo](#)<sup>19</sup> &
- [Appolinaire Djikeng](#)<sup>20</sup>



The red mangrove tree is indigenous to Africa and is being sequenced as part of the African BioGenome pilot project. Credit: Shane Gross/Nature Picture Library

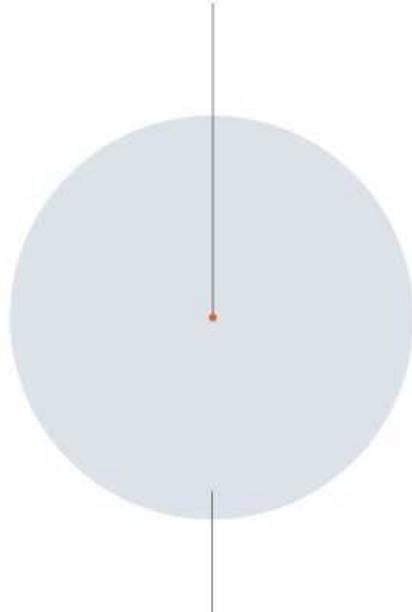
Sleeper fish (*Bostrychus africanus*) are a staple food in West Africa. Harvesting them provides an important source of income for hundreds of communities across the Gulf of Guinea in the Atlantic Ocean. Yet little is known about the genetics of this fish — information that is crucial to safeguarding its genetic diversity, and to enhancing its resilience in the face of climate change and other pressures.

This situation is all too familiar across Africa. Consider orphan crops, which have a crucial role in regional food security, even though they are not typically traded internationally. More than 50% of these have not had their genomes sequenced — from the fluted pumpkin (*Telfairia occidentalis*) to the marama bean (*Tylosema esculentum*). The same is true of more than 95% of the continent's known endangered species (see 'Africa's neglected genomes').

## AFRICA'S NEGLECTED GENOMES

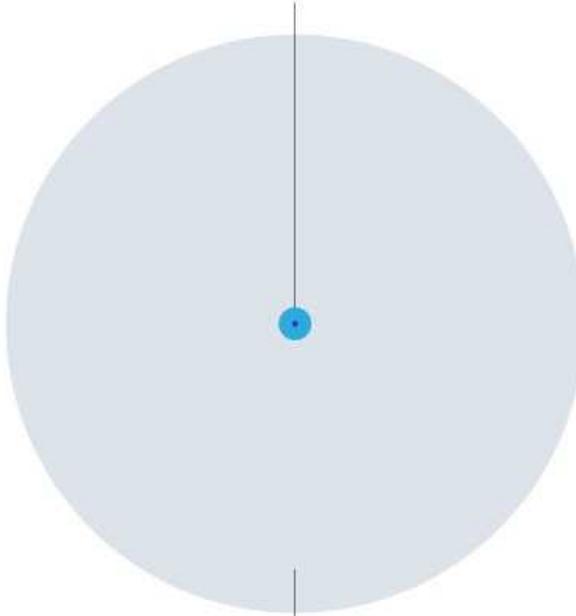
Around 95% of Africa's known endangered species are currently unsequenced.

**20**  
African plants sequenced\*,  
of which **0** were sequenced in Africa.



Around **60,000** known species  
of plant are found in Africa.

**385**  
African animals sequenced\*,  
of which **11** were sequenced in Africa.



Around **123,000** known species  
of animal are found in Africa.

\*Of 798 plants and 3,278 animals sequenced globally.

©nature

Sources: Analysis by T. E. Ebenezer *et al.*/Ref. 1/S. Hotaling *et al.* *Proc. Natl Acad. Sci. USA* **118**, e2109019118 (2021)

What's more, by our estimate, around 70% of the 35 or so projects that have focused on studying, conserving or improving biological diversity in Africa over the past 15 years have been led from outside the continent. In fact, among the plant genomes sequenced globally over the past 20 years, almost all of the African species were sequenced elsewhere — mainly in the United States, China and Europe<sup>1</sup>. This offshoring slows down the much-needed building of expertise and resources in genomics and bioinformatics in Africa (see 'Africa left out of global genomics efforts').

The [African BioGenome Project](#) (AfricaBP) is an effort to sequence the genomes of 105,000 endemic species: plants, animals, fungi, protists and

other eukaryotes. It currently involves 109 African scientists (87 of whom work in Africa) and 22 African organizations.

This store of reference genomes — built in Africa, for Africa — will help plant and animal breeders to produce resilient and sustainable food systems. It will inform biodiversity conservation across the continent. And it will strengthen Africa's ability to deliver on the goals of the post-2020 global biodiversity framework of the Convention on Biodiversity (CBD). These goals, one of which is to maintain at least 90% of genetic diversity for all known species by 2030, are to be agreed on next month at a meeting in Kunming, China.

## Africa left out of global genomics efforts

Most projects that aim to study, conserve or improve biological diversity in Africa have been led by researchers outside the continent.

Projects to sequence biodiversity rarely meet the needs of people in Africa or align with its countries' science agendas<sup>4,14–16</sup> (such as on agricultural technologies<sup>15</sup>). Take the Human Genome Project. Less than 2% of genomes analysed in the two decades since the project began are from African individuals, even though [Africa harbours more human genetic diversity than any other continent](#).

African researchers who contribute to data collection in such projects are not always credited for their work. A 2021 study<sup>17</sup> revealed that about 15% of 32,061 articles on global health research conducted in sub-Saharan Africa had no authors based in the country in which the research took place.

Currently, the International Nucleotide Sequence Database Collaboration, the core infrastructure for the collection and sharing of the world's nucleotide sequence data and metadata, names only those who have submitted samples or sequence data, not the primary owners or custodians of the sample. In practice, this means that if an African scientist collects samples from Bioko squeaker frogs (*Athroleptis bioko*) in Equatorial Guinea, for example, and sends them to a colleague in Canada who then submits a sequence to the database, only the Canadian researcher will receive recognition for the data. Recent efforts by the consortium and

others<sup>18</sup> will help to address some of these gaps. By December this year, the consortium will make it mandatory for those submitting sequence data to declare the country or region in which the sample was collected. But it is still unclear whether the credit given to sample custodians will be similar to that of sample submitters.

Besides the lack of recognition, African researchers rarely retain access to the data they help to collect, nor do they receive related benefits — either from royalties resulting from specific discoveries in genetics, or those stemming from technological advances and growth in scientific capability that such projects can bring.

For instance, during the 2014–16 Ebola epidemic in West Africa, around 269,000 blood samples were obtained from patients for diagnosis.

Thousands of those samples were shipped overseas, including to Europe and North America. None of the genomics researchers working in Africa knows where these samples are now housed<sup>19</sup> and, as far as the African human-genetics community knows, the sample providers never received the results of their blood collections.

An AfricaBP pilot project [was launched in June 2021](#). In this, researchers are sequencing 2,500 indigenous African species, including the Boyle's beaked blind snake (*Rhinotyphlops boylei*) from southern Africa and the red mangrove tree (*Rhizophora mangle*) from Nigeria. They are also mapping out the ethical, legal and social issues raised by a major biodiversity sequencing project — because of cultural sensitivities around certain species, or questions around who has access to the data and who benefits from any resulting discoveries.

For AfricaBP to be scaled up and sustained over the next decade, agencies and organizations need to allocate long-term investments to the project. Such groups include the African Union Commission, national and regional scientific agencies (such as the African Academy of Sciences), and international partners and organizations, including the US National Science Foundation and the UK research funder Wellcome. By our calculations, this will require at least US\$100 million per year for the next 10 years (see ‘AfricaBP: structure and costs’).

Some might argue that \$1 billion would be better spent on combating malnutrition and disease in impoverished communities across Africa. Yet consider the Human Genome Project, which cost around \$3 billion in 2003. By 2019, the human genetics and genomics sector alone was contributing \$265 billion annually to the US economy<sup>2</sup>. Likewise, the World Bank invested millions of dollars in outbreak preparedness from 2017, some of which was used to fund the African Centre of Excellence for Genomics of Infectious Diseases in Ede, Nigeria. This investment meant that Africa was much better equipped to meet the [challenges presented by the COVID-19 pandemic.](#)

## AfricaBP: structure and costs

The [African BioGenome Project](#) (AfricaBP) will involve researchers and organizations from all economic regions in the African Union, and will cost US\$100 million per year.

AfricaBP will convene 55 African researchers and policymakers from genomics, bioinformatics, biodiversity and agriculture — 11 for each of the 5 African Union geographical regions (northern, eastern, southern, central and western Africa). Another 165 people will be involved in the project (33 for each geographical region), including academic and industrial researchers, policymakers, and staff from governmental organizations, such as the National Institute of Agricultural Research of Morocco.

Ultimately, these people will feed genome sequences into various national or regional facilities. These include the National Gene Bank of Tunisia, which is using genetics to promote the conservation and sustainable use of Africa's plants, animals, fungi and protists, and the International Center for Research and Development on Livestock in the Subhumid Zone in Bobo-Dioulasso, Burkina Faso, which was established in 1994 to reduce poverty by improving food and nutritional security.

We estimate that producing high-quality reference genomes for around 105,000 endemic African species will cost around \$850 million to sequence, and around \$20 million to store, download, transfer and process the data (using high-performance computing and a mix of cloud platforms).

We reach this sum using our estimate of average genome sizes for plants and animals — 2.5 and 1.5 gigabases, respectively — and because the average cost per species per gigabase is \$4,200 (taking into account the price differences between North America and Africa for consumables, shipments and other overheads). We estimate the costs of sample collection, including permits, consultations and workshops, at \$41 million. Lastly, using the Newton International Fellowship as a benchmark, AfricaBP’s early-career research fellowships will cost roughly \$90 million over a 10-year period.

## Species sidelined

Thousands of African species have been ignored by the global genomics community. Only 20 of the 798 plant genomes sequenced globally over the past 20 years are native to Africa<sup>1</sup>, for example. Yet sub-Saharan Africa alone, which is home to at least 45,000 plant species<sup>3</sup>, is the second-largest contributor to global plant diversity after South America. Last year, researchers reported that 60% of these species are endemic, and that many could have potential applications in agriculture or drug development<sup>4</sup>. Evidence suggests, for instance, that African ginger (*Siphonochilus aethiopicus*) could be used to treat asthma and influenza, among other conditions<sup>5,6</sup>.

Most of the genomics and bioinformatics expertise that does exist across Africa, including the sequencing facilities, is concentrated in private and non-governmental organizations, such as Inqaba Biotechnical Industries in Pretoria, South Africa, and Redeemer’s University in Nigeria. This means that, although the national research institutes are given the responsibility of setting the country’s scientific agenda, the tools needed to actually improve public health, agriculture and conservation are outside their control<sup>7</sup>.

AfricaBP will focus on endemic African species that have economic, scientific and cultural significance for African communities.

Sustained government investment in genomics — including the creation of permanent university positions — will help to ensure that African scientists who have received training through African-coordinated genomics projects stay in Africa.

National and regional expansion of tissue-sample collection, taxonomic identification, biobanking of samples and cataloguing of metadata will make it much easier for researchers to monitor species — and ultimately to protect them. Species discovered as a result of the genomics project could be added to the CBD 2030 targets.

Lastly, if the African Union Commission includes AfricaBP in the suite of schemes it is currently backing, the project could enable the commission to achieve at least three of the development goals encapsulated in the African Union Agenda 2063: The Africa We Want. These are: the use of modern techniques and technology to increase agricultural productivity sustainably; the sustainable use of ocean resources to drive economic growth; and the development of environmentally sustainable and climate-resilient economies. (Agenda 2063 is the blueprint for the continent's transformation into a global powerhouse, as laid out by leaders of the 55 African Union member states in 2013.)

## Key priorities

AfricaBP will bring together national and regional institutions, countries and corporations, including already recognized genomics infrastructures, such as the National Institute for Biomedical Research in Kinshasa in the Democratic Republic of the Congo. The project has three main goals.

**Improve food systems.** The first goal is to provide a resource that enables plant and animal breeders to use various approaches (from conventional breeding to gene editing) to build resilient and sustainable food systems. A 2021 genome analysis<sup>8</sup> of 245 Ethiopian indigenous chickens, for instance, revealed the genetic basis of various adaptations that enable the chickens to tolerate harsh environmental conditions (from cold temperatures to water scarcity) — crucial information for poultry producers worldwide. To help achieve this goal, AfricaBP will partner with the African Plant Breeding Academy and the African Animal Breeding Network, both of which were established in the past decade to improve African breeders' training and research practices.

**Improve conservation.** The second goal is to make it easier for researchers to identify species and populations that are at risk of extinction, and to design and implement effective conservation strategies. A 2020 study<sup>9</sup> on the genetic structure of African savannah elephant populations, for example, revealed that the long-term survival of the elephants requires establishing at least 14 wildlife corridors between 16 of the protected areas in Tanzania.

Similarly, a genome study<sup>10</sup> of 13 individuals representing 2 subspecies of eastern gorilla showed that inbreeding has led to the purging of severely harmful recessive mutations from one of the subspecies (*Gorilla beringei beringei*, or mountain gorillas). The accumulation of such damaging mutations in eastern gorillas over the past 100,000 years has reduced their resilience to environmental change and pathogen evolution.



A technician checks cassava plants in a research laboratory near Abidjan, Côte d'Ivoire. Credit: Sia Kambou/AFP via Getty

**Improve sharing of data and benefits.** The third goal is to kick-start a process in which existing multilateral agreements around data sharing are

improved and harmonized across the continent — to ensure that the benefits derived from genetic resources are shared equitably across Africa.

In 2010, nations adopted the Nagoya Protocol on Access and Benefits Sharing to ensure that the benefits arising from the use of biological resources are shared fairly. Certainly, any benefit derived from the genetic resources obtained through AfricaBP should be shared by the people of Africa — whether it be a superior strain of drought-resistant sugar beet (*Beta macrocarpa* Guss) or a new drug derived from the rooibos plant (*Aspalathus linearis*).

As written, however, the Nagoya Protocol has gaps when it comes to Africa. It fails to take into account the customs and practices of the diverse ethnic groups across the continent. These might not be documented or written into law, but have shaped how people interact with certain plants or animals for hundreds — sometimes thousands — of years. In West Africa, for example, some communities forbid the cutting down or harming of iroko trees, which are thought to have supernatural powers.

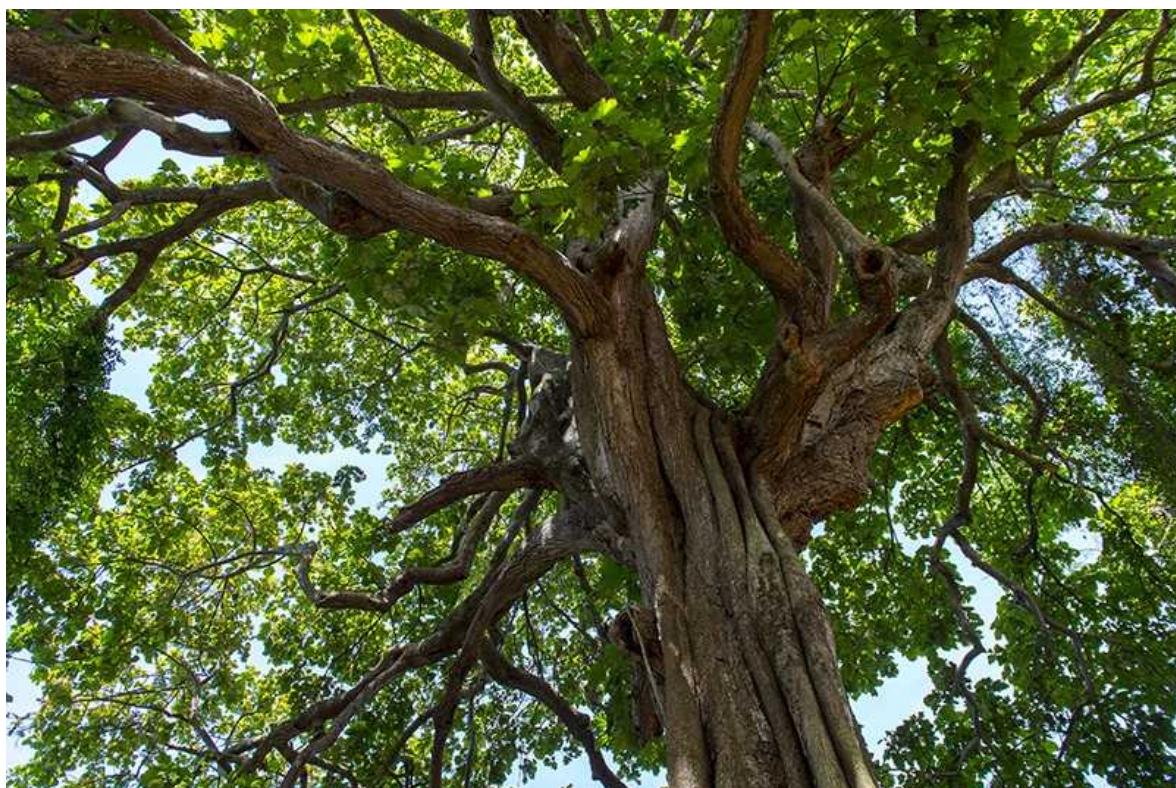
There are also inconsistencies in how the Nagoya Protocol is applied in different countries. The African Union guidelines for the implementation of the Nagoya Protocol in Africa states that those countries that are not parties to the Nagoya Protocol should be refused access to the genetic resources of other African member states. But only some countries follow this; South Africa grants non-parties access to the nation's genetic resources, whereas Ethiopia does not.

Likewise, not all countries require researchers wanting to extract genetic resources to consult community protocols. These include the rules and standards around the handling of biological specimens — as laid out by communities under the guidance of the custodians of customary laws (local chiefs and community heads). These custodians, in turn, work closely with state and national governments; sometimes, community protocols will refer to state, national or international laws. In Benin, for example, such protocols state that researchers cannot enter Gbévozoun forest or take any specimens from it because it houses the deity Gbévo, which protects the community.

Ultimately, it is the responsibility of the African Union Commission to improve and harmonize the treaties and guidelines around data and benefit sharing. Doing this would make it easier for AfricaBP researchers to obtain sampling permits, in accordance with the Nagoya Protocol and material transfer agreements (the legal documents required to send biological materials from one organization to another, or from one country to another).

But AfricaBP will enable the African Union, the CBD and other African agencies, such as the African Academy of Sciences, to integrate genomic information into their policymaking around biological diversity across Africa. This in itself will raise awareness about the Nagoya Protocol, and so encourage greater harmonization in its use.

Furthermore, the 109 scientists championing AfricaBP will coordinate with the African Group of Negotiators on Biodiversity (researchers, policymakers and other stakeholders who represent the continent in CBD negotiations) to ensure that sequencing information is specifically included in the post-2020 global biodiversity framework.



An iroko tree in Benin. Some West African communities forbid the cutting down of these trees, which locally are thought to have supernatural powers. Credit: Wolfgang Kaehler

Currently, the Nagoya Protocol specifies that ‘biological samples’ can be exchanged for scientific training or technology transfer. The inclusion of sequencing information would mean that early-career researchers who are members of an Indigenous community, such as the Amhara people in Ethiopia, could negotiate to receive training in genome sequencing and analysis if researchers from South Africa, say, wanted to collect tissue samples from their country.

Lastly, everyone involved in the AfricaBP project — now and over the next decade — will engage local chiefs and other custodians of traditional knowledge in the project from the outset. One way for researchers to engage with local communities or Indigenous peoples is through monthly meetings with government officials involved in Africa’s Access and Benefit Sharing National Focal Points. These individuals are specifically tasked with guiding compliance between the producers of biological resources, such as the Bedouin community in Egypt, and the users of those resources, such as researchers at the Pasteur Institute of Tunis in Tunisia. Another way this could be achieved is through AfricaBP ethics committees surveying thousands of people in a particular community — such as through town-hall meetings, electronic messages or telecommunications.

## Making it happen

Since 2009, \$22 million has been spent on building bioinformatics capacity across Africa through the Pan African Bioinformatics Network for H3Africa (H3ABioNet) project — including through training 150 researchers in core bioinformatics approaches and technologies. But around 10–15% of the trainees in this Africa-led project have relocated to North America or Europe, and there is no guarantee that they will return. What’s more, H3ABioNet funding winds down this year, and there are few permanent positions for trained bioinformatics personnel in African institutions. Because of this, up to 50% of the researchers who have received training through H3ABioNet could leave Africa.

In the case of AfricaBP, around 600 eligible early-career African researchers (those pursuing PhDs or postdocs) will be granted 3-year fellowships over the next 10 years. They will be able to work with AfricaBP's global partners<sup>11</sup>, such as the Wellcome Sanger Institute in Hinxton, UK, through exchange programmes. But they will be based mainly in national and regional AfricaBP facilities, to ensure that any skills they acquire are fed back into the continent.

Cloud-based computing and data storage will need to be coordinated to meet regional needs. Exchange programmes involving AfricaBP partners could help those regions or countries that lack resources; there are currently 87 genomic infrastructures in southern Africa, but only 8 in Central Africa<sup>7</sup>, for instance. These would be similar to the Newton International Fellowships, which enable early-career researchers from overseas to work for two years at a UK institution.

The 374 state-of-the-art Pacific Biosciences HiFi genome-sequencing machines that currently exist worldwide (as of 31 December 2021) can produce high-quality sequence data for more than 350 species per day<sup>12</sup>. But although the city of Cambridge, UK, alone has 12 of these machines, there are only 2 in the entire African continent. Building genomics capacity on the ground is a huge challenge in Africa because of the difficulty of transporting intact samples in countries that have poor transport infrastructure and hot climates, and because of Africa's expensive and low-quality Internet service.

To achieve such a massive sequencing feat, African researchers need state-of-the-art genome technologies. They also need mobile (albeit less accurate<sup>13</sup>) sequencing technologies that are less reliant on electricity and Internet connectivity, such as the Oxford Nanopore Technologies MinION machine. These are easily transportable and can be used in remote areas; they are roughly the size of a mobile phone<sup>13</sup>, whereas the Pacific Biosciences HiFi machines are about the size of a household refrigerator.

The 109 scientists spearheading AfricaBP are currently in discussion with leading institutions about the development of mobile sequencing platforms and integrated mobile laboratories. Encouragingly, portable, low-cost computing platforms, such as Raspberry Pi and eBioKit, are already being

used in Africa, for instance at Makerere University in Kampala, Uganda, in bioinformatics training programmes.

We ask all African life-science agencies to join AfricaBP. We also ask the African Union Commission and the African Academy of Sciences to provide the core funds — US\$100 million per year for the next 10 years. In our view, this investment will be dwarfed by the economic and other pay-offs that will stem from AfricaBP-enabled innovations and discoveries.

*Nature* **603**, 388–392 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00712-4>

## References

1. Marks, R. A., Hotaling, S., Frandsen, P. B. & VanBuren, R. *Nature Plants* **7**, 1571–1578 (2021).
2. Tripp, S. & Grueber, M. *The Economic Impact and Functional Applications of Human Genetics and Genomics* (American Society of Human Genetics & TEconomy Partners, 2021).
3. Linder, H. P. *Front. Ecol. Evol.* <https://doi.org/10.3389/fevo.2014.00038> (2014).
4. Ghazal, H. *et al.* *Plant J.* **107**, 21–36 (2021).
5. Street, R. A. & Prinsloo, G. *J. Chem.* **2013**, 205048 (2013).
6. Gericke, N. S. *Afr. J. Bot.* **77**, 850–856 (2011).
7. Inzaule, S. C., Tessema, S. K., Kebede, Y., Ogwell Ouma, A. E. & Nkengasong, J. N. *Lancet Infect. Dis.* **21**, E281–E289 (2021).
8. Gheyas, A. A. *et al.* *Mol. Biol. Evol.* **38**, 4268–4285 (2021).
9. Lohay, G. G., Weathers, T. C., Estes, A. B., McGrath, B. C. & Cavener, D. R. *Ecol. Evol.* **10**, 11069–11089 (2020).

10. Xue, Y. *et al.* *Science* **348**, 242–245 (2015).
11. Lewin, H. A. *et al.* *Proc. Natl Acad. Sci. USA* **119**, e2115635118 (2022).
12. Rhie, A. *et al.* *Nature* **592**, 737–746 (2021).
13. Tedersoo, L., Albertsen, M., Anslan, S. & Callahan, B. *Appl. Environ. Microbiol.* **87**, e0062621 (2021).
14. Dandara, C. *et al.* *AAS Open Res.* **2**, 19 (2019).
15. Marshall, K. *et al.* *Front. Genet.* **10**, 297 (2019).
16. Mulder, N. *et al.* *Annu. Rev. Biomed. Data Sci.* **4**, 57–81 (2021).
17. Rees, C. A. *et al.* *BMJ Glob. Health* **6**, e006982 (2021).
18. Lange, M. *et al.* *GigaScience* **10**, giab084 (2021).
19. Yakubu, A., Munung, N. S. & De Vries, J. *AMA J. Ethics* **22**, E156–E163 (2020).

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00712-4>

- CORRESPONDENCE
- 08 March 2022

# Ukraine: thousands sign plea for scientific sanctions against Russia

- [Dmytro Chumachenko](#) ✉

I am one of the initiators of an open letter from Ukraine's scientific community (see <https://doi.org/hjw9>) that calls for wide-ranging academic sanctions in response to Russia's shameful military attack on Ukraine on 24 February (*Nature* <https://doi.org/hjxc>; 2022). So far, more than 6,000 scientists from Ukraine and worldwide have signed it (see [go.nature.com/3hoh](http://go.nature.com/3hoh)).

Such sanctions would effectively cut off Russia's access to new technologies, scientific research and information support. Examples include blocking access to all scientometric databases and international grant programmes, and ceasing to index scientific publications from Russia.

We beg for strong support from the international research community at this terrible time in Ukraine (*Nature* <https://doi.org/hjxb>; 2022). Only with unity shall we be able to survive this terrible war. To sign the petition, go to [go.nature.com/3cmwp](http://go.nature.com/3cmwp).

*Nature* **603**, 393 (2022)

doi: <https://doi.org/10.1038/d41586-022-00695-2>

- CORRESPONDENCE
- 15 March 2022

# Secure molybdenum isotope supplies for diagnostics

- [Antonino Pietropaolo](#)<sup>0</sup>,
- [Marco Capogni](#)<sup>1</sup> &
- [Lina Quintieri](#)<sup>2</sup>

Molybdenum-99 is the precursor of the short-lived radioactive isotope technetium-99m, which is widely used in nuclear medicine for imaging diagnostics. Currently in short supply worldwide, most <sup>99</sup>Mo — a by-product of nuclear fission — is produced in four reactors, located in Belgium, the Netherlands, Australia and Poland. It is usually transported to hospitals by air, so deliveries have been delayed during the COVID-19 pandemic. Local, resilient and unbreakable supply chains are urgently needed.

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

[Buy](#)

All prices are NET prices.

### **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

*Nature* **603**, 393 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00705-3>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00705-3>

- CORRESPONDENCE
- 15 March 2022

# Predatory journals and conferences — three drivers

- [Tracey Elliott ORCID: http://orcid.org/0000-0001-9282-0855](#)<sup>0</sup> &
- [Teresa M. de la Puente ORCID: http://orcid.org/0000-0002-6138-2079](#)<sup>1</sup>

The InterAcademy Partnership (IAP), a global network of more than 140 scientific academies, has just issued a report on predatory academic journals and conferences (see [go.nature.com/3mdmshb](#)). Instances of such malpractice are rising at an alarming rate and require concerted action by all stakeholders — including universities and research funders — to avoid compromising the integrity of the global research enterprise.

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

[Buy](#)

All prices are NET prices.

### **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

*Nature* **603**, 393 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00704-4>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00704-4>

- CORRESPONDENCE
- 15 March 2022

# Forest protection: invest in professionals and their careers

- [Douglas Sheil](#) ORCID: <http://orcid.org/0000-0002-1166-6591> <sup>0</sup> &
- [J. Doland Nichols](#) <sup>1</sup>

The protection and restoration of forests has major implications for the world's climate, biodiversity and economic and societal development. But high-level commitments — such as those made by 141 nations at the COP26 climate summit in Glasgow, UK, last year — are being undermined by a dearth of people trained to fulfil those pledges.

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

## Subscribe

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

## Buy

All prices are NET prices.

## **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

*Nature* **603**, 393 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00706-2>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00706-2>

# Work

- **[How hybrid working took hold in science](#)** [ 11 March 2022]  
Career Feature • Two years since COVID-19 forced labs to shut down, group leaders describe how academic research has changed, perhaps forever.
- **[Campaign supernova: why I need a finely tuned antenna](#)**  
[ 14 March 2022]  
Where I Work • Julia Casanueva enhances the sensitivity of the Virgo interferometer to better detect distant stellar explosions.

- CAREER FEATURE
- 11 March 2022

# How hybrid working took hold in science

Two years since COVID-19 forced labs to shut down, group leaders describe how academic research has changed, perhaps forever.

- [Kendall Powell](#) 0

[Find a new job](#)



Núria López-Bigas has a one-to-one meeting with PhD student Hanna Kranas.Credit: Erika Lopez Arribillaga/IRB Barcelona

On 17 March 2020, Serena Nik-Zainal's laboratory shut down after two members of her group came down with COVID-19.

Announcing the closure to her team at the University of Cambridge's Medical Research Council Cancer Unit, UK, she said: "Take everything with you, take your screens and your good office chairs — there is no chance this is going to end in three weeks.

"As a PI holding lots of careers in my hands, the hardest thing was managing their expectations," says Nik-Zainal, whose priorities suddenly switched to addressing worries about well-being, lost experiments, staying connected and how to move research forward from various remote locations.

In the two years since the World Health Organization characterized COVID-19 as a pandemic, many research groups have established a 'new normal' routine that blends working from home with time in the laboratory. It's clear that many pre-pandemic work patterns are gone for good.

Indeed, academic supervisors are following the same trends as employers in other sectors. Last year, the World Economic Forum (WEF) asked 12,500 employees from 29 countries about their views on the future of work, post-pandemic (see [go.nature.com/3sdh6fj](https://go.nature.com/3sdh6fj)). Two-thirds responded that they wanted more flexibility in the amount of time they spent in the office, and 65% reported that they were more productive when they had a flexible schedule that allowed them to do some work from home. On average, respondents wanted to spend 2.5 days per week working from home once pandemic restrictions are lifted. The WEF survey and others show that the predictions that remote work would be disruptive, unproductive or lead to burnout were largely unfounded, and that workers' desire for flexibility greatly outweighs these concerns (see 'What hybrid workers want').

## What hybrid workers want

Last year, several global and US surveys asked employees across all sectors about work patterns once the pandemic is over or restrictions are lifted.

Respondents' desire for flexibility, in terms of both when and where they worked, trumped most concerns about the lack of face time with colleagues, maintaining productivity and even pay, in some cases. At the same time, the data show that happier workers would allow employers to reap cost savings and other benefits.

According to a 2021 World Economic Forum (WEF) survey of workers in 29 nations, those in China, France and Belgium want to work an average of 1.9 days a week from home, whereas those in India would prefer 3.4 days at home (see [go.nature.com/3sdh6fj](https://go.nature.com/3sdh6fj)).

The WEF also found that before the pandemic, 24% of employees globally worked mostly from home. By July 2021, the proportion working from home had risen to 39%, and a further 22% were working from another non-office location.

In a 2021 survey of 2,050 US workers, carried out by consultancy firm Global Workplace Analytics in Carlsbad, California, 90% of respondents said they were as productive — or more so — working remotely as in the office, and 84% of respondents said working remotely after the pandemic would make them happier (see [go.nature.com/3mf3dbb](https://go.nature.com/3mf3dbb)). Nearly 40% would be willing to take a 5% pay cut to work remotely at least part of the time.

The consultancy, which specializes in flexible workplace strategies, has also studied the benefits to employers and employees of higher levels of remote working. In a 2021 study on telecommuting trends, it estimated that a typical US employer could save, on average, US\$11,000 per year per employee working remotely half-time, in part because of increased productivity, lower office costs, and decreased staff turnover (see [go.nature.com/34w7hqf](https://go.nature.com/34w7hqf)). Employees working at home half the time could save between \$600 and \$6,000 per year and the equivalent of 11 full work days in commuting costs and time.

Nik-Zainal's team has embraced hybrid working. Those who need to do bench work have sorted themselves out according to rotation schedules and culture-room capacities. Those who do computer-based analysis have more flexibility, but generally work in the lab either twice a week or every

weekday during the core hours of 10 a.m. to 3 p.m.. Team meetings are held between those hours, too.

Nik-Zainal sees value on both sides of the hybrid equation: “Being in the same space is important for sparking creativity and discussing certain issues in person, but we do not need to be nine-to-five in each other’s company all the time,” she says.

Principal investigators (PIs), including those who started research groups during the pandemic, are now incorporating the best parts of pandemic flexibility into the future of research. “It’s hard to see any good when we are heading toward six million deaths,” says Mark Gerstein, a computational biologist at Yale University in New Haven, Connecticut. “But the pandemic has taught us new ways of thinking about things.”

For example, Gerstein has learnt that some group members work most efficiently at home, whereas others really need to come in to work. “I have been a little surprised that the tails of that spectrum have been so wide.”

Gerstein says that increased flexibility should also help to ease some of the thorniest problems that early-career researchers can face, such as [childcare support](#) and [the two-body problem](#) — the challenge of two partners needing to find a job in the same geographical location. “I want to be very flexible,” Gerstein says. “That’s what talented people want in their workplace.”

## People and project movers

Many lab leaders were quick to adopt online tools that could help them to organize and communicate with their teams. Cell biologists Eugenia Piddini and Rafael Carazo Salas at the University of Bristol, UK, for instance, started to use these techniques as soon they realized that home-working was inevitable, and Carazo Salas says that “many of those quick structures we put in place, we still keep to this day”. These include a Google Sheets booking system for lab spaces and devices that people can use to take data home and analyse them there.

Hybrid lab working has also changed the dynamics of groups. Gerstein’s weekly Zoom meeting with his 40-strong team can last for several hours, but

he's fine with a healthy dose of zoning out, turning cameras off and multitasking for those who don't need to engage in the main conversation. His group uses a Google Doc to draw up the agenda and the members share screens to annotate it in real time. He then saves the final document to the lab's Dropbox account.

"It is efficient and works even better than in-person meetings," says Gerstein, who plans to retain video meetings to accommodate childcare responsibilities, illness and scheduling conflicts. "Now, everyone is equal, even our collaborators in Europe or China. I don't think we'll ever go back to a large in-person lab meeting."

Likewise, Adam Steinbrenner's plant-immunology group holds a 30-person joint meeting with two other labs at University of Washington in Seattle. After someone presents their work, there is a 3-minute pause for people to post their thoughts or questions in a shared Google Doc before the meeting resumes with a question-and-answer session.

Almost all of the PIs *Nature* spoke to now schedule either weekly or monthly video calls with individual team members. Jean Fan, a computational biologist at Johns Hopkins University in Baltimore, Maryland, schedules these on Fridays. "I ask my students to teach me what they've learnt, where they might have gotten stuck and what their plan is for the next week," she says.

And theoretical physicist David Weir is thankful that they share their mentoring load with other group leaders in the computational-field-theory group at the University of Helsinki. They find their hour-long meetings with individual students easier to manage mentally when other supervisors are on hand to field questions and direct the conversation.



Adam Steinbrenner wears a GoPro camera to record the steps involved in a Western blot.Credit: Adam Steinbrenner

For similar reasons, Federica Di Nicolantonio no longer tries to squeeze her monthly one-to-one meetings into one day, instead conducting them over the course of a week. “My brain can’t manage meetings from 9 a.m. to 7 p.m., and this way I am fresher for each person,” says Di Nicolantonio, a cancer researcher at the Candiolo Cancer Institute in Turin, Italy.

Inducting and training new lab members can be especially challenging when done remotely. Rather than sitting alongside the trainees to troubleshoot coding errors, as they would have done in previous years, the mentors in Weir’s group had to share their screens over video calls. That was certainly more awkward and inefficient, Weir says, because without in-person expressions and body language, it was difficult to assess trainees’ comprehension.

Di Nicolantonio experienced similar inefficiencies when troubleshooting stalled research: “Recently, I ran into five people [in the lab] and managed to fix things with each of them in 2 minutes. That would have taken me half a day [virtually].”

In March 2020, Steinbrenner had a fledgling group that needed training in several experimental protocols. He took inspiration from Kenji López-Alt, a US chef and food writer who wears a GoPro camera on his head during YouTube demonstrations to explain the scientific basis of cooking. “I totally used his idea and did simple video editing,” he says. “Now, we have internal video files that show the million little steps of doing a Western blot.”

The beauty of such visual protocols, he says, is that they capture lab-specific details (such as which refrigerator shelf holds the correct buffer), and can be consulted when trainer and trainee are not in the lab at the same time. So far, he’s generated five visual protocols for common molecular-biology procedures, and plans to get students to make videos for horticulture techniques.

Likewise, researchers at the Allen Institute for Cell Science in Seattle use webcams set up next to tissue-culture hoods to live-stream training sessions,

and still help trainees to safely learn cell-culture techniques up close, yet remotely. The ability to see how the pipette should be held and to ask questions in real time makes training someone in another room, or another country, nearly seamless, says executive director Ru Gunawardane.

## Back to the lab

Some aspects of graduate education do require in-person training, argues Karla Neugebauer, director of graduate studies for Yale School of Medicine's Molecular Biophysics and Biochemistry PhD programme.

She's been gently but firmly advocating a return to in-person meetings when it is safe to do so, after noticing that early-career researchers needed more in-person opportunities. "The new students don't know their classmates, or the class above them or the department faculty members," Neugebauer says. Concerned that the programme's students weren't seeking enough advice from their peers, she holds get-togethers outside at her home so that the second-year and third-year cohorts can meet face-to-face, literally in opposing lines, for 5-minute introductions. "They are thirsty for it," she says.

She has, however, taken extra steps to ensure that the in-person events exceed COVID-19 safety regulations. "As a PI or director, you don't want to be the head of a superspreader event," she says. "So I put in a huge amount of effort." That includes communicating plans to the director of Yale's environmental health and safety office, and rethinking meeting spaces, capacities and logistics.

For a poster-session event, held in a large lobby area at Yale last December, numbers were limited to about 100 faculty members and graduate students. People were socially distanced when viewing posters, and the fonts were enlarged. "It was fabulous. People were having such a good time because students haven't had that kind of casual engagement to talk about their science," she says.



David Weir likes to conduct some group meetings while hiking along the edge of Holma-Saarijärvi in Nuuksio National Park in Finland. Credit: David Weir

For Neugebauer's dozen-member group, Zoom simply cannot replace her favourite mode of brainstorming: a small in-person group session, nowadays with masks. "There's a moment when I'm ready to write on the whiteboard, with lots of coloured pens," she says. It might be to outline a new manuscript or to go over someone's next experiment. "It's me seeing if I understand what my lab member means. It's also physical and joking around." Gerstein, however, prefers to do a small outdoor group meeting. "We just leave my office and walk and talk with no computer. It's fun to get outside, and that is something worth preserving," he says.

Weir says group meetings walking around Helsinki's parks and forests are good for free-flowing catch-ups. They are also good for one-to-one meetings that involve sensitive conversations not suited to a video call, for example, to address human-resources issues or a misunderstanding.

## The future of work

As the pandemic drags on, PIs are struggling to inject the organic cohesion that arises from a team working together in person. Piddini notes that some people do their experimental shifts so as to get in and out quickly, and as a result their interactions have become “transactional” and very business-like. For many PIs, the future will be dominated by finding ways to balance the desire for flexible working with the need to spur innovative thinking.

Slowdowns in experimental work can allow projects that were simmering on the back burner to flourish, leading to some unexpected discoveries. For example, a member of Nik-Zainal’s team decided to use the time he would usually have spent in the lab to dive into genomic and other data sets, and managed to verify a key genetic mutation in lab-cultured stem cells. “He wouldn’t have found it if he didn’t have that time,” she says. After experiments confirmed the mutation’s impact, the finding significantly improved a paper that the group was planning to submit.

Last August, Carazo Salas founded CellVoyant, a biotechnology company in Bristol that will use artificial intelligence and microscopy to optimize the manufacturing of human nerve, heart and other tissues for use in medicine and research. “I’ve wanted to do this for the last five years, and in one of the troughs of the pandemic waves, there was a window to meet with a lead investor,” he says. He then went on a four-continent virtual fundraising tour — something he could not have done in person with a toddler at home.

Virtual conferencing has opened up schedules in other productive ways, with academics able to teach undergraduates in Barcelona, Spain, in the morning and deliver a keynote address in New York City that afternoon with no jet lag.

“I’m not going back to that old life of flying around the world every two weeks,” says Nik-Zainal. “I literally watched my son grow about a foot and it’s been so lovely to be near my children.” She has slept better, had more time for regular exercise, and spent more time with team members working on manuscripts and poring over data. “In some ways, I get into the science more.”

When two major conferences were cancelled outright, Núria López-Bigas planned two half-day mini-conferences on Zoom with two other cancer-genomics group leaders, one in the United States and the other in the Netherlands. “We had common interests and thought it would be useful to cross ideas,” says López-Bigas, a group leader at the Institute for Research in Biomedicine in Barcelona. “It was a very good way for the people in my lab to get a chance to present their work.”

There were other unexpected benefits of virtual conferencing, too. At the 2020 and 2021 virtual meetings of the European Society of Human Genetics, Nik-Zainal noticed higher numbers of participants and better quality in the pre-recorded presentations. Virtual conferences allow parents of young children, disabled researchers, scientists who would otherwise need to travel and those with lower incomes to participate on an equal footing. As Neugebauer notes: “The person who comes up with an app that allows us to do hybrid meetings better is going to be a winner.”

Gerstein has also been rethinking his computational group’s workspace. “Do we want that same traditional look where people come in every day and sit at desks?” he asks. “I’m sceptical — no one wants to be in open-plan cubes.”

Instead, he sees his lab of the future as being one in which, ideally, researchers have their own office and can close the door when they need to think, code or write. There also needs to be a room big enough for three or four people, to host meetings or conference calls. Hybrid working could mean a lot of unoccupied space on certain days. He’s considering a ‘hotelling’ option, with lab members booking larger office spaces in advance as needed, alongside everyone having a smaller dedicated workstation in the group’s shared space.

Most PIs agree that hybrid working is here to stay in some guise. The flexibility for everyone to be included online, no matter their circumstances, far outweighs the problems of having to deal with technical glitches and the awkwardness of reminding someone, “You’re muted.”

But there are challenging kinks left to work out, too. “I’ve not been as good at motivating people over Zoom,” says Piddini. She calls in-person chats

over coffee with trainees “an essential part of the scientific endeavour, because we cast off from there with a lot of momentum”. Two years into a pandemic, energy is low and people have mental fatigue. When COVID-19 restrictions allow for doing science in person, she says, “You have to do that.”

Gunawardane agrees that “finding some ways to come back together physically is really important in team science”. But, she says, “I do not think we’ll go back to the pre-pandemic ways completely.” The younger generation of scientists expects some flexibility, and this period has proved that scientists can be very productive without rigid, nine-to-five laboratory schedules. “If we truly believe that we want to be inclusive and we value different ways of doing things, then flexibility has to be part of how we move forward.”

*Nature* **603**, 537-539 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00729-9>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00729-9>

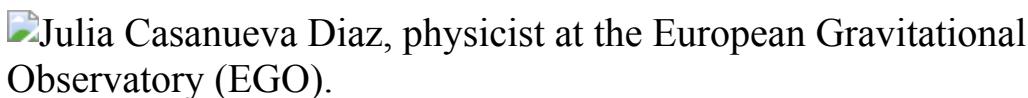
| [Section menu](#) | [Main menu](#) |

- WHERE I WORK
- 14 March 2022

# Campaign supernova: why I need a finely tuned antenna

Julia Casanueva enhances the sensitivity of the Virgo interferometer to better detect distant stellar explosions.

- [Virginia Gewin](#) <sup>0</sup>



Julia Casanueva is a researcher at the Virgo interferometer, outside Pisa, Italy. Credit: Enrico Sacchetti for *Nature*

Throughout my graduate studies, I enjoyed experimental physics and working on the instrumentation much more than doing data analysis. In 2014, I jumped at the chance to do my PhD at the European Space Agency's interferometric gravitational-wave antenna, known as Virgo, located outside Pisa, Italy. I arrived at the facility in November 2015.

The interferometer is essentially a series of 40-kilogram mirrors, each 35 centimetres in diameter, suspended throughout a pair of 3-kilometre vacuum-sealed tunnels that form an L shape. Twin laser beams travel along these arms, and multiple reflections amplify the detector's sensitivity to tiny variations in the distances between mirrors, caused by gravitational waves. When I started my PhD, it was one of only three such facilities in the world, and these waves had not yet been detected.

In September 2015, gravitational waves were observed by both Virgo and the Laser Interferometer Gravitational-Wave Observatory (LIGO) in the

United States, but this news had to be kept secret until all the measurements had been thoroughly confirmed and replicated. There was talk of little else here until the announcement was finally made in February 2016.

With my colleagues, I'm trying to improve Virgo's sensitivity further. Last July, we added a seventh mirror to increase the signal-to-noise ratio. If all goes well, this and other upgrades could double the maximum distance over which we can detect certain events in space.

In this image, I'm working on a piece of equipment that's helping to stabilize four of the existing mirrors while we finish testing the latest. My work is part of a huge collective effort involving four interferometers around the world. I might be alone in the picture, but my life is far from solitary.

It's always exciting to achieve new milestones, and I hope one day to detect a supernova, or perhaps something we don't expect: that would be even cooler.

*Nature* **603**, 542 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00730-2>

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00730-2>

# Research

- **[Savannahs store carbon despite frequent fires](#)** [ 16 March 2022]  
News & Views • An analysis of carbon stored in the plants and soil of an African savannah suggests that atmospheric carbon dioxide concentrations — and thus global warming — might be less affected by frequent fires than we thought.
- **[Methane might be made by all living organisms](#)** [ 09 March 2022]  
News & Views • It is textbook knowledge that some bacteria can generate methane enzymatically. A study now provides evidence that an alternative, non-enzymatic mode of methane production could occur in all metabolically active cells.
- **[Superfluid confines exotic atoms without disrupting precision measurements](#)** [ 16 March 2022]  
News & Views • High-precision measurements of exotic atoms containing antimatter are shown to be equally precise when the atoms are immersed in superfluid helium. Such immersion could be used for other atoms in studies of dark matter.
- **[AI predicts the effectiveness and evolution of gene promoter sequences](#)** [ 09 March 2022]  
News & Views • A long-standing goal of biology is the ability to predict gene expression from DNA sequence. A type of artificial intelligence known as a neural network, combined with high-throughput experiments, now brings this goal a step closer.
- **[The land-to-ocean loops of the global carbon cycle](#)** [ 16 March 2022]  
Perspective • An assessment of the land-to-ocean cycling of carbon through inland waters, estuaries, tidal wetlands and continental shelf waters provides a perspective on the global carbon cycle and identifies key knowledge gaps.
- **[High-resolution laser resonances of antiprotonic helium in superfluid 4He](#)** [ 16 March 2022]  
Article • The spectral lines of antiprotonic helium atoms are shown to retain their sub-gigahertz linewidth upon submersion in a bath of superfluid helium, enabling the hyperfine structure to be resolved.
- **[Unbiasing fermionic quantum Monte Carlo with a quantum computer](#)** [ 16 March 2022]

Article • A hybrid quantum-classical algorithm for solving many-electron problems is developed, enabling the simulation, with the aid of 16 qubits on a quantum processor, of chemical systems with up to 120 orbitals.

- **Steady Floquet–Andreev states in graphene Josephson junctions** [ 16 March 2022]

Article • Using continuous microwave application without substantial heating, Floquet–Andreev states in graphene Josephson junctions are realized, and their energy spectra are measured directly by superconducting tunnelling spectroscopy.

- **Wind dispersal of battery-free wireless devices** [ 16 March 2022]

Article • A dandelion-inspired wireless solar-powered sensing device weighing 30 milligrams that transmits data through radio backscatter achieves dispersal over a wide area by travelling on the breeze, and successfully lands upright.

- **Flat-surface-assisted and self-regulated oxidation resistance of Cu(111)** [ 16 March 2022]

Article • The fabrication of copper thin films with ultraflat surfaces and only occasional mono-atomic steps, which show semi-permanent resistance to oxidation over long periods, is reported and the mechanism explained using first-principles calculations.

- **A kinase-cGAS cascade to synthesize a therapeutic STING activator** [ 16 March 2022]

Article • Leveraging enzymatic selectivity, a single reaction stream provides a single diastereomer of the cyclic dinucleotide MK-1454, a promising immune-oncology drug candidate, without the use of protecting groups or chiral auxiliaries.

- **Limited increases in savanna carbon stocks over decades of fire suppression** [ 16 March 2022]

Article • A direct estimate is provided of the whole-ecosystem carbon response to fire suppression in a mesic African savanna, showing limited increase in carbon storage despite a large increase in tree cover.

- **New land-use-change emissions indicate a declining CO<sub>2</sub> airborne fraction** [ 16 March 2022]

Article • By generating a land use and land cover change emissions dataset using visibility data from two key deforestation regions, analysis of the data suggests a decrease in the CO<sub>2</sub> airborne fraction since 1959.

- **The evolution, evolvability and engineering of gene regulatory DNA** [ 09 March 2022]

Article • A framework for studying and engineering gene regulatory DNA sequences, based on deep neural sequence-to-expression models trained on large-scale libraries of random DNA, provides insight into the evolution, evolvability and fitness landscapes of regulatory DNA.

- **Cortical processing of flexible and context-dependent sensorimotor sequences** [ 09 March 2022]  
Article • A sequence licking task reveals a set of key cortical regions for the coding of movement sequences in mice.
- **FSH blockade improves cognition in mice with Alzheimer's disease** [ 02 March 2022]  
Article • Follicle-stimulating hormone acts directly on hippocampal and cortical neurons to accelerate amyloid- $\beta$  and Tau deposition and impair cognition in mice displaying features of Alzheimer's disease.
- **A non-canonical tricarboxylic acid cycle underlies cellular identity** [ 09 March 2022]  
Article • A non-canonical tricarboxylic acid cycle is required for changes in cell state.
- **Methane formation driven by reactive oxygen species across all living organisms** [ 09 March 2022]  
Article • Methane formation by a ROS-mediated process is linked to metabolic activity and is identified as a conserved feature across living systems.
- **T cell responses to SARS-CoV-2 spike cross-recognize Omicron** [ 31 January 2022]  
Article • T cell responses to spike protein from the SARS-CoV-2 Omicron variant (B.1.1.529) are broadly similar to the responses to ancestral, Beta (B.1.351) and Delta (B.1.617.2) spike protein in vaccinated, infected and unvaccinated individuals.
- **Vaccines elicit highly conserved cellular immunity to SARS-CoV-2 Omicron** [ 31 January 2022]  
Article • Current vaccines induce broadly cross-reactive cellular immunity against SARS-CoV-2 variants, including Omicron, and provide protection against severe disease despite a substantially reduced neutralizing antibody response.
- **Establishment of fetomaternal tolerance through glycan-mediated B cell suppression** [ 02 March 2022]  
Article • Pathways of glycan-mediated B cell suppression during pregnancy are important for promoting fetomaternal tolerance.
- **Ribosome collisions induce mRNA cleavage and ribosome rescue in bacteria** [ 09 March 2022]  
Article • In bacteria, a newly identified endonuclease activated by ribosome collisions truncates mRNA to trigger rescue of stalled ribosomes.

- **Bacterial ribosome collision sensing by a MutS DNA repair ATPase parologue** [ 09 March 2022]  
Article • Bacterial MutS2, a parologue of the DNA mismatch-repair protein MutS, is found to bind collided ribosomes and function in translational quality control.
- **Discovery of a Ni<sup>2+</sup>-dependent guanidine hydrolase in bacteria** [ 09 March 2022]  
Article • A bacterial enzyme is characterized and demonstrated to have Ni<sup>2+</sup>-dependent activity and high specificity for free guanidine enabling the bacteria to use guanidine as the sole nitrogen source for growth.
- **Overcoming universal restrictions on metal selectivity by protein design** [ 02 March 2022]  
Article • An alternative approach to metalloprotein design shows that it is possible to overcome the restrictions of the Irving–Williams series and achieve both flexibility and specificity in the binding of metal ions.
- **Structure determination of high-energy states in a dynamic protein ensemble** [ 02 March 2022]  
Article • Combining NMR spectroscopy-derived pseudocontact shifts (PCSSs) with Carr–Purcell–Meiboom–Gill (CPMG) relaxation dispersion enables protein structure determination of lowly populated high-energy states that are essential for macromolecular function.
- **The size of the land carbon sink in China** [ 16 March 2022]  
Matters Arising •
- **Reply to: The size of the land carbon sink in China** [ 16 March 2022]  
Matters Arising •
- **On the role of atmospheric model transport uncertainty in estimating the Chinese land carbon sink** [ 16 March 2022]  
Matters Arising •
- **Reply to: On the role of atmospheric model transport uncertainty in estimating the Chinese land carbon sink** [ 16 March 2022]  
Matters Arising •

- NEWS AND VIEWS
- 16 March 2022

# Savannahs store carbon despite frequent fires

An analysis of carbon stored in the plants and soil of an African savannah suggests that atmospheric carbon dioxide concentrations — and thus global warming — might be less affected by frequent fires than we thought.

- [Niall P. Hanan](#) <sup>0</sup> &
- [Anthony M. Swemmer](#) <sup>1</sup>

Savannahs burn more frequently than any other biome, and tropical savannahs alone account for 62% of the carbon dioxide emitted from fires globally<sup>1</sup>. Strategies involving fire suppression<sup>2</sup> or the planting of trees<sup>3</sup> in savannahs have therefore been proposed as a means of reducing CO<sub>2</sub> emissions and increasing carbon sequestration, thus potentially contributing to the mitigation of global climate change. But it remains unclear whether these measures would make a substantial difference to the accumulation of CO<sub>2</sub> in the atmosphere. [Writing Nature](#), Zhou *et al.*<sup>4</sup> analyse a long-term fire experiment in Kruger National Park, South Africa, and reveal that the total amount of carbon stored in the ecosystem increases more slowly than expected in the absence of fire — challenging our assumptions about how fire affects carbon storage in savannahs.

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

[Buy](#)

All prices are NET prices.

### **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

*Nature* **603**, 395-396 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00689-0>

# References

1. van der Werf, G. R. *et al.* *Earth Syst. Sci. Data* **9**, 697–720 (2017).
2. Russell-Smith, J. *et al.* *Clim. Change* **140**, 47–61 (2017).
3. Bastin, J.-F. *et al.* *Science* **365**, 76–79 (2019).
4. Zhou, Y. *et al.* *Nature* **603**, 445–449 (2022).
5. Pausas, J. G. & Bond, W. J. *Trends Ecol. Evol.* **35**, 767–775 (2020).
6. Hanan, N. P., Sea, W. B., Dangelmayr, G. & Govender, N. *Am. Nat.* **171**, 851–856 (2008).
7. Higgins, S. I. *et al.* *Ecology* **88**, 1119–1125 (2007).
8. Veldman, J. W. *et al.* *BioScience* **65**, 1011–1018 (2015).
9. Bayen, P., Lykke, A. M. & Thiombiano, A. *J. For. Res.* **27**, 313–320 (2016).
10. Govender, N., Trollope, W. S. W. & Van Wilgen, B. W. *J. Appl. Ecol.* **43**, 748–758 (2006).

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00689-0>

- NEWS AND VIEWS
- 09 March 2022

# Methane might be made by all living organisms

It is textbook knowledge that some bacteria can generate methane enzymatically. A study now provides evidence that an alternative, non-enzymatic mode of methane production could occur in all metabolically active cells.

- [Chang Liu](#)<sup>0</sup> &
- [Jingyao Zhang](#)<sup>1</sup>

Methane is mainly produced enzymatically by microorganisms called methanogenic archaea, in strictly oxygen-free (anoxic) conditions. In the past decade, it has become apparent that some oxygen-dependent species, including certain plants and fungi, can also produce methane<sup>1</sup>. However, the mechanisms that underlie this alternative mode of methane generation have been unclear. [Writing in Nature](#), Ernst *et al.*<sup>2</sup> reveal a process of methane production, driven by reactive oxygen species (ROS), that could occur in all living organisms, regardless of whether or not they exist in an anoxic environment. This process requires no specific enzymes — just ROS, free iron and suitable donors of methyl groups.

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

[Buy](#)

All prices are NET prices.

**Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

*Nature* **603**, 396-397 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00206-3>

**References**

1. Boros, M. & Keppler, F. *Front. Physiol.* **10**, 1244 (2019).
2. Ernst, L. *et al.* *Nature* **603**, 482–487 (2022).
3. Althoff, F. *et al.* *Nature Commun.* **5**, 4205 (2014).
4. Fenton, H. J. H. *J. Chem. Soc. Trans.* **65**, 899–910 (1894).
5. Lee, Y., Lee, C. & Yoon, J. *Wat. Res.* **38**, 2579–2588 (2004).
6. Giorgio, M., Trinei, M., Migliaccio, E. & Pelicci, P. G. *Nature Rev. Mol. Cell Biol.* **8**, 722–728 (2007).
7. Dunbar, K. L., Scharf, D. H., Litomska, A. & Hertweck, C. *Chem. Rev.* **117**, 5521–5577 (2017).
8. Muckenthaler, M. U., Rivella, S., Hentze, M. W. & Galy, B. *Cell* **168**, 344–361 (2017).
9. Sies, H. & Jones, D. P. *Nature Rev. Mol. Cell Biol.* **21**, 363–383 (2020).
10. Boros, M. *et al.* *Crit. Care Med.* **40**, 1269–1278 (2012).

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00206-3>

- NEWS AND VIEWS
- 16 March 2022

# Superfluid confines exotic atoms without disrupting precision measurements

High-precision measurements of exotic atoms containing antimatter are shown to be equally precise when the atoms are immersed in superfluid helium. Such immersion could be used for other atoms in studies of dark matter.

- [Yukari Matsuo](#) <sup>0</sup>

The structure and composition of matter can be revealed by examining how it interacts with light. Atoms absorb or emit a unique range of light frequencies appearing as lines on a spectrum, which can be measured with a technique known as laser spectroscopy. For isolated atoms, these lines are extremely narrow, offering a remarkable precision of up to 19 significant figures<sup>1</sup>. Confining atoms in fluid can make measurements more efficient by increasing the density of atoms, but it can also broaden and shift the spectral lines. [Writing in Nature](#), Sótér *et al.*<sup>2</sup> report a very narrow spectral line for an exotic form of helium engulfed in liquid helium, which might make such confinement useful for studies of the fundamental physics governing the subatomic Universe.

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

[Buy](#)

All prices are NET prices.

### **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

*Nature* **603**, 398-399 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00688-1>

# References

1. Bothwell, T. *et al.* *Metrologia* **56**, 065004 (2019).
2. Sótér, A. *et al.* *Nature* **603**, 411–415 (2022).
3. Hori, M. *et al.* *Science* **354**, 610–614 (2016).
4. Ulmer, S. *et al.* *Nature* **524**, 196–199 (2015).
5. Hori, M., Aghai-Khozani, H., Sótér, A., Dax, A. & Barna, D. *Nature* **581**, 37–41 (2020).
6. Hori, M. *et al.* *Phys. Rev. Lett.* **87**, 093401 (2001).
7. Toennies, J. P. & Vilesov. A. F. *Annu. Rev. Phys. Chem.* **49**, 1–41 (1998).
8. Moroshkin, P., Borel, A. & Kono, K. *Phys. Rev. B* **97**, 094504 (2018).
9. Cuoco, A., Krämer, M. & Korsmeier, M. *Phys. Rev. Lett.* **118**, 191102 (2017).

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00688-1>

- NEWS AND VIEWS
- 09 March 2022

# AI predicts the effectiveness and evolution of gene promoter sequences

A long-standing goal of biology is the ability to predict gene expression from DNA sequence. A type of artificial intelligence known as a neural network, combined with high-throughput experiments, now brings this goal a step closer.

- [Andreas Wagner](#) 0

Gene expression affects every aspect of life, from the survival of bacteria in specific environments to the anatomy and physiology of the human body. The ability to accurately predict how strongly a gene is expressed on the basis of the DNA sequences that regulate such expression would transform how researchers study biology. But the biochemical machinery that regulates gene expression is tremendously complex, and this goal has eluded biologists' best efforts for more than 50 years. [Writing in Nature](#), Vaishnav *et al.*<sup>1</sup> take advantage of two key technologies to produce a successful 'oracle' for gene expression in the yeast *Saccharomyces cerevisiae*.

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

[Buy](#)

All prices are NET prices.

### **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

*Nature* **603**, 399-400 (2022)

*doi:* <https://doi.org/10.1038/d41586-022-00384-0>

# References

1. Vaishnav, E. D. *et al.* *Nature* **603**, 455–463 (2022).
2. de Boer, C. G. *et al.* *Nature Biotechnol.* **38**, 56–65 (2020).
3. Jumper, J. *et al.* *Nature* **596**, 583–589 (2021).
4. Zhou, J. & Troyanskaya, O. G. *Nature Methods* **12**, 931–934 (2015).
5. Alipanahi, B., Delong, A., Weirauch, M. T. & Frey, B. J. *Nature Biotechnol.* **33**, 831–838 (2015).

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-022-00384-0>

| [Section menu](#) | [Main menu](#) |

- Perspective
- [Published: 16 March 2022](#)

# The land-to-ocean loops of the global carbon cycle

- [Pierre Regnier](#) ORCID: [orcid.org/0000-0002-4531-0868](https://orcid.org/0000-0002-4531-0868)<sup>1</sup>,
- [Laure Resplandy](#) ORCID: [orcid.org/0000-0002-1212-3943](https://orcid.org/0000-0002-1212-3943)<sup>2</sup>,
- [Raymond G. Najjar](#)<sup>3</sup> &
- [Philippe Ciais](#) ORCID: [orcid.org/0000-0001-8560-4943](https://orcid.org/0000-0001-8560-4943)<sup>4</sup>

[Nature](#) volume 603, pages 401–410 (2022)

- 1670 Accesses
- 119 Altmetric
- [Metrics details](#)

## Subjects

- [Carbon cycle](#)
- [Marine chemistry](#)

## Abstract

Carbon storage by the ocean and by the land is usually quantified separately, and does not fully take into account the land-to-ocean transport of carbon through inland waters, estuaries, tidal wetlands and continental shelf waters—the ‘land-to-ocean aquatic continuum’ (LOAC). Here we assess LOAC carbon cycling before the industrial period and perturbed by direct human interventions, including climate change. In our view of the

global carbon cycle, the traditional ‘long-range loop’, which carries carbon from terrestrial ecosystems to the open ocean through rivers, is reinforced by two ‘short-range loops’ that carry carbon from terrestrial ecosystems to inland waters and from tidal wetlands to the open ocean. Using a mass-balance approach, we find that the pre-industrial uptake of atmospheric carbon dioxide by terrestrial ecosystems transferred to the ocean and outgassed back to the atmosphere amounts to  $0.65 \pm 0.30$  petagrams of carbon per year ( $\pm 2$  sigma). Humans have accelerated the cycling of carbon between terrestrial ecosystems, inland waters and the atmosphere, and decreased the uptake of atmospheric carbon dioxide from tidal wetlands and submerged vegetation. Ignoring these changing LOAC carbon fluxes results in an overestimation of carbon storage in terrestrial ecosystems by  $0.6 \pm 0.4$  petagrams of carbon per year, and an underestimation of sedimentary and oceanic carbon storage. We identify knowledge gaps that are key to reduce uncertainties in future assessments of LOAC fluxes.

This is a preview of subscription content

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

## Subscribe

All prices are NET prices.  
VAT will be added later in the checkout.  
Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

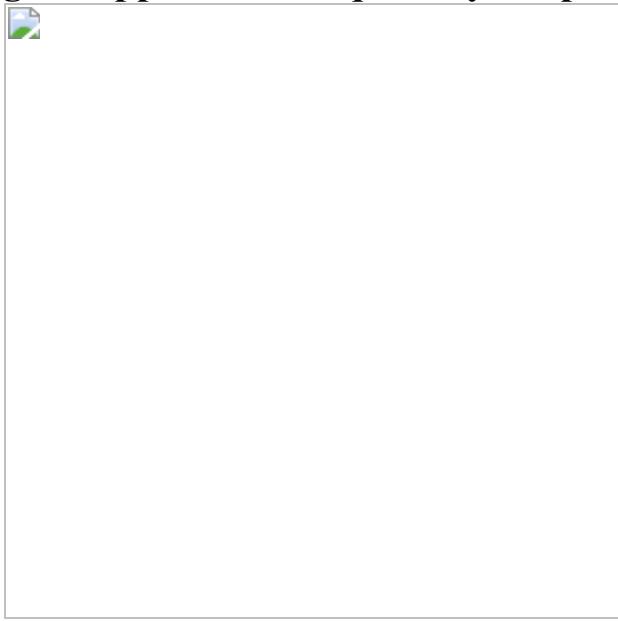
Buy

All prices are NET prices.

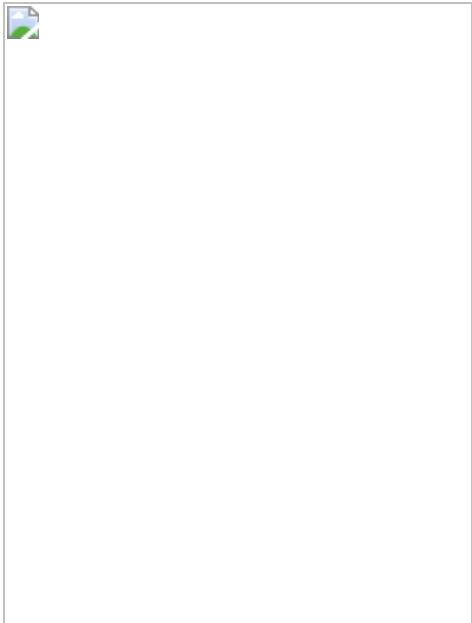
## **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

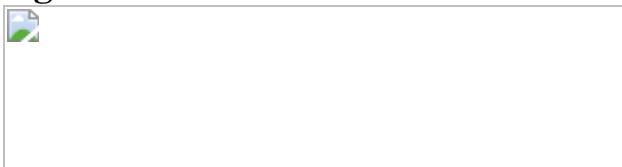
**Fig. 1: Approaches to quantify the pre-industrial carbon budget.**



**Fig. 2: The global carbon budget with LOAC fluxes.**



**Fig. 3: Bottom-up estimates of the pre-industrial open-ocean carbon budget in three latitudinal bands.**



## Data availability

Source data for Figs. 1–3, Supplementary Figs. 1, 2 are provided with the paper. All numbers and their associated uncertainties shown in Fig. 2 are synthesized in Supplementary Table 1 and described in detail in Supplementary Sections 1, 2.

## Code availability

No code was used to generate the figures in this study.

## References

1. IPCC Climate Change 2013: The Physical Science Basis (eds Stocker, T. F. et al.) (Cambridge Univ. Press 2013).

2. Ciais, P. et al. in Climate Change 2013: The Physical Science Basis (eds Stocker, T. F. et al.) Ch. 6 (IPCC, Cambridge Univ. Press, 2013).
3. Friedlingstein, P. et al. Global carbon budget 2019. *Earth Syst. Sci. Data* **11**, 1783–1838 (2019).
4. Billen, G., Lancelot, C. & Meybeck, M. in *Ocean Margin Processes in Global Change: Report of the Dahlem Workshop on Ocean Margin Processes in Global Change* (eds Mantoura, R. F. C. et al.) 19–44 (Wiley, 1991).
5. Ludwig, W., Probst, J. L. & Kempe, S. Predicting the oceanic input of organic carbon by continental erosion. *Glob. Biogeochem. Cycles* **10**, 23–41 (1996).
6. Mackenzie, F. T., De Carlo, E. H. & Lerman, A. in Treatise on Estuarine and Coastal Science (eds Middelburg, J. J. & Laane, R.) Ch. 5.10 (Elsevier, 2012).
7. Meybeck, M. Carbon, nitrogen, and phosphorus transport by world rivers. *Am. J. Sci.* **282**, 401–450 (1982).
8. Regnier, P. et al. Anthropogenic perturbation of the carbon fluxes from land to ocean. *Nat. Geosci.* **6**, 597–607 (2013). **A study quantifying the anthropogenic perturbation of the LOAC carbon fluxes, highlighting the need to include the LOAC in anthropogenic carbon budgets.**
9. Battin, T. J. et al. The boundless carbon cycle. *Nat. Geosci.* **2**, 598–600 (2009).
10. Borges, A. V., Dellile, B. & Frankignoulle, M. Budgeting sinks and sources of CO<sub>2</sub> in the coastal ocean: diversity of ecosystems counts. *Geophys. Res. Lett.* **32**, L14601 (2005).
11. Cai, W.-J., Dai, M. & Wang, Y. Air–sea exchange of carbon dioxide in ocean margins: a province-based synthesis. *Geophys. Res. Lett.* **33**, L12603 (2006).

12. Cole, J. J. et al. Plumbing the global carbon cycle: integrating inland waters into the terrestrial carbon budget. *Ecosystems* **10**, 171–184 (2007). **Pioneering study highlighting the much larger inland water carbon fluxes and the need to revise the ‘river pipeline’ model.**
13. Mulholland, P. J. & Elwood, J. W. The role of lake and reservoir sediments as sinks in the perturbed global carbon cycle. *Tellus* **34**, 490–499 (1982).
14. Richey, J. E. in The Global Carbon Cycle, Integrating Humans, Climate, and the Natural World Vol. 17 (eds Field, C. B. & Raupach, M. R.) 329–340 (Island Press, 2004).
15. Tranvik, L. J. et al. Lakes and reservoirs as regulators of carbon cycling and climate. *Limnol. Oceanogr.* **54**, 2298–2314 (2009).
16. Wollast, R. & Mackenzie, F. T. in Climate and Geo-Sciences (eds Berger, A. et al.) 453–473 (Kluwer Academic Publishers, 1989).
17. Mackenzie, F. T., Andersson, A. J., Lerman, A. & Ver, L. M. in The Sea Vol. 13 (eds Robinson, A. R. & Brink, K. H.) 193–225 (Harvard Univ. Press, 2005). **A landmark study revealing the quantitative significance of the LOAC for the global carbon budget.**
18. Ciais, P. et al. Current systematic carbon-cycle observations and the need for implementing a policy-relevant carbon observing system. *Biogeosciences* **11**, 3547–3602 (2014).
19. Raymond, P. A. et al. Global carbon dioxide emissions from inland waters. *Nature* **503**, 355–359 (2013). **A spatially resolved assessment of CO<sub>2</sub> emissions from the global inland water network, revealing the very efficient carbon turnover between terrestrial and freshwater ecosystems.**
20. Bauer, J. E. et al. The changing carbon cycle of the coastal ocean. *Nature* **504**, 61–70 (2013).

21. Sarmiento, J. L. & Sundquist, E. T. Revised budget for the oceanic uptake of anthropogenic carbon dioxide. *Nature* **356**, 589–593 (1992).  
**A study quantifying the pre-industrial land-to-ocean carbon transfers and the resulting open-ocean outgassing.**
22. Amiotte-Suchet, P. & Probst, J.-L. A global model for present day atmospheric/soil CO<sub>2</sub> consumption by chemical erosion of continental rocks (GEM-CO2). *Tellus B* **47**, 273–280 (1995).
23. Jacobson, A. R., Fletcher, S. E. M., Gruber, N., Sarmiento, J. L. & Gloor, M. A joint atmosphere–ocean inversion for surface fluxes of carbon dioxide: 1. Methods and global-scale fluxes. *Glob. Biogeochem. Cycles* **21**, GB1019 (2007).
24. Resplandy, L. et al. Revision of global carbon fluxes based on a reassessment of oceanic and riverine carbon transport. *Nat. Geosci.* **11**, 504–509 (2018).  
**A recent study advocating for an upward revision of the pre-industrial riverine and oceanic carbon transports, suggesting a tighter connection between the land and ocean carbon cycles.**
25. Le Quéré, C. et al. Global carbon budget 2017. *Earth Syst. Sci. Data* **10**, 405–448 (2018a).
26. Le Quéré, C. et al. Global carbon budget 2018. *Earth Syst. Sci. Data* **10**, 2141–2194 (2018b).
27. Friedlingstein, P. et al. Global carbon budget 2020. *Earth Syst. Sci. Data* **12**, 3269–3340 (2020).
28. Galy, V., Peucker-Ehrenbrink, B. & Eglinton, T. Global carbon export from the terrestrial biosphere controlled by erosion. *Nature* **521**, 204–207 (2015).
29. Lacroix, F., Ilyina, T. & Hartmann, J. Oceanic CO<sub>2</sub> outgassing and biological production hotspots induced by pre-industrial river loads of nutrients and carbon in a global modelling approach. *Biogeosciences* **17**, 55–88 (2020).

30. Li, M. et al. The carbon flux of global rivers: a re-evaluation of amount and spatial patterns. *Ecol. Indic.* **80**, 40–51 (2017).
31. Li, M. et al. Modeling global riverine DOC flux dynamics from 1951 to 2015. *J. Adv. Model. Earth Syst.* **11**, 514–530 (2019).
32. Luijendijk, E., Gleeson, T. & Moosdorf, N. Fresh groundwater discharge insignificant for the world's oceans but important for coastal ecosystems. *Nat. Commun.* **11**, 1260 (2020). **A global, spatially resolved quantitative assessment of carbon fluxes through the subsurface, suggesting a relatively minor contribution of the fresh groundwater pathway to the land–ocean exchanges.**
33. Wagner, S. et al. Soothsaying DOM: a current perspective on the future of oceanic dissolved organic carbon. *Front. Marine Sci.* **7**, 341 (2019).
34. Duarte, C. M. Reviews and syntheses: hidden forests, the role of vegetated coastal habitats in the ocean carbon budget. *Biogeosciences* **14**, 301–310 (2017).
35. Krause-Jensen, D. & Duarte, C. M. Substantial role of macroalgae in marine carbon sequestration. *Nat. Geosci.* **9**, 737–742 (2016).
36. Laruelle, G. G. et al. Global multi-scale segmentation of continental and coastal waters from the watersheds to the continental margins. *Hydrol. Earth Syst. Sci.* **17**, 2029–2051 (2013).
37. Roobaert, A. et al. The spatiotemporal dynamics of the sources and sinks of CO<sub>2</sub> in the global coastal ocean. *Glob. Biogeochem. Cycles* **33**, 1693–1714 (2019).
38. Windham-Myers, L. et al. in Second State of the Carbon Cycle Report (SOCCR2): A Sustained Assessment Report (eds Cavallaro, N. et al.) Ch. 15 (US Global Change Research Program, 2018).
39. Bourgeois, T. et al. Coastal-ocean uptake of anthropogenic carbon. *Biogeosciences* **13**, 4167–4185 (2016). **A study quantifying the**

**anthropogenic perturbation of CO<sub>2</sub> uptake by continental shelf waters, suggesting a small pre-industrial sink as further corroborated by the recent work by ref. <sup>55</sup>.**

40. Chmura, G. L., Anisfeld, S. C., Cahoon, D. R. & Lynch, J. C. Global carbon sequestration in tidal, saline wetland soils. *Glob. Biogeochem. Cycles* **17**, 1111 (2003).
41. Duarte, C. M., Middelburg, J. J. & Caraco, N. Major role of marine vegetation on the oceanic carbon cycle. *Biogeosciences* **2**, 1–8 (2005). **Pioneering study highlighting the quantitative role of coastal vegetation in the fixation, burial and lateral carbon exports to the open ocean and their anthropogenic perturbation.**
42. LaRowe, D. E. et al. Organic carbon and microbial activity in marine sediments on a global scale throughout the quaternary. *Geochim. Cosmochim. Acta* **286**, 227–247 (2020).
43. Smith, R. W., Bianchi, T. S., Allison, M., Savage, C. & Galy, V. High rates of organic carbon burial in fjord sediments globally. *Nat. Geosci.* **8**, 450–453 (2015).
44. O’Mara, N. & Dunne, J. Hot spots of carbon and alkalinity cycling in the coastal oceans. *Sci. Rep.* **9**, 4434 (2019).
45. Ouyang, X. & Lee, S. Updated estimates of carbon accumulation rates in coastal marsh sediments. *Biogeosciences* **11**, 5057–5071 (2014).
46. Holgerson, M. A. & Raymond, P. A. Large contribution to inland water CO<sub>2</sub> and CH<sub>4</sub> emissions from very small ponds. *Nat. Geosci.* **9**, 222–226 (2016).
47. Lauerwald, R., Laruelle, G. G., Hartmann, J., Ciais, P. & Regnier, P. A. G. Spatial patterns in CO<sub>2</sub> evasion from the global river network. *Glob. Biogeochem. Cycles* **29**, 534–554 (2015).

48. Marx, A. et al. A review of CO<sub>2</sub> and associated carbon dynamics in headwater streams: a global perspective. *Rev. Geophys.* **55**, 560–585 (2017).
49. Mendonça, R. et al. Organic carbon burial in global lakes and reservoirs. *Nat. Commun.* **8**, 1694 (2017). **A study proposing a significant downward revision of the global inland water carbon burial, corroborating the model results on the anthropogenic perturbation by ref. <sup>52</sup>.**
50. Lauerwald, R., Regnier, P., Guenet, B., Friedlingstein, P. & Ciais, P. How simulations of the land carbon sink are biased by ignoring fluvial carbon transfers: a case study for the Amazon Basin. *One Earth* **3**, 226–236 (2020).
51. Lapierre, J.-F., Guillemette, F., Berggren, M. & del Giorgio, P. A. Increases in terrestrially derived carbon stimulate organic carbon processing and CO<sub>2</sub> emissions in boreal aquatic ecosystems. *Nat. Commun.* **4**, 2972 (2013).
52. Maavara, T., Lauerwald, R., Regnier, P. & Van Cappellen, P. Global perturbation of organic carbon cycling by river damming. *Nat. Commun.* **8**, 15347 (2017).
53. Andersson, A. J., MacKenzie, F. T. & Lerman, A. Coastal ocean and carbonate systems in the high CO<sub>2</sub> world of the anthropocene. *Am. J. Sci.* **305**, 875–918 (2005).
54. Hastie, A., Lauerwald, R., Ciais, P., Papa, F. & Regnier, P. Historical and future contributions of inland waters to the Congo Basin carbon balance. *Earth Syst. Dyn.* **12**, 37–62 (2021).
55. Lacroix, F., Ilyina, T., Laruelle, G. G., & Regnier, P. Reconstructing the preindustrial coastal carbon cycle through a global ocean circulation model: was the global continental shelf already both autotrophic and a CO<sub>2</sub> sink? *Glob. Biogeochem. Cycles* **35**, e2020GB006603 (2021)

56. Landschützer, P., Gruber, N., Bakker, D. C. E. & Schuster, U. Recent variability of the global ocean carbon sink. *Glob. Biogeochem. Cycles* **28**, 927–949 (2014).
57. Rödenbeck, C. et al. Global surface-ocean  $p\text{CO}_2$  and sea-air  $\text{CO}_2$  flux variability from an observation-driven ocean mixed-layer scheme. *Ocean Sci.* **9**, 193–216 (2013).
58. Chau, T. T. T., Gehlen, M. & Chevallier, F. A seamless ensemble-based reconstruction of surface ocean  $p\text{CO}_2$  and air-sea  $\text{CO}_2$  fluxes over the global coastal and open oceans. *Biogeosciences* **19**, 1087–1109 (2022).
59. DeVries, T. et al. Decadal trends in the ocean carbon sink. *Proc. Natl Acad. Sci. USA* **116**, 11646–11651 (2019).
60. Gruber, N. et al. The oceanic sink for anthropogenic  $\text{CO}_2$  from 1994 to 2007. *Science* **363**, 1193–1199 (2019).
61. Gaillardet, J., Dupré, B., Louvat, P. & Allègre, C. J. Global silicate weathering and  $\text{CO}_2$  consumption rates deduced from the chemistry of large rivers. *Chem. Geol.* **159**, 3–30 (1999).
62. Hartmann, J., Jansen, N., Dürr, H. H., Kempe, S. & Köhler, P. Global  $\text{CO}_2$  consumption by chemical weathering: what is the contribution of highly active weathering regions? *Glob. Planet. Change* **69**, 185–194 (2009).
63. Aumont, O. et al. Riverine-driven interhemispheric transport of carbon. *Glob. Biogeochem. Cycles* **15**, 393–405 (2001).
64. Cai, W. J. Estuarine and coastal ocean carbon paradox:  $\text{CO}_2$  sinks or sites of terrestrial carbon incineration? *Annu. Rev. Mar. Sci.* **3**, 123–145 (2011).
65. Maher, D. T. & Eyre, B. D. Benthic fluxes of dissolved organic carbon in three temperate Australian estuaries: implications for global

- estimates of benthic DOC fluxes. *J. Geophys. Res.* **115**, G04039 (2010).
66. Duarte, C. M., Middelburg, J. J. & Caraco, N. Major role of marine vegetation on the oceanic carbon cycle. *Biogeosciences* **2**, 1–8 (2005).
  67. MacCreadie, P. et al. The future of blue carbon science. *Nat. Commun.* **10**, 3998 (2019).
  68. McLeod, E. et al. A blueprint for blue carbon: toward an improved understanding of the role of vegetated coastal habitats in sequestering CO<sub>2</sub>. *Front. Ecol. Environ.* **9**, 552–560 (2011).
  69. Pendleton, L. et al. Estimating global “blue carbon” emissions from conversion and degradation of vegetated coastal ecosystems. *PLoS ONE* **7**, e43542 (2012).
  70. Krumhansl, K. A. et al. Global patterns of kelp forest change over the past half-century. *Proc. Natl Acad. Sci. USA* **113**, 13785–13790 (2016).
  71. Duarte, C. M., Losada, I. J., Hendriks, I. E., Mazarrasa, I. & Marbà, N. The role of coastal plant communities for climate change mitigation and adaptation. *Nat. Clim. Change* **3**, 961–968 (2013).
  72. Liu, X. et al. Simulating water residence time in the coastal ocean: a global perspective. *Geophys. Res. Lett.* **46**, 13910–13919 (2019).
  73. Dittmar, T., Hertkorn, N., Kattner, G. & Lara, R. J. Mangroves, a major source of dissolved organic carbon to the oceans. *Glob. Biogeochem. Cycles* **20**, GB1012 (2006).
  74. Barrón, C., Apostolaki, E. T. & Duarte, C. M. Dissolved organic carbon fluxes by seagrass meadows and macroalgal beds. *Front. Mar. Sci.* **1**, 42 (2014).
  75. Maher, D. T., Call, M., Santos, I. R. & Sanders, C. J. Beyond burial: lateral exchange is a significant atmospheric carbon sink in mangrove forests. *Biol. Lett.* **14**, 20180200 (2018).

76. Bogard, M. J. et al. Hydrologic export is a major component of coastal wetland carbon budgets. *Glob. Biogeochem. Cycles* **34**, e2019GB006430 (2020).
77. Frischknecht, M., Münnich, M. & Gruber, N. Origin, transformation, and fate: the three-dimensional biological pump in the California Current System. *J. Geophys. Res. Oceans* **123**, 7939–7962 (2018).
78. Ciais, P. et al. Empirical estimates of regional carbon budgets imply reduced global soil heterotrophic respiration. *Natl Sci. Rev.* **8**, nwaa145 (2021).
79. Lovelock, C. E. & Reef, R. Variable impacts of climate change on blue carbon. *One Earth* **3**, 195–211 (2020).
80. Striegl, R. G., Dornblaser, M. M., McDonald, C. P., Rover, J. R. & Stets, E. G. Carbon dioxide and methane emissions from the Yukon River system. *Glob. Biogeochem. Cycles* **26**, GB0E05 (2012).
81. Butman, D. S. et al. Aquatic carbon cycling in the conterminous United States and implications for terrestrial carbon accounting. *Proc. Natl Acad. Sci. USA* **113**, 58–63 (2016).
82. Wallin, M. B. et al. Evasion of CO<sub>2</sub> from streams—the dominant component of the carbon export through the aquatic conduit in a boreal landscape. *Glob. Change Biol.* **19**, 785–797 (2013).
83. Drake, T. W., Raymond, P. A. & Spencer, R. G. M. Terrestrial carbon inputs to inland waters: a current synthesis of estimates and uncertainty. *Limnol. Oceanogr. Lett.* **3**, 132–142 (2018).
84. Maberly, S. C., Barker, P. A., Stott, A. W. & De Ville, M. M. Catchment productivity controls CO<sub>2</sub> emissions from lakes. *Nat. Clim. Change* **3**, 391–394 (2013).
85. Borges, A. V. et al. Globally significant greenhouse-gas emissions from African inland waters. *Nat. Geosci.* **8**, 637–642 (2015).

86. Tian, H. et al. Anthropogenic and climatic influences on carbon fluxes from eastern North America to the Atlantic Ocean: a process-based modeling study. *J. Geophys. Res. Biogeosci.* **120**, 757–772 (2015). **A pioneering study representing the land-to-ocean carbon transfers in a land-surface scheme of an Earth system model.**
87. Gommet, C. A. S. et al. Spatio-temporal patterns and drivers of terrestrial dissolved organic carbon (DOC) leaching to the European river network. *Earth Syst. Dyn.* **13**, 393–418 (2022).
88. Lauerwald, R. et al. ORCHILEAK (revision 3875): a new model branch to simulate carbon transfers along the terrestrial-aquatic continuum of the Amazon Basin. *Geosci. Model Dev.* **10**, 3821–3859 (2017).
89. Ciais, P. et al. The impact of lateral carbon fluxes on the European carbon balance. *Biogeosciences* **5**, 1259–1271 (2008).
90. Luysaert, S. et al. The European land and inland water CO<sub>2</sub>, CO, CH<sub>4</sub> and N<sub>2</sub>O balance between 2001 and 2005. *Biogeosciences* **9**, 3357–3380 (2012).
91. Cavallaro, N. G. et al. in Second State of the Carbon Cycle Report (SOCCR2): A Sustained Assessment Report (eds Cavallaro, N. et al.). 1–878 (US Global Change Research Program, 2018).
92. Hastie, A. et al. CO<sub>2</sub> evasion from boreal lakes: revised estimate, drivers of spatial variability, and future projections. *Glob. Change Biol.* **24**, 711–728 (2018).
93. Aufdenkampe, A. K. et al. Riverine coupling of biogeochemical cycles between land, oceans, and atmosphere. *Front. Ecol. Environ.* **9**, 53–60 (2011).
94. Richey, J. E., Melack, J. M., Aufdenkampe, A. K., Ballester, V. M. & Hess, L. L. Outgassing from Amazonian rivers and wetlands as a large tropical source of atmospheric CO<sub>2</sub>. *Nature* **416**, 617–620 (2002).

95. Abril, G. & Borges, A. V. Ideas and perspectives: carbon leaks from flooded land: do we need to replumb the inland water active pipe? *Biogeosciences* **16**, 769–784 (2019).
96. Hastie, A., Lauerwald, R., Ciais, P. & Regnier, P. Aquatic carbon fluxes dampen the overall variation of net ecosystem productivity in the Amazon Basin: an analysis of the interannual variability in the boundless carbon cycle. *Glob. Change Biol.* **25**, 2094–2111 (2019).
97. Gómez-Gener, L. et al. Global carbon dioxide efflux from rivers enhanced by high nocturnal emissions. *Nat. Geosci.* **14**, 289–294 (2021).
98. Abril, G. et al. Technical note: large overestimation of pCO<sub>2</sub> calculated from pH and alkalinity in acidic, organic-rich freshwaters. *Biogeosciences* **12**, 67–78 (2015).
99. Golub, M., Desai, A. R., McKinley, G. A., Remucal, C. K. & Stanley, E. H. Large uncertainty in estimating pCO<sub>2</sub> from carbonate equilibria in lakes. *J. Geophys. Res. Biogeosci.* **122**, 2909–2924 (2017).
100. Heathcote, A. J., Anderson, N. J., Prairie, Y. T., Engstrom, D. R. & del Giorgio, P. A. Large increases in carbon burial in northern lakes during the Anthropocene. *Nat. Commun.* **6**, 10016 (2015).
101. Kastowski, M., Hinderer, M. & Vecsei, A. Long-term carbon burial in European lakes: analysis and estimate. *Glob. Biogeochem. Cycles* **25**, GB3019 (2011).
102. Seitzinger, S. P. et al. Global river nutrient export: a scenario analysis of past and future trends. *Glob. Biogeochem. Cycles* **24**, GB0A08 (2010).
103. Mayorga, E. et al. Global Nutrient Export from WaterSheds 2 (NEWS 2): model development and implementation. *Environ. Model. Softw.* **25**, 837–853 (2010).

104. Ren, W. et al. Century-long increasing trend and variability of dissolved organic carbon export from the Mississippi River Basin driven by natural and anthropogenic forcing. *Glob. Biogeochem. Cycles* **30**, 1288–1299 (2016).
105. Jones, J. B., Stanley, E. H. & Mulholland, P. J. Long-term decline in carbon dioxide supersaturation in rivers across the contiguous United States. *Geophys. Res. Lett.* **30**, 1495 (2003).
106. Ran, L. et al. Substantial decrease in CO<sub>2</sub> emissions from Chinese inland waters due to global change. *Nat. Commun.* **12**, 1730 (2021).
107. Park, J. H. et al. Reviews and syntheses: anthropogenic perturbations to carbon fluxes in Asian river systems—concepts, emerging trends, and research challenges. *Biogeosciences* **15**, 3049–3069 (2018).
108. Kicklighter, D. W. et al. Insights and issues with simulating terrestrial DOC loading of Arctic river networks. *Ecol. Appl.* **23**, 1817–1836 (2013).
109. Bowring, S. P. K. et al. ORCHIDEE MICT-LEAK (r5459), a global model for the production, transport, and transformation of dissolved organic carbon from Arctic permafrost regions—Part 2: model evaluation over the Lena River Basin. *Geosci. Model Dev.* **13**, 507–520 (2020).
110. Laruelle, G. G., Goossens, N., Arndt, S., Cai, W.-J. & Regnier, P. Air–water CO<sub>2</sub> evasion from US East Coast estuaries. *Biogeosciences* **14**, 2441–2468 (2017).
111. St-Laurent, P. et al. Relative impacts of global changes and regional watershed changes on the inorganic carbon balance of the Chesapeake Bay. *Biogeosciences* **17**, 3779–3796 (2020).
112. Durr, H. H. et al. Worldwide typology of nearshore coastal systems: defining the estuarine filter of river inputs to the oceans. *Estuaries Coast.* **34**, 441–458 (2011).

113. Laruelle, G. G. et al. Continental shelves as a variable but increasing global sink for atmospheric carbon dioxide. *Nat. Commun.* **9**, 454 (2018).
114. Lacroix, F., Ilyina, T., Mathis, M., Laruelle, G. G. & Regnier, P. Historical increases in land-derived nutrient inputs may alleviate effects of a changing physical climate on the oceanic carbon cycle. *Glob. Change Biol.* **27**, 5491–5513 (2021).
115. Cotovicz, L. Jr, Knoppers, B., Brandini, N., Santos, S. & Abril, G. A strong CO<sub>2</sub> sink enhanced by eutrophication in a tropical coastal embayment (Guanabara Bay, Rio de Janeiro, Brazil). *Biogeosciences* **12**, 6125–6146 (2015).

## Acknowledgements

P.R. and P.C. acknowledge funding from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement number 643052 (C-CASCADES). P.R. also received financial support from BELSPO through the project ReCAP, which is part of the Belgian research programme FedTwin and from the European Union’s Horizon 2020 research and innovation programme under grant agreements number 776810 (VERIFY) and number 101003536 (ESM2025–Earth System Models for the Future). P.C. has been co-funded by the French Agence Nationale de la Recherche (ANR) Convergence Lab Changement climatique et usage des terres (CLAND), the European Space Agency Climate Change Initiative ESA-CCI RECCAP2 project 1190 (ESRIN/ 4000123002/18/I-NB) and Observation-based system for monitoring and verification of greenhouse gases (VERIFY, grant agreement number 776810). L.R. gratefully acknowledges support from the Alfred P. Sloan Foundation Research Fellowship, the Princeton Catalysis Initiative at Princeton University and the NASA OCO-2 Science Team Grant 80NSSC18K0893. R.G.N. acknowledges support from NASA Carbon Cycle Science and Interdisciplinary Science Programs and NSF Chemical Oceanography Program. This study benefitted from discussions with W.-J. Cai, A. Coppola, P. Friedlingstein and N. Gruber.

# Author information

## Affiliations

1. Biogeochemistry and Modelling of the Earth System-BGEOSYS,  
Department of Geoscience, Environment and Society, Université Libre  
de Bruxelles, Brussels, Belgium

Pierre Regnier

2. Department of Geosciences, High Meadows Environmental Institute,  
Princeton University, Princeton, NJ, USA

Laure Resplandy

3. Department of Meteorology and Atmospheric Science, The  
Pennsylvania State University, University Park, PA, USA

Raymond G. Najjar

4. Laboratoire des Sciences du Climat et de l'Environnement – UVSQ,  
UPSAclay, Gif sur Yvette, France

Philippe Ciais

## Contributions

P.R. and P.C. initiated the design of the study that led to this paper. P.R. and L.R. directed the analysis and coordinated the conception and writing of the paper. L.R. designed all the figures. P.R. and R.G.N. co-led the synthesis of LOAC fluxes. P.C., R.G.N., L.R. and P.R. contributed to the budget analysis and the writing of the paper. P.R. and L.R. have contributed equally to this study.

## Corresponding author

Correspondence to [Pierre Regnier](#).

# Ethics declarations

## Competing interests

The authors declare no competing interests.

## Peer review

### Peer review information

*Nature* thanks David Butman, Galen McKinley and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

## Additional information

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Supplementary information

### [Supplementary Information](#)

This file contains Supplementary text, Figs. 1, 2, Tables 1–3 and References

## Rights and permissions

### [Reprints and Permissions](#)

## About this article

## Cite this article

Regnier, P., Resplandy, L., Najjar, R.G. *et al.* The land-to-ocean loops of the global carbon cycle. *Nature* **603**, 401–410 (2022).

<https://doi.org/10.1038/s41586-021-04339-9>

- Received: 22 December 2020
- Accepted: 11 December 2021
- Published: 16 March 2022
- Issue Date: 17 March 2022
- DOI: <https://doi.org/10.1038/s41586-021-04339-9>

## Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04339-9>

| [Section menu](#) | [Main menu](#) |

- Article
- Open Access
- [Published: 16 March 2022](#)

# High-resolution laser resonances of antiprotonic helium in superfluid $^4\text{He}$

- [Anna Sótér<sup>1</sup> nAff7](#),
- [Hossein Aghai-Khozani<sup>1</sup> nAff8](#),
- [Dániel Barna<sup>2,3</sup>](#),
- [Andreas Dax<sup>2</sup> nAff9](#),
- [Luca Venturelli<sup>4,5</sup>](#) &
- [Masaki Hori](#) [ORCID: orcid.org/0000-0002-4999-3141<sup>1,6</sup>](#)

*Nature* volume **603**, pages 411–415 (2022)

- 1349 Accesses
- 231 Altmetric
- [Metrics details](#)

## Subjects

- [Atomic and molecular collision processes](#)
- [Exotic atoms and molecules](#)
- [Experimental particle physics](#)

## Abstract

When atoms are placed into liquids, their optical spectral lines corresponding to the electronic transitions are greatly broadened compared to those of single, isolated atoms. This linewidth increase can often reach a factor of more than a million, obscuring spectroscopic structures and preventing high-resolution spectroscopy, even when superfluid helium, which is the most transparent, cold and chemically inert liquid, is used as the host material<sup>1,2,3,4,5,6</sup>. Here we show that when an exotic helium atom with a constituent antiproton<sup>7,8,9</sup> is embedded into superfluid helium, its visible-

wavelength spectral line retains a sub-gigahertz linewidth. An abrupt reduction in the linewidth of the antiprotonic laser resonance was observed when the liquid surrounding the atom transitioned into the superfluid phase. This resolved the hyperfine structure arising from the spin–spin interaction between the electron and antiproton with a relative spectral resolution of two parts in  $10^6$ , even though the antiprotonic helium resided in a dense matrix of normal matter atoms. The electron shell of the antiprotonic atom retains a small radius of approximately 40 picometres during the laser excitation<sup>7</sup>. This implies that other helium atoms containing antinuclei, as well as negatively charged mesons and hyperons that include strange quarks formed in superfluid helium, may be studied by laser spectroscopy with a high spectral resolution, enabling the determination of the particle masses<sup>9</sup>. The sharp spectral lines may enable the detection of cosmic-ray antiprotons<sup>10,11</sup> or searches for antideuterons<sup>12</sup> that come to rest in liquid helium targets.

[Download PDF](#)

## Main

Laser spectroscopy of antihydrogen<sup>13,14</sup> and antiprotonic helium atoms ( $(\bar{p}H)^+$ ) have recently been carried out to investigate the symmetry between matter and antimatter. These experiments are complementary to some precision measurements on the properties of single antiprotons<sup>15,16</sup>. The high precision of these experiments could only be achieved by reducing or eliminating the collisions with normal atoms that were found to annihilate the antiprotons in the exotic atoms or perturb their atomic energy levels and strongly broaden the laser resonances<sup>17,18</sup>. This required forming the antihydrogen or  $(\bar{p}H)^+$  in magnetic traps or gaseous helium targets of extremely low atomic density  $\rho < 10^{18} \text{ cm}^{-3}$ , so that sharp spectral lines of effectively single isolated atoms were resolved from which the atomic transition frequencies were precisely determined. Other exotic atoms were found to accelerate and heat up during their formation, and during collisions with other molecules<sup>19,20,21</sup> so that their X-ray spectral lines were broadened. In this work, we instead observed the surprising phenomenon wherein, in contrast to the above previous results on exotic atoms, the  $(\bar{p}H)^+$  embedded in superfluid helium (He II) showed visible-wavelength spectral lines that are narrower than those of many other implanted stable atoms made of normal matter reported so far. At the superfluid phase transition temperature where the atomic density of the liquid is largest, the antiproton laser resonance abruptly narrowed to sub-GHz linewidths corresponding to a relative spectral resolution of  $2 \times 10^{-6}$ , which is more than a factor of 10 narrower than the same  $(\bar{p}H)^+$  spectra observed in supercritical-

phase helium of lower density. This implies that other varieties of helium atoms containing antideuterons—or negatively charged mesons and hyperons<sup>22</sup> that include the strange quark which cannot be readily decelerated and cooled using synchrotrons or isolated in ion traps—may instead be stopped in He II and measured with a high spectral resolution by laser spectroscopy. The fact that the lineshapes are so sensitive to the liquid temperature and phase suggests that  $\langle\bar{p}\rangle\langle e\rangle^4\langle H\rangle$  may be used to study some condensed-matter effects in superfluid helium<sup>23,24,25,26,27</sup>.

Laser spectroscopy of various atoms implanted into bulk He II began in the 1980s with the expectation that He II, being cold and inert, would constitute an ideal host material where the distortions of the atomic spectral lines would be small compared to other liquids. Measurements quickly showed, however, that the visible-wavelength spectral lines involving the outermost valance orbitals of alkaline<sup>2,3</sup>, alkaline earth<sup>4,5</sup> and lanthanide<sup>6</sup> atoms embedded in normal liquid helium (He I) and He II are nevertheless highly shifted ( $|\Delta\nu| \approx 10^4$  GHz) and asymmetrically broadened ( $\Gamma = 10^3$ – $10^4$  GHz) by a factor greater than or equal to  $10^6$  compared to the typical natural widths of several MHz or less. Other specific resonances involving electronic excitations from the inner shells of Tm<sup>28</sup>, Eu<sup>29</sup>, Cu, Au<sup>30</sup> and Dy<sup>31,32</sup> atoms in He II showed widths of  $\Gamma = 4$ – $80$  GHz with additional wing structures that were typically located 150–300 GHz away from the main spectral line. The linewidths of some of these resonances decreased in superfluid helium<sup>29,32</sup>. The broad, complex lineshapes have been interpreted by effective models, some of which involve the formation of bubble-like defects around various impurities (particularly alkaline, alkaline earth and rare earth) atoms and molecules.

The neutral  $\langle\bar{p}\rangle\langle e\rangle^4\langle H\rangle$  studied here consists of a helium nucleus, an electron in the  $1s$  ground state, and an antiproton occupying a Rydberg state of large principal  $n$  and orbital angular momentum  $\ell \approx n - 1$  quantum numbers. The atom's longevity in liquid helium targets has been theoretically and experimentally studied<sup>33,34,35</sup>. The  $n > 41$  antiproton orbitals that extend outside the electron shell with a root-mean-square radius  $r_e \approx 40$  pm are easily destroyed in collisions with other atoms and have never been detected. By comparison, the  $n = 30$ – $40$  antiproton orbitals lie well within the electron shell (Fig. 1a) and should in principle be better protected, but numerous states were found to likewise be destroyed for atoms synthesized in gas targets of moderate density  $\rho = 10^{20}$ – $10^{21}$  cm $^{-3}$  (ref. 36) so that laser spectroscopy of antiprotonic atoms suspended in liquid targets has not been achieved so far. In this work we nevertheless detected two transitions  $(n, \ell) = (37, 35) \rightarrow (38, 34)$  and  $(39, 35) \rightarrow (38, 34)$  at the visible wavelengths  $\lambda = 726$  nm and 597 nm, respectively, that survived in He I and He II targets. The resonance parent states  $(37, 35)$  and  $(39, 35)$  have microsecond-scale lifetimes, whereas the daughter state  $(38, 34)$

has an Auger width  $\Gamma_A \approx 21$  MHz (ref. 7; Fig. 1b). As the radius of a single isolated  $(\bar{p}\{\}^4H\{e\})^+$  atom is an order of magnitude smaller (Fig. 1a) than the valance orbitals of the above-mentioned impurity atoms, and the optical transitions of the massive antiproton involve remarkably small changes in the radius  $\Delta r_e \leq 2$  pm (ref. 7) of the electron and the related  $(\bar{p}\{\}^4H\{e\})^+$  pairwise potentials<sup>17,18</sup>, we may expect lineshapes that are quantitatively different from those of other many impurity atoms.

**Fig. 1: Laser spectroscopy of  $(\bar{p}\{\}^4H\{e\})^+$  synthesized in He I and He II targets.**

figure 1

**a**, Radial distribution functions of the antiproton (red curves) and electron (blue curves) for states  $(n, \ell) = (37, 35)$  and  $(38, 34)$  of a single isolated  $(\bar{p}\{\}^4H\{e\})^+$  atom. Here  $\langle R_{n\ell}(r) \rangle$  denotes the radial component of the antiproton or electron orbital at a distance  $r$  from the helium nucleus. Laser transitions between the two states involve an unusually small ( $\Delta r_e \leq 2$  pm) change in the root-mean-square radius of the  $1s$  electronic orbital. **b**, Energy level diagram of  $(\bar{p}\{\}^4H\{e\})^+$  indicating the positions of the transitions  $(37, 35) \rightarrow (38, 34)$  and  $(39, 35) \rightarrow (38, 34)$ . **c**, Signal of the Cherenkov detector, which corresponds to the time distribution of  $(\bar{p}\{\}^4H\{e\})^+$  annihilations. The laser resonance is induced at  $t \approx 1.4$   $\mu$ s after  $(\bar{p}\{\}^4H\{e\})^+$  formation. The large number of annihilations at  $t = 0$  is not shown. **d**, Experimental layout. Antiprotons emerging from a radio frequency quadrupole (RFQ) decelerator came to rest in either a gaseous or supercritical phase helium, He I, or He II target. The produced  $(\bar{p}\{\}^4H\{e\})^+$  atoms were irradiated with a resonant laser pulse (see text).

In this work the Antiproton Decelerator of CERN provided a pulsed beam containing  $\langle N \rangle_{\bar{p}} = (2-3) \times 10^7$  antiprotons with an energy  $E_k = 5.3$  MeV and repetition rate  $f_r \approx 8$  mHz (Fig. 1d). The beam was allowed to traverse a radio frequency quadrupole (RFQ) decelerator which slowed down approximately 20% of the

antiprotons to  $E_k = 75$  keV. A dipole magnet diverted the slow antiprotons to another experiment. The remaining approximately 80% missed the longitudinal acceptance of the RFQ decelerator and emerged with an energy  $E_k = 3.2\text{--}5.3$  MeV. We constructed a beamline to transport the higher-energy antiprotons to the position of the helium target and collected spectroscopic data over several months. The  $(\bar{p}\{\{4\}\rm{H}\}\{\rm{e}\})^+$  were produced by allowing the antiprotons to come to rest in a 35-mm diameter, hermetically sealed chamber filled with He I or He II at temperatures between  $T = 1.49(3)$  K and  $4.15(2)$  K (see Methods for a discussion on the experimental uncertainty of the target temperature and pressure). A separate target filled with helium of  $T = 5.97(6)\text{--}6.4(4)$  K and maximum pressure  $p = 555(2)$  kPa, which corresponds to the supercritical phase with approximately 70% of the He I density, was also used.

The 726-nm transition with a small dipole moment  $d_m = 0.018$  a.u. (a.u., 1 atomic unit =  $8.48 \times 10^{-30}$  C m) was excited by irradiating the  $(\bar{p}\{\{4\}\rm{H}\}\{\rm{e}\})^+$  with a 50-ns-long laser pulse of linewidth  $\Gamma_{\text{las}} = 60$  MHz and fluence  $\varepsilon = 4$  mJ cm $^{-2}$  (Fig. 1b). For this we used a Ti:sapphire pulsed laser which was injection-seeded by a continuous-wave (CW) laser. The 597-nm resonance of  $d_m = 0.24$  a.u. was excited by a 50-ns-long laser pulse of  $\Gamma_{\text{las}} = 80$  MHz and  $\varepsilon = 0.12$  mJ cm $^{-2}$ , which was generated by a CW pulse-amplified dye laser<sup>9</sup>. The optical frequencies of the seed lasers were measured with a precision of better than 1 MHz using a femtosecond frequency comb. A mechanical shutter prevented the residual seed beam from entering and heating up the target during the intervals between the antiproton arrivals. The two-body antiprotonic helium ion  $(\bar{p}\{\{4\}\rm{H}\}\{\rm{e}\})^{2+}$  that remained after Auger emission was rapidly destroyed by collisions with the surrounding liquid. The resulting spike (Fig. 1c) in the flux of charged pions that emerged from the antiproton annihilations was measured by an acrylic Cherenkov detector.

Fig. 2 shows the spectral profiles of the 726-nm resonance measured at six pressures of the gaseous and supercritical helium targets, which were obtained by plotting the intensity of the induced annihilation signals as a function of the laser frequency. Each point represents data collected from 2–5 antiproton pulses. The lineshapes contain contributions from the 21-MHz natural width<sup>7</sup> of the daughter state, the 60–80-MHz laser linewidth, the hyperfine structure arising from the spin–spin interaction between the electron and antiproton<sup>37</sup>, power broadening effects, and the complicated (see below) effects of the interactions with the surrounding helium, combined with the motions of the  $(\bar{p}\{\{4\}\rm{H}\}\{\rm{e}\})^+$  atoms. The apparent linewidths were obtained by fitting the spectral profiles with four overlapping Lorentzian functions that were fixed to the relative positions of the calculated hyperfine intervals<sup>37</sup>. This simplified definition of the full width at half maximum (FWHM)

Lorentzian linewidth  $\Gamma_L$  avoided the ambiguities that may arise from a more specific lineshape model with many parameters. The total uncertainty was taken as the quadratic sum of the statistical uncertainty arising from the finite number of excited atoms and the systematic uncertainty. The latter includes contributions from the calibration and fluctuations of the target temperature and shifts  $|\Delta\nu_{\text{las}}| \leq 30\text{--}60$  MHz in the laser frequency that are due to spurious modulations that occurred during the amplification of the laser pulse. The heating effect of the laser on the target was investigated using finite-element simulations.

**Fig. 2: Resonance profiles of the transition  $(n, \ell) = (37, 35) \rightarrow (38, 34)$  synthesized in various targets.**

 figure 2

Spectral lines measured in gaseous and supercritical helium (green data points) at pressures  $p = 101$  kPa, 231 kPa, 343 kPa and 424 kPa, and temperature  $T = 6.0\text{--}6.4$  K show progressively larger collisional shifts and broadenings relative to the data<sup>9</sup> measured in gas of  $p = 0.1$  kPa and  $T \approx 1.6$  K (grey) with a linewidth  $\Gamma_L = 0.19(3)$  GHz. The best fits of four overlapping Lorentzian functions are shown superimposed on each spectrum. Spectral lines observed in He I (blue) of  $T = 4.01$  K, 3.51 K, 3.01 K and 2.20 K became narrower as the liquid temperature was reduced. A rapid reduction of the linewidth below the He II transition temperature revealed the hyperfine structure, with the sharpest spectral lines observed between  $T = 1.78(2)$  K and 1.88(2) K (red). Error bars, 1 s.d.

As shown in Fig. 2, the resonance profile measured in a dilute gas target of pressure  $p = 0.1$  kPa and temperature  $T \approx 1.6$  K (ref. 9) resolves the hyperfine structure as two distinct peaks. The best fit of four overlapping Lorentzian functions is indicated by the grey curve with a linewidth of  $\Gamma_L \approx 0.19(3)$  GHz. When the pressure was increased by three orders of magnitude to  $p = 101.4(1.4)$  kPa at  $T = 6.4(4)$  K, collisions shifted the resonance centroid by  $\Delta\nu = -4.94(5)$  GHz and broadened the linewidth to  $\Gamma_L \approx 1.24(11)$  GHz, as expected. With a further, twofold increase of the target pressure to  $p = 231.1(1.4)$  kPa, the broadening became so great that the hyperfine structure could no longer be resolved (green curves). When the helium transitioned to the supercritical

phase upon increasing the pressure between  $p = 343(2)$  kPa and  $555(2)$  kPa (see Extended Data Table 1), the resonance further shifted from  $\Delta\nu = -36.7(4)$  GHz to  $-64.4(4)$  GHz and broadened to such a degree  $\Gamma_L = 15(2)$  GHz that the spectral line could no longer be resolved with a high signal-to-noise ratio. We used laser beams of fluence  $\varepsilon = 8\text{--}9$  mJ cm $^{-2}$  to compensate, and normalized the spectra of Fig. 2 to correct for the corresponding increase in the laser power.

In liquid helium, however, the spectral lines became far narrower despite the higher atomic density (Fig. 2). As the He I target was cooled from  $T = 4.157(15)$  K to  $2.201(17)$  K, the apparent linewidth rapidly decreased from  $\Gamma_L = 7.7(7)$  GHz to  $2.21(17)$  GHz (Fig. 3a). A similar reduction from  $\Gamma_L = 8.4(1.1)$  GHz to  $3.8(4)$  GHz was observed for the 597-nm resonance (Fig. 3b). This temperature dependence of  $\Gamma_L$  could be approximated by a single exponential (blue lines); the best fit yielded reduced  $\chi^2$  values of approximately 0.3 and 0.4 for the 726-nm and 597-nm resonances, respectively.

**Fig. 3: Linewidths of the spectral profiles of two  $\langle \bar{p} \rangle^4 \rm{H}$  laser transitions as a function of the He I or He II target temperature.**

 figure 3

**a, b,** The FWHM Lorentzian linewidths of the transition  $(n, \ell) = (37, 35) \rightarrow (38, 34)$  (a) and  $(n, \ell) = (39, 35) \rightarrow (38, 34)$  (b). In normal liquid helium the temperature dependence of the linewidth was well approximated by a single exponential (indicated by the blue curve), whereas in superfluid helium a characteristic dependence with the smallest linewidths at temperatures between  $T = 1.7$  K and  $1.9$  K was observed. Error bars, 1 s.d.

An abrupt further narrowing of the linewidth was observed below the He II phase transition temperature  $T_\lambda = 2.18$  K (Fig. 3a), which unexpectedly revealed the

hyperfine structure (Fig. 2). The highest resolution of  $\Gamma_L = 0.91(4)\text{--}0.92(4)$  GHz for the 726-nm resonance was observed at  $T = 1.78(2)\text{--}1.88(2)$  K, whereas the 597-nm resonance narrowed to  $\Gamma_L = 1.07(11)$  GHz at  $T = 1.73(2)$  K (Fig. 3b). At a slightly lower temperature  $T = 1.49(3)$  K the linewidths increased by factor of approximately 2 to  $\Gamma_L = 1.72(9)$  GHz and 2.04(15) GHz, respectively. This resulted in the characteristic temperature dependencies shown in Fig. 3a,b, which were fitted with parabolas. The spectral resolutions are an order of magnitude higher than the limits that would be expected according to the predictions of ab initio theoretical calculations for atoms of temperature  $T \approx 5.4$  K using the impact approximation of binary collisions<sup>17,18</sup>, which were based on pairwise potentials derived from the highly precise wavefunctions of  $(\bar{p})^4 H^{\dagger} e^{\dagger})^2 (+)$  (refs. 7,38). Taken together with the abrupt narrowing observed at  $T_\lambda$ , this implies that collective effects in the superfluid narrow the laser resonances below the limit expected from simple binary atomic collisions.

In gaseous and supercritical helium, the centroids of the two resonances shifted linearly with target density relative to the zero-density frequencies<sup>9</sup> (Fig. 4a and Extended Data Fig. 1a) with gradients  $d\nu/d\rho \approx -(4.0\text{--}4.5) \times 10^{-21}$  GHz cm<sup>3</sup>. This agrees with the results of previous experiments that were carried out in gas targets of much lower density<sup>36</sup>, and with the predictions<sup>17,18</sup> of the above binary collision calculations (Table 1). Similar gradients were observed in He I of temperatures between  $T = 4.2$  K and 2.8 K. When cooled below 2.5 K, however, the gradients began to increase in a nonlinear way relative to density, before abruptly changing sign at the He II transition temperature  $T_\lambda$  (Figs. 4b,c and Extended Data Fig. 1b,c). The temperature dependence of the frequency shifts likewise appear to be roughly linear between  $T = 2.2$  K and 2.55 K (Fig. 4d and Extended Data Fig. 1d) with gradients  $d\nu/dT = (3.5\text{--}3.9)$  GHz K<sup>-1</sup>, but in He II the gradients abruptly decreased by a factor of 5–10 (Table 1). The onset of superfluidity thus affects both the shift and linewidth of the atomic resonances.

**Fig. 4: Collisional shifts of the resonance frequency of the transition  $(n, \ell) = (37, 35) \rightarrow (38, 34)$  observed in gas, supercritical phase, He I and He II targets.**

---

 **figure 4**

**a**, Shifts in gaseous and supercritical helium (indicated by green squares), He I (blue triangles) and He II (red circles) targets relative to the zero-density transition frequency of  $\nu_0 = 412,885.133(1)$  GHz (ref. 9) as a function of the atomic density. **b, c**, Magnified views of **a**, showing a deviation from a linear density dependence at target temperature  $T < 2.5$  K. The results of the best fits of linear functions on the gaseous and supercritical helium (green line) and He I (blue line) data at  $T \geq 2.7$  K are shown superimposed. The size of the correlated uncertainty arising from the conversion of the liquid helium pressure and temperature into atomic density based on the International Temperature Scale of 1990 is indicated separately by the single error bar labelled as c.e. (see [Methods](#)). **d**, The temperature dependence shows an abrupt change in the gradient  $d\nu/dT$  at the He II transition temperature  $T_\lambda = 2.18$  K. Error bars, 1 s.d.

**Table 1 Gradients of the collisional shifts of the laser resonance frequencies of  $(\bar{\rho} \boldsymbol{p})^4 H_e^+$  against target density ( $d\nu/d\rho$ ) and temperature ( $d\nu/dT$ )**

The spectral lineshapes of other neutral atoms implanted into He I and He II have been interpreted by simplified effective models in which these impurities reside in bubble-like defects<sup>1,2,3,4,5,6</sup> of 1-nm-scale diameters  $d_b$ . The Franck–Condon principle implies that the excitation of the D1 transition of an impurity Cs atom, for example, occurs within a fixed configuration of the defect. This is rapidly followed by a more than  $\Delta d_b = 200$  pm expansion of the bubble diameter and some vibrations that occur on the surface of the bubble with a total energy  $E_b \approx 20$  meV (refs. 3,30). These vibrations are subsequently damped as some of the characteristic frequencies  $\nu_b = 100\text{--}500$  GHz

coincide with those of elementary roton ( $\nu_r \approx 180$  GHz) and maxon ( $\nu_m \approx 290$  GHz) excitations of He II. Because approximately 10 phonons with millielectronvolt energies are emitted into the surrounding liquid, a continuous optical spectrum of width  $\Gamma = 10^3\text{--}10^4$  GHz is observed. By comparison, laser transitions involving the inner electron orbitals of Tm, Eu, Cu or Au atoms cause smaller expansions in  $d_b$ , and so the resulting spectra contain sharp ‘zero phonon lines’ and relatively broad ‘phonon wings’ corresponding to zero and 1–2 phonon emissions, respectively<sup>29,30,31,39</sup>. The microscopic structure of bulk He II that causes these effects is not fully understood. Neutron diffraction experiments indicate that less than 10% of the  ${}^4\text{He}$  atoms occupy the ground state that corresponds to a Bose–Einstein condensate in a temperature-dependent way<sup>23,24,25,26</sup>, whereas the remainder occupy excited states that allow the phonons and rotons to propagate and interact with the impurities. For  $\text{SF}_6$  (ref. 40) and OCS molecules<sup>41</sup> attached to He II clusters cooled to  $T < 400$  mK, sharp rotational resonances at infrared wavelengths  $\lambda \approx 10.6$   $\mu\text{m}$  and  $4.85$   $\mu\text{m}$  have been observed. This implies that the molecules can rotate relatively freely in the superfluidity that arises within the small ( $10^2\text{--}10^4$ ) number of  ${}^4\text{He}$  atoms that comprise these clusters, though the moment of inertia of the corresponding transitions appear increased by factor 2.7–2.8 compared to a single isolated molecule. Compared to the visible-wavelength spectral lines of many other atoms and molecules embedded in He II, the antiprotonic E1 resonances of  $\langle\langle\bar{p}\rangle\rangle^4\langle\langle\text{rm}\{\text{H}\}\rangle\rangle\langle\langle\text{rm}\{\text{e}\}\rangle\rangle^+\rangle$  were distinguished by symmetrical and sharp lines that abruptly became narrow at temperatures below  $T_\lambda$ . This is believed to arise from the fact that the antiprotonic atom retains a radius  $r_e \approx 40$  pm of the electron shell during the laser excitation<sup>7</sup> which is an order of magnitude smaller than many other normal neutral atoms. The perturbations to the surrounding superfluid helium as evidenced by the calculated  $\langle\langle\bar{p}\rangle\rangle^4\langle\langle\text{rm}\{\text{H}\}\rangle\rangle\langle\langle\text{rm}\{\text{e}\}\rangle\rangle^+\rangle - \langle\langle\bar{p}\rangle\rangle^4\langle\langle\text{rm}\{\text{H}\}\rangle\rangle\langle\langle\text{rm}\{\text{e}\}\rangle\rangle\rangle$  pairwise potentials are so small that the spectral lineshapes become sensitive to minute effects, including the abrupt changes in the number densities of elementary excitations that occur during the superfluid phase transition.

A broad laser resonance of linewidth  $\Gamma \approx 100$  GHz corresponding to a relative resolution  $\Gamma/\nu \approx 6 \times 10^{-4}$  was recently observed in pionic helium ( $\pi^4\text{He}^+ \equiv {}^4\text{He}^{2+} + e^- + \pi^-$ ) atoms that were formed in He II (ref. 42). This linewidth mostly arises from the picosecond-scale Auger lifetime of the resonance daughter state. The present work implies that other spectral lines involving  $\pi^4\text{He}^+$  states with longer lifetimes may in principle reach much higher resolutions of approximately  $10^{-6}$ . Other three-body exotic helium atoms such as neutral kaonic helium ( $K^4\text{He}^+ \equiv {}^4\text{He}^{2+} + e^- + K^-$ ) with a mean lifetime of  $\tau = 10$  ns, and antideuteronic helium( $\langle\langle\bar{d}\rangle\rangle^4\langle\langle\text{rm}\{\text{H}\}\rangle\rangle\langle\langle\text{rm}\{\text{e}\}\rangle\rangle^+\rangle \equiv \langle\langle\bar{d}\rangle\rangle^4\langle\langle\text{rm}\{\text{H}\}\rangle\rangle\langle\langle\text{rm}\{\text{e}\}\rangle\rangle^+\rangle^2 + \langle\langle\bar{d}\rangle\rangle^4\langle\langle\text{rm}\{\text{H}\}\rangle\rangle\langle\langle\text{rm}\{\text{e}\}\rangle\rangle^-\rangle^2 + \langle\langle\bar{d}\rangle\rangle^4\langle\langle\text{rm}\{\text{H}\}\rangle\rangle\langle\langle\text{rm}\{\text{e}\}\rangle\rangle^+\rangle\langle\langle\bar{d}\rangle\rangle^4\langle\langle\text{rm}\{\text{H}\}\rangle\rangle\langle\langle\text{rm}\{\text{e}\}\rangle\rangle^-\rangle$ ) with a microsecond-scale lifetime are expected to have atomic structures and diameters that

are comparable to  $\langle \bar{p} \rangle \langle \{ \}^4 \{ \text{rm}{H} \} \{ \text{rm}{e} \} \rangle^{\{+\}}$ . Collisional shifts of the transition frequencies may limit the precision of future laser spectroscopy experiments on such atoms, but the shifts can be extrapolated to zero density as shown in Fig. 4.

Any anomalous excess in the flux of antiprotons<sup>10,11</sup> or antideuterons<sup>12,43</sup> of low kinetic energy  $E \leq 200$  MeV in cosmic rays is predicted to constitute an important indication of the possible dark matter that decay or annihilate in the halo of the Milky Way<sup>44,45</sup>, or primordial black hole evaporation by Hawking radiation<sup>46,47,48</sup>. The 30–300 litres of coolant He I or He II stored in some satellites over many months<sup>49,50</sup>, or alternatively in high-altitude balloons, correspond to an effective target thickness in which  $E \leq 100$  MeV antiprotons readily come to rest and form  $\langle \bar{p} \rangle \langle \{ \}^4 \{ \text{rm}{H} \} \{ \text{rm}{e} \} \rangle^{\{+\}}$ . Some 3% of the antiprotons are captured into long-lived states of  $\langle \bar{p} \rangle \langle \{ \}^4 \{ \text{rm}{H} \} \{ \text{rm}{e} \} \rangle^{\{+\}}$  that are formed in gaseous and liquid helium targets<sup>35,51</sup>. On the basis of the results of the experiment presented here, we estimate that the sharp spectral lines of these  $\langle \bar{p} \rangle \langle \{ \}^4 \{ \text{rm}{H} \} \{ \text{rm}{e} \} \rangle^{\{+\}}$  may be detected with an efficiency of approximately  $10^{-3}$  per stopped antiproton using laser spectroscopy, depending on the transitions that may be simultaneously interrogated.

## Methods

### Gaseous and supercritical phase helium target

The measurements involving the gaseous and supercritical phase helium targets (Fig. 1d) were carried out using a 35-mm-diameter chamber made of stainless steel that was designed to withstand inner pressures  $p > 1$  MPa at a temperature  $T \approx 4\text{--}6$  K. The antiprotons entered through a  $t_r = 75\text{-}\mu\text{m}$ -thick window made of annealed titanium that was attached to one side of the chamber by vacuum brazing, the laser beam through a 28-mm-diameter, 5-mm-thick UV-grade sapphire window affixed on the opposite side. The chamber was mounted on a liquid helium constant-flow cryostat. The cryostat was shielded from external thermal radiation by two layers of Cu enclosures, which were each cooled by either the vapour of the coolant helium or with liquid nitrogen.

The target pressure  $p$  was taken as the average between the readings of two capacitive manometers of precision  $\epsilon_{\text{prec}} = 1.4\text{--}2.0$  kPa that were in pressure equilibrium with the cryogenic target gas. During the measurement runs that each lasted 8–14 h, drifts of  $\epsilon_{\text{fluc}} = 0.2$  kPa were observed when the target pressure was adjusted to  $p \approx 100$  kPa. Larger fluctuations  $\epsilon_{\text{fluc}} = 1$  kPa were seen at a higher target pressure  $p \approx 560$  kPa. The

total uncertainty  $\epsilon_p = 1.4\text{--}2.2$  kPa on the target pressure was taken as the quadratic sum of  $\epsilon_{\text{fluc}}$  and  $\epsilon_{\text{prec}}$ .

A carbon-ceramic sensor<sup>52</sup> was mounted on the outer wall of the stainless steel chamber which had a specified precision  $\epsilon_c = 10$  mK. Its reading was stabilized to a value  $T_2 = 6.30$  K by regulating the current applied to a nichrome ribbon heater mounted on the heat exchanger of the cryostat using a proportional–integral–derivative (PID) controller. The temperature of the target gaseous or supercritical helium was measured by a second sensor of reading  $T_1$  which was suspended in the target helium. This  $T_1$  reading fluctuated by  $\epsilon_{\text{fluc}} = 10\text{--}20$  mK during the experimental runs. Both sensors were excited by currents  $I = 10$   $\mu\text{A}$  with a precision of  $\leq 10$  nA.

We calibrated the  $T_1$  sensor reading at nine target conditions of the liquid sealed in the chamber at temperatures between  $T_1 = 3.04$  K and  $4.53$  K by measuring the corresponding He I saturated vapour pressures between  $p = 34.9(1.4)$  kPa and  $142.2(1.4)$  kPa using the two manometers. At vapour pressures  $p > 120$  kPa the  $T_1$  reading deviated by less than 30 mK from the expected temperatures  $T(p)$  which were calculated using the programme HEPAK<sup>53</sup>. The programme was based on the parameterized state equations of helium according to the International Temperature Scale of 1990 (ITS-90)<sup>54</sup>. This value  $\epsilon_{\text{prec}} = 30$  mK was taken to be the uncertainty of measuring the gaseous or supercritical helium temperature at the position of the sensor. Deviations between the readings of sensors  $T_1$  and  $T_2$  arose owing to the differences in the thermal conductivities of the stainless steel chamber and the gaseous or supercritical helium. From this we estimated that the upper limit of the thermal gradient in the portion of the helium target where  $\bar{p}^{(4)} \{H\} \{e\}^{(+)} \}$  were produced varied between  $\epsilon_{\text{grad}} = 15$  mK and  $400$  mK depending on the target pressure. The uncertainty became particularly large ( $\epsilon_{\text{grad}} = 400$  mK) at the lowest pressure  $p \approx 101$  kPa used in this experiment. The total uncertainty  $\epsilon_T = 40\text{--}400$  mK of the temperature  $T$  of the gaseous and supercritical phase targets was taken to be the quadratic sum of the uncertainties  $\epsilon_c$ ,  $\epsilon_{\text{prec}}$ ,  $\epsilon_{\text{fluc}}$ , and  $\epsilon_{\text{grad}}$ . The target pressure and temperature were converted to atomic density  $\rho$  with an average uncertainty of 0.1% and a maximum uncertainty of 0.5%<sup>54,55</sup> using the HEPAK programme<sup>53</sup>.

## Thermometry of He I and He II target

The He I and He II target was a diameter  $d = 35$  mm chamber made of oxygen-free high conductivity (OFHC) copper. The antiprotons entered through a  $t_r = 50\text{-}\mu\text{m}$ -thick Ti window, whereas the laser beam entered through a 35-mm-diameter fused silica

window that was vacuum-brazed on the chamber. The chamber was mounted on the above constant-flow cryostat, and coolant liquid helium was circulated and evaporated by mechanical booster and rotary vane pumps with pumping speeds of  $500 \text{ m}^3 \text{ h}^{-1}$  and  $200 \text{ m}^3 \text{ h}^{-1}$ , respectively. A carbon-ceramic sensor<sup>52</sup> was suspended in the target liquid helium (Fig. 1d) which was isolated from the coolant liquid, and excited with a current  $I = 10 \mu\text{A}$ . During the spectroscopy experiments involving the He I target, the sensor readout was stabilized within  $\varepsilon_{\text{fluc}} = 2\text{--}11 \text{ mK}$  by either regulating the current applied to a nichrome ribbon heater mounted on the heat exchanger, or by adjusting the flow conductance through a butterfly throttle valve placed upstream of the pumps with a PID controller. The readout fluctuation measured in the He II target was  $\varepsilon_{\text{fluc}} = 2\text{--}5 \text{ mK}$ .

The 7th-degree polynomial calibration function used to convert the sensor resistance to temperature was obtained from the best fit on calibration data measured at 24 temperatures between  $T_1 = 1.5 \text{ K}$  and  $297 \text{ K}$  some 36 months prior to the  $(\bar{p})^{(\rm H-e)^+}$  experiments. Every  $\sim 24 \text{ h}$  during the laser spectroscopy measurements, the sensor reading  $T_1$  was calibrated at the vapour–He I–He II triple point which was taken to be  $T_\lambda = 2.1768 \text{ K}$  (ref. <sup>56</sup>). A second sensor of identical design, which was calibrated 3–4 months prior to the experiment, was placed at roughly the same position in the target liquid. As its reading  $T_2$  was within  $3 \text{ mK}$  of the literature value of  $T_\lambda$ , we used  $T_2$  to calibrate the  $T_1$  reading. In some of the calibration measurements, we stabilized the target temperature and pressure on the vapour saturation line of helium. The sensor readings were found to agree with the expected literature values within the  $\varepsilon_c = 10 \text{ mK}$  precision specified by the manufacturer and the uncertainty of the measured pressure. The temperature-dependent part of the uncertainty related to this calibration procedure involving the two sensors varied between  $\varepsilon_a = 4 \text{ mK}$  at a target temperature  $(T=4.16, \text{K})$ , and  $\varepsilon_a = 24 \text{ mK}$  at  $(T=1.49, \text{K})$ .

The total uncertainty  $\varepsilon_T = 12\text{--}27 \text{ mK}$  of the liquid target temperature  $T$  was obtained as the quadratic sum of the above uncertainties  $\varepsilon_a$ ,  $\varepsilon_c$ ,  $\varepsilon_{\text{fluc}}$  and  $\varepsilon_{\text{heat}}$ . Here  $\varepsilon_{\text{heat}} = 5 \text{ mK}$  denotes the maximum temperature gradient that may arise owing to the heating effect of the laser in the portion of the target volume where antiprotons came to rest (see below). The calibrated temperatures and uncertainties are provided in Extended Data Tables 1, 2.

## Heating of He I and He II target by laser beam

The pair of sensors suspended in the He I and He II target detected no substantial temperature excursions that coincided with the single laser pulse that arrived every

110–120 s. In Extended Data Fig. 2a the spatial distribution of antiprotons that came to rest in the He II target calculated by a Monte Carlo simulation based on the GEANT4 package<sup>57</sup> is shown. The emittance and energy distributions of the antiproton beam emerging from the RFQ decelerator was estimated by a particle-tracing programme<sup>58</sup> and used in the simulation. The  $\langle\bar{p}\rangle\langle\rm{H}\rangle\langle e\rangle^+\rangle$  atoms were found to form within a volume located 2–3 mm away from the Ti window.

We then carried out finite-element simulations based on the COMSOL package<sup>59,60</sup> to estimate the upper limits of the thermal gradients that arose in the He I target owing to the heating of the laser beam. We chose to simulate He I of temperature  $T = 2.3$  K which has a particularly small specific heat, conductivity and viscosity<sup>53</sup>. The dominant heat transport process in the He I was due to convection, and so an infinite-plate approximation including the effects of gravity and temperature-dependent density was used to simulate the laminar convective flows.

In the simulation, a 70-ns-long laser pulse deposited 70% of its  $E = 10$  mJ energy into the entrance foil. Some  $\sim 0.8$  s prior to the arrival of the antiprotons, a mechanical shutter opened and allowed part of the CW seed laser beam of power  $I = 200$  mW to enter the target. The actual intensity of the CW laser beam during the experiment is estimated to be less than 10% of this value. The shutter was closed some 1 s after antiproton arrival, and was not reopened until the next arrival 119 s later. Small integration steps in time  $\Delta t = 0.01$  s were used when the lasers were either turned on or off. Meshes with cell sizes of less than 50  $\mu\text{m}$  were used to calculate the flow and heat transport of the He I located near the Ti window.

The simulations showed that the largest heating occurred in a thin layer of He I that was located within a distance of  $d = 0.7$  mm from the entrance foil. Less than 1% of the  $\langle\bar{p}\rangle\langle\rm{H}\rangle\langle e\rangle^+\rangle$  were estimated to form in this region, whereas the remaining atoms were distributed in areas with much less heating. The  $\langle\bar{p}\rangle\langle\rm{H}\rangle\langle e\rangle^+\rangle$  with microsecond-scale lifetimes were destroyed well before the volume that includes the atoms could be considerably heated; the residual temperature excursion when the next antiproton pulse arrived at the target was  $\Delta T < 4$  mK (Extended Data Fig. 2b).

## Particle detectors

A  $300 \times 100 \times 20$  mm<sup>3</sup> Cherenkov detector made of UV-transparent acrylic was mounted near the cryostat so that it covered a solid angle of  $\sim 1.6\pi$  steradian seen from the experimental target (Fig. 1d). The charged particles (mostly pions) that emerged from the antiproton annihilations in the target traversed the detector. The resulting flash of Cherenkov light was detected by a gateable fine-mesh photomultiplier<sup>61</sup> with a 17.5-mm-diameter bialkali photocathode and high dynamic range. The waveform of

this signal was recorded by a digital oscilloscope of vertical resolution 8 bits, analogue bandwidth  $f_b = 4$  GHz, and digital sampling rate  $f_s = 5$  gigasamples per s.

The horizontal and vertical spatial profiles of the antiproton beam were measured by a pair of beam profile monitors<sup>62</sup> which were positioned between the RFQ decelerator and target. A small fraction of the beam was intercepted by a grid of 20-μm-diameter tungsten–rhenium wires that were plated with gold. The resulting secondary electron emission was measured by charge-sensitive preamplifiers. Pairs of dipole magnets were tuned to carefully steer the antiprotons into the experimental target.

## Laser systems

The 726-nm laser pulses were generated by an injection-seeded Ti:sapphire laser<sup>63</sup>. The system was based on a CW Ti:sapphire ring laser that was pumped by the second harmonic of a single-mode Nd:YVO<sub>4</sub> laser. The optical frequency of the CW laser was stabilized against a Ti:sapphire femtosecond frequency comb with a precision of better than 1 MHz (ref. <sup>64</sup>). The CW seed beam was injected into a triangular Ti:sapphire cavity of 0.8-m circumference, which was pumped by a Q-switched Nd:YAG laser to produce 40–50-ns-long laser pulses of energy  $E \approx 10$  mJ. The laser pulses of wavelengths  $\lambda = 842$  nm and 471 nm which were used to search for the resonances  $(n, \ell) = (38, 35) \rightarrow (39, 34)$  and  $(37, 34) \rightarrow (36, 33)$ , respectively, were also generated by this laser system. No statistically significant signal corresponding to the two transitions were observed for  $\langle \bar{p} \rangle^4 \langle H_e \rangle^4$  produced in He I targets, presumably owing to the high rate of collisions in He I that destroyed the antiproton populations in the related states.

The 597-nm laser pulses were generated by a CW pulse-amplified dye laser. A ring CW dye laser utilizing a rhodamine 6G dye solution dissolved in ethylene glycol was pumped by an argon ion laser. This seed beam was then amplified to  $E = 2\text{--}3$  mJ in an 18-mm-long rectangular dye cell, followed by a 30-mm-long Bethune dye cell filled with rhodamine B dye dissolved in methanol. The cells were pumped from the transverse direction by a Q-switched, single-longitudinal-mode Nd:YAG laser<sup>9</sup> of  $E = 180$  mJ. The frequency modulation induced in the optical frequencies of the Ti:sapphire and dye lasers during pulse amplification was not more than  $|\Delta\nu_{\text{las}}| = 30$  MHz<sup>63</sup> and 60 MHz, respectively.

The laser beams were expanded by a telescope and collimated by a pair of 25-mm-diameter irises placed at a distance of about 3 m from each other. The laser fluence of the 25-mm-diameter laser beam was measured behind the downstream iris.

## Resonance spectra

The resonance profiles of the transition  $(n, \ell) = (37, 35) \rightarrow (38, 34)$  at a wavelength  $\lambda = 726$  nm measured in gaseous and supercritical helium, He I, and He II targets are shown in Extended Data Fig. 3. The temperature and pressure of the target used in each measurement are indicated in the corresponding plot. Each data point in the spectra represents data collected from 2–5 antiproton pulses. The best fit of four overlapping Lorentzian functions are shown superimposed in the spectra measured in gaseous or supercritical helium (indicated using green curves), He I (blue curves) or He II (red curves) targets. The relative positions of the Lorentzian functions  $\Delta\nu_{\text{HFS}} = -0.9070$  GHz,  $-0.8808$  GHz,  $0.8690$  GHz and  $0.8783$  GHz were fixed to the theoretical values corresponding to the hyperfine intervals<sup>37</sup>, whereas the antiproton populations were assumed to be equally distributed among the magnetic sublevels. The corresponding profiles of the  $(39, 35) \rightarrow (38, 34)$  resonance at  $\lambda = 597$  nm are shown in Extended Data Fig. 4 with the hyperfine intervals fixed to  $\Delta\nu_{\text{HFS}} = -0.2795$  GHz,  $-0.2386$  GHz,  $0.2409$  GHz and  $0.2546$  GHz<sup>37</sup>.

The full width at half maximum (FWHM) Lorentzian widths  $\Gamma_L$  obtained from these fits (see Extended Data Tables 1, 2) avoided the ambiguities that may arise from a more specific and complicated lineshape model with numerous parameters. The total experimental uncertainty denoted as ‘(total)’ in Extended Data Tables 1, 2 is the quadratic sum of the statistical uncertainty ‘(stat)’ arising from the finite number of excited atoms in the experimental target that were detected by the Cherenkov detector, and systematic uncertainties. The systematic uncertainty is taken as the quadratic sum of the uncertainty ‘(fluc)’ that arose from fluctuations of the target temperature, and shifts ‘(las)’ of up to 30 or 60 MHz in the laser frequency that are due to spurious modulations that were induced during the pulsed amplification in the Ti:sapphire and dye lasers, respectively.

The spin-averaged transition frequencies that were determined from the best fit of the above four overlapping Lorentzian functions are shown in Extended Data Tables 1, 2. The systematic uncertainty ‘(fit)’ related to the selection of this simplified fit function was estimated by fitting the spectra with an alternative and more complicated model involving four overlapping Voigt functions that were fixed to the relative positions of the above hyperfine intervals. The Gaussian width of the Voigt function was varied between zero and the Doppler width arising from the Maxwellian thermal motions of the atoms that corresponded to the target temperature  $T$ . The maximum difference of the centroid frequencies determined by the Lorentzian and Voigt functions were taken as the systematic uncertainty, ‘(fit)’. The total uncertainty ‘(total)’ on the transition frequency was then taken as the quadratic sum of the statistical uncertainty ‘(stat)’ and the systematic uncertainties that include the contribution from ‘(fit)’ and the above laser modulation ‘(las)’ = 30 or 60 MHz.

## Data availability

All data are available from the corresponding author on reasonable request.

## References

1. Toennies, J. P. & Vilesov, A. F. Spectroscopy of atoms and molecules in liquid helium. *Annu. Rev. Phys. Chem.* **49**, 1–41 (1998).
2. Takahashi, Y., Sano, K., Kinoshita, T. & Yabuzaki, T. Spectroscopy of alkali atoms and molecules in superfluid helium. *Phys. Rev. Lett.* **71**, 1035–1038 (1993).
3. Kinoshita, T., Fukuda, K., Takahashi, Y. & Yabuzaki, T. Optical properties of alkali-metal atoms in pressurized liquid helium. *Phys. Rev. A* **52**, 2707–2716 (1995).
4. Bauer, H. et al. Laser spectroscopy of alkaline earth atoms in He II. *Phys. Lett. A* **146**, 134–140 (1990).
5. Kanorsky, S. I., Arndt, M., Dziewior, R., Weis, A. & Hänsch, T. W. Optical spectroscopy of atoms trapped in solid helium. *Phys. Rev. B* **49**, 3645–3647 (1994).
6. Hui, Q., Persson, J. L., Beijersbergen, J. H. M. & Takami, M. Spectroscopy and dynamics of neutral atoms in superfluid helium. *Z. Phys. B* **98**, 353–357 (1995).
7. Korobov, V. I., Hilico, L. & Karr, J.-P.  $m\alpha^7$ -order corrections in the hydrogen molecular ions and antiprotonic helium. *Phys. Rev. Lett.* **112**, 103003 (2014).
8. Hori, M. et al. Two-photon laser spectroscopy of antiprotonic helium and the antiproton-to-electron mass ratio. *Nature* **475**, 484–488 (2011).
9. Hori, M. et al. Buffer-gas cooling of antiprotonic helium to 1.5 to 1.7 K, and antiproton-to-electron mass ratio. *Science* **354**, 610–614 (2016).
10. Abe, K. et al. Measurement of the cosmic-ray antiproton spectrum at solar minimum with a long-duration balloon flight over Antarctica. *Phys. Rev. Lett.* **108**, 051102 (2012).
11. Aguilar, M. et al. Antiproton flux, antiproton-to-proton flux ratio, and properties of elementary particle fluxes in primary cosmic rays measured with the Alpha

- Magnetic Spectrometer on the International Space Station. *Phys. Rev. Lett.* **117**, 091103 (2016).
12. Aramaki, T. et al. Antideuteron sensitivity for the GAPS experiment. *Astropart. Phys.* **74**, 6–13 (2016).
  13. Ahmadi, M. et al. Characterization of the 1S–2S transition in antihydrogen. *Nature* **557**, 71–78 (2018).
  14. Baker, C. et al. Laser cooling of antihydrogen atoms. *Nature* **592**, 35–42 (2021).
  15. DiSciacca, J. et al. One-particle measurement of the antiproton magnetic moment. *Phys. Rev. Lett.* **110**, 130801 (2013).
  16. Ulmer, S. et al. High-precision comparison of the antiproton-to-proton charge-to-mass ratio. *Nature* **524**, 196–199 (2015).
  17. Bakalov, D., Jeziorski, B., Korona, T., Szalewicz, K. & Tchoukova, E. Density shift and broadening of transition lines in antiprotonic helium. *Phys. Rev. Lett.* **84**, 2350–2353 (2000).
  18. Bakalov, D. Density shift and broadening of dipole transitions in antiprotonic helium. *Hyperfine Interact.* **209**, 25–28 (2012).
  19. Badertscher, A. et al. Experimental determination of the kinetic energy distribution of  $\pi^- p$  atoms in liquid hydrogen. *Phys. Lett. B* **392**, 278–282 (1997).
  20. Siems, T. et al. First direct observation of Coulomb explosion during the formation of exotic atoms. *Phys. Rev. Lett.* **84**, 4573–4576 (2000).
  21. Jensen, T. S. & Markushin, V. E. Collisional deexcitation of exotic hydrogen atoms in highly excited states. *Eur. Phys. J. D* **21**, 271–283 (2002).
  22. Fetkovich, J. G., McKenzie, J., Riley, B. R. & Wang, I.-T. Measurement of the cascade time of  $\Sigma^-$  in liquid helium. *Nucl. Phys. A* **240**, 485–492 (1975).
  23. Glyde, H. R., Azuah, R. T. & Stirling, W. G. Condensate, momentum distribution, and final-state effects in liquid  $^4\text{He}$ . *Phys. Rev. B* **62**, 14337–14349 (2000).
  24. Zheng-Johansson, J. X. & Johansson, P.-I. *The Microscopic Theory of Superfluid He II and its QCE Superfluidity Mechanism Applied to Superconductors: Theory of Condensed Matter Expounded Through the System He II* (Nova Science Publishers, 2004).

25. Diallo, S. O. et al. Evidence for a common physical origin of the Landau and BEC theories of superfluidity. *Phys. Rev. Lett.* **113**, 215302 (2014).
26. Dmowski, W. et al. Observation of dynamic atom–atom correlation in liquid helium in real space. *Nat. Commun.* **8**, 15294 (2017).
27. Lemeshko, M. Quasiparticle approach to molecules interacting with quantum solvents. *Phys. Rev. Lett.* **118**, 095301 (2017).
28. Ishikawa, K. et al. Laser spectroscopy of thulium atoms implanted in liquid and solid  $^4\text{He}$ . *Phys. Rev. B* **56**, 780–787 (1997).
29. Hui, Q. & Takami, M. Phonon bands associated with the inner-shell electronic absorption lines of Eu atoms in bulk liquid helium. *J. Low Temp. Phys.* **119**, 393–401 (2000).
30. Moroshkin, P., Lebedev, V. & Weis, A. Phonon generation in condensed  $^4\text{He}$  by laser-excited atomic bubbles. *Europhys. Lett.* **96**, 26002 (2011).
31. Moroshkin, P., Borel, A. & Kono, K. Laser spectroscopy of phonons and rotons in superfluid helium doped with Dy atoms. *Phys. Rev. B* **97**, 094504 (2018).
32. Moroshkin, P. & Kono, K. Zero-phonon lines in the spectra of dysprosium atoms in superfluid helium. *Phys. Rev. B* **99**, 104512 (2019).
33. Russell, J. E. Metastable states of  $\alpha\pi^-e^-$ ,  $\alpha K^-e^-$ , and  $\backslash(\backslash\alpha \backslash\bar{p}\{e\}^{-})$  atoms. *Phys. Rev. Lett.* **23**, 63–64 (1969).
34. Fetkovich, J. G., Riley, B. R. & Wang, I. T. The atomic cascade of negative particles in liquid helium. *Phys. Lett. B* **35**, 178–180 (1971).
35. Iwasaki, M. et al. Discovery of antiproton trapping by long-lived metastable states in liquid helium. *Phys. Rev. Lett.* **67**, 1246–1249 (1991).
36. Hori, M. et al. Sub-ppm laser spectroscopy of antiprotonic helium and a CPT-violation limit on the antiprotonic charge and mass. *Phys. Rev. Lett.* **87**, 093401 (2001).
37. Bakalov, D. & Korobov, V. Hyperfine structure of antiprotonic helium energy levels. *Phys. Rev. A* **57**, 1662–1667 (1998).
38. Korobov, V. I., Zhong, Z.-X. & Tian, Q.-L. Leading term of the  $\backslash(\backslash\rm{He}\}-\backslash\bar{p}\{\backslash\rm{He}\}^+)$  long-range interaction. *Phys. Rev. A* **92**, 052517 (2015).

39. Hizhnyakov, V., Boltrushko, V. & Benedek, G. Thermal broadening of the zero-phonon line in superfluid helium. *Phys. Rev. B* **103**, 214515 (2021).
40. Hartmann, M., Miller, R. E., Toennies, J. P. & Vilesov, A. Rotationally resolved spectroscopy of SF<sub>6</sub> in liquid helium clusters: a molecular probe of cluster temperature. *Phys. Rev. Lett.* **75**, 1566–1569 (1995).
41. Grebenev, S., Toennies, J. P. & Vilesov, A. F. Superfluidity within a small helium-4 cluster: the microscopic Andronikashvili experiment. *Science* **279**, 2083–2086 (1998).
42. Hori, M., Aghai-Khozani, H., Sótér, A., Dax, A. & Barna, D. Laser spectroscopy of pionic helium atoms. *Nature* **581**, 37–41 (2020).
43. Nozzoli, F., Dimiccoli, F., Iuppa, R., Riccia, E. & Zuccona, P. An helium calorimeter for antideuteron identification in cosmic rays. In *Proc. 37th International Cosmic Ray Conf. (ICRC2021)* (IUPAP, 2021).
44. Donato, F., Maurin, D., Brun, P., Delahaye, T. & Salati, P. Constraints on WIMP dark matter from the high energy PAMELA  $\bar{p}/p$  data. *Phys. Rev. Lett.* **102**, 071301 (2009).
45. Cuoco, A., Krämer, M. & Korsmeier, M. Novel dark matter constraints from antiprotons in light of AMS-02. *Phys. Rev. Lett.* **118**, 191102 (2017).
46. Turner, M. S. Could primordial black holes be the source of the cosmic ray antiprotons? *Nature* **297**, 379–381 (1982).
47. Maki, K., Mitsui, T. & Orito, S. Local flux of low-energy antiprotons from evaporating primordial black holes. *Phys. Rev. Lett.* **76**, 3474–3477 (1996).
48. Barrau, A. et al. Antiprotons from primordial black holes. *Astron. Astrophys.* **388**, 676–687 (2002).
49. Fujimoto, R. et al. Performance of the helium dewar and the cryocoolers of the Hitomi soft x-ray spectrometer. *J. Astron. Telesc. Instrum. Syst.* **4**, 011208 (2017).
50. Gehrz, R. D. et al. The NASA Spitzer Space Telescope. *Rev. Sci. Instrum.* **78**, 011302 (2007).
51. Hori, M. et al. Primary populations of metastable antiprotonic <sup>4</sup>He and <sup>3</sup>He atoms. *Phys. Rev. Lett.* **89**, 093401 (2002).

52. Filippov, Y. P. & Miklyaev, V. M. A comparison of two kinds of TVO cryogenic temperature sensors. *Cryogenics* **100**, 85–91 (2019).
53. Arp, V. D., McCarty, R. D. & Friend, D. G. Thermophysical Properties of Helium-4 from 0.8 to 1500 K with Pressures to 2000 MPa. Technical note 1334 (NIST, 1998).
54. Preston-Thomas, H. The International Temperature Scale of 1990 (ITS-90). *Metrologia* **27**, 3–10 (1990); erratum **27**, 107 (1990).
55. McCarty, R. D. & Arp, V. D. A new wide range equation of state for helium. In Advances in Cryogenic Engineering (ed. Fast, R. W.) 1465–1475 (Springer, 1990).
56. Donnelly, R. J. & Barenghi, C. F. The observed properties of liquid helium at the saturated vapor pressure. *J. Phys. Chem. Ref. Data* **27**, 1217–1274 (1998).
57. Agostinelli, S. et al. GEANT4—a simulation toolkit. *Nucl. Instrum. Methods Phys. Res. A* **506**, 250–303 (2003).
58. Bylinsky, Y., Lombardi, A. M. & Pirkl, W. RFQD - a decelerating radio frequency quadrupole for the CERN antiproton facility. In *Proc. 20th Intl. Linear Accelerator Conf.* (ed. Chao, A. W.) 554–556 (SLAC, 2000)
59. Bielert, E. R., Verweij, A. P. & ten Kate, H. H. J. Implementation of the superfluid helium phase transition using finite element modeling: simulation of transient heat transfer and He-I/He-II phase front movement in cooling channels of superconducting magnets. *Cryogenics* **53**, 78–85 (2013).
60. Bielert, E. R., ten Kate, H. H. J. & Verweij, A. P. A structured approach to analyze the influence of channel dimensions on heat extraction via superfluid helium. *Phys. Proc.* **67**, 815–821 (2015).
61. Hori, M., Yamashita, K., Hayano, R. S. & Yamazaki, T. Analog Cherenkov detectors used in laser spectroscopy experiments on antiprotonic helium. *Nucl. Instrum. Meth. A* **496**, 102–122 (2003).
62. Hori, M. Photocathode microwire monitor for nondestructive and highly sensitive spatial profile measurements of ultraviolet, x-ray, and charged particle beams. *Rev. Sci. Instrum.* **76**, 113303 (2005).
63. Hori, M. & Dax, A. Chirp-corrected, nanosecond Ti:Sapphire laser with 6 MHz linewidth for spectroscopy of antiprotonic helium. *Opt. Lett.* **34**, 1273–1275 (2009).

64. Diddams, S. A., Vahala, K. & Udem, T. Optical frequency combs: coherently uniting the electromagnetic spectrum. *Science* **369**, eaay3676 (2020).

## Acknowledgements

We are indebted to the CERN Antiproton Decelerator and Proton Synchrotron operational staff, the CERN cryogenics laboratory, and the members of the ASACUSA collaboration. We thank A. Adamczak, D. Bakalov, S. Balibar, D. M. Ceperley, T. W. Hänsch, R. Holzwarth, V. I. Korobov, M. Lemeshko, A. Lombardi, W. Pirkl, O. Pirotte, P. A. Posocco and T. Udem. This work was supported by the Max-Planck-Gesellschaft, the European Research Council (ERC-Stg), and the Ludwig-Maximilians-Universität München.

## Funding

Open access funding provided by Max Planck Society.

## Author information

### Author notes

1. Anna Sótér

Present address: ETH Zurich, Zurich, Switzerland

2. Hossein Aghai-Khozani

Present address: McKinsey and Company, Munich, Germany

3. Andreas Dax

Present address: Paul Scherrer Institut, Villigen, Switzerland

## Affiliations

1. Max-Planck-Institut für Quantenoptik, Garching, Germany

Anna Sótér, Hossein Aghai-Khozani & Masaki Hori

2. CERN, Geneva, Switzerland

Dániel Barna & Andreas Dax

3. Institute for Particle and Nuclear Physics, Wigner Research Center for Physics,  
Budapest, Hungary

Dániel Barna

4. Dipartimento di Ingegneria dell'Informazione, Università di Brescia, Brescia,  
Italy

Luca Venturelli

5. Istituto Nazionale di Fisica Nucleare, Sezione di Pavia, Pavia, Italy

Luca Venturelli

6. Fakultät für Physik, Ludwig-Maximilians-Universität München, Munich,  
Germany

Masaki Hori

## Contributions

M.H. proposed the experiment and developed the Cherenkov detectors, antiproton beam profile monitors, and data acquisition. M.H., A.D. and A.S. developed the laser systems. A.S. and M.H. developed the cryogenic target, carried out the offline data analysis, and wrote the manuscript. M.H. and A.S. constructed the magnetic beamline for antiprotons connected to the RFQ decelerator. A.S. carried out the temperature calibration of the target and the simulations. All authors participated in the data taking, discussed the results, and provided manuscript comments.

## Corresponding author

Correspondence to [Masaki Hori](#).

## Ethics declarations

## Competing interests

The authors declare no competing interests.

## Peer review

## Peer review information

*Nature* thanks Andrzej Adamczak, Makoto Fujiwara, Yukari Matsuo and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

## Additional information

### Extended data

is available for this paper at <https://doi.org/10.1038/s41586-022-04440-7>.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Extended data figures and tables

### Extended Data Fig. 1 Collisional shifts of the resonance frequency of the transition $(n, \ell) = (39, 35) \rightarrow (38, 34)$ observed in gas and supercritical phase helium, He I and He II targets.

**a**, Shifts in gaseous and supercritical helium (indicated by green square markers), He I (blue triangles), and He II (red circles) targets relative to the zero-density transition frequency of  $\nu_0 = 501,948.753(2)$  GHz<sup>9</sup> as a function of the atomic density. **b,c**

Magnified views of a, showing a deviation from a linear density dependence at target temperature  $T < 2.5$  K. The results of the best fits of linear functions on the gaseous and supercritical helium (green line) and He I (blue line) data at  $T \geq 2.7$  K are shown superimposed. The size of the correlated uncertainty arising from the conversion of the liquid helium pressure and temperature into atomic density based on the International Temperature Scale of 1990 is indicated separately by the error bar labelled as c.e. (see text). **d**, The temperature dependence shows an abrupt change in the gradient  $d\nu/dT$  at the He II transition temperature  $T_\lambda = 2.18$  K. Error bars, 1 s.d.

### Extended Data Fig. 2 Simulated spatial distributions of antiprotons and the temperature in the He I target.

**a**, Spatial distribution of antiprotons that come to rest in the He I target calculated by a GEANT4 Monte Carlo simulation (not to scale). **b**, Spatial distribution of temperature in the He I target prior to the arrival of the laser pulse, simulated by COMSOL. In the portion of the target where the  $\backslash(\backslash bar\{p\} \backslash\{rm\{H\}\} \backslash\{rm\{e\}\})^+\backslash$  atoms were formed, the laser pulse arriving 119 s earlier is estimated to increase the temperature of the He I by not more than 4 mK.

**Extended Data Fig. 3 Resonance profiles of the  $(n, \ell) = (37, 35) \rightarrow (38, 34)$  transition of wavelength  $\lambda = 726$  nm as a function of the optical frequency of the laser observed in gas and supercritical phase helium, He I and He II targets.**

The temperature and atomic density of the helium target that was used in each measurement are indicated (see text). Error bars, 1 s.d.

**Extended Data Fig. 4 Resonance profiles of the  $(n, \ell) = (39, 35) \rightarrow (38, 34)$  transition of wavelength  $\lambda = 597$  nm as a function of the optical frequency of the laser observed in gas and supercritical phase helium, He I and He II targets.**

The temperature and atomic density of the helium target used in each measurement are indicated (see text). Error bars, 1 s.d.

**Extended Data Table 1 Frequency centroids and FWHM Lorentzian linewidths of the  $(n, \ell) = (37, 35) \rightarrow (38, 34)$  resonance of wavelength  $\lambda = 726$  nm at various target phases and densities**

**Extended Data Table 2 Frequency centroids and FWHM Lorentzian linewidths of the  $(n, \ell) = (39, 35) \rightarrow (38, 34)$  resonance of wavelength  $\lambda = 597$  nm at various target phases and densities**

## Rights and permissions

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

**Reprints and Permissions**

## About this article

## Cite this article

Sótér, A., Aghai-Khozani, H., Barna, D. *et al.* High-resolution laser resonances of antiprotonic helium in superfluid  $^4\text{He}$ . *Nature* **603**, 411–415 (2022).  
<https://doi.org/10.1038/s41586-022-04440-7>

- Received: 14 June 2021
- Accepted: 18 January 2022
- Published: 16 March 2022
- Issue Date: 17 March 2022
- DOI: <https://doi.org/10.1038/s41586-022-04440-7>

## Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

## [Superfluid confines exotic atoms without disrupting precision measurements](#)

- Yukari Matsuo

News & Views 16 Mar 2022

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-022-04440-7>

- Article
- Open Access
- [Published: 16 March 2022](#)

# Unbiasing fermionic quantum Monte Carlo with a quantum computer

- [William J. Huggins](#)<sup>1</sup>,
- [Bryan A. O’Gorman](#)<sup>2</sup>,
- [Nicholas C. Rubin](#)<sup>1</sup>,
- [David R. Reichman](#) [ORCID: orcid.org/0000-0001-8424-902X](#)<sup>3</sup>,
- [Ryan Babbush](#)<sup>1</sup> &
- [Joonho Lee](#) [ORCID: orcid.org/0000-0002-9667-1081](#)<sup>1,3</sup>

*Nature* volume 603, pages 416–420 (2022)

- 118 Altmetric
- [Metrics details](#)

## Subjects

- [Chemical physics](#)
- [Quantum physics](#)
- [Theoretical chemistry](#)

## Abstract

Interacting many-electron problems pose some of the greatest computational challenges in science, with essential applications across many fields. The solutions to these problems will offer accurate predictions of chemical reactivity and kinetics, and other properties of quantum systems<sup>1,2,3,4</sup>. Fermionic quantum Monte Carlo (QMC) methods<sup>5,6</sup>, which use a statistical sampling of the ground state, are among the most powerful approaches to these problems. Controlling the fermionic sign problem with constraints ensures the efficiency of QMC at the expense of potentially significant biases owing to the limited flexibility of classical computation. Here we propose an approach that combines constrained QMC with quantum computation to reduce such

biases. We implement our scheme experimentally using up to 16 qubits to unbias constrained QMC calculations performed on chemical systems with as many as 120 orbitals. These experiments represent the largest chemistry simulations performed with the help of quantum computers, while achieving accuracy that is competitive with state-of-the-art classical methods without burdensome error mitigation. Compared with the popular variational quantum eigensolver<sup>7,8</sup>, our hybrid quantum-classical computational model offers an alternative path towards achieving a practical quantum advantage for the electronic structure problem without demanding exceedingly accurate preparation and measurement of the ground-state wavefunction.

[Download PDF](#)

## Main

The complexity of finding an accurate solution of the Schrödinger equation seemingly grows exponentially with the number of electrons in the system. This fact has greatly hindered progress towards an efficient means of accurately calculating ground-state quantum mechanical properties of complex systems. Over the last century, substantial research effort has been devoted to the development of new algorithms for solution of this many-electron problem. At present, all available general-purpose methods can be grouped into two categories: (1) methods that scale exponentially with system size while yielding numerically exact answers, and (2) methods for which the cost scales polynomially with system size, but that are only approximate by construction.

Approaches in this second category are the only methods that can feasibly be applied to large systems at present. The accuracy of the solutions obtained by these methods is often unsatisfactory and is almost always difficult to assess.

Quantum computing has arisen as an alternative model for the calculation of quantum properties that may complement, and potentially surpass, classical methods in terms of efficiency<sup>9,10</sup>. Although the ultimate ambition of this field is to construct a universal fault-tolerant quantum computer<sup>11</sup>, the experimental devices of today are limited to noisy intermediate-scale quantum (NISQ) computers<sup>12</sup>. NISQ algorithms for the computation of ground states have largely centred around the variational quantum eigensolver (VQE) framework<sup>7,8</sup>, which necessitates coping with optimization difficulties, measurement overhead and circuit noise. As an alternative, algorithms based on imaginary-time evolution have been put forward, which, in principle, avoid the optimization problem<sup>13,14</sup>. However, because of the non-unitary nature of imaginary-time evolution, one must resort to heuristics to achieve reasonable scaling with system size. New strategies that avoid these limitations may help to enable the first practical quantum advantage in fermionic simulations. In this work, we propose and experimentally demonstrate a class of quantum-classical hybrid algorithms that

offer a different route to addressing these challenges. We do not attempt to represent the ground-state wavefunction using our quantum processor, choosing instead to use it to guide a quantum Monte Carlo (QMC) calculation performed on a classical coprocessor. Using this approach, our experimental demonstration surpasses the scale of previous experimental work on quantum simulation in chemistry<sup>15,16,17</sup>.

## Theory and algorithms

QMC approaches<sup>5,6</sup> target the exact ground-state wavefunction,  $\langle|\{\Psi\}_0\rangle$ , of a many-body Hamiltonian,  $\langle\hat{H}\rangle$ , via imaginary-time evolution of an initial state  $\langle|\{\Phi\}_0\rangle$  with a non-zero overlap with  $\langle|\{\Psi\}_0\rangle$ :

$$\langle|\{\varPsi\}_0\rangle \xrightarrow{\text{propto}} \lim_{\tau \rightarrow \infty} \langle|\varPsi(\tau)\rangle, \quad \langle|\varPsi(\tau)\rangle \equiv \langle|\rm{e}^{-\tau\hat{H}}|\varPhi_0\rangle \quad (1)$$

where  $\langle|\tau\rangle$  is imaginary time and  $\langle|\Psi(\tau)\rangle$  denotes the time-evolved wavefunction from  $\langle|\Phi_0\rangle$  by  $\tau$  (Fig. 1a). In QMC, the imaginary-time evolution in equation (1) is implemented stochastically, which can enable a polynomial scaling algorithm to sample an estimate for the exact ground-state energy by avoiding the explicit storage of high-dimensional objects, such as  $\langle\hat{H}\rangle$  and  $\langle|\{\Psi\}_0\rangle$ . The ground-state energy,  $\langle E \rangle = E(\tau = \infty)$ , is estimated from averaging a time series of  $\langle|\{E\}^i(\tau)\rangle$ , given by a weighted average over  $M$  statistical samples,

$$\langle E \rangle = \frac{1}{M} \sum_i w_i E^{(i)} \quad (2)$$

where  $E^{(i)}(\tau)$  is the  $i$ th statistical sample for the energy and  $w_i(\tau)$  is the corresponding normalized weight for that sample at imaginary time  $\tau$ . Although formally exact, such a stochastic imaginary-time evolution algorithm will generically run into the fermionic sign problem<sup>18</sup>, which manifests as a result of alternating signs in the weights of each statistical sample used in equation (2). In the worst case, the fermionic sign problem causes the estimator of the energy in equation (2) to have exponentially large variance (Fig. 1b, top), necessitating that one averages exponentially many samples so as to obtain a target precision. Accordingly, exact, unbiased QMC approaches are only applicable to small systems<sup>19,20,21</sup> or those lacking a sign problem<sup>22</sup>.

**Fig. 1: Imaginary-time evolution, sign problem and our quantum-classical hybrid algorithm.**

---

 **figure 1**

**a**, Depiction of the imaginary-time evolution, which shows an exponential convergence to the ground state as a function of imaginary time,  $\tau$ . **b**, Illustration of the fermionic sign problem. Exact deterministic imaginary-time evolution and an unconstrained QMC calculation, which is exact on average but has a signal-to-noise ratio that diverges with increasing  $\tau$  due to the sign problem (top). Constrained QMC calculations with classical and quantum constraints. The use of quantum constraint helps to reduce the bias that is non-negligible when using the classical constraint (bottom). **c**, Overview of the QC-QMC algorithm. Stochastic wavefunction samples, represented as  $\langle \langle \varphi_i | \rangle \rangle_\tau$ , are evolved in time along with associated weights  $\{w_i\}_\tau$ . Throughout the time evolution, queries to the quantum processor about the overlap value between the quantum trial wavefunction  $\langle |\Psi_T\rangle |$  and a stochastic wavefunction sample  $\langle \langle \varphi_i | \rangle \rangle_\tau$  are made while updating the gate parameters to describe  $\langle \langle \varphi_i | \rangle \rangle_\tau$ . Our quantum processor uses  $N$  qubits to efficiently estimate the overlap, which is then used to evolve  $w_i$  and to discard stochastic wavefunction samples with  $w_i < 0$ . Finally, observables, such as  $\langle E(\tau) \rangle$ , are computed on the classical computer using overlap queries to the quantum processor ([Supplementary Section C](#)).

The sign problem can be controlled to give an estimator of the ground-state energy with polynomially bound variance by imposing constraints on the imaginary-time evolution of each statistical sample represented by a trial wavefunction,  $\langle \langle \varphi_i | \rangle \rangle_\tau$ . These constraints (which include prominent examples such as the fixed node<sup>6,23</sup> and phaseless approximations<sup>24,25</sup>) are imposed by demanding that the overlaps of the trial wavefunction  $\langle \langle |\Psi_T\rangle | \rangle$  (where T denotes trial) with the stochastic samples  $\langle \langle \varphi_i | \rangle \rangle_\tau$  remain

positive during the imaginary-time propagation. Although constrained QMC calculations are typically much more accurate than those using the bare trial wavefunction directly (Fig. 1b, bottom), the remaining bias of the constrained QMC results is wholly determined by the choice of the trial wavefunction. Imposing these constraints necessarily introduces a potentially significant bias in the final ground-state energy estimate, which can be removed in the limit that the trial wavefunction approaches the exact ground state. Alternatively, the bias can be removed by releasing the constraints during propagation, at the expense of suffering an uncontrolled sign problem<sup>26</sup>.

Classically, computationally tractable options for trial wavefunctions are limited to states such as a single mean-field determinant (for example, a Hartree–Fock state), a linear combination of mean-field states, a simple form of the electron–electron pair (two-body) correlator (usually called a Jastrow factor) applied to mean-field states or some other physically motivated transformations applied to mean-field states, such as backflow approaches<sup>27</sup>. On the other hand, any wavefunction that can be prepared with a quantum circuit is a candidate for a trial wavefunction on a quantum computer, including more general two-body correlators. These trial wavefunctions will be referred to as ‘quantum’ trial wavefunctions.

At present, there is no efficient classical algorithm to estimate (to additive error) the overlap between  $\langle|\{\varphi\}_i(\tau)\rangle$  and various quantum trial wavefunctions  $\langle|\{\Psi\}_{\text{T}}\rangle$ , such as unitary coupled-cluster with singles and doubles<sup>28</sup>, qubit coupled-cluster methods<sup>29</sup>, wavefunctions constructed by adiabatic state preparation<sup>30</sup> or the multiscale entanglement renormalization ansatz<sup>31</sup>. This is true even when  $\langle|\{\varphi\}_i(\tau)\rangle$  is simply a computational basis state or a Slater determinant. As quantum computers can efficiently approximate  $\langle\langle\Psi_{\text{T}}|\{\varphi\}_i(\tau)\rangle\rangle$ , there is a potential quantum advantage in this task, as well as its particular use in QMC. This offers a different route towards quantum advantage in ground-state fermion simulations compared with VQE, which instead seeks an advantage in the variational energy evaluation. We expand on this discussion of quantum advantage in Supplementary Section F.

Our quantum-classical hybrid QMC algorithm (QC-QMC) utilizes quantum trial wavefunctions while performing the majority of imaginary-time evolution on a classical computer, and is summarized in Fig. 1c. In essence, on a classical computer one performs imaginary-time evolution for each wavefunction statistical sample,  $\langle|\{\varphi\}_i(\tau)\rangle$ , and collects observables such as the ground-state energy estimate,  $\langle|E^i(\tau)|\rangle$ . During this procedure, a constraint associated with the quantum trial wavefunction is imposed to control the sign problem. To perform the constrained time evolution, the only quantity that needs to be calculated on the quantum computer is the overlap between the trial wavefunction,  $\langle|\{\Psi\}$

$\}_{\text{range}} \rangle$ , and the statistical sample of the wavefunction at imaginary time  $\tau$ ,  $\langle |\{\varphi_i\}(\tau)\rangle$ . Although our approach applies generally to any form of constrained QMC, here we discuss an experimental demonstration of the algorithm that uses an implementation of QMC known as auxiliary-field QMC (AFQMC), which will be referred to as QC-AFQMC (see [Methods](#) for more details). As a single determinant mean-field trial wavefunction is the most widely used classical form of the trial function for AFQMC owing to its efficiency, here we use ‘AFQMC’ to denote AFQMC with a mean-field trial wavefunction.

## Discussion

As the first example, in Fig. 2 we illustrate the quantum primitive used to perform the experiment on an  $H_4$  molecule involving 8 qubits (see [Methods](#) for more details). Our eight spin-orbital quantum trial wavefunction consists of a valence bond wavefunction known as a perfect pairing state<sup>32,33</sup> and a hardware-efficient quantum circuit<sup>15</sup> with an offline single-particle rotation, which would be classically difficult to use as a trial wavefunction for AFQMC. The state preparation circuit in Fig. 2a shows how this trial wavefunction can be efficiently prepared on a quantum computer.

**Fig. 2: 8-qubit experiment.**

 figure 2

- 
- a.** Circuit used for the 8-qubit  $H_4$  experiment over a  $2 \times 4$  qubit grid (from  $Q_{1,1}$  to  $Q_{2,1}$ ) on the Sycamore quantum processor<sup>43</sup>. In the circuit diagram, H denotes the Hadamard gate, G denotes a Givens rotation gate (generated by  $XX + YY$ ), P denotes a single-qubit Clifford gate and  $\langle |\{\Psi\}_{\text{range}} \rangle$  denotes the quantum trial wavefunction. Note that the ‘offline’ orbital rotation is not present in the actual quantum circuit because it is handled via classical post-processing, as discussed in [Supplementary Section C](#). **b, c,** Convergence of the atomization energy of  $H_4$  as a function of the number of measurements. A minimal basis set (STO-3G) with four orbitals total from four independent experiments (exp.) (**b**) and a quadruple-zeta basis set (cc-pVQZ) with 120 orbitals total from two independent experiments (**c**). The different symbols in **b** and **c** show independent experimental results. Note that the ideal (that is, noiseless) atomization energy of the quantum trial (Q. trial) in **b** is

precisely on top of the exact one and that the QC-AFQMC energy would likewise be exact in the absence of noise. For the system in **c**, QC-AFQMC with this quantum trial would yield an error of 0.2 kcal mol<sup>-1</sup> despite a much larger error in the variational energy of the quantum trial. Further note that the quantum resource used in **c** is 8 qubit, but, as shown in [Supplementary Section C](#), our algorithm enables the addition of ‘virtual’ electron correlation in a much larger Hilbert space. The top panels of **b** and **c** magnify the energy range near the exact answer. See Extended Data Tables [1–8](#) for the raw data for **b**, **c**, as well as other relevant data.

### [Source data](#)

In this 8-qubit experiment, we consider H<sub>4</sub> in a square geometry with side lengths of 1.23 Å and its dissociation into four hydrogen atoms. This system is often used as a test bed for electron correlation methods in quantum chemistry<sup>[34,35](#)</sup>. We perform our calculations using two Gaussian basis sets: the minimal (STO-3G) basis set<sup>[36](#)</sup> and the correlation consistent quadruple-zeta (cc-pVQZ) basis set<sup>[37](#)</sup>. The latter basis set is of the size and accuracy required to make a direct comparison with laboratory experiments. When describing the ground state of this system, there are two equally important, degenerate mean-field states. This makes AFQMC with a single mean-field trial wavefunction highly unreliable. In addition, a method often referred to as a ‘gold standard’ classical approach (that is, coupled-cluster with singles, doubles and perturbative triples, CCSD(T)<sup>[38](#)</sup>) also performs poorly for this system.

In Table [1](#), the difficulties of AFQMC and CCSD(T) are well illustrated by comparing their atomization energies with exact values in two different basis sets. Both approaches show errors that are significantly larger than ‘chemical accuracy’ (1 kcal mol<sup>-1</sup>). The variational energy of the quantum trial reconstructed from experiment has a bias that can be as large as 33 kcal mol<sup>-1</sup>. The noise on our quantum device makes the quality of our quantum trial far from that of the ideal (that is, noiseless) ansatz, as shown in Fig. [2b,c](#), resulting in an error as large as 10 kcal mol<sup>-1</sup> in the atomization energy. Nonetheless, QC-AFQMC reduces this error significantly, and achieves chemical accuracy in both bases. Notably, we achieve this accuracy even in the larger basis, where the variational energy of the quantum trial in the absence of noise is far from exact.

**Table 1** Atomization energy (kcal mol<sup>-1</sup>) of H<sub>4</sub>

As shown in [Supplementary Section C](#), for the larger basis set we obtain a residual ‘virtual’ correlation energy by using the quantum resources on a smaller number of orbitals to unbias an AFQMC calculation on a larger number of orbitals, with no additional overhead to the quantum computer. This capability makes our implementation competitive with state-of-the-art classical approaches. Similar virtual

correlation energy strategies have been previously discussed within the framework of VQE<sup>39</sup>, but, unlike our approach, those strategies come with a significant measurement overhead. To unravel the QC-AFQMC results on H<sub>4</sub> further, in Fig. 2b,c we illustrate the evolution of trial and QC-AFQMC energies as a function of the number of measurements made on the device. Despite the presence of significant noise in approximately 10<sup>5</sup> measurements, QC-AFQMC achieves chemical accuracy while coping with a sizeable residual bias in the underlying quantum trial.

Next, we move to a larger example, N<sub>2</sub>, which requires a total of 12 qubits in our quantum experiment. Here a simpler quantum trial is used for QC-AFQMC by taking just the valence bond part of the wavefunction depicted in Fig. 2a. We examine the potential energy surface of N<sub>2</sub> from compressed to elongated geometries, which is another common benchmark problem for classical quantum chemistry methods<sup>35,40</sup>. In Fig. 3a, the QC-AFQMC result is shown for the calculations performed in a triple-zeta basis (cc-pVTZ) set<sup>37</sup>, which corresponds to a 60-orbital or 120-qubit Hilbert space. All examined methods, CCSD(T), AFQMC and QC-AFQMC, perform well near the equilibrium geometry, but CCSD(T) and AFQMC deviate from the exact results significantly as the bond distance is stretched. As a result, the error for ‘gold standard’ CCSD(T) can be as large as 14 kcal mol<sup>-1</sup>, and the error for AFQMC with a classical trial wavefunction can be as large as -8 kcal mol<sup>-1</sup>. The error in the QC-AFQMC computation ranges from -2 kcal mol<sup>-1</sup> to 1 kcal mol<sup>-1</sup> depending on the bond distance. Thus, although we do not achieve chemical accuracy with QC-AFQMC, we note that, even with a simple quantum trial wavefunction, we produce energies that are competitive with state-of-the-art classical approaches. Idealized (that is, noiseless) VQE experiments for the same trial wavefunction would yield similar results to our quantum trial results Fig. 3a (within 4.5 kcal mol<sup>-1</sup>), which are much worse than our QC-AFQMC results with an error as large as 50 kcal mol<sup>-1</sup>.

**Fig. 3: 12-qubit and 16-qubit experiments.**

---

 **figure 3**

**a**, Circuit layout showing spin-up and spin-down qubits for the 12-qubit experiment (top). Potential energy surface of N<sub>2</sub> in a triple-zeta basis set (cc-pVTZ<sup>37</sup>; 60 orbital) (bottom). The relative energies are shifted to zero at 2.25 Å. Inset shows the error in total energy relative to the exact results in kcal mol<sup>-1</sup>. The shaded region in the inset shows the bounds for chemical accuracy (1 kcal mol<sup>-1</sup>). Neither the variational energy of the quantum trial nor the statistical error bars of the AFQMC methods are visible on this scale. **b**, Circuit layout showing spin-up and spin-down qubits for the 16-qubit experiment (top). Error in total energy as a function of lattice constant of diamond in a double-zeta basis (DZVP-GTH; 26 orbitals) (bottom). The shaded region shows the bounds for chemical accuracy. Our quantum trial results are not visible on this scale. Inset shows a supercell structure of diamond in which two highlighted atoms form the minimal unit cell. See Extended Data Tables [9](#), [10](#) for the raw data for **a**, **b**, respectively.

[Source data](#)

Finally, we present a 16-qubit experiment result for the ground-state simulation of a minimal unit cell (two-atom) model of periodic solid diamond in a double-zeta basis set (DZVP-GTH<sup>41</sup>; 26 orbitals). Although at this level of theory the model exhibits significant finite-size effects and does not predict the correct experimental lattice constant, we aim to illustrate the utility of our algorithm in materials science

applications. We emphasize that this is the largest quantum simulation of chemistry on a quantum processor so far (detailed resource counts and comparison with prior works are available in Extended Data Tables 11, 12). We again use the simple perfect pairing state as our quantum trial wavefunction and demonstrate the improvement over a range of lattice parameters compared with classical AFQMC and CCSD(T) in Fig. 3b. There is a substantial improvement in the error going from AFQMC to QC-AFQMC, showing the increased accuracy due to better trial wavefunctions. At the same time, QC-AFQMC performed using the idealized quantum trial produces results comparable to our experimental energies, suggesting that the error in our QC-AFQMC energies is mainly due to the use of an insufficiently accurate trial wavefunction rather than experimental error. Our accuracy is limited by the simple form of our quantum trial and yet we achieve accuracy nearly on a par with the classical gold standard method, CCSD(T).

## Conclusion and outlook

In summary, we propose a scalable, noise-resilient quantum-classical hybrid algorithm that seamlessly embeds a special-purpose quantum primitive into an accurate quantum computational many-body method, namely QMC. Our work offers a computational strategy that effectively unbiases fermionic QMC approaches by leveraging state-of-the-art quantum information tools. We have realized this algorithm for a specific QMC algorithm known as AFQMC, and demonstrated its performance in experiments as large as 16 qubit on a NISQ processor, producing electronic energies that are competitive with state-of-the-art classical quantum chemistry methods. Our algorithm also enables the incorporation of the electron correlation energy outside the space that is handled by the quantum computer without increasing quantum resources or measurement overheads. In [Supplementary Section F](#), we discuss issues related to asymptotic scaling and the potential for quantum advantage in our algorithm. Although we have yet to achieve practical quantum advantage over available classical algorithms, the flexibility and scalability of our proposed approach in the construction of quantum trial functions, and its inherent noise resilience, promise a path forward for the simulation of chemistry in the NISQ era and beyond.

*Note added in proof:* After this work was nearly complete, a theory paper by Yang et al. appeared on arXiv<sup>42</sup>, describing a quantum algorithm for assisting real-time dynamics with unconstrained QMC.

## Methods

### Wavefunction overlap estimation

In this work, we estimate the overlap between the trial wavefunction and the statistical samples using a technique known as shadow tomography<sup>43,44</sup>. Experimentally, this entails performing randomly chosen measurements of a reference state related to  $\langle \Psi | \text{state}(T) \rangle$  before beginning the QMC calculation, yielding the representation of  $\langle \Psi | \text{state}(T) \rangle$  in the computational basis for subsequent overlap evaluations. In this formulation of QC-QMC, there is no need for the QMC calculation to iteratively query the quantum processor, despite the fact that the details of the statistical samples are not determined in advance. By disentangling the interaction between the quantum and classical computer we avoid feedback latency, an appealing feature on early NISQ platforms that comes at the cost of requiring potentially expensive classical post-processing (see [Supplementary Section D](#) for more details). Furthermore, our algorithm naturally achieves some degree of noise robustness, as explained in [Supplementary Section D](#), because the quantity directly used in QC-QMC is the ratio between overlap values, which is inherently resilient to the estimates of the overlaps being rescaled. We highlight the challenges posed by the need to measure wavefunction overlaps precisely and the trade-offs involved in the use of shadow tomography (see also [Supplementary Section D](#)), while giving our perspective on the most promising paths forward.

## Phaseless constraints in AFQMC

In AFQMC, the  $\langle \varphi_i(\tau) \rangle$  take the form of Slater determinants in arbitrary single-particle bases, enabling us to express the energy estimator (presented in Supplementary equation (3)) in terms of a modest number of wavefunction overlaps that we can evaluate efficiently on the quantum processor ([Supplementary Section C](#)). The phaseless constraint is imposed to control the sign problem and, likewise, only requires calculating the overlaps between  $\langle \Psi | \text{state}(T) \rangle$  and  $\langle \varphi_i(\tau) \rangle$ , as detailed in Supplementary equation (6). AFQMC has been shown to be accurate in a number of cases even with classically available trial wavefunctions<sup>45,46</sup>; however, the bias incurred from the phaseless constraint cannot be overlooked.

## Quantum processor

The experiments in this work were carried out on the Google 54-qubit quantum processor known as Sycamore<sup>47</sup>. The circuits were compiled using hardware-native conditional Z gates with typical error rates of  $\approx 0.5\%$  (ref. [48](#)).

## Data availability

The datasets generated and/or analysed during the current study are available from the corresponding authors on reasonable request. [Source data](#) are provided with this paper.

## Code availability

We used available packages such as Q-Chem <sup>49</sup> and [Cirq](#) (see <https://github.com/quantumlib/Cirq> for details on obtaining the source code); more details are available in Supplementary Section E. Other codes used herein are available from the corresponding authors on reasonable request.

## References

1. Friesner, R. A. Ab initio quantum chemistry: methodology and applications. *Proc. Natl Acad. Sci. USA* **102**, 6648–6653 (2005).
2. Helgaker, T., Klopper, W. & Tew, D. P. Quantitative quantum chemistry. *Mol. Phys.* **106**, 2107–2143 (2008).
3. Cao, Y. et al. Quantum chemistry in the age of quantum computing. *Chem. Rev.* **119**, 10856–10915 (2019).
4. Bauer, B., Bravyi, S., Motta, M. & Kin-Lic Chan, G. Quantum algorithms for quantum chemistry and quantum materials science. *Chem. Rev.* **120**, 12685–12717 (2020).
5. Acioli, P. H. Review of quantum Monte Carlo methods and their applications. *J. Mol. Struct. THEOCHEM* **394**, 75–85 (1997).
6. Foulkes, W. M. C., Mitas, L., Needs, R. J. & Rajagopal, G. Quantum Monte Carlo simulations of solids. *Rev. Mod. Phys.* **73**, 33 (2001).
7. Peruzzo, A. et al. A variational eigenvalue solver on a photonic quantum processor. *Nat. Commun.* **5**, 4213 (2014).
8. McClean, J. R., Romero, J., Babbush, R. & Aspuru-Guzik, A. The theory of variational hybrid quantum-classical algorithms. *New J. Phys.* **18**, 23023 (2016).
9. Feynman, R. P. Simulating physics with computers. *Int. J. Theor. Phys.* **21**, 467–488 (1982).
10. Lloyd, S. Universal quantum simulators. *Science* **273**, 1073–1078 (1996).

11. Shor, P. W. in *Proc. 37th Conf. on Foundations of Computer Science* 56–65 (IEEE Comput. Soc. Press, 1996).
12. Preskill, J. Quantum computing and the entanglement frontier. Preprint at <https://arxiv.org/abs/1203.5813> (2012).
13. McArdle, S. et al. Variational ansatzbased quantum simulation of imaginary time evolution. *npj Quantum Inf.* **5**, 75 (2019).
14. Motta, M. et al. Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution. *Nat. Phys.* **16**, 205–210 (2020).
15. Kandala, A. et al. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature* **549**, 242–246 (2017).
16. Nam, Y. et al. Ground-state energy estimation of the water molecule on a trapped-ion quantum computer. *npj Quantum Inf.* **6**, 33 (2020).
17. Google AI Quantum et al. Hartree-Fock on a superconducting qubit quantum computer. *Science* **369**, 1084–1089 (2020).
18. Troyer, M. & Wiese, U.-J. Computational complexity and fundamental limitations to fermionic quantum Monte Carlo simulations. *Phys. Rev. Lett.* **94**, 170201 (2005).
19. Blankenbecler, R., Scalapino, D. J. & Sugar, R. L. Monte Carlo calculations of coupled boson-fermion systems. I. *Phys. Rev. D* **24**, 2278–2286 (1981).
20. Chang, C.-C., Gogolenko, S., Perez, J., Bai, Z. & Scalettar, R. T. Recent advances in determinant quantum Monte Carlo. *Philos. Mag.* **95**, 1260–1281 (2015).
21. Mahajan, A. & Sharma, S. Taming the sign problem in auxiliary-field quantum Monte Carlo using accurate wave functions. *J. Chem. Theory Comput.* **17**, 4786–4798 (2021).
22. Li, Z.-X. & Yao, H. Sign-problem-free fermionic quantum Monte Carlo: developments and applications. *Annu. Rev. Condens. Matter Phys.* **10**, 337–356 (2019).
23. Moskowitz, J. W., Schmidt, K. E., Lee, M. A. & Kalos, M. H. A new look at correlation energy in atomic and molecular systems. II. The application of the Green’s function Monte Carlo method to LiH. *J. Chem. Phys.* **77**, 349–355 (1982).

24. Zhang, S., Carlson, J. & Gubernatis, J. E. Constrained path Monte Carlo method for fermion ground states. *Phys. Rev. B* **55**, 7464 (1997).
25. Zhang, S. & Krakauer, H. Quantum Monte Carlo method using phase-free random walks with Slater determinants. *Phys. Rev. Lett.* **90**, 136401 (2003).
26. Sorella, S. Linearized auxiliary fields Monte Carlo technique: efficient sampling of the fermion sign. *Phys. Rev. B* **84**, 241110 (2011).
27. Becca, F. & Sorella, S. *Quantum Monte Carlo Approaches for Correlated Systems* (Cambridge Univ. Press, 2017).
28. Bartlett, R. J., Kucharski, S. A. & Noga, J. Alternative coupled-cluster ansätze II. The unitary coupled-cluster method. *Chem. Phys. Lett.* **155**, 133–140 (1989).
29. Ryabinkin, I. G., Yen, T.-C., Genin, S. N. & Izmaylov, A. F. Qubit coupled cluster method: a systematic approach to quantum chemistry on a quantum computer. *J. Chem. Theory Comput.* **14**, 6317–6326 (2018).
30. Ortiz, G., Gubernatis, J. E., Knill, E. & Laamme, R. Quantum algorithms for fermionic simulations. *Phys. Rev. A* **64**, 022319 (2001).
31. Evenbly, G. & Vidal, G. Tensor network renormalization yields the multiscale entanglement renormalization ansatz. *Phys. Rev. Lett.* **115**, 200401 (2015).
32. Goddard, W. A., Dunning, T. H., Hunt, W. J. & Hay, P. J. Generalized valence bond description of bonding in low-lying states of molecules. *Acc. Chem. Res.* **6**, 368–376 (1973).
33. Cullen, J. Generalized valence bond solutions from a constrained coupled cluster method. *Chem. Phys.* **202**, 217–229 (1996).
34. Paldus, J., Piecuch, P., Pylypow, L. & Jeziorski, B. Application of Hilbert-space coupled-cluster theory to simple ( $H_2$ )<sub>2</sub> model systems: planar models. *Phys. Rev. A* **47**, 2738–2782 (1993).
35. Lee, J., Huggins, W. J., Head-Gordon, M. & Birgitta Whaley, K. Generalized unitary coupled cluster wave functions for quantum computation. *J. Chem. Theory Comput.* **15**, 311–324 (2019).
36. Hehre, W. J., Stewart, R. F. & Pople, J. A. Self-consistent molecular-orbital methods. I. Use of Gaussian expansions of Slater-type atomic orbitals. *J. Chem. Phys.* **51**, 2657–2664 (1969).

37. Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* **90**, 1007–1023 (1989).
38. Raghavachari, K., Trucks, G. W., Pople, J. A. & Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* **157**, 479–483 (1989).
39. Takeshita, T. et al. Increasing the representation accuracy of quantum simulations of chemistry without extra quantum resources. *Phys. Rev. X* **10**, 011004 (2020).
40. Siegbahn, P. E. M. The externally contracted CI method applied to N<sub>2</sub>. *Int. J. Quantum Chem.* **23**, 1869–1889 (1983).
41. VandeVondele, J. & Hutter, J. Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases. *J. Chem. Phys.* **127**, 114105 (2007).
42. Yang, Y., Lu, B.-N. & Li, Y. Accelerated quantum Monte Carlo with mitigated error on noisy quantum computer. *PRX Quantum* **2**, 040361 (2021).
43. Aaronson, S. Shadow tomography of quantum states. *SIAM J. Comput.* **49**, STOC18-368–STOC18-394 (2020).
44. Huang, H.-Y., Kueng, R. & Preskill, J. Predicting many properties of a quantum system from very few measurements. *Nat. Phys.* **16**, 1050–1057 (2020).
45. Zheng, B.-X. et al. Stripe order in the underdoped region of the two-dimensional Hubbard model. *Science* **358**, 1155–1160 (2017).
46. Arute, F. et al. Quantum supremacy using a programmable superconducting processor. *Nature* **574**, 505–510 (2019).
47. Williams, K. T. et al. Direct comparison of many-body methods for realistic electronic hamiltonians. *Phys. Rev.* **10**, 011041 (2020).
48. Chen, Z. et al. Exponential suppression of bit or phase errors with cyclic error correction. *Nature* **595**, 383–387 (2021).
49. Epifanovsky, E. et al. Software for the frontiers of quantum chemistry: an overview of developments in the Q-Chem 5 package. *J. Chem. Phys.* **155**, 084801 (2021).

## Acknowledgements

We thank members of the Google Quantum AI theory team and F. Malone for helpful discussions. J.L. and D.R.R. acknowledge the support of NSF CHE-1954791. B.O. is supported by a NASA Space Technology Research Fellowship and the NSF QLCI program through grant number OMA-2016245. The quantum hardware used for this experiment was developed by the Google Quantum AI hardware team, under the direction of A. Megrant, J. Kelly and Y. Chen. Theoretical foundations for device calibrations were provided by the physics team lead by V. Smelyanskiy. Initial data collection was enabled by cloud access to these devices as part of Google Quantum AI's Quantum Computing Service Early Access Program. P. Roushan and C. Neill from the Google team helped to execute the experiment on hardware and design figures.

## Author information

### Affiliations

1. Google Quantum AI, Mountain View, CA, USA

William J. Huggins, Nicholas C. Rubin, Ryan Babbush & Joonho Lee

2. Berkeley Quantum Information & Computation Center, University of California, Berkeley, CA, USA

Bryan A. O'Gorman

3. Department of Chemistry, Columbia University, New York, NY, USA

David R. Reichman & Joonho Lee

### Contributions

J.L. conceived the quantum-classical hybrid QMC algorithm, performed QMC calculations and, with contribution from others, drafted the manuscript. W.J.H. proposed the use of shadow tomography and designed the experiment, with contributions from others. B.O. helped with theoretical analysis and the compilation of circuits. N.C.R. helped with the presentation of figures. J.L. and R.B. managed the scientific collaboration. All authors participated in discussions, writing the manuscript and analysis of the data. J.L. and W.J.H. contributed equally to this work.

### Corresponding authors

Correspondence to [William J. Huggins](#) or [Joonho Lee](#).

## Ethics declarations

### Competing interest

The authors declare no competing interests.

### Peer review

#### Peer review information

*Nature* thanks Cornelius Hempel, Brenda Rubenstein, James Whitfield and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

### Additional information

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### Extended data figures and tables

**Extended Data Table 1 Experimental data of variational energy for  $H_4$  in STO-3G with partitioned tomography**

**Extended Data Table 2 Experimental data of variational energy for  $H_4$  in STO-3G with unpartitioned shadow tomography**

**Extended Data Table 3 Experimental data of variational energy for  $H_4$  in cc-pVQZ with partitioned tomography**

**Extended Data Table 4 Experimental data of variational energy for  $H_4$  in STO-3G with unpartitioned tomography**

**Extended Data Table 5 Experimental data of QC-AFQMC energy for  $H_4$  in STO-3G with partitioned shadow tomography**

**Extended Data Table 6 Experimental data of QC-AFQMC energy for  $H_4$  in STO-3G with unpartitioned shadow tomography**

**Extended Data Table 7 Experimental data of QC-AFQMC energy for  $H_4$  in cc-pVQZ with partitioned shadow tomography**

**Extended Data Table 8 Experimental data of QC-AFQMC energy for  $H_4$  in cc-pVQZ with unpartitioned shadow tomography**

**Extended Data Table 9** Raw data for N<sub>2</sub> potential energy surface for seven bond distances ( $R$ )

**Extended Data Table 10** Raw data for the diamond cold curve for five lattice constants ( $R$ )

**Extended Data Table 11** Resource counts for the QC-AFQMC experiments realized in this work

**Extended Data Table 12** Resource estimates from prior fermionic simulations using gate-model quantum computers on more than four qubits

## Supplementary information

### Supplementary Information

Supplementary text, equations, tables and references

### Peer Review File

## Source data

### Source Data Fig. 2

### Source Data Fig. 3

## Rights and permissions

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

### Reprints and Permissions

## About this article

## Cite this article

Huggins, W.J., O’Gorman, B.A., Rubin, N.C. *et al.* Unbiasing fermionic quantum Monte Carlo with a quantum computer. *Nature* **603**, 416–420 (2022).  
<https://doi.org/10.1038/s41586-021-04351-z>

- Received: 14 July 2021
- Accepted: 14 December 2021
- Published: 16 March 2022
- Issue Date: 17 March 2022
- DOI: <https://doi.org/10.1038/s41586-021-04351-z>

## Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04351-z>

| [Section menu](#) | [Main menu](#) |

- Article
- [Published: 16 March 2022](#)

# Steady Floquet–Andreev states in graphene Josephson junctions

- [Sein Park<sup>1 na1</sup>](#),
- [Wonjun Lee<sup>1,2 na1</sup>](#),
- [Seong Jang<sup>1</sup>](#),
- [Yong-Bin Choi<sup>1</sup>](#),
- [Jinho Park<sup>1</sup>](#),
- [Woochan Jung<sup>1</sup>](#),
- [Kenji Watanabe](#) ORCID: [orcid.org/0000-0003-3701-8119<sup>3</sup>](https://orcid.org/0000-0003-3701-8119),
- [Takashi Taniguchi](#) ORCID: [orcid.org/0000-0002-1467-3105<sup>4</sup>](https://orcid.org/0000-0002-1467-3105),
- [Gil Young Cho](#) ORCID: [orcid.org/0000-0003-4957-3013<sup>1,2,5</sup>](https://orcid.org/0000-0003-4957-3013) &
- [Gil-Ho Lee](#) ORCID: [orcid.org/0000-0002-7619-8979<sup>1,5</sup>](https://orcid.org/0000-0002-7619-8979)

*Nature* volume 603, pages 421–426 (2022)

- 1476 Accesses
- 13 Altmetric
- [Metrics details](#)

## Subjects

- [Electronic properties and materials](#)
- [Quantum mechanics](#)
- [Superconducting devices](#)

## Abstract

Engineering quantum states through light–matter interaction has created a paradigm in condensed-matter physics. A representative example is the Floquet–Bloch state, which is generated by time-periodically driving the Bloch wavefunctions in crystals. Previous attempts to realize such states in condensed-matter systems have been limited by the transient nature of the Floquet states produced by optical pulses<sup>1,2,3</sup>, which masks the universal properties of non-equilibrium physics. Here we report the generation of steady Floquet–Andreev states in graphene Josephson junctions by continuous microwave application and direct measurement of their spectra by superconducting tunnelling spectroscopy. We present quantitative analysis of the spectral characteristics of the Floquet–Andreev states while varying the phase difference of the superconductors, the temperature, the microwave frequency and the power. The oscillations of the Floquet–Andreev-state spectrum with phase difference agreed with our theoretical calculations. Moreover, we confirmed the steady nature of the Floquet–Andreev states by establishing a sum rule of tunnelling conductance<sup>4</sup>, and analysed the spectral density of Floquet states depending on Floquet interaction strength. This study provides a basis for understanding and engineering non-equilibrium quantum states in nanodevices.

This is a preview of subscription content

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

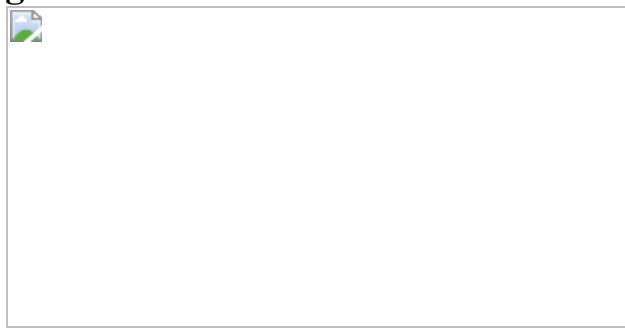
[Buy](#)

All prices are NET prices.

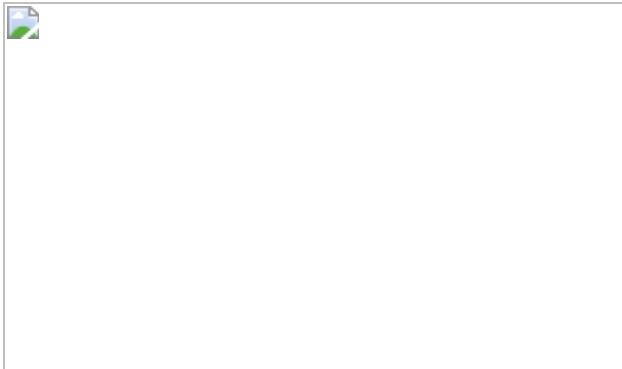
### **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

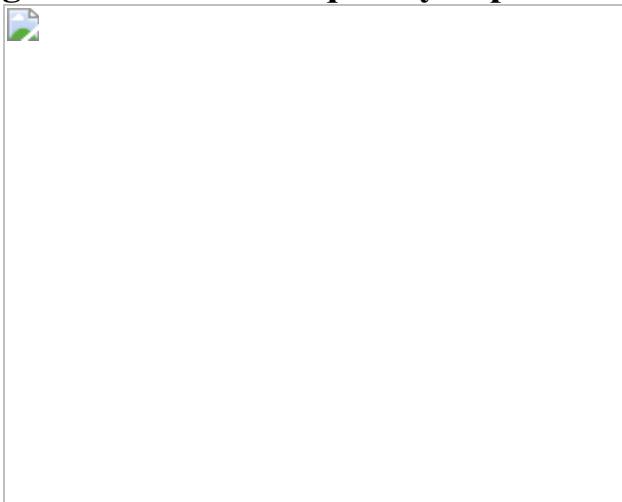
**Fig. 1: Schematics of Andreev bound state and device geometry.**



**Fig. 2: Microwave power dependence of F-A states.**



**Fig. 3: Phase and frequency dependence of Floquet–Bloch states.**



## Data availability

The data supporting the findings of this study are available from the corresponding authors upon reasonable request.

## References

1. Wang, Y. H., Steinberg, H., Jarillo-Herrero, P. & Gedik, N. Observation of Floquet–Bloch states on the surface of a topological insulator. *Science* **342**, 453–457 (2013).
2. Mahmood, F. et al. Selective scattering between Floquet–Bloch and Volkov states in a topological insulator. *Nat. Phys.* **12**, 306–310 (2016).

3. Mciver, J. W. et al. Light-induced anomalous Hall effect in graphene. *Nat. Phys.* **16**, 38–41 (2020).
4. Uhrig, G. S., Kalthoff, M. H. & Freericks, J. K. Positivity of the spectral densities of retarded Floquet Green functions. *Phys. Rev. Lett.* **122**, 130604 (2019).
5. Fausti, D. et al. Light-induced superconductivity in a stripe-ordered cuprate. *Science* **331**, 189–191 (2011).
6. Matsunaga, R. et al. Light-induced collective pseudospin precession resonating with Higgs mode in a superconductor. *Science* **345**, 1145–1149 (2014).
7. Rudner, M. S. & Lindner, N. H. Band structure engineering and non-equilibrium dynamics in Floquet topological insulators. *Nat. Rev. Phys.* **2**, 229–244 (2020).
8. Jiang, L. et al. Majorana fermions in equilibrium and in driven cold-atom quantum wires. *Phys. Rev. Lett.* **106**, 220402 (2011).
9. Bauer, B. et al. Topologically protected braiding in a single wire using Floquet Majorana modes. *Phys. Rev. B* **100**, 041102(R) (2019).
10. Clark, L. W. et al. Interacting Floquet polaritons. *Nature* **571**, 532–536 (2019).
11. Wintersperger, K. et al. Realization of an anomalous Floquet topological system with ultracold atoms. *Nat. Phys.* **16**, 1058–1063 (2020).
12. Freericks, J. K., Krishnamurthy, H. R. & Pruschke, T. Theoretical description of time-resolved photoemission spectroscopy: application to pump-probe experiments. *Phys. Rev. Lett.* **102**, 136401 (2009).
13. D'alessio, L. & Rigol, M. Long-time behavior of isolated periodically driven interacting lattice systems. *Phys. Rev. X* **4**, 041048 (2014).

14. Abanin, D. A., De Roeck, W., Ho, W. W. & Huvaneers, F. Effective Hamiltonians, prethermalization, and slow energy absorption in periodically driven many-body systems. *Phys. Rev. B* **95**, 014112 (2017).
15. Mori, T., Ikeda, T. N., Kaminishi, E. & Ueda, M. Thermalization and prethermalization in isolated quantum systems: a theoretical overview. *J. Phys. B* **51**, 112001 (2018).
16. Ponte, P., Chandran, A., Papić, Z. & Abanin, D. A. Periodically driven ergodic and many-body localized quantum systems. *Ann. Phys.* **353**, 196–204 (2015).
17. Deng, C., Orgiazzi, J.-L., Shen, F., Ashhab, S. & Lupascu, A. Observation of Floquet states in a strongly driven artificial atom. *Phys. Rev. Lett.* **115**, 133601 (2015).
18. Fuchs, G. D., Dobrovitski, V. V., Toyli, D. M., Heremans, F. J. & Awschalom, D. D. Gigahertz dynamics of a strongly driven single quantum spin. *Science* **326**, 1520–1522 (2009).
19. Koski, J. V. et al. Floquet spectroscopy of a strongly driven quantum dot charge qubit with a microwave resonator. *Phys. Rev. Lett.* **121**, 043603 (2018).
20. Jamali, S. et al. Floquet spin states in OLEDs. *Nat. Commun.* **12**, 465 (2021).
21. Huang, K.-F. et al. Interference of Cooper quartet Andreev bound states in a multi-terminal graphene-based Josephson junction. Preprint at <https://arxiv.org/abs/2008.03419> (2020).
22. Melin, R., Danneau, R., Yang, K., Caputo, J. G. & Doucot, B. Engineering the Floquet spectrum of superconducting multiterminal quantum dots. *Phys. Rev. B* **100**, 035450 (2019).
23. Melin, R., Caputo, J. G., Yang, K. & Doucot, B. Simple Floquet–Wannier–Stark–Andreev viewpoint and emergence of low-energy

- scales in a voltage-biased three-terminal Josephson junction. *Phys. Rev. B* **95**, 085415 (2017).
24. Nichele, F. et al. Relating Andreev bound states and supercurrents in hybrid Josephson junctions. *Phys. Rev. Lett.* **124**, 226801 (2020).
  25. Bretheau, L. et al. Tunnelling spectroscopy of Andreev states in graphene. *Nat. Phys.* **13**, 756–760 (2017).
  26. Pillet, J. D. et al. Andreev bound states in supercurrent-carrying carbon nanotubes revealed. *Nat. Phys.* **6**, 965–969 (2010).
  27. Giazotto, F., Peltonen, J. T., Meschke, M. & Pekola, J. P. Superconducting quantum interference proximity transistor. *Nat. Phys.* **6**, 254–259 (2010).
  28. Lee, G.-H., Kim, S., Jhi, S. H. & Lee, H.-J. Ultimately short ballistic vertical graphene Josephson junctions. *Nat. Commun.* **6**, 6181 (2015).
  29. Choi, D. H. et al. Van-der-Waals-gap tunneling spectroscopy for single-wall carbon nanotubes. *Carbon* **113**, 237–242 (2017).
  30. Le Sueur, H., Joyez, P., Pothier, H., Urbina, C. & Esteve, D. Phase controlled superconducting proximity effect probed by tunneling spectroscopy. *Phys. Rev. Lett.* **100**, 197002 (2008).
  31. Viljas, J. K. & Heikkila, T. T. Electron-phonon heat transfer in monolayer and bilayer graphene. *Phys. Rev. B* **81**, 245454 (2010).
  32. Walsh, E. D. et al. Graphene-based Josephson-junction single-photon detector. *Phys. Rev. Appl.* **8**, 024022 (2017).
  33. Tikhonov, K. S., Skvortsov, M. A. & Klapwijk, T. M. Superconductivity in the presence of microwaves: full phase diagram. *Phys. Rev. B* **97**, 184516 (2018).
  34. Kitagawa, T., Berg, E., Rudner, M. & Demler, E. Topological characterization of periodically driven quantum systems. *Phys. Rev. B* **82**, 235114 (2010).

35. Dehghani, H., Oka, T. & Mitra, A. Out-of-equilibrium electrons and the Hall conductance of a Floquet topological insulator. *Phys. Rev. B* **91**, 155422 (2015).
36. He, Q. L. et al. Chiral Majorana fermion modes in a quantum anomalous Hall insulator–superconductor structure. *Science* **357**, 294–299 (2017).
37. Verdeny, A., Puig, J. & Mintert, F. Quasi-periodically driven quantum systems. *Z. Naturforsch. A* **71**, 897–907 (2016).
38. Crowley, P. J. D., Martin, I. & Chandran, A. Topological classification of quasiperiodically driven quantum systems. *Phys. Rev. B* **99**, 064306 (2019).
39. Sentef, M. A. et al. Theory of Floquet band formation and local pseudospin textures in pump-probe photoemission of graphene. *Nat. Commun.* **6**, 7047 (2015).
40. Wang, L. et al. One-dimensional electrical contact to a two-dimensional material. *Science* **342**, 614–617 (2013).
41. Beenakker, C. W. J. Colloquium: Andreev reflection and Klein tunneling in graphene. *Rev. Mod. Phys.* **80**, 1337–1354 (2008).
42. Titov, M. & Beenakker, C. W. J. Josephson effect in ballistic graphene. *Phys. Rev. B* **74**, 041401(R) (2006).
43. Beenakker, C. W. Specular Andreev reflection in graphene. *Phys. Rev. Lett.* **97**, 067007 (2006).
44. Bardeen, J. Tunnelling from a many-particle point of view. *Phys. Rev. Lett.* **6**, 57–59 (1961).
45. Dynes, R. C., Narayanamurti, V. & Gurno, J. P. Direct measurement of quasiparticle-lifetime broadening in a strong-coupled superconductor. *Phys. Rev. Lett.* **41**, 1509–1512 (1978).

46. Semenov, A. V., Devyatov, I. A., De Visser, P. J. & Klapwijk, T. M.  
Coherent excited states in superconductors due to a microwave field.  
*Phys. Rev. Lett.* **117**, 047002 (2016).

## Acknowledgements

We thank H.-J. Lee, K. W. Kim, J. C. W. Song, D. Cho, K. C. Fong and C. Lee for reading the manuscript, and C. B. Winkelmann for a discussion about the Tien–Gordon model. S.P., S.J., Y.-B.C. and G.-H.L. acknowledge the support of the Samsung Science and Technology Foundation (project number SSTF-BA1702-05) for device fabrications and low-temperature measurements. W.J. and G.-H.L. acknowledge the support of National Research Foundation of Korea (NRF) funded by the Korean Government (grant numbers 2016R1A5A1008184, 2020R1C1C1013241 and 2020M3H3A1100839) for data analysis. W.L. and G.Y.C. acknowledge the support of the National Research Foundation of Korea (NRF) funded by the Korean Government (grant numbers 2020R1C1C1006048 and 2020R1A4A3079707), as well as grant number IBS-R014-D1. W.L. and G.Y.C. are also supported by the Air Force Office of Scientific Research under award number FA2386-20-1-4029. K.W. and T.T. acknowledge support from the Elemental Strategy Initiative conducted by the MEXT, Japan, grant number JPMXP0112101001 and JSPS KAKENHI grant number JP20H00354.

## Author information

### Author notes

1. These authors contributed equally: Sein Park, Wonjun Lee

## Affiliations

1. Department of Physics, Pohang University of Science and Technology, Pohang, Republic of Korea

Sein Park, Wonjun Lee, Seong Jang, Yong-Bin Choi, Jinho Park, Woochan Jung, Gil Young Cho & Gil-Ho Lee

2. Center for Artificial Low Dimensional Electronic Systems, Institute for Basic Science (IBS), Pohang, Republic of Korea

Wonjun Lee & Gil Young Cho

3. Research Center for Functional Materials, National Institute for Materials Science, Tsukuba, Japan

Kenji Watanabe

4. International Center for Materials Nanoarchitectonics, National Institute for Materials Science, Tsukuba, Japan

Takashi Taniguchi

5. Asia Pacific Center for Theoretical Physics, Pohang, Republic of Korea

Gil Young Cho & Gil-Ho Lee

## Contributions

G.-H.L. and G.Y.C. conceived and supervised the project. S.P. designed and fabricated the devices. T.T. and K.W. provided the hBN crystal. S.P., S.J. and Y.-B.C. performed the measurements. W.L. and G.Y.C. carried out theoretical calculations. S.P., W.L., S.J., J.P., W.J., G.Y.C. and G.-H.L. performed the data analysis. S.P., W.L., S.J., G.Y.C. and G.-H.L. wrote the paper.

## Corresponding authors

Correspondence to [Gil Young Cho](#) or [Gil-Ho Lee](#).

## Ethics declarations

## **Competing interests**

The authors declare no competing interests.

## **Peer review**

### **Peer review information**

*Nature* thanks Romain Danneau, James McIver and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

## **Additional information**

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## **Supplementary information**

### **Supplementary Information**

This file contains Supplementary text, figures, tables, equations and references.

### **Peer Review File**

## **Rights and permissions**

### **Reprints and Permissions**

## **About this article**

### **Cite this article**

Park, S., Lee, W., Jang, S. *et al.* Steady Floquet–Andreev states in graphene Josephson junctions. *Nature* **603**, 421–426 (2022).  
<https://doi.org/10.1038/s41586-021-04364-8>

- Received: 30 April 2021
- Accepted: 16 December 2021
- Published: 16 March 2022
- Issue Date: 17 March 2022
- DOI: <https://doi.org/10.1038/s41586-021-04364-8>

## Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

## [Long-lived non-equilibrium quantum states could lead to improved electronic devices](#)

Research Briefing 16 Mar 2022

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04364-8>

- Article
- [Published: 16 March 2022](#)

# Wind dispersal of battery-free wireless devices

- [Vikram Iyer](#) [ORCID: orcid.org/0000-0002-3025-7953<sup>1</sup>](#),
- [Hans Gaensbauer<sup>2</sup>](#),
- [Thomas L. Daniel<sup>3</sup>](#) &
- [Shyamnath Gollakota](#) [ORCID: orcid.org/0000-0002-9863-3054<sup>1,2</sup>](#)

[Nature](#) volume **603**, pages 427–433 (2022)

- 2011 Accesses
- 300 Altmetric
- [Metrics details](#)

## Subjects

- [Computer science](#)
- [Electrical and electronic engineering](#)

## Abstract

Plants cover a large fraction of the Earth's land mass despite most species having limited to no mobility. To transport their propagules, many plants have evolved mechanisms to disperse their seeds using the wind<sup>1,2,3,4</sup>. A dandelion seed, for example, has a bristly filament structure that decreases its terminal velocity and helps orient the seed as it wafts to the ground<sup>5</sup>. Inspired by this, we demonstrate wind dispersal of battery-free wireless

sensing devices. Our millimetre-scale devices weigh 30 milligrams and are designed on a flexible substrate using programmable, off-the-shelf parts to enable scalability and flexibility for various sensing and computing applications. The system is powered using lightweight solar cells and an energy harvesting circuit that is robust to low and variable light conditions, and has a backscatter communication link that enables data transmission. To achieve the wide-area dispersal and upright landing that is necessary for solar power harvesting, we developed dandelion-inspired, thin-film porous structures that achieve a terminal velocity of  $0.87 \pm 0.02$  metres per second and aerodynamic stability with a probability of upright landing of over 95%. Our results in outdoor environments demonstrate that these devices can travel 50–100 metres in gentle to moderate breeze. Finally, in natural systems, variance in individual seed morphology causes some seeds to fall closer and others to travel farther. We adopt a similar approach and show how we can modulate the porosity and diameter of the structures to achieve dispersal variation across devices.

This is a preview of subscription content

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

## Subscribe

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

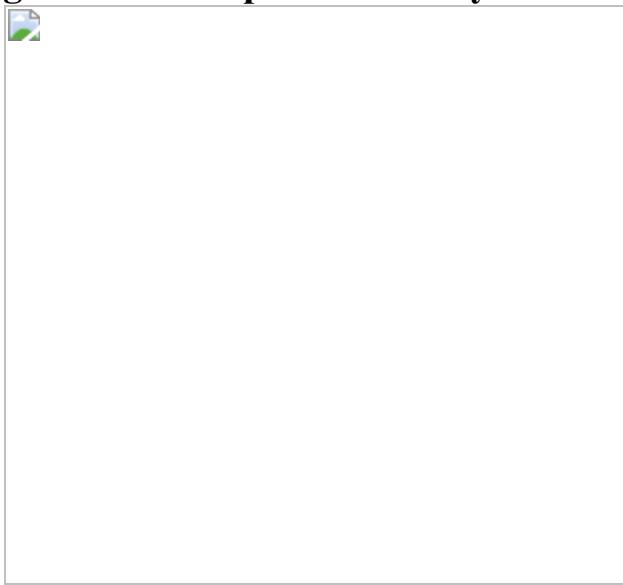
Buy

All prices are NET prices.

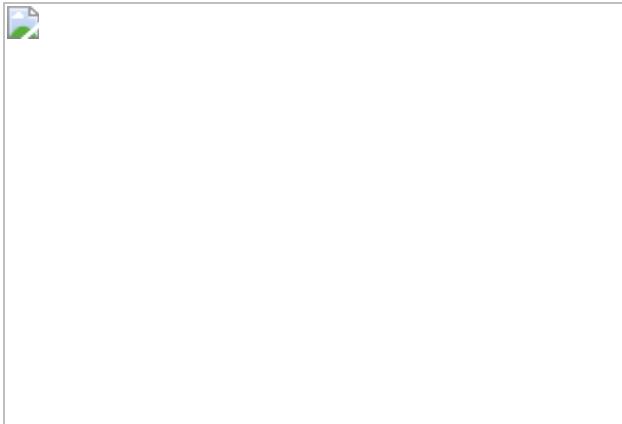
## **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

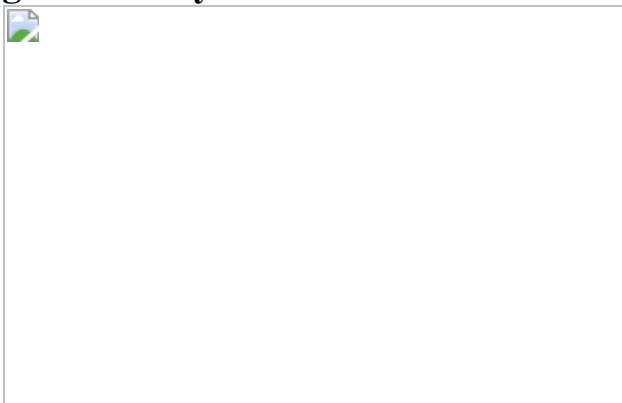
**Fig. 1: Wind-dispersed battery-free wireless devices.**



**Fig. 2: Wind-dispersal mechanism.**



**Fig. 3: Battery-free wireless sensor.**



**Fig. 4: Outdoor evaluation.**

## Data availability

All data needed to evaluate the conclusions of the paper are available in the paper or in the Extended Data and Supplementary Information.

## Code availability

Source code is available on a public GitHub repository for download:  
[https://github.com/uw-x/wind\\_dispersal/](https://github.com/uw-x/wind_dispersal/).

## References

1. Cummins, C. et al. A separated vortex ring underlies the flight of the dandelion. *Nature* **562**, 414–418 (2018).
2. Lentink, D., Dickson, W. B., van Leeuwen, J. L. & Dickinson, M. H. Leading-edge vortices elevate lift of autorotating plant seeds. *Science* **324**, 1438–1440 (2009).
3. Greene, D. F. The role of abscission in long-distance seed dispersal by the wind. *Ecology* **86**, 3105–3110 (2005).
4. Greene, D. F. & Johnson, E. A. The aerodynamics of plumed seeds. *Funct. Ecol.* **4**, 117–125 (1990).
5. Sheldon, J. C. & Burrows, F. M. The dispersal effectiveness of the achene–pappus units of selected Compositae in steady winds with convection. *New Phytol.* **72**, 665–675 (1973).
6. Andersen, M. C. Diaspore morphology and seed dispersal in several wind-dispersed Asteraceae. *Am. J. Bot.* **80**, 487–492 (1993).
7. Casseau, V., De Croon, G., Izzo, D. & Pandolfi, C. Morphologic and aerodynamic considerations regarding the plumed seeds of *Tragopogon pratensis* and their implications for seed dispersal. *PLoS ONE* **10**, e0125040 (2015).
8. Augspurger, C. K. & Franson, S. E. Wind dispersal of artificial fruits varying in mass, area, and morphology. *Ecology* **68**, 27–42 (1987).

9. Andersen, M. C. An analysis of variability in seed settling velocities of several wind-dispersed Asteraceae. *Am. J. Bot.* **79**, 1087–1091 (1992).
10. Edwards, D. J. et al. CICADA flying circuit board unmanned aerial vehicle. In *2018 AIAA Aerospace Sciences Meeting* 1008 (AIAA, 2018).
11. Pounds, P. & Singh, S. Samara: biologically inspired self-deploying sensor networks. *IEEE Potentials* **34**, 10–14 (2015).
12. Iyer, V., Kim, M., Xue, S., Wang, A. & Gollakota, S. Airdropping sensor networks from drones and insects. In *MobiCom '20: Proc. 26th Annu. Intl Conf. Mobile Computing and Networking* 61 (ACM, 2020).
13. Vogel, S. *Life in Moving Fluids: The Physical Biology of Flow* 2nd edn (Princeton Univ. Press, 2020).
14. Seo, D. et al. Wireless recording in the peripheral nervous system with ultrasonic neural dust. *Neuron* **91**, 529–539 (2016).
15. Kahn, J. M., Katz, R. H. & Pister, K. S. J. Next century challenges: mobile networking for “Smart Dust”. In *MobiCom '99: Proc. 5th Annu. Intl Conf. Mobile Computing and Networking* (eds Kodesh, H. et al.) 271–278 (ACM, 1999).
16. Lee, Y. et al. A modular 1 mm<sup>3</sup> die-stacked sensing platform with low power I<sup>2</sup>C inter-die communication and multi-modal energy harvesting. *IEEE J. Solid-State Circ.* **48**, 229–243 (2013).
17. Chen, Y. et al. Energy-autonomous wireless communication for millimeter scale Internet-of-Things sensor nodes. *IEEE J. Sel. Area Commun.* **34**, 3962–3977 (2016).
18. Chuo, L.-X. et al. A 915 MHz asymmetric radio using *Q*-enhanced amplifier for a fully integrated 3 × 3 × 3 mm<sup>3</sup> wireless sensor node with 20 m non-line-of-sight communication. In *IEEE Intl Solid-State Circuits Conf.* 132–133 (IEEE, 2017).

19. Iyer, V., Najafi, A., James, J., Fuller, S. & Gollakota, S. Wireless steerable vision for live insects and insect-scale robots. *Sci. Robotics* **5**, abb0839 (2020).
20. Rostami, M., Sundaresan, K., Chai, E., Rangarajan, S. & Ganesan, D. Redefining passive in backscattering with commodity devices. *MobiCom '20: Proc. 26th Annu. Intl Conf. Mobile Computing and Networking* 3 (ACM, 2020).
21. Iyer, V., Nandakumar, R., Wang, A., Fuller, S. B. & Gollakota, S. Living IoT: a flying wireless platform on live insects. In *MobiCom '19: 25th Annu. Intl Conf. Mobile Computing and Networking* (eds Agarwal, S. et al.) 5 (ACM, 2019).
22. Talla, V. et al. LoRa backscatter: enabling the vision of ubiquitous connectivity. In *Proc. ACM Interactive, Mobile, Wearable and Ubiquitous Technologies* 105 (ACM, 2017).
23. Kim, B. H. et al. Three-dimensional electronic microfliers inspired by wind-dispersed seeds. *Nature* **597**, 503–510 (2021).
24. Zussman, E., Yarin, A. & Weihs, D. A micro-aerodynamic decelerator based on permeable surfaces of nanofiber mats. *Exp. Fluids* **33**, 315–320 (2002).
25. Lentink, D., Dickson, W. B., van Leeuwen, J. L. & Dickinson, M. H. Leading-edge vortices elevate lift of autorotating plant seeds. *Science* **324**, 1438–1440 (2009).
26. James, J., Iyer, V., Chukewad, Y., Gollakota, S. & Fuller, S. B. Liftoff of a 190 mg laser-powered aerial vehicle: the lightest wireless robot to fly. In *2018 IEEE Intl Conf. Robotics and Automation (ICRA)* (ed. Lynch, K.) 3587–3594 (IEEE, 2018).
27. Katanbaf, M., Weinand, A. & Talla, V. Simplifying backscatter deployment: full-duplex LoRa backscatter. In *18th Symp. Networked Systems Design and Implementation* 955–972 (USENIX, 2021).

28. Shen, S. et al. An LC passive wireless gas sensor based on PANI/CNT composite. *Sensors* **18**, 3022 (2018).
29. Nandakumar, R., Iyer, V. & Gollakota, S. 3D Localization for Sub-Centimeter Sized Devices. In *SenSys '18: Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems* (eds Zhang, L. et al.) 108–119 (ACM 2018).
30. Riley, J. R. et al. Tracking bees with harmonic radar. *Nature* **379**, 29–30 (1996).

## Acknowledgements

We thank S. Kaplan and J. Smith for feedback on the manuscript and K. Johnson, V. Arroyos and B. Nguyen for help with experiments. This research is funded by Moore Inventor Fellow award #10617 and grant FA9550-14-1-0398 from the US Air Force Office of Scientific Research. S.G. is also funded by the National Science Foundation.

## Author information

### Affiliations

1. Paul G. Allen School of Computer Science and Engineering,  
University of Washington, Seattle, WA, USA  
  
Vikram Iyer & Shyamnath Gollakota
2. Department of Electrical and Computer Engineering, University of  
Washington, Seattle, WA, USA  
  
Hans Gaensbauer & Shyamnath Gollakota
3. Department of Biology, University of Washington, Seattle, WA, USA  
  
Thomas L. Daniel

## Contributions

V.I. designed and fabricated the hardware and performed experiments. S.G. and V.I. designed the overall system, experiments and evaluation. T.D. advised on the design and characterization of the drag-enhancing structures. H.G. implemented the backscatter code and sensor release mechanism used for outdoor testing. V.I. and S.G. wrote the manuscript; T.D. edited the manuscript. S.G. and V.I conceptualized the work.

## Corresponding authors

Correspondence to [Vikram Iyer](#) or [Shyamnath Gollakota](#).

## Ethics declarations

## Competing interests

S.G. is a co-founder of Jeeva Wireless, Wavely Diagnostics, and Sound Life Sciences.

## Peer review

## Peer review information

*Nature* thanks Elizabeth Farrell Helbing and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

## Additional information

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Extended data figures and tables

## Extended Data Fig. 1 Effect of outer ring on terminal velocity.

**a**, Terminal velocities of 26% fill structures with different diameters using a 30-mg payload. **b**, Terminal velocities of 55% fill structures with different diameters with and without an outer ring using a 30-mg payload.

## Extended Data Fig. 2 Additional fill patterns.

Additional fill patterns used to vary porosity. Changes include concentric circles connecting horizontal spokes extending out from the centre with various spacings, as well as additional cuts in the concentric circles between spokes.

## Extended Data Fig. 3 Fabrication process.

Fabrication process used to produce the curved drag enhancing structures.

## Extended Data Fig. 4 SX1276 receiver on–off keying packet structure.

**a**, Packet structure for the SX1276 receiver operating in on–off keying mode. **b**, Minimal packet structure used for testing the lower bit rates of 1.5 and 4.8 kbps. CRC, cyclic redundancy check.

## Extended Data Fig. 5 Supporting transmissions from multiple devices.

**a**, Natural variance of start-up time for outdoor measurements. **b**, Two waveforms recorded under similar light conditions show that natural variation in start-up time and time between packets reduces the probability of collisions. **c**, Addition of a random delay between each packet to reduce collisions. **d**, Network simulation of up to 1,000 nodes each transmit 100 packets. Node start and inter-packet transmission times are normally distributed based on power-harvesting data ( $N = 100$  packets per node  $\times 10$  repetitions = 1,000). **e**, Network simulation with normally distributed start-up time assuming 1.4 klx followed by fixed 10-s delay between packets ( $N$

$= 1000$ ). **f**, Result of applying pseudorandom delays of up to 8 packet lengths with random start times assuming 1.4 klx ( $N = 1,000$ ). The plot shows that this approach can support large networks where the data are transmitted at rates less than once every 10 s.

### **Extended Data Fig. 6 Additional power harvesting results.**

**a**, Cold-start time required to charge the storage capacitor from zero to the time at which it can backscatter its first packet; measured using indoor LED lights ( $N \geq 5$ , error bars  $= \pm\sigma$ ). **b**, Throughput of the sensor device after start up at different illumination levels indoors ( $N \geq 5$ , error bars  $= \pm\sigma$ ). **c**, Measurements of power harvested at different light levels outdoors from our solar cells. ( $N \geq 3$ , error bars  $= \pm\sigma$ ).

### **Extended Data Fig. 7 Drone mechanism to drop the sensors.**

**a**, Small quadrotor with the mechanism used to drop devices from altitudes over 5 m outdoors. **b**, Drop mechanism consisting of a solenoid that can be wirelessly triggered by a Bluetooth transmission. Upon triggering, the supercapacitor keeps the solenoid retracted for a few seconds allowing the wire holding the sensor in place to fall freely. Because our evaluation requires tracking each individual structure, the drone attachment currently drops a single device. It can be extended to drop multiple devices simultaneously using a similar mechanism.

### **Extended Data Fig. 8 Sensor calibration set-up using a weather station.**

**a**, Picture of weather station used for reference temperature, humidity, solar radiation and pressure measurements, and placement of the test wireless sensor node. **b**, Correlation of small onboard light sensor readings with reference solar radiation data. **c**, Correlation of small, onboard temperature sensor (HDC2010) readings with reference solar radiation data. **d**, Correlation of onboard temperature sensor (BME280) readings with reference solar radiation data.

## **Extended Data Fig. 9 Prototype full-duplex cancellation of our narrowband signals.**

**a**, Full-duplex cancellation block diagram. **b**, Results of applying full-duplex cancellation to reduce transmitter interference for backscatter.

## **Extended Data Table 1 Sensor power measurements and part numbers**

## **Supplementary information**

### **Supplementary Information**

This file contains three Supplementary Tables and their accompanying legends; and the legends for Supplementary Videos 1–6.

### **Supplementary Video 1**

Vertical wind tunnel test.

### **Supplementary Video 2**

Effect of outer ring.

### **Supplementary Video 3**

Effect of stiffness.

### **Supplementary Video 4**

Unstable descent of solid discs.

### **Supplementary Video 5**

Self-righting behaviour.

## Supplementary Video 6

Outdoor wind dispersal.

## Rights and permissions

### Reprints and Permissions

## About this article

### Cite this article

Iyer, V., Gaensbauer, H., Daniel, T.L. *et al.* Wind dispersal of battery-free wireless devices. *Nature* **603**, 427–433 (2022).  
<https://doi.org/10.1038/s41586-021-04363-9>

- Received: 09 December 2020
- Accepted: 16 December 2021
- Published: 16 March 2022
- Issue Date: 17 March 2022
- DOI: <https://doi.org/10.1038/s41586-021-04363-9>

### Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04363-9>

| [Section menu](#) | [Main menu](#) |

- Article
- Open Access
- [Published: 16 March 2022](#)

# Flat-surface-assisted and self-regulated oxidation resistance of Cu(111)

- [Su Jae Kim](#) ORCID: [orcid.org/0000-0002-6033-8551<sup>1</sup>](https://orcid.org/0000-0002-6033-8551)<sup>nal</sup>,
- [Yong In Kim](#) ORCID: [orcid.org/0000-0002-9540-5996<sup>2</sup>](https://orcid.org/0000-0002-9540-5996)<sup>nal</sup>,
- [Bipin Lamichhane](#) ORCID: [orcid.org/0000-0002-9503-083X<sup>3</sup>](https://orcid.org/0000-0002-9503-083X)<sup>nal</sup>,
- [Young-Hoon Kim](#) ORCID: [orcid.org/0000-0001-7343-1512<sup>2</sup>](https://orcid.org/0000-0001-7343-1512),
- [Yousil Lee<sup>1</sup>](#),
- [Chae Ryong Cho](#) ORCID: [orcid.org/0000-0002-6685-4012<sup>4</sup>](https://orcid.org/0000-0002-6685-4012),
- [Miyeon Cheon<sup>1</sup>](#),
- [Jong Chan Kim](#) ORCID: [orcid.org/0000-0003-4101-0590<sup>5</sup>](https://orcid.org/0000-0003-4101-0590),
- [Hu Young Jeong](#) ORCID: [orcid.org/0000-0002-5550-5298<sup>6</sup>](https://orcid.org/0000-0002-5550-5298),
- [Taewoo Ha](#) ORCID: [orcid.org/0000-0002-4089-1982<sup>7</sup>](https://orcid.org/0000-0002-4089-1982),
- [Jungdae Kim](#) ORCID: [orcid.org/0000-0002-8567-1529<sup>8</sup>](https://orcid.org/0000-0002-8567-1529),
- [Young Hee Lee](#) ORCID: [orcid.org/0000-0001-7403-8157<sup>2,7,9</sup>](https://orcid.org/0000-0001-7403-8157),
- [Seong-Gon Kim](#) ORCID: [orcid.org/0000-0002-1629-0319<sup>3</sup>](https://orcid.org/0000-0002-1629-0319),
- [Young-Min Kim](#) ORCID: [orcid.org/0000-0003-3220-9004<sup>2,7</sup>](https://orcid.org/0000-0003-3220-9004) &
- [Se-Young Jeong](#) ORCID: [orcid.org/0000-0003-1019-4403<sup>10,11</sup>](https://orcid.org/0000-0003-1019-4403)

[Nature](#) volume 603, pages 434–438 (2022)

- 2093 Accesses
- 17 Altmetric
- [Metrics details](#)

## Subjects

- [Surfaces, interfaces and thin films](#)

## Abstract

Oxidation can deteriorate the properties of copper that are critical for its use, particularly in the semiconductor industry and electro-optics applications<sup>1,2,3,4,5,6,7</sup>. This has prompted numerous studies exploring copper oxidation and possible passivation strategies<sup>8</sup>. In situ observations have, for example, shown that oxidation involves stepped surfaces: Cu<sub>2</sub>O growth occurs on flat surfaces as a result of Cu adatoms detaching from steps and diffusing across terraces<sup>9,10,11</sup>. But even though this mechanism explains why single-crystalline copper is more resistant to oxidation than polycrystalline copper, the fact that flat copper surfaces can be free of oxidation has not been explored further. Here we report the fabrication of copper thin films that are semi-permanently oxidation resistant because they consist of flat surfaces with only occasional mono-atomic steps. First-principles calculations confirm that mono-atomic step edges are as impervious to oxygen as flat surfaces and that surface adsorption of O atoms is suppressed once an oxygen face-centred cubic (fcc) surface site coverage of 50% has been reached. These combined effects explain the exceptional oxidation resistance of ultraflat Cu surfaces.

[Download PDF](#)

## Main

Given that the step edge is vulnerable to oxidation because surface steps act as the dominant source of Cu adatoms for oxide growth on surface terraces<sup>2,11</sup>, oxidation resistance requires that surface step edges are avoided<sup>6,12,13</sup>. In this regard, the close-packed Cu(111) surface is superior to other Cu surfaces<sup>14,15</sup> and our experimental demonstration thus used a single-crystal Cu(111) film (SCCF) grown by atomic sputtering epitaxy (ASE)<sup>16</sup> to show that a tightly coordinated flat surface can remain semi-permanently stable against oxidation. Theoretical calculations show the atomic-scale oxidation-resistant mechanism of the flat copper surface from the perspective of the likely pathways for oxygen atoms into the viable structures of the copper surface with a discovery of the self-regulated protection layer at elevated oxygen coverages. The implication is that the atomically flat surface of the SCCF shows oxidation-resistant properties owing to the high energy barrier for oxygen infiltration and self-regulation owing to high oxygen coverages.

## Flatness of Cu(111) surface

The surface and structural characteristics of a 110-nm-thick SCCF with an ultraflat surface (see Extended Data Fig. 1 for a large-scale characterization) are examined using high-resolution (scanning) transmission electron microscopy (HR(S)TEM) combined with geometrical phase analysis (GPA)<sup>17,18</sup> (Fig. 1). The cross-sectional (S)TEM images (Fig. 1a,e) show that the copper film grew along the [111] direction,

thus creating an exposed surface (111) plane with mono-atomic step-edge structures. Typical multi-atomic step edges and intrinsic defects such as grain boundaries and stacking faults are rarely detected. It is remarkable that the outermost copper surface layer has the same atomic configuration as the interior copper without evidence of surface relaxation or structural changes by surface oxidation, even at the step-edge positions. To examine the local strain behaviour near the surface region, lattice distortions along the in-plane ( $x$ ) and out-of-plane ( $y$ ) directions relative to the inside of the SCCF were measured by the GPA technique (Fig. 1b,c). The resulting strain field maps ( $E_{xx}$  and  $E_{yy}$ ) clearly show that no noticeable change in lattice strain is observed throughout the entire surface region. This means that the SCCF has a nearly perfect atomic structure up to its outermost surface layer without any structural defects, such as vacancies or dislocations. The simulated HRTEM image using an amorphous carbon/flat copper surface model matches well with the experimental HRTEM image (Fig. 1a). By comparing the layer spacings of the (111) stacking planes ( $d_{(111)} = 0.21$  nm) between the simulated and experimental images (Fig. 1d), it is evident that the Cu surface is undistorted and ultraflat, and has the same structure as bulk Cu. Annular dark-field (ADF) and annular bright-field (ABF) STEM images of the SCCF surface are complementary to the HRTEM observation results (Fig. 1e). We can discriminate this remarkable ultraflat surface structure of the SCCF from that of the conventional polycrystalline Cu thin film (PCCF) and the Cu(111) surface of a bulk Cu single crystal with a notable oxidized copper layer on top of them (Extended Data Fig. 2 and Fig. 3). Notably, we found that the SCCF sample maintained its ultraflat and pristine surface even after more than a year of air exposure (Fig. 2a-d), suggesting that our SCCF has exceptional oxidation-resistant properties (also see Extended Data Fig. 4 for a large-scale scanning tunnelling microscope (STM) topography).

**Fig. 1: Surface of a single-crystal copper thin film grown by ASE.**



figure 1

**a**, Cross-sectional HRTEM image of the surface region of the copper thin film observed in the  $\langle\overline{1}\overline{1}0\rangle$  orientation. The simulated HRTEM image and the corresponding model of the carbon–copper supercell are presented alongside the experimental HRTEM image. Scale bar is 2 nm. **b, c**, In-plane ( $E_{xx}$ ) and out-of-plane ( $E_{yy}$ ) strain field maps obtained by GPA for the experimental HRTEM image. The reference area of Cu chosen for GPA is marked by the white square boxes in each map. Note that complex patterns for parts of the carbon film are shaded by a grey colour for clarity, as those are not relevant for the strain behaviour of the SCCF. Scale bars are 2 nm. **d**, Comparison of the experimental and simulated intensity profiles obtained along the out-of-plane direction for (111) layer spacings ( $d_{(111)} = 0.21$  nm). **e**, Simultaneously acquired ADF-STEM and ABF-STEM images for the surface region of the copper thin film. Scale bars are 2 nm.

[Source data](#)

**Fig. 2: Long-term oxidation resistance and structural stability of the SCCF surface.**

 [figure 2](#)

**a**, Low-magnification BF-TEM image of the SCCF sample exposed to ambient air at room temperature for about 1 year (top). The result shows that the atomically flat surface morphology over the entire SCCF film has remained almost unchanged. HRTEM image of the surface region (marked by a red rectangle in the top panel) of the 1-year-old SCCF sample (bottom). The sample was oriented in the  $\langle [1\overline{1}0] \rangle$  direction. Scale bars are 100 nm (top) and 2 nm (bottom). **b**, Comparison of the intensity profiles for the (111) plane between the as-deposited (red) and 1-year-old (blue) samples. **c**, EBSD map showing perfect alignment along the (111) plane. Scale bar is 3  $\mu$ m. **d**, IPF with a sole spot associated with the (111) plane. The inset image is the enlarged image of the sole-spot area. **e**, HRTEM image of the surface region of the 3-year-old SCCF sample observed in the  $\langle [1\overline{1}0] \rangle$  orientation (top) and in-plane ( $E_{xx}$ , middle) and out-of-plane ( $E_{yy}$ , bottom) strain field maps obtained by GPA for the HRTEM image of the 3-year-old SCCF sample. Scale bars are 2 nm. **f**, FFT patterns of region 1 (top), region 2 (middle) and both regions (bottom). **g**, Thermogravimetric analysis for the PCCF and SCCF samples with different surface roughness. r.m.s., root mean square.

### Source data

The low-magnification bright-field TEM (BF-TEM) image (Fig. 2a, top) and the cross-sectional HRTEM image (Fig. 2a, bottom) show that the mono-atomic step-edge structure is maintained even after a year. The two intensity profiles for the (111) plane between the as-deposited (red) and the 1-year-old (blue) samples (Fig. 2b) show a change in the planar spacing of the (111) planes up to the uppermost surface layer, and they agree well with each other, suggesting that no notable oxidation has occurred even in the SCCF under long-term exposure to air. The electron backscatter diffraction (EBSD) map (Fig. 2c) and inverse pole figure (IPF) (Fig. 2d) show that no misalignment of crystal lattices deviating from the (111) plane has occurred after a year. Oxidized portions were rare on the 3-year-old SCCF sample oriented along  $\langle [1\overline{1}0] \rangle$  in the HRTEM image, although they were found on the sample edge where the sample was cut (Fig. 2e, top). The resulting strain field maps ( $E_{xx}$  and  $E_{yy}$ ) prepared using the GPA technique (Fig. 2e, middle and bottom) show that the overlayer (region 1) has a new lattice structure that is mismatched with the SCCF. The fast Fourier transform (FFT) patterns of region 1 (Fig. 2f, top), region 2 (Fig. 2f, middle) and both regions (Fig. 2f, bottom) indicate that the two regions belong to the Cu<sub>2</sub>O and Cu phases, respectively, and suggest that part of the surface is oxidized. However, the oxidized surface only reaches a few Cu<sub>2</sub>O layers, which is thin compared with the thickness of natural oxide layers in polycrystalline Cu (Extended Data Fig. 2). Comparative thermogravimetric analysis of PCCF and SCCF samples having different surface roughness clearly demonstrates that the SCCF with mono-atomic step edges shows exceptional resistance to initial oxidation at elevated temperatures compared with the other samples (Fig. 2g).

## Growth condition for the flat surface

The flatness of a surface is decisively influenced by the interface structure between the film and the substrate<sup>19,20</sup>, which can be relaxed by structural defects such as dislocations. The interface structure between the Cu film and the Al<sub>2</sub>O<sub>3</sub> substrate is characterized by HR(S)TEM (Fig. 3). The overall interface structure viewed at the  $\langle 11\bar{2} \rangle$  orientation of the Cu film and the FFT pattern of the imageshow that the crystallographic orientation relationship (OR) is  $\langle (111) \rangle_{Cu} / \langle (001) \rangle_{Al} / \langle (110) \rangle_{Al}$  (Fig. 3a,b). The copper lattices seem to flawlessly adjoin with the Al<sub>2</sub>O<sub>3</sub> substrate without interfacial misfit defects, suggesting that the Cu film grows metamorphically on the substrate<sup>21</sup> (Extended Data Fig. 5). The enlarged image of the interface (Fig. 3c) shows a detailed lattice mismatch between the two materials. The in-plane atomic distance mismatch  $(f) = (\langle d \rangle_{Cu} - \langle d \rangle_{O}) / \langle d \rangle_{O}$  between Cu atoms in the film and O atoms in Al<sub>2</sub>O<sub>3</sub> is estimated to be 6.9%. However, considering the extended atomic distance mismatch (EADM)<sup>22,23</sup>, the large mechanical misfit strain can be relieved if the mismatch for a relatively long period of atomic distance is extremely small. The EADM is defined as  $(ID - I'D') / ID'$ , in which  $D$  and  $D'$  are the Cu–Cu distance in the Cu (111) epilayer and the O–O distance in the substrate, respectively, and  $I$  and  $I'$  are smallest non-reducible integers determined by the relation  $D:D' \sim I:I'$ . Given the  $D_{Cu-Cu} = (14 \times \langle d \rangle_{Cu} - \langle d \rangle_{O})^{[110]}$  interatomic spacing (3.578 nm) of Cu atoms in the film and the  $\langle D' \rangle = (13 \times \langle d \rangle_{O} - \langle d \rangle_{O})^{[110]}$  interatomic spacing (3.575 nm) of O atoms in Al<sub>2</sub>O<sub>3</sub>, the EADM of the Cu–Al<sub>2</sub>O<sub>3</sub> interface is about 0.1% (Fig. 3c). The light-element-sensitive ABF-STEM imaging<sup>24</sup> (Fig. 3d) shows that the interface model of Cu grown on the oxygen-terminated Al<sub>2</sub>O<sub>3</sub> surface matches well with the experimental heterostructure. The inverted intensity profile obtained across the interface (Fig. 3e) clearly corroborates the presence of an oxygen layer between the Cu and Al layers, thus indicating the existence of Cu–O interactions at the interface, which can be stabilized on a typical Al<sub>2</sub>O<sub>3</sub> surface terminated with oxygen<sup>25</sup>. Owing to the in-plane lattice mismatch, a large-scale interference pattern, that is, a moiré pattern<sup>26</sup>, can be observed in the plane view of the Cu–Al<sub>2</sub>O<sub>3</sub> heterostructure (Fig. 3f). Indeed, the hexagonal moiré pattern with a dimension ( $d_m$ ) of 1.83 nm is observed owing to the different in-plane lattice periodicities in the vertical OR of  $[111]_{Cu}/[001]_{Al_2O_3}$ , which is confirmed by the FFT pattern analysis (inset in Fig. 3f). The simulated moiré pattern

generated by the epitaxy model of Cu–Al<sub>2</sub>O<sub>3</sub> with the same OR is consistent with the experimental moiré pattern, showing the repeating large-scale contrast feature (Fig. 3g, h). This vertical OR observation corroborates our EADM analysis and indicates that the growth mechanism of the SCCF on the Al<sub>2</sub>O<sub>3</sub> substrate can be understood on the basis of the large-scale mismatch epitaxial relationship, rather than the atomic-scale lattice interrelation. The detailed chemical nature of the SCCF at the surface and the interface regions are investigated by the combined spectroscopic approaches of energy-dispersive X-ray spectroscopy (EDX), electron energy loss spectroscopy (EELS) in ADF-STEM imaging mode and X-ray photoelectron spectroscopy<sup>27,28</sup> (Extended Data Fig. 6).

**Fig. 3: Interface structure and crystallographic relation.**

---

 **figure 3**

**a**, Cross-sectional HRTEM image of the Cu–Al<sub>2</sub>O<sub>3</sub> heterointerface having the orientation relation of  $\langle\{111\}\rangle_{\rm Cu} \langle\overline{112}\rangle_{\rm Cu} // \langle\{001\}\rangle_{\rm Al} \langle\{2\}\rangle_{\rm O} \langle\{3\}\rangle_{\rm O}$ . Scale bar is 2 nm. **b**, A composite pattern of FFTs for the regions of Cu and Al<sub>2</sub>O<sub>3</sub> in the HRTEM image. **c**, Enlarged HRTEM image for the region marked by the dashed box in **a**. Scale bar is 1 nm. **d**, ABF-STEM image and superimposed atomic model of the Cu–Al<sub>2</sub>O<sub>3</sub> heterointerface. Scale bar is 5 Å. **e**, Inverted intensity profile obtained across the interface between Cu and Al<sub>2</sub>O<sub>3</sub> in **d**. **f**, Plan-view HRTEM image of the Cu–Al<sub>2</sub>O<sub>3</sub> heterointerface. The inset is the FFT pattern of the HRTEM image. Scale bar is 2 nm. **g, h**, Side view

(g) and plan view (h) of the deduced epitaxy model of Cu grown on an Al<sub>2</sub>O<sub>3</sub> substrate. The yellow diamonds with unit length of  $d_m$  in f and h indicate the 2D unit cell of the moiré pattern.

## Calculation of oxidation resistance

We can understand the exceptional oxidation resistance of our films using a microscopic model of copper oxidation on the basis of first-principles density functional theory (DFT) calculations<sup>29,30,31</sup>. The main reason for the suppression of oxidation is that our atomically flat film is free of one critical feature, namely, multi-atomic step edges, as shown in Fig. 4. The energy profile in Fig. 4a shows that the penetration of an O atom through a pristine Cu(111) surface needs an activation energy of more than 1.4 eV (Fig. 4b), and the O atom becomes stable only in the second subsurface interlayer space with a further energy barrier of 1.3 eV (Fig. 4c). One of the main reasons for the strong resistance of a flat surface to oxidation is that, when Cu atoms are oxidized, the out-of-plane distance between Cu layers increases from 2.10 to 2.48 Å, and the volume expands by 18% (Extended Data Fig. 7). The volume increase is even more pronounced in the early stage of oxidation: the Cu layer distance in the monolayer of Cu<sub>2</sub>O on the Cu surface is 3.26 Å, and the volume is greater by 55% (Extended Data Fig. 7). Given that one exposed Cu surface is insufficient to initiate the oxidation process, we examine a structure wherein two crystallographically different planes meet, namely, the edge of a multi-atomic step. Notably O<sub>2</sub> molecules physisorbed on a Cu(111) surface readily dissociate into O ions with a small activation energy of 0.027 eV (Extended Data Fig. 8). Figure 4d–f shows the initiation of oxidation on the edge of such multi-atomic steps according to our DFT calculation; Fig. 4a shows that there is a critical difference between a mono-atomic step and a multi-atomic step in terms of the oxidation resistance. Whereas the penetration of O atoms at a bi-atomic and tri-atomic step is an exothermic reaction ( $\Delta E = -0.90$  and  $-0.86$  eV, respectively) with smaller activation energies ( $E_a = 0.97$  and  $0.71$  eV, respectively), the penetration of O at the edge of a mono-atomic step is highly endothermic ( $\Delta E = +0.80$  eV), requiring a much larger activation energy ( $E_a = 1.56$  eV), and shows strong oxidation resistance. Our DFT calculation indicates that, as the coverage of oxygen increases, the incremental adsorption energy for the next O atom becomes smaller and eventually becomes negative beyond 50% of oxygen coverage, rendering the adsorption process energetically unfavourable (Fig. 4g). This self-regulation of oxygen suppresses further adsorption of O atoms on an ultraflat Cu surface at elevated oxygen coverages and increasingly enhances the oxidation resistance of the surface. Our microscopic model of oxidation is well supported by HRTEM images, which show that atomically flat Cu thin films with occasional presence of mono-atomic steps are highly resistant to oxidation over an extended period ( $\geq 3$  years) (Fig. 2), whereas Cu films with multi-atomic steps show substantial

oxidation (Extended Data Figs. 2 and 3). The oxidation resistance of a Cu(111) surface is greatly affected by the type of surface defects<sup>32,33,34,35</sup>, which suggests that an atomically flat Cu(111) surface without multi-atom steps is essential to achieve strong oxidation resistance.

**Fig. 4: Theoretical analysis and a model of oxidation of the copper surface.**

 figure 4

**a–f**, Energy profile of an O atom along various penetration paths (a): from the Cu(111) surface into the first Cu substrate interstitial layer (black solid circles) (b); from the first to the second Cu substrate interstitial layer (black open circles) (c); from the outside to the inside of the mono-atomic step edge (red solid circles) (d); from the outside to the inside of the bi-atomic step edge (blue open squares) (e); and from the outside to the inside of the tri-atomic step edge (pink open squares) (f). Blue spheres represent Cu atoms in the bulk or substrate, green spheres represent Cu atoms in the steps and dark blue spheres represent Cu atoms on the step edge. Orange spheres represent adsorbed O atoms and red spheres represent O atoms infiltrating into the interstitial region. Panel f also shows the initiation of oxidation on the edge of a multi-atomic step: (1) given a pristine step edge and ambient oxygen conditions, (2) two O atoms (orange spheres) adsorb on each side of the Cu on the edge (Cu-1, dark blue sphere); (3) the third adsorbed O atom (O-2, red sphere) causes the local structure of Cu-1 to be similar to that of Cu in a monolayer of  $\langle\{\rm{Cu}\}\rangle_2\langle\rm{O}\rangle$ , and the expansion of volume causes Cu-1 to move upward and open a pathway for O-2; and (4) O-2 passes through the opening and binds with Cu atoms in the next row to push those Cu atoms upward to sustain the oxidation process. **g**, Incremental adsorption energy of O atom as a function of oxygen coverage of the fcc sites on the Cu(111) surface. The red and blue shading in a and g represents the endothermic and exothermic reactions, respectively.

[Source data](#)

# Methods

## ASE for the preparation of single-crystal Cu thin films

Cu thin films were grown as nearly defect-free and grain-boundary-free single crystals using the ASE technique achieved by addressing the problems of conventional sputter systems<sup>16</sup>. Because the ASE system controlling atomic-level growth is very sensitive to environmental factors, it requires three key instrumental modifications compared with a commercially available sputtering system. First, we used a single-crystal target instead of a polycrystalline one; second, all conducting wires of the wiring network were replaced with single-crystal copper wires; finally, a mechanical noise-reduction system was installed to suppress the mechanical vibrations from the surroundings. The idea of this technology is on the basis of completely eliminating the noise caused by electron-grain boundary scattering in the conduction network in the device, which, to our knowledge, has been completely ignored until now. Once the system was appropriately constructed, we confirmed that the quality of the Cu films was greatly improved, with a high level of reproducibility. The schematic diagram of the ASE system with the three key features is shown in Extended Data Fig. 9a and the detailed descriptions are as follows.

### Single-crystal Cu sputtering target

For a polycrystalline target with a high density of surface steps and grain boundaries, atoms on the edges or grain boundaries have a lower binding energy than that of atoms arranged on a flat plane of the grain. Hence the atoms at the structural defects are prone to being sputtered as atomic clusters for the RF power set to remove atoms from a flat plane, which will eventually be deposited on the substrate as randomly oriented clusters. To realize the practicability of ASE growth, the use of a single-crystal Cu target with a (111) surface is essential because Cu atoms are sputtered from the target as individual atoms. Thus the uniform stacking of Cu atoms on the substrate is empirically achieved for the growth of an ultraflat film. The single-crystal Cu(111) target can be obtained from the single-crystal Cu ingot grown by the Czochralski method by cutting using wire electrical discharge machining as a 6-mm-thick disc with a 2-inch diameter (see the two images on the upper-left side of Extended Data Fig. 9a). Even though Cu single-crystal ingots are commercially available, we grew them using our own apparatus in this study, and there were no differences between the two in the resulting improvement of film quality.

### Electrical noise reduction using single-crystal Cu wires

To reduce electrical noise interference, we replaced the electrical networks made of conducting wires in conventional sputtering systems with single-crystal Cu wires as much as possible (see the three images on the lower-left side of Extended Data Fig. 9a). The effectiveness of this modification was previously demonstrated in a Hall measurement kit with circuitry and connecting components made of single-crystal Cu, notably improving the measurement precision of the electrical coefficients, such as carrier density and mobility<sup>36</sup>. Single-crystal Cu wires can be prepared by the wire electrical discharge machining cutting of a single-crystal Cu disc in a spiral manner, as reported in our previous study<sup>37</sup>. To ensure further reduction of electrical noise, we replaced a typical RF power cable with a single crystal. To monitor how effectively the single-crystal power cable improves the RF power stability in the sputtering system, we measured the change in the RF power over time with and without the single-crystal Cu power cable (Extended Data Fig. 10a,b). Single-crystal wiring has already been adapted in the sputtering system used in the test. Therefore it is evident that the output RF power is more stabilized in a narrower power range ( $12.61 \pm 0.005$  W) after changing the original power cable to a single crystal. Although the RF power stability of the sputtering system before modification is also of good quality compared with the conventional ones, we empirically confirmed that such a level of stability is not sufficient for the growth of the Cu film with an atomic-level flatness.

### Mechanical noise reduction using mechanical diodes

Although mechanical noise is not the main source of interference in the ASE system, the electrical noise reduction on the basis of the single-crystal Cu wiring cannot be effective unless the mechanical noise is effectively screened. After the tests to reduce the mechanical noise with several choices, including absorbers, barriers, vibration isolators and vibration dampers, we found that the application of a mechanical diode consisting of a set of metal spikes and pads is very effective and economical in shielding the mechanical interferences transmitted through the wall and the floor. As depicted in Extended Data Fig. 9a, we designed mechanical diodes and installed them on every device, including the chamber and the vacuum pumps. The growth of the atomically flat metal films was not successful without this shield against mechanical noise.

To verify the reproducibility of our ASE approach for the growth of the ultraflat Cu(111) film, we measured the root mean square (r.m.s.) roughness of many samples as a function of the film thickness and compared these values with those of the samples grown by a conventional sputtering system equipped with only the single-crystal Cu target (Extended Data Fig. 10c). The averaged r.m.s. value for the 29 samples grown by the ASE system was estimated to be around  $0.20 \pm 0.1$  nm (red dashed line), which is similar to the planar spacing of Cu(111). Notably, it can be further decreased to about  $0.17 \pm 0.1$  nm (purple dashed line) when considering thinner

Cu films below 35 nm, suggesting their reliable applications to ultrathin electronic devices. However, the averaged r.m.s. value for the samples grown by the conventional sputtering system equipped with the single-crystal Cu target was around  $0.66 \pm 0.2$  nm, which is good but not enough for the growth of the ultraflat Cu films with only the mono-atomic steps.

### Optimized sputtering conditions using the ASE system

A double-side polished (001) Al<sub>2</sub>O<sub>3</sub> wafer with a thickness of 430  $\mu\text{m}$  was used as the substrate material. The optimized deposition temperature and RF (13.56 MHz in frequency) power were about 170 °C and 30 W, respectively, and varied slightly, depending on the ASE systems. The target-to-substrate distance was set at 95 mm. The base pressure was maintained at under  $2 \times 10^{-7}$  torr and the working pressure at  $5.4 \times 10^{-3}$  torr with an Ar gas flow of 50 sccm. Ar gas with a purity of 99.9999% (6N) was used. The relationship between the deposition time and the thickness of the thin film (or the average growth rate) was determined from the average deposition time of a 200-nm-thick film grown under the optimum conditions. The determined average growth rate of roughly 4.3 nm min<sup>-1</sup> is fairly reliable above a film thickness of 10 nm. The kinetic energy of sputtered Cu atoms depends on the incident ion energy of Ar<sup>+</sup> and the binding energy of Cu atoms at the surface of the target. The crystallographically different surfaces of Cu have different surface binding energies ( $E_{\text{b,Cu}}$ ), which were reported to be 4.62, 4.26 and 4.65 eV for the Cu(100), Cu(110) and Cu(111) planes, respectively<sup>38</sup>. Considering the potential energy (\(\{E\}\_{\{\{\rm{Ar}\}\}}^{\{+\}}=15\{\rm{eV}\}\)) of the accelerated Ar<sup>+</sup> ion at the maximum current of 1 A in our ASE equipment, the kinetic energy of the sputtered Cu atoms can be narrowly distributed at around 10.35 eV, which is roughly calculated from \(\{E\}\_{\{\{\rm{Ar}\}\}}^{\{+\}}-\{E\}\_{\{\{\rm{b}\},\{\rm{Cu}\}(111)\}}\)) in the case of the Cu(111) single-crystal target used in this study. By contrast, in the case of polycrystalline Cu target dominantly having a mixture of Cu(100), Cu(110) and Cu(111) exposed planes at the surface, the kinetic energy of the sputtered Cu atoms is expected to be spread between 10.35 and 10.74 eV. When considering the surface defects and grain boundaries at the target, the kinetic energy would be distributed more widely up to 11.52 eV (ref. <sup>39</sup>). The radial distribution of the incident flux of the sputtered Cu atoms at the substrate is known to determine the uniformity of the deposition thickness of Cu film<sup>40</sup>. To check the thickness uniformity of the grown SCCF, we measured the thickness at five different points from centre to edge in a 2-inch wafer using an atomic force microscope (AFM). As a result, the thickness uniformity was estimated to be about 99.8% (Extended Data Fig. [9b,c](#)). This result suggests that the diffusive flux of Cu atoms is purely uniform at the position of the substrate. Note that the substrate of our system is rotated at 30 rpm.

## Structural and chemical characterizations

X-ray diffraction  $\theta$ - $2\theta$  measurements were performed using a PANalytical Empyrean Series 2 instrument equipped with a Cu  $K\alpha$  source (40 kV, 30 mA). Data were collected in the range  $20^\circ < 2\theta < 90^\circ$ , with a step size of  $0.0167^\circ$  and a dwell time of 0.5 s per point in all cases. AFM measurements were carried out using anXE-100 system (Park Systems, Inc.). Scanning electron microscopy, EBSD, pole figure (PF) and IPF measurements were performed with a Zeiss Supra 40VP with a scanning electron microprobe. An STM surface analysis was conducted using a custom-built STM system installed at the University of Ulsan, Korea. Electron-transparent cross-sectional TEM samples were prepared by the Ga ion beam milling and lift-out method in focused ion beam systems (FIB, Helios NanoLab 450, FEI and AURIGA CrossBeam Workstation, Carl Zeiss) and the possible damaged surface layers on the samples were removed by subsequent low-energy Ar ion beam surface milling at 700 eV for 15 min (Model 1040 NanoMill, Fischione). The plan-view TEM specimen was prepared by mechanical polishing and dimple grinding, followed by ion milling with Ar ions. Double  $C_s$ -corrected (S)TEM systems (JEM-ARM200CF, JEOL) equipped with EELS (Quantum ER965, Gatan) and EDX (JED-2300T, JEOL) were used for atomic-scale structure imaging and chemical analysis of the samples at an accelerating voltage of 200 kV. The inner and outer angle ranges for ADF-STEM and ABF-STEM imaging were 45–180 and 12–24 mrad, respectively. The HRTEM, ADF-STEM and ABF-STEM images were denoised by a local 2D difference image filter that is implemented in commercial software (HREM-Filters Pro, HREM Research Inc.). The HRTEM simulation was carried out for the amorphous carbon/copper \([1\overline{1}0]\) supercell structure ( $4.8 \times 4.8 \text{ nm}^2$ ) using the multislice method, which is implemented in the commercial software MacTempas (Total Resolution LLC), with the following microscope and imaging parameters: accelerating voltage ( $V = 200 \text{ kV}$ ), spherical aberration coefficient ( $C_s = 0.4 \mu\text{m}$ ), chromatic aberration coefficient ( $C_c = 1.1 \text{ mm}$ ), convergence semi-angle ( $\alpha = 0.5 \text{ mrad}$ ), sample thickness ( $t = 20.5 \text{ nm}$ ) and defocus ( $\Delta f = +14 \text{ nm}$ ). The simulated HRTEM image was estimated to have a correlation of 0.98 with the experimental HRTEM image as a cross-correlation factor. For the quantitative analysis of local strain components in the copper thin film, the GPA technique was used, which allows mapping two-dimensional local displacement fields by analysing the phase shift between non-collinear Fourier components of the lattice vectors  $g_1$  and  $g_2$ . For EELS measurements of the Cu–Al<sub>2</sub>O<sub>3</sub> interface and the Cu surface, the core-loss EELS spectra of the Cu  $L$  edge were obtained from the interface to the Cu surface using the line scan function of the scanning step (0.72 nm) for a 37.83-nm length with an energy dispersion of 0.5 eV pix<sup>-1</sup> and a dwell time of 2.0 s pix<sup>-1</sup>. For core-loss EELS spectrum imaging for the comparison of surface structures between the SCCF and the PCCF samples, surface regions of both samples were selected by  $20 \times 16$  pixels that can be translated as 9.87

$\times 7.9 \text{ nm}^2$  and scanned with the step size of 0.493 nm to obtain EELS spectrum imaging dataset. The selected range of energy loss was set to be 477–988 eV including O *K* and Cu *L*<sub>2,3</sub> edges. Nanoscale STEM-EDX maps of the constituent elements of the Cu–Al<sub>2</sub>O heterostructure were obtained for a 256 × 256 pixel resolution with a high-efficiency dual silicon drift detector X-ray detector system having a wide collection window of 100 mm<sup>2</sup> for each detector, and the sample drift during the acquisition was corrected by tracking the reference area assigned at the acquisition setup.

## Thermogravimetric analysis for the PCCF and SCCF samples with different surface roughness

The thermal mass change was measured using a thermogravimetric measurement system (TG-DTA 2000S, MAC Science). We prepared two SCCF samples with different values of surface roughness and a PCCF sample. The gravimetric changes of the three samples were measured in the temperature range from room temperature to 700 °C at a heating rate of 20 °C min<sup>-1</sup> under air atmosphere. The two SCCF samples with different values of r.m.s. roughness, that is, 0.3 nm (SCCF-0.3, corresponding to a mono-atomic step) and 0.7 nm (SCCF-0.7, corresponding to a bi-atomic or tri-atomic step), were tested to ascertain the effect of the surface steps on oxidation in comparison with the PCCF sample (blue) (r.m.s. roughness around 10 nm). Although the SCCF sample with a roughness of 0.3 nm (red) was synthesized by ASE and had an ultraflat surface with occasional mono-atomic steps, the SCCF sample with a roughness of 0.7 nm (green), which was synthesized using a single-crystal target in a conventional sputtering system, had occasional multi-atomic steps of 2 or 3 atomic layers in height.

## Theoretical calculations

All ab initio total energy calculations and geometry optimizations were performed with DFT in the generalized gradient approximation Perdew–Burke–Ernzerhof functional<sup>29</sup> and with the projected augmented-plane-wave method<sup>30</sup>, as implemented by Kresse and Joubert<sup>31</sup>. The Cu substrate was represented by slabs of six layers with the theoretical equilibrium lattice constant. A vacuum length of 15 Å was used, and the bottom two layers of the slab were fixed in their bulk positions. The electron wave functions were expanded in a plane-wave basis set with a cut-off energy of 400 eV. The Brillouin zone for the slabs was sampled using *k*-point sets equivalent to at least a (5 × 5 × 1) Monkhorst–Pack grid<sup>41</sup> for the primitive fcc unit cell. The climbing image-nudged elastic band method<sup>42</sup> was used to calculate activation energies. The local atomic charge was computed using Bader’s charge decomposition method<sup>43</sup>, which divides the total volume into individual atomic volumes for each atom as the one

containing a single charge density maximum and separated from the other volumes by a zero-flux surface of the gradients of the charge density magnitude. The incremental adsorption energy of an oxygen atom as a function of oxygen coverage  $\langle\Theta\rangle = n/\{n_0\}$  for  $n$  O atoms on  $n_0$  fcc sites on a Cu(111) substrate is calculated as the energy change when a further O atom is adsorbed on the substrate,  $\langle E_{ad} \rangle = E[\{rm{Cu}\} + (n+1)\{rm{O}\}] - E[\{rm{Cu}\} + n\{rm{O}\}] - \frac{1}{2}E[\{rm{O}_2\}]$ , in which  $E[\{rm{Cu} + nO\}]$  is the total energy of  $n$  oxygen atoms on the Cu substrate and  $E[\{O_2\}]$  is the total energy of an isolated oxygen molecule.

## Data availability

The data that support the findings of this study are available from the corresponding authors on reasonable request. [Source data](#) are provided with this paper.

## References

1. Gattinoni, C. & Michaelides, A. Atomistic details of oxide surfaces and surface oxidation: the example of copper and its oxides. *Surf. Sci. Rep.* **70**, 424–447 (2015).
2. Zhu, Q., Zou, L., Zhou, G., Saidi, W. A. & Yang, J. C. Early and transient stages of Cu oxidation: atomistic insights from theoretical simulations and in situ experiments. *Surf. Sci.* **652**, 98–113 (2016).
3. Lahtonen, K., Hirsimaki, M., Lampimaki, M. & Valden, M. Oxygen adsorption-induced nanostructures and island formation on Cu{100}: bridging the gap between the formation of surface confined oxygen chemisorption layer and oxide formation. *J. Chem. Phys.* **129**, 124703 (2008).
4. Ahonen, M. et al. Adsorption dynamics of O<sub>2</sub> on Cu(100): the role of vacancies, steps and adatoms in dissociative chemisorption of O<sub>2</sub>. *Chem. Phys. Lett.* **456**, 211–214 (2008).
5. Moritani, K., Okada, M., Teraoka, Y., Yoshigoe, A. & Kasai, T. Reconstruction of Cu(111) induced by a hyperthermal oxygen molecular beam. *J. Phys. Chem. C* **112**, 8662–8667 (2008).
6. Yang, J. C., Kolasa, B., Gibson, J. M. & Yeadon, M. Self-limiting oxidation of copper. *Appl. Phys. Lett.* **73**, 2841–2843 (1998).

7. Huang, C. L., Weng, W. L., Liao, C. N. & Tu, K. N. Suppression of interdiffusion-induced voiding in oxidation of copper nanowires with twin-modified surface. *Nat. Commun.* **9**, 340 (2018).
8. Bellchambers, P. et al. Elucidating the exceptional passivation effect of 0.8 nm evaporated aluminium on transparent copper films. *Front. Mater.* **5**, 71 (2018).
9. Li, L. et al. Surface-step-induced oscillatory oxide growth. *Phys. Rev. Lett.* **113**, 136104 (2014).
10. Zhou, G. et al. In situ atomic-scale visualization of oxide islanding during oxidation of Cu surfaces. *Chem. Commun.* **49**, 10862–10864 (2013).
11. Zhou, G. et al. Step-edge-induced oxide growth during the oxidation of Cu surfaces. *Phys. Rev. Lett.* **109**, 235502 (2012).
12. Baber, A. E. et al. In situ imaging of Cu<sub>2</sub>O under reducing conditions: formation of metallic fronts by mass transfer. *J. Am. Chem. Soc.* **135**, 16781–16784 (2013).
13. Fujita, K., Ando, D., Uchikoshi, M., Mimura, K. & Isshiki, M. New model for low-temperature oxidation of copper single crystal. *Appl. Surf. Sci.* **276**, 347–358 (2013).
14. Zhang, X. et al. Nanocrystalline copper films are never flat. *Science* **357**, 397–400 (2017).
15. Schiøtz, J. & Jacobsen, K. W. Nanocrystalline metals: roughness in flatland. *Nat. Mater.* **16**, 1059–1060 (2017).
16. Kim, S. J. et al. Color of copper/copper oxide. *Adv. Mater.* **33**, 2007345 (2021).
17. Hýtch, M. J., Snoeck, E. & Kilaas, R. Quantitative measurement of displacement and strain fields from HREM micrographs. *Ultramicroscopy* **74**, 131–146 (1998).
18. Hýtch, M. J., Putaux, J.-L. & Pénisson, J.-M. Measurement of the displacement field of dislocations to 0.03 Å by electron microscopy. *Nature* **423**, 270–273 (2003).
19. Sasaki, T. et al. Atomic and electronic structures of Cu/a-Al<sub>2</sub>O<sub>3</sub> interfaces prepared by pulsed-laser deposition. *Sci. Technol. Adv. Mater.* **4**, 575–584 (2003).
20. Ernst, F., Pirouz, P. & Heuer, A. H. HRTEM study of a Cu/Al<sub>2</sub>O<sub>3</sub> interface. *Philos. Mag. A* **63**, 259–277 (1991).

21. Richardson, C. J. K. & Lee, M. L. Metamorphic epitaxial materials. *MRS Bull.* **41**, 193–198 (2016).
22. Kato, T. et al. Simultaneous growth of two differently oriented GaN epilayers on (11.0) sapphire II. A growth model of (00.1) and (10.0) GaN. *J. Cryst. Growth* **183**, 131–139 (1998).
23. Liu, X. Y. et al. Perfect strain relaxation in metamorphic epitaxial aluminum on silicon through primary and secondary interface misfit dislocation arrays. *ACS Nano* **12**, 6843–6850 (2018).
24. Kim, Y. M., Pennycook, S. J. & Borisevich, A. Y. Quantitative comparison of bright field and annular bright field imaging modes for characterization of oxygen octahedral tilts. *Ultramicroscopy* **181**, 1–7 (2017).
25. Kung, P., Sun, C. J., Saxler, A., Ohsato, H. & Razeghi, M. Crystallography of epitaxial growth of wurtzite-type thin films on sapphire substrates. *J. Appl. Phys.* **75**, 4515–4519 (1994).
26. Williams, D. B. & Carter, C. B. *Transmission Electron Microscopy* (Plenum Press, 1996).
27. Ngantcha, J. P., Gerland, M., Kihn, Y. & Riviere, A. Correlation between microstructure and mechanical spectroscopy of a Cu-Cu<sub>2</sub>O alloy between 290 K and 873 K. *Eur. Phys. J. App. Phys.* **29**, 83–89 (2005).
28. Ewels, P., Sikora, T., Serin, V., Ewels, C. P. & Lajaunie, L. A complete overhaul of the electron energy-loss spectroscopy and X-ray absorption spectroscopy database: eelsdb.eu. *Microsc. Microanal.* **22**, 717–724 (2016).
29. Perdew, J. P., Burke, K. & Ernzerof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
30. Blochl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17979 (1994).
31. Kresse, G. & Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **59**, 1758–1775 (1999).
32. Coulman, D. J., Wintterlin, J., Behm, R. J. & Ertl, G. Novel mechanism for the formation of chemisorption phases: the (2×1)O-Cu(110) “added row” reconstruction. *Phys. Rev. Lett.* **64**, 1761–1764 (1990).

33. Sun, L. D., Hohage, M., Denk, R. & Zeppenfeld, P. Oxygen adsorption on Cu(110) at low temperature. *Phys. Rev. B* **76**, 245412 (2007).
34. Lian, X., Xiao, P., Liu, R. & Henkelman, G. Communication: calculations of the  $(2 \times 1)$ -O reconstruction kinetics on Cu(110). *J. Chem. Phys.* **146**, 111101 (2017).
35. Li, L., Liu, Q., Li, J., Saidi, W. A. & Zhou, G. Kinetic barriers of the phase transition in the oxygen chemisorbed Cu(110)- $(2 \times 1)$ -O as a function of oxygen coverage. *J. Phys. Chem. C* **118**, 20858–20866 (2014).
36. Cha, S. Y. et al. Improving the precision of Hall effect measurements using a single-crystal copper probe. *Rev. Sci. Instrum.* **83**, 013901 (2012).
37. Cho, Y. C. et al. Copper better than silver: electrical resistivity of the grain-free single-crystal copper wire. *Cryst. Growth Des.* **10**, 2780–2784 (2010).
38. Jackson, D. P. Binding energies in cubic metal surfaces. *Radiat. Eff.* **18**, 185–189 (1973).
39. Kudriavtsev, Y., Villegas, A., Godines, A. & Asomoza, R. Calculation of the surface binding energy for ion sputtered particles. *Appl. Surf. Sci.* **239**, 273–278 (2005).
40. Yagisawa, T. & Makabe, T. Modeling of dc magnetron plasma for sputtering: transport of sputtered copper atoms. *J. Vac. Sci. Technol. A* **24**, 908–913 (2006).
41. Monkhorst, H. J. & Pack, J. D. Special points for Brillouin-zone integrations. *Phys. Rev. B* **13**, 5188–5192 (1976).
42. Henkelman, G., Uberuaga, B. P. & Jonsson, H. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *J. Chem. Phys.* **113**, 9901–9904 (2000).
43. Bader, R. F. W. A quantum theory of molecular structure and its applications. *Chem. Rev.* **91**, 893–928 (1991).

## Acknowledgements

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (nos. NRF-2017R1A2B3011822 and NRF-2020R1A4A4078780), the Creative Materials Discovery Program (nos. NRF-2015M3D1A1070672 and NRF-2016M3D1A1919181) through the NRF, and the Institute for Basic Science (IBS-

R011-D1). Computer time allocation was provided by the High Performance Computing Center (HPCC) at Mississippi State University.

## Author information

### Author notes

1. These authors contributed equally: Su Jae Kim, Yong In Kim, Bipin Lamichhane.

### Affiliations

1. Crystal Bank Research Institute, Pusan National University, Busan, Republic of Korea

Su Jae Kim, Yousil Lee & Miyeon Cheon

2. Department of Energy Science, Sungkyunkwan University, Suwon, Republic of Korea

Yong In Kim, Young-Hoon Kim, Young Hee Lee & Young-Min Kim

3. Department of Physics and Astronomy, Mississippi State University, Mississippi State, MS, USA

Bipin Lamichhane & Seong-Gon Kim

4. Department of Nanoenergy Engineering, Pusan National University, Busan, Republic of Korea

Chae Ryong Cho

5. School of Materials Science and Engineering, Ulsan National Institute of Science and Technology, Ulsan, South Korea

Jong Chan Kim

6. UNIST Central Research Facilities (UCRF), Ulsan National Institute of Science and Technology, Ulsan, South Korea

Hu Young Jeong

7. Center for Integrated Nanostructure Physics, Institute for Basic Science (IBS), Sungkyunkwan University, Suwon, Republic of Korea

Taewoo Ha, Young Hee Lee & Young-Min Kim

8. Department of Physics, University of Ulsan, Ulsan, Republic of Korea

Jungdae Kim

9. Department of Physics, Sungkyunkwan University, Suwon, Republic of Korea

Young Hee Lee

10. Department of Optics and Mechatronics Engineering, Pusan National University, Busan, Republic of Korea

Se-Young Jeong

11. Department of Cogno-Mechatronics Engineering, Pusan National University, Busan, Republic of Korea

Se-Young Jeong

## Contributions

S.-Y.J., Y.-M.K. and S.-G.K. conceived this work and wrote the manuscript. S.-Y.J. and Y.-M.K. supervised the experiments. S.J.K. and Y.L. prepared the SCCF samples. Y.-M.K., Y.I.K., Y.-H.K., J.C.K. and H.Y.J. performed TEM analyses and sample preparations. J.K. conducted the STM analysis. S.J.K., Y.L., T.H., C.R.C. and M.C. performed EBSD and AFM experiments and assisted with the data analyses. S.-Y.J. and S.-G.K. established the theoretical model, and S.-G.K. and B.L. carried out first-principles calculations. S.-Y.J. and Y.H.L. supervised and coordinated this work. All authors contributed to the discussion and analysis of the results.

## Corresponding authors

Correspondence to [Seong-Gon Kim](#), [Young-Min Kim](#) or [Se-Young Jeong](#).

## Ethics declarations

## Competing interests

The authors declare no competing interests.

## Peer review

## Peer review information

*Nature* thanks Ross Hatton and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

## Additional information

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Extended data figures and tables

### [Extended Data Fig. 1 Pristine SCCFs grown by ASE.](#)

**a**, Photo of a 2-inch SCCF grown on  $\text{Al}_2\text{O}_3$ . **b**,  $\theta$ - $2\theta$  X-ray diffraction data taken at different positions marked in **a**. **c**, Surface morphologies at 20-nm (upper) and 1.5-nm (lower) vertical scales of AFM images with a root mean square (RMS) surface roughness of 0.28 nm. **d**, EBSD map showing perfect alignment along the (111) plane. **e**, IPF with a sole spot associated with the (111) plane. **f**, [100] PF showing the sixfold symmetry of the {100} PF. The inset images in **e** and **f** are the enlarged images of the sole-spot areas. **g**, Scanning electron microscopy images of the sample at different magnifications. **h**, Low-magnification cross-sectional BF-TEM image of the SCCF sample.

### [Source data](#)

### [Extended Data Fig. 2 Structural and chemical characterizations of a conventional PCCF grown on a glass substrate.](#)

**a**, Low-magnification cross-sectional BF-TEM image of the PCCF sample showing a typical polycrystalline structure. **b**, Magnified image of the region marked by the yellow box in **a**, which shows the presence of a thick oxidized Cu layer on the surface of the PCCF. **c**, **d**, HRTEM images of oxidized Cu nanoparticles formed in the regions denoted by white boxes in **b**. The measured planar spacings of the (111) planes are 2.32 and 2.46 Å, which represent typical values of CuO and Cu<sub>2</sub>O phases, respectively. **e**, ADF-STEM image of the surface region of the PCCF. **f**, **g**, STEM-EDX elemental maps of Cu (**f**) and O (**g**) in the surface region. **h**, Vertically averaged intensity profiles of Cu (blue) and O (red) across the surface. Note that the intensity of each element was normalized with respect to its intensity maximum. From this result, it is evident that the highly oxidized Cu surface layer (indicated by the pink shaded

region in the graph) can be formed on the PCCF with a thickness of tens of nanometres.

[Source data](#)

[Extended Data Fig. 3 Surface oxidation of bulk Cu single crystal.](#)

**a, g**, Low-magnification micrographs of the non-polished and polished Cu single-crystal samples prepared by FIB milling. The insets show the two Cu single crystals sliced and cut to dimensions of  $5 \times 5 \text{ mm}^2$ . **b, h**, HRTEM images of the two Cu surfaces. Note that the yellow arrows indicate the thin  $\text{Cu}_2\text{O}$  layers that formed on the two surfaces. **c, d, i, j**, FFT patterns of the two regions of  $\text{Cu}_2\text{O}$  (red box, labelled ‘1’) and Cu (green box, labelled ‘2’) depicted in each HRTEM image. **e, k**, Composite of the two FFT patterns for comparison. **f, l**, Measurements of the layer spacings across the interface between Cu and  $\text{Cu}_2\text{O}$  (red dotted arrows).

[Source data](#)

[Extended Data Fig. 4 STM topography images of the clean Cu\(111\) single-crystal film at room temperature.](#)

**a**, Configuration of step edges for an imaging area of  $200 \times 200 \text{ nm}^2$ . **b**, Atomic-resolution surface structure (image size around  $3 \times 3 \text{ nm}^2$ ). The arrows indicate the three corresponding  $<110>$  orientations on the (111) plane. **c**, Profile of step height obtained from the line scan marked with a black arrow in **a**, which shows a series of mono-atomic steps corresponding to the planar spacing of the Cu(111).

[Extended Data Fig. 5 Interface structure and strain distribution.](#)

**a**, Cross-sectional HRTEM image of the interface region of the  $\text{Cu}-\text{Al}_2\text{O}_3$  heterostructure with an orientation relation of  $\langle\langle(111)\rangle\rangle_{\text{Cu}} \parallel \langle\langle\overline{11}\overline{2}\rangle\rangle_{\text{Cu}} \parallel \langle\langle001\rangle\rangle_{\text{Al}} \langle\langle\overline{110}\rangle\rangle_{\text{Al}}$ . **b**, Enlarged image showing an abrupt interface structure. **c, d**, In-plane ( $E_{xx}$ ) and out-of-plane ( $E_{yy}$ ) strain field maps obtained by GPA of the HRTEM image. Note that the region enclosed by the white dotted line indicates the strained interface region of about 1–2 nm in thickness. The colour scale indicates the magnitude of the strain relative to the reference area marked by the white box. Scale bars are 2 nm.

## Extended Data Fig. 6 Chemical analysis and electronic structure of the grown SCCF.

**a, b**, Nanoscale STEM-EDX elemental maps of Al, O and Cu in the regions on the surface and at the interface of the Cu–Al<sub>2</sub>O<sub>3</sub> heterostructure, respectively. **c**, Comparison of the Cu  $L_{2,3}$  electron energy-loss near-edge structure profiles obtained at four different positions (marked by the numbers 1–4 in the ADF-STEM images) and the reference profiles obtained from copper metal and copper oxides, CuO and Cu<sub>2</sub>O (lower graphs). **d**, ADF-STEM image, core-loss EELS spectrum imaging data and the constructed elemental map of Cu and O for the surface region of the SCCF sample. The red and blue profiles on the right side are core-loss EEL spectra of O  $K$  and Cu  $L_{2,3}$  edges extracted from the white dotted rectangle denoted in the EELS map. **e**, Equivalent data arrangement for the PCCF sample for comparison of surface oxidation. **f**, X-ray photoelectron spectroscopy depth profile depending on the etching time for the SCCF sample.

Source data

## Extended Data Fig. 7 Structural and electronic properties of Cu, Cu<sub>2</sub>O and Cu<sub>2</sub>O monolayers on the Cu(111) surface.

**a–c**, Structural model of Cu, Cu<sub>2</sub>O and Cu<sub>2</sub>O monolayers on the Cu(111) surface. Cu and O atoms are represented by blue and red spheres, respectively. Only the bonds between O and Cu atoms are drawn explicitly for clarity. **d**, Structural parameters of Cu, Cu<sub>2</sub>O and Cu<sub>2</sub>O monolayers on the Cu(111) surface. **e**, Degree of oxidation of atoms in various configurations.  $n_e$  is the electron charge of each atom (the number of valence electrons attributed to each atom computed using the Bader decomposition method<sup>43</sup>) and  $n_0$  is the number of valence electrons of the corresponding isolated neutral atom.

## Extended Data Fig. 8 Relative energy per oxygen atom for the dissociation of the O<sub>2</sub> molecule approaching the Cu(111) surface as a function of O–O bond length.

Blue spheres represent Cu atoms and red spheres represent O atoms. The insets show the side and top views of the configurations, with corresponding letters on the curve. **a**, O<sub>2</sub> molecule far away from the surface. **b**, O<sub>2</sub> molecule in the physisorbed state. **c**, O<sub>2</sub> molecule in the transition state. **d**, Two O atoms adsorbed on the surface. **e**, O atoms diffused on the surface away from each other.

## [Source data](#)

### [Extended Data Fig. 9 Structure of the ASE system and uniformity of the grown Cu film.](#)

**a**, Schematic diagram of the construction of the ASE system used in this study. The three key technological modifications made to differ from a commercially available sputtering system are as follows: single-crystal Cu sputtering target, electrical wiring with single-crystal Cu cables and mechanical noise-reduction parts, such as vibration absorbers and mechanical diodes. **b**, Two-inch SCCF grown wafer with a target thickness of 95 nm. To check the uniformity of the film thickness, the wafer was cut into pieces along the red lines and the five pieces marked by numbers were mounted to our AFM instrument to measure their thickness. **c**, Thickness profiles of the five samples. One side of each SCCF film was mechanically removed to expose the bare surface of the Al<sub>2</sub>O<sub>3</sub> substrate to measure the film thickness from AFM edge profiling. Note that the mid-value in the thickness profile was set to be zero in AFM. The average thickness of the five samples was estimated to be  $95 \pm 0.18$  nm, which corresponds to about 99.8% in uniformity.

### [Extended data Fig. 10 Monitoring RF power stability and the effect of ASE on surface roughness.](#)

**a, b**, Plots showing the output RF power fluctuations measured over time in the sputtering system without (**a**, black) and with a single-crystal Cu power cable (**b**, red). **c**, Plot of the measured root mean square (RMS) roughness of many samples as a function of the film thickness. The RMS values of the blue squares and the red circles were measured from the samples grown by the conventional sputtering system equipped with a single-crystal Cu target only and those grown using the ASE system, respectively. The thickness-dependent average RMS surface roughness of the blue squares (blue dashed line) and the red circles (red dashed line, ASE) are  $0.66 \pm 0.2$  nm and  $0.20 \pm 0.1$  nm, respectively. The average RMS surface roughness of the single-crystal films grown by the ASE method for <35 nm (purple dashed line) is  $0.17 \pm 0.1$  nm.

## [Source data](#)

## **Supplementary information**

### [Peer Review File](#)

## Source data

[Source Data Fig. 1](#)

[Source Data Fig. 2](#)

[Source Data Fig. 4](#)

[Source Data Extended Data Fig. 1](#)

[Source Data Extended Data Fig. 2](#)

[Source Data Extended Data Fig. 3](#)

[Source Data Extended Data Fig. 6](#)

[Source Data Extended Data Fig. 8](#)

[Source Data Extended Data Fig. 10](#)

## Rights and permissions

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

[Reprints and Permissions](#)

## About this article

[Cite this article](#)

Kim, S.J., Kim, Y.I., Lamichhane, B. *et al.* Flat-surface-assisted and self-regulated oxidation resistance of Cu(111). *Nature* **603**, 434–438 (2022).  
<https://doi.org/10.1038/s41586-021-04375-5>

- Received: 08 March 2021
- Accepted: 20 December 2021
- Published: 16 March 2022
- Issue Date: 17 March 2022
- DOI: <https://doi.org/10.1038/s41586-021-04375-5>

## Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04375-5>

| [Section menu](#) | [Main menu](#) |

- Article
- [Published: 16 March 2022](#)

# A kinase-cGAS cascade to synthesize a therapeutic STING activator

- [John A. McIntosh](#) ORCID: [orcid.org/0000-0002-9487-490X<sup>1</sup>](https://orcid.org/0000-0002-9487-490X),
- [Zhijian Liu](#)<sup>1</sup>,
- [Brian M. Andresen](#) ORCID: [orcid.org/0000-0002-9645-3359<sup>2</sup>](https://orcid.org/0000-0002-9645-3359),
- [Nastaran Salehi Marzijarani](#) ORCID: [orcid.org/0000-0001-8940-4674<sup>1</sup>](https://orcid.org/0000-0001-8940-4674),
- [Jeffrey C. Moore](#) ORCID: [orcid.org/0000-0002-9807-6315<sup>1</sup>](https://orcid.org/0000-0002-9807-6315),
- [Nicholas M. Marshall](#)<sup>1</sup>,
- [Margie Borra-Garske](#)<sup>3</sup>,
- [Jennifer V. Obligacion](#)<sup>1</sup>,
- [Patrick S. Fier](#)<sup>1</sup>,
- [Feng Peng](#)<sup>1</sup>,
- [Jacob H. Forstater](#)<sup>1</sup>,
- [Matthew S. Winston](#)<sup>1</sup>,
- [Chihui An](#)<sup>1</sup>,
- [Wonsuk Chang](#)<sup>2</sup>,
- [Jongwon Lim](#)<sup>2</sup>,
- [Mark A. Huffman](#)<sup>1</sup>,
- [Steven P. Miller](#) ORCID: [orcid.org/0000-0001-8615-274X<sup>1</sup>](https://orcid.org/0000-0001-8615-274X),
- [Fuh-Rong Tsay](#)<sup>4</sup>,
- [Michael D. Altman](#)<sup>2</sup>,
- [Charles A. Lesburg](#) ORCID: [orcid.org/0000-0001-7245-7331<sup>2</sup>](https://orcid.org/0000-0001-7245-7331),
- [Dietrich Steinhuebel](#)<sup>1</sup>,
- [B. Wesley Trotter](#) ORCID: [orcid.org/0000-0002-6780-0358<sup>2</sup>](https://orcid.org/0000-0002-6780-0358),

- [Jared N. Cumming](#)<sup>2</sup>,
- [Alan Northrup](#)<sup>2</sup>,
- [Xiaodong Bu](#)<sup>4</sup>,
- [Benjamin F. Mann](#)<sup>4</sup>,
- [Mirlinda Biba](#)<sup>4</sup>,
- [Kaori Hiraga](#)<sup>1</sup>,
- [Grant S. Murphy](#) ORCID: [orcid.org/0000-0003-0388-0309](https://orcid.org/0000-0003-0388-0309)<sup>1</sup>,
- [Joshua N. Kolev](#)<sup>1</sup>,
- [Amanda Makarewicz](#)<sup>1</sup>,
- [Weilan Pan](#)<sup>1</sup>,
- [Iman Farasat](#)<sup>1</sup>,
- [Rachel S. Bade](#) ORCID: [orcid.org/0000-0002-5384-9687](https://orcid.org/0000-0002-5384-9687)<sup>1</sup>,
- [Kevin Stone](#)<sup>1</sup>,
- [Da Duan](#)<sup>3</sup>,
- [Oscar Alvizo](#)<sup>3</sup>,
- [Donovan Adpressa](#)<sup>2</sup>,
- [Erik Guetschow](#)<sup>4</sup>,
- [Erik Hoyt](#)<sup>4</sup>,
- [Erik L. Regalado](#) ORCID: [orcid.org/0000-0002-7352-6391](https://orcid.org/0000-0002-7352-6391)<sup>1,4</sup>,
- [Steve Castro](#)<sup>4</sup>,
- [Nelo Rivera](#)<sup>4</sup>,
- [Joseph P. Smith](#) ORCID: [orcid.org/0000-0002-0062-2534](https://orcid.org/0000-0002-0062-2534)<sup>4</sup>,
- [Fengqiang Wang](#)<sup>4</sup>,
- [Alejandro Crespo](#)<sup>1</sup>,
- [Deeptak Verma](#) ORCID: [orcid.org/0000-0003-0740-0624](https://orcid.org/0000-0003-0740-0624)<sup>1</sup>,
- [Stephanus Axnanda](#) ORCID: [orcid.org/0000-0002-7239-9220](https://orcid.org/0000-0002-7239-9220)<sup>4</sup>,
- [Zachary E. X. Dance](#) ORCID: [orcid.org/0000-0003-1807-7930](https://orcid.org/0000-0003-1807-7930)<sup>4</sup>,
- [Paul N. Devine](#)<sup>1</sup>,
- [David Tschaen](#)<sup>1</sup>,
- [Keith A. Canada](#)<sup>1</sup>,
- [Paul G. Bulger](#)<sup>1</sup>,
- [Benjamin D. Sherry](#)<sup>1</sup>,
- [Matthew D. Truppo](#)<sup>1</sup>,
- [Rebecca T. Ruck](#) ORCID: [orcid.org/0000-0001-9980-9675](https://orcid.org/0000-0001-9980-9675)<sup>1</sup>,

- [Louis-Charles Campeau](#) ORCID: [orcid.org/0000-0002-2373-802X<sup>1</sup>](https://orcid.org/0000-0002-2373-802X),
- [David Jonathan Bennett](#) ORCID: [orcid.org/0000-0003-0163-7479<sup>2</sup>](https://orcid.org/0000-0003-0163-7479),
- [Guy R. Humphrey<sup>1</sup>](#),
- [Kevin R. Campos<sup>1</sup>](#) &
- [Matthew L. Maddess<sup>1</sup>](#)

[Nature](#) volume **603**, pages 439–444 (2022)

- 3946 Accesses
- 66 Altmetric
- [Metrics details](#)

## Subjects

- [Biosynthesis](#)
- [Process chemistry](#)

## Abstract

The introduction of molecular complexity in an atom- and step-efficient manner remains an outstanding goal in modern synthetic chemistry. Artificial biosynthetic pathways are uniquely able to address this challenge by using enzymes to carry out multiple synthetic steps simultaneously or in a one-pot sequence<sup>1,2,3</sup>. Conducting biosynthesis *ex vivo* further broadens its applicability by avoiding cross-talk with cellular metabolism and enabling the redesign of key biosynthetic pathways through the use of non-natural cofactors and synthetic reagents<sup>4,5</sup>. Here we describe the discovery and construction of an enzymatic cascade to MK-1454, a highly potent stimulator of interferon genes (STING) activator under study as an immuno-oncology therapeutic<sup>6,7</sup> (ClinicalTrials.gov study [NCT04220866](#)). From two non-natural nucleotide monothiophosphates, MK-1454 is assembled diastereoselectively in a one-pot cascade, in which two thiophosphate nucleotides are simultaneously generated biocatalytically, followed by coupling and cyclization catalysed by an engineered animal

cyclic guanosine-adenosine synthase (cGAS). For the thiotriphosphate synthesis, three kinase enzymes were engineered to develop a non-natural cofactor recycling system in which one thiotriphosphate serves as a cofactor in its own synthesis. This study demonstrates the substantial capacity that currently exists to use biosynthetic approaches to discover and manufacture complex, non-natural molecules.

This is a preview of subscription content

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

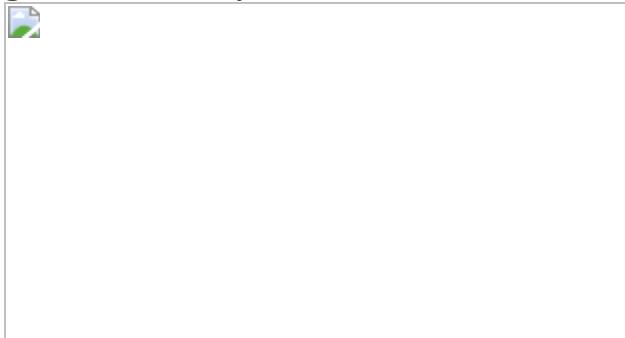
[Buy](#)

All prices are NET prices.

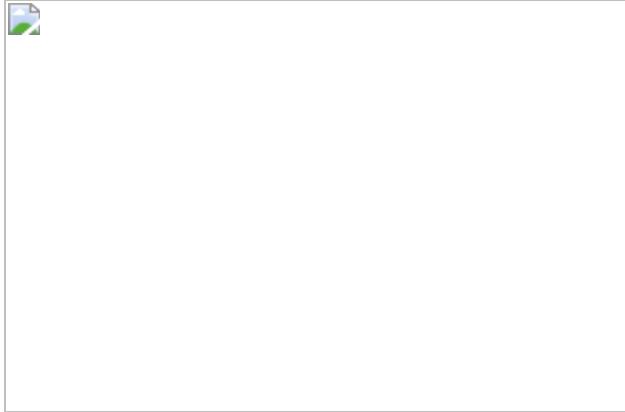
### **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

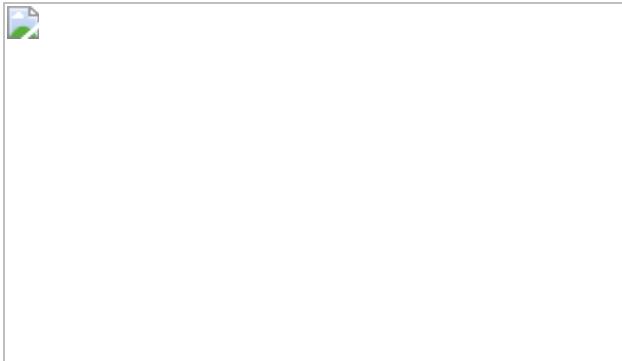
**Fig. 1: Discovery of MK-1454.**



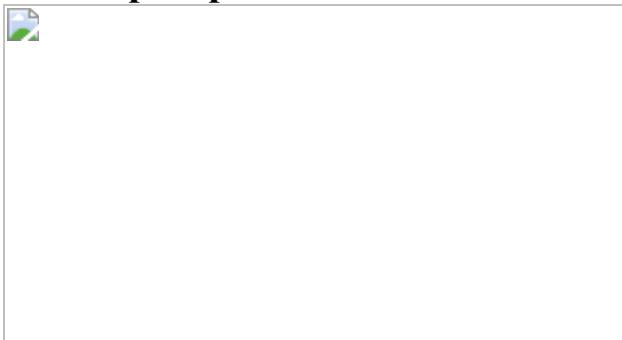
**Fig. 2: Diastereoselective CDN synthesis catalysed by cGAS.**



**Fig. 3: Assembly of thiotriphosphate biocatalytic cascade.**



**Fig. 4: Cascade process to MK-1454 from nucleotide monothiophosphates without isolated intermediates.**



## Data availability

The data supporting the findings of this study are available within the paper and its [Supplementary Information](#). Coordinates for the STING structure have been deposited in the Protein Data Bank (PDB ID [7MHC](#)). DNA sequences of wild-type and evolved enzymes used in this study are available in the Supplementary Data files and have been deposited in Genbank (accession codes [OL362244–OL362267](#)). Gene sequences are available in the Supplementary Data files. The enzymes are commercially available from Codexis, Inc., subject to existing license obligations and restrictions.

## References

1. Schrittwieser, J. H., Velikogne, S., Hall, M. & Kroutil, W. Artificial biocatalytic linear cascades for preparation of organic molecules. *Chem. Rev.* **118**, 270–348 (2018).

2. Huffman, M. A. et al. Design of an in vitro biocatalytic cascade for the manufacture of islatravir. *Science* **366**, 1255–1259 (2019).
3. Luo, X. et al. Complete biosynthesis of cannabinoids and their unnatural analogues in yeast. *Nature* **567**, 123–126 (2019).
4. Arnold, F. H. Directed evolution: bringing new chemistry to life. *Angew Chem. Int. Ed. Engl.* **57**, 4143–4148 (2018).
5. Bowie, J. U. et al. Synthetic biochemistry: the bio-inspired cell-free approach to commodity chemical production. *Trends Biotechnol.* **38**, 766–778 (2020).
6. Altman, M. D. et al. Cyclic di-nucleotide compounds as STING agonists. Patent WO2017027646A1 (2016).
7. Walsh, C. T., Tu, B. P. & Tang, Y. Eight kinetically stable but thermodynamically activated molecules that power cell metabolism. *Chem. Rev.* **118**, 1460–1494 (2018).
8. Cross, R. STING fever is sweeping through the cancer immunotherapy world. *Chem. Eng. News* **96**, 24–26 (2018).
9. Burdette, D. L. et al. STING is a direct innate immune sensor of cyclic di-GMP. *Nature* **478**, 515–518 (2011).
10. Sun, L., Wu, J., Du, F., Chen, X. & Chen, Z. J. Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. *Science* **339**, 786–791 (2013).
11. Knouse, K. W. et al. Unlocking P(V): reagents for chiral phosphorothioate synthesis. *Science* **361**, 1234–1238 (2018).
12. Lioux, T. et al. Design, synthesis, and biological evaluation of novel cyclic adenosine-inosine monophosphate (cAIMP) analogs that activate stimulator of interferon genes (STING). *J. Med. Chem.* **59**, 10253–10267 (2016).

13. Featherston, A. L. et al. Catalytic asymmetric and stereodivergent oligonucleoside synthesis. *Science* **371**, 702–707 (2021).
14. Yan, H., Wang, X., KuoLee, R. & Chen, W. Synthesis and immunostimulatory properties of the phosphorothioate analogues of cdiGMP. *Bioorg. Med. Chem. Lett.* **18**, 5631–5634 (2008).
15. Gaffney, B. L., Veliath, E., Zhao, J. & Jones, R. A. One-flask syntheses of c-di-GMP and the [Rp,Rp] and [Rp,Sp] thiophosphate analogues. *Org. Lett.* **12**, 3269–3271 (2010).
16. Zhang, X. et al. Cyclic GMP-AMP containing mixed phosphodiester linkages is an endogenous high-affinity ligand for STING. *Mol. Cell* **51**, 226–235 (2013).
17. Gao, P. et al. Structure-function analysis of STING activation by c[G(2',5')pA(3',5')p] and targeting by antiviral DMXAA. *Cell* **154**, 748–762 (2013).
18. Li, L. et al. Hydrolysis of 2'3'-cGAMP by ENPP1 and design of nonhydrolyzable analogs. *Nat. Chem. Biol.* **10**, 1043–1048 (2014).
19. Ablasser, A. et al. cGAS produces a 2'-5'-linked cyclic dinucleotide second messenger that activates STING. *Nature* **498**, 380–384 (2013).
20. Diner, E. J. et al. The innate immune DNA sensor cGAS produces a noncanonical cyclic dinucleotide that activates human STING. *Cell. Rep.* **3**, 1355–1361 (2013).
21. Gao, P. et al. Cyclic [G(2',5')pA(3',5')p] is the metazoan second messenger produced by DNA-activated cyclic GMP-AMP synthase. *Cell* **153**, 1094–1107 (2013).
22. Eckstein, F. Nucleoside phosphorothioates. *Annu. Rev. Biochem.* **54**, 367–402 (1985).
23. Du, M. & Chen, Z. J. DNA-induced liquid phase condensation of cGAS activates innate immune signaling. *Science* **361**, 704–709

(2018).

24. Thillier, V., Sallamand, C. C. B., Vasseur, J. J. & Debart, F. Solid-phase synthesis of oligonucleotide 5'-( $\alpha$ -P-Thio)triphosphates and 5'-( $\alpha$ -P-Thio)( $\beta,\gamma$ -methylene)triphosphates. *Eur. J. Org. Chem.* **2015**, 302–308 (2015).
25. Ludwig, J. & Eckstein, F. Rapid and efficient synthesis of nucleoside 5'-O-(1-thiotriphosphates), 5'-triphosphates, and 2',3'-cyclophosphorothioates using 2-chloro-4H-1,3,2,-benzodioxaphosphorin-4-one. *J. Org. Chem.* **54**, 631–635 (1989).
26. Moran, J. R. & Whitesides, G. M. A practical enzymatic synthesis of (Sp)-adenosine 5'-O-(1-thiotriphosphate) ((Sp)-ATP- $\alpha$ -S)). *J. Org. Chem.* **49**, 1984 (1984).
27. Jaffe, E. K. & Cohn, M.  $^{31}$ P nuclear magnetic resonance spectra of the thiophosphate analogues of adenine nucleotides; effects of pH and Mg $^{2+}$  binding. *Biochemistry* **17**, 652–657 (1978).
28. Rex Sheu, K. F. & Frey, P. A. Enzymatic and 32P nuclear magnetic resonance study of adenylate kinase-catalyzed stereospecific phosphorylation of adenosine 5'-phosphorothioate. *J. Biol. Chem.* **252**, 4445–4448 (1977).
29. Sandoval, B. A. & Hyster, T. K. Emerging strategies for expanding the toolbox of enzymes in biocatalysis. *Curr. Opin. Chem. Biol.* **55**, 45–51 (2020).
30. Ren, X. & Fasan, R. Engineered and artificial metalloenzymes for selective C–H functionalization. *Curr. Opin. Green Sustain. Chem.* **31**, 100494 (2021).
31. Brandenberg, O. F., Fasan, R. & Arnold, F. H. Exploiting and engineering hemoproteins for abiological carbene and nitrene transfer reactions. *Curr. Opin. Biotechnol.* **47**, 102–111 (2017).

32. Qu, G., Li, A., Acevedo-Rocha, C. G., Sun, Z. & Reetz, M. T. The crucial role of methodology development in directed evolution of selective enzymes. *Angew Chem. Int. Ed. Engl.* **59**, 13204–13231 (2020).
33. Lim, J. & Kim, H. Y. Novel applications of biocatalysis to stereochemistry determination of 2'3'-cGAMP bisphosphorothioate (2'3'-cG<sup>S</sup>A<sup>S</sup>MP). *ACS Omega* **5**, 14173–14179 (2020).
34. Crans, D. C. & Whitesides, G. M. A convenient synthesis of disodium acetyl phosphate for use in in situ ATP cofactor regeneration. *J. Org. Chem.* **48**, 3130–3132 (1983).
35. Gottlieb, H. E., Kotlyar, V. & Nudelman, A. J. NMR chemical shifts of common laboratory solvents as trace impurities. *J. Org. Chem.* **62**, 7512–7515 (1997).
36. Fulmer, G. R. et al. NMR chemical shifts of trace impurities: common laboratory solvents, organics, and gases in deuterated solvents relevant to the organometallic chemist. *Organometallics* **29**, 2176–2179 (2010).
37. Pan, B. S. et al. An orally available non-nucleotide STING agonist with antitumor activity. *Science* **369**, eaba6098 (2020).
38. Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41**, 207–234 (2005).
39. Abele, U. & Schulz, G. E. High-resolution structures of adenylate kinase from yeast ligated with inhibitor Ap5A, showing the pathway of phosphoryl transfer. *Protein Sci.* **4**, 1262–1271 (1995).
40. Sekulic, N., Shuvalova, L., Spangenberg, O., Konrad, M. & Lavie, A. Structural characterization of the closed conformation of mouse guanylate kinase. *J. Biol. Chem.* **277**, 30236–30243 (2002).

## Acknowledgements

This research used resources at the Industrial Macromolecular Crystallography Association Collaborative Access Team beamline 17-ID, supported by the companies of the Industrial Macromolecular Crystallography Association through a contract with Hauptman–Woodward Medical Research Institute. This research used resources of the Advanced Photon Source, a US Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under contract no. DE-AC02-06CH11357. We acknowledge the help and support of the following people: C. Prier and A. Fryszkowska for comments on earlier versions of this manuscript; W. Pinto and F. Tenkorang for analytical assistance; K. Sirk for reaction optimization support; J. Corry and E. Fisher for helpful discussions; S. (G.) Xu, E. Frank and A. Struck for support with biochemical and cell-based assays used in this work; M. Childers, M. Lu, R. Otte, A. Hidle, T. Henderson, J. Jewell, L. Nogle and A. Beard for contributions to synthesis, purification and characterization of the molecules described in the article; L. Miller, A. Petkova, J. Riggins and A. Sowell-Kantz for support of enzyme evolution.

## Author information

### Affiliations

1. Process Research & Development, Merck & Co., Inc., Rahway, NJ, USA

John A. McIntosh, Zhijian Liu, Nastaran Salehi Marzijarani, Jeffrey C. Moore, Nicholas M. Marshall, Jennifer V. Obligacion, Patrick S. Fier, Feng Peng, Jacob H. Forstater, Matthew S. Winston, Chihui An, Mark A. Huffman, Steven P. Miller, Dietrich Steinhuebel, Kaori Hiraga, Grant S. Murphy, Joshua N. Kolev, Amanda Makarewicz, Weilan Pan, Iman Farasat, Rachel S. Bade, Kevin Stone, Erik L. Regalado, Alejandro Crespo, Deeptak Verma, Paul N. Devine, David Tschaen, Keith A. Canada, Paul G. Bulger, Benjamin D. Sherry, Matthew D. Truppo, Rebecca T. Ruck, Louis-Charles Campeau, Guy R. Humphrey, Kevin R. Campos & Matthew L. Maddess

2. Discovery Chemistry, Merck & Co., Inc., Kenilworth, NJ, USA

Brian M. Andresen, Wonsuk Chang, Jongwon Lim, Michael D. Altman, Charles A. Lesburg, B. Wesley Trotter, Jared N. Cumming, Alan Northrup, Donovan Adpressa & David Jonathan Bennett

3. Codexis, Inc., Redwood City, CA, USA

Margie Borra-Garske, Da Duan & Oscar Alvizo

4. Analytical Research & Development, Merck & Co., Inc., Rahway, NJ, USA

Fuh-Rong Tsay, Xiaodong Bu, Benjamin F. Mann, Mirlinda Biba, Erik Guetschow, Erik Hoyt, Erik L. Regalado, Steve Castro, Nelo Rivera, Joseph P. Smith, Fengqiang Wang, Stephanus Axnanda & Zachary E. X. Dance

## Contributions

J.A.M., J.C.M., N.M.M., M.B.-G., M.A.H., C.A., D.D., B.F.M., K.H., G.S.M., J.N.K., A.M., W.P., I.F., O.A., A.C., D.V., M.B., K.A.C. and M.D.T. contributed to enzyme discovery and engineering. Z.L., N.S.M., J.V.O., P.S.F., F.P., J.A.M., M.S.W., C.A., F.-R.T., J.H.F., X.B., R.S.B., J.H.F., K.S., E.G., E.H., E.L.R., S.C., N.R., J.P.S., F.W., S.A. and Z.E.X.D. contributed to reaction optimization, characterization and/or process development. B.M.A., W.C., J.L., M.D.A., C.A.L., D.S., B.W.T., J.N.C., A.N., D.A. and M.L.M. contributed to CDN analogue synthesis and characterization. M.A.H., S.P.M., P.N.D., D.T., P.G.B., B.D.S., R.T.R., L.C.C., D.J.B., G.R.H., K.R.C. and M.L.M. provided intellectual contributions. J.A.M., Z.L., B.M.A., M.B.-G., M.A.H., M.L.M., R.T.R. and L.C.C. prepared the manuscript.

## Corresponding authors

Correspondence to [John A. McIntosh](#), [Zhijian Liu](#) or [Brian M. Andresen](#).

# Ethics declarations

## Competing interests

The authors are current or former employees of Merck Sharp & Dohme Corp., a subsidiary of Merck & Co., Inc., Kenilworth, NJ, USA or Codexis, Inc., which are assignees for patents governing chemical matter, processes and enzyme sequences reported in the article.

## Peer review

### Peer review information

*Nature* thanks Elaine O'Reilly and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

## Additional information

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Supplementary information

### Supplementary Information

Abbreviations, synthetic methods, Supplementary Figs. 1–23, Tables 1 and 2, NMR spectra and references.

### Supplementary Data 1

Wild-type cGAS variant sequences.

### Supplementary Data 2

Wild-type adenylate and guanylate kinase sequences.

### **Supplementary Data 3**

Acetate kinase sequences.

### **Supplementary Data 4**

Sequences of evolved cGAS, GK, AK and AcK variants.

## **Rights and permissions**

### **Reprints and Permissions**

## **About this article**

### **Cite this article**

McIntosh, J.A., Liu, Z., Andresen, B.M. *et al.* A kinase-cGAS cascade to synthesize a therapeutic STING activator. *Nature* **603**, 439–444 (2022).  
<https://doi.org/10.1038/s41586-022-04422-9>

- Received: 18 May 2021
- Accepted: 10 January 2022
- Published: 16 March 2022
- Issue Date: 17 March 2022
- DOI: <https://doi.org/10.1038/s41586-022-04422-9>

### **Share this article**

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-022-04422-9>

| [Section menu](#) | [Main menu](#) |

- Article
- [Published: 16 March 2022](#)

# Limited increases in savanna carbon stocks over decades of fire suppression

- [Yong Zhou](#) [ORCID: orcid.org/0000-0003-2546-8462<sup>1,2</sup>](#),
- [Jenia Singh<sup>3</sup>](#),
- [John R. Butnor<sup>4</sup>](#),
- [Corli Coetsee](#) [ORCID: orcid.org/0000-0001-6797-9311<sup>5,6</sup>](#),
- [Peter B. Boucher<sup>3</sup>](#),
- [Madelon F. Case<sup>2</sup>](#) [nAff7](#),
- [Evan G. Hockridge](#) [ORCID: orcid.org/0000-0001-8139-5427<sup>3</sup>](#),
- [Andrew B. Davies<sup>3</sup>](#) &
- [A. Carla Staver](#) [ORCID: orcid.org/0000-0002-2384-675X<sup>1,2</sup>](#)

[Nature](#) volume **603**, pages 445–449 (2022)

- 1485 Accesses
- 194 Altmetric
- [Metrics details](#)

## Subjects

- [Carbon cycle](#)
- [Fire ecology](#)
- [Tropical ecology](#)

# Abstract

Savannas cover a fifth of the land surface and contribute a third of terrestrial net primary production, accounting for three-quarters of global area burned and more than half of global fire-driven carbon emissions<sup>1,2,3</sup>. Fire suppression and afforestation have been proposed as tools to increase carbon sequestration in these ecosystems<sup>2,4</sup>. A robust quantification of whole-ecosystem carbon storage in savannas is lacking however, especially under altered fire regimes. Here we provide one of the first direct estimates of whole-ecosystem carbon response to more than 60 years of fire exclusion in a mesic African savanna. We found that fire suppression increased whole-ecosystem carbon storage by only  $35.4 \pm 12\%$  (mean  $\pm$  standard error), even though tree cover increased by  $78.9 \pm 29.3\%$ , corresponding to total gains of  $23.0 \pm 6.1 \text{ Mg C ha}^{-1}$  at an average of about  $0.35 \pm 0.09 \text{ Mg C ha}^{-1} \text{ year}^{-1}$ , more than an order of magnitude lower than previously assumed<sup>4</sup>. Frequently burned savannas had substantial belowground carbon, especially in biomass and deep soils. These belowground reservoirs are not fully considered in afforestation or fire-suppression schemes but may mean that the decadal sequestration potential of savannas is negligible, especially weighed against concomitant losses of biodiversity and function.

This is a preview of subscription content

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

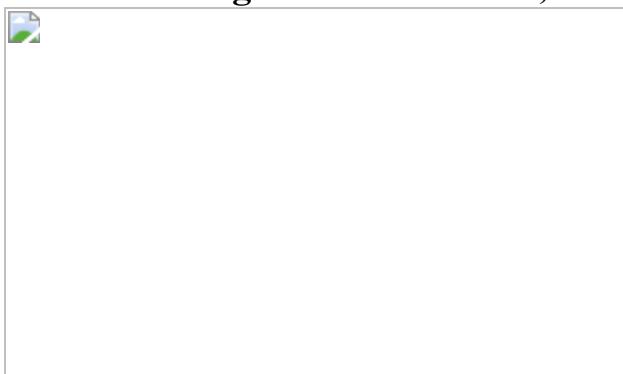
[Buy](#)

All prices are NET prices.

### **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

**Fig. 1: Vegetation characteristics of experimental burn plots in a mesic savanna in Kruger National Park, South Africa.**



**Fig. 2: Changes in whole-ecosystem carbon storage across fire treatments.**



**Fig. 3: Changes in SOC with tree cover and soil sand content.**



## Data and code availability

Data and code are available in the Dryad Digital Repository:  
<https://doi.org/10.5061/dryad.pg4f4qrr5>. Source data are provided with this paper.

## References

1. Giglio, L., Schroeder, W. & Justice, C. O. The collection 6 MODIS active fire detection algorithm and fire products. *Remote Sens. Environ.* **178**, 31–41 (2016).
2. Grace, J., José, J. S., Meir, P., Miranda, H. S. & Montes, R. A. Productivity and carbon fluxes of tropical savannas. *J. Biogeogr.* **33**, 387–400 (2006).
3. Van Der Werf, G. R. et al. Global fire emissions estimates during 1997–2016. *Earth Syst. Sci. Data* **9**, 697–720 (2017).
4. Bastin, J.-F. et al. The global tree restoration potential. *Science* **365**, 76–79 (2019).

5. Russell-Smith, J. et al. Opportunities and challenges for savanna burning emissions abatement in southern Africa. *J. Environ. Manage.* **288**, 112414 (2021).
6. Andela, N. et al. A human-driven decline in global burned area. *Science* **356**, 1356–1362 (2017).
7. Wu, C. et al. Historical and future global burned area with changing climate and human demography. *One Earth* **4**, 517–530 (2021).
8. Pellegrini, A. F. A. et al. Fire frequency drives decadal changes in soil carbon and nitrogen and ecosystem productivity. *Nature* **553**, 194–198 (2018).
9. Higgins, S. I. et al. Effects of four decades of fire manipulation on woody vegetation structure in savanna. *Ecology* **88**, 1119–1125 (2007).
10. Staver, A. C., Archibald, S. & Levin, S. A. The global extent and determinants of savanna and forest as alternative biome states. *Science* **334**, 230–232 (2011).
11. Shi, Z. et al. The age distribution of global soil carbon inferred from radiocarbon measurements. *Nat. Geosci.* **13**, 555–559 (2020).
12. Pellegrini, A. F. A., Hedin, L. O., Staver, A. C. & Govender, N. Fire alters ecosystem carbon and nutrients but not plant nutrient stoichiometry or composition in tropical savanna. *Ecology* **96**, 1275–1285 (2015).
13. Tilman, D. et al. Fire suppression and ecosystem carbon storage. *Ecology* **81**, 2680–2685 (2000).
14. Mokany, K., Raison, R. J. & Prokushkin, A. S. Critical analysis of root:shoot ratios in terrestrial biomes. *Glob. Change Biol.* **12**, 84–96 (2006).
15. de Miranda, S. D. C. et al. Regional variations in biomass distribution in Brazilian savanna woodland. *Biotropica* **46**, 125–138 (2014).

16. Wigley, B. J., Cramer, M. D. & Bond, W. J. Sapling survival in a frequently burnt savanna: mobilisation of carbon reserves in *Acacia karroo*. *Plant Ecol.* **203**, 1 (2009).
17. Sankaran, M. et al. Determinants of woody cover in African savannas. *Nature* **438**, 846–849 (2005).
18. Staver, A. C., Botha, J. & Hedin, L. Soils and fire jointly determine vegetation structure in an African savanna. *New Phytol.* **216**, 1151–1160 (2017).
19. Zhou, Y., Wigley, B. J., Case, M. F., Coetsee, C. & Staver, A. C. Rooting depth as a key woody functional trait in savannas. *New Phytol.* **227**, 1350–1361 (2020).
20. Govender, N., Trollope, W. S. W., Van, & Wilgen, B. W. The effect of fire season, fire frequency, rainfall and management on fire intensity in savanna vegetation in South Africa. *J. Appl. Ecol.* **43**, 748–758 (2006).
21. Colgan, M. S., Asner, G. P. & Swemmer, T. Harvesting tree biomass at the stand level to assess the accuracy of field and airborne biomass estimation in savannas. *Ecol. Appl.* **23**, 1170–1184 (2013).
22. Davies, A. B. & Asner, G. P. Elephants limit aboveground carbon gains in African savannas. *Glob. Change Biol.* **25**, 1368–1382 (2019).
23. Butnor, J. R. et al. Utility of ground-penetrating radar as a root biomass survey tool in forest systems. *Soil Sci. Soc. Am. J.* **67**, 1607–1615 (2003).
24. Staver, A. C., Wigley-Coetsee, C. & Botha, J. Grazer movements exacerbate grass declines during drought in an African savanna. *J. Ecol.* **107**, 1482–1491 (2019).
25. Ryan, C. M., Williams, M. & Grace, J. Above- and belowground carbon stocks in a miombo woodland landscape of Mozambique. *Biotropica* **43**, 423–432 (2011).

26. Swezy, D. M. & Agee, J. K. Prescribed-fire effects on fine-root and tree mortality in old-growth ponderosa pine. *Can. J. For. Res.* **21**, 626–634 (1991).
27. Canadell, J. et al. Maximum rooting depth of vegetation types at the global scale. *Oecologia* **108**, 583–595 (1996).
28. Coetsee, C., Bond, W. J. & February, E. C. Frequent fire affects soil nitrogen and carbon in an African savanna by changing woody cover. *Oecologia* **162**, 1027–1034 (2010).
29. Holdo, R. M., Mack, M. C. & Arnold, S. G. Tree canopies explain fire effects on soil nitrogen, phosphorus and carbon in a savanna ecosystem. *J. Veg. Sci.* **23**, 352–360 (2012).
30. Lloyd, J. et al. Contributions of woody and herbaceous vegetation to tropical savanna ecosystem productivity: a quasi-global estimate. *Tree Physiol.* **28**, 451–468 (2008).
31. Wigley, B. J., Augustine, D. J., Coetsee, C., Ratnam, J. & Sankaran, M. Grasses continue to trump trees at soil carbon sequestration following herbivore exclusion in a semiarid African savanna. *Ecology* **101**, e03008 (2020).
32. Khomo, L., Trumbore, S., Bern, C. R. & Chadwick, O. A. Timescales of carbon turnover in soils with mixed crystalline mineralogies. *Soil* **3**, 17–30 (2017).
33. Six, J., Conant, R. T., Paul, E. A. & Paustian, K. Stabilization mechanisms of soil organic matter: implications for C-saturation of soils. *Plant Soil* **241**, 155–176 (2002).
34. Abreu, R. C. R. et al. The biodiversity cost of carbon sequestration in tropical savanna. *Sci. Adv.* **3**, e1701284 (2017).
35. Bond, W. J., Stevens, N., Midgley, G. F. & Lehmann, C. E. The trouble with trees: afforestation plans for Africa. *Trends Ecol. Evol.* **34**, 963–965 (2019).

36. West, T. A., Börner, J. & Fearnside, P. M. Climatic benefits from the 2006–2017 avoided deforestation in Amazonian Brazil. *Front. For. Glob. Change* **2**, 52 (2019).
37. Aleman, J. C., Blarquez, O. & Staver, C. A. Land-use change outweighs projected effects of changing rainfall on tree cover in sub-Saharan Africa. *Glob. Change Biol.* **22**, 3013–3025 (2016).
38. Huang, J., Yu, H., Guan, X., Wang, G. & Guo, R. Accelerated dryland expansion under climate change. *Nat. Clim. Change* **6**, 166–171 (2016).
39. Ratajczak, Z., Nippert, J. B. & Collins, S. L. Woody encroachment decreases diversity across North American grasslands and savannas. *Ecology* **93**, 697–703 (2012).
40. Smit, I. P. & Prins, H. H. Predicting the effects of woody encroachment on mammal communities, grazing biomass and fire frequency in African savannas. *PLoS One* **10**, e0137857 (2015).
41. Huxman, T. E. et al. Ecohydrological implications of woody plant encroachment. *Ecology* **86**, 308–319 (2005).
42. Hermoso, V., Regos, A., Morán-Ordóñez, A., Duane, A. & Brotons, L. Tree planting: a double-edged sword to fight climate change in an era of megafires. *Glob. Change Biol.* **27**, 3001–3003 (2021).
43. Venter F. A. *Classification of Land for Management Planning in the Kruger National Park*. PhD thesis, Univ. South Africa (1990).
44. Biggs, R., Biggs, H. C., Dunne, T. T., Govender, N. & Potgieter, A. L. F. Experimental burn plot trial in the Kruger National Park: history, experimental design and suggestions for data analysis. *Koedoe* **46**, 15 (2003).
45. Codron, J. et al. Taxonomic, anatomical, and spatio-temporal variations in the stable carbon and nitrogen isotopic compositions of

- plants from an African savanna. *J. Archaeol. Sci.* **32**, 1757–1772 (2005).
46. Zhou, Y., Boutton, T. W. & Ben Wu, X. Soil carbon response to woody plant encroachment: importance of spatial heterogeneity and deep soil storage. *J. Ecol.* **105**, 1738–1749 (2017).
  47. Sheldrick B. & Wang C. In *Soil Sampling and Methods of Analysis* (ed. Carter, M. R.) 499–511 (CRC Press, 1993).
  48. Butnor, J. R. et al. Surface-based GPR underestimates below-stump root biomass. *Plant Soil* **402**, 47–62 (2016).
  49. Pau, G., Fuchs, F., Sklyar, O., Boutros, M. & Huber, W. EBImage—an R package for image processing with applications to cellular phenotypes. *Bioinformatics* **26**, 979–981 (2010).
  50. Hirano, Y. et al. Limiting factors in the detection of tree roots using ground-penetrating radar. *Plant Soil* **319**, 15–24 (2009).
  51. Popescu, S. C. & Wynne, R. H. Seeing the trees in the forest. *Photogramm. Eng. Remote Sensing* **70**, 589–604 (2004).
  52. Case, M. F., Wigley-Coetsee, C., Nzima, N., Scogings, P. F. & Staver, A. C. Severe drought limits trees in a semi-arid savanna. *Ecology* **100**, e02842 (2019).
  53. Beucher S. & Meyer F. In *Mathematical Morphology in Image Processing* (ed. Dougherty, E. R.) 433–481 (CRC Press, 1993).
  54. Nickless, A., Scholes, R. J. & Archibald, S. A method for calculating the variance and confidence intervals for tree biomass estimates obtained from allometric equations. *S. Afr. J. Sci.* **107**, 1–10 (2011).
  55. Plowright A. & Roussel J.-R. ForestTools: analyzing remotely sensed forest data. R package version 0.2.1. <https://CRAN.R-project.org/package=ForestTools> (2020).

56. Hijmans R. J. raster: geographic data analysis and modeling. R package version 3.3-7. <https://CRAN.R-project.org/package=raster> (2020).
57. Penman J. et al. (eds) *Good Practice Guidance for Land Use, Land-Use Change and Forestry* (Intergovernmental Panel on Climate Change, 2003).
58. Kuznetsova, A., Brockhoff, P. & Christensen, R. lmerTest package: tests in linear mixed effects models. *J. Stat. Softw.* **82**, 1–26 (2017).

## Acknowledgements

We gratefully acknowledge the logistical support provided by South African National Parks staff in Kruger National Park. Y.Z. was supported by a G. Evelyn Hutchinson Environmental Postdoctoral Fellowship from the Yale Institute for Biospheric Studies, A.C.S. was partially supported by a grant from the United States National Science Foundation (NSF MSB-1802453) and by funding from Yale University, J.S., P.B.B., E.G.H. and A.B.D. from Harvard University, and J.R.B. from the USDA Forest Service, Southern Research Station.

## Author information

### Author notes

#### 1. Madelon F. Case

Present address: Institute of Ecology and Evolution, University of Oregon, Eugene, OR, USA

## Affiliations

#### 1. Yale Institute for Biospheric Studies, Yale University, New Haven, CT, USA

Yong Zhou & A. Carla Staver

2. Department of Ecology and Evolutionary Biology, Yale University,  
New Haven, CT, USA

Yong Zhou, Madelon F. Case & A. Carla Staver

3. Department of Organismic and Evolutionary Biology, Harvard  
University, Cambridge, MA, USA

Jenia Singh, Peter B. Boucher, Evan G. Hockridge & Andrew B.  
Davies

4. USDA Forest Service, Southern Research Station, University of  
Vermont, Burlington, VT, USA

John R. Butnor

5. Scientific Services, SANParks, Skukuza, South Africa

Corli Coetsee

6. School of Natural Resource Management, Nelson Mandela University,  
George, South Africa

Corli Coetsee

## Contributions

Conceptualization: Y.Z. and A.C.S.; methodology: Y.Z., J.S., J.R.B., P.B.B.,  
A.B.D. and A.C.S.; investigation: Y.Z., C.C., M.F.C., E.G.H., A.B.D. and  
A.C.S.; visualization: Y.Z. and A.C.S.; funding acquisition: A.B.D. and  
A.C.S.; project administration: A.B.D. and A.C.S.; supervision: A.B.D. and  
A.C.S.; writing — original draft: Y.Z. and A.C.S.; writing — review and  
editing: Y.Z., J.S., J.R.B., C.C., P.B.B., M.F.C., E.G.H., A.B.D. and A.C.S.

## Corresponding authors

Correspondence to [Yong Zhou](#) or [A. Carla Staver](#).

# Ethics declarations

## Competing interests

The authors declare no competing interests.

## Peer review

### Peer review information

*Nature* thanks Sebastian Dötterl, Niall Hanan and Douglas Morton for their contribution to the peer review of this work. Peer reviewer reports are available.

## Additional information

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Extended data figures and tables

[\*\*Extended Data Fig. 1 An example showing belowground to aboveground biomass allocation for resprouting \*Terminalia sericea\*.\*\*](#)

**a, b**, Five *T. sericea* individuals that have experienced annual burning were excavated in the Pretoriuskop landscape in Kruger National Park, South Africa. **c**, The difference between aboveground and belowground biomass and the ratio of belowground to aboveground biomass was 19.5. The box plots show medians (that is, 50th percentile), 25th and 75th percentiles, and the approximate 95% confidence interval for five replicates.

[\*\*Extended Data Fig. 2 Maps showing the study site.\*\*](#)

Maps showing the locations of different fire treatments (that is, annual, triennial and unburned) examined in this study and located in each string (Fayi, Kambeni, Numbi and Shabeni) across the Pretoriuskop landscape at Kruger National Park, South Africa. Base map for South Africa modified from Natural Earth.

**Extended Data Fig. 3 Changes in SOC storage and soil  $\delta^{13}\text{C}$  across different fire treatments throughout the 60-cm soil column.**

Effects of fire treatments on total SOC storage ( $\text{Mg C ha}^{-1}$ ) (a), soil  $\delta^{13}\text{C}$  (‰) (b),  $\text{C}_3$ -derived SOC storage (that is, from woody plants) ( $\text{Mg C ha}^{-1}$ ) (c) and  $\text{C}_4$ -derived SOC storage (that is, from grasses) ( $\text{Mg C ha}^{-1}$ ) (d). Values are mean  $\pm$  standard errors ( $n = 4$ ).

**Extended Data Fig. 4 Long-term monitoring of grass fuel loads and their correlation to LiDAR-derived mean grass height.**

a–c, Grass fuel loads ( $\text{kg ha}^{-1}$ ) for annual (a), triennial (b) and April B2 (that is, burning in April for every two years, as a proxy for unburned) (c) treatments from 1982 to 2009 for different strings at the Pretoriuskop landscape in Kruger National Park, South Africa. Disconnected lines indicate missing data for specific years. d, The correlation between averaged grass fuel loads from 1982 to 2009 and LiDAR-derived mean grass heights (m) ( $R^2 = 0.38$ ,  $P = 0.03$ ). The mean grass height was calculated by averaging heights of pixels that range from 0.05 to 0.5 m in the CHM derived from LiDAR. Please note especially that, in panel d, LiDAR-derived mean grass height was estimated from the unburned treatment itself, but that field-estimated grass fuel load was estimated from the April B2 treatment as a proxy (as grass fuel load is not routinely measured for the unburned treatment).

**Extended Data Fig. 5 The uncertainty of coarse lateral and taproot biomass estimates.**

**a**, The uncertainty of coarse lateral and taproot biomass for each treatment replicate. Error bars indicate the 95% confidence interval for coarse lateral and taproot biomass estimates derived from fitting regression lines (see Supplementary Figs. 5 and 10). Coarse-lateral-root biomass estimates were significantly correlated with taproot biomass estimates ( $R^2 = 0.75$ ,  $P < 0.001$ ). Letters F, K, N and S indicate Fayi, Kambeni, Numbi and Shabeni strings at the Pretoriuskop landscape in Kruger National Park, South Africa; letters A, T and U indicate annual, triennial and unburned treatments. **b, c** The uncertainty of (that is, lower bound, mean and upper bound) coarse lateral and taproot biomass across different fire treatments. The box plots show medians (that is, 50th percentile), 25th and 75th percentiles, and the 95% confidence interval for four replicates. Points in **b** and **c** indicate outliers.

### Extended Data Fig. 6 Depth distribution of coarse-lateral-root biomass across fire treatments and soil sand content.

**a**, Depth distribution of the GPR index (% in the number of pixels above the threshold for root detections) as an indicator of coarse-lateral-root biomass allocation throughout the soil column across different fire treatments at each string. Horizontal lines indicate the depth (cm) at which the GPR index reaches 50% of the total detections in the 60-cm soil column. **b**, Effects of fire treatment on the depth distribution of coarse-lateral-root biomass ( $P = 0.51$ ). The box plots show medians (that is, 50th percentile), 25th and 75th percentiles, and the 95% confidence interval for four replicates. **c**, The correlation between soil sand content (%) and depth distribution of coarse-lateral-root biomass ( $R^2 = 0.61$ ,  $P = 0.003$ ). The regression line indicates the significant linear fit and the shaded bands illustrate the 95% confidence interval of the linear fit.

### Extended Data Fig. 7 The correlation between ratio of belowground to aboveground carbon storage and tree cover (%) ( $R^2 = 0.83$ , $P < 0.0001$ ).

The regression line indicates the significant linear fit and the shaded bands illustrate the 95% confidence interval of the linear fit.

## Extended Data Fig. 8 The validation of the object-based method to estimate aboveground woody biomass.

**a**, The correlation between LiDAR-derived stem density for trees with height > 5m (trees ha<sup>-1</sup>) and field-measured stem density (trees ha<sup>-1</sup>). The field-measured stem density was from ref. [52](#), which surveyed tree heights in eight 10-m-radius plots at each annual, triennial and unburned treatment in Kambeni, Numbi and Shabeni strings at the Pretoriuskop landscape in Kruger National Park, South Africa. The regression line indicates the significant linear fit and the shaded bands illustrate the 95% confidence interval of the linear fit. The dashed line indicates the 1:1 line. **b**, Differences in aboveground woody biomass between allometric-derived, object-based and plot-averaged estimates. The allometric-derived biomass estimation was on the basis of species-specific allometric equations developed in ref. [54](#), which predict aboveground woody biomass from DBH. This estimation was calculated for trees with DBH > 5 cm in each 10 × 10-m plot. The plot-averaged LiDAR biomass was estimated using an allometric equation derived from on-the-ground plot-level sampling relating aboveground woody biomass to LiDAR-derived canopy height and canopy area (aboveground woody biomass = -11.5 + 25.8 \* canopy height \* canopy area); please refer to ref. [21](#) for more details. The canopy height and canopy area were averaged across pixels with height > 0.5 m in each 30-m-radius plot. The box plots show medians (that is, 50th percentile), 25th and 75th percentiles, and the 95% confidence interval for four replicates. Points in **b** indicate outliers.

**Extended Data Table 1 Results comparing the relative effect of different fire treatments on each component of ecosystem carbon storage from linear mixed-effects models using fire treatments as the fixed effect and string as a random effect**

## **Supplementary information**

### Supplementary Information

This file contains Supplementary Figures 1–10 and Supplementary Tables 1–4.

## [Peer Review File](#)

## **Source data**

### [Source Data Fig. 1](#)

### [Source Data Fig. 2](#)

### [Source Data Fig. 3](#)

## **Rights and permissions**

### [Reprints and Permissions](#)

## **About this article**

### **Cite this article**

Zhou, Y., Singh, J., Butnor, J.R. *et al.* Limited increases in savanna carbon stocks over decades of fire suppression. *Nature* **603**, 445–449 (2022).  
<https://doi.org/10.1038/s41586-022-04438-1>

- Received: 24 June 2021
- Accepted: 14 January 2022
- Published: 16 March 2022
- Issue Date: 17 March 2022
- DOI: <https://doi.org/10.1038/s41586-022-04438-1>

## Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

## Savannahs store carbon despite frequent fires

- Niall P. Hanan
- Anthony M. Swemmer

News & Views 16 Mar 2022

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-022-04438-1>

| [Section menu](#) | [Main menu](#) |

- Article
- [Published: 16 March 2022](#)

# New land-use-change emissions indicate a declining CO<sub>2</sub> airborne fraction

- [Margreet J. E. van Marle](#) ORCID: [orcid.org/0000-0001-7473-5550](https://orcid.org/0000-0001-7473-5550)<sup>1,2 na1</sup>,
- [Dave van Wees](#) ORCID: [orcid.org/0000-0001-5565-7155](https://orcid.org/0000-0001-5565-7155)<sup>1 na1</sup>,
- [Richard A. Houghton](#) ORCID: [orcid.org/0000-0002-3298-7028](https://orcid.org/0000-0002-3298-7028)<sup>3</sup>,
- [Robert D. Field](#) ORCID: [orcid.org/0000-0003-3001-5580](https://orcid.org/0000-0003-3001-5580)<sup>4,5</sup>,
- [Jan Verbesselt](#) ORCID: [orcid.org/0000-0001-7923-4309](https://orcid.org/0000-0001-7923-4309)<sup>6</sup> &
- [Guido. R. van der Werf](#) ORCID: [orcid.org/0000-0001-9042-8630](https://orcid.org/0000-0001-9042-8630)<sup>1</sup>

[Nature](#) volume 603, pages 450–454 (2022)

- 1467 Accesses
- 183 Altmetric
- [Metrics details](#)

## Subjects

- [Climate and Earth system modelling](#)
- [Projection and prediction](#)

## Abstract

About half of the anthropogenic CO<sub>2</sub> emissions remain in the atmosphere and half are taken up by the land and ocean<sup>1</sup>. If the carbon uptake by land and ocean sinks becomes less efficient, for example, owing to warming oceans<sup>2</sup> or thawing permafrost<sup>3</sup>, a larger fraction of anthropogenic emissions will remain in the atmosphere, accelerating climate change. Changes in the efficiency of the carbon sinks can be estimated indirectly by analysing trends in the airborne fraction, that is, the ratio between the atmospheric growth rate and anthropogenic emissions of CO<sub>2</sub> (refs.

[4,5,6,7,8,9,10](#)). However, current studies yield conflicting results about trends in the airborne fraction, with emissions related to land use and land cover change (LULCC) contributing the largest source of uncertainty<sup>7,11,12</sup>. Here we construct a LULCC emissions dataset using visibility data in key deforestation zones. These visibility observations are a proxy for fire emissions<sup>13,14</sup>, which are — in turn — related to LULCC<sup>15,16</sup>. Although indirect, this provides a long-term consistent dataset of LULCC emissions, showing that tropical deforestation emissions increased substantially (0.16 Pg C decade<sup>-1</sup>) since the start of CO<sub>2</sub> concentration measurements in 1958. So far, these emissions were thought to be relatively stable, leading to an increasing airborne fraction<sup>4,5</sup>. Our results, however, indicate that the CO<sub>2</sub> airborne fraction has decreased by  $0.014 \pm 0.010$  decade<sup>-1</sup> since 1959. This suggests that the combined land–ocean sink has been able to grow at least as fast as anthropogenic emissions.

This is a preview of subscription content

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

[Subscribe to Journal](#)

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

[Buy article](#)

Get time limited or full article access on ReadCube.

\$32.00

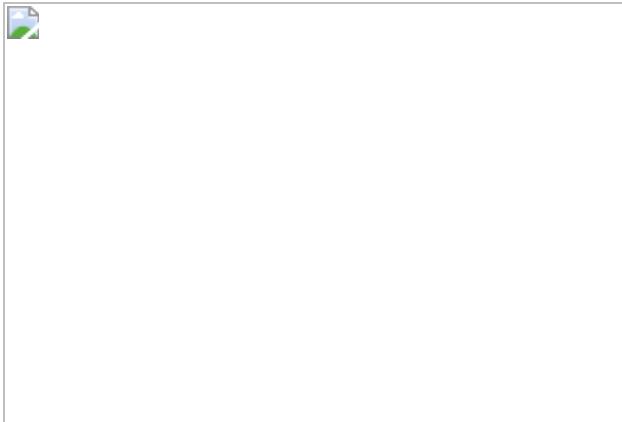
[Buy](#)

All prices are NET prices.

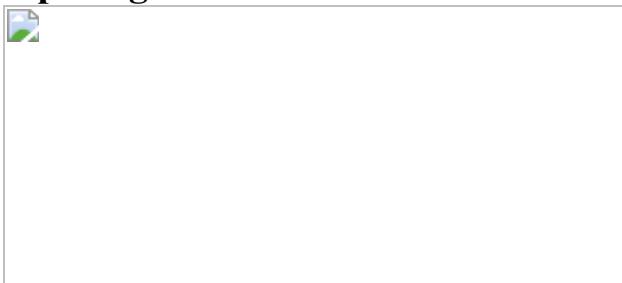
### **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

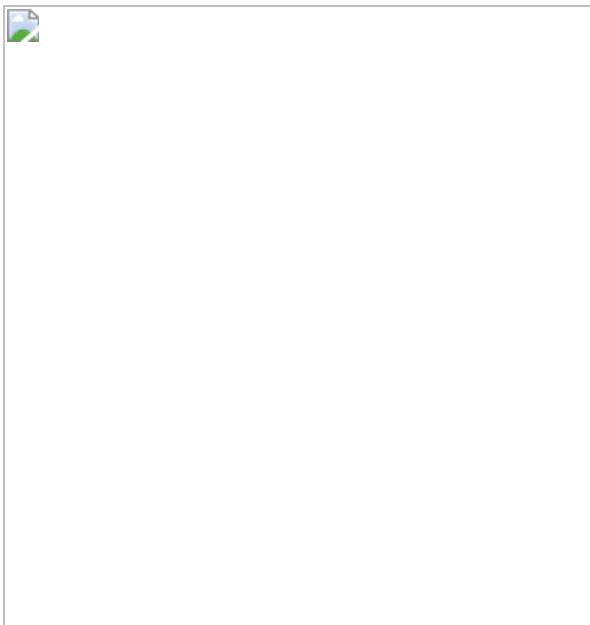
**Fig. 1: Visibility-derived LULCC emissions for two key deforestation regions.**



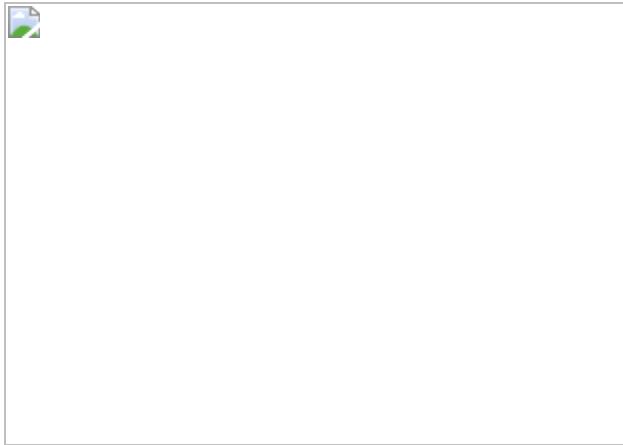
**Fig. 2: Global LULCC emissions and other time series required for computing the AF.**



**Fig. 3: AF trend estimates on the basis of three different LULCC emission datasets.**



**Fig. 4: Sensitivity of the AF trend to average annual LULCC emissions and its slope.**



## Data availability

GCP data are publicly available at

<https://www.globalcarbonproject.org/carbonbudget>. The ENSO index was based on anomalies in Niño 3 SSTs, which are publicly available at [https://psl.noaa.gov/gcos\\_wgsp/Timeseries /Data/nino3.long.anom.data](https://psl.noaa.gov/gcos_wgsp/Timeseries /Data/nino3.long.anom.data).

Gridded visibility-based fire emissions can be found at the CMIP6 forcing data repository (<https://esgf-node.llnl.gov/search/input4mips/>). All input data, including the Houghton and Nassikas dataset<sup>27</sup>, have been made available at <https://doi.org/10.5281/zenodo.5617953>.

## Code availability

The Python code that was used to assimilate the raw data and perform the analyses is available at <https://doi.org/10.5281/zenodo.5617953>.

## References

1. Canadell, J. G. et al. 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* (eds Masson-Delmotte, V. et al.) (Cambridge Univ. Press, 2021) (in the press).
2. McKinley, G. A., Fay, A. R., Takahashi, T. & Metzl, N. Convergence of atmospheric and North Atlantic carbon dioxide trends on

- multidecadal timescales. *Nat. Geosci.* **4**, 606–610 (2011).
3. Schuur, E. A. G. et al. Climate change and the permafrost carbon feedback. *Nature* **520**, 171–179 (2015).
  4. Le Quéré, C., Raupach, M. R., Canadell, J. G. & Al, G. M. Trends in the sources and sinks of carbon dioxide. *Nat. Geosci.* **2**, 831–836 (2009).
  5. Raupach, M. R. et al. The declining uptake rate of atmospheric CO<sub>2</sub> by land and ocean sinks. *Biogeosciences* **11**, 3453–3475 (2014).
  6. Knorr, W. Is the airborne fraction of anthropogenic CO<sub>2</sub> emissions increasing? *Geophys. Res. Lett.* **36**, L21710 (2009).
  7. Gloor, M., Sarmiento, J. L. & Gruber, N. What can be learned about carbon cycle climate feedbacks from the CO<sub>2</sub> airborne fraction? *Atmos. Chem. Phys.* **10**, 7739–7751 (2010).
  8. Keenan, T. F. et al. Recent pause in the growth rate of atmospheric CO<sub>2</sub> due to enhanced terrestrial carbon uptake. *Nat. Commun.* **7**, 13428 (2016).
  9. Sarmiento, J. L. et al. Trends and regional distributions of land and ocean carbon sinks. *Biogeosciences* **7**, 2351–2367 (2010).
  10. Ballantyne, A. P., Alden, C. B., Miller, J. B., Tans, P. P. & White, J. W. C. Increase in observed net carbon dioxide uptake by land and oceans during the past 50 years. *Nature* **488**, 70–72 (2012).
  11. Friedlingstein, P. et al. Global carbon budget 2020. *Earth Syst. Sci. Data* **12**, 3269–3340 (2020).
  12. Mahowald, N. M. et al. Interactions between land use change and carbon cycle feedbacks. *Global Biogeochem. Cycles* **31**, 96–113 (2017).

13. Field, R. D., van der Werf, G. R. & Shen, S. S. P. Human amplification of drought-induced biomass burning in Indonesia since 1960. *Nat. Geosci.* **2**, 185–188 (2009).
14. van Marle, M. J. E. et al. Fire and deforestation dynamics in Amazonia (1973–2014). *Global Biogeochem. Cycles* **31**, 24–38 (2017).
15. Morton, D. C. et al. Cropland expansion changes deforestation dynamics in the southern Brazilian Amazon. *Proc. Natl Acad. Sci. USA* **103**, 14637–14641 (2006).
16. van Wees, D. et al. The role of fire in global forest loss dynamics. *Glob. Change Biol.* **27**, 2377–2391 (2021).
17. Cox, P. M., Betts, R. A., Jones, C. D., Spall, S. A. & Totterdell, I. J. Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature* **408**, 184–187 (2000).
18. Friedlingstein, P. & Prentice, I. Carbon–climate feedbacks: a review of model and observation based estimates. *Curr. Opin. Environ. Sustain.* **2**, 251–257 (2010).
19. Pan, Y. et al. A large and persistent carbon sink in the world’s forests. *Science* **333**, 988–993 (2011).
20. Piao, S. et al. Net carbon dioxide losses of northern ecosystems in response to autumn warming. *Nature* **451**, 49–52 (2008).
21. Hubau, W. et al. Asynchronous carbon sink saturation in African and Amazonian tropical forests. *Nature* **579**, 80–87 (2020).
22. Le Quéré, C. et al. Saturation of the Southern Ocean CO<sub>2</sub> sink due to recent climate change. *Science* **316**, 1735–1738 (2007).
23. Miettinen, J., Shi, C. & Liew, S. C. Deforestation rates in insular Southeast Asia between 2000 and 2010. *Glob. Change Biol.* **17**, 2261–2270 (2011).

24. Morton, D. C. et al. Agricultural intensification increases deforestation fire activity in Amazonia. *Glob. Change Biol.* **14**, 2262–2275 (2008).
25. Otón, G., Lizundia-Loiola, J., Pettinari, M. L. & Chuvieco, E. Development of a consistent global long-term burned area product (1982–2018) based on AVHRR-LTDR data. *Int. J. Appl. Earth Obs. Geoinf.* **103**, 102473 (2021).
26. van der Werf, G. R. et al. Global fire emissions estimates during 1997–2016. *Earth Syst. Sci. Data* **9**, 697–720 (2017).
27. Houghton, R. A. & Nassikas, A. A. Global and regional fluxes of carbon from land use and land-cover change 1850–2015. *Global Biogeochem. Cycles* **31**, 456–472 (2017).
28. Brondizio, E. S. & Moran, E. F. Level-dependent deforestation trajectories in the Brazilian Amazon from 1970 to 2001. *Popul. Environ.* **34**, 69–85 (2012).
29. Houghton, R. A. How well do we know the flux of CO<sub>2</sub> from land-use change? *Tellus B Chem. Phys. Meteorol.* **62**, 337–351 (2010).
30. van der Werf, G. R. et al. Continental-scale partitioning of fire emissions during the 1997 to 2001 El Niño/La Niña period. *Science* **303**, 73–76 (2004).
31. Tans, P. & Keeling, R. Trends in atmospheric carbon dioxide. *National Oceanic & Atmospheric Administration, Earth System Research Laboratories (NOAA/ESRL) and Scripps Institution of Oceanography*. <http://www.esrl.noaa.gov/gmd/ccgg/trends/> and <http://scrippsc02.ucsd.edu/> (accessed 24 January 2021).
32. Boden, T. A., Marland, G. & Andres, R. J. Global, regional, and national fossil-fuel CO<sub>2</sub> emissions. *Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy*. [http://cdiac.ornl.gov/trends/emis/overview\\_2013.html](http://cdiac.ornl.gov/trends/emis/overview_2013.html) [https://doi.org/10.3334/CDIAC/00001\\_V2016](https://doi.org/10.3334/CDIAC/00001_V2016) (2016).

33. Ramo, R. et al. African burned area and fire carbon emissions are strongly impacted by small fires undetected by coarse resolution satellite data. *Proc. Natl Acad. Sci. USA* **118**, e2011160118 (2021).
34. Booth, B. B. B. et al. Narrowing the range of future climate projections using historical observations of atmospheric CO<sub>2</sub>. *J. Clim.* **30**, 3039—3053 (2017).
35. Hausfather, Z. & Peters, G. P. Emissions – the ‘business as usual’ story is misleading. *Nature* **577**, 618–620 (2020).
36. Schwalm, C. R., Glendon, S. & Duffy, P. B. RCP8.5 tracks cumulative CO<sub>2</sub> emissions. *Proc. Natl Acad. Sci. USA* **117**, 19656–19657 (2020).
37. Jones, C. et al. Twenty-first-century compatible CO<sub>2</sub> emissions and airborne fraction simulated by CMIP5 earth system models under four representative concentration pathways. *J. Clim.* **26**, 4398–4413 (2013).
38. Gruber, N. Carbon at the coastal interface. *Nature* **517**, 148–149 (2015).
39. Zaehle, S. & Friend, A. D. Carbon and nitrogen cycle dynamics in the O-CN land surface model: 1. Model description, site-scale evaluation, and sensitivity to parameter estimates. *Global Biogeochem. Cycles* **24**, GB1005 (2010).
40. Mercado, L. M. et al. Impact of changes in diffuse radiation on the global land carbon sink. *Nature* **458**, 1014–1017 (2009).
41. Ballantyne, A. et al. Accelerating net terrestrial carbon uptake during the warming hiatus due to reduced respiration. *Nat. Clim. Change* **7**, 148–152 (2017).
42. Watson, A. J. et al. Revised estimates of ocean-atmosphere CO<sub>2</sub> flux are consistent with ocean carbon inventory. *Nat. Commun.* **11**, 4422 (2020).

43. Wang, S. et al. Recent global decline of CO<sub>2</sub> fertilization effects on vegetation photosynthesis. *Science* **370**, 1295–1300 (2020).
44. Trade. FAOSTAT online database. *Food and Agriculture Organization of the United Nations* <http://www.fao.org/faostat/en/#data> (accessed 19 August 2021).
45. Gilfillan, D. & Marland, G. CDIAC-FF: global and national CO<sub>2</sub> emissions from fossil fuel combustion and cement manufacture: 1751–2017. *Earth Syst. Sci. Data* **13**, 1667–1680 (2021).
46. Liu, Z. et al. Reduced carbon emission estimates from fossil fuel combustion and cement production in China. *Nature* **524**, 335–338 (2015).
47. Houghton, R. A. Revised estimates of the annual net flux of carbon to the atmosphere from changes in land use and land management 1850–2000. *Tellus B Chem. Phys. Meteorol.* **55**, 378–390 (2003).
48. Hooijer, A. et al. Current and future CO<sub>2</sub> emissions from drained peatlands in Southeast Asia. *Biogeosciences* **7**, 1505–1514 (2010).
49. Hansis, E., Davis, S. J. & Pongratz, J. Relevance of methodological choices for accounting of land use change carbon fluxes. *Global Biogeochem. Cycles* **29**, 1230–1246 (2015).
50. Gasser, T. et al. Historical CO<sub>2</sub> emissions from land use and land cover change and their uncertainty. *Biogeosciences* **17**, 4075–4101 (2020).
51. Field, R. D. et al. Indonesian fire activity and smoke pollution in 2015 show persistent nonlinear sensitivity to El Niño-induced drought. *Proc. Natl Acad. Sci. USA* **113**, 9204–9209 (2016).
52. van Marle, M. J. E. et al. Historic global biomass burning emissions for CMIP6 (BB4CMIP) based on merging satellite observations with proxies and fire models (1750–2015). *Geosci. Model Dev.* **10**, 3329–3357 (2017).

53. van der Werf, G. R. et al. CO<sub>2</sub> emissions from forest loss. *Nat. Geosci.* **2**, 737–738 (2009).
54. Dlugokencky, E. & Tans, P. Trends in atmospheric carbon dioxide. *National Oceanic & Atmospheric Administration, Earth System Research Laboratory (NOAA/ESRL)* [http://www.esrl.noaa.gov/gmd/ccgg/trends/gl\\_gr.html](http://www.esrl.noaa.gov/gmd/ccgg/trends/gl_gr.html) (accessed 19 August 2021).
55. Gregg, J. S., Andres, R. J. & Marland, G. China: emissions pattern of the world leader in CO<sub>2</sub> emissions from fossil fuel consumption and cement production. *Geophys. Res. Lett.* **35**, L08806 (2008).
56. Raupach, M. R., Canadell, J. G. & Le Quéré, C. Anthropogenic and biophysical contributions to increasing atmospheric CO<sub>2</sub> growth rate and airborne fraction. *Biogeosciences* **5**, 1601–1613 (2008).
57. Ammann, C. M. A monthly and latitudinally varying volcanic forcing dataset in simulations of 20th century climate. *Geophys. Res. Lett.* **30**, 483–487 (2003).
58. Riahi, K. et al. RCP 8.5—a scenario of comparatively high greenhouse gas emissions. *Clim. Change* **109**, 33–57 (2011).

## Acknowledgements

This research was funded by the European Research Council (ERC) grant number 280061 and the Netherlands Organization for Scientific Research (NWO) (Vici scheme research programme, no. 016.160.324).

## Author information

### Author notes

1. These authors contributed equally: Margreet J. E. van Marle, Dave van Wees

## Affiliations

1. Department of Earth Sciences, Faculty of Science, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

Margreet J. E. van Marle, Dave van Wees & Guido. R. van der Werf

2. Deltares, Delft, The Netherlands

Margreet J. E. van Marle

3. Woodwell Climate Research Center, Falmouth, MA, USA

Richard A. Houghton

4. Department of Physics and Applied Mathematics, Columbia University, New York, NY, USA

Robert D. Field

5. NASA Goddard Institute for Space Studies, New York, NY, USA

Robert D. Field

6. Laboratory of Geo-information Science and Remote Sensing, Wageningen University, Wageningen, The Netherlands

Jan Verbesselt

## Contributions

G.R.vdW. designed the research, M.J.E.vM., R.D.F. and R.A.H. constructed the LULCC time series, D.vW. performed the AF trend analysis with help from J.V., the paper was written by D.vW., M.J.E.vM. and G.R.vdW.

## Corresponding author

Correspondence to [Guido. R. van der Werf](#).

# Ethics declarations

## Competing interests

The authors declare no competing interests.

## Peer review

### Peer review information

*Nature* thanks Emilio Chuvieco, Chris Jones and Tom Oda for their contribution to the peer review of this work.

## Additional information

### Extended data

is available for this paper at <https://doi.org/10.1038/s41586-021-04376-4>.

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Extended data figures and tables

### [Extended Data Fig. 1 RCP8.5-projected and observed evolution of atmospheric CO<sub>2</sub> growth using our LULCC data.](#)

RCP8.5-projected (background) and observed (forefront) atmospheric CO<sub>2</sub> growth over 2010–2020 on the basis of observed concentrations, sources from fossil fuel burning, cement manufacturing and LULCC based on this study. Sink strength is computed as the residual. AF is short for airborne fraction and the numbers indicate what the difference is between observed values and RCP8.5 projections for each component.

## Extended Data Fig. 2 Evolution of RCP8.5-projected and observed anthropogenic emissions and atmospheric CO<sub>2</sub> growth rate over 2000–2019.

Fossil fuel emissions increased less than projected in RCP8.5 after 2012, but this was partly compensated for by higher-than-projected LULCC emissions in most years.

## Extended Data Fig. 3 RCP8.5-projected and observed evolution of atmospheric CO<sub>2</sub> growth on the basis of other LULCC datasets.

RCP8.5-projected (background) and observed (forefront) evolution of atmospheric CO<sub>2</sub> growth over 2010–2020 on the basis of observed concentrations, sources from fossil fuel burning, cement manufacturing and LULCC on the basis of the GCP (**a**) and H&N (**b**). Sink strength is computed as the residual. AF is short for airborne fraction and the numbers indicate what the difference is between observed values and RCP8.5 projections for each component.

## Extended Data Fig. 4 Schematic overview for production of LULCC emissions.

This overview shows our method to construct LULCC emissions on the basis of fire emissions in key deforestation zones of GFED4s (1997–2019) and visibility-based  $B_{ext}$  anchored to GFED4s for the preceding period. These were supplemented by non-fire emissions including those stemming from peat dynamics in EQAS.

## Rights and permissions

[Reprints and Permissions](#)

## About this article

## Cite this article

van Marle, M.J.E., van Wees, D., Houghton, R.A. *et al.* New land-use-change emissions indicate a declining CO<sub>2</sub> airborne fraction. *Nature* **603**, 450–454 (2022). <https://doi.org/10.1038/s41586-021-04376-4>

- Received: 26 April 2021
- Accepted: 20 December 2021
- Published: 16 March 2022
- Issue Date: 17 March 2022
- DOI: <https://doi.org/10.1038/s41586-021-04376-4>

## Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

## [Net carbon uptake has kept pace with increasing fossil-fuel emissions](#)

Research Briefing 16 Mar 2022

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04376-4>

- Article
- [Published: 09 March 2022](#)

# The evolution, evolvability and engineering of gene regulatory DNA

- [Eeshit Dhaval Vaishnav](#) ORCID: [orcid.org/0000-0003-3720-8051<sup>1,2 na1</sup>](https://orcid.org/0000-0003-3720-8051),
- [Carl G. de Boer](#) ORCID: [orcid.org/0000-0001-8935-5921<sup>3,4 na1</sup>](https://orcid.org/0000-0001-8935-5921),
- [Jennifer Molinet](#) ORCID: [orcid.org/0000-0001-9716-4455<sup>5,6</sup>](https://orcid.org/0000-0001-9716-4455),
- [Moran Yassour](#)<sup>4,7,8</sup>,
- [Lin Fan](#)<sup>2</sup>,
- [Xian Adiconis](#)<sup>4,9</sup>,
- [Dawn A. Thompson](#)<sup>2</sup>,
- [Joshua Z. Levin](#) ORCID: [orcid.org/0000-0002-0170-3598<sup>4,9</sup>](https://orcid.org/0000-0002-0170-3598),
- [Francisco A. Cubillos](#) ORCID: [orcid.org/0000-0003-3022-469X<sup>5,6</sup>](https://orcid.org/0000-0003-3022-469X) &
- [Aviv Regev](#) ORCID: [orcid.org/0000-0003-3293-3158<sup>4,10 nAff11</sup>](https://orcid.org/0000-0003-3293-3158)

*Nature* volume 603, pages 455–463 (2022)

- 21k Accesses
- 1 Citations
- 479 Altmetric
- [Metrics details](#)

## Subjects

- [Computational models](#)
- [Genomics](#)
- [Machine learning](#)
- [Molecular evolution](#)
- [Synthetic biology](#)

## Abstract

Mutations in non-coding regulatory DNA sequences can alter gene expression, organismal phenotype and fitness<sup>1,2,3</sup>. Constructing complete fitness landscapes, in which DNA sequences are mapped to fitness, is a long-standing goal in biology, but has remained elusive because it is challenging to generalize reliably to vast sequence spaces<sup>4,5,6</sup>. Here we build sequence-to-expression models that capture fitness landscapes and use them to decipher principles of regulatory evolution. Using millions of randomly sampled promoter DNA sequences and their measured expression levels in the yeast *Saccharomyces cerevisiae*, we learn deep neural network models that generalize with excellent prediction performance, and enable sequence design for expression engineering. Using our models, we study expression divergence under genetic drift and strong-selection weak-mutation regimes to find that regulatory evolution is rapid and subject to diminishing returns epistasis; that conflicting expression objectives in different environments constrain expression adaptation; and that stabilizing selection on gene expression leads to the moderation of regulatory complexity. We present an approach for using such models to detect signatures of selection on expression from natural variation in regulatory sequences and use it to discover an instance of convergent regulatory evolution. We assess mutational robustness, finding that regulatory mutation effect sizes follow a power law, characterize regulatory evolvability, visualize promoter fitness landscapes, discover evolvability archetypes and illustrate the mutational robustness of natural regulatory sequence populations. Our work provides a general framework for designing regulatory sequences and addressing fundamental questions in regulatory evolution.

This is a preview of subscription content

## **Access options**

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

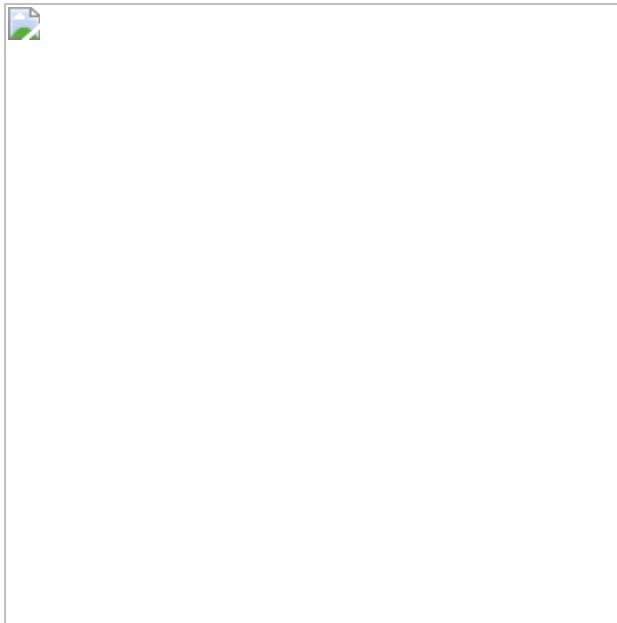
[Buy](#)

All prices are NET prices.

**Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

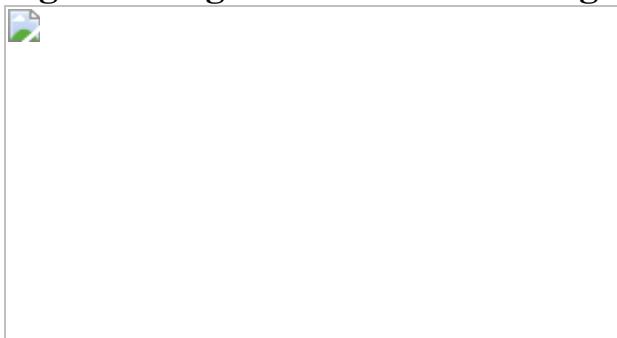
**Fig. 1: The evolution, evolvability and engineering of gene regulatory DNA.**



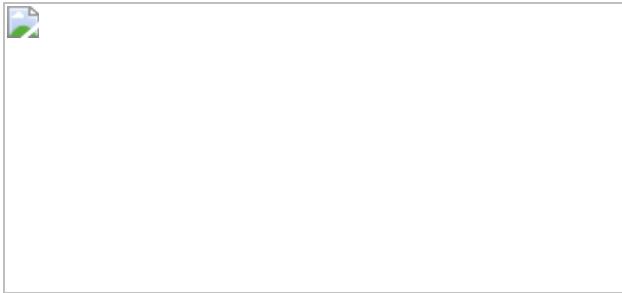
**Fig. 2: The evolutionary malleability of gene expression.**



**Fig. 3: The ECC detects signatures of selection on gene expression using natural genetic variation in regulatory DNA.**



**Fig. 4: The evolvability vector captures fitness landscapes.**



## Data availability

Data generated for this study are available at the NCBI GEO with accession numbers [GSE163045](#) and [GSE163866](#). All models and processed data are available on Zenodo at <https://zenodo.org/record/4436477>.

## Code availability

Code is available on GitHub at <https://github.com/ledv/evolution> and CodeOcean at <https://codeocean.com/capsule/8020974/tree>. A web app is available at <https://ledv.github.io/evolution/>.

## References

1. Wittkopp, P. J. & Kalay, G. *Cis*-regulatory elements: molecular mechanisms and evolutionary processes underlying divergence. *Nat. Rev. Genet.* **13**, 59–69 (2011).
2. Hill, M. S., Vande Zande, P. & Wittkopp, P. J. Molecular and evolutionary processes generating variation in gene expression. *Nat. Rev. Genet.* **22**, 203–215 (2021).
3. Fuqua, T. et al. Dense and pleiotropic regulatory information in a developmental enhancer. *Nature* **587**, 235–239 (2020).
4. de Visser, J. A. G. M. & Krug, J. Empirical fitness landscapes and the predictability of evolution. *Nat. Rev. Genet.* **15**, 480–490 (2014).

5. Kondrashov, D. A. & Kondrashov, F. A. Topological features of rugged fitness landscapes in sequence space. *Trends Genet.* **31**, 24–33 (2015).
6. de Visser, J. A. G. M., Elena, S. F., Fragata, I. & Matuszewski, S. The utility of fitness landscapes and big data for predicting evolution. *Heredity* **121**, 401–405 (2018).
7. Weirauch, M. T. & Hughes, T. R. Conserved expression without conserved regulatory sequence: the more things change, the more they stay the same. *Trends Genet.* **26**, 66–74 (2010).
8. Orr, H. A. The genetic theory of adaptation: a brief history. *Nat. Rev. Genet.* **6**, 119–127 (2005).
9. Weinreich, D. M., Lan, Y., Wylie, C. S. & Heckendorn, R. B. Should evolutionary geneticists worry about higher-order epistasis? *Curr. Opin. Genet. Dev.* **23**, 700–707 (2013).
10. Venkataram, S. et al. Development of a comprehensive genotype-to-fitness map of adaptation-driving mutations in yeast. *Cell* **166**, 1585–1596 (2016).
11. Keren, L. et al. Massively parallel interrogation of the effects of gene expression levels on fitness. *Cell* **166**, 1282–1294 (2016).
12. Sarkisyan, K. S. et al. Local fitness landscape of the green fluorescent protein. *Nature* **533**, 397–401 (2016).
13. Ogden, P. J., Kelsic, E. D., Sinai, S. & Church, G. M. Comprehensive AAV capsid fitness landscape reveals a viral gene and enables machine-guided design. *Science* **366**, 1139–1143 (2019).
14. Pitt, J. N. & Ferré-D’Amaré, A. R. Rapid construction of empirical RNA fitness landscapes. *Science* **330**, 376–379 (2010).
15. Shultzaberger, R. K., Malashock, D. S., Kirsch, J. F. & Eisen, M. B. The fitness landscapes of *cis*-acting binding sites in different promoter

and environmental contexts. *PLoS Genet.* **6**, e1001042 (2010).

16. Mustonen, V., Kinney, J., Callan, C. G. & Lässig, M. Energy-dependent fitness: a quantitative model for the evolution of yeast transcription factor binding sites. *Proc. Natl Acad. Sci. USA* **105**, 12376–12381 (2008).
17. Hartl, D. L. What can we learn from fitness landscapes? *Curr. Opin. Microbiol.* **0**, 51–57 (2014).
18. Otwinowski, J. & Nemenman, I. Genotype to phenotype mapping and the fitness landscape of the *E. coli* lac promoter. *PLoS ONE* **8**, e61570 (2013).
19. Sinai, S. & Kelsic, E. D. A primer on model-guided exploration of fitness landscapes for biological sequence design. Preprint at <https://arxiv.org/abs/2010.10614> (2020).
20. Zhou, J. & Troyanskaya, O. G. Predicting effects of noncoding variants with deep learning-based sequence model. *Nat. Methods* **12**, 931–934 (2015).
21. Avsec, Ž. et al. Base-resolution models of transcription-factor binding reveal soft motif syntax. *Nat. Genet.* **53**, 354–366 (2021).
22. Shrikumar, A., Greenside, P. & Kundaje, A. Learning important features through propagating activation differences. *Proc. 34th International Conference on Machine Learning* 3145–3153 (2017).
23. Avsec, Ž. et al. Effective gene expression prediction from sequence by integrating long-range interactions. *Nat. Methods* **18**, 1196–1203 (2021).
24. Fragata, I., Blanckaert, A., Louro, M. A. D., Liberles, D. A. & Bank, C. Evolution in the light of fitness landscape theory. *Trends Ecol. Evol.* **34**, 69–82 (2019).

25. Payne, J. L. & Wagner, A. The causes of evolvability and their evolution. *Nat. Rev. Genet.* **20**, 24–38 (2019).
26. de Boer, C. G. et al. Deciphering eukaryotic gene-regulatory logic with 100 million random promoters. *Nat. Biotechnol.* **38**, 56–65 (2020).
27. Crocker, J. et al. Low affinity binding site clusters confer hox specificity and regulatory robustness. *Cell* **160**, 191–203 (2015).
28. Habib, N., Wapinski, I., Margalit, H., Regev, A. & Friedman, N. A functional selection model explains evolutionary robustness despite plasticity in regulatory networks. *Mol. Syst. Biol.* **8**, 619 (2012).
29. Gillespie, J. H. Molecular evolution over the mutational landscape. *Evolution* **38**, 1116–1129 (1984).
30. Jerison, E. R. & Desai, M. M. Genomic investigations of evolutionary dynamics and epistasis in microbial evolution experiments. *Curr. Opin. Genet. Dev.* **35**, 33–39 (2015).
31. Sæther, B.-E. & Engen, S. The concept of fitness in fluctuating environments. *Trends Ecol. Evol.* **30**, 273–281 (2015).
32. Vaswani, A. et al. in *Advances in Neural Information Processing Systems 30* (eds. Guyon, I. et al.) 5998–6008 (Curran Associates, 2017).
33. Weirauch, M. T. et al. Evaluation of methods for modeling transcription factor sequence specificity. *Nat. Biotechnol.* **31**, 126–134 (2013).
34. Yang, N. & Bielawski, N. Statistical methods for detecting molecular adaptation. *Trends Ecol. Evol.* **15**, 496–503 (2000).
35. Moses, A. M. Statistical tests for natural selection on regulatory regions based on the strength of transcription factor binding sites. *BMC Evol. Biol.* **9**, 286 (2009).

36. Rifkin, S. A., Houle, D., Kim, J. & White, K. P. A mutation accumulation assay reveals a broad capacity for rapid evolution of gene expression. *Nature* **438**, 220–223 (2005).
37. Peter, J. et al. Genome evolution across 1,011 *Saccharomyces cerevisiae* isolates. *Nature* **556**, 339–344 (2018).
38. Erb, I. & van Nimwegen, E. Transcription factor binding site positioning in yeast: proximal promoter motifs characterize TATA-less promoters. *PLoS One* **6**, e24279 (2011).
39. Gilad, Y., Oshlack, A. & Rifkin, S. A. Natural selection on gene expression. *Trends Genet.* **22**, 456–461 (2006).
40. Alhusaini, N. & Coller, J. The deadenylase components Not2p, Not3p, and Not5p promote mRNA decapping. *RNA* **22**, 709–721 (2016).
41. Yang, J.-R., Maclean, C. J., Park, C., Zhao, H. & Zhang, J. Intra and interspecific variations of gene expression levels in yeast are largely neutral: (Nei Lecture, SMBE 2016, Gold Coast). *Mol. Biol. Evol.* **34**, 2125–2139 (2017).
42. Chen, J. et al. A quantitative framework for characterizing the evolutionary history of mammalian gene expression. *Genome Res.* **29**, 53–63 (2019).
43. Payne, J. L. & Wagner, A. Mechanisms of mutational robustness in transcriptional regulation. *Front. Genet.* **6**, 322 (2015).
44. Shoval, O. et al. Evolutionary trade-offs, Pareto optimality, and the geometry of phenotype space. *Science* **336**, 1157–1160 (2012).
45. van Dijk, D. et al. Finding archetypal spaces using neural networks. *IEEE International Conference on Big Data* 2634-2643 (2019).
46. He, X., Duque, T. S. P. C. & Sinha, S. Evolutionary origins of transcription factor binding site clusters. *Mol. Biol. Evol.* **29**, 1059–1070 (2012).

47. Cliften, P. F. et al. Surveying *Saccharomyces* genomes to identify functional elements by comparative DNA sequence analysis. *Genome Res.* **11**, 1175–1186 (2001).
48. Heinz, S., Romanoski, C. E., Benner, C. & Glass, C. K. The selection and function of cell type-specific enhancers. *Nat. Rev. Mol. Cell Biol.* **16**, 144–154 (2015).
49. Lehner, B. Selection to minimise noise in living systems and its implications for the evolution of gene expression. *Mol. Syst. Biol.* **4**, 170 (2008).
50. Metzger, B. P. H., Yuan, D. C., Gruber, J. D., Duveau, F. & Wittkopp, P. J. Selection on noise constrains variation in a eukaryotic promoter. *Nature* **521**, 344–347 (2015).
51. Kosuri, S. et al. Composability of regulatory sequences controlling transcription and translation in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*. **110**, 14024–14029 (2013).
52. Shalem, O. et al. Systematic dissection of the sequence determinants of gene 3' end mediated expression control. *PLoS Genet.* **11**, e1005147 (2015).
53. Kinney, J. B., Murugan, A., Callan, C. G. Jr & Cox, E. C. Using deep sequencing to characterize the biophysical mechanism of a transcriptional regulatory sequence. *Proc. Natl Acad. Sci. USA*. **107**, 9158–9163 (2010).
54. Sharon, E. et al. Inferring gene regulatory logic from high-throughput measurements of thousands of systematically designed promoters. *Nat. Biotechnol.* **30**, 521–530 (2012).
55. Melnikov, A. et al. Systematic dissection and optimization of inducible enhancers in human cells using a massively parallel reporter assay. *Nat. Biotechnol.* **30**, 271–277 (2012).

56. Kwasnieski, J. C., Mogno, I., Myers, C. A., Corbo, J. C. & Cohen, B. A. Complex effects of nucleotide variants in a mammalian *cis*-regulatory element. *Proc. Natl Acad. Sci. USA* **109**, 19498–19503 (2012).
57. Kircher, M. et al. Saturation mutagenesis of twenty disease-associated regulatory elements at single base-pair resolution. *Nat. Commun.* **10**, 3583 (2019).
58. Townsley, K. G., Brennand, K. J. & Huckins, L. M. Massively parallel techniques for cataloguing the regulome of the human brain. *Nat. Neurosci.* **23**, 1509–1521 (2020).
59. Renganaath, K. et al. Systematic identification of *cis*-regulatory variants that cause gene expression differences in a yeast cross. *eLife* **9**, e62669 (2020).
60. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
61. Alipanahi, B., Delong, A., Weirauch, M. T. & Frey, B. J. Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning. *Nat. Biotechnol.* **33**, 831–838 (2015).
62. Travers, C. et al. Opportunities and obstacles for deep learning in biology and medicine. *J. R. Soc. Interface* **15**, 20170387 (2018).
63. Avsec, Ž. et al. The Kipoi repository accelerates community exchange and reuse of predictive models for genomics. *Nat. Biotechnol.* **37**, 592–600 (2019).
64. Quang, D. & Xie, X. FactorNet: a deep learning framework for predicting cell type specific transcription factor binding from nucleotide-resolution sequential data. *Methods* **166**, 40–47 (2019).
65. Zhou H. et al. Towards a better understanding of reverse-complement equivariance for deep learning models in genomics. *Proc. 16th*

*Machine Learning in Computational Biology meeting* **165**, 1–33 (2022).

66. Morrow, A. et al. Convolutional kitchen sinks for transcription factor binding site prediction. Preprint at <https://arxiv.org/abs/1706.00125> (2017).
67. Kelley, D. R., Snoek, J. & Rinn, J. L. Bassett: learning the regulatory code of the accessible genome with deep convolutional neural networks. *Genome Res.* **26**, 990–999 (2016).
68. Koo, P. K., Majdandzic, A., Ploenzke, M., Anand, P. & Paul, S. B. Global importance analysis: an interpretability method to quantify importance of genomic features in deep neural networks. *PLoS Comput. Biol.* **17**, e1008925 (2021).
69. Quang, D. & Xie, X. DanQ: a hybrid convolutional and recurrent deep neural network for quantifying the function of DNA sequences. *Nucleic Acids Res.* **44**, e107 (2016).
70. Kingma, D. P. & Ba, J. Adam: a method for stochastic optimization. *International Conference on Learning Representations* (Poster) (2015).
71. Abadi, M. et al. TensorFlow: large-scale machine learning on heterogenous systems. Software available from <https://www.tensorflow.org/> (2015).
72. Jouppi, N. P. et al. In-datacenter performance analysis of a tensor processing unit. In *Proc. 44th Annual International Symposium on Computer Architecture* 1–12 (2017).
73. Li, J., Pu, Y., Tang, J., Zou, Q. & Guo, F. DeepATT: a hybrid category attention neural network for identifying functional effects of DNA sequences. *Brief. Bioinform.* **22**, bbaa159 (2020).
74. Ullah, F. & Ben-Hur, A. A self-attention model for inferring cooperativity between regulatory features. *Nucleic Acids Res.* **49**, e77 (2021).

75. Clauwaert, J., Menschaert, G. & Waegeman, W. Explainability in transformer models for functional genomics. *Brief. Bioinform.* **22**, bbab060 (2021).
76. Hinton, G. & Tieleman, T. Lecture 6.5—RmsProp: divide the gradient by a running average of its recent magnitude. *COURSERA: Neural Networks for Machine Learning* **4**, 26–31 (2012).
77. Sinai, S. et al. AdaLead: a simple and robust adaptive greedy search algorithm for sequence design. Preprint at <https://arxiv.org/abs/2010.02141> (2020).
78. Linder, J., Bogard, N., Rosenberg, A. B. & Seelig, G. A generative neural network for maximizing fitness and diversity of synthetic DNA and protein sequences. *Cell Syst.* **11**, 49–62 (2020).
79. Brookes, D., Park, H. & Listgarten, J. Conditioning by adaptive sampling for robust design. *Proc. Mach. Learn. Res.* **97**, 773–782 (2019).
80. Killoran, N., Lee, L. J., Delong, A., Duvenaud, D. & Frey, B. J. Generating and designing DNA with deep generative models. *Neurips Computational Biology Workshop* (2017).
81. Fortin, F.-A., Rainville, F.-M. D., Gardner, M.-A., Parizeau, M. & Gagné, C. DEAP: evolutionary algorithms made easy. *J. Mach. Learn. Res.* **13**, 2171–2175 (2012).
82. Jaeger, S. A. et al. Conservation and regulatory associations of a wide affinity range of mouse transcription factor binding sites. *Genomics* **95**, 185–195 (2010).
83. Tanay, A. Extensive low-affinity transcriptional interactions in the yeast genome. *Genome Res.* **16**, 962–972 (2006).
84. Sniegowski, P. D. & Gerrish, P. J. Beneficial mutations and the dynamics of adaptation in asexual populations. *Phil. Trans. R. Soc. B* **365**, 1255–1263 (2010).

85. Szendro, I. G., Franke, J., de Visser, J. A. & Krug, J. Predictability of evolution depends nonmonotonically on population size. *Proc. Natl Acad. Sci. USA* **110**, 571–576 (2013).
86. Orr, H. A. The population genetics of adaptation: the adaptation of DNA Sequences. *Evolution* **56**, 1317–1330 (2002).
87. Bailey, T. L. DREME: motif discovery in transcription factor ChIP-seq data. *Bioinformatics* **27**, 1653–1659 (2011).
88. de Boer, C. G. & Hughes, T. R. YeTFaSCo: a database of evaluated yeast transcription factor sequence specificities. *Nucleic Acids Res.* **40**, D169–D179 (2012).
89. Kent, W. J. BLAT—the BLAST-Like Alignment Tool. *Genome Res.* **12**, 656–664 (2002).
90. Cherry, J. M. et al. Saccharomyces Genome Database: the genomics resource of budding yeast. *Nucleic Acids Res.* **40**, D700–D705 (2012).
91. Smith, J. D., McManus, K. F. & Fraser, H. B. A novel test for selection on *cis*-regulatory elements reveals positive and negative selection acting on mammalian transcriptional enhancers. *Mol. Biol. Evol.* **30**, 2509–2518 (2013).
92. Liu, J. & Robinson-Rechavi, M. Robust inference of positive selection on regulatory sequences in the human brain. *Sci. Adv.* **6**, eabc9863 (2020).
93. Rice, D. P. & Townsend, J. P. A test for selection employing quantitative trait locus and mutation accumulation data. *Genetics* **190**, 1533–1545 (2012).
94. Denver, D. R., Morris, K., Lynch, M. & Thomas, W. K. High mutation rate and predominance of insertions in the *Caenorhabditis elegans* nuclear genome. *Nature* **430**, 679–682 (2004).

95. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).
96. Thompson, D. A. et al. Evolutionary principles of modular gene regulation in yeasts. *eLife* **2**, e00603 (2013).
97. Yassour, M. et al. Strand-specific RNA sequencing reveals extensive regulated long antisense transcripts that are conserved across yeast species. *Genome Biol.* **11**, R87 (2010).
98. Grabherr, M. G. et al. Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).
99. Wu, T. D. & Watanabe, C. K. GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics* **21**, 1859–1875 (2005).
100. Wapinski, I., Pfeffer, A., Friedman, N. & Regev, A. Natural history and evolutionary principles of gene duplication in fungi. *Nature* **449**, 54–61 (2007).
101. Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
102. Yates, A. D. et al. Ensembl 2020. *Nucleic Acids Res.* **48**, D682–D688 (2020).
103. DiCarlo, J. E. et al. Genome engineering in *Saccharomyces cerevisiae* using CRISPR–Cas systems. *Nucleic Acids Res.* **41**, 4336–4343 (2013).
104. Fleiss, A. et al. Reshuffling yeast chromosomes with CRISPR/Cas9. *PLoS Genet.* **15**, e1008332 (2019).
105. Horwitz, A. A. et al. Efficient multiplexed integration of synergistic alleles and metabolic pathways in yeasts via CRISPR–Cas. *Cell Syst.*

**1**, 88–96 (2015).

106. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the  $2^{-\Delta\Delta CT}$  method. *Methods* **25**, 402–408 (2001).
107. Vandesompele, J. et al. Accurate normalization of real-time quantitative RT–PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **3**, research0034.1 (2002).
108. Teste, M.-A., Duquenne, M., François, J. M. & Parrou, J.-L. Validation of reference genes for quantitative expression analysis by real-time RT–PCR in *Saccharomyces cerevisiae*. *BMC Mol. Biol.* **10**, 99 (2009).
109. Mardones, W. et al. Rapid selection response to ethanol in *Saccharomyces eubayanus* emulates the domestication process under brewing conditions. *Microb. Biotechnol.* <https://doi.org/10.1111/1751-7915.13803> (2021).
110. Ibstedt, S. et al. Concerted evolution of life stage performances signals recent selection on yeast nitrogen use. *Mol. Biol. Evol.* **32**, 153–161 (2015).
111. Rich, M. S. et al. Comprehensive analysis of the SUL1 promoter of *Saccharomyces cerevisiae*. *Genetics* **203**, 191–202 (2016).
112. Rest, J. S. et al. Nonlinear fitness consequences of variation in expression level of a eukaryotic gene. *Mol. Biol. Evol.* **30**, 448–456 (2013).
113. Bergen, A. C., Olsen, G. M. & Fay, J. C. Divergent MLS1 promoters lie on a fitness plateau for gene expression. *Mol. Biol. Evol.* **33**, 1270–1279 (2016).
114. Alstott, J., Bullmore, E. & Plenz, D. Powerlaw: a Python package for analysis of heavy-tailed distributions. *PLoS One* **9**, e85777 (2014).

## Acknowledgements

We thank Google TPU Research Cloud for TPU access, L. Gaffney for help with figure preparation, Broad Genomics Platform for sequencing work, J.-C. Hütter for advice on fitness responsivity, J. Pfiffner-Borges for help with RNA-seq, R. Yu, B. Lee and N. Jaberi for manuscript feedback and members of the A.R. laboratory for discussions. E.D.V. was supported by the MIT Presidential Fellowship; C.G.d.B. was supported by a Canadian Institutes for Health Research Fellowship and the NIH (K99-HG009920-01); and F.A.C. and J.M. were supported by ANID (Programa Iniciativa Científica Milenio, ICN17\_022). Work was supported by the Klarman Cell Observatory, Howard Hughes Medical Institute (HHMI) and Google TPU Research Cloud (<https://sites.research.google/trc/about/>). A.R. was an Investigator of the HHMI.

## Author information

### Author notes

1. Aviv Regev

Present address: Genentech, South San Francisco, CA, USA

2. These authors contributed equally: Eeshit Dhaval Vaishnav, Carl G. de Boer

## Affiliations

1. Massachusetts Institute of Technology, Cambridge, MA, USA

Eeshit Dhaval Vaishnav

2. Broad Institute of MIT and Harvard, Cambridge, MA, USA

Eeshit Dhaval Vaishnav, Lin Fan & Dawn A. Thompson

3. School of Biomedical Engineering, University of British Columbia, Vancouver, British Columbia, Canada

Carl G. de Boer

4. Klarman Cell Observatory, Broad Institute of MIT and Harvard, Cambridge, MA, USA

Carl G. de Boer, Moran Yassour, Xian Adiconis, Joshua Z. Levin & Aviv Regev

5. Departamento de Biología, Facultad de Química y Biología, Universidad de Santiago de Chile, Santiago, Chile

Jennifer Molinet & Francisco A. Cubillos

6. ANID—Millennium Science Initiative Program, Millennium Institute for Integrative Biology (iBio), Santiago, Chile

Jennifer Molinet & Francisco A. Cubillos

7. Faculty of Medicine, The Hebrew University of Jerusalem, Jerusalem, Israel

Moran Yassour

8. The Rachel and Selim Benin School of Computer Science and Engineering, The Hebrew University of Jerusalem, Jerusalem, Israel

Moran Yassour

9. Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA, USA

Xian Adiconis & Joshua Z. Levin

10. Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA

Aviv Regev

## Contributions

E.D.V., C.G.d.B. and A.R. conceived, designed and supervised the study. E.D.V. and C.G.d.B. performed the analyses. M.Y., L.F., X.A. and D.A.T. performed and D.A.T., J.Z.L. and A.R. supervised the Ascomycota cross-species RNA-seq experiments. J.M. performed and F.A.C. supervised the *CDC36* experiments. E.D.V. and C.G.d.B. performed the rest of the experiments. E.D.V., C.G.d.B. and A.R. wrote the manuscript.

## Corresponding authors

Correspondence to [Eeshit Dhaval Vaishnav](#), [Carl G. de Boer](#) or [Aviv Regev](#).

## Ethics declarations

### Competing interests

A.R. is a co-founder and equity holder of Celsius Therapeutics and Immunitas and until 31 July 2020 was a member of the scientific advisory board of Thermo Fisher Scientific, Syros Pharmaceuticals, Neogene Therapeutics and Asimov. As of 1 August 2020, A.R. is an employee of Genentech. The other authors declare no competing interests.

## Peer review

### Peer review information

*Nature* thanks Martin Taylor and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

## Additional information

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Extended data figures and tables

## Extended Data Fig. 1 The convolutional sequence-to-expression model generalizes reliably and characterizes sequence trajectories under different evolutionary regimes.

**a–d**, Prediction of expression from sequence in complex (YPD) (**a, b**) and defined (SD-Uracil) (**c, d**) medium. Predicted (*x* axis) and experimentally measured (*y* axis) expression for (**a, c**) random test sequences (sampled separately from and not overlapping with the training data) and (**b, d**) native yeast promoter sequences containing random single base mutations. Top left: Pearson's *r* and associated two-tailed *P* value. Compression of predictions in the lower left results from binning differences during cell sorting in different experiments ([Supplementary Notes](#)). **e**, Experimental validation of trajectories from simulations of random genetic drift. Distribution of measured (light grey) and predicted (dark grey) changes in expression in the defined medium (SD-Uracil) (*y* axis) for the synthesized randomly designed sequences (*n* = 2,986) at each mutational step (*x* axis). Midline: median; boxes: interquartile range; whiskers: 5<sup>th</sup> and 95<sup>th</sup> percentile range. **f, g**, Simulation and validation of expression trajectories under SSWM in defined medium (SD-Uracil). **f**, Distribution of predicted expression levels (*y* axis) in defined medium at each evolutionary time step (*x* axis) for sequences under SSWM favouring high (red) or low (blue) expression, starting with native promoter sequences (*n* = 5,720). Midline: median; boxes: interquartile range; whiskers: 5<sup>th</sup> and 95<sup>th</sup> percentile range. **g**, Experimentally measured expression distribution in defined medium (*y* axis) for the synthesized sequences (*n* = 6,304 sequences; 637 trajectories) at each mutational step (*x* axis) from predicted mutational trajectories under SSWM, favouring high (red) or low (blue) expression. Midline: median; boxes: interquartile range; whiskers: 5<sup>th</sup> and 95<sup>th</sup> percentile range. **h–o**, Experimental validation of predicted expression for sequences from the random genetic drift and SSWM simulations. Experimentally measured (*y* axis) and predicted (*x* axis) expression level (**l–o**) or expression change from the starting sequence (**h–k**) in complex (**h, j, l, n**) or defined (**i, k, m, o**) medium using sequences from the random genetic drift (Fig. [2e](#), Extended Data Fig [1e](#), **h, i, l, m** here) and SSWM (Fig. [2g](#), Extended Data Fig [1g](#), **j, k, n, o** here) validation experiments. Top left: Pearson's *r* and associated two-tailed *P* values.

## Extended Data Fig. 2 Characterization of sequence trajectories under strong competing selection pressures using the convolutional model.

**a, b**, Expression is highly correlated between defined and complex medium. Measured (**a**) and predicted (**b**) expression in defined (*x* axis) and complex (*y* axis) medium for a set of test sequences measured in both media. Top left: Pearson's *r* and associated two-tailed *P* values. **c**, Opposing relationships between organismal fitness and *URA3* expression in two environments. Measured expression (*x* axis, using a YFP reporter) and fitness (*y* axis; when used as the promoter sequence for the *URA3* gene) for yeast with each of 11 promoters predicted to span a wide range of expression levels in complex medium with 5-FOA (red), where higher expression of *URA3* is toxic owing to *URA3*-mediated conversion of 5-FOA to 5-fluorouracil, and in defined medium lacking uracil (blue), where *URA3* is required for uracil synthesis. Error bars: Standard error of the mean ( $n = 3$  replicate experiments). **d–f**, Competing expression objectives constrain adaptation. **d, e**, Difference in predicted expression (*y* axis) at each evolutionary time step (*x* axis) under selection to maximize (red) or minimize (blue) the difference between expression in defined and complex medium, starting with either native sequences (**d**, as Fig. 2*h*,  $n = 5,720$ ) or random sequences (**e**,  $n = 10,000$ ). **f**, Distribution of predicted expression (*y* axis) in complex (blue) and defined (red) medium at each evolutionary time step (*x* axis) for a starting set of random sequences ( $n = 10,000$ ). Midline: median; boxes: interquartile range; whiskers: 5<sup>th</sup> and 95<sup>th</sup> percentile range. **g**, Motifs enriched within sequences evolved for competing objectives in different environments. Top five most enriched motifs, found using DREME<sup>87</sup> (Methods) within sequences computationally evolved from a starting set of random sequences to either maximize (left) or minimize (right) the difference in expression between defined and complex medium, along with DREME E-values, the corresponding rank of the same motif when using native sequences as a starting point, the probable cognate transcription factor and that transcription factor's known motif.

## Extended Data Fig. 3 The transformer sequence-to-expression model generalizes reliably and characterizes sequence

## trajectories under different evolutionary regimes.

**a–d**, Prediction of expression from sequence in the complex (**a, b**) and defined (**c, d**) medium. Predicted (*x* axis) and experimentally measured (*y* axis) expression for (**a, c**) random test sequences (sampled separately from and not overlapping with the training data) and (**b, d**) native yeast promoter sequences containing random single base mutations. Top left: Pearson's *r* and associated two-tailed *P* value. Compression of predictions in the lower left results from binning differences during cell sorting in different experiments ([Supplementary Notes](#)). **e**, Predicted (*x* axis) and experimentally measured (*y* axis) expression in complex medium (YPD) for all native yeast promoter sequences. Pearson's *r* and associated two-tailed *P* values are shown. **f**, Predicted expression divergence under random genetic drift. Distribution of the change in predicted expression (*y* axis) for random starting sequences ( $n = 5,720$ ) at each mutational step (*x* axis) for trajectories simulated under random genetic drift. Silver bar: differences in expression between unrelated sequences. **g, h**, Comparison of the distribution of measured (light grey) and transformer model predicted (dark grey) changes in expression (*y* axis) in complex medium (**g**,  $n = 2,983$ ) and defined medium (**h**,  $n = 2,986$ ) for synthesized randomly designed sequences at each mutational step (*x* axis). **i, j**, Predicted expression evolution under SSWM. Distribution of predicted expression levels (*y* axis) in complex medium (**i**,  $n = 10,322$ ) and defined medium (**j**,  $n = 6,304$ ) at each mutational step (*x* axis) for sequence trajectories under SSWM favouring high (red) or low (blue) expression, starting with 5,720 native promoter sequences. (**f–j**) Midline: median; boxes: interquartile range; whiskers: 5<sup>th</sup> and 95<sup>th</sup> percentile range. **k–r**, Comparison of model predicted expression for sequences synthesized previously for the random genetic drift and SSWM analyses. Experimentally measured (*y* axis) and transformer model predicted (*x* axis) expression level (**o–r**) or expression change from the starting sequence (**k–n**) in complex (**k, m, o, q**) or defined (**l, n, p, r**) medium using sequences from the random genetic drift (Fig. [2c](#), Extended Data Fig. [1e](#); **k, l, o, p** here) and SSWM (Fig. [2g](#), Extended Data Fig. [1g](#); **m, n, q, r** here) validation experiments. Top left: Pearson's *r* and associated two-tailed *P* values.

## Extended Data Fig. 4 Signatures of stabilizing selection on gene expression detected from regulatory DNA across natural populations.

**a**, Expression-altering alleles in the CDC36 promoter are attributed primarily to altered UPC2 binding. Transcription factor interaction strength<sup>26</sup> (expression attributable to each transcription factor) difference between the high and low alleles (each point is a transcription factor) for each of two low expression alleles (allele 1:  $x$  axis; allele 2:  $y$  axis). Each low-expressing allele is compared to the high-expression allele with the most similar sequence (across all promoter sequences analysed from the 1,011 strains;  $\{\{e\}_{\{\{\backslash rm\{TF\}\},\{A\}_{\{high\}}-\{e\}_{\{\{\backslash rm\{TF\}\},\{A\}_{\{low\}}\}}\}}$ ). **b**, Distribution of ECC (y axis, calculated from 1,011 *S. cerevisiae* genomes, top left) for *S. cerevisiae* genes whose orthologues have divergent (blue) or conserved (purple) expression (within *Saccharomyces* (left,  $n = 4,191$ ), Ascomycota (middle,  $n = 4,910$ ), or mammals (right,  $n = 199$ ) (as determined by cross species RNA-seq, top right).  $P$  values: two-sided Wilcoxon rank-sum test. Midline: median; boxes: interquartile range; whiskers: 5<sup>th</sup> and 95<sup>th</sup> percentile range. **c**, Determination of expression change threshold for defining a ‘tolerated mutation’ to compute mutational robustness. We used all genes with an ECC consistent with stabilizing selection ( $ECC > 0$ ; left), calculated the variance in predicted expression across the 1011 yeast strains for each gene, and chose the tolerable mutation threshold,  $\{\{\backslash epsilon\}\}$ , as two standard deviations of the distribution of the variance (right). ~73% of genes with  $ECC > 0$  had an expression variation lower than  $\{\{\backslash epsilon\}\}$ . **d**, Distribution of the effects (magnitude;  $y$  axis) of mutations (rank ordered;  $x$  axis) on expression for all native regulatory sequences follows a power law with an exponent of 2.252. Shaded regions are equal in area.

## Extended Data Fig. 5 Fitness responsivity of a gene as the total variation of its expression-to-fitness relationship $F_{GENE}$ curves.

Expression ( $x$  axis) and fitness ( $y$  axis) level curves for each select gene, fit from experimental measurements of expression and fitness across promoter

variants by Keren et al<sup>11</sup>. Fitness responsivity calculated as the total variation in each curve is noted above each panel.

## Extended Data Fig. 6 Analysis of regulatory evolvability reveals sequence-encoded signatures of expression conservation from solitary sequences.

**a**, Selection of optimal number of archetypes. Mean-square-reconstruction error (y axis) for reconstructing the evolvability vectors from the embeddings learned by the autoencoder for an increasing number of archetypes (x axis). Red circle: optimal number of archetypes selected as prescribed<sup>45</sup> by the ‘elbow method’. **b**, The archetypal embeddings learned by the autoencoder accurately capture evolvability vectors. Original (y axis) and reconstructed (x axis) expression changes (the values in the evolvability vectors) for each native sequence (none seen by the autoencoder in training). Top left: Pearson’s  $r$  and associated two-tailed  $P$  values. **c–f**, Evolvability space captures regulatory sequences’ evolutionary properties. Proximity to the malleable archetype ( $A_{\text{malleable}}$ ) (x axis) and mutational robustness (**c, e** y axis) or ECC (**d, f** y axis) for all yeast genes (**e, f**) or the gene for which fitness responsivity was quantified (**c, d**). Top right: Spearman’s  $\rho$  and associated two-sided  $P$  value. ‘L’-shape of relationship in **e** results from the robust cleft,  $A_{\text{maxima}}$ , and  $A_{\text{minima}}$  all being distal to  $A_{\text{malleable}}$  (left side of plot). **g**, All native (S288C reference) promoter sequences (points) projected onto the archetypal evolvability space learned from random sequences; coloured by their ECC. Large coloured circles: evolvability archetypes. **h**, The proximity to the malleable archetype (x axis) and fitness responsivity (y axis) for the 80 genes with measured fitness responsivity. Top right: Spearman’s  $\rho$  and associated two-tailed  $P$  values. Light blue error band: 95% confidence interval. **i**, All native (S288C reference) promoter sequences (points) projected on the evolvability space learned from random sequences; coloured by their mean pairwise distance in the archetypal evolvability space between all promoter alleles across the 1,011 yeast isolates for that gene (orthologue evolvability dispersion). Large coloured circles: evolvability archetypes.

## Extended Data Fig. 7 Visualizing promoter fitness landscapes in sequence space.

Visualizing the fitness landscapes for the promoters of *HXT3* (a), *ADH1* (b), *GCN4* (c), *RPL3* (d), *FBA1* (e), *TUB3* (f), *URA3* (in defined medium) (g), *URA3* (in complex medium + 5FOA) (h). 1,000 promoter sequences represented by their evolvability vectors projected onto the 2D archetypal evolvability space and coloured by their associated fitness as reflected by their predicted growth rate relative to wild type (colour, Methods), estimated by first mapping sequences to expression with our model and then expression to fitness as measured and estimated previously<sup>11</sup>.

## Extended Data Fig. 8 In silico mutagenesis of malleable and robust promoters.

SSWM trajectories for (a) *DBP7*, a malleable promoter, and (b) *UTH1*, a robust promoter. Each subplot shows the *in silico* mutagenesis effects for how expression level (colour) changes when mutating each position (*x* axis) to each of the four bases (*y* axis) of each sequence (subplots) in the trajectories. The DNA sequence is indicated above each wild-type subplot (indicated with ‘WT’ at left). Arrows indicate the mutations selected at each step, which always correspond to the mutation of maximal effect; increasing expression goes up the figure from wild type and decreasing expression goes down. Part of the malleability of the *DBP7* promoter results from an intermediate-affinity Rap1p-binding site (grey bar). The first mutations in increasing- and decreasing-expression trajectories either increase or decrease (respectively) the affinity of this site. The *UTH1* promoter changes gradually in expression and evolves proximal repressor binding sites to dampen expression (grey bars).

## Supplementary information

### Supplementary Information

This file contains Supplementary Notes, Supplementary Figures 1–21, legends for Supplementary Tables 1 and 2, Supplementary Tables 3 and 4,

and additional references.

## **Reporting Summary**

## **Supplementary Tables**

This file contains Supplementary Tables 1 and 2; see main Supplementary Information PDF for legends.

## **Rights and permissions**

### **Reprints and Permissions**

## **About this article**

### **Cite this article**

Vaishnav, E.D., de Boer, C.G., Molinet, J. *et al.* The evolution, evolvability and engineering of gene regulatory DNA. *Nature* **603**, 455–463 (2022).  
<https://doi.org/10.1038/s41586-022-04506-6>

- Received: 08 February 2021
- Accepted: 02 February 2022
- Published: 09 March 2022
- Issue Date: 17 March 2022
- DOI: <https://doi.org/10.1038/s41586-022-04506-6>

## **Share this article**

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

## Further reading

- [AI predicts the effectiveness and evolution of gene promoter sequences](#)

- Andreas Wagner

*Nature* (2022)

## [AI predicts the effectiveness and evolution of gene promoter sequences](#)

- Andreas Wagner

News & Views 09 Mar 2022

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-022-04506-6>

- Article
- [Published: 09 March 2022](#)

# Cortical processing of flexible and context-dependent sensorimotor sequences

- [Duo Xu<sup>1</sup>](#),
- [Mingyuan Dong](#) ORCID: [orcid.org/0000-0001-9120-1315<sup>1</sup>](https://orcid.org/0000-0001-9120-1315),
- [Yuxi Chen](#) ORCID: [orcid.org/0000-0002-7826-4212<sup>2</sup>](https://orcid.org/0000-0002-7826-4212),
- [Angel M. Delgado<sup>2</sup>](#),
- [Natasha C. Hughes<sup>2</sup>](#),
- [Linghua Zhang<sup>1</sup>](#) &
- [Daniel H. O'Connor](#) ORCID: [orcid.org/0000-0002-9193-6714<sup>1</sup>](https://orcid.org/0000-0002-9193-6714)

[Nature](#) volume 603, pages 464–469 (2022)

- 4238 Accesses
- 48 Altmetric
- [Metrics details](#)

## Subjects

- [Motor cortex](#)
- [Sensory processing](#)

## Abstract

The brain generates complex sequences of movements that can be flexibly configured based on behavioural context or real-time sensory feedback<sup>1</sup>, but how this occurs is not fully understood. Here we developed a ‘sequence licking’ task in which mice directed their tongue to a target that moved through a series of locations. Mice could rapidly branch the sequence online based on tactile feedback. Closed-loop optogenetics and electrophysiology revealed that the tongue and jaw regions of the primary somatosensory (S1TJ) and motor (M1TJ) cortices<sup>2</sup> encoded and controlled tongue kinematics at the level of individual licks. By contrast, the tongue ‘premotor’ (anterolateral motor) cortex<sup>3,4,5,6,7,8,9,10</sup> encoded latent variables including intended lick angle, sequence identity and progress towards the reward that marked successful sequence execution. Movement-nonspecific sequence branching signals occurred in the anterolateral motor cortex and M1TJ. Our results reveal a set of key cortical areas for flexible and context-informed sequence generation.

This is a preview of subscription content

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

## Subscribe

All prices are NET prices.  
VAT will be added later in the checkout.  
Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

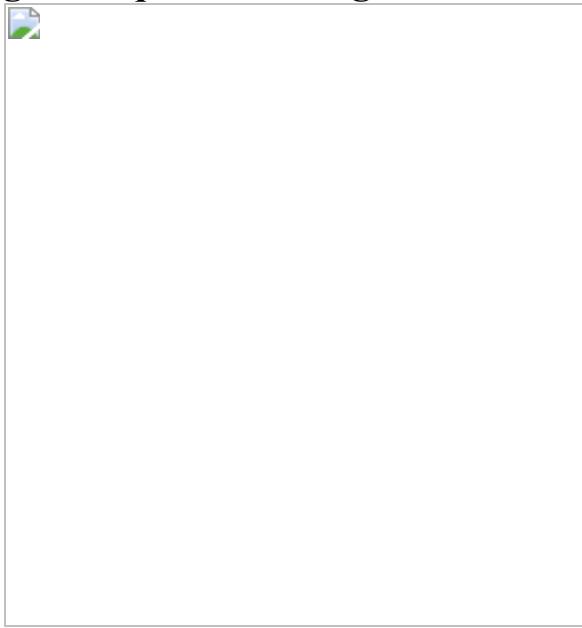
Buy

All prices are NET prices.

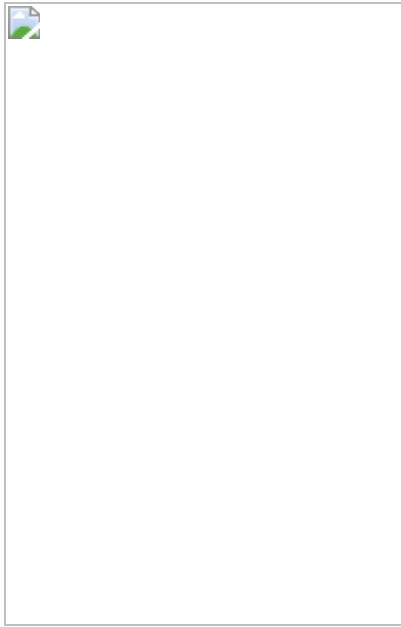
## **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

**Fig. 1: Sequence licking task.**



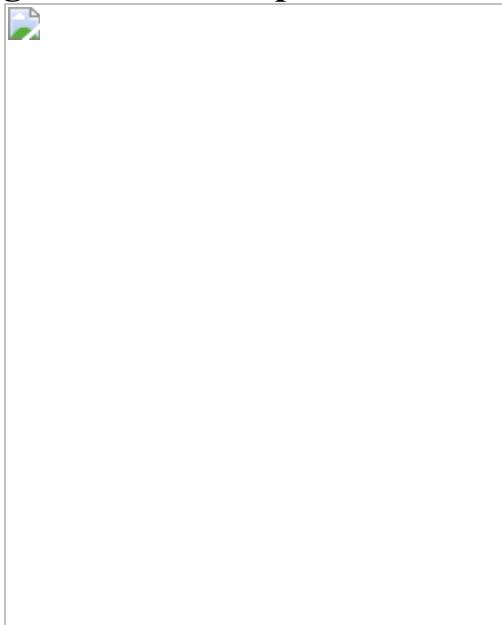
**Fig. 2: Optogenetic inhibition and single-unit activity survey across cortical regions during sequence execution.**



**Fig. 3: Neuronal populations code with increasing levels of abstraction across cortical areas.**



**Fig. 4: Context-dependent coding of subsequences in ALM.**



## Data availability

Data are available from the corresponding author upon request.

## Code availability

The MATLAB code used to analyse the data is available at GitHub and from the corresponding author upon request.

## References

1. Rosenbaum, D. A. *Human Motor Control* (Elsevier, 2010).
2. Mayrhofer, J. M. et al. Distinct contributions of whisker sensory cortex and tongue-jaw motor cortex in a goal-directed sensorimotor transformation. *Neuron* **103**, 1034–1043.e5 (2019).
3. Chen, T.-W., Li, N., Daie, K. & Svoboda, K. A map of anticipatory activity in mouse motor cortex. *Neuron* **94**, 866–879.e4 (2017).
4. Economo, M. N. et al. Distinct descending motor cortex pathways and their roles in movement. *Nature* **563**, 79–84 (2018).
5. Gao, Z. et al. A cortico-cerebellar loop for motor planning. *Nature* **563**, 113–116 (2018).
6. Guo, Z. V. et al. Flow of cortical activity underlying a tactile decision in mice. *Neuron* **81**, 179–194 (2014).
7. Inagaki, H. K., Fontolan, L., Romani, S. & Svoboda, K. Discrete attractor dynamics underlies persistent activity in the frontal cortex. *Nature* **566**, 212–217 (2019).
8. Li, N., Chen, T.-W., Guo, Z. V., Gerfen, C. R. & Svoboda, K. A motor cortex circuit for motor planning and movement. *Nature* **519**, 51–56 (2015).

9. Li, N., Daie, K., Svoboda, K. & Druckmann, S. Robust neuronal dynamics in premotor cortex during motor planning. *Nature* **532**, 459–464 (2016).
10. Komiyama, T. et al. Learning-related fine-scale specificity imaged in motor cortex circuits of behaving mice. *Nature* **464**, 1182–1186 (2010).
11. Kurnikova, A., Moore, J. D., Liao, S.-M., Deschênes, M. & Kleinfeld, D. Coordination of orofacial motor actions into exploratory behavior by rat. *Curr. Biol.* **27**, 688–696 (2017).
12. McElvain, L. E. et al. Circuits in the rodent brainstem that control whisking in concert with other orofacial motor actions. *Neuroscience* **368**, 152–170 (2018).
13. Welker, W. I. Analysis of sniffing of the albino rat 1). *Behaviour* **22**, 223–244 (1964).
14. Chartier, J., Anumanchipalli, G. K., Johnson, K. & Chang, E. F. Encoding of articulatory kinematic trajectories in human speech sensorimotor cortex. *Neuron* **98**, 1042–1054.e4 (2018).
15. Svoboda, K. & Li, N. Neural mechanisms of movement planning: motor cortex and beyond. *Curr. Opin. Neurobiol.* **49**, 33–41 (2018).
16. Ayling, O. G. S., Harrison, T. C., Boyd, J. D., Goroshkov, A. & Murphy, T. H. Automated light-based mapping of motor cortex by photoactivation of channelrhodopsin-2 transgenic mice. *Nat. Methods* **6**, 219–224 (2009).
17. Guo, J.-Z. et al. Cortex commands the performance of skilled movement. *eLife* **4**, e10774 (2015).
18. Clemens, A. M., Fernandez Delgado, Y., Mehlman, M. L., Mishra, P. & Brecht, M. Multisensory and motor representations in rat oral somatosensory cortex. *Sci. Rep.* **8**, 13556 (2018).

19. Proske, U. & Gandevia, S. C. The proprioceptive senses: their roles in signaling body shape, body position and movement, and muscle force. *Physiol. Rev.* **92**, 1651–1697 (2012).
20. Franklin, D. W. & Wolpert, D. M. Computational mechanisms of sensorimotor control. *Neuron* **72**, 425–442 (2011).
21. Shadmehr, R., Smith, M. A. & Krakauer, J. W. Error correction, sensory prediction, and adaptation in motor control. *Annu. Rev. Neurosci.* **33**, 89–108 (2010).
22. Chesler, A. T. et al. The role of PIEZO2 in human mechanosensation. *N. Engl. J. Med.* **375**, 1355–1364 (2016).
23. Inagaki, H. K., Inagaki, M., Romani, S. & Svoboda, K. Low-dimensional and monotonic preparatory activity in mouse anterior lateral motor cortex. *J. Neurosci.* **38**, 4163–4185 (2018).
24. Stapleton, J. R. Rapid taste responses in the gustatory cortex during licking. *J. Neurosci.* **26**, 4126–4138 (2006).
25. Accolla, R., Bathellier, B., Petersen, C. C. H. & Carleton, A. Differential spatial representation of taste modalities in the rat gustatory cortex. *J. Neurosci.* **27**, 1396–1404 (2007).
26. Jin, X. & Costa, R. M. Start/stop signals emerge in nigrostriatal circuits during sequence learning. *Nature* **466**, 457–462 (2010).
27. Kriegeskorte, N. & Douglas, P. K. Interpreting encoding and decoding models. *Curr. Opin. Neurobiol.* **55**, 167–179 (2019).
28. Russo, A. A. et al. Neural trajectories in the supplementary motor area and motor cortex exhibit distinct geometries, compatible with different classes of computation. *Neuron* **107**, 745–758.e6 (2020).
29. Russo, A. A. et al. Motor cortex embeds muscle-like commands in an untangled population response. *Neuron* **97**, 953–966.e8 (2018).

30. Evarts, E. V. & Tanji, J. Reflex and intended responses in motor cortex pyramidal tract neurons of monkey. *J. Neurophysiol.* **39**, 1069–1080 (1976).
31. Heindorf, M., Arber, S. & Keller, G. B. Mouse motor cortex coordinates the behavioral response to unpredicted sensory feedback. *Neuron* **99**, 1040–1054.e5 (2018).
32. Pruszynski, J. A. et al. Primary motor cortex underlies multi-joint integration for fast feedback control. *Nature* **478**, 387–390 (2011).
33. Scott, S. H., Cluff, T., Lowrey, C. R. & Takei, T. Feedback control during voluntary motor actions. *Curr. Opin. Neurobiol.* **33**, 85–94 (2015).
34. Stavisky, S. D., Kao, J. C., Ryu, S. I. & Shenoy, K. V. Motor cortical visuomotor feedback activity is initially isolated from downstream targets in output-null neural state space dimensions. *Neuron* **95**, 195–208.e9 (2017).
35. Bollu, T. et al. Cortex-dependent corrections as the tongue reaches for and misses targets. *Nature* **594**, 82–87 (2021).
36. Tanji, J. Sequential organization of multiple movements: involvement of cortical motor areas. *Annu. Rev. Neurosci.* **24**, 631–651 (2001).
37. Desrochers, T. M., Burk, D. C., Badre, D. & Sheinberg, D. L. The monitoring and control of task sequences in human and non-human primates. *Front. Syst. Neurosci.* **9**, 185 (2016).
38. Shima, K. & Tanji, J. Neuronal activity in the supplementary and presupplementary motor areas for temporal organization of multiple movements. *J. Neurophysiol.* **84**, 2148–2160 (2000).
39. Tanji, J. & Shima, K. Role for supplementary motor area cells in planning several movements ahead. *Nature* **371**, 413–416 (1994).

40. Sohn, J.-W. & Lee, D. Order-dependent modulation of directional signals in the supplementary and presupplementary motor areas. *J. Neurosci.* **27**, 13655–13666 (2007).
41. Chabrol, F. P., Blot, A. & Mrsic-Flogel, T. D. Cerebellar contribution to preparatory activity in motor neocortex. *Neuron* **103**, 506–519.e4 (2019).
42. Vong, L. et al. Leptin action on GABAergic neurons prevents obesity and reduces inhibitory tone to POMC neurons. *Neuron* **71**, 142–154 (2011).
43. Madisen, L. et al. A toolbox of Cre-dependent optogenetic transgenic mice for light-induced activation and silencing. *Nat. Neurosci.* **15**, 793–802 (2012).
44. Zhao, S. et al. Cell type-specific channelrhodopsin-2 transgenic mice for optogenetic dissection of neural circuitry function. *Nat. Methods* **8**, 745–752 (2011).
45. Savitt, J. M. Bcl-x is required for proper development of the mouse substantia nigra. *J. Neurosci.* **25**, 6721–6728 (2005).
46. Taniguchi, H. et al. A resource of Cre driver lines for genetic targeting of GABAergic neurons in cerebral cortex. *Neuron* **71**, 995–1013 (2011).
47. Zhou, X. et al. Deletion of PIK3C3/Vps34 in sensory neurons causes rapid neurodegeneration by disrupting the endosomal but not the autophagic pathway. *Proc. Natl Acad. Sci. USA* **107**, 9424–9429 (2010).
48. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. *Proc. IEEE Conf. Computer Vision and Pattern Recognition* 770–778 (2016).
49. Badrinarayanan, V., Kendall, A. & Cipolla, R. SegNet: a deep convolutional encoder-decoder architecture for image segmentation.

*IEEE Trans. Pattern Anal. Mach. Intell.* **39**, 2481–2495 (2017).

50. Mowery, T. M., Kotak, V. C. & Sanes, D. H. Transient hearing loss within a critical period causes persistent changes to cellular properties in adult auditory cortex. *Cereb. Cortex* **25**, 2083–2094 (2015).
51. Pachitariu, M., Steinmetz, N. A., Kadir, S. N., Carandini, M. & Harris, K. D. Fast and accurate spike sorting of high-channel count probes with KiloSort. *Adv. Neural Inf. Process. Syst.* <https://papers.nips.cc/paper/2016/file/1145a30ff80745b56fb0cecf65305017-Paper.pdf> (2016).
52. Hill, D. N., Mehta, S. B. & Kleinfeld, D. Quality metrics to accompany spike sorting of extracellular signals. *J. Neurosci.* **31**, 8699–8705 (2011).
53. Navratilova, Z., Godfrey, K. B. & McNaughton, B. L. Grids from bands, or bands from grids? An examination of the effects of single unit contamination on grid cell firing fields. *J. Neurophysiol.* **115**, 992–1002 (2016).
54. Saravanan, V., Berman, G. J. & Sober, S. J. Application of the hierarchical bootstrap to multi-level data in neuroscience. Preprint at <https://arxiv.org/abs/2007.07797> (2020).
55. Williams, A. H. & Linderman, S. W. Statistical neuroscience in the single trial limit. *Curr. Opin. Neurobiol.* **70**, 193–205 (2021).
56. Zou, H. & Hastie, T. Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **67**, 301–320 (2005).

## Acknowledgements

We thank W. Olson, R. Dasgupta, Y.-T. Chang, V. Chokshi, J. Cohen, M. Economo and K. Svoboda for comments on the manuscript; A. Aly, Y. Hwang and H. Deng for assistance with experiments and data curation; V. Chokshi and W. Olson for sharing animals; B. Bari and J. Cohen for VGAT-

Cre;Ai32 mice; and T. Babola for suggestions on the hearing loss experiment. This work was supported by NIH grants R01NS089652 and 1R01NS104834-01 to D.H.O., and NIH center grant P30NS050274.

## Author information

### Affiliations

1. The Solomon H. Snyder Department of Neuroscience, Krieger Mind/Brain Institute, Kavli Neuroscience Discovery Institute, Brain Science Institute, The Johns Hopkins University School of Medicine, Baltimore, MD, USA

Duo Xu, Mingyuan Dong, Linghua Zhang & Daniel H. O'Connor

2. Undergraduate Studies, Krieger School of Arts and Sciences, Johns Hopkins University, Baltimore, MD, USA

Yuxi Chen, Angel M. Delgado & Natasha C. Hughes

### Contributions

D.X., M.D., Y.C., A.M.D., N.C.H. and L.Z. performed the experiments. D.X. developed custom software, hardware, the analysis code and analysed the data, with input from all authors. D.X., M.D., N.C.H. and D.H.O. wrote the paper with input from all authors.

### Corresponding author

Correspondence to [Daniel H. O'Connor](#).

### Ethics declarations

### Competing interests

The authors declare no competing interests.

## Peer review information

*Nature* thanks the anonymous reviewers for their contribution to the peer review of this work.

## Additional information

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Extended data figures and tables

### [Extended Data Fig. 1 Behavioral measurements, performance, and control experiments.](#)

**a**, Confusion matrix showing the performance of the classification network. The numbers represent percentages within each (true) class ( $n = 1696$  frames). **b**, Performance of the regression network. Top, the gray probability distribution shows how  $L$  from five human individuals varied from the mean  $L$  across the five. The red distribution shows how predicted  $L$  varied from the human mean. Bottom, similar quantification as the top but for  $\theta$ .  $n = 573$  frames. **c**, CAD images of the sensor core (left) and the assembly (right) with a lick tube. **d**, Linear relationship between the applied force and the sensor output voltage. **e**, Two example trials showing the trajectories of the tongue tip when a mouse sequentially reached the 7 port positions, for both sequence directions. Arrows indicate the direction of time within each trajectory. **f**, Patterns of kinematics and forces of single licks at each port position ( $n = 25683$  trials from 17 mice; mean  $\pm$  95% bootstrap confidence interval). The duration of individual licks was normalized. **g**, Top, the pattern of angle deviation from midline ( $|\theta|$ ) of single licks pooled from R3 and L3. The vertical line indicates maximum  $|\theta|$  ( $|\theta|_{max}$ ). Middle, tongue length ( $L$ ) expressed as a fraction of its maximum ( $L_{max}$ ). The horizontal line indicates, on average, the fraction where  $|\theta|_{max}$  occurred. Bottom, time aligned probability distributions showing when touch onset,  $|\theta|_{max}$ ,  $L_{max}$ , or  $\theta_{shoot}$  occurred. Red lines mark quartiles.  $n =$

25683 trials from 17 mice. Lick patterns show mean  $\pm$  95% bootstrap confidence interval. **h**, Top, probability distributions of  $L_{max}$  and  $\theta_{touch}$  for licks at each port position. Bottom, probability distributions of the change in  $\Theta_{touch}$  ( $\Delta\Theta_{touch}$ ) and instantaneous sequence speed (Methods) for each interval separating port positions. Distributions show mean  $\pm$  s.d. across n = 17 mice. **i**, Median time to first touch (top) and the average number of missed licks during sequence performance (bottom) in control (Sham) versus hearing loss (Earplug) conditions. Bars show group means and lines show data from individual mice. \*\*\* p < 0.001, n.s. p > 0.05, paired one-tailed bootstrap test, n = 5 mice. **j**, Average number of missed licks before first touch (top) and during sequence performance (bottom) in control (Normal) versus odor masking (Masked) conditions. Same statistical tests as in (i), n = 6 mice. **k**, Similar to (j) but comparing control (Saline) versus tongue numbing (Lidocaine) conditions. n = 5 mice. **l**, Learning curves for 15 individual mice (gray) and the mean (black) showing a reduction in sequence initiation time (left) in response to the auditory cue and an increase in sequence speed (right). The three red asterisks correspond to the three examples of sequence performance shown in (n). **m**, Gradual increase in task difficulty (Methods) accompanying the improved performance shown in (l). **n**, Depiction of example sequences performed by a mouse in alternating directions across consecutive trials at different stages of learning. Trial onsets are marked by yellow bars. Port positions shown in the black trace are overlaid with touch onsets (dots).

## Extended Data Fig. 2 Performance in backtracking sequences.

**a**, Learning curves for 9 individual mice (gray) and the mean (black) showing the duration of time spent to perform standard (left) and backtracking (right) sequences. **b**, Similar to (a) but limited to the interval following the middle lick in standard (left) or backtracking (right) sequences. **c**,  $L$ ,  $L'$  and  $\Theta$  patterns for seven consecutive licks aligned at the Mid touch (number 0). Licks in standard sequences (n = 7458 trials) are shown in black, those in backtracking sequences (n = 2695 trials) are in green. Mean  $\pm$  s.d. **d**, Probability distributions of instantaneous lick rate for each interval separating consecutive pairs of the seven licks during standard (black) or backtracking (green) sequences (n = 8 mice; mean  $\pm$  s.d.). **e**, Top, time to locate the port at its next position during the 4th interval, for

standard sequences (black) or for sequences when the port backtracked (green). Bottom, the number of missed licks during the 4th interval. Mean  $\pm$  95% bootstrap confidence interval. n = 7458 standard and 2695 backtracking sequences from 47 total sessions.

### **Extended Data Fig. 3 Closed-loop optogenetic inhibition defines cortical areas involved in sequence control.**

**a**, Left, dorsal view of an example “clear-skull” preparation. Right, table shows the center coordinates used for illumination for each target region. **b**, Triggering scheme for photoinhibition at sequence initiation, mid-sequence and water consumption. **c**, Average spike waveform of putative pyramidal cells (black; n = 224) and putative FS neurons (blue; n = 117), normalized to the amplitude of negative peaks. **d**, Relationship between spike widths (defined as the trough to peak time of average waveform) and changes in mean spike rate under opto illumination (4 mW, within 1 mm) relative to baseline. Pyramidal cells (black; n = 42) and FS neurons (blue; n = 41) were classified by the two thresholds (dashed lines at 0.4 and 0.5 ms) with ambiguous units (gray; n = 6) in the middle. **e**, Distributions of spike widths from neurons in **(d)** (filled bars; n = 89) and from all neurons (empty bars; n = 414) including those where illuminations were not at recording sites. Classification thresholds are shown in dashed lines. **f**, Left, inhibition efficiency of putative pyramidal cells as a function of light power and distance away from the center of illumination (n = 224 units total). Right, similar to left but showing the excitation efficiency of putative FS neurons (n = 117 units total). Mean  $\pm$  95% hierarchical bootstrap confidence interval. **g**, Example trial with S1TJ inhibition triggered at mid-sequence. Instantaneous tongue angle ( $\Theta$ ) and length (L) are shown in lighter traces. Shooting angles ( $\Theta_{shoot}$ ) and maximum length ( $L_{max}$ ) of each lick are marked using stems on top of the instantaneous traces. The blue waveform indicates photostimulation. Traces and markers during photostimulation are colored blue. **h**, Similar to **(g)** but inhibiting ALM–M1TJ. **i**, Similar to **(g)** but inhibiting S1BF. **j**, Changes in licking kinematics (rows) when inhibiting each of the five brain regions (columns), quantified across all three inhibition periods (Methods). Bar plots show mean  $\pm$  99% hierarchical bootstrap confidence interval. Gray lines show the data of individual mice. Two-tailed hierarchical bootstrap test, \*\*\* p < 0.001, \*\* p < 0.01, \* p

$< 0.05$ , n.s.  $p \geq 0.05$ , after Bonferroni correction for 15 comparisons. **k**, Changes in the rate of lick (solid bars) and touch (dashed bars) at each of the inhibition periods (rows) when inhibiting each of the five brain regions (columns). Plot style and statistical tests are the same as in (j) but using Bonferroni correction for 30 comparisons. **l**, Same convention as in (j) but showing results with half-power (2 mW) inhibition. **m**, Same convention as in (k) but showing results with half-power (2 mW) inhibition.

## Extended Data Fig. 4 Characterization of single-unit responses.

**a**, Contamination rates and refractory period violation rates of all recorded single- (green) and multi-units (black). The shaded region shows the thresholds for assignment as multi- vs single-unit. **b**, CDF of contamination rate including single- (green) and multi-units (gray). **c**, Same as (b) but for refractory period violation rate. **d**, The number of single- (green) and multi-units (black) recorded in each session, grouped by brain area. **e**, ISI histograms of randomly selected single-units. Refractory period violation rates (RPV) and contamination rates (Con) are labeled on the top (in percent). **f**, Responses of three simultaneously recorded S1TJ neurons during right-to-left (top half) or left-to-right (bottom half) licking sequences, aligned at cue onset (left column), middle touch (middle column), and the last consummatory touch (right column). For each sequence direction, the first row shows rasters of lick times (touches in black and misses in gray) from 10 selected trials (Methods). Stacked below are spike rasters and the corresponding PETHs (mean  $\pm$  s.e.) from the same 10 trials for each example neuron. **g**, Same as (f) but for three example neurons from M1TJ. **h**, Same as (f) but for three example neurons from ALM. **i**, Uncertainty in mean spike rate (normalized to peak) estimated by bootstrap crossvalidation (Methods). Each data point is the bootstrap average value of the root mean squared error (RMSE) for a single neuron. Data ( $n = 804$  neurons) are grouped by brain region and presented in whisker-box plots (centre mark: median, bounds of box: 25th and 75th percentiles, max whisker length: 1.5 times IQR, no max or min limit). **j**, The probability (mean  $\pm$  95% bootstrap confidence interval) of a PETH being consistently grouped into the same cluster across bootstrap iterations for different total numbers of clusters. Maximal consistency was achieved when using thirteen clusters for NNMF (arrow). **k**, NNMF components that

represent each of the thirteen PETH clusters. Right-to-left (blue) and left-to-right (red) activities (mean  $\pm$  95% bootstrap confidence interval) are overlaid together. The vertical lines are located at time zero in each period. The height of the lines represents the scale of normalized neuronal activity from 0 to 1. **l**, Histograms of PETH peak times. Plot organization and time alignment are the same as in (f). **m**, Proportions of neurons from different clusters at different cortical depths. Some clusters with similar types of response were grouped together for better readability. ALM ( $n = 324$ ), M1TJ ( $n = 233$ ) and S1TJ ( $n = 119$ ). **n**, Quantification of rhythmicity in PETHs. Black traces are mid-sequence PETHs of three example neurons in (f), (g), and (h). Colored traces show the best fit licking rhythms (6.5 Hz sinusoids). Average Pearson's correlation coefficients ( $r_{avg}$ ) of the left-to-right and right-to-left fits are shown beneath neuron IDs. **o**, Empirical CDFs of  $r_{avg}$  for neurons in S1TJ, M1TJ, and ALM. Circles mark the values of the 9 example neurons in (f), (g), and (h).

### Extended Data Fig. 5 Additional analyses for population coding.

**a**, Relative goodness of fit of tongue angle regressions with a range of shifts in spike times. Black traces and error bars show mean  $\pm$  95% bootstrap confidence interval. Lighter traces show individual recordings. (S1TJ,  $n = 8$  sessions; M1TJ,  $n = 9$  sessions; ALM,  $n = 13$  sessions). **b**, Absolute pairwise Pearson's correlation coefficients among the five behavioral variables (mean;  $n = 35$  sessions). **c**, Single-trial decoding of the five behavioral variables (rows; black traces) from 29 simultaneously recorded S1TJ units in a right-to-left (left) and a left-to-right (right) sequence. **d**, Same as (c) but decoding from 58 simultaneously recorded units in ALM. **e**, Total percent variance explained (VE) by the first five principal components (left in each region) versus that by the five coding axes (right in each region) during sequence execution. Lighter lines show individual recording sessions and thicker lines show the means. **f**, Absolute pairwise cosine values among coding axes (mean;  $n = 35$  sessions). **g**, Cumulative time histograms showing the fraction of trials that could be correctly classified as a standard vs backtracking sequence as time progresses. Two-tailed

bootstrap test, \*\*\*  $p \approx 0$ , n.s.  $p = 0.91$ . **h**, Same as sequence progress in Fig. 3a, b, but for “zigzag” sequences.

## Extended Data Fig. 6 Reward modulation of activity in ALM.

**a**, Responses of two simultaneously recorded ALM neurons (number 1 and number 2) aligned at the first lick (specifically the middle of a tongue-out period) that touched water reward. For each sequence direction, shown at top are rasters of lick times (touches in black and misses in gray) and the duration of water delivery (blue) from 20 selected trials (Methods). Stacked below are spike rasters and the corresponding PETHs from the same 20 trials for each example neuron. **b**, The probability of licking (i.e. tongue-out) as a function of time. Licks are sequentially indexed with respect to the first lick (number 0) touching the water. **c**, Patterns of kinematics and force for single licks around the first lick (number 0) touching water ( $n = 25289$  trials; mean  $\pm$  95% bootstrap confidence interval). The duration of individual licks was normalized. The total force ( $F_{total}$ ) is the vector sum of vertical and lateral forces. **d**, Decoding of  $\tau$ ,  $I$  and  $\theta$  (mean  $\pm$  99% bootstrap confidence interval) from neuronal populations recorded in ALM ( $n = 13$  sessions), M1TJ ( $n = 9$  sessions), and S1TJ ( $n = 8$  sessions) in right-to-left (blue) or left-to-right (red) trials around the consumption period. **e**, The difference between the decoded  $\theta$  traces in right-to-left versus left-to-right trials. Same data source, mean and error presentation as in (d).

## Extended Data Fig. 7 Coding of upcoming sequences in ALM.

**a**, Depiction of sequences performed by a mouse in alternating directions across 14 consecutive trials. Trial onsets are marked by yellow lines. Port positions shown in the black trace are overlaid with touch onsets (dots). **b**, Probability distributions of  $\Theta_{shoot}$  (left) and  $L_{max}$  (middle) for the first 3 licks at the start of a sequence ( $n = 8$  mice; mean  $\pm$  s.d.). The negative y-axis of  $\Theta_{shoot}$  points to the side at which the port is located. The CDF (right; 8 individual mice in gray and the mean in black) of the maximal  $\Theta_{shoot}$  explored before touching the port (at the side of negative  $\Theta_{shoot}$ ). The blue line shows the probability of successfully locating the port without exploring beyond the midline. **c**, Top, rasters of two example neurons which

had persistent and target position (*TP*) selective firing during the 14 consecutive trials in (a). Bottom, normalized and smoothed (0.25 s s.d. Gaussian kernel) spike rates of the two neurons. d, Decoded instantaneous *TP* (dark trace) from 58 simultaneously recorded units in ALM, overlaid with normalized port position (light trace). e, Decoding of *TP* from ALM (mean  $\pm$  99% bootstrap confidence interval) before upcoming right-to-left trials (blue) or left-to-right trials (red). Crossvalidated  $R^2$  is shown (mean  $\pm$  s.d.; n = 13 sessions). f, Goodness of fit for linear models that predict *TP* during ITIs, quantified by crossvalidated  $R^2$ . g, Using the same linear models in (e) to decode *TP* during execution of standard right-to-left (blue) or left-to-right (red) sequences (mean  $\pm$  99% bootstrap confidence interval). h, Same as (f) but for  $\tau$ .

## Supplementary information

### Reporting Summary

### Supplementary Table 1

**Mouse information.** Excel workbook detailing assignment of individual mice to experiments and figures.

### Supplementary Video 1

**Example performance in a standard sequence.** Top, high-speed video capturing the bottom and side views of the mouse. The tracked base and the tip of the tongue are labeled by red asterisks. Bottom, time-aligned behavioral variables and events similar to Fig. 1d. A moving bar indicates the position of the current frame. The playback was slowed down 5-fold.

### Supplementary Video 2

**Example performance in a backtracking sequence.** Same conventions as for Supplementary Video 1.

### Supplementary Video 3

**Example trial with inhibition in S1TJ at sequence initiation.** Similar conventions as for Supplementary Video 1 but showing the waveform of optogenetic inhibition (instead of the vertical and lateral lick forces).

## Supplementary Video 4

**Example trial with inhibition in ALM at sequence initiation.** Same conventions as for Supplementary Video 3.

## Supplementary Video 5

**Example trial with low-power inhibition in M1B at consumption.** Same conventions as for Supplementary Video 3.

## Supplementary Video 6

**Example performance in a zigzag sequence.** Same conventions as for Supplementary Video 1.

## Rights and permissions

## Reprints and Permissions

## About this article

### Cite this article

Xu, D., Dong, M., Chen, Y. *et al.* Cortical processing of flexible and context-dependent sensorimotor sequences. *Nature* **603**, 464–469 (2022). <https://doi.org/10.1038/s41586-022-04478-7>

- Received: 10 January 2020
- Accepted: 26 January 2022
- Published: 09 March 2022

- Issue Date: 17 March 2022
- DOI: <https://doi.org/10.1038/s41586-022-04478-7>

## Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-022-04478-7>

| [Section menu](#) | [Main menu](#) |

- Article
- [Published: 02 March 2022](#)

# FSH blockade improves cognition in mice with Alzheimer's disease

- [Jing Xiong<sup>1,2</sup> na1](#),
- [Seong Su Kang<sup>1</sup> na1](#),
- [Zhihao Wang<sup>1</sup>](#),
- [Xia Liu<sup>1</sup>](#),
- [Tan-Chun Kuo](#) [ORCID: orcid.org/0000-0001-5301-755X<sup>3</sup>](#),
- [Funda Korkmaz](#) [ORCID: orcid.org/0000-0002-9174-8369<sup>3</sup>](#),
- [Ashley Padilla](#) [ORCID: orcid.org/0000-0003-4877-5465<sup>3</sup>](#),
- [Sari Miyashita<sup>3</sup>](#),
- [Pokman Chan<sup>4</sup>](#),
- [Zhaohui Zhang<sup>2</sup>](#),
- [Pavel Katsel](#) [ORCID: orcid.org/0000-0001-8076-0162<sup>5</sup>](#),
- [Jocoll Burgess<sup>3,5</sup>](#),
- [Anisa Gumerova<sup>3</sup>](#),
- [Kseniia Ievleva<sup>3</sup>](#),
- [Damini Sant<sup>3</sup>](#),
- [Shan-Ping Yu](#) [ORCID: orcid.org/0000-0003-1335-1398<sup>6</sup>](#),
- [Valeriia Muradova](#) [ORCID: orcid.org/0000-0003-3257-0780<sup>3</sup>](#),
- [Tal Frolinger<sup>3</sup>](#),
- [Daria Lizneva<sup>3</sup>](#),
- [Jameel Iqbal<sup>3</sup>](#),
- [Ki A. Goosens](#) [ORCID: orcid.org/0000-0002-5246-2261<sup>3,5</sup>](#),
- [Sakshi Gera<sup>3</sup>](#),
- [Clifford J. Rosen<sup>7</sup>](#),
- [Vahram Haroutunian](#) [ORCID: orcid.org/0000-0001-5860-2512<sup>5</sup>](#),

- [Vitaly Ryu<sup>3</sup>](#),
- [Tony Yuen<sup>3</sup>](#),
- [Mone Zaidi<sup>3</sup>](#) &
- [Keqiang Ye](#) [ORCID: orcid.org/0000-0002-7657-8154<sup>1,8</sup>](#)

*Nature* volume **603**, pages 470–476 (2022)

- 15k Accesses
- 2 Citations
- 399 Altmetric
- [Metrics details](#)

## Subjects

- [Alzheimer's disease](#)
- [Translational research](#)

## Abstract

Alzheimer’s disease has a higher incidence in older women, with a spike in cognitive decline that tracks with visceral adiposity, dysregulated energy homeostasis and bone loss during the menopausal transition<sup>1,2</sup>. Inhibiting the action of follicle-stimulating hormone (FSH) reduces body fat, enhances thermogenesis, increases bone mass and lowers serum cholesterol in mice<sup>3,4,5,6,7</sup>. Here we show that FSH acts directly on hippocampal and cortical neurons to accelerate amyloid- $\beta$  and Tau deposition and impair cognition in mice displaying features of Alzheimer’s disease. Blocking FSH action in these mice abrogates the Alzheimer’s disease-like phenotype by inhibiting the neuronal C/EBP $\beta$ - $\delta$ -secretase pathway. These data not only suggest a causal role for rising serum FSH levels in the exaggerated Alzheimer’s disease pathophysiology during menopause, but also reveal an opportunity for treating Alzheimer’s disease, obesity, osteoporosis and dyslipidaemia with a single FSH-blocking agent.

This is a preview of subscription content

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

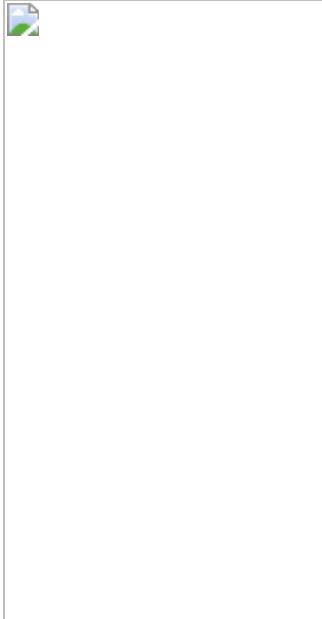
[Buy](#)

All prices are NET prices.

## **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

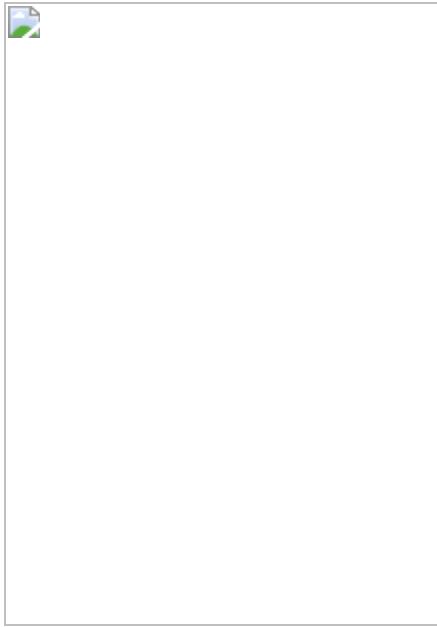
**Fig. 1: FSH-blocking antibody reverses AD neuropathology and cognitive decline in Alzheimer's mice.**



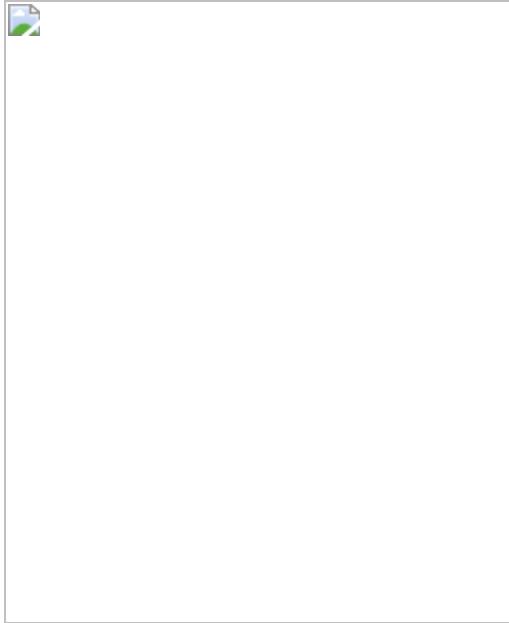
**Fig. 2: Neuronal FSH receptors in mouse and human brain.**



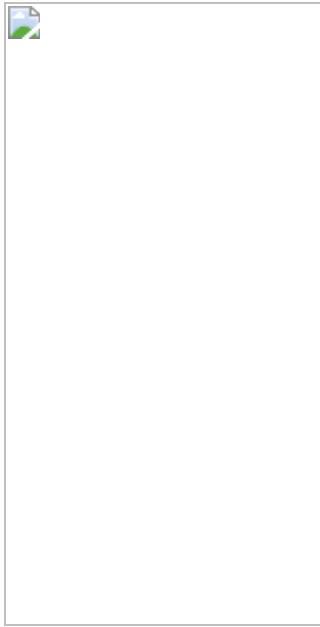
**Fig. 3: Targeted *Fshr* knockdown in the hippocampus ameliorates AD neuropathology and impaired spatial memory.**



**Fig. 4: Recombinant FSH induces AD pathologies and cognitive decline in 3xTg Mice.**



**Fig. 5: FSH-induced AD pathology is dampened in *Cebpb*<sup>+/-</sup> 3xTg mice.**



## Data availability

All original western blots are provided in the [Supplementary Information](#). The original unedited camera images for histology and immunohistochemistry are available online (<https://osf.io/9hp8r/>). There are no restrictions on data availability. Unique biological material will be made available to other investigators on request. [Source data](#) are provided with this paper.

## References

1. Sowers, M. et al. Changes in body composition in women over six years at midlife: ovarian and chronological aging. *J. Clin. Endocrinol. Metab.* **92**, 895–901 (2007).
2. Sowers, M. R. et al. Hormone predictors of bone mineral density changes during the menopausal transition. *J. Clin. Endocrinol. Metab.* **91**, 1261–1267 (2006).
3. Guo, Y. et al. Blocking FSH inhibits hepatic cholesterol biosynthesis and reduces serum cholesterol. *Cell Res.* **29**, 151–166 (2019).

4. Han, X. et al. A novel follicle-stimulating hormone vaccine for controlling fat accumulation. *Theriogenology* **148**, 103–111 (2020).
5. Ji, Y. et al. Epitope-specific monoclonal antibodies to FSH $\beta$  increase bone mass. *Proc. Natl Acad. Sci. USA* **115**, 2192–2197 (2018).
6. Liu, P. et al. Blocking FSH induces thermogenic adipose tissue and reduces body fat. *Nature* **546**, 107–112 (2017).
7. Geng, W. et al. Immunization with FSH $\beta$  fusion protein antigen prevents bone loss in a rat ovariectomy-induced osteoporosis model. *Biochem. Biophys. Res. Commun.* **434**, 280–286 (2013).
8. Fisher, D. W., Bennett, D. A. & Dong, H. Sexual dimorphism in predisposition to Alzheimer's disease. *Neurobiol. Aging* **70**, 308–324 (2018).
9. Andersen, K. et al. Gender differences in the incidence of AD and vascular dementia: the EURODEM studies. EURODEM Incidence Research Group. *Neurology* **53**, 1992–1997 (1999).
10. Marongiu, R. Accelerated ovarian failure as a unique model to study peri-menopause influence on Alzheimer's disease. *Front. Aging Neurosci.* **11**, 242 (2019).
11. Matyi, J. M., Rattiner, G. B., Schwartz, S., Buhusi, M. & Tschanz, J. T. Lifetime estrogen exposure and cognition in late life: the Cache County study. *Menopause* **26**, 1366–1374 (2019).
12. Zandi, P. P. et al. Hormone replacement therapy and incidence of Alzheimer disease in older women: the Cache County Study. *JAMA* **288**, 2123–2129 (2002).
13. O'Brien, J., Jackson, J. W., Grodstein, F., Blacker, D. & Weuve, J. Postmenopausal hormone therapy is not associated with risk of all-cause dementia and Alzheimer's disease. *Epidemiol. Rev.* **36**, 83–103 (2014).

14. Shumaker, S. A. et al. Conjugated equine estrogens and incidence of probable dementia and mild cognitive impairment in postmenopausal women: Women's Health Initiative Memory Study. *JAMA* **291**, 2947–2958 (2004).
15. Short, R. A., Bowen, R. L., O'Brien, P. C. & Graff-Radford, N. R. Elevated gonadotropin levels in patients with Alzheimer disease. *Mayo Clin. Proc.* **76**, 906–909 (2001).
16. Bowen, R. L., Isley, J. P. & Atkinson, R. L. An association of elevated serum gonadotropin concentrations and Alzheimer disease? *J. Neuroendocrinol.* **12**, 351–354 (2000).
17. Randolph, J. F. Jr et al. Change in follicle-stimulating hormone and estradiol across the menopausal transition: effect of age at the final menstrual period. *J. Clin. Endocrinol. Metab.* **96**, 746–754 (2011).
18. Epperson, C. N., Sammel, M. D. & Freeman, E. W. Menopause effects on verbal memory: findings from a longitudinal community cohort. *J. Clin. Endocrinol. Metab.* **98**, 3829–3838 (2013).
19. Greendale, G. A. et al. Effects of the menopause transition and hormone use on cognitive performance in midlife women. *Neurology* **72**, 1850–1857 (2009).
20. Meyer, P. M. et al. A population-based longitudinal study of cognitive functioning in the menopausal transition. *Neurology* **61**, 801–806 (2003).
21. Zhu, L. L. et al. Blocking antibody to the β-subunit of FSH prevents bone loss by inhibiting bone resorption and stimulating bone synthesis. *Proc. Natl Acad. Sci. USA* **109**, 14574–14579 (2012).
22. Oddo, S. et al. Triple-transgenic model of Alzheimer's disease with plaques and tangles: intracellular A $\beta$  and synaptic dysfunction. *Neuron* **39**, 409–421 (2003).

23. Webster, S. J., Bachstetter, A. D., Nelson, P. T., Schmitt, F. A. & Van Eldik, L. J. Using mice to model Alzheimer's dementia: an overview of the clinical disease and the preclinical behavioral changes in 10 mouse models. *Front. Genet.* **5**, 88 (2014).
24. Carroll, J. C. et al. Progesterone and estrogen regulate Alzheimer-like neuropathology in female 3xTg-AD mice. *J. Neurosci.* **27**, 13357–13365 (2007).
25. Zhang, Z. et al. Cleavage of tau by asparagine endopeptidase mediates the neurofibrillary pathology in Alzheimer's disease. *Nat. Med.* **20**, 1254–1262 (2014).
26. Zhang, Z. et al. Delta-secretase cleaves amyloid precursor protein and regulates the pathogenesis in Alzheimer's disease. *Nat. Commun.* **6**, 8762 (2015).
27. Rocca, W. A. et al. Increased risk of cognitive impairment or dementia in women who underwent oophorectomy before menopause. *Neurology* **69**, 1074–1083 (2007).
28. Tokuyama, N. et al. Individual and combining effects of anti-RANKL monoclonal antibody and teriparatide in ovariectomized mice. *Bone Rep.* **2**, 1–7 (2015).
29. Rosen, C. J. & Zaidi, M. Contemporaneous reproduction of preclinical science: a case study of FSH and fat. *Ann. N. Y. Acad. Sci.* **1404**, 17–19 (2017).
30. Minkeviciene, R. et al. Age-related decrease in stimulated glutamate release and vesicular glutamate transporters in APP/PS1 transgenic and wild-type mice. *J. Neurochem.* **105**, 584–594 (2008).
31. Onos, K. D. et al. Enhancing face validity of mouse models of Alzheimer's disease with natural genetic variation. *PLoS Genet.* **15**, e1008155 (2019).

32. Volianskis, A., Kostner, R., Molgaard, M., Hass, S. & Jensen, M. S. Episodic memory deficits are not related to altered glutamatergic synaptic transmission and plasticity in the CA1 hippocampus of the APPswe/PS1deltaE9-deleted transgenic mice model of ss-amyloidosis. *Neurobiol. Aging* **31**, 1173–1187 (2010).
33. Araujo, A. B. & Wittert, G. A. Endocrinology of the aging male. *Best Pract. Res. Clin. Endocrinol. Metab.* **25**, 303–319 (2011).
34. Casadesus, G. et al. Increases in luteinizing hormone are associated with declines in cognitive performance. *Mol. Cell. Endocrinol.* **269**, 107–111 (2007).
35. Berry, A., Tomidokoro, Y., Ghiso, J. & Thornton, J. Human chorionic gonadotropin (a luteinizing hormone homologue) decreases spatial memory and increases brain amyloid- $\beta$  levels in female rats. *Horm. Behav.* **54**, 143–152 (2008).
36. Liu, T., Wimalasena, J., Bowen, R. L. & Atwood, C. S. Luteinizing hormone receptor mediates neuronal pregnenolone production via up-regulation of steroidogenic acute regulatory protein expression. *J. Neurochem.* **100**, 1329–1339 (2007).
37. Gera, S. et al. First-in-class humanized FSH blocking antibody targets bone and fat. *Proc. Natl Acad. Sci. USA* **117**, 28971–28979 (2020).
38. Straccia, M. et al. Pro-inflammatory gene expression and neurotoxic effects of activated microglia are attenuated by absence of CCAAT/enhancer binding protein  $\beta$ . *J. Neuroinflammation* **8**, 156 (2011).
39. Ramji, D. P. & Foka, P. CCAAT/enhancer-binding proteins: structure, function and regulation. *Biochem. J.* **365**, 561–575 (2002).
40. Wang, Z. H. et al. Delta-secretase phosphorylation by SRPK2 enhances its enzymatic activity, provoking pathogenesis in Alzheimer's disease. *Mol. Cell* **67**, 812–825 (2017).

41. Sun, L. et al. FSH directly regulates bone mass. *Cell* **125**, 247–260 (2006).
42. Hammond, S. L., Leek, A. N., Richman, E. H. & Tjalkens, R. B. Cellular selectivity of AAV serotypes for gene delivery in neurons and astrocytes by neonatal intracerebroventricular injection. *PLoS ONE* **12**, e0188830 (2017).
43. von Jonquieres, G. et al. Recombinant human myelin-associated glycoprotein promoter drives selective AAV-mediated transgene expression in oligodendrocytes. *Front. Mol. Neurosci.* **9**, 13 (2016).
44. Randolph, J. F. Jr. et al. Reproductive hormones in the early menopausal transition: relationship to ethnicity, body size, and menopausal status. *J. Clin. Endocrinol. Metab.* **88**, 1516–1522 (2003).
45. Ashe, K. H. & Zahs, K. R. Probing the biology of Alzheimer’s disease in mice. *Neuron* **66**, 631–645 (2010).
46. Millward, C. A. et al. Mice with a deletion in the gene for CCAAT/enhancer-binding protein  $\beta$  are protected against diet-induced obesity. *Diabetes* **56**, 161–167 (2007).
47. Zaidi, M. et al. Actions of pituitary hormones beyond traditional targets. *J. Endocrinol.* **237**, R83–R98 (2018).
48. Silverman, E., Eimerl, S. & Orly, J. CCAAT enhancer-binding protein  $\beta$  and GATA-4 binding regions within the promoter of the steroidogenic acute regulatory protein (StAR) gene are required for transcription in rat ovarian cells. *J. Biol. Chem.* **274**, 17987–17996 (1999).
49. Sirois, J. & Richards, J. S. Transcriptional regulation of the rat prostaglandin endoperoxide synthase 2 gene in granulosa cells. Evidence for the role of a cis-acting C/EBP $\beta$  promoter element. *J. Biol. Chem.* **268**, 21931–21938 (1993).

50. Wang, H., Liu, X., Chen, S. & Ye, K. Spatiotemporal activation of the C/EBP $\beta$ / $\delta$ -secretase axis regulates the pathogenesis of Alzheimer's disease. *Proc. Natl Acad. Sci. USA* **115**, E12427–E12434 (2018).
51. Sterneck, E., Tessarollo, L. & Johnson, P. F. An essential role for C/EBP $\beta$  in female reproduction. *Genes Dev.* **11**, 2153–2162 (1997).
52. Minkeviciene, R. et al. Amyloid  $\beta$ -induced neuronal hyperexcitability triggers progressive epilepsy. *J. Neurosci.* **29**, 3453–3462 (2009).
53. Zhang, Z. et al. 7,8-dihydroxyflavone prevents synaptic loss and memory deficits in a mouse model of Alzheimer's disease. *Neuropsychopharmacology* **39**, 638–650 (2014).
54. Xiang, J. et al. Delta-secretase-cleaved Tau antagonizes TrkB neurotrophic signalings, mediating Alzheimer's disease pathologies. *Proc. Natl Acad. Sci. USA* **116**, 9094–9102 (2019).
55. Leger, M. et al. Object recognition test in mice. *Nat. Protoc.* **8**, 2531–2537 (2013).
56. Ioannidis, J. P. Why most published research findings are false. *PLoS Med.* **2**, e124 (2005).
57. Collins, F. S. & Tabak, L. A. Policy: NIH plans to enhance reproducibility. *Nature* **505**, 612–613 (2014).
58. McNutt, M. Reproducibility. *Science* **343**, 229 (2014).
59. Mullard, A. Cancer reproducibility project yields first results. *Nat. Rev. Drug Discov.* **16**, 77 (2017).
60. Horrigan, S. K. et al. Replication study: melanoma genome sequencing reveals frequent PREX2 mutations. *eLife* **6**, e21634 (2017).
61. Horrigan, S. K., Reproducibility Project: Cancer Biology. Replication study: the CD47-signal regulatory protein alpha (SIRPa) interaction is a therapeutic target for human solid tumors. *eLife* **6**, e18173 (2017).

## Acknowledgements

This work performed at Emory University School of Medicine was supported by NIH grants RF1 AG051538 and R01 AG065177 to K.Y. K.Y. thanks the staff at the Alzheimer's Disease Research Center at Emory University for human samples; the Rodent Behavioral Core (RBC), one of the Emory Integrated Core Facilities; the Viral Vector Core of the Emory Neuroscience NINDS Core Facility (P30 NS055077); and the NIH's Georgia Clinical and Translational Science Alliance (UL1 TR002378). Work at Icahn School of Medicine at Mount Sinai performed at the Center for Translational Medicine and Pharmacology (CeTMaP) was supported by U19 AG060917 to M.Z. and C.J.R.; R01 DK113627 to M.Z. and J.I.; and R01 AG071870, R01 AG074092 and U01 AG073148 to T.Y. and M.Z. M.Z. thanks the Harrington Discovery Institute for the Innovator–Scholar Award towards development of the FSH-Ab. C.J.R. acknowledges support from the NIH (P20 GM121301 to C.J.R.). We thank M. Ehrlich and S. Gandy for their intellectual contributions, and S. Babunovic for proofreading the final version of the paper.

## Author information

### Author notes

1. These authors contributed equally: Jing Xiong, Seong Su Kang

## Affiliations

1. Department of Pathology and Laboratory Medicine, Emory University School of Medicine, Atlanta, GA, USA

Jing Xiong, Seong Su Kang, Zhihao Wang, Xia Liu & Keqiang Ye

2. Department of Neurology, Renmin Hospital of Wuhan University, Wuhan, China

Jing Xiong & Zhaohui Zhang

3. Center for Translational Medicine and Pharmacology and Departments of Pharmacological Sciences and Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Tan-Chun Kuo, Funda Korkmaz, Ashley Padilla, Sari Miyashita, Jocoll Burgess, Anisa Gumerova, Kseniia Ievleva, Damini Sant, Valeriia Muradova, Tal Frolinger, Daria Lizneva, Jameel Iqbal, Ki A. Goosens, Sakshi Gera, Vitaly Ryu, Tony Yuen & Mone Zaidi

4. Alamak Biosciences, Beverly, MA, USA

Pokman Chan

5. Department of Psychiatry and Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Pavel Katsel, Jocoll Burgess, Ki A. Goosens & Vahram Haroutunian

6. Department of Anesthesiology, Emory University School of Medicine, Atlanta, GA, USA

Shan-Ping Yu

7. Maine Medical Center Research Institute, Scarborough, ME, USA

Clifford J. Rosen

8. Faculty of Life and Health Sciences, and Brain Cognition and Brain Disease Institute (BCBDI), Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China

Keqiang Ye

## Contributions

K.Y. conceived the idea that FSH may be a mediator of AD in post-menopausal women and, together with M.Z., jointly proposed that blocking FSH using M.Z.'s polyclonal FSH-Ab could prevent AD. Thereafter, V.R.,

M.Z. and T.Y. designed the experiments, analysed the data and jointly wrote the manuscript. J.X., S.S.K. and Z.W. designed and performed most of the experiments and analysed the data. X.L. prepared primary neurons and assisted with mouse experiments. P.C. performed RNAscope studies. V.R. and A.G. recorded and analysed RNAscope data. F.K., J.B. and S.M. conducted studies with *APP/PS1* mice for contemporaneous replication<sup>6,29</sup>. K.I. and D.L. performed the *Fshr* mRNA expression studies. D.S., A.P. and S.G. generated FSH-Ab and Hu6. T.-C.K. and S.G. performed antibody distribution studies. P.K. and V.H. conceived and performed the ViewRNA studies with human brain. S.-P.Y. and Z.Z. assisted with data analysis and interpretation. In summary, K.Y.'s laboratory was primarily responsible for generating data in Figs. 1a–e, 2a–c, f, 3, 4 and 5 and Extended Data Figs. 1a–e, 2a–c, f and 3–10. M.Z.'s laboratory produced data in Figs. 1f and 2d–e, g and Extended Data Fig. 1i–p. Moreover, J.I., K.A.G. and C.J.R. assisted with the conception of experiments and manuscript preparation. T.Y., V.M. and V.R. checked image integrity with the help of J.X., S.S.K. and K.Y.; A.G. and T.F. edited and revised the manuscript. T.-C.K. rechecked the raw data files. T.Y., V.R., M.Z. and K.Y. oversaw overall data management and provenance.

## Corresponding authors

Correspondence to [Mone Zaidi](#) or [Keqiang Ye](#).

## Ethics declarations

## Competing interests

M.Z. is listed as an inventor on issued patents on inhibiting FSH for the prevention and treatment of osteoporosis and obesity: US patent numbers 5,436,285 (1995, awarded to Icahn School of Medicine at Mount Sinai (ISMMS)), 5,674,887 (1997, awarded to ISMMS and University of Pittsburgh), 8,435,948 (2013, awarded to ISMMS) and 11,034,761 (2021, awarded to ISMMS). M.Z. is also listed as an inventor on a pending patent application on composition and use of humanized monoclonal anti-FSH antibodies. These patents are owned by ISMMS, and M.Z. would be

recipient of royalties according to institutional policy. M.Z. and K.Y. are listed as inventors of a pending patent application on the use of FSH as a target for preventing Alzheimer's disease. The latter patent is jointly owned by ISMMS and Emory University, and M.Z. and K.Y. would be recipient of royalties according to institutional policy. The other authors declare no competing interests.

## Peer review

### Peer review information

*Nature* thanks Li Gan and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

## Additional information

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Extended data figures and tables

### [Extended Data Fig. 1 Effects of Anti-FSH \$\beta\$ Antibody in Reversing Ovariectomy-Induced Neuropathology in 3xTg Mice.](#)

**a**, Ovariectomized 3xTg mice displayed hypoplastic thread-like uteri and elevated serum FSH and LH levels. FSH-Ab (200  $\mu$ g/mouse, every 2 days, i.p., 8 weeks) did not alter total serum FSH, LH or 17 $\beta$ -estradiol levels. Statistics: mean  $\pm$  s.e.m.,  $N=8$  mice per group, one-way ANOVA. **b**, Immunofluorescent micrographs showing enhanced labelling in the following pairs: amyloid  $\beta$  (A $\beta$ , red) and thioflavin-S (Thio-S, green); A $\beta$  (red) and cleaved APP $C^{586}$  (green); and pTau (red) and Tau $^{1-368}$  (green) in the hippocampus after OVX, and its amelioration with FSH-Ab (scale bar, 20  $\mu$ m). **c**, Upregulation of *Cebpb*, *Lgnm*, *App* and *Mapt* in OVX mouse brains, with reversal to near-baseline with FSH-Ab. Statistics: mean  $\pm$

s.e.m.,  $N=3$  mice *per* group, one-way AVOVA. **d**, Immunofluorescence micrographs showing that OVX induces apoptosis (TUNEL, green) in hippocampal NeuN-positive neurons (red); this apoptosis is abolished by FSH-Ab (scale bar, 20  $\mu\text{m}$ ). **e**, Golgi staining on brain sections from CA1 region post-OVX shows a substantial reduction in spine numbers, which is corrected with the FSH-Ab (scale bar, 5  $\mu\text{m}$ ). Statistics: mean  $\pm$  s.e.m.,  $N=3$  mice *per* group (10 sections), one-way AVOVA. **f**, Transmission electron micrographic images and quantitative analysis of synapses in hippocampal sections post-OVX treated with IgG or FSH-Ab (scale bar, 1  $\mu\text{m}$ ). Statistics: mean  $\pm$  s.e.m.,  $N=3$  mice *per* group (8 sections), one-way AVOVA. **g**, Morris Water Maze testing shows no differences in swim speed. Statistics: mean  $\pm$  s.e.m.,  $N=9$  mice *per* group, one-way ANOVA. **h**, Cognitive testing using the Novel Object Recognition test revealed the absence of significant difference between *APP/PS1* and non-transgenic mice in Discrimination Index [(Novel Object Head Entry – Familiar Object Head Entry)/Total Head Entry]; the result is expected at 9 months of age in *APP/PS1* mice. Thus, no effect of FSH-Ab was noted at this age, despite the reduction in A $\beta$ 40 and A $\beta$ 42 accumulation shown in Fig. 1f. Statistics: mean  $\pm$  s.e.m., mice *per* group 9, 10, 4 and 4 from left to right; Whisker plot, upper and lower ends of the whiskers show maxima and minima, line in box shows median, and upper and lower box boundaries show 75<sup>th</sup> and 25<sup>th</sup> percentile, respectively; unpaired two-tailed Student's *t*-test; **i**, ELISA showing no cross-reactivity of FSH-Ab with LH. **j**, Western immunoblot showing no change in expression of the LHCGR in whole brain lysates upon OVX or FSH-Ab treatment ( $N=2$  mice *per* group). **k**, IVIS imaging of isolated tissues from mice injected with AlexaFluor750–FSH, i.v., showing localization of FSH in the brain ( $N=3$  mice *per* group). **l**, Western immunoblots of whole brain lysates showing that i.p. injection of human FSH (5 IU) causes an elevation of brain FSH ( $N=3$  mice *per* group). **m**, Immunofluorescence micrographs showing the detection of peripherally injected (i.p.) biotinylated FSH-Ab (red) and biotinylated goat IgG (red) in brain sections (scale bar, 20  $\mu\text{m}$ ). Note the absence of cellular or nuclear co-localization with MAP2 or DAPI, respectively. **n**, Representative PET image shows that  $^{89}\text{Zr}$ -labelled humanized monoclonal FSH-Ab ( $^{89}\text{Zr}$ -Hu6), injected i.v., is localized to live brain (arrows).  $\gamma$ -counting in perfused tissue shows presence of  $^{89}\text{Zr}$ -Hu6 in dissected brain tissue at 24 and 48 h post-injection ( $N=4$  mice). **o**, IVIS imaging and quantitation with

AlexaFluor750-labelled Hu6, given i.v. shows localization in perfused whole brain tissue;  $N=3$  mice per group. Control (Ctrl): phosphate-buffered saline (PBS). **p**, Confirmatory immunofluorescence on the same mice (**o**) using anti-human IgG showing Hu6 localization (red) in proximity to CD31<sup>+</sup> endothelial cells (green) (scale bar, 100  $\mu$ m). For gel source data, see Supplementary Fig. 1.

[Source data](#)

## **Extended Data Fig. 2 FSHR Activation Triggers Amyloidogenic Protein Accumulation.**

**a**, Western immunoblots showing the effect of activating neuronal FSHRs by FSH (30 ng/mL) in human SH-SY5Y and primary rat neuronal cells on the expression of C/EBP $\beta$ , AEP, as well as the cleavage of amyloid precursor protein (APP) and Tau using antibodies noted in ‘Methods’. FSH (30 ng/mL) likewise stimulated the expression of *CEBPB*, *LGMN*, *APP* and *MAPT* (qPCR) (**b**); AEP activity (**c**); and certain inflammatory cytokines (ELISA), namely IL-6 and IL-1 $\beta$  (**d**). Statistics: mean  $\pm$  s.e.m.; Mice per group, (**b**) 3, (**c**) 6, and (**d**) 6; one-way ANOVA. **e**, Western immunoblotting showing C/EBP $\beta$ , AEP, APP, cleaved APP $^{1-585}$ , total Tau, and cleaved Tau $^{1-368}$  in response to FSH or PBS following transfection with human *FSHR* siRNA (*si-FSHR*) for SH-SY5Y cells or rat *Fshr* siRNA (*si-Fshr*) for primary rat neurons, or appropriate scrambled siRNAs. **f**, mRNA levels of *CEPB*, *LGMN*, *APP* and *MAPT* in SH-SY5Y cells incubated with FSH after control or *si-FSHR* transfection. **g**, AEP activity after incubation with FSH in control or *si-FSHR*-transfected SH-SY5Y cells. **h**, IL-6 and IL-1 $\beta$  levels (ELISAs) in SH-SY5Y cells incubated with FSH following control or *si-FSHR* infection. Statistics: mean  $\pm$  s.e.m.; (**f**) 3 biological replicates; (**g**, **h**) 6 mice per group; one-way ANOVA.

[Source data](#)

## **Extended Data Fig. 3 FSH Induces APP and Tau Cleavage Through C/EBP $\beta$ and AEP/ $\delta$ -Secretase Activation in Human SH-SY5Y Cells and Rat Cortical Neurons.**

**a**, Western immunoblots showing the effect of FSH (30 ng/mL) on Tau, APP, AEP and FSHR of knocking down C/EBP $\beta$  expression by lentiviral infection with shRNA-*Cebpb* (sh-*Cebpb*) or reducing  $\delta$ -secretase activity by adeno-associated virus infection of AEP<sup>C189S</sup> in both human SH-SY5Y cells and rat cortical neurons. The stimulatory action of FSH was reversed at 48 h. **b, c**, Effect of FSH (30 ng/mL) on APP, APP<sup>C586</sup>, pTau and Tau<sup>1–368</sup> accumulation (immunofluorescence, scale bar, 40  $\mu$ m, **b**) and AEP activity (**c**) in rat cortical neurons infected with sh-*Cebpb* or AAV-AEP<sup>C189S</sup>. Statistics: (**c**) mean  $\pm$  s.e.m.;  $N=6$  mice *per* group; one-way ANOVA. **d**, Western immunoblots showing the time course of FSH effects on C/EBP $\beta$ , phosphorylated C/EBP $\beta$  (pC/EBP $\beta$ ), AEP, pAEP<sup>S226</sup>, total AKT, pAKT<sup>S473</sup>, total ERK1/2, pERK1/2, total SRPK2, pSRPK2<sup>T492</sup> and pNF $\kappa$ B-p65. **e**, Western immunoblots showing the effect of a 30-minute incubation with FSH (30 ng/mL) on levels of C/EBP $\beta$ , AEP, pAEPS<sup>S226</sup>, total AKT and pAKT<sup>S473</sup>, total ERK1/2 and pERK1/2, total SRPK2 and pSRPK2<sup>T492</sup> in the presence or absence of the cAMP inhibitor SQ22536 (100  $\mu$ M), G $\alpha_i$  inhibitor pertussis toxin (PTX, 50 ng/ml), AKTi-1/2 inhibitor (10  $\mu$ M) and ERK1/2 inhibitor PD98059 (10  $\mu$ M).

[Source data](#)

## Extended Data Fig. 4 Targeted Knockdown of *Fshr* in the Hippocampus Diminishes AD Pathologies.

**a**, Quantitative PCR shows significantly reduced expression of *Fshr*, *Cebpb*, *Lgmn*, *App* and *Mapt*. **b**, Immunohistochemistry of the hippocampus shows reduced accumulation of A $\beta$  and pTau, as well as of proteinaceous deposits (silver staining) in si-*Fshr*-injected OVX mice (scale bar, 50  $\mu$ m). **c**, The two isoforms of A $\beta$ , namely A $\beta$ 40 and A $\beta$ 42, were also reduced. **d**, Notable is the marked increase in dendritic spines (Golgi staining) (scale bar, 5  $\mu$ m). Statistics: mean  $\pm$  s.e.m., (**a**) 3 biological replicates; (**c**) 5 mice *per* group; (**d**) 10 sections from 3 mice *per* group; unpaired two-tailed Student's *t*-test.

[Source data](#)

## Extended Data Fig. 5 Recombinant FSH Triggers AD Pathology in 3xTg Mice.

**a**, Serum FSH levels—both mouse (endogenous) and human (exogenous)—24 h after i.p. injection of 2, 5 or 10 IU human recombinant FSH. **b**, Serum LH levels also shown. Female 3xTg mice were injected with recombinant FSH (5 IU *per* mouse, daily, i.p., 3 months). **c**, Immunohistochemistry for A $\beta$  or pTau in hippocampus post-FSH injection (scale bar, 50  $\mu$ m). **d**, Silver staining of the prefrontal cortex, and hippocampal CA1 and dentate gyrus (DG) regions showing enhanced proteinaceous deposits in FSH-injected mice (scale bar, 50  $\mu$ m). **e**, Brain mRNA levels of *Cebpb*, *Lgmn*, *App* and *Mapt*. **f**, Golgi staining of brain sections from the CA1 region shows reduced spine numbers in FSH-injected mice (scale bar, 5  $\mu$ m). **g**, Transmission electron micrographs of hippocampal sections showing reduced synapse numbers post-FSH (scale bar, 1  $\mu$ m). Immunofluorescence micrographs showing the following image pairs in the hippocampus and/or cortex post-FSH: (**h**) A $\beta$  (red) and cleaved APP<sup>C586</sup> (green); (**i**) pTau (red) and cleaved Tau<sup>1–368</sup> (green); (**j**) A $\beta$  (red) and thioflavin-S (green); and (**k**) NeuN (red) and TUNEL (green) (scale bar, 20  $\mu$ m). **l**, Immunofluorescence showing co-localization of C/EBP $\beta$ , AEP, A $\beta$  and pTau to NeuN-positive neurons upon FSH stimulation [10x (scale bar, 300  $\mu$ m) and 40x (scale bar, 50  $\mu$ m) magnifications]. Statistics: mean  $\pm$  s.e.m., (**a**, **b**) 3 mice *per* group; (**e**) 3 biological replicates, (**f**, **g**) 10 sections from 3 mice *per* group; unpaired two-tailed Student's *t*-test.

[Source data](#)

## Extended Data Fig. 6 Effect of Recombinant FSH in Triggering AD Pathology in Ovariectomized 3xTg Mice With Oestrogen Replacement.

3xTg mice were ovariectomized at 3 months and supplemented with 17 $\beta$ -estradiol using 90-day-release pellets (E2, 0.36 mg) to render them biochemically eugonadal. The mice were randomly divided to be injected with PBS or recombinant human FSH (5 IU *per* mouse, daily, i.p., 3 months). **a**, Serum level of FSH and 17 $\beta$ -estradiol. **b**, Western

immunoblotting showing increased C/EBP $\beta$ , AEP, cleaved APP $^{1-373}$  and APP $^{1-585}$ , total Tau, cleaved Tau $^{1-368}$  and pTau in the brain after FSH injection. **c**, Brain AEP enzymatic activity also shown. **d**, Immunohistochemistry of the hippocampus shows increased expression of A $\beta$  and pTau in the FSH group. Silver staining showed increased proteinaceous deposits in FSH-treated mice (scale bar, 50  $\mu$ m). Statistics: mean  $\pm$  s.e.m., mice *per* group; **(a)** 4 and **(c)** 5; unpaired two-tailed Student's *t*-test.

### Source data

### Extended Data Fig. 7 Effect of Recombinant FSH in Triggering AD Pathology and Cognitive Decline in Male Mice.

Male 3xTg mice were injected with recombinant FSH at 5 IU *per* mouse daily, i.p. for 3 months. **a**, Western immunoblots showing increased C/EBP $\beta$ , AEP, cleaved APP $^{1-373}$  and APP $^{1-585}$ , total Tau, cleaved Tau $^{1-368}$  and pTau in the brain (3 mice *per* group). **b**, **c**, Brain AEP activity (**b**) and A $\beta$  isoforms, A $\beta$ 40 and A $\beta$ 42 (**c**) were also increased with FSH. **d**, Morris Water Maze test shows enhanced escape latency to mount the platform (seconds). Also shown are integrated escape latency (area under the curve, AUC) and percentage of time spent in the target quadrant (Probe Trial Test). **e**, Silver staining of the prefrontal cortex, and hippocampus CA1 and dentate gyrus (DG) regions showing enhanced proteinaceous deposits in FSH-injected mice (scale bar, 50  $\mu$ m). **f**, Immunohistochemistry for A $\beta$  or pTau in the hippocampus post-FSH injection (scale bar, 50  $\mu$ m). **g**, Brain mRNA levels of *Cebpb*, *Lgmn*, *App* and *Mapt*. **h**, Golgi staining of brain sections from the CA1 region of the hippocampus showing reduced spine numbers in FSH-injected mice (scale bar, 5  $\mu$ m). **i**, Transmission electron micrographs of hippocampal sections showing reduced synapse numbers post-FSH (scale bar, 1  $\mu$ m). Immunofluorescence micrographs showing the following image pairs: **(j)** A $\beta$  (red) and cleaved APP $^{C586}$  (green); **(k)** pTau (red) and cleaved Tau $^{1-368}$  (green); **(l)** A $\beta$  (red) and thioflavin-S (green); and **(m)** NeuN (red) and TUNEL (green) in the hippocampus and/or cortex of male 3xTg mice after FSH (scale bar, 20  $\mu$ m). Statistics: mean  $\pm$  s.e.m.,

mice *per* group, (b, c) 5, (d) 7, (g) 3, (h, i), 3 (10 sections); unpaired two-tailed Student's *t*-test.

[Source data](#)

**Extended Data Fig. 8 Effect of FSH in Triggering AD Pathology and Cognitive Decline in Female Wild Type and APP-KI Mice.**

In *APP-KI* mice, three amino acid substitutions (G601R;F606Y;R609H) are knocked into A $\beta$ -coding exon 14 of the *APP* gene—this results in the *non-transgenic* expression at basal levels of oligomerizable *human* A $\beta$ . Female wild type and *APP-KI* mice were injected with recombinant FSH (5 IU, daily, i.p. 3 months). **a–d**, In wild type mice, Western immunoblotting showing increased C/EBP $\beta$ , AEP, cleaved APP<sup>1–373</sup> and APP<sup>1–585</sup>, total Tau, and cleaved Tau<sup>1–368</sup> in whole brain (3 mice *per* group) (a), as well as increased silver staining (b), elevated AEP activity (c) and increases A $\beta$  isoforms, A $\beta$ 40 and A $\beta$ 42 (d) upon FSH treatment. **e**, Morris Water Maze test, however, showed no difference in escape latency to mount the platform (seconds). Also shown are no differences in integrated escape latency (area under the curve, AUC) and percentage of time spent in the target quadrant (Probe Trial Test). **f–l**, Western immunoblotting showing elevations in C/EBP $\beta$ , AEP, cleaved APP<sup>1–373</sup> and APP<sup>1–585</sup>, and cleaved Tau<sup>1–368</sup> in whole brain (f), along with enhancements in silver staining (g), AEP activity (h), and A $\beta$  isoforms (i) in *APP-KI* mice in response to FSH injection. **j**, There was also a significant spatial memory deficit on the Morris Water Maze test. **k**, Immunofluorescence micrographs showed increases in the hippocampus and/or cortex of female *APP-KI* mice post-FSH injected in the following pairs: A $\beta$  (red) and cleaved APP<sup>C586</sup> (green); A $\beta$  (red) and thioflavin-S (green); and NeuN (red) and TUNEL (green). Scale bar: (b, g) 50  $\mu$ m, (k) 100  $\mu$ m (magnified view, 10  $\mu$ m). Statistics: mean  $\pm$  s.e.m.; mice *per* group, (c, d, h, i) 5, (e, j) 8; unpaired two-tailed Student's *t*-test.

[Source data](#)

## Extended Data Fig. 9 C/EBP $\beta$ Mediates FSH-Induced AD Neuropathology and Cognitive Decline in 3xTg Mice.

**a**, *Cebpb*, *Lgmn*, *App* and *Mapt* mRNA expression following FSH injection to 3xTg or *Cebpb*<sup>+/−</sup> 3xTg mice. Statistics: mean\(\pm\)\ s.e.m.,  $N=3$  biological replicates, one-way ANOVA. **b**, Immunohistochemistry for A $\beta$  and pTau and silver staining for proteinaceous deposits (scale bar, 50  $\mu$ m). **c–e**, Immunofluorescence staining for A $\beta$  (red) and C/EBP $\beta$  (green) (**c**) and for pTau (red) and C/EBP $\beta$  (green) (**d**) (scale bar, 20  $\mu$ m), and Golgi staining for dendritic spines (**e**) in the hippocampus in female 3xTg or *Cebpb*<sup>+/−</sup> 3xTg mice, post-FSH (scale bar, 5  $\mu$ m). Statistics: mean  $\pm$  s.e.m., 10 sections from 3 mice *per* group, one-way ANOVA. **f**, Morris Water Maze testing showed no difference in swim speed. Statistics: 9 mice *per* group, one-way ANOVA.

[Source data](#)

## Extended Data Fig. 10 C/EBP $\beta$ Mediates Ovariectomy-Induced AD Neuropathology and Cognitive Decline in 3xTg Mice.

**a**, *Cebpb*, *Lgmn*, *App* and *Mapt* mRNA expression following ovariectomy of 3xTg or *Cebpb*<sup>+/−</sup> 3xTg mice. Statistics: mean  $\pm$  s.e.m., 3 biological replicates, one-way ANOVA. **b**, Immunohistochemistry for A $\beta$  and pTau and silver staining for proteinaceous deposits (scale bar, 50  $\mu$ m). **c–e**, Immunofluorescence staining for A $\beta$  (red) and C/EBP $\beta$  (green) (**c**) and for pTau (red) and C/EBP $\beta$  (green) (**d**) (scale bar, 20  $\mu$ m), and Golgi staining for dendritic spines (**e**) in the hippocampus in female 3xTg or *Cebpb*<sup>+/−</sup> 3xTg mice (scale bar, 5  $\mu$ m), post-OVX. Statistics: mean  $\pm$  s.e.m., 10 sections from 3 mice *per* group, one-way ANOVA. **f**, Morris Water Maze test showed no difference in the swim speed between 3xTg and *Cebpb*<sup>+/−</sup> 3xTg mice. Statistics: left to right: 7, 8, 8 mice *per* group, one-way ANOVA.

[Source data](#)

## Supplementary information

## **Supplementary Information**

Supplementary Fig. 1: uncropped agarose gels and western blots.

Supplementary Table 1: information about the key antibodies used in this study.

## **Reporting Summary**

### **Supplementary Video 1**

PET scan of a mouse injected with  $^{89}\text{Zr}$ -labelled FSH antibody. Four C57BL/6 mice were injected through the tail vein with 100  $\mu\text{Ci}$  of the  $^{89}\text{Zr}$ -radiolabelled monoclonal FSH antibody Hu6. After 24 h, the mice were anaesthetized and then imaged using a Mediso NanoScan PET/CT scanner. In this representative video,  $^{89}\text{Zr}$ -labelled Hu6 was noted in the brain. These data provide compelling evidence that the FSH antibody can cross the blood–brain barrier.

## **Source data**

### **Source Data Fig. 1**

### **Source Data Fig. 2**

### **Source Data Fig. 3**

### **Source Data Fig. 4**

### **Source Data Fig. 5**

### **Source Data Extended Data Fig. 1**

### **Source Data Extended Data Fig. 2**

### **Source Data Extended Data Fig. 3**

[\*\*Source Data Extended Data Fig. 4\*\*](#)

[\*\*Source Data Extended Data Fig. 5\*\*](#)

[\*\*Source Data Extended Data Fig. 6\*\*](#)

[\*\*Source Data Extended Data Fig. 7\*\*](#)

[\*\*Source Data Extended Data Fig. 8\*\*](#)

[\*\*Source Data Extended Data Fig. 9\*\*](#)

[\*\*Source Data Extended Data Fig. 10\*\*](#)

## Rights and permissions

[Reprints and Permissions](#)

## About this article

### Cite this article

Xiong, J., Kang, S.S., Wang, Z. *et al.* FSH blockade improves cognition in mice with Alzheimer's disease. *Nature* **603**, 470–476 (2022).  
<https://doi.org/10.1038/s41586-022-04463-0>

- Received: 16 July 2020
- Accepted: 25 January 2022
- Published: 02 March 2022
- Issue Date: 17 March 2022
- DOI: <https://doi.org/10.1038/s41586-022-04463-0>

## Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

## Further reading

- [\*\*FSH provides link between menopause and AD\*\*](#)
  - Sarah Lemprière  
*Nature Reviews Neurology* (2022)
  - M. Teresa Villanueva  
*Nature Reviews Drug Discovery* (2022)
- [\*\*Hooking FSH as a potential target for Alzheimer disease\*\*](#)

## [\*\*A pituitary hormone has a key role in Alzheimer's disease\*\*](#)

Research Briefing 02 Mar 2022

## [\*\*FSH provides link between menopause and AD\*\*](#)

- Sarah Lemprière

Research Highlight 15 Mar 2022

| [Section menu](#) | [Main menu](#) |

- Article
- [Published: 09 March 2022](#)

# A non-canonical tricarboxylic acid cycle underlies cellular identity

- [Paige K. Arnold](#) ORCID: [orcid.org/0000-0002-0851-2057](https://orcid.org/0000-0002-0851-2057)<sup>1,2</sup> na1,
- [Benjamin T. Jackson](#) ORCID: [orcid.org/0000-0003-1089-4362](https://orcid.org/0000-0003-1089-4362)<sup>1,2</sup> na1,
- [Katrina I. Paras](#) ORCID: [orcid.org/0000-0001-9743-5710](https://orcid.org/0000-0001-9743-5710)<sup>1,3</sup>,
- [Julia S. Brunner](#) ORCID: [orcid.org/0000-0002-3195-147X](https://orcid.org/0000-0002-3195-147X)<sup>1</sup>,
- [Madeleine L. Hart](#) ORCID: [orcid.org/0000-0001-9125-0627](https://orcid.org/0000-0001-9125-0627)<sup>4</sup>,
- [Oliver J. Newsom](#)<sup>4</sup>,
- [Sydney P. Alibekoff](#)<sup>4</sup>,
- [Jennifer Endress](#) ORCID: [orcid.org/0000-0002-7538-6302](https://orcid.org/0000-0002-7538-6302)<sup>1,3</sup>,
- [Esther Drill](#) ORCID: [orcid.org/0000-0002-3315-4538](https://orcid.org/0000-0002-3315-4538)<sup>5</sup>,
- [Lucas B. Sullivan](#) ORCID: [orcid.org/0000-0002-6745-8222](https://orcid.org/0000-0002-6745-8222)<sup>4</sup> &
- [Lydia W. S. Finley](#) ORCID: [orcid.org/0000-0003-4023-7574](https://orcid.org/0000-0003-4023-7574)<sup>1</sup>

*Nature* volume 603, pages 477–481 (2022)

- 19k Accesses
- 464 Altmetric
- [Metrics details](#)

## Subjects

- [Cancer metabolism](#)
- [Cell biology](#)
- [Cell growth](#)
- [Embryonic stem cells](#)

- [Metabolomics](#)

## Abstract

The tricarboxylic acid (TCA) cycle is a central hub of cellular metabolism, oxidizing nutrients to generate reducing equivalents for energy production and critical metabolites for biosynthetic reactions. Despite the importance of the products of the TCA cycle for cell viability and proliferation, mammalian cells display diversity in TCA-cycle activity<sup>1,2</sup>. How this diversity is achieved, and whether it is critical for establishing cell fate, remains poorly understood. Here we identify a non-canonical TCA cycle that is required for changes in cell state. Genetic co-essentiality mapping revealed a cluster of genes that is sufficient to compose a biochemical alternative to the canonical TCA cycle, wherein mitochondrially derived citrate exported to the cytoplasm is metabolized by ATP citrate lyase, ultimately regenerating mitochondrial oxaloacetate to complete this non-canonical TCA cycle. Manipulating the expression of ATP citrate lyase or the canonical TCA-cycle enzyme aconitase 2 in mouse myoblasts and embryonic stem cells revealed that changes in the configuration of the TCA cycle accompany cell fate transitions. During exit from pluripotency, embryonic stem cells switch from canonical to non-canonical TCA-cycle metabolism. Accordingly, blocking the non-canonical TCA cycle prevents cells from exiting pluripotency. These results establish a context-dependent alternative to the traditional TCA cycle and reveal that appropriate TCA-cycle engagement is required for changes in cell state.

This is a preview of subscription content

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

## Subscribe

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

## Subscribe

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

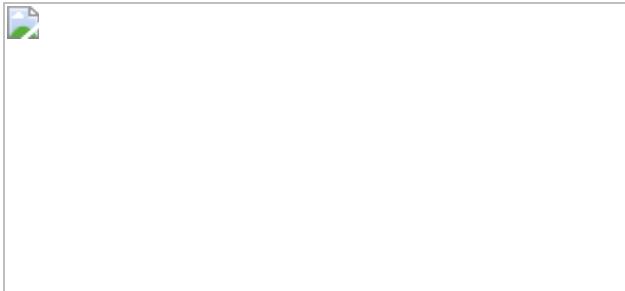
## Buy

All prices are NET prices.

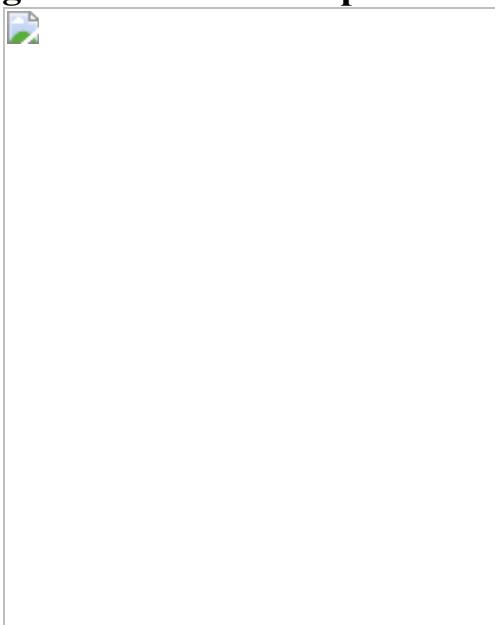
## **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

**Fig. 1: Genetic co-essentiality mapping of metabolic enzymes reveals two TCA-cycle modules.**



**Fig. 2: ACL loss disrupts TCA-cycle metabolism in ES cells.**



**Fig. 3: Engagement of the non-canonical TCA cycle is cell-state dependent.**



**Fig. 4: Exit from naive pluripotency requires engagement of the non-canonical TCA cycle.**



## Data availability

RNA-seq data supporting the findings of this study have been deposited in the Gene Expression Omnibus under the accession code [GSE183434](#).

Alignment was performed against the mouse mm10 genome assembly. Gene essentiality data and NSCLC gene expression data are available from the DepMap Portal (<https://depmap.org/portal/>). Isotopologue distributions from all MS isotope tracing experiments are provided in [Supplementary Table 4. Source data](#) are provided with this paper.

## Code availability

The code used to perform gene essentiality correlation and network modelling is provided at GitHub (<https://github.com/finley-lab/coessentiality-network>).

## References

1. DeBerardinis, R. J. & Chandel, N. S. Fundamentals of cancer metabolism. *Sci. Adv.* **2**, e1600200 (2016).
2. Kim, J. & DeBerardinis, R. J. Mechanisms and implications of metabolic heterogeneity in cancer. *Cell Metab.* **30**, 434–446 (2019).
3. Muir, A., Danai, L. V. & Vander Heiden, M. G. Microenvironmental regulation of cancer cell metabolism: implications for experimental design and translational studies. *Dis. Model. Mech.* **11**, dmm035758 (2018).

4. Tsherniak, A. et al. Defining a cancer dependency map. *Cell* **170**, 564–576 (2017).
5. Wainberg, M. et al. A genome-wide atlas of co-essential modules assigns function to uncharacterized genes. *Nat. Genet.* **53**, 638–649 (2021).
6. Hatzivassiliou, G. et al. ATP citrate lyase inhibition can suppress tumor cell growth. *Cancer Cell* **8**, 311–321 (2005).
7. Zhao, S. et al. ATP-citrate lyase controls a glucose-to-acetate metabolic switch. *Cell Rep.* **17**, 1037–1052 (2016).
8. Chen, P. H. et al. Metabolic diversity in human non-small cell lung cancer cells. *Mol. Cell* **76**, 838–851 (2019).
9. Alves, T. C. et al. Integrated, step-wise, mass-isotopomeric flux analysis of the TCA cycle. *Cell Metab.* **22**, 936–947 (2015).
10. Sullivan, L. B. et al. Aspartate is an endogenous metabolic limitation for tumour growth. *Nat. Cell Biol.* **20**, 782–788 (2018).
11. Lewis, C. A. et al. Tracing compartmentalized NADPH metabolism in the cytosol and mitochondria of mammalian cells. *Mol. Cell* **55**, 253–263 (2014).
12. Williamson, D. H., Lund, P. & Krebs, H. A. The redox state of free nicotinamide-adenine dinucleotide in the cytoplasm and mitochondria of rat liver. *Biochem. J.* **103**, 514–527 (1967).
13. Birsoy, K. et al. An essential role of the mitochondrial electron transport chain in cell proliferation is to enable aspartate synthesis. *Cell* **162**, 540–551 (2015).
14. Krebs, H. A. & Eggleston, L. V. The oxidation of pyruvate in pigeon breast muscle. *Biochem. J.* **34**, 442–459 (1940).
15. Shintaku, J. et al. MyoD regulates skeletal muscle oxidative metabolism cooperatively with alternative NF-κB. *Cell Rep.* **17**, 514–

526 (2016).

16. Stacpoole, P. W. The pharmacology of dichloroacetate. *Metabolism* **38**, 1124–1144 (1989).
17. Kalkan, T. et al. Tracking the embryonic stem cell transition from ground state pluripotency. *Development* **144**, 1221–1234 (2017).
18. Vardhana, S. A. et al. Glutamine independence is a selectable feature of pluripotent stem cells. *Nat. Metab.* **1**, 676–687 (2019).
19. Pereira, L., Yi, F. & Merrill, B. J. Repression of Nanog gene transcription by Tcf3 limits embryonic stem cell self-renewal. *Mol. Cell. Biol.* **26**, 7479–7491 (2006).
20. Wray, J. et al. Inhibition of glycogen synthase kinase-3 alleviates Tcf3 repression of the pluripotency network and increases embryonic stem cell resistance to differentiation. *Nat. Cell Biol.* **13**, 838–845 (2011).
21. Moussaieff, A. et al. Glycolysis-mediated changes in acetyl-CoA and histone acetylation control the early differentiation of embryonic stem cells. *Cell Metab.* **21**, 392–402 (2015).
22. Sullivan, L. B. et al. Supporting aspartate biosynthesis is an essential function of respiration in proliferating cells. *Cell* **162**, 552–563 (2015).
23. Borst, P. The malate-aspartate shuttle (Borst cycle): how it started and developed into a major metabolic pathway. *IUBMB Life* **72**, 2241–2259 (2020).
24. Assmann, N. et al. Srebp-controlled glucose metabolism is essential for NK cell functional responses. *Nat. Immunol.* **18**, 1197–1206 (2017).
25. Luengo, A. et al. Increased demand for NAD<sup>+</sup> relative to ATP drives aerobic glycolysis. *Mol. Cell* **81**, 691–707 (2020).

26. Davidson, S. M. et al. Environment impacts the metabolic dependencies of Ras-driven non-small cell lung cancer. *Cell Metab.* **23**, 517–528 (2016).
27. Marin-Valencia, I. et al. Analysis of tumor metabolism reveals mitochondrial glucose oxidation in genetically diverse human glioblastomas in the mouse brain *in vivo*. *Cell Metab.* **15**, 827–837 (2012).
28. Zhu, X. G. et al. Functional genomics *in vivo* reveal metabolic dependencies of pancreatic cancer cells. *Cell Metab.* **33**, 211–221 (2021).
29. Dempster, J. M. et al. Extracting biological insights from the project achilles genome-scale CRISPR screens in cancer cell lines. Preprint at <https://doi.org/10.1101/720243> (2019).
30. Meyers, R. M. et al. Computational correction of copy number effect improves specificity of CRISPR-Cas9 essentiality screens in cancer cells. *Nat. Genet.* **49**, 1779–1784 (2017).
31. Ashburner, M. et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25–29 (2000).
32. The Gene Ontology Consortium. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.* **47**, D330–D338 (2019).
33. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).
34. Ghandi, M. et al. Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* **569**, 503–508 (2019).
35. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl Acad. Sci. USA* **102**, 15545–15550 (2005).

36. Carey, B. W., Finley, L. W., Cross, J. R., Allis, C. D. & Thompson, C. B. Intracellular  $\alpha$ -ketoglutarate maintains the pluripotency of embryonic stem cells. *Nature* **518**, 413–416 (2015).
37. Dow, L. E. et al. Inducible in vivo genome editing with CRISPR-Cas9. *Nat. Biotechnol.* **33**, 390–394 (2015).
38. Ran, F. A. et al. Genome engineering using the CRISPR-Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).
39. MacDougall, M. S., Clarke, R. & Merrill, B. J. Intracellular  $\text{Ca}^{2+}$  homeostasis and nuclear export mediate exit from naive pluripotency. *Cell Stem Cell* **25**, 210–224 (2019).
40. Morris, J. P. T. et al.  $\alpha$ -Ketoglutarate links p53 to cell fate during tumour suppression. *Nature* **573**, 595–599 (2019).
41. Fellmann, C. et al. An optimized microRNA backbone for effective single-copy RNAi. *Cell Rep.* **5**, 1704–1713 (2013).
42. Millard, P. et al. IsoCor: isotope correction for high-resolution MS labeling experiments. *Bioinformatics* **35**, 4484–4487 (2019).
43. Chen, S., Zhou, Y., Chen, Y. & Gu, J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).
44. Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
45. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
46. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).

## Acknowledgements

We thank the members of the Finley laboratory for discussion, and A. Boire, A. Intlekofer, S. Vardhana, E. Reznik and K. S. Tan for feedback. *Rex1::GFPd2* cells were a gift from A. Smith. P.K.A. is an NICHD Ruth L. Kirschstein Predoctoral fellow (F31HD098824). B.T.J. is a Gerstner Sloan Kettering Grayer Fellow and is supported by a Medical Scientist Training Program grant from the NIGMS of the National Institutes of Health under award number T32GM007739 to the Weill Cornell/Rockefeller/Sloan Kettering Tri-Institutional MD-PhD Program. K.I.P. is supported by a Bruce Charles Forbes Pre-Doctoral Fellowship (MSKCC). J.S.B. is supported by a Human Frontier Science Program Fellowship (LT000200/2021-L). L.B.S. is supported by R00CA218679 and P30CA015704. L.W.S.F. is a Searle Scholar. This research was also supported by grants to L.W.S.F. from the Pershing Square Sohn Prize for Cancer Research, the Starr Foundation (I12-0051), the NIH/NCI (R37 CA252305) and the Anna Fuller Fund as well as the Memorial Sloan Kettering Cancer Center Support Grant P30CA008748.

## Author information

### Author notes

1. These authors contributed equally: Paige K. Arnold, Benjamin T. Jackson

## Affiliations

1. Cell Biology Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA

Paige K. Arnold, Benjamin T. Jackson, Katrina I. Paras, Julia S. Brunner, Jennifer Endress & Lydia W. S. Finley

2. Louis V. Gerstner Jr. Graduate School of Biomedical Sciences, New York, NY, USA

Paige K. Arnold & Benjamin T. Jackson

3. Weill Cornell Graduate School of Medical Sciences, Cornell University, New York, NY, USA

Katrina I. Paras & Jennifer Endress

4. Human Biology Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

Madeleine L. Hart, Oliver J. Newsom, Sydney P. Alibekoff & Lucas B. Sullivan

5. Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, NY, USA

Esther Drill

## Contributions

P.K.A. and L.W.S.F. conceived the study. P.K.A. and B.T.J. performed all of the experiments with assistance from K.I.P., J.S.B. and J.E.; M.L.H., O.J.N. and S.P.A. performed LC–MS experiments under the guidance of L.B.S.; B.T.J. performed genetic co-essentiality mapping and network modelling with guidance from E.D. K.I.P. performed RNA-seq. L.B.S. provided additional study guidance. L.W.S.F. supervised the project. P.K.A. and L.W.S.F. wrote the manuscript with input from all of the authors.

## Corresponding author

Correspondence to [Lydia W. S. Finley](#).

## Ethics declarations

## Competing interests

P.K.A., B.T.J. and L.W.S.F. are listed as inventors on a provisional patent application (US provisional application no. 63/272,940) filed by the Memorial Sloan Kettering Cancer Center. The patent application covers the

use of ACL inhibitors to modify the self-renewal potential of ES cells. The other authors declare no competing interests.

## Peer review

### Peer review information

*Nature* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

## Additional information

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Extended data figures and tables

### [Extended Data Fig. 1 Metabolic gene essentiality correlations across cancer cell lines.](#)

Heatmap depicting hierarchical clustering of pairwise gene essentiality score correlations of core metabolic pathway genes derived from four GO terms: tricarboxylic acid (TCA) cycle, canonical glycolysis, one-carbon metabolic process and fatty-acyl-CoA metabolic process. Genes are colour coded to the left of the heatmap according to the GO term. TCA cycle genes are highlighted (red) in the dendrogram. Gene names and correlation scores can be found in [Supplementary Table 1](#).

### [Extended Data Fig. 2 Effect of ACL inhibition on <sup>13</sup>C labelling of TCA cycle metabolites.](#)

**a**, Two-dimensional network diagram representing gene essentiality score correlations between TCA cycle genes and their top co-dependencies. The strength of the correlation between genes is represented by both the length and thickness of the connecting edge. Correlation scores are shown

in [Supplementary Table 1](#). **b, c**, Fractional enrichment of citrate (left) and malate (right) in three NSCLC cell lines cultured in medium containing [ $U-^{13}C$ ]glucose (**b**) or [ $U-^{13}C$ ]glutamine (**c**) and treated with vehicle or 50  $\mu M$  BMS-303141(ACLi) for 24 h. **d**, Schematic depicting [ $U-^{13}C$ ]asparagine labelling of aspartate and citrate in cells expressing guinea pig asparaginase (ASNase). Asparagine-derived aspartate will generate M+4-labelled citrate. Top, M+4-labelled citrate metabolized via the canonical TCA cycle will lose two labelled carbons as  $CO_2$ , ultimately regenerating citrate that retains two labelled carbons (M+2). Bottom, M+4-labelled citrate metabolized by ACL will yield M+4 labelled oxaloacetate that will ultimately regenerate M+4-labelled citrate. **e, f**, Fractional labelling of aspartate (left) and citrate (right) (**e**) or citrate M+2 relative to citrate M+4 (cit+2/cit+4) (**f**) in ASNase-expressing 143B human osteosarcoma cells cultured in medium containing [ $U-^{13}C$ ]asparagine and treated with vehicle or 50  $\mu M$  ACLi for 24 h. Data are mean  $\pm$  SD,  $n = 3$  independent replicates. Significance was assessed in comparison to vehicle treatment by two-way ANOVA with Sidak's multiple comparisons post-test (**b-c, e**) or using unpaired two-tailed Student's *t*-test (**f**).

#### [Source data](#)

#### [Extended Data Fig. 3 ACO2 and ACL disruption in embryonic stem cells.](#)

**a**, Fractional M+2 enrichment of citrate and malate in mouse ES cells cultured in medium containing [ $U-^{13}C$ ]glucose. **b**, Fractional enrichment of malate M+2 relative to citrate M+2 (mal+2/cit+2) derived from [ $U-^{13}C$ ]glucose in ES cells following treatment with vehicle or 50  $\mu M$  BMS-303141 (ACLi) for 24 h. **c, d**, Immunoblot of clonal mouse ES cells in which CRISPR/Cas9-mediated editing was used to target either a non-genic region of chromosome 8 (Ctrl) and *Acly* (ACLY-1 and ACLY-2) (**c**) or *Aco2* (ACO2-1 and ACO2-2) (**d**). **e, f**, Assessment of the [ $U-^{13}C$ ]glucose-derived mal+2/cit+2 ratio (**e**) or steady-state levels of TCA cycle metabolites represented as the fold change (expressed in  $\log_2$ ) relative to Ctrl (**f**) in control and *Aco2*-edited ES cells. **g, h, j, k**, Fractional M+1 enrichment of NADH (**g**), lactate (**h**), fumarate (**j**) and succinate (**k**) in control and *Acly*-

edited ES cells cultured in medium containing [ $4\text{-}^2\text{H}$ ]glucose. **i**, Schematic depicting deuterium transfer from [ $4\text{-}^2\text{H}$ ]glucose first onto malate in the cytoplasm then onto TCA cycle metabolites in the mitochondria. **l**, Quantification of the lactate over pyruvate ratio in control and *Acly*-edited ES cells. **m**, The baseline oxygen consumption rate (OCR) in control and *Acly*-edited ES cells normalized to protein content. Twelve technical replicates were averaged for each of three independent experiments. Data are mean  $\pm$  SD,  $n = 3$  independent replicates unless otherwise noted. Significance was assessed using unpaired two-tailed Student's *t*-test (**a, b**) or in comparison to control cells by one-way ANOVA with Sidak's multiple comparisons post-test for all other panels.

[Source data](#)

**Extended Data Fig. 4 SLC25A1 and MDH1 contribute to TCA cycle metabolism in embryonic stem cells.**

**a, b**, Immunoblot of clonal mouse ES cells in which CRISPR/Cas9-mediated editing was used to target either a non-genic region of chromosome 8 (Ctrl) and *Slc25a1* (SLC25A1-1 and SLC25A1-2) (**a**) or *Mdh1* (MDH1-1 and MDH1-2) (**b**). **c, d**, Fractional M+1 enrichment of malate (Mal), fumarate (Fum), aspartate (Asp) and citrate (Cit) in control (Ctrl) and *Slc25a1*-edited ES cells (**c**) or *Mdh1*-edited ES cells (**d**) cultured in medium containing [ $4\text{-}^2\text{H}$ ]glucose. **e, f**, Fractional M+2 enrichment of citrate, fumarate, malate and aspartate derived from [ $\text{U-}^{13}\text{C}$ ]glucose in control and *Slc25a1*-edited (**e**) or *Mdh1*-edited (**f**) ES cells. **g**, Steady-state levels of TCA cycle metabolites in *Slc25a1*-edited or *Mdh1*-edited ES cells. Levels are represented as the fold change (expressed in  $\log_2$ ) relative to chromosome 8-targeted control cells. Data are mean  $\pm$  SD,  $n = 3$  independent replicates. Significance was assessed in comparison to control cells by two-way ANOVA (**c-f**) with Sidak's multiple comparisons post-test.

[Source data](#)

## Extended Data Fig. 5 Effect of myogenic differentiation on <sup>13</sup>C-glucose labelling of TCA cycle intermediates.

**a**, Immunoblot comparing expression of myogenesis markers MYOG and MYH3 between proliferating (Prolif) and 100% confluent (Conf) myoblasts and myotubes differentiated for 3, 5 or 7 days. **b**, Fractional labelling of citrate (left) and malate (right) in proliferating and confluent myoblasts and myotubes differentiated for 3, 5 or 7 days cultured in medium containing [ $\text{U-}^{13}\text{C}$ ]glucose. **c**, Fractional M+1 enrichment from [ $4\text{-}^2\text{H}$ ]glucose of malate (Mal), fumarate (Fum), aspartate (Asp) and citrate (Cit) in myoblasts and myotubes differentiated for 5 days. **d**, Immunoblot comparing expression of ACL and ACO2 in C2C12 cells expressing doxycycline-inducible shRNAs targeting *Acly* (shAcly-1 and shAcly-2), *Aco2* (shAco2-1 and shAco2-2) or Renilla luciferase (shRen, used as a control). Cells were cultured on doxycycline for two days to induce shRNA expression. **e-h**, Fractional M+2 enrichment of citrate (left) and malate (right) or malate M+2 relative to citrate M+2 (mal+2/cit+2) in myoblasts (**e, f**) or myotubes (**g, h**) expressing doxycycline-inducible shRNAs targeting *Acly*, *Aco2* or Renilla luciferase cultured in medium containing [ $\text{U-}^{13}\text{C}$ ]glucose. Myoblasts and myotubes were cultured on doxycycline for 2 or 4 days, respectively, to induce shRNA expression. Data are mean  $\pm$  SD,  $n = 3$  independent replicates. In **b**, significance was assessed using one-way ANOVA with Sidak's multiple comparisons post-test to compare total metabolite fraction labelled relative to proliferating myoblasts. In remaining panels, significance was assessed by two-way ANOVA in comparison to myoblasts (**c**) or by one-way ANOVA in comparison to shRen-expressing myoblasts (**e-f**) or myotubes (**g-h**) with Sidak's multiple comparisons post-test.

[Source data](#)

## Extended Data Fig. 6 Transcriptional profiles associated with TCA cycle choice.

**a**, Gene set enrichment analysis showing that genes positively correlated with fractional enrichment of malate M+2 relative to citrate M+2

(mal+2/cit+2) derived from [ $^{13}\text{C}$ ]glucose in 68 NSCLC cell lines are enriched for KEGG citric acid (TCA) cycle-associated genes. **b**, **c**, Fractional M+2 enrichment of citrate (Cit), fumarate (Fum), malate (Mal) and aspartate (Asp) (**b**) or mal+2/cit+2 (**c**) derived from [ $^{13}\text{C}$ ]glucose in mouse ES cells following treatment with vehicle, 5 mM DCA or 10  $\mu\text{M}$  MPCi for 24 h. Data are mean  $\pm$  SD,  $n = 3$  independent samples. In **b-c**, significance was assessed in comparison to vehicle treatment by two-way ANOVA (**b**) or one-way ANOVA (**c**) with Sidak's multiple comparisons post-test.

[Source data](#)

[Extended Data Fig. 7 ACL loss blunts exit from naive pluripotency.](#)

**a**, Experimental setup for cell fate transitions. Mouse ES cells cultured in serum and leukemia inhibitory factor (LIF) are a heterogenous population that can be converted to the naive, ground state of pluripotency by addition of MEK and GSK3 $\beta$  inhibitors (2i) in either serum replete (serum/LIF+2i, **d-f**) or serum-free (2i/LIF, **g-i**) media formulations. Transition to serum-free medium lacking 2i/LIF ( $-2i/\text{LIF}$ ) induces exit from the naive pluripotent state, enabling ES cells to gain differentiation competence. **b**, RT-qPCR of pluripotency-associated (*Nanog*, *Esrrb*, *Klf2*, *Rex1* and *Oct4*) and early differentiation-associated (*Fgf5*, *Otx2* and *Sox1*) genes in 2i/LIF-cultured ES cells subjected to 2i/LIF withdrawal for 12, 24 or 40 h. Levels are represented as the fold change (expressed in  $\log_2$ ) relative to naive, 2i/LIF-cultured ES cells (0 h). **c**, Quantification of alkaline phosphatase (AP)-positive colonies representing ES cells that failed to exit from the pluripotent state. 2i/LIF-cultured ES cells were subjected to 2i/LIF withdrawal for 0, 12, 24 or 40 h and then reseeded at clonal density into medium containing 2i and LIF. **d-f**, Fractional labelling of citrate (Cit), malate (Mal) and aspartate (Asp) in serum/LIF+2i-cultured ES cells incubated with [ $^{13}\text{C}$ ]glucose (**d**), [ $^{13}\text{C}$ ]glutamine (**e**) or [ $4\text{-}^2\text{H}$ ]glucose (**f**) subjected to exit from pluripotency for the indicated times. **g**, Fractional enrichment of glucose-derived malate M+2 relative to citrate M+2 (mal+2/cit+2) in 2i/LIF-cultured ES cells subjected to 2i/LIF withdrawal

for the indicated times. **h, i**, Fractional labelling of citrate, malate and aspartate in 2i/LIF-cultured ES cells cultured in medium containing [U-<sup>13</sup>C]glucose (**h**) or [U-<sup>13</sup>C]glutamine (**i**) subjected to 2i/LIF withdrawal for the indicated times. **j**, Immunoblot of polyclonal ES cells in which CRISPR/Cas9-mediated editing was used to target either a non-genic region of chromosome 8 (sgChr8) or *Tcf7l1* (sgTcf7l1). **k**, RT-qPCR of pluripotency-associated (*Nanog*, *Esrrb*, and *Rex1*) and early differentiation-associated (*Sox1*) genes in control and *Tcf7l1*-edited ES cells adapted to the naive, ground state and subjected to 2i/LIF withdrawal for the indicated times. Levels are represented as the fold change (expressed in log<sub>2</sub>) relative to chromosome 8-targeted control cells in the naive, ground state (0 h). **l, m**, Fractional labelling of citrate (left) and malate (right) (**l**) and glucose-derived mal+2/cit+2 ratio (**m**) in chromosome 8-targeted control or *Tcf7l1*-edited ES cells cultured in medium containing [U-<sup>13</sup>C]glucose subjected to 2i/LIF withdrawal for the indicated times. Data are mean ± SD,  $n = 3$  independent replicates. In **d-e**, **h-i**, and **l**, significance was assessed using one-way ANOVA (**d-e**, **h-i**) or two-way ANOVA (**l**) with Sidak's multiple comparisons post-test to compare total metabolite fraction labelled relative to the 0 h timepoint (**d-e**, **h-i**) or control cells (**l**). In remaining panels, significance was assessed relative to the 0 h timepoint using one-way ANOVA (**c, f-g**) or chromosome 8-targeted control cells at each time point using two-way ANOVA (**k, m**) with Sidak's multiple comparisons post-test.

### Source data

### Extended Data Fig. 8 Acetate does not reverse the effects of ACL loss on exit from pluripotency.

**a, b**, Fractional labelling of citrate (Cit), malate (Mal) and aspartate (Asp) in control and *Acly*-edited ES cells cultured in medium containing [U-<sup>13</sup>C]glucose (**a**) or [U-<sup>13</sup>C]glutamine (**b**) following 40 h of 2i/LIF withdrawal. **c, d**, Fractional enrichment of malate M+2 relative to citrate M+2 (mal+2/cit+2) derived from [U-<sup>13</sup>C]glucose (**c**) or steady-state levels of TCA cycle metabolites (**d**) in naive, 2i-adapted control (Ctrl) and *Acly*-edited (ACLY-1 and ACLY-2) ES cells. Steady-state levels are represented as the fold change (expressed in log2) relative to control cells. **e, f**,

Assessment of the [ $\text{U}-^{13}\text{C}$ ]glucose-derived mal+2/cit+2 ratio (**e**) and steady-state levels of TCA cycle metabolites (**f**) in control and *Acly*-edited ES cells subjected to 2i/LIF withdrawal for 40 h. **g**, Relative viability (measured by PI exclusion) of control and *Acly*-edited ES cells maintained in the naive pluripotent state (+2i/LIF, left) or subjected to 2i/LIF withdrawal for 40 h (-2i/LIF, right). **h**, Immunoblot showing expression of ACSS2, the enzyme that converts acetate to acetyl-CoA in the cytosol, in naive, ground state ES cells subjected to 2i/LIF withdrawal for the indicated times. **i**, Fractional labelling of palmitate in control and *Acly*-edited ES cells cultured in medium containing [ $\text{U}-^{13}\text{C}$ ]acetate following 40 h of 2i/LIF withdrawal. Each bar represents one independent sample. **j**, Immunoblot comparing levels of acetylation (ac) at indicated histone lysine residues in control and *Acly*-edited ES cells subjected to 2i/LIF withdrawal for 40 h in the presence of vehicle or 5 mM sodium acetate. **k**, Relative viability of control and *Acly*-edited ES cells subjected to 2i/LIF withdrawal for 40 h in the presence of vehicle or 5 mM sodium acetate. **l**, Quantification of GFP mean fluorescence intensity (MFI) encoded by the *Rex1::GFPd2* reporter in ES cells subjected to 2i/LIF withdrawal for 40 h in the presence of vehicle or 50  $\mu\text{M}$  BMS-303141 (ACLi). Naïve ES cells (+2i/LIF) are included as a control. Representative histograms are shown in Fig. [4d](#). **m**, RT-qPCR of pluripotency-associated (*Nanog*, *Esrrb* and *Rex1*) and early differentiation-associated (*Sox1*) genes in control and *Acly*-edited ES cells subjected to 2i/LIF withdrawal for 40 h. Levels are represented as the fold change (expressed in  $\log_2$ ) relative to chromosome 8-targeted control cells. **n**, Alkaline phosphatase (AP) staining of colony formation assay representing control and *Acly*-edited ES cells that failed to exit the naive pluripotent state. 2i-adapted ES cells were subjected to 2i/LIF withdrawal for 40 h and then reseeded at clonal density into medium containing 2i/LIF. Quantification is shown in Fig. [4e](#). **o**, RT-qPCR of pluripotency-associated genes in control and *Acly*-edited ES cells subjected to 2i/LIF withdrawal for 40 h in the presence of vehicle or 5 mM sodium acetate. **p**, Quantification of GFP MFI encoded by the *Rex1::GFPd2* reporter in ES cells subjected to 2i/LIF withdrawal for 40 h in the presence of DMSO or 50  $\mu\text{M}$  BMS-303141(ACLi) and vehicle or 5 mM sodium acetate. **q**, Quantification of AP-positive colonies representing control and *Acly*-edited ES cells that failed to exit from the pluripotent state. ES cells were subjected to 2i/LIF withdrawal for 40 h in the presence of vehicle or 5 mM sodium acetate

prior to reseeding at clonal density into medium containing 2i and LIF. Data are mean  $\pm$  SD,  $n = 5$  (**p**),  $n = 4$  (**g, k, l**) or  $n = 3$  (all other experiments) independent replicates. In **a-b**, significance was assessed using one-way ANOVA with Sidak's multiple comparisons post-test to compare total metabolite fraction labelled relative to control cells. In remaining panels, significance was assessed by two-way ANOVA relative to control cells (**k, q**) or DMSO treatment (**p**) with Sidak's multiple comparisons post-test, or by one-way ANOVA in comparison to control cells (**c, e, g**) with Sidak's multiple comparisons post-test or in the indicated comparisons (**l**) with Tukey's multiple comparisons post-test.

### Source data

### Extended Data Fig. 9 Effect of SLC25A1 and MDH1 loss in exit from naive pluripotency.

**a, b**, Relative viability (measured by PI exclusion) of control and *Slc25a1*-edited (left) and *Mdh1*-edited (right) ES cells maintained in the naive pluripotent state (+2i/LIF, **a**) or subjected to 2i/LIF withdrawal for 40 h (-2i/LIF, **b**). **c**, Steady-state levels of TCA cycle metabolites in control and *Slc25a1*-edited (left) and *Mdh1*-edited (right) ES cells subjected to 2i/LIF withdrawal for 40 h. Steady-state levels are represented as the fold change (expressed in  $\log_2$ ) relative to control cells. **d, e**, Relative O-propargyl-puromycin (OP-puro) mean fluorescence intensity (MFI) in control, *Slc25a1*-edited, *Acly*-edited and *Mdh1*-edited ES cells that have been maintained in the naive pluripotent state (**d**) or subjected to 2i/LIF withdrawal for 40 h (**e**). Dotted line represents OP-puro MFI following cycloheximide (CHX) treatment as a control. **f, g**, Population doublings of control, *Slc25a1*-edited, *Acly*-edited and *Mdh1*-edited ES cells that have been maintained in the naive pluripotent state (**f**) or subjected to 2i/LIF withdrawal for 40 h (**g**). **h, i**, RT-qPCR of pluripotency-associated (*Nanog*, *Esrrb* and *Rex1*) and early differentiation-associated (*Sox1*) genes in control and *Slc25a1*-edited (**h**) and *Mdh1*-edited (**i**) ES cells subjected to 2i/LIF withdrawal for 40 h. Data are mean  $\pm$  SD,  $n = 3$  independent samples. Significance was assessed in comparison to control cells by one-way ANOVA with Sidak's multiple comparisons post-test.

[Source data](#)

[Extended Data Fig. 10 Mode of TCA cycle engagement regulates cell fate.](#)

**a**, Population doublings of control and *Aco2*-edited ES cells cultured in metastable (serum/LIF) conditions. **b**, Cumulative population doublings over the indicated passages of control and *Aco2*-edited ES cells upon conversion to the naive, ground state of pluripotency via addition of MEK and GSK3 $\beta$  inhibitors (+2i). **c**, RT–qPCR of pluripotency-associated genes at the indicated passages in control and *Aco2*-edited ES cells following addition of 2i. Gene expression at every passage was normalized to passage 0 (p0). Data are mean  $\pm$  SD,  $n = 1$  (**b**) or  $n = 3$  (**a, c**) independent replicates. Significance was assessed in comparison to control cells by one-way ANOVA with Sidak's multiple comparisons post-test (**a**) or relative to control cells at each timepoint with  $P$  values coloured according to comparison by two-way ANOVA with Sidak's multiple comparisons post-test (**c**).

[Source data](#)

## Supplementary information

### [Supplementary Information](#)

Supplementary Fig. 1: uncropped scans of source data for immunoblots.  
Supplementary Fig. 2: gating strategy for FACS analysis.

### [Reporting Summary](#)

### [Peer Review File](#)

### [Supplementary Table 1](#)

Source data for Fig. 1a and Extended Data Figs. 1 and 2a. Correlations derived from gene co-essentiality analysis. Data underlie the 2D network

diagrams shown in Fig. 1a and Extended Data Fig. 2a. and the heat map shown in Extended Data Fig. 1.

## **Supplementary Table 2**

All sgRNA, shRNA and RT–qPCR primer sequences. Primer sequences for sgRNAs used to perform CRISPR–Cas9-mediated editing in mouse ES cells to generate clonal lines, shRNAs used to infect C2C12 mouse cells and gene expression analysis using RT–qPCR.

## **Supplementary Table 3**

CRISPR-sequencing primers and derived amplicons. The primer sequences used to amplify genomic fragments containing edited loci of interest for sgRNA editing analysis in clonal mouse ES cell lines and the associated amplicons derived from this analysis. Amplicons are annotated to show editing of each allele for all clonal ES cell lines.

## **Supplementary Table 4**

Source data for all isotope tracing MS experiments. Full isotopologue distributions generated by isotope-tracing MS experiments.

## **Source data**

### **Source Data Fig. 1**

### **Source Data Fig. 2**

### **Source Data Fig. 3**

### **Source Data Fig. 4**

### **Source Data Extended Data Fig. 2**

[\*\*Source Data Extended Data Fig. 3\*\*](#)

[\*\*Source Data Extended Data Fig. 4\*\*](#)

[\*\*Source Data Extended Data Fig. 5\*\*](#)

[\*\*Source Data Extended Data Fig. 6\*\*](#)

[\*\*Source Data Extended Data Fig. 7\*\*](#)

[\*\*Source Data Extended Data Fig. 8\*\*](#)

[\*\*Source Data Extended Data Fig. 9\*\*](#)

[\*\*Source Data Extended Data Fig. 10\*\*](#)

## Rights and permissions

[Reprints and Permissions](#)

## About this article

### Cite this article

Arnold, P.K., Jackson, B.T., Paras, K.I. *et al.* A non-canonical tricarboxylic acid cycle underlies cellular identity. *Nature* **603**, 477–481 (2022).  
<https://doi.org/10.1038/s41586-022-04475-w>

- Received: 19 February 2021
- Accepted: 26 January 2022
- Published: 09 March 2022
- Issue Date: 17 March 2022

- DOI: <https://doi.org/10.1038/s41586-022-04475-w>

## Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

## [Alternative cycle for citrate](#)

- Paulina Strzyz

Research Highlight 16 Mar 2022

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-022-04475-w>

| [Section menu](#) | [Main menu](#) |

- Article
- [Published: 09 March 2022](#)

# Methane formation driven by reactive oxygen species across all living organisms

- [Leonard Ernst](#) [ORCID: orcid.org/0000-0003-3994-8192<sup>1,2,3</sup>](#),
- [Benedikt Steinfeld<sup>1,2,4 na1</sup>](#),
- [Uladzimir Barayeu<sup>5,6 na1</sup>](#),
- [Thomas Klintzsch](#) [ORCID: orcid.org/0000-0002-0006-2944<sup>3,7</sup>](#),
- [Markus Kurth<sup>8</sup>](#),
- [Dirk Grimm](#) [ORCID: orcid.org/0000-0001-6227-5665<sup>1,9</sup>](#),
- [Tobias P. Dick<sup>5,6</sup>](#),
- [Johannes G. Rebelein](#) [ORCID: orcid.org/0000-0003-2560-716X<sup>2</sup>](#),
- [Ilka B. Bischofs](#) [ORCID: orcid.org/0000-0001-5905-7106<sup>1,2,6</sup>](#) &
- [Frank Keppler](#) [ORCID: orcid.org/0000-0003-2766-8812<sup>3,10</sup>](#)

*Nature* volume 603, pages 482–487 (2022)

- 6246 Accesses
- 1 Citations
- 412 Altmetric
- [Metrics details](#)

## Subjects

- [Carbon cycle](#)

# Abstract

Methane ( $\text{CH}_4$ ), the most abundant hydrocarbon in the atmosphere, originates largely from biogenic sources<sup>1</sup> linked to an increasing number of organisms occurring in oxic and anoxic environments. Traditionally, biogenic  $\text{CH}_4$  has been regarded as the final product of anoxic decomposition of organic matter by methanogenic archaea. However, plants<sup>2,3</sup>, fungi<sup>4</sup>, algae<sup>5</sup> and cyanobacteria<sup>6</sup> can produce  $\text{CH}_4$  in the presence of oxygen. Although methanogens are known to produce  $\text{CH}_4$  enzymatically during anaerobic energy metabolism<sup>7</sup>, the requirements and pathways for  $\text{CH}_4$  production by non-methanogenic cells are poorly understood. Here, we demonstrate that  $\text{CH}_4$  formation by *Bacillus subtilis* and *Escherichia coli* is triggered by free iron and reactive oxygen species (ROS), which are generated by metabolic activity and enhanced by oxidative stress. ROS-induced methyl radicals, which are derived from organic compounds containing sulfur- or nitrogen-bonded methyl groups, are key intermediates that ultimately lead to  $\text{CH}_4$  production. We further show  $\text{CH}_4$  production by many other model organisms from the Bacteria, Archaea and Eukarya domains, including in several human cell lines. All these organisms respond to inducers of oxidative stress by enhanced  $\text{CH}_4$  formation. Our results imply that all living cells probably possess a common mechanism of  $\text{CH}_4$  formation that is based on interactions among ROS, iron and methyl donors, opening new perspectives for understanding biochemical  $\text{CH}_4$  formation and cycling.

This is a preview of subscription content

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

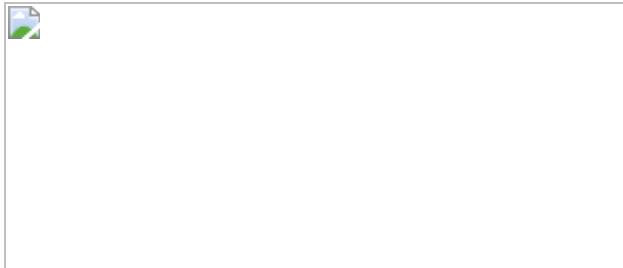
[Buy](#)

All prices are NET prices.

**Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

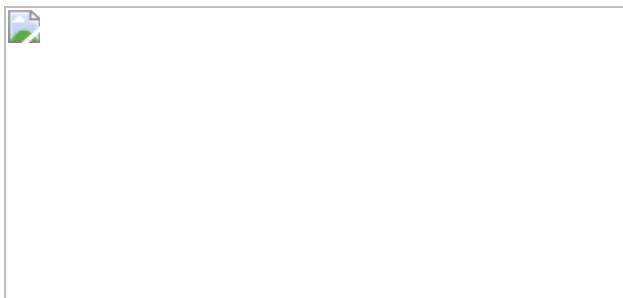
**Fig. 1: Proposed mechanism of ROS-driven CH<sub>4</sub> formation in living systems.**



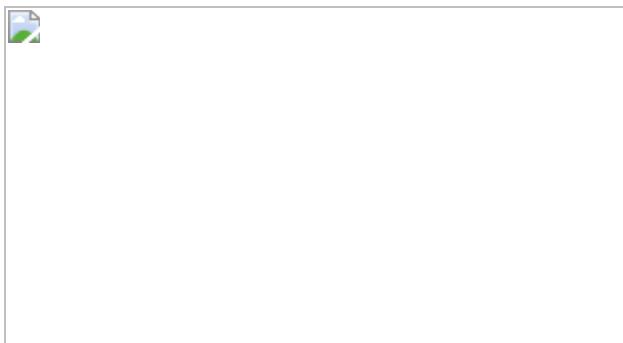
**Fig. 2: *B. subtilis* forms CH<sub>4</sub> by a ROS-driven pathway.**



**Fig. 3: Mechanistic insights into cellular CH<sub>4</sub> formation.**



**Fig. 4: Common non-methanogenic CH<sub>4</sub> formation across all lifeforms.**



## Data availability

The data reported in this paper are available in the [Supplementary Information](#) provided with this paper and on request from the corresponding authors. [Source data](#) are provided with this paper.

## References

1. Saunois, M. et al. The global methane budget 2000–2017. *Earth Syst. Sci. Data* **12**, 1561–1623 (2020).
2. Keppler, F., Hamilton, J. T. G., Braß, M. & Röckmann, T. Methane emissions from terrestrial plants under aerobic conditions. *Nature* **439**, 187–191 (2006).
3. McLeod, A. R. et al. Ultraviolet radiation drives methane emissions from terrestrial plant pectins. *New Phytol.* **180**, 124–132 (2008).
4. Lenhart, K. et al. Evidence for methane production by saprotrophic fungi. *Nat. Commun.* **3**, 1046 (2012).
5. Klintzsch, T. et al. Methane production by three widespread marine phytoplankton species: release rates, precursor compounds, and potential relevance for the environment. *Biogeosciences* **16**, 4129–4144 (2019).
6. Bižić, M. et al. Aquatic and terrestrial cyanobacteria produce methane. *Sci. Adv.* **6**, eaax5343 (2020).
7. Thauer, R. K. Methyl (alkylalkyl)-coenzyme M reductases: nickel F-430-containing enzymes involved in anaerobic methane formation and in anaerobic oxidation of methane or of short chain alkanes. *Biochemistry* **58**, 5198–5220 (2019).
8. Conrad, R. The global methane cycle: recent advances in understanding the microbial processes involved. *Env. Microbiol. Rep.* **1**, 285–292 (2009).

9. DeLong, E. F. Exploring marine planktonic archaea: then and now. *Front. Microbiol.* **11**, 3527 (2021).
10. Vorholt, J., Kunow, J., Stetter, K. O. & Thauer, R. K. Enzymes and coenzymes of the carbon monoxide dehydrogenase pathway for autotrophic CO<sub>2</sub> fixation in *Archaeoglobus lithotrophicus* and the lack of carbon monoxide dehydrogenase in the heterotrophic *A. profundus*. *Arch. Microbiol.* **163**, 112–118 (1995).
11. Hartmann, J. F. et al. High spatiotemporal dynamics of methane production and emission in oxic surface water. *Environ. Sci. Technol.* **54**, 1451–1463 (2020).
12. Kamat, S. S., Williams, H. J., Dangott, L. J., Chakrabarti, M. & Raushel, F. M. The catalytic mechanism for aerobic formation of methane by bacteria. *Nature* **497**, 132–136 (2013).
13. Metcalf, W. W. et al. Synthesis of methylphosphonic acid by marine microbes: a source for methane in the aerobic ocean. *Science* **337**, 1104–1107 (2012).
14. Zheng, Y. et al. A pathway for biological methane production using bacterial iron-only nitrogenase. *Nat. Microbiol.* **3**, 281–286 (2018).
15. North, J. A. et al. A nitrogenase-like enzyme system catalyzes methionine, ethylene, and methane biogenesis. *Science* **369**, 1094–1098 (2020).
16. Wang, Q. et al. Aerobic bacterial methane synthesis. *Proc. Natl Acad. Sci. USA* **118**, e2019229118 (2021).
17. Postgate, J. R. Methane as a minor product of pyruvate metabolism by sulphate-reducing and other bacteria. *J. Gen. Microbiol.* **57**, 293–302 (1969).
18. Althoff, F. et al. Abiotic methanogenesis from organosulphur compounds under ambient conditions. *Nat. Commun.* **5**, 4205 (2014).

19. Enami, S., Sakamoto, Y. & Colussi, A. J. Fenton chemistry at aqueous interfaces. *Proc. Natl Acad. Sci. USA* **111**, 623–628 (2014).
20. Mittler, R. ROS are good. *Trends Plant Sci.* **22**, 11–19 (2017).
21. Braun, V. & Hantke, K. Recent insights into iron import by bacteria. *Curr. Opin. Chem. Biol.* **15**, 328–334 (2011).
22. Dunbar, K. L., Scharf, D. H., Litomska, A. & Hertweck, C. Enzymatic carbon–sulfur bond formation in natural product biosynthesis. *Chem. Rev.* **117**, 5521–5577 (2017).
23. Wax, R. & Freese, E. Initiation of the germination of *Bacillus subtilis* spores by a combination of compounds in place of l-alanine. *J. Bacteriol.* **95**, 433–438 (1968).
24. Ewing, D. The effects of dimethylsulfoxide (DMSO) on the radiation sensitivity of bacterial spores. *Radiat. Res.* **90**, 348–355 (1982).
25. Setlow, B., Melly, E. & Setlow, P. Properties of spores of *Bacillus subtilis* blocked at an intermediate stage in spore germination. *J. Bacteriol.* **183**, 4894–4899 (2001).
26. Candeias, L. P., Stratford, M. R. L. & Wardman, P. Formation of hydroxyl radicals on reaction of hypochlorous acid with ferrocyanide, a model iron(II) complex. *Free Radical Res.* **20**, 241–249 (2009).
27. Bruskov, V. I., Masalimov, Z. K. & Chernikov, A. V. Heat-induced generation of reactive oxygen species in water. *Dokl. Biochem. Biophys.* **384**, 181–184 (2002).
28. Foyer, C. H. & Noctor, G. Ascorbate and glutathione: the heart of the redox hub. *Plant Physiol.* **155**, 2–18 (2011).
29. Rush, J. D. & Koppenol, W. H. Reactions of Fe(II)–ATP and Fe(II)–citrate complexes with *t*-butyl hydroperoxide and cumyl hydroperoxide. *FEBS Lett.* **275**, 114–116 (1990).

30. Wongnate, T. et al. The radical mechanism of biological methane synthesis by methyl-coenzyme M reductase. *Science* **352**, 953–958 (2016).
31. Ross, M. O. & Rosenzweig, A. C. A tale of two methane monooxygenases. *J. Biol. Inorg. Chem.* **22**, 307–319 (2016).
32. Mols, M. & Abee, T. Primary and secondary oxidative stress in *Bacillus*. *Environ. Microbiol.* **13**, 1387–1394 (2011).
33. Wishkerman, A. et al. Enhanced formation of methane in plant cell cultures by inhibition of cytochrome *c* oxidase. *Plant Cell Environ.* **34**, 457–464 (2011).
34. Tuboly, E. et al. Methane biogenesis during sodium azide-induced chemical hypoxia in rats. *Am. J. Physiol. Cell Physiol.* **304**, 207–214 (2013).
35. Klintzsch, T. et al. Effects of temperature and light on methane production of widespread marine phytoplankton. *J. Geophys. Res. Biogeosci.* **125**, e2020JG005793 (2020).
36. Polag, D., Leiß, O. & Keppler, F. Age dependent breath methane in the German population. *Sci. Total Environ.* **481**, 582–587 (2014).
37. Zhang, X. et al. Methane limit LPS-induced NF-κB/MAPKs signal in macrophages and suppress immune response in mice by enhancing PI3K/AKT/GSK-3β-mediated IL-10 expression. *Sci. Rep.* **6**, 293591 (2016).
38. Qaderi, M. M. & Reid, D. M. Methane emissions from six crop species exposed to three components of global climate change: temperature, ultraviolet-B radiation and water stress. *Physiol. Plant* **137**, 139–147 (2009).
39. Brüggemann, N. et al. Nonmicrobial aerobic methane emission from poplar shoot cultures under low-light conditions. *New Phytol.* **182**, 912–918 (2009).

40. Harwood, C. R. & Cutting, S. M. (eds) *Molecular Biological Methods for Bacillus*. Vol. 1 (John Wiley & Sons, 1990).
41. Mutlu, A. et al. Phenotypic memory in *Bacillus subtilis* links dormancy entry and exit by a spore quantity–quality tradeoff. *Nat. Commun.* **9**, 69 (2018).

## Acknowledgements

We thank M. Schneider, M. Schroll, J. Hädeler, I. Ralenekova, M. Greule, B. Knape, S. Rheinberger, C. Kaspar, Y. Gietz, C. Walter, E. Wiedtke, N. Fischer and L. Dietz for technical support; T. Erb, S. Greiner, M. Baumgart, M. Knop and R. Fischer for providing strains (affiliations included in the Supplementary information); R. Thauer for providing critical comments; and P. Hardy and J. Hamilton for editing the manuscript. Figures [1](#), [2a](#) and [3b,c](#), Extended Data Fig. [6](#) and Supplementary Fig. [7a](#) were created with [BioRender.com](#). This work was supported by funds from the Max Planck Society (I.B.B., J.G.R., L.E., B.S.), Friedrich Naumann Foundation (L.E.) and the Landesgraduiertenstiftung Baden-Württemberg (B.S.). L.E., J.G.R., F.K., T.K. and D.G. were supported by the German Research Foundation (DFG; 446841743, KE 884/8-2 and KE 884/16-2). The EPR core facility was supported by 3DMM2O–Cluster of Excellence (EXC-2082/1–390761711) at Heidelberg University.

## Author information

### Author notes

1. These authors contributed equally: Benedikt Steinfeld, Uladzimir Barayeu

## Affiliations

1. BioQuant Center, Heidelberg University, Heidelberg, Germany  
Leonard Ernst, Benedikt Steinfeld, Dirk Grimm & Ilka B. Bischofs

2. Max-Planck-Institute for Terrestrial Microbiology, Marburg, Germany

Leonard Ernst, Benedikt Steinfeld, Johannes G. Rebelein & Ilka B. Bischofs

3. Institute of Earth Sciences, Heidelberg University, Heidelberg, Germany

Leonard Ernst, Thomas Klintzsch & Frank Keppler

4. Zentrum für Molekulare Biologie Heidelberg (ZMBH), Heidelberg University, Heidelberg, Germany

Benedikt Steinfeld

5. Division of Redox Regulation, German Cancer Research Center (DKFZ), DKFZ-ZMBH Alliance, Heidelberg, Germany

Uladzimir Barayeu & Tobias P. Dick

6. Faculty of Biosciences, Heidelberg University, Heidelberg, Germany

Uladzimir Barayeu, Tobias P. Dick & Ilka B. Bischofs

7. Department for Plant Nutrition, Gießen University, Gießen, Germany

Thomas Klintzsch

8. Heidelberg Institute for Theoretical Studies, Heidelberg, Germany

Markus Kurth

9. Department of Infectious Diseases/Virology, Medical Faculty, Heidelberg University, Heidelberg, Germany

Dirk Grimm

10. Heidelberg Center for the Environment (HCE), Heidelberg University, Heidelberg, Germany

Frank Keppler

## Contributions

L.E., I.B.B. and F.K. designed the research. L.E. performed the experiments. U.B., B.S. and M.K. contributed to the experiments and analysed the following data: Fig. 2c, Extended Data Fig. 4 and Supplementary Figs. 5 and 9 (U.B.); Extended Data Fig. 1 (B.S.); and Supplementary Fig. 5 (M.K.). L.E., B.S., U.B., T.K., J.G.R., I.B.B. and F.K. analysed all other data. I.B.B., T.P.D., D.G., J.G.R. and F.K. supervised the research. L.E., I.B.B. and F.K. wrote the manuscript.

## Corresponding authors

Correspondence to [Leonard Ernst](#), [Ilka B. Bischofs](#) or [Frank Keppler](#).

## Ethics declarations

## Competing interests

The authors declare no competing interests.

## Peer review

## Peer review information

*Nature* thanks Stephen Ragsdale, Marc Strous, David Valentine and Jingyao Zhang for their contribution to the peer review of this work. Peer reviewer reports are available.

## Additional information

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Extended data figures and tables

### Extended Data Fig. 1 Germinated, but not dormant, *B. subtilis* spores generate CH<sub>4</sub> from <sup>2</sup>H-DMSO.

Scatterplot depicting <sup>2</sup>H % excess and amounts of CH<sub>4</sub> formed. The second y-axis shows the calculated corresponding fraction of methyl precursor compound (FMPC [%]) involved in CH<sub>4</sub> generation. Insets show brightfield micrographs of refractile dormant spores and non-refractile germinated spores (see [Methods](#)). DMSO was 95 % unlabeled (~natural abundance) and 5 % <sup>2</sup>H labelled. A <sup>2</sup>H % excess of 3.75 % correlates with 100 % FMPC (as CH<sub>4</sub> is formed from a <sup>2</sup>H-labelled methyl group, containing three hydrogen atoms, and a fourth, unlabelled hydrogen atom). Amounts of CH<sub>4</sub> emitted were calculated by subtracting the media controls from the spore cultures. Germination was induced by adding 3 mL of AGFK (19.8 mM L-asparagine, 33.6 mM D-glucose, 33.6 mM D-fructose and 60 mM KCl). Data points represent individual measurements from N = 3 technical replicates from one experiment.

#### Source data

### Extended Data Fig. 2 Methane formation and population growth of *B. subtilis* correlate with supply of fresh air.

While population growth (brown line) and CH<sub>4</sub> formation (yellow bars) stop under anoxic conditions, both are initiated again upon supply of fresh air (air: ~21 % O<sub>2</sub> and 78 % N<sub>2</sub>). Here, *B. subtilis* was grown at 37 °C in LB medium (starting OD<sub>600nm</sub> = 0.12), supplemented with 200 mM DMSO, in atmospherically-sealed glass vials with a volume of 20 mL including 10 mL bacterial culture and 10 mL headspace. At 1 h incubation intervals, the headspace gas was sampled and analysed. Anoxic conditions were generated by exchanging the headspace gas with a nitrogen atmosphere after drawing vacuum four times. Oxic conditions were restored by removing the vial seals for 3 min allowing entry of atmospheric air.

Amounts of CH<sub>4</sub> emitted were calculated by subtracting atmospheric background CH<sub>4</sub> levels from measured CH<sub>4</sub> levels. Six rounds of anoxia-reoxygenation were performed in order to establish that the process could occur repeatedly. Data points represent individual measurements, bars represent means from N = 2 technical replicates as guide to the eye.

[Source data](#)

**Extended Data Fig. 3 Methane formation by *B. subtilis* is enhanced by oxidative and environmental stressors.**

In comparison to unstressed cultures, CH<sub>4</sub> levels (yellow) increase significantly upon treatment with oxidants HOCl or H<sub>2</sub>O<sub>2</sub>. Environmental stressors NaCl (salt stress) and heat (50 °C) also significantly enhance CH<sub>4</sub> formation by *B. subtilis*. Pre-cultures were grown at 37 °C in LB media (starting OD<sub>600nm</sub> = 0.01) for 10 h and subsequently supplemented with 100 mM DMSO and, optionally, 300 µg mL<sup>-1</sup> HOCl, 1 mM H<sub>2</sub>O<sub>2</sub> or 4 % NaCl. 30 mL cultures and corresponding media controls were incubated at 37 °C or 50 °C (heat stress) in 60 mL sealed glass vials for 8 h. Amounts of CH<sub>4</sub> emitted were calculated by subtracting media controls from culture samples. Bars represent means ± SD from N = 3 technical replicates from one experiment, respectively. Statistical analysis was performed using paired two-tailed *t*-tests, \*\*\*: *p* ≤ 0.001. Data points represent individual measurements.

[Source data](#)

**Extended Data Fig. 4 *B. subtilis* ROS levels are promoted by oxidative stress and reduced by antioxidants.**

In *B. subtilis* cultures treated with iron (F) and DMSO (S), bacterial ROS levels are enhanced upon oxidative stress induction (O) and reduced upon antioxidant addition. *B. subtilis* pre-cultures were grown in S750 media (starting OD<sub>600nm</sub> = 0.01), containing 50 nM FeSO<sub>4</sub> and, optionally, 500 µM butylated hydroxytoluene (A) or Trolox (6-hydroxy-2,5,7,8-

tetramethylchroman-2-carboxylic acid) acting as assay validation antioxidant (V) for 6 h under standard conditions ( $37^{\circ}\text{C}$ , 180 rpm). Subsequently, cultures were supplemented with  $4.95\text{ }\mu\text{M FeSO}_4$  and 200 mM DMSO and grown for 2 h. Cultures were next diluted 10-fold, supplemented with  $25\text{ }\mu\text{M 5-(and-6)-chloromethyl 2',7'-dichlorofluorescin diacetate (CM-H2DCFDA)}$  and incubated for 30 min in the dark. Finally, oxidative stress was induced by adding  $150\text{ }\mu\text{g mL}^{-1}\text{ HOCl}$ . After 10 min incubation, DCF fluorescence was measured from live cell population with FACS Canto II (BD) at green channel (excitation: 480 nm, emission: 520 nm) and analyzed with analyzed using FlowJo<sup>TM</sup> v10.8 Software (BD Life Sciences). **(A)** Gating strategy for flow cytometry. For all samples, the identical gating strategy was applied. **(B)** The assay was validated by supplying cultures with CM-H2DCFDA, HOCl and Trolox. **(C)** Detected cellular fluorescence distributions with and without BHT treatment, indicating respective ROS levels of cells stressed with HOCl. **(D)** Median values indicating the different cellular ROS levels from the obtained distributions. Numbers in brackets denote the respective analyzed cell counts.

#### [Source data](#)

#### [Extended Data Fig. 5 \*B. subtilis\* forms CH<sub>4</sub> from methylated S-/N-compounds.](#)

$^{2}\text{H}$  % excess, the corresponding fraction methyl precursor compound (FMPC [%]) and amounts of CH<sub>4</sub> produced in the presence of the indicated substrates provide unambiguous evidence for CH<sub>4</sub> formation from methionine (yellow), trimethylamine (green) and DMSO (red). In contrast, no significant CH<sub>4</sub> formation from the quaternary compound choline (blue) was detected. In comparison to cultures grown under standard conditions (circles), all of the three measured readouts increased in stressed cultures (squares). All examined compounds except DMSO were 95 % unlabelled and 5 %  $^{2}\text{H}$ -labelled. Among the 5 % labelled compounds, all methyl groups were fully  $^{2}\text{H}$ -labelled, implying that a  $^{2}\text{H}$  % excess of 3.75 % correlates with 100 % FMPC (as CH<sub>4</sub> is formed from a  $^{2}\text{H}$ -labelled methyl

group, containing three hydrogen atoms, and a fourth, unlabelled hydrogen atom). As DMSO was 2 % labelled, a  $^{2}\text{H}$  % excess of 1.5 % correlates with 100 % FMPC. The reason for addition of only 2 % labelled DMSO was due to the large amounts of  $\text{CH}_4$  formed and practical reasons for the IRMS measurements (limitations of measuring  $\text{CH}_4$  highly enriched in  $^{2}\text{H}$ ). A pre-culture (Start OD<sub>600nm</sub> = 0.01) was grown in S750 medium, supplemented with 50  $\mu\text{M}$  FeCl<sub>3</sub>, incubated for 8 h at 37 °C and 180 rpm. The pre-culture was subsequently split into 30 mL fractions. Both these fractions and media controls were incubated in sealed 60 mL glass vials for 24 h under identical conditions and supplemented with 50 mM substrate and, optionally, 300  $\mu\text{g}$  mL<sup>-1</sup> HOCl in order to induce oxidative stress. Methane amounts were measured by GC-FID and isotope values by GC-TC-IRMS (see SI Methods section).  $\text{CH}_4$  emission was calculated by subtracting media controls from corresponding bacterial cultures. It is proposed that methionine, trimethylamine and DMSO serve as substrates for ROS-driven  $\text{CH}_4$  formation (see Extended Data Fig. 6).

[Source data](#)

[Extended Data Fig. 6 Model of ROS-driven  \$\text{CH}\_4\$  formation from endogenous substrates.](#)

Metabolic pathways and biochemical C1-transfer exist in a variety of different organisms and facilitate the production of substrates suitable for ROS-driven  $\text{CH}_4$  formation from endogenous precursor compounds. The methyl group of pyruvate (labelled in red) can be transferred to substrates for ROS-driven  $\text{CH}_4$  formation, e.g. methionine, or natural products, e.g. DMSO or TMA, which are produced by several organisms.

[Extended Data Fig. 7 Methane formation \*in vitro\* is driven by  \$\text{H}\_2\text{O}\_2\$  and enhanced by  \$\text{Fe}^{3+}\$ -reductants and Fenton-activating  \$\text{Fe}^{2+}\$ -chelators.](#)

S750 media (grey), supplemented with 50 µM FeCl<sub>3</sub> and 200 mM DMSO, was treated with 50 mM H<sub>2</sub>O<sub>2</sub> (dashed lines), 1 mM ascorbate (red), glutathione (blue) or NADH (green) and 2 mM ATP, citrate or EDTA and incubated for 24 h at a neutral pH, 37 °C and 180 rpm. Upon H<sub>2</sub>O<sub>2</sub> supplementation, CH<sub>4</sub> formation in the non-supplemented media increased by ~19-fold. Addition of Fe<sup>3+</sup>-reductants ascorbate, glutathione or NADH enhanced CH<sub>4</sub> formation by factors of ~83, ~13 or ~2, respectively. Upon addition of Fenton-activating Fe<sup>2+</sup>-chelators ATP, citrate or EDTA, the observed CH<sub>4</sub> formation in media supplemented with H<sub>2</sub>O<sub>2</sub> and ascorbate further increased by factors of ~8 (ATP), ~2 (citrate) or ~22 (EDTA). In media supplemented with H<sub>2</sub>O<sub>2</sub> and glutathione, CH<sub>4</sub> formation increased by factors of ~7 (ATP), ~3 (citrate) or ~26 (EDTA). In media supplemented with H<sub>2</sub>O<sub>2</sub> and NADH, CH<sub>4</sub> formation increased by factors of ~5 (ATP), ~4 (citrate) or ~19 (EDTA). The dashed red line shows background (laboratory air) CH<sub>4</sub> content. Sealed vials with a volume of 40 mL including 20 mL sample and 20 mL headspace were used for incubations. Data points represent individual measurements, bars represent means from N = 2 technical replicates as guide to the eye.

[Source data](#)

**Extended Data Fig. 8 Bacteria facilitate ROS-driven CH<sub>4</sub> formation.**

Overall, of the 19 different species of bacteria investigated all form CH<sub>4</sub> (yellow) which is enhanced upon HOCl-induced oxidative stress induction (yellow, dashed lines). Bacteria were cultivated in Terrific Broth (TB) medium, supplemented with 1 % Glucose, 0.1 % K-Glutamate, 50 µg mL<sup>-1</sup> L-Tryptophan, 50 µM MnCl<sub>2</sub>, 5 µM FeSO<sub>4</sub>, 2 mM MgSO<sub>4</sub>, 2 µM Thiamine, 1 µM ZnCl<sub>2</sub>, 0.7 mM CaCl<sub>2</sub>, 1.5 mM NaCl, 100 mM DMSO and, optionally, 150 µg mL<sup>-1</sup> HOCl. Furthermore, *Thermus thermophilus* HB27 and *Halomonas GFAJ-1* were supplemented with additional 30 mM and 500 mM NaCl, respectively. Stationary-phase bacterial pre-cultures (100 µL) were added to 9.9 mL media in 20 mL vials and incubated for 24 h at

37 °C with shaking at 180 r.p.m. For statistical analysis, averages for each duplicate were calculated and a two-tailed Wilcoxon signed rank test was performed ( $H_0$  = no difference between stressed and unstressed cultures). Thus, the observed differences between stressed and unstressed cultures could be demonstrated to be significant ( $p \leq 0.001$ ). Data points represent individual measurements, bars represent means from  $N = 2$  technical replicates as guide to the eye.

[Source data](#)

**Extended Data Fig. 9 Human HEK293T cells convert  $^2\text{H}$ -DMSO into  $\text{CH}_4$ .**

$^2\text{H}$  % excess, the corresponding fraction methyl precursor compound (FMPC [%]) and amounts of  $\text{CH}_4$  produced in the presence of  $^2\text{H}$ -DMSO provide unambiguous evidence for  $\text{CH}_4$  formation by HEK293T cells. 30 mL HEK293T cells (starting from  $6 * 10^5$  cells  $\text{mL}^{-1}$ ) and media controls were grown in Dulbecco's Modified Eagle's Medium (DMEM), supplemented with 10 % Fetal Calf Serum (FBS), 50 nM  $\text{FeSO}_4$  and 50 mM DMSO and incubated at 37 °C with shaking at 50 r.p.m. for 48 h in 60 mL atmospherically-sealed glass vials. DMSO was 95 % unlabelled and 5 % labelled. Among the 5 %  $^2\text{H}$ -DMSO, all methyl groups were fully  $^2\text{H}$ -labelled, implying that a  $^2\text{H}$  % excess of 3.75 % correlates with 100 % FMPC (as  $\text{CH}_4$  is formed from a  $^2\text{H}$ -labelled methyl group, containing three hydrogen atoms, and a fourth, unlabelled hydrogen atom).  $\text{CH}_4$  emission was calculated by subtracting media control from cell culture amounts. Methane amounts were measured by GC-FID and isotope values determined by GC-TC-IRMS (see [Methods](#) section).

[Source data](#)

**Extended Data Fig. 10 ROS-driven  $\text{CH}_4$  formation by mammalian cell lines.**

30 mL HEK293T, HeLa and A549 cells (starting from  $6 \times 10^5$  cells  $\text{mL}^{-1}$ ) and media controls were grown in Dulbecco's Modified Eagle's Medium (DMEM), supplemented with 10 % Fetal Calf Serum (FBS) and 50 mM DMSO at 37 °C with shaking at 50 rpm for 96 h in 60 mL atmospherically-sealed glass vials. HL60 cells (starting from  $1 \times 10^6$  cells  $\text{mL}^{-1}$ ) were grown in RPMI medium with identical supplementations and incubation conditions. Optionally, oxidative stress conditions were induced by supplementing cultures and media controls with 150  $\mu\text{g mL}^{-1}$  HOCl (orange, dashed lines). Amounts of  $\text{CH}_4$  emitted were calculated by subtracting media controls from respective cell culture samples. Bars represent means  $\pm$  SD from  $N = 3$  technical replicates from one experiment. Statistical analysis was performed using paired two-tailed *t*-tests, \*\*:  $p \leq 0.01$ , \*\*\*:  $p \leq 0.001$ . Data points represent individual measurements.

## [Source data](#)

# **Supplementary information**

## [Supplementary Information](#)

This file contains Supplementary Figs. 1–10, a list of the organisms used and the Supplementary References.

## [Reporting Summary](#)

## [Supplementary Data](#)

Data underlying Supplementary Figs. 1–10.

## [Peer Review File](#)

# **Source data**

## [Source Data Fig. 2](#)

[\*\*Source Data Fig. 3\*\*](#)

[\*\*Source Data Fig. 4\*\*](#)

[\*\*Source Data Extended Data Fig. 1\*\*](#)

[\*\*Source Data Extended Data Fig. 2\*\*](#)

[\*\*Source Data Extended Data Fig. 3\*\*](#)

[\*\*Source Data Extended Data Fig. 4\*\*](#)

[\*\*Source Data Extended Data Fig. 5\*\*](#)

[\*\*Source Data Extended Data Fig. 7\*\*](#)

[\*\*Source Data Extended Data Fig. 8\*\*](#)

[\*\*Source Data Extended Data Fig. 9\*\*](#)

[\*\*Source Data Extended Data Fig. 10\*\*](#)

## Rights and permissions

[Reprints and Permissions](#)

## About this article

### Cite this article

Ernst, L., Steinfeld, B., Barayeau, U. *et al.* Methane formation driven by reactive oxygen species across all living organisms. *Nature* **603**, 482–487 (2022). <https://doi.org/10.1038/s41586-022-04511-9>

- Received: 23 January 2021
- Accepted: 03 February 2022
- Published: 09 March 2022
- Issue Date: 17 March 2022
- DOI: <https://doi.org/10.1038/s41586-022-04511-9>

## Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

## Further reading

- [\*\*Methane might be made by all living organisms\*\*](#)

- Chang Liu
- Jingyao Zhang

*Nature* (2022)

## [\*\*Methane might be made by all living organisms\*\*](#)

- Chang Liu
- Jingyao Zhang

News & Views 09 Mar 2022

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-022-04511-9>

| [Section menu](#) | [Main menu](#) |

- Article
- Open Access
- [Published: 31 January 2022](#)

# T cell responses to SARS-CoV-2 spike cross-recognize Omicron

- [Roanne Keeton<sup>1,2</sup>,](#)
- [Marius B. Tincho ORCID: orcid.org/0000-0002-1293-5933<sup>1,2</sup>,](#)
- [Amkele Ngomti<sup>1,2</sup>,](#)
- [Richard Baguma ORCID: orcid.org/0000-0002-5106-021X<sup>1,2</sup>,](#)
- [Ntombi Benede ORCID: orcid.org/0000-0002-3319-8215<sup>1,2</sup>,](#)
- [Akiko Suzuki ORCID: orcid.org/0000-0001-8885-0189<sup>1,2</sup>,](#)
- [Khadija Khan ORCID: orcid.org/0000-0001-7565-7400<sup>3,4</sup>,](#)
- [Sandile Cele<sup>3,4</sup>,](#)
- [Mallory Bernstein<sup>3,4</sup>,](#)
- [Farina Karim<sup>3,4</sup>,](#)
- [Sharon V. Madzorera<sup>5,6</sup>,](#)
- [Thandeka Moyo-Gwete<sup>5,6</sup>,](#)
- [Mathilda Mennen<sup>7</sup>,](#)
- [Sango Skelem<sup>7</sup>,](#)
- [Marguerite Adriaanse<sup>7</sup>,](#)
- [Daniel Mutithu<sup>7</sup>,](#)
- [Olukayode Aremu<sup>7</sup>,](#)
- [Cari Stek<sup>1,7</sup>,](#)
- [Elsa du Bruyn<sup>1,7</sup>,](#)
- [Mieke A. Van Der Mescht ORCID: orcid.org/0000-0002-7521-1756<sup>8</sup>,](#)
- [Zelda de Beer<sup>9</sup>,](#)
- [Talita R. de Villiers<sup>9</sup>,](#)
- [Annie Bodenstein<sup>9</sup>,](#)
- [Gretha van den Berg<sup>9</sup>,](#)
- [Adriano Mendes<sup>10</sup>,](#)
- [Amy Strydom<sup>10</sup>,](#)
- [Marietjie Venter<sup>10</sup>,](#)
- [Jennifer Giandhari<sup>11</sup>,](#)

- [Yeshnee Naidoo<sup>11</sup>](#),
- [Sureshnee Pillay<sup>11</sup>](#),
- [Houriiyah Tegally<sup>11</sup>](#),
- [Alba Grifoni ORCID: orcid.org/0000-0002-2209-5966<sup>12</sup>](#),
- [Daniela Weiskopf<sup>12</sup>](#),
- [Alessandro Sette<sup>12,13</sup>](#),
- [Robert J. Wilkinson<sup>1,7,14,15,16</sup>](#),
- [Tulio de Oliveira<sup>11,17</sup>](#),
- [Linda-Gail Bekker<sup>1,7,18</sup>](#),
- [Glenda Gray<sup>19</sup>](#),
- [Veronica Ueckermann ORCID: orcid.org/0000-0002-4419-3583<sup>20</sup>](#),
- [Theresa Rossouw<sup>8</sup>](#),
- [Michael T. Boswell ORCID: orcid.org/0000-0003-2152-1617<sup>20</sup>](#),
- [Jinal Bihman ORCID: orcid.org/0000-0001-6354-4003<sup>5,6</sup>](#),
- [Penny L. Moore ORCID: orcid.org/0000-0001-8719-4028<sup>1,5,6,21</sup>](#),
- [Alex Sigal ORCID: orcid.org/0000-0001-8571-2004<sup>3,4,22</sup>](#),
- [Ntobeko A. B. Ntusi<sup>1,7,14,23</sup>](#),
- [Wendy A. Burgers<sup>1,2,14 na1</sup>](#) &
- [Catherine Riou ORCID: orcid.org/0000-0001-9679-0745<sup>1,2,14 na1</sup>](#)

*Nature* volume **603**, pages 488–492 (2022)

- 44k Accesses
- 8 Citations
- 1112 Altmetric
- [Metrics details](#)

## Subjects

- [Lymphocyte activation](#)
- [SARS-CoV-2](#)
- [Viral infection](#)

## Abstract

The SARS-CoV-2 Omicron variant (B.1.1.529) has multiple spike protein mutations<sup>1,2</sup> that contribute to viral escape from antibody neutralization<sup>3,4,5,6</sup> and reduce vaccine

protection from infection<sup>7,8</sup>. The extent to which other components of the adaptive response such as T cells may still target Omicron and contribute to protection from severe outcomes is unknown. Here we assessed the ability of T cells to react to Omicron spike protein in participants who were vaccinated with Ad26.CoV2.S or BNT162b2, or unvaccinated convalescent COVID-19 patients ( $n = 70$ ). Between 70% and 80% of the CD4<sup>+</sup> and CD8<sup>+</sup> T cell response to spike was maintained across study groups. Moreover, the magnitude of Omicron cross-reactive T cells was similar for Beta (B.1.351) and Delta (B.1.617.2) variants, despite Omicron harbouring considerably more mutations. In patients who were hospitalized with Omicron infections ( $n = 19$ ), there were comparable T cell responses to ancestral spike, nucleocapsid and membrane proteins to those in patients hospitalized in previous waves dominated by the ancestral, Beta or Delta variants ( $n = 49$ ). Thus, despite extensive mutations and reduced susceptibility to neutralizing antibodies of Omicron, the majority of T cell responses induced by vaccination or infection cross-recognize the variant. It remains to be determined whether well-preserved T cell immunity to Omicron contributes to protection from severe COVID-19 and is linked to early clinical observations from South Africa and elsewhere<sup>9,10,11,12</sup>.

[Download PDF](#)

## Main

The newest SARS-CoV-2 variant of concern, designated Omicron<sup>1</sup>, was first described on 26 November 2021 from sequences from Botswana, Hong Kong and South Africa<sup>2</sup>. Omicron is responsible for the current surge of infections in South Africa, and is becoming globally dominant. The variant has more than 30 mutations in the spike protein compared with the ancestral strain and a substantial ability to evade the neutralizing antibody response<sup>3,4,5,6</sup>. This is associated with greater capacity for reinfection<sup>13</sup>, as well as lower early estimates of vaccine effectiveness against symptomatic disease<sup>7,8</sup>. SARS-CoV-2-specific T cells have a role in modulating COVID-19 severity. A study of acute COVID-19 using combined measurement of CD4<sup>+</sup> T cells, CD8<sup>+</sup> T cells and neutralizing antibodies has suggested that coordination of these three arms of the adaptive response leads to lower disease severity<sup>14</sup>. A greater CD8<sup>+</sup> T cell response in blood and highly clonally expanded CD8<sup>+</sup> T cells in bronchoalveolar lavage were observed in convalescent patients who experienced mild or moderate disease compared with severe disease<sup>15,16</sup>, and CD8<sup>+</sup> T cells provided partial protective immunity in the context of suboptimal antibody titres in macaques<sup>17</sup>. In this study, we include 138 participants grouped according to their vaccination and COVID-19 status, in order to determine whether T cells generated in response to vaccination or previous SARS-CoV-2 infection could cross-recognize

Omicron and to define the profile of T cell responses in Omicron-infected patients compared with those infected with other variants of concern.

## T cell cross-reactivity to Omicron

We examined T cell responses in participants who had received one or two doses of the Ad26.COV2.S vaccine (Johnson and Johnson/Janssen,  $n = 20$  per group), two doses of the BNT162b2 mRNA vaccine (Pfizer–BioNTech,  $n = 15$  per group), or who had recovered from infection ( $n = 15$  per group) (Fig 1a, Extended Data Table 1a, b). Convalescent donors were examined a median of 1.4 months (interquartile range (IQR): 1.3–6 months) after mild or asymptomatic infection. More than 85% of vaccinees generated a T cell response to vaccination, measured 22–32 days after the last dose (Fig. 1b). Both vaccination and infection induced spike-specific CD4<sup>+</sup> T cell responses, whereas a CD8 response was less consistently detected (Fig. 1c). We measured cytokine production (IFN- $\gamma$ , IL-2 and TNF) by intracellular cytokine staining in response to peptide pools covering the full Wuhan-1 spike protein (ancestral) and the Omicron spike (Fig. 1d, Extended Data Fig. 1a).

**Fig. 1: T cell response to the ancestral and Omicron SARS-CoV-2 spike after vaccination and in unvaccinated COVID-19 convalescent patients.**

---

 **figure 1**

**a.** Clinical characteristics of the study groups. \*Data from after Covid-19 infection were available for only 6 out of the 13 participants who received one dose of Ad26.COV2.S. **b.** The proportion of participants exhibiting an ancestral spike-specific CD4<sup>+</sup> T cell response after vaccination with one or two doses of Ad26.COV2.S or two doses of BNT162b2. **c.** The profile of the ancestral spike-specific T cell response in vaccinees and convalescent (conv) individuals. **d.** Representative examples of IFN- $\gamma$  production in response to ancestral and Omicron spike in two individuals who received two doses of Ad26.COV2.S. **e, g.** Frequency of spike-specific CD4<sup>+</sup> (**e**) and CD8<sup>+</sup> T cells (**g**) producing any of the measured cytokines (IFN- $\gamma$ , IL-2 or TNF) in response to peptide pools representing ancestral and Omicron spike protein. Bars represent the median of responders. Differences between SARS-CoV-2 variants were calculated using a two-tailed Wilcoxon paired test. **f, h.** Fold change in the frequency

of spike-specific CD4<sup>+</sup> (**f**) and CD8<sup>+</sup> T cells (**h**) between ancestral and Omicron spike responses. Bars represent medians. No significant differences were observed between groups using a Kruskal–Wallis test with Dunn’s multiple comparisons post test. The number of participants included in each analysis is indicated on the graphs.

### Source Data

The levels of CD4<sup>+</sup> T cell responses to Omicron spike were consistently and significantly lower than those responsive to ancestral spike in all groups tested (Fig. [1e](#)). This translated to a median decrease of 14–30% of the CD4 response to Omicron, as demonstrated by fold change (Fig. [1f](#)). Similar results were observed for the CD8<sup>+</sup> T cell response (Fig. [1g,h](#)): vaccinees who had received two doses of Ad26.COV2.S and convalescent donors exhibited a significantly lower frequency of Omicron spike-specific CD8<sup>+</sup> T cells, although the other groups did not. There was a median reduction of 17–25% in the CD8 response to Omicron compared with the ancestral virus. Of note, a fraction of responders (5 out of 32, 15%) exhibited a loss of CD8<sup>+</sup> T cell recognition of Omicron (Figure [1g](#), Extended Data Fig. [1b](#)), probably reflecting specific human leukocyte antigen (HLA) molecules being adversely affected by mutations in particular CD8 epitopes<sup>[18](#)</sup>.

In parallel, we measured the neutralizing activity against ancestral and Omicron spike from the plasma of the same participants who received BNT162b2 ( $n = 10$ ) or two doses of Ad26.COV.S ( $n = 19$ ) (Extended Data Fig. [2](#)). As previously described<sup>[3,5,6](#)</sup>, Omicron escapes the SARS-CoV-2 neutralizing antibodies generated after BNT162b2 vaccination. Here we present neutralizing responses to Omicron after two doses of Ad26.COV2.S (Extended Data Fig. [2b](#)), demonstrating diminished neutralization capacity compared with D614G ancestral virus and the Beta variant. Comparison of the fold change in T cell responses and neutralizing antibodies targeting ancestral or Omicron spike further emphasizes the preservation of the T cell response, even when neutralization is severely reduced.

Mutations in variant epitopes have the potential to affect the functional capacity of cells<sup>[19](#)</sup>. Thus, we compared the polyfunctional profiles of T cells in vaccinees and convalescent individuals and demonstrate similar capacities for cytokine co-expression across all groups for both ancestral and Omicron-specific T cells (Extended Data Figs. [3a, b](#), [4a,b](#)). Notably, there were also no differences in the polyfunctional profiles between ancestral and Omicron spike for either CD4<sup>+</sup> or CD8<sup>+</sup> T cells (Extended Data Figs. [3c](#), [4c](#)), indicating the absence of a functional deficit in cross-reactive Omicron T cell responses. We also compared Omicron spike responses to other variants of concern in Ad26.COV2.S vaccinees, by testing spike peptide pools corresponding to the viral sequences of the Beta and Delta strains (Extended Data Fig. [5a](#)). There were no significant differences in cross-reactive CD4<sup>+</sup> and CD8<sup>+</sup> T cell responses between

Beta, Delta and Omicron (Extended Data Fig. 5b), with the exception of a greater decrease in the Omicron CD4 response compared with Beta in people who had received two doses of Ad26.COV2.S. Of note, whereas previous SARS-CoV-2 infection in vaccinees was associated with a higher frequency of spike-specific T cells (Extended Data Fig. 6a), it had no impact on Omicron cross-reactivity (Extended Data Fig. 6b). Overall, these results show that CD4<sup>+</sup> and CD8<sup>+</sup> T cell recognition of Omicron spike is largely preserved compared with the ancestral strain, and is similar to other variants of concern carrying fewer mutations.

## T cell response to different variants

The SARS-CoV-2 epidemic in South Africa has been characterized by four virologically distinct infection waves (Fig. 2b). This enabled us to compare T cell responses in patients infected with SARS-CoV-2 during the current fourth epidemic wave, dominated by Omicron, with those infected in previous waves dominated by ancestral (wave 1,  $n = 17$ ), Beta (wave 2,  $n = 16$ ) and Delta (wave 3,  $n = 16$ ) variants (Fig. 2a). In addition to extensive mutations in spike, Omicron has 20 additional mutations in other proteins which could also result in T cell escape. Therefore, we measured the frequency of CD4<sup>+</sup> and CD8<sup>+</sup> T cells reactive towards ancestral spike, nucleocapsid and membrane proteins, all major targets of the T cell response<sup>20</sup>. We studied SARS-CoV-2-infected patients who were hospitalized with COVID-19 (Fig. 2a). These recently hospitalized patients, recruited between 1 December 2021 and 15 December 2021 ( $n = 19$ ), had no previous history of COVID-19 and were unvaccinated. Omicron infection was inferred by spike gene target failure (SGTF)<sup>21</sup> in nine of these patients. Although swabs were unavailable for the remainder, with Omicron accounting for more than 90% of sequences from South Africa at the time of recruitment and 98% in Tshwane from where the samples originated (Fig. 2b), there was a high probability of Omicron infection in all of these patients.

**Fig. 2: T cell response to ancestral SARS-CoV-2 in unvaccinated hospitalized patients with COVID-19 who were infected with the ancestral, Beta, Delta or Omicron SARS-CoV-2 variants.**

---

 **figure 2**

**a**, Clinical characteristics of the study groups. Severe disease was defined on the basis of oxygen therapy requirement according to the WHO ordinal scale scoring system ( $\geq 5$ ; O<sub>2</sub> via high flow to extracorporeal membrane oxygenation). **b**, SARS-CoV-2 epidemiological dynamics in South Africa showing the prevalence of different SARS-CoV-2 strains (based on 24,762 sequences; left axis) and the number of COVID-19 cases (right axis). The bars on the top of the graph indicate the periods when samples were collected for each epidemic wave. **c, d**, Frequency of SARS-CoV-2-specific CD4<sup>+</sup> (**c**) and CD8<sup>+</sup> T cells (**d**) producing any of the measured cytokines (IFN- $\gamma$ , IL-2 or TNF) in response to ancestral SARS-CoV-2 spike (S), nucleocapsid (N) and membrane (M) peptide pools. Pies depict the proportion of participants exhibiting a detectable T cell response to each protein. **e**, Comparison of T cell response to ancestral or Omicron spike in Omicron-infected patients. Bars represent medians of

responders. No significant differences were observed between antigens amongst responders using a Kruskal–Wallis test with Dunn’s multiple comparisons post test. The number of participants included in each analysis is indicated on the graphs.

### Source Data

Despite differences in age, disease severity and co-morbidities across the infection waves (Fig. 2a, Extended Data Table 1c), T cell responses directed at spike, nucleocapsid and membrane proteins in wave 4 patients were of similar magnitude as those in patients infected with other SARS-CoV-2 variants in previous waves (Fig. 2c, d). The frequency of responders also did not differ markedly across the waves. Of note, we did not find any association between the absence of detectable CD4<sup>+</sup> T cell responses and the time post COVID-19 diagnosis or disease severity. Furthermore, the magnitude of Omicron spike-specific CD4 responses mounted by those infected in wave 4 was highly comparable to those against ancestral spike (Fig. 2e), suggesting that the CD4 responses mostly target conserved epitopes in spike. Using data from the Immune Epitope Database (<https://www.iedb.org>), we assessed the frequency of T cell recognition of experimentally-confirmed epitopes spanning the entire spike protein. Data show that Omicron spike mutations occur in regions poorly targeted by CD4<sup>+</sup> T cells, but are more common in regions frequently targeted by CD8<sup>+</sup> T cells (Extended Data Fig. 7).

To gain deeper insight into the recognition of variable spike epitopes by CD8<sup>+</sup> T cells, we also performed in silico analysis to define predicted HLA class I restriction for Omicron variable epitopes (Extended Data Table 2). Six confirmed spike epitopes containing Omicron mutations (A67V/Δ69–70, G142D/ Δ143–145, S373P, S375F, D614G, P681H and N764K) would be detrimentally affected for binding to specific class I alleles, four of which were located at positions that recorded a frequency of recognition greater than 10%. However, we also found another seven confirmed epitopes that contained Omicron mutations (T95I, S371L/S373P/S375F, K417N, G446S, Q493R, N764K and L981F) but had no effect on class I binding compared with the ancestral sequence, five of which were located at positions with a frequency of recognition greater than 10%. Overall, this suggests that although some Omicron mutations may mediate escape from specific HLA-restricted CD8<sup>+</sup> T cells, not all mutations appear to have an impact on class I binding.

## Discussion

Here we measured the ability of individuals to cross-recognize Omicron spike following vaccination, prior infection or both. We also studied unvaccinated individuals with no history of previous infection, whose first encounter with spike was with the Omicron variant. We demonstrate that vaccination and infection induce robust

CD4<sup>+</sup> and CD8<sup>+</sup> T cell responses that largely cross-react with Omicron, consistent with recent work from our laboratory and others on limited T cell escape by Beta, Delta and other variants<sup>22,23,24</sup>. Despite extensive neutralization escape against Omicron<sup>5</sup>, 70–80% of the T cell response is cross-reactive. In contrast to neutralizing antibody epitopes, T cell epitopes are abundant and located across the entire spike protein<sup>20</sup>, suggesting that the majority of SARS-CoV-2 spike-specific T cell responses are directed against conserved epitopes and that SARS-CoV-2 viral evasion from T cells may be limited.

Of note, Omicron mutations appear to abolish CD8<sup>+</sup> T cell recognition in 5 out of 32 participants (15%), in agreement with a recent report<sup>25</sup>. This loss of cross-reactive CD8<sup>+</sup> T cell responses could have pathological consequences for some individuals. Further analyses are required to define specific HLA class I profiles and epitopes linked to loss of T cell responses.

T cells are crucial components of the antiviral immune response. Although they do not prevent infection, CD4<sup>+</sup> T cells are indispensable for the generation of protective antibody responses and supporting the maturation of CD8<sup>+</sup> T cells. Hence, given the ability of variants of concern to escape neutralization, the generation and maintenance of robust SARS-CoV-2-specific T cell responses could contribute to long-term vaccine efficacy against severe disease. Several studies have reported a waning of the neutralizing response after vaccination or infection<sup>26,27,28</sup>. However, humoral responses can be enhanced upon booster vaccination, improving Omicron neutralization<sup>3,6,29,30</sup>. Vaccine- and infection-induced T cell responses also decay after antigen clearance<sup>31,32</sup>, but SARS-CoV-2-specific CD8<sup>+</sup> T cells exhibit the hallmarks of long-lived cells<sup>33</sup>, and T cell responses to SARS-CoV-1 infection were detectable 17 years later<sup>34</sup>. The longer-term durability of SARS-CoV-2-specific T cells and whether vaccine boosters can further enhance cellular immunity remain to be determined.

Despite the sharp increase in cases in South Africa in the current surge<sup>35</sup>, this has not translated into the expected increase in hospitalization or deaths, compared with previous waves<sup>12</sup>. This uncoupling of caseloads and severe outcomes could be attributed to population immunity, including maintenance of cross-reactive T cell responses observed in our study and/or intrinsic differences in Omicron severity. South Africa has high levels of SARS-CoV-2 seropositivity, driven mainly by previous infection (estimated at more than 60%) and a modest proportion of vaccinated people<sup>36</sup> (40%). Emerging data hint at reduced intrinsic severity of Omicron, including reduced infection of lower airway cells<sup>37,38</sup>. The relative contribution of high levels of immunity and potential changes in intrinsic virulence on clinical outcomes are difficult to disentangle. Moreover, it remains to be determined whether the apparently milder

outcomes at a population level will be observed in other contexts with different exposure histories and vaccination coverage, or whether the higher transmissibility of Omicron and the expected massive increase in cases in a short period will offset any gains. So far, immune correlates of protection from disease are not clearly defined and large-scale prospective studies would be necessary to evaluate correlates of protection and define the role of T cell responses in disease.

Our study had several limitations. We studied Omicron cross-reactivity of vaccine responses approximately one month after vaccination. Since T cell responses decline over time, the detection of continued cross-reactivity with variants over time will be related to the durability of the T cell response. Recall memory responses *in vivo* are likely to expand rapidly upon viral infection and contribute to limiting viral replication. We also focused on cytokine production by T helper 1 ( $T_{H1}$ ) cells to quantify CD4 $^{+}$  and CD8 $^{+}$  T cell responses. Additional approaches such as the activation-induced marker assay may capture the cellular immune response in a more comprehensive manner<sup>39</sup>. The use of 15mer peptides will have underestimated SARS-CoV-2 specific CD8 $^{+}$  T cells, as 9mer or 10mer peptides are optimal for HLA class I binding, and it has been estimated that 15mer peptides capture 77% of the frequency of CD8 $^{+}$  T cells when compared with shorter peptides<sup>40</sup>. Moreover, the saturating concentration of peptides used in these studies may underestimate the effect of mutations on T cells. In addition, the use of peptides does not allow us to define the potential effect of mutations on antigen processing and presentation, thus underestimating the effect of Omicron mutations on T cell cross-recognition. Finally, confirmation of our results from cohorts in other geographical areas and exposure to other vaccines would offer further reassurance of the maintenance of T cell responses against Omicron. Indeed, emerging data suggest this to be the case<sup>25,41,42,43,44,45</sup>.

Overall, our data show that unlike neutralizing antibodies, the SARS-CoV-2 T cell responses generated upon vaccination or previous infection are highly cross-reactive with Omicron. Early reports emerging from South Africa, England and Scotland have reported a lower risk of hospitalization and severe disease compared with the previous Delta wave<sup>9,10,11,12</sup>. It remains to be defined whether cell-mediated immunity provides protection from severe disease and contributes to the apparent milder outcomes for Omicron. Moreover, the resilience of the T cell response demonstrated here also bodes well in the event that more highly mutated variants emerge in future.

## Methods

### Human participants

A total of 138 participants were included in this study and grouped according to their vaccination and COVID-19 status. Participants were selected based on availability of peripheral blood mononuclear cells (PBMC) and clinical data were recorded by trained clinicians using RedCap (v9.5.36). The study was approved by the University of Cape Town Human Research Ethics Committee (ref: HREC 190/2020, 207/2020 and 209/2020) and the University of the Witwatersrand Human Research Ethics Committee (Medical) (ref. M210429 and M210752), the Biomedical Research Ethics Committee at the University of KwaZulu-Natal (ref. BREC/00001275/2020) and the University of Pretoria Health Sciences Research Ethics Committee (ref. 247/2020). Written informed consent was obtained from all participants.

### **Participants vaccinated with Ad26.COV2.S (one or two doses) or BNT162b2 (two doses)**

PBMC samples from 40 participants (20 who received one dose of Ad26.COV2.S vaccine and 20 who received two doses) were included in this study. These participants are enrolled in the Sisonke phase IIIb trial, an implementation trial of Ad26.COV2.S in healthcare workers. Recruitment took place at Groote Schuur Hospital (Cape Town, Western Cape, South Africa) between July 2020 and December 2021. Prior COVID-19 infection was recorded in 13 out of the 20 participants who had received one dose of the Ad26.COV2.S vaccine and in 14 out of 20 participants who had received two doses. Additionally, we also included samples from 15 participants vaccinated with two doses of BNT162b2 (Pfizer), enrolled in a prospective cohort study in KwaZulu Natal (South Africa). Prior COVID-19 infection was recorded for 6 out of 15 participants. The demographic and clinical characteristics of vaccinated participants are summarized in Extended Data Table [1a](#), with individual participant details presented in Extended Data Table [4](#).

### **Convalescent COVID-19 participants**

COVID-19 convalescent volunteers ( $n = 15$ ) were recruited from Groote Schuur Hospital in Cape Town (Western Cape, South Africa). Based on the reported date of infection, seven were probably infected with ancestral SARS-CoV-2 (before August 2020), whereas for the other 8, the infection date occurred in December 2020, suggesting an infection with the Beta variant. Samples were obtained between 19 January and 15 February 2021 before SARS-CoV-2 vaccination became available in South Africa. All had a documented positive SARS-CoV-2 PCR swab result or a positive SARS-CoV-2 nucleocapsid-specific antibody result (Roche Elecsys assay). The median time post positive test was 1.4 months, ranging from 1 to 7 months. The demographic and clinical characteristics of convalescent volunteers are summarized in Extended Data Table [1b](#), with individual participant details presented in Extended Data Table [3](#).

## Hospitalized COVID-19 patients

Sixty-eight hospitalized COVID-19 patients were included in this study. These participants were grouped according to the time of their hospitalization, reflecting four distinct infection waves in South Africa, each dominated by a different SARS-CoV-2 strain (Fig. [2b](#)). Wave 1, 2 and 3 participants were recruited from Groote Schuur Hospital in Cape Town (Western Cape, South Africa) and wave 4 patients were recruited from Groote Schuur Hospital and Tshwane District Hospital in Tshwane (Gauteng, South Africa). Wave 1 patients ( $n = 17$ ) were enrolled between 11 June and 24 July 2020, at a time when ancestral (Wuhan-1 D614G)-related SARS-CoV-2 strains were circulating. No viral sequences are available for these patients, but we assumed that all were infected with a virus closely related to the ancestral virus, as sampling occurred almost three months before the emergence of the Beta variant in South Africa. Wave 2 patients ( $n = 16$ ) were recruited between 31 December 2020 and 15 January 2021, when the Beta variant dominated. Viral sequences were available for 6 wave 2 participants, all of whom had confirmed Beta infection (GISAID accession numbers: EPI\_ISL\_1040693, 1040658, 1040661, 1040685, 1040657 and 1040663). Wave 3 patients ( $n = 16$ ) were recruited between 14 July and 21 July 2021. Wave 3 was dominated by the Delta variant. Viral sequences were available for 7 wave 3 participants, all of which were confirmed to be Delta infection (GISAID accession numbers: EPI\_ISL\_3506484, 3506367, 3957813, 3506504, 3506512 and 3506518). Wave 4 patients ( $n = 19$ ) were recruited between 1 December and 15 December 2021. The SARS-CoV-2 Omicron variant was dominant during this current wave. Amongst those patients, nine had a Taqpath PCR test performed (Thermofisher), all of which were characterized by SGTF, highly suggestive of an Omicron infection. Although we did not have confirmation of Omicron for the remaining samples, they were recruited at a time when wave 4 was driven by Omicron infection (Fig. [2b](#); there was no concomitant Delta wave in South Africa as has occurred elsewhere), with the prevalence of Omicron in South Africa at the time of recruitment being over 90% by whole-genome sequencing (WGS). Moreover, in Tshwane, from where the remainder of the samples originated, Omicron was responsible for 98% of infections sequenced at the time of sampling (61 out of 62 samples sequenced).

All hospitalized patients from waves 1, 2 and 4 were unvaccinated at the time of sampling. Third wave participants with known vaccination status were all unvaccinated ( $n = 8$ ), and the remainder ( $n = 8$ ) had unknown vaccination status. Moreover, all hospitalized patients from wave 1, 2 and 4 had no clinical record of a previous symptomatic COVID-19 episode, apart from one Wave 4 participant with an unknown history. The majority of wave 3 patients had an unknown history of prior COVID-19. The demographic and clinical characteristics of hospitalized COVID-19 participants are summarized in Extended Data Table [1c](#), and individual patient clinical data are presented in Extended Data Table [3](#).

## SARS-CoV-2 spike, WGS and phylogenetic analysis

WGS of SARS-CoV-2 was performed from nasopharyngeal swabs. Sequencing was performed as previously described<sup>2</sup>. In brief, RNA was extracted on an automated Chemagic 360 instrument, using the CMG-1049 kit (Perkin Elmer). Libraries for WGS were prepared using either the Oxford Nanopore Midnight protocol with Rapid Barcoding or the Illumina COVIDseq Assay. The quality control checks on raw sequence data and the genome assembly were performed using Genome Detective 1.133 (<https://www.genomedetective.com>) which was updated for the accurate assembly and variant calling of tiled primer amplicon Illumina or Oxford Nanopore reads, and the Coronavirus Typing Tool. Phylogenetic classification of the genomes was done using the widespread dynamic lineage classification method from the PANGOLIN software suite (v1.2.106) (<https://github.com/hCoV-2019/pangolin>).

## Isolation of PBMC

Blood was collected in heparin tubes and processed within 4 h of collection. PBMC were isolated by density gradient sedimentation using Ficoll-Paque (Amersham Biosciences) as per the manufacturer's instructions and cryopreserved in freezing media consisting of heat-inactivated fetal bovine serum (FBS, Thermo Fisher Scientific) containing 10% DMSO and stored in liquid nitrogen until use.

## SARS-CoV-2 antigens

For T cell assays on hospitalized patients, we used commercially available peptide pools (15mer sequences with 11 amino acids of overlap) covering the full length of the Wuhan-1 SARS-CoV-2 nucleocapsid, membrane and near full-length spike proteins (PepTivator, Miltenyi Biotech). For spike, we combined (1) a pool of peptides (15-mer sequences with 11 amino acids overlap) covering the ancestral N-terminal S1 domain of SARS-CoV-2 spike (GenBank MN908947.3, Protein QHD43416.1) from amino acids 1 to 692, and (2) a pool of peptides (15-mer sequences with 11 amino acids overlap) covering the immunodominant sequence domains of the ancestral C-terminal S2 domain of SARS-CoV-2 (GenBank MN908947.3, Protein QHD43416.1) including the sequence domains spanning residues 683–707, 741–770, 785–802 and 885–1273. Pools were resuspended in distilled water at a concentration of 50 µg ml<sup>-1</sup> and used at a final concentration of 1 µg ml<sup>-1</sup>. To determine T cell responses to SARS-CoV-2 variants in vaccinated and convalescent volunteers, we used custom mega pools of peptides. These peptides (15-mers overlapping by 10 amino acids) spanned the entire spike protein corresponding to the ancestral Wuhan sequence (GenBank: MN908947), Beta (B.1.351; GISAID: EPI\_ISL\_736932), Delta SARS-CoV-2 variants (B.1.617.2; GISAID: EPI\_ISL\_2020950) or Omicron (B.1.1.529), carrying in the spike sequence all the 38 currently described mutations (A67V, H69del, V70del, T95I, G142D,

V143del, Y144del, Y145del, S152W, N211del, L212I, ins214EPE, G339D, S371L, S373P, S375F, K417N, N440K, G446S, S477N, T478K, E484A, Q493R, G496S, Q498R, N501Y, Y505H, T547K, D614G, H655Y, N679K, P681H, N764K, D796Y, N856K, Q954H, N969K and L981F). In brief, peptides were synthesized as crude material (TC Peptide Lab). All individual peptides included in each mega pool are listed in Supplementary Table 1. All peptides were individually resuspended in dimethyl sulfoxide (DMSO) at a concentration of 10–20 mg ml<sup>-1</sup>. Megapools for each antigen were created by pooling aliquots of these individual peptides in the respective SARS-CoV-2 spike sequences, followed by sequential lyophilization steps, and resuspension in DMSO at 1 mg ml<sup>-1</sup>. There were 253 peptides in the ancestral, Beta and Delta variant pool, and 254 peptides in the Omicron pool. Pools were used at a final concentration of 1 µg ml<sup>-1</sup> with an equimolar DMSO concentration in the non-stimulated control.

## Cell stimulation and flow cytometry staining

Cryopreserved PBMC were thawed, washed and rested in RPMI 1640 containing 10% heat-inactivated FCS for 4 h prior to stimulation. PBMC were seeded in a 96-well V-bottom plate at approximately  $2 \times 10^6$  PBMC per well and stimulated with either the commercial ancestral SARS-CoV-2 spike (S), Nucleocapsid (N) or membrane protein (M) peptide pools (1 µg ml<sup>-1</sup>) obtained from Miltenyi or custom spike mega pools corresponding to the ancestral (Wuhan-1), Beta, Delta or Omicron variants (1 µg ml<sup>-1</sup>). All stimulations were performed in the presence of brefeldin A (10 µg ml<sup>-1</sup>, Sigma-Aldrich) and co-stimulatory antibodies against CD28 (clone 28.2) and CD49d (clone L25) (1 µg ml<sup>-1</sup> each; BD Biosciences). As a negative control, PBMC were incubated with co-stimulatory antibodies, Brefeldin A and an equimolar amount of DMSO. After 16 h of stimulation, cells were washed, stained with LIVE/DEAD Fixable VIVID Stain (1/2,500, Invitrogen, Carlsbad, CA, USA) and subsequently surface stained with the following antibodies: CD14 Pac Blue (1/100, TuK4, Invitrogen Thermofisher Scientific), CD19 Pac Blue (1/100, SJ25-C1, Invitrogen Thermofisher Scientific), CD4 PERCP-Cy5.5 (1/100, L200, BD Biosciences), CD8 BV510 (1/100, RPA-8, Biolegend). Cells were then fixed and permeabilized using a Cytofix/Cytoperm buffer (BD Biosciences) and stained with CD3 BV650 (1/100, OKT3) IFN-γ Alexa 700 (1/250, B27), TNF BV786 (1/100, Mab11) and IL-2 APC (1/100, MQ1-17H12) from Biolegend. Finally, cells were washed and fixed in CellFIX (BD Biosciences). Samples were acquired on a BD Fortessa flow cytometer and analyzed using FlowJo (v10.8, FlowJo) and Pestle and Spice v6.1 (<https://niaid.github.io/spice>). A gating strategy is provided in Extended Data Fig. 1. Results are expressed as the frequency of CD4<sup>+</sup> or CD8<sup>+</sup> T cells expressing IFN-γ, TNF or IL-2. Due to high TNF backgrounds, cells producing TNF alone were excluded from the analysis. All data are presented after background subtraction.

## **Live virus neutralization assay**

A live neutralization assay was performed on plasma obtained from 10 out of the 15 participants vaccinated with BNT162b2 included in this study. H1299-E3 cells were plated in a 96-well plate (Corning) at 30,000 cells per well 1 day pre-infection. Plasma was separated from EDTA-anticoagulated blood by centrifugation at 500g for 10 min and stored at -80 °C. Aliquots of plasma samples were heat-inactivated at 56 °C for 30 min and clarified by centrifugation at 10,000g for 5 min. Virus stocks were used at approximately 50–100 focus-forming units per microwell and added to diluted plasma. Antibody-virus mixtures were incubated for 1 h at 37 °C, 5% CO<sub>2</sub>. Cells were infected with 100 µl of the virus–antibody mixtures for 1 h, then 100 µl of 1× RPMI 1640 (Sigma-Aldrich, R6504), 1.5% carboxymethylcellulose (Sigma-Aldrich, C4888) overlay was added without removing the inoculum. Cells were fixed 18 h post-infection using 4% PFA (Sigma-Aldrich) for 20 min. Foci were stained with a rabbit anti-spike monoclonal antibody (BS-R2B12, GenScript A02058) at 0.5 µg ml<sup>-1</sup> in a permeabilization buffer containing 0.1% saponin (Sigma-Aldrich), 0.1% BSA (Sigma-Aldrich) and 0.05% Tween-20 (Sigma-Aldrich) in PBS. Plates were incubated with primary antibody overnight at 4 °C, then washed with wash buffer containing 0.05% Tween-20 in PBS. Secondary goat anti-rabbit horseradish peroxidase (Abcam ab205718) antibody was added at 1 µg ml<sup>-1</sup> and incubated for 2 h at room temperature with shaking. TrueBlue peroxidase substrate (SeraCare 5510–0030) was then added at 50 µl per well and incubated for 20 min at room temperature. Plates were imaged in an ELISPOT instrument with built-in image analysis (C.T.L.).

## **SARS-CoV-2 pseudovirus-based neutralization assay**

A pseudovirus-based neutralization assay was performed on plasma obtained from all participants vaccinated with two doses of Ad26.COV2.S ( $n = 20$ ). SARS-CoV-2 pseudotyped lentiviruses were prepared by co-transfecting the HEK 293T cell line with the SARS-CoV-2 D614G spike (D614G) or SARS-CoV-2 Beta spike (L18F, D80A, D215G, K417N, E484K, N501Y, A701V and 242–244 del) plasmids with a firefly luciferase encoding lentivirus backbone plasmid. The parental plasmids were provided by E. Landais and D. Sok. For the neutralization assays, heat-inactivated plasma samples were incubated with SARS-CoV-2 pseudotyped virus for 1 h at 37 °C, 5% CO<sub>2</sub>. Subsequently,  $1 \times 10^4$  HEK 293T cells engineered to overexpress ACE-2, provided by M. Farzan, were added and the incubated at 37 °C, 5% CO<sub>2</sub> for 72 h, upon which the luminescence of the luciferase gene was measured. CB6 and CA1 monoclonal antibodies were used as controls.

## **Statistical analysis**

Statistical analyses were performed in Prism (v9; GraphPad Software). Non-parametric tests were used for all comparisons. The Kruskal–Wallis and Mann–Whitney tests were used for unmatched samples, and the Friedman and Wilcoxon tests for paired samples. *P* values less than 0.05 were considered statistically significant. Details of analysis performed for each experiment are described in the figure legends.

## Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

## Data availability

Complete genome sequences for the viral isolates were deposited in GISAID. [Source data](#) are provided with this paper.

## References

1. WHO. Classification of Omicron (B.1.1.529): SARS-CoV-2 variant of concern. [https://www.who.int/news/item/26-11-2021-classification-of-omicron-\(b.1.1.529\)-sars-cov-2-variant-of-concern](https://www.who.int/news/item/26-11-2021-classification-of-omicron-(b.1.1.529)-sars-cov-2-variant-of-concern) (accessed 22 December 2021).
2. Viana, R. et al. Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern Africa. *Nature* <https://doi.org/10.1038/s41586-022-04411-y> (2022).
3. Cameroni, E. et al. Broadly neutralizing antibodies overcome SARS-CoV-2 Omicron antigenic shift. *Nature* <https://doi.org/10.1038/s41586-021-04386-2> (2021).
4. Cao, Y. et al. Omicron escapes the majority of existing SARS-CoV-2 neutralizing antibodies. *Nature* <https://doi.org/10.1038/s41586-021-04385-3> (2021).
5. Cele, S. et al. Omicron extensively but incompletely escapes Pfizer BNT162b2 neutralization. *Nature* <https://doi.org/10.1038/s41586-021-04387-1> (2021).
6. Planas, D. et al. Considerable escape of SARS-CoV-2 Omicron to antibody neutralization. *Nature* <https://doi.org/10.1038/s41586-021-04389-z> (2021).
7. Andrews, N. et al. Effectiveness of COVID-19 vaccines against the Omicron (B.1.1.529) variant of concern. Preprint at <https://doi.org/10.1101/2021.12.14.21267615> (2021).

8. Holm Hansen, C. et al. Vaccine effectiveness against SARS-CoV-2 infection with the Omicron or Delta variants following a two-dose or booster BNT162b2 or mRNA-1273 vaccination series: a Danish cohort study. Preprint at <https://doi.org/10.1101/2021.12.20.21267966> (2021).
9. Goga, A. et al. Breakthrough Covid-19 infections during periods of circulating Beta, Delta and Omicron variants of concern, among health care workers in the Sisonke Ad26.COV2.S vaccine trial, South Africa. Preprint at <https://doi.org/10.1101/2021.12.21.21268171> (2021).
10. Mahase, E. Covid-19: hospital admission 50–70% less likely with omicron than Delta, but transmission a major concern. *BMJ* **375**, n3151 (2021).
11. Sheikh, A. et al. Severity of Omicron variant of concern and vaccine effectiveness against symptomatic disease: national cohort with nested test negative design study in Scotland. Preprint at [https://www.pure.ed.ac.uk/ws/portalfiles/portal/245818096/Severity\\_of\\_Omicron\\_variant\\_of\\_concern\\_and\\_vaccine\\_effectiveness\\_against\\_symptomatic\\_disease.pdf](https://www.pure.ed.ac.uk/ws/portalfiles/portal/245818096/Severity_of_Omicron_variant_of_concern_and_vaccine_effectiveness_against_symptomatic_disease.pdf) (2021).
12. Wolter, N. et al. Early assessment of the clinical severity of the SARS-CoV-2 Omicron variant in South Africa. Preprint at <https://doi.org/10.1101/2021.12.21.21268116> (2021).
13. Pulliam, J. R. C. et al. Increased risk of SARS-CoV-2 reinfection associated with emergence of the Omicron variant in South Africa. Preprint at <https://doi.org/10.1101/2021.11.11.21266068> (2021).
14. Rydzynski Moderbacher, C. R. et al. Antigen-specific adaptive immunity to SARS-CoV-2 in acute COVID-19 and associations with age and disease severity. *Cell* **183**, 996–1012 (2020).
15. Peng, Y. et al. Broad and strong memory CD4<sup>+</sup> and CD8<sup>+</sup> T cells induced by SARS-CoV-2 in UK convalescent individuals following COVID-19. *Nat. Immunol.* **21**, 1336–1345 (2020).
16. Liao, M. et al. Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nat. Med.* **26**, 842–844 (2020).
17. McMahan, K. et al. Correlates of protection against SARS-CoV-2 in rhesus macaques. *Nature* **590**, 630–634 (2021).

18. Faraz Ahmed, S. et al. SARS-CoV-2 T cell responses elicited by COVID-19 vaccines or infection are expected to remain robust against Omicron. *Viruses* **14**, 79–84 (2022).
19. Holler, P. D. & Kranz, D. M. Quantitative analysis of the contribution of TCR/pepMHC affinity and CD8 to T cell activation. *Immunity* **18**, 255–264 (2003).
20. Grifoni, A. et al. SARS-CoV-2 human T cell epitopes: adaptive immune response against COVID-19. *Cell Host Microbe* **29**, 1076–1092 (2021).
21. Scott, L. et al. Track Omicron’s spread with molecular data. *Science* **374**, 1454–1455 (2021).
22. Keeton, R. et al. Prior infection with SARS-CoV-2 boosts and broadens Ad26.COV2.S immunogenicity in a variant-dependent manner. *Cell Host Microbe* **29**, 1611–1619 (2021).
23. Riou, C. et al. Escape from recognition of SARS-CoV-2 Beta variant spike epitopes but overall preservation of T cell immunity. *Sci. Transl. Med.* **14**, abj6824 (2021).
24. Tarke, A. et al. Impact of SARS-CoV-2 variants on the total CD4<sup>+</sup> and CD8<sup>+</sup> T cell reactivity in infected or vaccinated individuals. *Cell Rep. Med.* **2**, 100355 (2021).
25. Naranbhai, V. et al. T cell reactivity to the SARS-CoV-2 Omicron variant is preserved in most but not all prior infected and vaccinated individuals. Preprint at <https://doi.org/10.1101/2022.01.04.21268586> (2022).
26. Doria-Rose, N. A. et al. Booster of mRNA-1273 strengthens SARS-CoV-2 Omicron neutralization. Preprint at <https://doi.org/10.1101/2021.12.15.21267805> (2021).
27. Dupont, L. et al. Neutralizing antibody activity in convalescent sera from infection in humans with SARS-CoV-2 and variants of concern. *Nat. Microbiol.* **6**, 1433–1442 (2021).
28. Levin, E. G. et al. Waning immune humoral response to BNT162b2 Covid-19 vaccine over 6 months. *N. Engl. J. Med.* **385**, e84 (2021).
29. Doria-Rose, N. et al. Antibody persistence through 6 months after the second dose of mRNA-1273 vaccine for Covid-19. *N. Engl. J. Med.* **384**, 2259–2261 (2021b).

30. Garcia-Beltran, W. F. et al. mRNA-based COVID-19 vaccine boosters induce neutralizing immunity against SARS-CoV-2 Omicron variant. *Cell* **185**, 457–466.e4 (2022).
31. Cohen, K. W. et al. Longitudinal analysis shows durable and broad immune memory after SARS-CoV-2 infection with persisting antibody responses and memory B and T cells. *Cell Rep. Med.* **2**, 100354 (2021).
32. Dan, J. M. et al. Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection. *Science* **371**, eabf4063 (2021).
33. Adamo, S. et al. Signature of long-lived memory CD8<sup>+</sup> T cells in acute SARS-CoV-2 infection. *Nature* **602**, 148–155 (2022).
34. Le Bert, N. et al. SARS-CoV-2-specific T cell immunity in cases of COVID-19 and SARS, and uninfected controls. *Nature* **584**, 457–462 (2020).
35. Karim, S. S. A. Omicron SARS-CoV-2 variant: a new chapter in the COVID-19 pandemic. *Lancet* **398**, 2126–2128 (2021).
36. Madhi, S. A. et al. South African population immunity and severe COVID-19 with Omicron variant. Preprint at <https://doi.org/10.1101/2021.12.20.21268096> (2021).
37. Meng, B. et al. SARS-CoV-2 Omicron spike mediated immune escape, infectivity and cell-cell fusion. Preprint at <https://doi.org/10.1101/2021.12.17.473248> (2021).
38. Zhao, H. et al. SARS-CoV-2 Omicron variant shows less efficient replication and fusion activity when compared with delta variant in TMPRSS2-expressed cells: Omicron variant replication kinetics. *Emerg. Microbes Infect.* **11**, 277–283 (2021).
39. Tarke, A. et al. Comprehensive analysis of T cell immunodominance and immunoprevalence of SARS-CoV-2 epitopes in COVID-19 cases. *Cell Rep. Med.* **2**, 100204 (2021).
40. Kiecker, F. et al. Analysis of antigen-specific T-cell responses with synthetic peptides - what kind of peptide for which purpose? *Hum. Immunol.* **65**, 523–536 (2004).
41. Tarke, A. et al. SARS-CoV-2 vaccination induces immunological T cell memory able to cross-recognize variants from Alpha to Omicron. *Cell* <https://doi.org/10.1016/j.cell.2022.01.015> (2022).

42. De Marco, L. et al. Preserved T cell reactivity to the SARS-CoV-2 Omicron variant indicates continued protection in vaccinated individuals. Preprint at <https://doi.org/10.1101/2021.12.30.474453> (2021).
43. GeurtsvanKessel, C.H. et al. Divergent SARS CoV-2 Omicron-reactive T- and B cell responses in COVID-19 vaccine recipients. *Sci. Immunol.* <https://doi.org/10.1126/scimmunol.abo2202> (2022).
44. Gao, Y. et al. Ancestral SARS-CoV-2-specific T cells cross-recognize Omicron (B.1.1.529). *Nat. Med.* <https://doi.org/10.1038/s41591-022-01700-x> (2022).
45. Liu, J. et al. Vaccines elicit highly conserved cellular immunity to SARS-CoV-2 Omicron. *Nature* <https://doi.org/10.1038/s41586-022-04465-y>. (2022).

## Acknowledgements

We thank the participants who volunteered for this study. We thank W. van Hougenhouck-Tulleken for assistance with database management; R. Ramlall and J. Semugga for assistance with participant recruitment at Tshwane District Hospital; C. Scheepers and J. Everatt for generous assistance with plotting lineage prevalence; L. Singh and the NHLS Greenpoint Laboratory for PCR testing; and Biocair for assistance with shipping delays of critical reagents. Research reported in this publication was supported by the South African Medical Research Council (SA-MRC) with funds received from the South African Department of Science and Innovation (DSI), including grants 96825, SHIPNCD 76756 and DST/CON 0250/2012. This work was also supported by the Poliomyelitis Research Foundation (21/65) and Wellcome CIDRI-Africa, which is supported by core funding from Wellcome Trust (203135/Z/16/Z and 222574). This project has been funded in part with Federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, under contract no. 75N93021C00016 to A. Sette and contract no. 75N9301900065 to A. Sette and D.W. P.L.M. is supported by the SA Research Chairs Initiative of the DSI and the National Research Foundation (NRF; 9834). W.A.B. and C.R. are supported by the EDCTP2 programme of the European Union's Horizon 2020 programme (TMA2017SF-1951-TB-SPEC to C.R. and TMA2016SF-1535-CaTCH-22 to W.A.B.). N.A.B.N. acknowledges funding from the SA-MRC, MRC UK and NRF. A. Sette acknowledges funding from the Bill and Melinda Gates award INV-018944, the NIH (AI138546) and the SA-MRC. R.J.W. acknowledges funding from the Francis Crick Institute, which receives funding from Wellcome FC0010218, UKRI FC0010218, CRUK FC0010218 and Rosetrees Trust (M926).

## Author information

## Author notes

1. These authors contributed equally: Wendy A. Burgers, Catherine Riou

## Affiliations

1. Institute of Infectious Disease and Molecular Medicine, University of Cape Town, Observatory, Cape Town, South Africa

Roanne Keeton, Marius B. Tincho, Amkele Ngomti, Richard Baguma, Ntombi Benede, Akiko Suzuki, Cari Stek, Elsa du Bruyn, Robert J. Wilkinson, Linda-Gail Bekker, Penny L. Moore, Ntobeko A. B. Ntusi, Wendy A. Burgers & Catherine Riou

2. Division of Medical Virology, Department of Pathology, University of Cape Town; Observatory, Cape Town, South Africa

Roanne Keeton, Marius B. Tincho, Amkele Ngomti, Richard Baguma, Ntombi Benede, Akiko Suzuki, Wendy A. Burgers & Catherine Riou

3. Africa Health Research Institute, Durban, South Africa

Khadija Khan, Sandile Cele, Mallory Bernstein, Farina Karim & Alex Sigal

4. School of Laboratory Medicine and Medical Sciences, University of KwaZulu-Natal, Durban, South Africa

Khadija Khan, Sandile Cele, Mallory Bernstein, Farina Karim & Alex Sigal

5. National Institute for Communicable Diseases of the National Health Laboratory Service, Johannesburg, South Africa

Sharon V. Madzorera, Thandeka Moyo-Gwete, Jinal Bihman & Penny L. Moore

6. SA MRC Antibody Immunity Research Unit, School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

Sharon V. Madzorera, Thandeka Moyo-Gwete, Jinal Bihman & Penny L. Moore

7. Department of Medicine, University of Cape Town and Groote Schuur Hospital; Observatory, Cape Town, South Africa

Mathilda Mennen, Sango Skelem, Marguerite Adriaanse, Daniel Mutithu, Olukayode Aremu, Cari Stek, Elsa du Bruyn, Robert J.

Wilkinson, Linda-Gail Bekker & Ntobeko A. B. Ntusi

8. Department of Immunology, University of Pretoria, Pretoria, South Africa

Mieke A. Van Der Mescht & Theresa Rossouw

9. Tshwane District Hospital, Tshwane, South Africa

Zelda de Beer, Talita R. de Villiers, Annie Bodenstein & Gretha van den Berg

10. Centre for Viral Zoonoses, Department of Medical Virology, University of Pretoria, Pretoria, South Africa

Adriano Mendes, Amy Strydom & Marietjie Venter

11. KwaZulu-Natal Research Innovation and Sequencing Platform, University of KwaZulu-Natal, Durban, South Africa

Jennifer Giandhari, Yeshnee Naidoo, Sureshnee Pillay, Houriiyah Tegally & Tulio de Oliveira

12. Center for Infectious Disease and Vaccine Research, La Jolla Institute for Immunology, La Jolla, CA, USA

Alba Grifoni, Daniela Weiskopf & Alessandro Sette

13. Department of Medicine, Division of Infectious Diseases and Global Public Health, University of California, San Diego (UCSD), La Jolla, CA, USA

Alessandro Sette

14. Wellcome Centre for Infectious Diseases Research in Africa, University of Cape Town, Observatory, Cape Town, South Africa

Robert J. Wilkinson, Ntobeko A. B. Ntusi, Wendy A. Burgers & Catherine Riou

15. Department of Infectious Diseases, Imperial College London, London, UK

Robert J. Wilkinson

16. The Francis Crick Institute, London, UK

Robert J. Wilkinson

17. Centre for Epidemic Response and Innovation, Stellenbosch University,  
Stellenbosch, South Africa

Tulio de Oliveira

18. Desmond Tutu HIV Centre, University of Cape Town, Cape Town, South Africa

Linda-Gail Bekker

19. South African Medical Research Council, Cape Town, South Africa

Glenda Gray

20. Department of Internal Medicine, University of Pretoria and Steve Biko  
Academic Hospital, Pretoria, South Africa

Veronica Ueckermann & Michael T. Boswell

21. Centre for the AIDS Programme of Research in South Africa, Durban, South  
Africa

Penny L. Moore

22. Max Planck Institute for Infection Biology, Berlin, Germany

Alex Sigal

23. Cape Heart Institute, Faculty of Health Sciences, University of Cape Town;  
Observatory, Cape Town, South Africa

Ntobeko A. B. Ntusi

## Contributions

W.A.B, C.R. and R.K. designed the study. C.R. and W.A.B analysed the data and wrote the manuscript. K.K., S.C., M.B., F.K. and S.V.M. performed neutralization assays and analysed the data. R.K., M.B.T., A.N., R.B., N.B. and A. Suzuki performed flow cytometry assays. M.M., S.S., M.A., D.M., O.A., C.S., E.d.B., M.A.V.D.M., Z.d.B., T.R.d.V. A.B., G.v.d.B, A.M., A. Strydom and M.V. recruited patients and managed cohorts. A.G., D.W. and A. Sette designed and provided variant peptide pools. T.d.O., J.G., S.P., Y.N. and H.T. generated and analysed sequence data. T.M.-G., J.B., P.L.M. and A. Sigal supervised the neutralization assay. N.A.B.N., R.J.W., V.U., M.T.B and T.R. led the clinical cohorts. L.-G.B. and G.G. led the Sisonke Ad26.COV2.S vaccination trial. All authors reviewed and edited the manuscript.

## Corresponding authors

Correspondence to [Wendy A. Burgers](#) or [Catherine Riou](#).

## Ethics declarations

### Competing interests

A. Sette is a consultant for Gritstone Bio, Flow Pharma, Arcturus Therapeutics, ImmunoScape, CellCarta, Avalia, Moderna, Fortress and Repertoire. All of the other authors declare no competing interests. La Jolla Institute for Immunology has filed a patent for protection for various aspects of vaccine design and identification of specific epitopes.

## Peer review

### Peer review information

*Nature* thanks Antonio Bertoletti, Katherine Kedzierska and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

## Additional information

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Extended data figures and tables

### [Extended Data Fig. 1 Gating strategy and examples of flow cytometry plots.](#)

**a**, Gating strategy and representative examples of SARS-CoV-2 spike-specific IFN- $\gamma$ , IL-2 and TNF- $\alpha$  production. **b**, Spike-specific expression of IFN- $\gamma$  in the T cell compartment of the three BNT162b2-vaccinated participants where Omicron-specific CD8+ T cells were undetectable.

### [Extended Data Fig. 2 Neutralization of Omicron compared to the ancestral SARS-CoV-2 \(D614G\) by plasma from participants vaccinated](#)

## with two doses of BNT162b2 or Ad26.COV2.S.

**a**, Neutralization by BNT162b2 plasma (n = 10), 6 with prior COVID-19 infection and 4 without) was performed using a live virus neutralization assay. The reciprocal plasma dilution ( $FRNT_{50}$ ) resulting in 50% reduction in the number of infection foci is reported. The threshold of detection was set at a  $FRNT_{50}$  of 20. A two-tailed paired Wilcoxon test was used to compare ancestral and Omicron titers. Comparison of the fold change in SARS-CoV-2-specific CD8+ and CD4+ T cell responses and neutralization titers (Omicron/ancestral) is depicted in the right panel. Bars represent medians. **b**, Neutralization against ancestral, Beta and Omicron variants by plasma from Ad26.COV2.S vaccinees (two doses; n = 19), including 14 with prior COVID-19 infection and 5 without, was performed using a SARS-CoV-2 pseudovirus-based neutralization assay. The threshold of detection was a 50% inhibitory dilution ( $ID_{50}$ ) of 20. A Friedman test with Dunn's multiple comparisons post-test was used to compare the titers of the three variants tested. Comparison of the fold change in SARS-CoV-2-specific CD8+ and CD4+ T cell response and neutralization titers (Omicron/ancestral) is depicted in the right panel. Bars represent medians.

### Source Data

## Extended Data Fig. 3 Polyfunctional profiles of SARS-CoV-2-specific CD4+ T cells after vaccination and in unvaccinated convalescent volunteers.

**a, b**, Comparison of the polyfunctional profile of ancestral (**a**) and Omicron (**b**) spike-specific CD4+ T cells between the four groups (Ad26.COV2.S-one dose, Ad26.COV.S-two doses, BNT162b2-two doses and unvaccinated convalescent volunteers). **c**, Comparison of the polyfunctional profile between ancestral and Omicron spike-specific CD4+ T cells including all CD4+ T cell responding participants, irrespective of their clinical grouping. The medians and IQR are shown. Each response pattern (i.e., any possible combination of IFN- $\gamma$ , IL-2 or TNF- $\alpha$  expression) is color-coded, and data are summarized in the pie charts. No significant differences were observed between pies using a permutation test. The number of participants included in each analysis is indicated on the graphs.

### Source Data

## Extended Data Fig. 4 Polyfunctional profiles of SARS-CoV-2-specific CD8+ T cells after vaccination and in unvaccinated convalescent volunteers.

**a, b**, Comparison of the polyfunctional profile of ancestral (**a**) and Omicron (**b**) spike-specific CD8+ T cells between the four groups (Ad26.COV2.S-one dose, Ad26.COV2.S-two doses, BNT162b2-two doses and unvaccinated convalescent COVID-19 volunteers). **c**, Comparison of the polyfunctional profile between ancestral spike and Omicron spike-specific CD8+ T cells including all CD8+ T cell responding participants, irrespective of their clinical grouping. The medians and IQR are shown. Each response pattern (i.e., any possible combination of IFN- $\gamma$ , IL-2 or TNF- $\alpha$  expression) is color-coded, and data are summarized in the pie charts. No significant differences were observed between pies using a permutation test. The number of participants included in each analysis is indicated on the graphs.

[Source Data](#)

[Extended Data Fig. 5 T cell responses to the ancestral, Beta, Delta and Omicron SARS-CoV-2 spike in participants who received Ad26.COV2.S \(one or two doses\).](#)

**a**, Frequency of spike-specific CD4+ (left panel) and CD8+ T cells (right panel) producing any of the measured cytokines (IFN- $\gamma$ , IL-2 or TNF- $\alpha$ ) in response to ancestral, Beta, Delta and Omicron spike peptide pools. Bars represent median of responders. No significant differences were observed between variants using a Kruskal-Wallis test with Dunn's multiple comparisons post-test. **b**, Fold change in the frequency of spike-specific CD4+ (left panel) and CD8+ T cells (right panel) between ancestral and Omicron spike responses. Bars represent medians. Differences between SARS-CoV-2 variants were calculated using a Kruskal-Wallis test with Dunn's multiple comparisons post-test. Median fold changes are indicated at the bottom of each graph. The number of participants included in each analysis is indicated on the graphs.

[Source Data](#)

[Extended Data Fig. 6 Impact of prior COVID-19 infection on T cell responses to the ancestral and Omicron SARS-CoV-2 spike in vaccinated participants.](#)

**a**, Comparison of the frequency of ancestral spike-specific T cell responses in vaccinated participants who had (Y) or did not have (N) prior SARS-CoV-2 infection. Pies depict the proportion of participants exhibiting a detectable CD8+ T cell response. Bars represent medians. Statistical differences were calculated using a two-tailed Mann-Whitney test. **b**, Fold change in the frequency of spike-specific CD4+ T cells between ancestral and Omicron spike responses in the three vaccine groups. Bars represent medians. Statistical differences were calculated using a two-tailed Mann-

Whitney test. The number of participants included in each analysis is indicated on the graphs.

[Source Data](#)

[\*\*Extended Data Fig. 7 Distribution of spike SARS-CoV-2 epitopes targeted by CD4+ and CD8+ T cells.\*\*](#)

**a**, Schematic of SARS-CoV-2 spike protein primary structure colored by domain. NTD: N-terminal domain, RBD: receptor binding domain, SD1: Sub-domain 1, SD2: Sub-domain 2. **b**, Distribution and frequency of recognition of confirmed CD4+ (top) and CD8+ T cell epitopes (bottom) across the entire spike protein. Data represent experimentally confirmed epitopes from the Immune Epitope Database and Analysis Resource ([www.iedb.org](http://www.iedb.org)). Red lines depict the position of Omicron mutations that recorded a frequency of recognition > 10% and blue lines < 10%. The position of variable epitopes associated with specific HLA-class I (see Extended Data Table 2) is indicated by a triangle. Mutations with a detrimental or neutral impact for HLA binding are depicted in orange and green, respectively.

[Source Data](#)

**Extended Data Table 1 Clinical characteristics of vaccinated, convalescent, and hospitalized COVID-19 participants**

**Extended Data Table 2 In silico analysis of the impact of Omicron mutations on epitope recognition by MHC Class I**

**Extended Data Table 3 Clinical characteristics of each hospitalized and convalescent COVID-19 participant**

**Extended Data Table 4 Clinical characteristics of each vaccinated participant**

## Supplementary information

[\*\*Supplementary Table 1\*\*](#)

List of 15mer peptides included in the Ancestral, Beta, Delta and Omicron mega pools. The Ancestral, Beta and Delta pools consisted of 253 peptides and the Omicron pool consists of 254 peptides.

[\*\*Reporting Summary\*\*](#)

[\*\*Peer Review File\*\*](#)

## Source data

[Source Data Fig. 1](#)

[Source Data Fig. 2](#)

[Source Data Extended Data Fig. 2](#)

[Source Data Extended Data Fig. 3](#)

[Source Data Extended Data Fig. 4](#)

[Source Data Extended Data Fig. 5](#)

[Source Data Extended Data Fig. 6](#)

[Source Data Extended Data Fig. 7](#)

## Rights and permissions

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

[Reprints and Permissions](#)

## About this article

### Cite this article

Keeton, R., Tincho, M.B., Ngomti, A. *et al.* T cell responses to SARS-CoV-2 spike cross-recognize Omicron. *Nature* **603**, 488–492 (2022).  
<https://doi.org/10.1038/s41586-022-04460-3>

- Received: 27 December 2021
- Accepted: 20 January 2022
- Published: 31 January 2022
- Issue Date: 17 March 2022
- DOI: <https://doi.org/10.1038/s41586-022-04460-3>

## Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

## Further reading

- [\*\*The emergence of COVID-19 vaccine resistance depends on human choices\*\*](#)

*Nature Human Behaviour* (2022)

- [\*\*Cross reactive T cells hold up against Omicron\*\*](#)

◦ Alexandra Flemming

*Nature Reviews Immunology* (2022)

- [\*\*COVID-19 vaccines in patients with cancer: immunogenicity, efficacy and safety\*\*](#)

◦ Annika Fendler  
◦ Elisabeth G. E. de Vries  
◦ Marie von Lilienfeld-Toal

*Nature Reviews Clinical Oncology* (2022)

- **Rapid epidemic expansion of the SARS-CoV-2 Omicron variant in southern Africa**

- Raquel Viana
- Sikhulile Moyo
- Tulio de Oliveira

*Nature* (2022)

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-022-04460-3>

| [Section menu](#) | [Main menu](#) |

- Article
- Open Access
- [Published: 31 January 2022](#)

# Vaccines elicit highly conserved cellular immunity to SARS-CoV-2 Omicron

- [Jinyan Liu<sup>1</sup>](#),  
<sup>1</sup> [na1](#)
- [Abishek Chandrashekhar](#) [ORCID: orcid.org/0000-0001-7821-5552](#)<sup>1</sup>  
<sup>1</sup> [na1](#)
- [Daniel Sellers<sup>1</sup>](#)  
<sup>1</sup> [na1](#)
- [Julia Barrett<sup>1</sup>](#),
- [Catherine Jacob-Dolan](#) [ORCID: orcid.org/0000-0001-6641-0342](#)<sup>1,2</sup>,
- [Michelle Lifton](#) [ORCID: orcid.org/0000-0003-0279-2198](#)<sup>1</sup>,
- [Katherine McMahan<sup>1</sup>](#),
- [Michaela Sciacca<sup>1</sup>](#),
- [Haley VanWyk<sup>1</sup>](#),
- [Cindy Wu<sup>1</sup>](#),
- [Jingyou Yu<sup>1</sup>](#),
- [Ai-iris Y. Collier](#) [ORCID: orcid.org/0000-0002-8539-3960](#)<sup>1</sup> &
- [Dan H. Barouch](#) [ORCID: orcid.org/0000-0001-5127-4659](#)<sup>1,2</sup>

*Nature* volume 603, pages 493–496 (2022)

- 37k Accesses
- 5 Citations
- 839 Altmetric
- [Metrics details](#)

## Subjects

- [SARS-CoV-2](#)
- [T cells](#)
- [Vaccines](#)

## Abstract

The highly mutated SARS-CoV-2 Omicron (B.1.1.529) variant has been shown to evade a substantial fraction of neutralizing antibody responses elicited by current vaccines that encode the WA1/2020 spike protein<sup>1</sup>. Cellular immune responses, particularly CD8<sup>+</sup> T cell responses, probably contribute to protection against severe SARS-CoV-2 infection<sup>2,3,4,5,6</sup>. Here we show that cellular immunity induced by current vaccines against SARS-CoV-2 is highly conserved to the SARS-CoV-2 Omicron spike protein. Individuals who received the Ad26.COV2.S or BNT162b2 vaccines demonstrated durable spike-specific CD8<sup>+</sup> and CD4<sup>+</sup> T cell responses, which showed extensive cross-reactivity against both the Delta and the Omicron variants, including in central and effector memory cellular subpopulations. Median Omicron spike-specific CD8<sup>+</sup> T cell responses were 82–84% of the WA1/2020 spike-specific CD8<sup>+</sup> T cell responses. These data provide immunological context for the observation that current vaccines still show robust protection against severe disease with the SARS-CoV-2 Omicron variant despite the substantially reduced neutralizing antibody responses<sup>7,8</sup>.

[Download PDF](#)

## Main

Recent studies have shown that vaccine-elicited neutralizing antibodies (NAbs) are substantially reduced to the highly mutated SARS-CoV-2 Omicron variant<sup>1</sup>. To evaluate the cross-reactivity of vaccine-elicited cellular immune responses against the SARS-CoV-2 Omicron variant, we assessed CD8<sup>+</sup> and CD4<sup>+</sup> T cell responses in 47 individuals who were vaccinated with the adenovirus vector-based Ad26.COV2.S vaccine<sup>9</sup> (Johnson & Johnson;  $n = 20$ ) or the mRNA-based BNT162b2 vaccine<sup>10</sup> (Pfizer;  $n = 27$ ) in Boston, MA, USA (Extended Data Table 1).

## Humoral immune responses

All individuals were SARS-CoV-2 naive by history and also had negative antibody responses to nucleocapsid (Extended Data Fig. 1). Following vaccination with BNT162b2, we observed high WA1/2020-specific pseudovirus NAb responses at month 1, followed by a sharp decline by month 8 ( $P < 0.0001$ , two-tailed Mann–Whitney test), as expected<sup>11,12</sup> (Fig. 1a). Following vaccination with Ad26.COV2.S, there were initial substantially lower WA1/2020-specific pseudovirus NAb responses at month 1, but these responses were more durable and persisted at month 8 (refs. 11,13) (Fig. 1a). However, minimal cross-reactive Omicron-specific NAbs were

observed for both vaccines ( $P < 0.0001$  for both, two-tailed Mann–Whitney tests) (Fig. 1a), consistent with recent data in the absence of additional boosting<sup>1</sup>. The responses of receptor-binding domain-specific binding antibodies were assessed by ELISA and showed similar trends, with minimal cross-reactive Omicron receptor-binding domain-specific binding antibodies (Fig. 1b, Extended Data Fig. 2).

**Fig. 1: Humoral immune responses to Omicron.**

 figure 1



Antibody responses at months 1 and 8 following final vaccination with Ad26.COV2.S ( $n = 20$ ) or BNT162b2 ( $n = 27$ ). **a**, Neutralizing antibody (NAb) titres by a luciferase-based pseudovirus neutralization assay. **b**, Receptor-binding domain (RBD)-specific binding antibody titres by ELISA. Responses were measured against the SARS-CoV-2 WA1/2020 (WA), B.1.617.2 (Delta), B.1.351 (Beta) and B.1.1.529 (Omicron) variants. Medians (red bars) are depicted and numerically shown.

## Cellular immune responses

In contrast to antibody responses, spike-specific cellular immune responses assessed by pooled peptide IFN $\gamma$  ELISPOT assays showed substantial cross-reactivity to Omicron (Extended Data Fig. 3, Supplementary Table 1). We next assessed spike-specific CD8 $^{+}$  and CD4 $^{+}$  T cell responses by intracellular cytokine staining assays (Extended Data Figs. 4, 5, Supplementary Table 1). Ad26.COV2.S induced median spike-specific IFN $\gamma$  CD8 $^{+}$  T cell responses of 0.061%, 0.062% and 0.051% against WA1/2020, Delta and Omicron, respectively, at month 8 following vaccination (Fig. 2a). BNT162b2 induced median spike-specific IFN $\gamma$  CD8 $^{+}$  T cell responses of 0.028% and 0.023% against WA1/2020 and Omicron, respectively, at month 8 following vaccination (Fig. 2a). These data suggest that median Omicron-specific CD8 $^{+}$  T cell responses were 82–84% cross-reactive with WA1/2020-specific CD8 $^{+}$  T cell responses (the *P* value was not significant; two-tailed Mann–Whitney test). Spike-specific IFN $\gamma$  CD4 $^{+}$  T cell responses elicited by Ad26.COV2.S were a median of 0.026%, 0.030% and 0.029% against WA1/2020, Delta and Omicron, respectively, and by BNT162b2 were a median of 0.033% and 0.027% against WA1/2020 and Omicron, respectively, at month 8, indicating that median Omicron-specific CD4 $^{+}$  T cell responses were 82–100% cross-reactive with WA1/2020-specific CD4 $^{+}$  T cell responses (the *P* value was not significant; two-tailed Mann–Whitney test) (Fig. 2b). These data demonstrate substantial CD8 $^{+}$  and CD4 $^{+}$  T cell cross-reactivity to Omicron, although responses may be impacted more in select individuals (Fig. 3a). Substantial Omicron cross-reactivity was also observed for spike-specific IFN $\gamma$ -secreting, TNF-secreting and IL-2-secreting CD8 $^{+}$  and CD4 $^{+}$  T cell responses (Extended Data Fig. 6). By contrast, unvaccinated, uninfected individuals had negligible spike-specific CD8 $^{+}$  and CD4 $^{+}$  T cell responses (Fig. 2a,b).

**Fig. 2: Cellular immune responses to Omicron.**

---

 **figure 2**

T cell responses at months 1 and 8 following final vaccination with Ad26.COV2.S ( $n = 20$ ) or BNT162b2 ( $n = 27$ ). **a, b**, Pooled peptide spike-specific IFN $\gamma$  CD8 $^{+}$  T cell responses (**a**) and CD4 $^{+}$  T cell responses (**b**) by intracellular cytokine staining assays. Responses were measured against the SARS-CoV-2 WA1/2020, B.1.617.2 (Delta) and B.1.1.529 (Omicron) variants. Responses in five unvaccinated, uninfected individuals are also shown. Media backgrounds were subtracted from the specific values. Medians (red bars) are depicted and numerically shown.

**Fig. 3: Correlations of variant-specific and WA1/2020-specific cellular immune responses.**

---

 **figure 3**

**a**, Ratio of Omicron to WA1/2020 CD8<sup>+</sup> (top) and CD4<sup>+</sup> (bottom) T cell responses in individual participants. **b**, **c**, Correlations of log Delta-specific and Omicron-specific to log WA1/2020-specific CD8<sup>+</sup> T cell responses (**b**) and CD4<sup>+</sup> T cell responses (**c**) by intracellular cytokine staining assays. Two-sided unadjusted *P* and *R* values for linear regression correlations are shown, and lines of best fit and slopes are depicted.

Omicron-specific CD8<sup>+</sup> T cell responses correlated with WA1/2020-specific CD8<sup>+</sup> T cell responses for the Ad26.COV2.S vaccine for both timepoints ( $R = 0.78, P < 0.0001$ , slope = 0.75) and the BNT162b2 vaccine ( $R = 0.56, P < 0.0001$ , slope = 0.81), although

two individuals had undetectable Omicron-specific CD8<sup>+</sup> T cell responses following vaccination with BNT162b2 (Fig. 3b). Similarly, Omicron-specific CD4<sup>+</sup> T cell responses correlated with WA1/2020-specific CD4<sup>+</sup> T cell responses for both the Ad26.COV2.S vaccine ( $R = 0.79$ ,  $P < 0.0001$ , slope = 0.83) and the BNT162b2 vaccine ( $R = 0.90$ ,  $P < 0.0001$ , slope = 0.88) (Fig. 3c).

Spike-specific IFN $\gamma$  CD8<sup>+</sup> and CD4<sup>+</sup> T cell central (CD45RA<sup>-</sup>CD27<sup>+</sup>) and effector (CD45RA<sup>-</sup>CD27<sup>-</sup>) memory subpopulations elicited by Ad26.COV2.S also showed extensive cross-reactivity to Delta and Omicron variants. At month 8, CD8<sup>+</sup> T cell central memory responses were 0.076%, 0.054% and 0.075%, CD8<sup>+</sup> T cell effector memory responses were 0.168%, 0.143% and 0.146%, CD4<sup>+</sup> T cell central memory responses were 0.030%, 0.035% and 0.038%, and CD4<sup>+</sup> T cell effector memory responses were 0.102%, 0.094% and 0.083% against WA1/202, Delta and Omicron, respectively (Fig. 4).

**Fig. 4: Cellular immune memory subpopulations to Omicron.**



Pooled peptide spike-specific IFN $\gamma$  CD8 $^{+}$  and CD4 $^{+}$  central memory (CD45RA $^{-}$ CD27 $^{+}$ ) and effector memory (CD45RA $^{-}$ CD27 $^{-}$ ) T cell responses by intracellular cytokine staining assays at months 1 and 8 following final vaccination with Ad26.COV2.S ( $n = 20$ ). Responses were measured against the SARS-CoV-2 WA1/2020, B.1.617.2 (Delta) and B.1.1.529 (Omicron) variants. Medians (red bars) are depicted and numerically shown.

## Discussion

Our data demonstrate that Ad26.COV2.S and BNT162b2 elicit broadly cross-reactive cellular immunity against SARS-CoV-2 variants including Omicron. The consistency of these observations across two different vaccine platform technologies (viral vector and mRNA) suggests the generalizability of these findings. The extensive cross-reactivity of Omicron-specific CD8 $^{+}$  and CD4 $^{+}$  T cell responses contrasts sharply with the marked reduction of Omicron-specific antibody responses. These data are consistent with previous studies that have shown greater cross-reactivity of vaccine-elicited cellular immune responses than humoral immune responses against the SARS-CoV-2 Alpha, Beta and Gamma variants<sup>14</sup>. T cell responses target multiple regions in the spike protein, consistent with the largely preserved cellular immune responses to Omicron<sup>6,14</sup>. The 82–84% cross-reactivity of CD8 $^{+}$  T cell responses to Omicron is also consistent with theoretical predictions based on the Omicron mutations. Limitations of our study include the use of high doses of peptides with costimulation in the intracellular cytokine staining assays, and the lack of assessing the effect of mutations on antigen processing.

Preclinical studies have shown that CD8 $^{+}$  T cells contribute to protection against SARS-CoV-2 in rhesus macaques, particularly when antibody responses are suboptimal<sup>5</sup>. Durable CD8 $^{+}$  and CD4 $^{+}$  T cell responses have also been reported following infection and vaccination<sup>2,3,4,6,11,13,15,16</sup>. Given the role of CD8 $^{+}$  T cells in clearance of viral infections, it is likely that cellular immunity contributes substantially to vaccine protection against severe SARS-CoV-2 disease. This may be particularly relevant for Omicron, which dramatically evades neutralizing antibody responses. Recent studies have shown that Ad26.COV2.S and BNT162b2 provided 85% and 70% protection, respectively, against hospitalization due to the Omicron variant in South Africa<sup>7,8</sup>. Our data provide immunological context for the observation that current vaccines still provide robust protection against severe disease due to the SARS-CoV-2 Omicron variant despite substantially reduced neutralizing antibody responses.

## Methods

## **Study population**

Samples from individuals who received the BNT162b2 vaccine were obtained from the Beth Israel Deaconess Medical Center (BIDMC) specimen biorepository. Samples from individuals who received Ad26.COV2.S were obtained from the COV1001 study (NCT04436276). Both studies were approved by the BIDMC Institutional Review Board. All participants provided informed consent. Individuals were excluded from this study if they had a history of SARS-CoV-2 infection, received other COVID-19 vaccines or received immunosuppressive medications.

## **Pseudovirus NAb assay**

The SARS-CoV-2 pseudoviruses expressing a luciferase reporter gene were used to measure pseudovirus NAbs. In brief, the packaging construct psPAX2 (AIDS Resource and Reagent Program), the luciferase reporter plasmid pLenti-CMV Puro-Luc (Addgene) and the spike protein expressing pcDNA3.1-SARS-CoV-2 S $\Delta$ CT were co-transfected into HEK293T cells (American Type Culture Collection (ATCC) CRL\_3216) with Lipofectamine 2000 (Thermo Fisher Scientific). Pseudoviruses of SARS-CoV-2 variants were generated by using the WA1/2020 strain (Wuhan/WIV04/2019, GISAID accession ID: EPI\_ISL\_402124), the B.1.1.7 variant (Alpha, GISAID accession ID: EPI\_ISL\_601443), the B.1.351 variant (Beta, GISAID accession ID: EPI\_ISL\_712096), the B.1.617.2 variant (Delta, GISAID accession ID: EPI\_ISL\_2020950) or the B.1.1.529 variant (Omicron, GISAID ID: EPI\_ISL\_7358094.2). The supernatants containing the pseudotype viruses were collected 48 h after transfection; pseudotype viruses were purified by filtration with a 0.45- $\mu$ m filter. To determine the neutralization activity of human serum, HEK293T-hACE2 cells were seeded in 96-well tissue culture plates at a density of  $1.75 \times 10^4$  cells per well overnight. Threefold serial dilutions of heat-inactivated serum samples were prepared and mixed with 50  $\mu$ l of pseudovirus. The mixture was incubated at 37 °C for 1 h before adding to HEK293T-hACE2 cells. After 48 h, cells were lysed in a Steady-Glo Luciferase Assay (Promega) according to the manufacturer's instructions. SARS-CoV-2 neutralization titres were defined as the sample dilution at which a 50% reduction (NT50) in relative light units was observed relative to the average of the virus control wells.

## **ELISA**

SARS-CoV-2 spike receptor-binding domain (RBD)-specific binding antibodies in serum were assessed by ELISA. Ninety-six-well plates were coated with 2  $\mu$ g ml $^{-1}$  of similarly produced SARS-CoV-2 WA1/2020, B.1.617.2 (Delta), B.1.351 (Beta) or B.1.1.529 (Omicron) RBD protein in 1 $\times$  Dulbecco's phosphate-buffered saline (DPBS) and incubated at 4 °C overnight. Assay performance was similar for these four

RBD proteins. After incubation, plates were washed once with wash buffer (0.05% Tween 20 in 1× DPBS) and blocked with 350 µl of casein block solution per well for 2–3 h at room temperature. Following incubation, block solution was discarded and plates were blotted dry. Serial dilutions of heat-inactivated serum diluted in casein block were added to wells, and plates were incubated for 1 h at room temperature, before three more washes and a 1-h incubation with a 1:4,000 dilution of anti-human IgG horseradish peroxidase (Invitrogen, Thermo Fisher Scientific) at room temperature in the dark. Plates were washed three times, and 100 µl of SeraCare KPL TMB SureBlue Start solution was added to each well; plate development was halted by adding 100 µl of SeraCare KPL TMB Stop solution per well. The absorbance at 450 nm, with a reference at 650 nm, was recorded with a VersaMax microplate reader (Molecular Devices). For each sample, the ELISA end point titre was calculated using a four-parameter logistic curve fit to calculate the reciprocal serum dilution that yields a corrected absorbance value (450–650 nm) of 0.2. Interpolated end point titres were reported.

### **Enzyme-linked immunospot assay**

Peptide pools were 16 amino acid peptides overlapping by 11 amino acids spanning the SARS-CoV-2 WA1/2020, B.1.617.2 (Delta) or B.1.1.529 (Omicron; GISAID ID: EPI\_ISL\_7358094.2) spike proteins (21st Century Biochemicals). Enzyme-linked immunospot (ELISPOT) plates were coated with mouse anti-human IFN $\gamma$  monoclonal antibody from MabTech at 1 µg per well and incubated overnight at 4 °C. Plates were washed with DPBS and blocked with R10 media (RPMI with 10% heat-inactivated FBS with 1% of 100× penicillin–streptomycin, 1 M HEPES, 100 mM sodium pyruvate, 200 mM l-glutamine and 0.1% of 55 mM 2-mercaptoethanol) for 2–4 h at 37 °C. SARS-CoV-2 pooled S peptides from SARS-CoV-2 WA1/2020, B.1.617.2 (Delta) or B.1.1.529 (Omicron) (21st Century Biochemicals) were prepared and plated at a concentration of 2 µg per well, and 100,000 cells per well were added to the plate. The peptides and cells were incubated for 15–20 h at 37 °C. All steps following this incubation were performed at room temperature. The plates were washed with ELISPOT wash buffer and incubated for 2–4 h with biotinylated mouse anti-human IFN $\gamma$  monoclonal antibody from MabTech (1 µg ml<sup>-1</sup>). The plates were washed a second time and incubated for 2–3 h with conjugated goat anti-biotin AP from Rockland, Inc. (1.33 µg ml<sup>-1</sup>). The final wash was followed by the addition of nitro-blue tetrazolium chloride/5-bromo-4-chloro 3 ‘indolylphosphate p-toluidine salt (NBT/BCIP chromagen) substrate solution for 7 min. The chromagen was discarded and the plates were washed with water and dried in a dim place for 24 h. Plates were scanned and counted on a Cellular Technologies Limited Immunospot Analyzer.

### **Intracellular cytokine staining assay**

CD4<sup>+</sup> and CD8<sup>+</sup> T cell responses were quantitated by pooled peptide-stimulated intracellular cytokine staining (ICS) assays. Peptide pools were 16 amino acid peptides overlapping by 11 amino acids spanning the SARS-CoV-2 WA1/2020, B.1.617.2 (Delta) or B.1.1.529 (Omicron; GISAID ID: EPI\_ISL\_7358094.2) spike proteins (21st Century Biochemicals). 10<sup>6</sup> peripheral blood mononuclear cells were resuspended in 100 µl of R10 media supplemented with CD49d monoclonal antibody (1 µg ml<sup>-1</sup>) and CD28 monoclonal antibody (1 µg ml<sup>-1</sup>). Each sample was assessed with mock (100 µl of R10 plus 0.5% DMSO; background control), peptides (2 µg ml<sup>-1</sup>) and/or 10 pg ml<sup>-1</sup> phorbol myristate acetate (PMA) and 1 µg ml<sup>-1</sup> ionomycin (Sigma-Aldrich) (100 µl; positive control) and incubated at 37 °C for 1 h. After incubation, 0.25 µl of GolgiStop (BD) and 0.25 µl of GolgiPlug (BD) in 50 µl of R10 was added to each well and incubated at 37 °C for 8 h and then held at 4 °C overnight. The next day, the cells were washed twice with DPBS, stained with aqua live/dead dye for 10 min and then stained with predetermined titres of monoclonal antibodies to CD279 (clone EH12.1, BB700), CD4 (clone L200, BV711), CD27 (clone M-T271, BUV563), CD8 (clone SK1, BUV805) and CD45RA (clone 5H9, APC H7) for 30 min. Cells were then washed twice with 2% FBS/DPBS buffer and incubated for 15 min with 200 µl of BD CytoFix/CytoPerm Fixation/Permeabilization solution. Cells were washed twice with 1X Perm Wash buffer (BD Perm/Wash Buffer 10X in the CytoFix/CytoPerm Fixation/Permeabilization kit diluted with MilliQ water and passed through a 0.22-µm filter) and stained intracellularly with monoclonal antibodies to Ki67 (clone B56, BB515), IL-21 (clone 3A3-N2.1, PE), CD69 (clone TP1.55.3, ECD), IL-10 (clone JES3-9D7, PE CY7), IL-13 (clone JES10-5A2, BV421), IL-4 (clone MP4-25D2, BV605), TNF (clone Mab11, BV650), IL-17 (clone N49-653, BV750), IFNγ (clone B27, BUV395), IL-2 (clone MQ1-17H12, BUV737), IL-6 (clone MQ2-13A5, APC) and CD3 (clone SP34.2, Alexa 700) for 30 min. Cells were washed twice with 1X Perm Wash buffer and fixed with 250 µl of freshly prepared 1.5% formaldehyde. Fixed cells were transferred to a 96-well round bottom plate and analysed by BD FACSymphony system. Data were analysed using FlowJo v9.9.

## Statistical analysis

Descriptive statistics and logistic regression were performed using GraphPad Prism 8.4.3, (GraphPad Software). Immunological data were generated in duplicate and were compared by Mann–Whitney tests. Correlations were evaluated by linear regression.  $P < 0.05$  were considered significant.

## Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

## Data availability

All data are available in the paper or the supplementary material.

## References

1. Cele, S. et al. Omicron extensively but incompletely escapes Pfizer BNT162b2 neutralization. *Nature* **602**, 654–656 (2022).
2. Dan, J. M. et al. Immunological memory to SARS-CoV-2 assessed for up to 8 months after infection. *Science* **371**, eabf4063 (2021).
3. Sette, A. & Crotty, S. Adaptive immunity to SARS-CoV-2 and COVID-19. *Cell* **184**, 861–880 (2021).
4. Goel, R. R. et al. mRNA vaccines induce durable immune memory to SARS-CoV-2 and variants of concern. *Science* **374**, abm0829 (2021).
5. McMahan, K. et al. Correlates of protection against SARS-CoV-2 in rhesus macaques. *Nature* **590**, 630–634 (2021).
6. Grifoni, A. et al. Targets of T cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals. *Cell* **181**, 1489–1501.e15 (2020).
7. Collie, S., Champion, J., Moultrie, H., Bekker, L. G. & Gray, G. Effectiveness of BNT162b2 vaccine against Omicron variant in South Africa. *N. Engl. J. Med.* **386**, 494–496 (2022).
8. Gray, G. E. et al. Vaccine effectiveness against hospital admission in South African health care workers who received a homologous booster of Ad26.COV2 during an Omicron COVID 19 wave: preliminary results of the Sisonke 2 Study. Preprint at *medRxiv* <https://doi.org/10.1101/2021.12.28.21268436> (2021).
9. Sadoff, J. et al. Safety and efficacy of single-dose Ad26.COV2.S vaccine against Covid-19. *N. Engl. J. Med.* **384**, 2187–2201 (2021).
10. Polack, F. P. et al. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. *N. Engl. J. Med.* **383**, 2603–2615 (2020).
11. Collier, A. Y. et al. Differential kinetics of immune responses elicited by Covid-19 vaccines. *N. Engl. J. Med.* **385**, 2010–2012 (2021).

12. Falsey, A. R. et al. SARS-CoV-2 neutralization with BNT162b2 vaccine dose 3. *N. Engl. J. Med.* **385**, 1627–1629 (2021).
13. Barouch, D. H. et al. Durable humoral and cellular immune responses 8 months after Ad26.COV2.S vaccination. *N. Engl. J. Med.* **385**, 951–953 (2021).
14. Alter, G. et al. Immunogenicity of Ad26.COV2.S vaccine against SARS-CoV-2 variants in humans. *Nature* **596**, 268–272 (2021).
15. Poon, M. M. L. et al. SARS-CoV-2 infection generates tissue-localized immunological memory in humans. *Sci. Immunol.* **6**, eabl9105 (2021).
16. Vidal, S. J. et al. Correlates of neutralization against SARS-CoV-2 variants of concern by early pandemic sera. *J. Virol.* **95**, e0040421 (2021).

## Acknowledgements

We acknowledge NIH grant CA260476, the Massachusetts Consortium for Pathogen Readiness, the Ragon Institute and the Musk Foundation (D.H.B.), as well as the Reproductive Scientist Development Program from the Eunice Kennedy Shriver National Institute of Child Health & Human Development and Burroughs Wellcome Fund HD000849 (A.Y.C.). The COV1001 study was funded in part by Biomedical Advanced Research and Development Authority HHSO100201700018C. We thank MesoScale Discovery for providing the ECLA kits.

## Author information

### Author notes

1. These authors contributed equally: Jinyan Liu, Abishek Chandrashekhar, Daniel Sellers

## Affiliations

1. Beth Israel Deaconess Medical Center, Boston, MA, USA

Jinyan Liu, Abishek Chandrashekhar, Daniel Sellers, Julia Barrett, Catherine Jacob-Dolan, Michelle Lifton, Katherine McMahan, Michaela Sciacca, Haley VanWyk, Cindy Wu, Jingyou Yu, Ai-ris Y. Collier & Dan H. Barouch

2. Ragon Institute of MGH, MIT, and Harvard, Cambridge, MA, USA

Catherine Jacob-Dolan & Dan H. Barouch

## Contributions

This study was designed by D.H.B. Samples were provided by A.Y.C. Cellular immune responses were assessed by J.L., A.C., D.S., J.B., M.L., M.S., H.V. and C.W. Humoral immune responses were assessed by C.J.-D., K.M. and J.Y.

## Corresponding author

Correspondence to [Dan H. Barouch](#).

## Ethics declarations

## Competing interests

D.H.B. is a co-inventor on provisional COVID-19 vaccine patents (63/121,482, 63/133,969 and 63/135,182). All other authors declare no competing interests.

## Peer review

## Peer review information

*Nature* thanks Antonio Bertoletti, Katherine Kedzierska and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

## Additional information

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Extended data figures and tables

### [Extended Data Fig. 1 Nucleocapsid antibody responses.](#)

Nucleocapsid antibody responses at month 8 following final vaccination with Ad26.COV2.S ( $N = 20$ ) or BNT162b2 ( $N = 27$ ) by meso-scale discovery (MSD) electrochemoluminescent assay. SARS-CoV-2 convalescent and pre-pandemic samples were included as positive and negative controls, respectively. Relative light units are shown.

## **Extended Data Fig. 2 ELISA reactivity against WA1/2020, Beta, Delta, and Omicron RBD proteins.**

Positive and negative control standards were assessed by ELISA against WA1/2020, Beta, Delta, and Omicron RBD proteins. The positive control standards were known to have 2-3 fold lower antibody titers to Omicron.

## **Extended Data Fig. 3 Cellular immune responses to Omicron by ELISPOT assays.**

Spike-specific IFN- $\gamma$  ELISPOT assays at month 1 and 8 following final vaccination with Ad26.COV2.S (N = 20) or BNT162b2 (N = 27). Responses were measured against the SARS-CoV-2 WA1/2020, B.1.617.2 (Delta), and B.1.1.529 (Omicron) variants. Medians (red bars) are depicted and numerically shown.

## **Extended Data Fig. 4 Representative CD8+ T cell responses by flow cytometry.**

Representative of 47 samples is shown.

## **Extended Data Fig. 5 Representative CD4+ T cell responses by flow cytometry.**

Representative of 47 samples is shown.

## **Extended Data Fig. 6 Cellular immune responses to Omicron by intracellular cytokine staining assays.**

Spike-specific IFN- $\gamma$ , TNF- $\alpha$ , and IL-2 CD8+ and CD4+ T cell responses by intracellular cytokine staining assays at month 8 following final vaccination with Ad26.COV2.S (N = 20). Responses were measured against the SARS-CoV-2 WA1/2020, B.1.617.2 (Delta), and B.1.1.529 (Omicron) variants. Medians (red bars) are depicted and numerically shown.

## **Extended Data Table 1 Characteristics of study population**

## **Supplementary information**

### **Reporting Summary**

## Supplementary Table 1

Peptide sequences for WA1/2020 and Omicron spike.

## Rights and permissions

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

## Reprints and Permissions

## About this article

### Cite this article

Liu, J., Chandrashekhar, A., Sellers, D. *et al.* Vaccines elicit highly conserved cellular immunity to SARS-CoV-2 Omicron. *Nature* **603**, 493–496 (2022).  
<https://doi.org/10.1038/s41586-022-04465-y>

- Received: 28 December 2021
- Accepted: 25 January 2022
- Published: 31 January 2022
- Issue Date: 17 March 2022
- DOI: <https://doi.org/10.1038/s41586-022-04465-y>

## Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

## Further reading

- [\*\*Cross reactive T cells hold up against Omicron\*\*](#)

- Alexandra Flemming

*Nature Reviews Immunology* (2022)

- [\*\*T cell responses to SARS-CoV-2 spike cross-recognize Omicron\*\*](#)

- Roanne Keeton
- Marius B. Tincho
- Catherine Riou

*Nature* (2022)

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-022-04465-y>.

| [Section menu](#) | [Main menu](#) |

- Article
- [Published: 02 March 2022](#)

# Establishment of fetomaternal tolerance through glycan-mediated B cell suppression

- [G. Rizzuto](#) ORCID: [orcid.org/0000-0003-3052-0253<sup>1</sup>](https://orcid.org/0000-0003-3052-0253),
- [J. F. Brooks<sup>2</sup>](#),
- [S. T. Tuomivaara<sup>3,4,5</sup>](#),
- [T. I. McIntyre](#) ORCID: [orcid.org/0000-0001-6670-9977<sup>6</sup>](https://orcid.org/0000-0001-6670-9977),
- [S. Ma](#) ORCID: [orcid.org/0000-0002-7150-0665<sup>7</sup>](https://orcid.org/0000-0002-7150-0665),
- [D. Rideaux<sup>7</sup>](#),
- [J. Zikherman](#) ORCID: [orcid.org/0000-0002-0873-192X<sup>2,6,8</sup>](https://orcid.org/0000-0002-0873-192X),
- [S. J. Fisher<sup>3,4,6</sup>](#) &
- [A. Erlebacher](#) ORCID: [orcid.org/0000-0003-2834-5943<sup>4,6,7,8</sup>](https://orcid.org/0000-0003-2834-5943)

*Nature* volume **603**, pages 497–502 (2022)

- 5199 Accesses
- 142 Altmetric
- [Metrics details](#)

## Subjects

- [Follicular B cells](#)
- [Glycobiology](#)
- [Peripheral tolerance](#)

## Abstract

Discrimination of self from non-self is fundamental to a wide range of immunological processes<sup>1</sup>. During pregnancy, the mother does not recognize the placenta as immunologically foreign because antigens expressed by trophoblasts, the placental cells that interface with the maternal immune system, do not activate maternal T cells<sup>2</sup>. Currently, these activation defects are thought to reflect suppression by regulatory T cells<sup>3</sup>. By contrast, mechanisms of B cell tolerance to trophoblast antigens have not been identified. Here we provide evidence that glycan-mediated B cell suppression has a key role in establishing fetomaternal tolerance in mice. B cells specific for a model trophoblast antigen are strongly suppressed through CD22–LYN inhibitory signalling, which in turn implicates the sialylated glycans of the antigen as key suppressive determinants. Moreover, B cells mediate the MHC-class-II-restricted presentation of antigens to CD4<sup>+</sup> T cells, which leads to T cell suppression, and trophoblast-derived sialoglycoproteins are released into the maternal circulation during pregnancy in mice and humans. How protein glycosylation promotes non-immunogenic placental self-recognition may have relevance to immune-mediated pregnancy complications and to tumour immune evasion. We also anticipate that our findings will bolster efforts to harness glycan biology to control antigen-specific immune responses in autoimmune disease.

This is a preview of subscription content

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

## Subscribe

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

## Subscribe

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

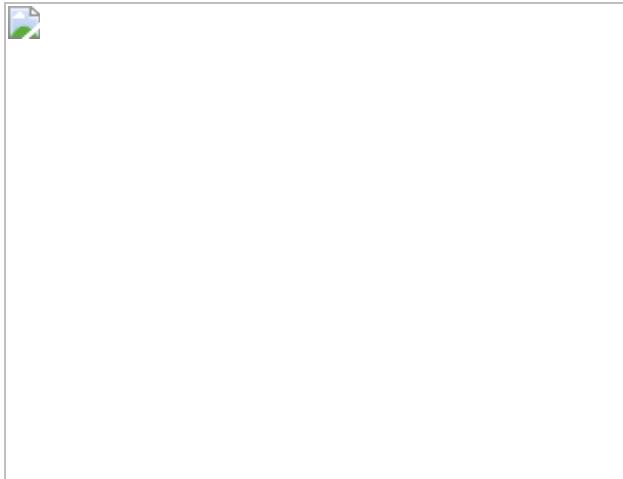
## Buy

All prices are NET prices.

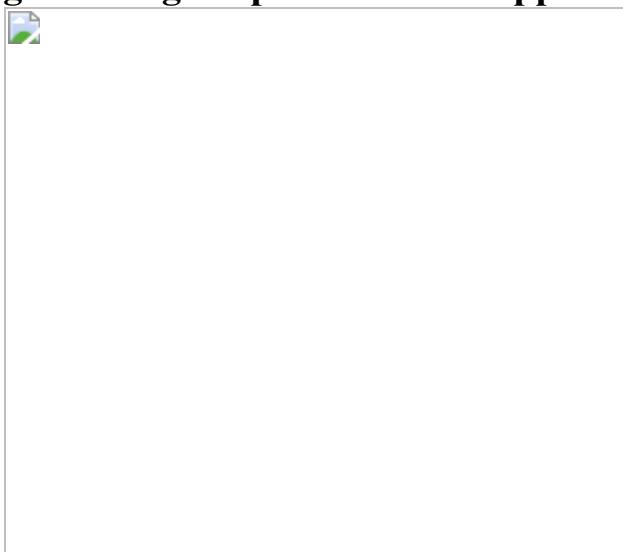
## **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

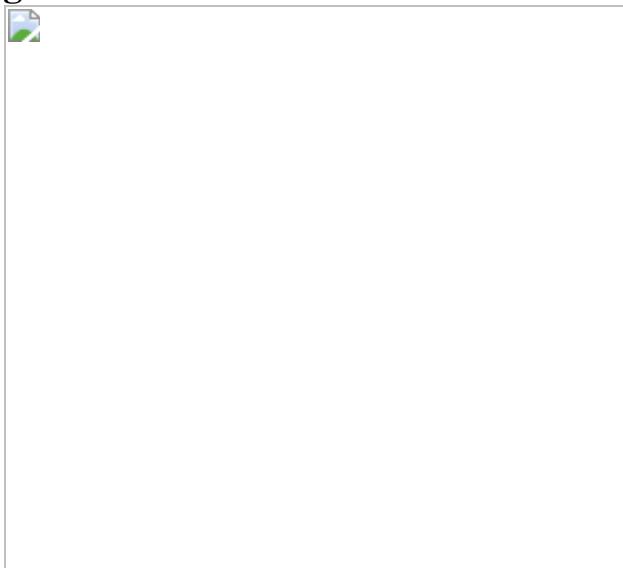
**Fig. 1: B cells present t-mOVA to CD4<sup>+</sup> T cells, which leads to impaired CD4<sup>+</sup> T-cell priming and suppressed responses to c-OVA.**



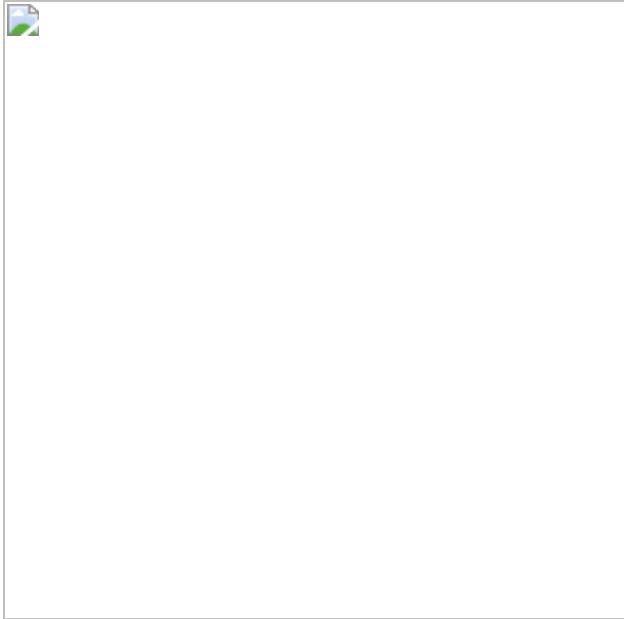
**Fig. 2: Antigen-specific B cell suppression by t-mOVA.**



**Fig. 3: CD22 and LYN mediate t-mOVA-induced B cell suppression.**



**Fig. 4: CD4<sup>+</sup> T-cell priming to t-mOVA in *Lyn*<sup>-/-</sup> mice.**



## Data availability

The MS proteomics data are publicly available within the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) under the dataset identifier [PXD029966](#). All other data generated and analysed during this study are included in this published article (and its Supplementary Information files). [Source data](#) are provided with this paper.

## References

1. Janeway, C. A. Jr. The immune system evolved to discriminate infectious nonself from noninfectious self. *Immunol. Today* **13**, 11–16 (1992).
2. Erlebacher, A. Mechanisms of T cell tolerance towards the allogeneic fetus. *Nat. Rev. Immunol.* **13**, 23–33 (2013).
3. Jiang, T. T. et al. Regulatory T cells: new keys for further unlocking the enigma of fetal tolerance and pregnancy complications. *J. Immunol.* **192**, 4949–4956 (2014).

4. Erlebacher, A., Vencato, D., Price, K. A., Zhang, D. & Glimcher, L. H. Constraints in antigen presentation severely restrict T cell recognition of the allogeneic fetus. *J. Clin. Invest.* **117**, 1399–1411 (2007).
5. Barton, B. M., Xu, R., Wherry, E. J. & Porrett, P. M. Pregnancy promotes tolerance to future offspring by programming selective dysfunction in long-lived maternal T cells. *J. Leukoc. Biol.* **101**, 975–987 (2017).
6. Jasti, S., Farahbakhsh, M., Nguyen, S., Petroff, B. K. & Petroff, M. G. Immune response to a model shared placenta/tumor-associated antigen reduces cancer risk in parous mice. *Biol. Reprod.* **96**, 134–144 (2017).
7. Kinder, J. M. et al. CD8<sup>+</sup> T cell functional exhaustion overrides pregnancy-induced fetal antigen alloimmunization. *Cell Rep.* **31**, 107784 (2020).
8. Tay, C. S., Tagliani, E., Collins, M. K. & Erlebacher, A. *Cis*-acting pathways selectively enforce the non-immunogenicity of shed placental antigen for maternal CD8 T cells. *PLoS ONE* **8**, e84064 (2013).
9. Ehst, B. D., Ingulli, E. & Jenkins, M. K. Development of a novel transgenic mouse for the study of interactions between CD4 and CD8 T cells during graft rejection. *Am. J. Transplant.* **3**, 1355–1362 (2003).
10. Filatenkov, A. A. et al. CD4 T cell-dependent conditioning of dendritic cells to produce IL-12 results in CD8-mediated graft rejection and avoidance of tolerance. *J. Immunol.* **174**, 6909–6917 (2005).
11. Moldenhauer, L. M. et al. Cross-presentation of male seminal fluid antigens elicits T cell activation to initiate the female immune response to pregnancy. *J. Immunol.* **182**, 8080–8093 (2009).
12. Rowe, J. H., Ertelt, J. M., Aguilera, M. N., Farrar, M. A. & Way, S. S. Foxp3<sup>+</sup> regulatory T cell expansion required for sustaining pregnancy compromises host defense against prenatal bacterial pathogens. *Cell Host Microbe* **10**, 54–64 (2011).

13. Rowe, J. H., Ertelt, J. M., Xin, L. & Way, S. S. Pregnancy imprints regulatory memory that sustains anergy to fetal antigen. *Nature* **490**, 102–106 (2012).
14. Bouaziz, J. D. et al. Therapeutic B cell depletion impairs adaptive and autoreactive CD4<sup>+</sup> T cell activation in mice. *Proc. Natl Acad. Sci. USA* **104**, 20878–20883 (2007).
15. Taylor, J. J. et al. Deletion and anergy of polyclonal B cells specific for ubiquitous membrane-bound self-antigen. *J. Exp. Med.* **209**, 2065–2077 (2012).
16. Harvey, D. J., Wing, D. R., Kuster, B. & Wilson, I. B. Composition of N-linked carbohydrates from ovalbumin and co-purified glycoproteins. *J. Am. Soc. Mass. Spectrom.* **11**, 564–571 (2000).
17. Perdicchio, M. et al. Sialic acid-modified antigens impose tolerance via inhibition of T-cell proliferation and de novo induction of regulatory T cells. *Proc. Natl Acad. Sci. USA* **113**, 3329–3334 (2016).
18. Meyer, S. J., Linder, A. T., Brandl, C. & Nitschke, L. B cell Siglecs—news on signaling and its interplay with ligand binding. *Front. Immunol.* **9**, 2820 (2018).
19. Hibbs, M. L. et al. Multiple defects in the immune system of Lyn-deficient mice, culminating in autoimmune disease. *Cell* **83**, 301–311 (1995).
20. Chu, C. L. & Lowell, C. A. The Lyn tyrosine kinase differentially regulates dendritic cell generation and maturation. *J. Immunol.* **175**, 2880–2889 (2005).
21. Nancy, P. et al. Chemokine gene silencing in decidual stromal cells limits T cell access to the maternal–fetal interface. *Science* **336**, 1317–1321 (2012).
22. Abeln, M. et al. Sialic acid is a critical fetal defense against maternal complement attack. *J. Clin. Invest.* **129**, 422–436 (2019).

23. Hromatka, B. S. et al. Polysialic acid enhances the migration and invasion of human cytotrophoblasts. *Glycobiology* **23**, 593–602 (2013).
24. Chen, Q. et al. Evidence for differential glycosylation of trophoblast cell types. *Mol. Cell Proteomics* **15**, 1857–1866 (2016).
25. Pfrengle, F., Macauley, M. S., Kawasaki, N. & Paulson, J. C. Copresentation of antigen and ligands of Siglec-G induces B cell tolerance independent of CD22. *J. Immunol.* **191**, 1724–1731 (2013).
26. Tokatlian, T. et al. Innate immune recognition of glycans targets HIV nanoparticle immunogens to germinal centers. *Science* **363**, 649–654 (2019).
27. Archambault, A. S. et al. Cutting edge: conditional MHC class II expression reveals a limited role for B cell antigen presentation in primary and secondary CD4 T cell responses. *J. Immunol.* **191**, 545–550 (2013).
28. Brodie, E. J., Infantino, S., Low, M. S. Y. & Tarlinton, D. M. Lyn, lupus, and (B) lymphocytes, a lesson on the critical balance of kinase signaling in immunity. *Front. Immunol.* **9**, 401 (2018).
29. Lateef, A. & Petri, M. Management of pregnancy in systemic lupus erythematosus. *Nat. Rev. Rheumatol.* **8**, 710–718 (2012).
30. Laubli, H. & Varki, A. Sialic acid-binding immunoglobulin-like lectins (Siglecs) detect self-associated molecular patterns to regulate immune responses. *Cell. Mol. Life Sci.* **77**, 593–605 (2020).
31. Hogquist, K. A. et al. T cell receptor antagonist peptides induce positive selection. *Cell* **76**, 17–27 (1994).
32. Barnden, M. J., Allison, J., Heath, W. R. & Carbone, F. R. Defective TCR expression in transgenic mice constructed using cDNA-based  $\alpha$ - and  $\beta$ -chain genes under the control of heterologous regulatory elements. *Immunol. Cell Biol.* **76**, 34–40 (1998).

33. Hildner, K. et al. Batf3 deficiency reveals a critical role for CD8 $\alpha$ <sup>+</sup> dendritic cells in cytotoxic T cell immunity. *Science* **322**, 1097–1100 (2008).
34. Kitamura, D., Roes, J., Kuhn, R. & Rajewsky, K. A B cell-deficient mouse by targeted disruption of the membrane exon of the immunoglobulin mu chain gene. *Nature* **350**, 423–426 (1991).
35. Kim, J. M., Rasmussen, J. P. & Rudensky, A. Y. Regulatory T cells prevent catastrophic autoimmunity throughout the lifespan of mice. *Nat. Immunol.* **8**, 191–197 (2007).
36. McKenna, H. J. et al. Mice lacking Flt3 ligand have deficient hematopoiesis affecting hematopoietic progenitor cells, dendritic cells, and natural killer cells. *Blood* **95**, 3489–3497 (2000).
37. Rickert, R. C., Roes, J. & Rajewsky, K. B lymphocyte-specific, Cre-mediated mutagenesis in mice. *Nucleic Acids Res.* **25**, 1317–1318 (1997).
38. Miyake, Y. et al. Critical role of macrophages in the marginal zone in the suppression of immune responses to apoptotic cell-associated antigens. *J. Clin. Invest.* **117**, 2268–2278 (2007).
39. Mason, D. Y., Jones, M. & Goodnow, C. C. Development and follicular localization of tolerant B lymphocytes in lysozyme/anti-lysozyme IgM/IgD transgenic mice. *Int. Immunol.* **4**, 163–175 (1992).
40. Lewandoski, M., Meyers, E. N. & Martin, G. R. Analysis of *Fgf8* gene function in vertebrate development. *Cold Spring Harb. Symp. Quant. Biol.* **62**, 159–168 (1997).
41. Chan, V. W., Meng, F., Soriano, P., DeFranco, A. L. & Lowell, C. A. Characterization of the B lymphocyte populations in Lyn-deficient mice and the role of Lyn in signal initiation and down-regulation. *Immunity* **7**, 69–81 (1997).

42. Otipoby, K. L. et al. CD22 regulates thymus-independent responses and the lifespan of B cells. *Nature* **384**, 634–637 (1996).
43. Tanigaki, K. et al. Endothelial Fc $\gamma$  receptor IIB activation blunts insulin delivery to skeletal muscle to cause insulin resistance in mice. *Diabetes* **65**, 1996–2005 (2016).
44. Quah, B. J., Warren, H. S. & Parish, C. R. Monitoring lymphocyte proliferation in vitro and in vivo with the intracellular fluorescent dye carboxyfluorescein diacetate succinimidyl ester. *Nat. Protoc.* **2**, 2049–2056 (2007).
45. Dahlgren, M. W. et al. T follicular helper, but not Th1, cell differentiation in the absence of conventional dendritic cells. *J. Immunol.* **194**, 5187–5199 (2015).
46. Okeke, E. B., Okwor, I. & Uzonna, J. E. Regulatory T cells restrain CD4 $^{+}$  T cells from causing unregulated immune activation and hypersensitivity to lipopolysaccharide challenge. *J. Immunol.* **193**, 655–662 (2014).
47. van Rooijen, N. & van Nieuwmeegen, R. Elimination of phagocytic cells in the spleen after intravenous injection of liposome-encapsulated dichloromethylene diphosphonate. An enzyme-histochemical study. *Cell Tissue Res.* **238**, 355–358 (1984).
48. Hong, S. et al. B cells are the dominant antigen-presenting cells that activate naive CD4 $^{+}$  T cells upon immunization with a virus-derived nanoparticle antigen. *Immunity* **49**, 695–708.e4 (2018).
49. Maley, F., Trimble, R. B., Tarentino, A. L. & Plummer, T. H. Jr. Characterization of glycoproteins and their associated oligosaccharides through the use of endoglycosidases. *Anal. Biochem.* **180**, 195–204 (1989).
50. Liu, H., Gaza-Bulseco, G., Chumsae, C. & Newby-Kew, A. Characterization of lower molecular weight artifact bands of

recombinant monoclonal IgG1 antibodies on non-reducing SDS–PAGE. *Biotechnol. Lett.* **29**, 1611–1622 (2007).

51. Drake, P. M. et al. A lectin affinity workflow targeting glycosite-specific, cancer-related carbohydrate structures in trypsin-digested human plasma. *Anal. Biochem.* **408**, 71–85 (2011).
52. Collins, B. E. et al. High-affinity ligand probes of CD22 overcome the threshold set by *cis* ligands to allow for binding, endocytosis, and killing of B cells. *J. Immunol.* **177**, 2994–3003 (2006).
53. Wu, C., Jin, X., Tsueng, G., Afrasiabi, C. & Su, A. I. BioGPS: building your own mash-up of gene annotations and expression profiles. *Nucleic Acids Res.* **44**, D313–D316 (2016).
54. Marsh, B. & Blelloch, R. Single nuclei RNA-seq of mouse placental labyrinth development. *eLife* **9**, e60266 (2020).
55. Dziadek, M. A. & Andrews, G. K. Tissue specificity of alpha-fetoprotein messenger RNA expression during mouse embryogenesis. *EMBO J.* **2**, 549–554 (1983).
56. Jones, E. A., Clement-Jones, M., James, O. F. & Wilson, D. I. Differences between human and mouse alpha-fetoprotein expression during early development. *J. Anat.* **198**, 555–559 (2001).
57. Moore, T. & Dveksler, G. S. Pregnancy-specific glycoproteins: complex gene families regulating maternal–fetal interactions. *Int. J. Dev. Biol.* **58**, 273–280 (2014).
58. Mendoza, M. et al. Glycan characterization of pregnancy-specific glycoprotein 1 and its identification as a novel Galectin-1 ligand. *Glycobiology* **30**, 895–909 (2020).
59. Zhu, B. C. et al. Human placental (fetal) fibronectin: increased glycosylation and higher protease resistance than plasma fibronectin. Presence of poly-lactosamine glycopeptides and properties of a 44-

kilodalton chymotryptic collagen-binding domain: difference from human plasma fibronectin. *J. Biol. Chem.* **259**, 3962–3970 (1984).

## Acknowledgements

This work was supported by NIH grant K08AI137209 (to G.R.); the UCSF Department of Laboratory Medicine and School of Medicine (to A.E.); and the UCSF Parnassus Flow Cytometry Core Facility with funding from DRC Center Grant NIH P30 DK063720. We thank D. Hirshhorn-Cyberman for helpful discussions, A. DeFranco for comments on the manuscript, C. Mineo for mice, D. Mueller for providing the tetramer synthesis protocol and S. Gaw for providing plasma collected from healthy mothers.

## Author information

### Affiliations

1. Department of Pathology, University of California San Francisco, San Francisco, CA, USA

G. Rizzuto

2. Department of Medicine, University of California San Francisco, San Francisco, CA, USA

J. F. Brooks & J. Zikherman

3. Department of Obstetrics, Gynecology and Reproductive Sciences, University of California San Francisco, San Francisco, CA, USA

S. T. Tuomivaara & S. J. Fisher

4. Center for Reproductive Sciences, University of California San Francisco, San Francisco, CA, USA

S. T. Tuomivaara, S. J. Fisher & A. Erlebacher

5. Sandler-Moore Mass Spectrometry Core Facility, University of California, San Francisco, CA, USA  
S. T. Tuomivaara
6. Biomedical Sciences Program, University of California San Francisco, San Francisco, CA, USA  
T. I. McIntyre, J. Zikherman, S. J. Fisher & A. Erlebacher
7. Department of Laboratory Medicine, University of California San Francisco, San Francisco, CA, USA  
S. Ma, D. Rideaux & A. Erlebacher
8. Bakar ImmunoX Initiative, University of California San Francisco, San Francisco, CA, USA  
J. Zikherman & A. Erlebacher

## Contributions

Conceptualization: G.R., S.J.F., J.Z. and A.E. Investigation: G.R., J.F.B., S.T.T., T.I.M., S.M. and D.R. Resources: S.J.F., J.Z. and A.E. Formal analysis: G.R., S.T.T. and A.E. Visualization: G.R. and A.E. Supervision: S.J.F., J.Z. and A.E. Writing (original draft): G.R. and A.E. Writing (review and editing): G.R., J.F.B., S.T.T., T.I.M., S.J.F., J.Z. and A.E.

## Corresponding author

Correspondence to [A. Erlebacher](#).

## Ethics declarations

## Competing interests

S.J.F. is a consultant for Novo Nordisk, and J.Z. is a consultant for Walking Fish Therapeutics. G.R., J.F.B., S.T.T., T.I.M., S.M., D.R. and A.E. declare no competing interests.

## Peer review

### Peer review information

*Nature* thanks Pascal Gagneux, Shiv Pillai and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

## Additional information

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Extended data figures and tables

### Extended Data Fig. 1 Experimental scheme.

To create an experimental model in which transmembrane OVA (mOVA) is expressed as a surrogate trophoblast antigen, we mated non-transgenic female mice (all C57BL/6-background) to C57BL/6 male mice hemizygous for the Act-mOVA (CAG-OVAL) transgene<sup>9</sup> (right). This creates a pregnancy in which, on average, 50% of the concepti will bear the transgene. Due to the promoter/enhancer sequences of the transgene, these transgenic concepti (red) will ubiquitously express mOVA, but there is particularly high expression levels by placental trophoblasts that have invaded into uterine blood vessels and are thus directly exposed to the maternal circulation<sup>4</sup>. Moreover, the topology of the mOVA constructs directs OVA expression to the external surface of the cell. Thus, the OVA protein itself is bathed in maternal blood. Although the mechanism remains elusive, previous work has established that mOVA is shed into the maternal circulation starting at about E10.5, and genetic experiments have

established that its presentation to maternal CD4<sup>+</sup> and CD8<sup>+</sup> T cells is mediated exclusively by maternal APCs<sup>4</sup>. As a negative control, females are mated to non-transgenic C57BL/6 males (left). At mid-gestation, the pregnant mice are injected with OVA-specific OT-I and/or OT-II TCR transgenic T cells, with or without adjuvants or c-OVA.

**Extended Data Fig. 2 xmOVA pregnant mice show a mild, OVA-dependent expansion of OVA-specific Tregs, but Treg depletion does not alter their suppressed CD4<sup>+</sup> T cell response to t-mOVA.**

Treg depletion was accomplished through use of the *Foxp3*<sup>DTR</sup> system, in which the gene for the diphtheria toxin receptor (DTR) is knocked into the X-linked *Foxp3* locus, thus rendering Tregs sensitive to diphtheria toxin-(DT-) induced ablation<sup>35</sup>. Since complete Treg ablation starting at mid-gestation is known to cause near-total pregnancy failure<sup>12</sup>, our experiments employed *Foxp3*<sup>DTRWT</sup> female mice in which, due to random X-inactivation, ~50% of CD4<sup>+</sup> T cells express a wild-type *Foxp3* allele and the other ~50% express the *DTR* knock-in allele. DT administration thus causes a ~50% acute reduction in Treg frequencies<sup>35</sup>. While this reduction is transient, it is still sufficient to induce a significant degree of fetal loss in allogeneic mating combinations<sup>12</sup>. By contrast, partial Treg ablation in the syngeneic mating combinations employed here (C57BL/6 x C57BL/6, aside from the mOVA transgene) did not induce fetal loss. To prevent the transferred OT-II cells themselves from generating an OVA-specific Treg population, we also employed OT-II *Foxp3*<sup>DTR/Y</sup> males as cell donors. All FOXP3<sup>+</sup> OT-II cells from these mice are ablatable since they all express the *Foxp3*<sup>DTR</sup> allele. **a**, Frequency of FOXP3<sup>+</sup> Treg OT-II cells among total splenic OT-II cells 6 days after adoptive transfer into virgin mice or mid-gestational (E12.5–15.5) WT or *Foxp3*<sup>DTRWT</sup> mice mated as indicated. The *Foxp3*<sup>DTRWT</sup> mice received OT-II *Foxp3*<sup>DTR/Y</sup> cells, and were injected daily with DT starting at E10.5, in line with previously work<sup>12</sup>. Note that virtually none of the transferred cells in this latter group converted into Tregs. Adjusted *P*-values were determined by ordinary one-way ANOVA with Šídák's multiple comparisons test applied to the four comparisons

shown. ns, not significant. Data were accumulated from 5 individual experiments and all mice are shown. **b**, Confirmation of partial, DT-induced depletion of host FOXP3<sup>+</sup> CD4<sup>+</sup> cells in *Foxp3*<sup>DTRWT</sup> female mice. The frequency of FOXP3<sup>+</sup> cells among total CD4<sup>+</sup> lymphocytes in the spleens of virgin *Foxp3*<sup>DTRWT</sup> female mice was measured 24 h after DT administration. *P*-value was determined by two-tailed, unpaired *t*-test. Data are from 1 experiment and all mice are shown. **c-f**, Proliferation index (**c**), fold expansion (**d**), activation marker expression (**e**), and IFN- $\gamma$  production (**f**) of CFSE-labeled *Foxp3*<sup>DTR/Y</sup> OT-II cells 6 days after adoptive transfer on E12.5–15.5 into *Foxp3*<sup>DTRWT</sup> mice mated as indicated. The mice were injected daily with DT starting on E10.5, thus partially depleting endogenous Tregs and completely ablating all OT-II cells that have converted into Tregs, as described above. Some groups received i.v. adjuvants (poly(I:C)+anti-CD40 antibodies)  $\pm$  c-OVA at the time of OT-II transfer. Adjusted *P*-values were determined by ordinary one-way ANOVA with Šídák's multiple comparisons test. Each group was compared to the xB6 control group, and the xB6+c-OVA/Adj group was compared to the xmOVA+c-OVA/Adj group. Bars show mean $\pm$ s.d. Data are from 4 independent experiments and all mice are shown.

[Source data](#)

[Extended Data Fig. 3 B cells present t-mOVA to maternal CD4<sup>+</sup> T cells; t-mOVA presentation to CD8<sup>+</sup> T cells is primarily mediated by dendritic cells.](#)

**a**, Representative flow cytometry ( $n>4$ /group) for MHCII expression on splenic CD19<sup>+</sup> B cells, CD11b<sup>+</sup> cells, and CD11c<sup>+</sup> cells in WT or *H2-Ab*<sup>f/f</sup>*Cd19-cre* mice, demonstrating that *H2-Ab*<sup>f/f</sup>*Cd19-cre* mice show loss of MHCII expression specifically on B cells. **b, c**, Proliferation index of CFSE-labeled OT-II (**b**) and OT-I (**c**) cells, measured 50 h after late-gestational transfer into xmOVA mated pregnant mice of the indicated genotypes. Adjusted *P*-values were determined by ordinary one-way ANOVA with Šídák's multiple comparisons test. Each group was compared to the WT group. Bars show mean $\pm$ s.d. Data were accumulated over 21 individual experiments and all mice are shown. **d, e**, Maternal CD4<sup>+</sup> and

CD8<sup>+</sup> T cell recognition of t-mOVA does not require macrophages; serum does not restore CD4<sup>+</sup> T cell recognition in B cell deficient mice.

Proliferation index of CFSE-labeled OT-II (**d**) and OT-I (**e**) cells, measured 50 h after late-gestational transfer into xmOVA pregnant mice of the indicated genotypes. CD169-DTR mice received a single depleting dose of DT<sup>38</sup> 48 h prior to cell transfer. Some WT mice received Clodronate Liposomes 24 h prior to cell transfer (WT + Clod Lip), and some μMT mice received 300  $\mu$ l serum from late-gestational xmOVA-mated pregnant females administered in two separate doses (150  $\mu$ l each) at 18 and 6 h prior to cell transfer (μMT + serum). Data for untreated WT and μMT mice are the same as in panels **b** and **c**. Adjusted *P*-values were determined by ordinary one-way ANOVA with Šídák's multiple comparisons test. Each group was compared to the WT group. Bars show mean±s.d. All mice are shown. WT + Clod Lip mice, CD169-DTR mice, and μMT + serum mice data were accumulated from 2, 5, and 1 independent experiment(s), respectively. **f**, *H2-Ab*<sup>fl/fl</sup>*Cd19-cre* mice retain the ability to present c-OVA to OT-II cells. Proliferation index of OT-II cells at 50 h after adoptive transfer into xB6-mated WT or *H2-Ab*<sup>fl/fl</sup>*Cd19-cre* mice on E12.5–16.5. The mice were injected i.v. with c-OVA at the time of transfer. *P*-values were determined by two-tailed, unpaired *t*-test. Bars show mean±s.d. Data were accumulated from 3 independent experiments and all mice are shown.

#### Source data

#### Extended Data Fig. 4 Identification of endogenous OVA-specific B cells using fluorescently-conjugated OVA-tetramers.

**a**, Representative flow plots showing the gating scheme used to identify OVA-specific B cells. Spleen cell suspensions from a WT mouse were stained with OVA- and control (Ctrl)-tetramers followed by magnetic bead enrichment of tetramer<sup>+</sup> cells (15, see [Methods](#)). Briefly, SSC-H/W and FSC-H/W were used to exclude doublets, DAPI allowed for dead cell exclusion, and a dump channel consisting of Gr-1, CD11c, F4/80 and Thy1.2 was used to exclude non-B cells. Gating on the OVA-Tet<sup>+</sup>Ctrl-Tet<sup>neg</sup> population identifies cells that bind to OVA and excludes those that recognize the non-OVA components of the tetramer reagent. **b**,

Confirmation that OVA-Tet<sup>+</sup>Ctrl-Tet<sup>neg</sup> B cells are OVA-specific. Whole spleen samples from WT mice were incubated in 300  $\mu$ M of monomeric BSA or OVA beginning 20 min prior to tetramer staining<sup>15</sup> (see [Methods](#)). Note the loss of the OVA-Tet<sup>+</sup>Ctrl-Tet<sup>neg</sup> population in the preparation preincubated with monomeric OVA.

### [Extended Data Fig. 5 Antigen-specific B cell recognition of t-mOVA during pregnancy.](#)

**a**, Total follicular (FO) splenic B cells from xB6 and xmOVA matings show a similar frequency of CD95<sup>hi</sup>MHCII<sup>hi</sup> cells at mid (E8.5–13.5) and late gestation (E14.5–18.5), demonstrating that the increased frequency of CD95<sup>hi</sup>MHCII<sup>hi</sup> cells among OVA-specific FO B cells in late gestation xmOVA mice (Fig. [2a,b](#)) was antigen-driven. *P*-values were determined by two-tailed, unpaired *t*-test. Data were accumulated over 22 independent experiments and all mice are displayed. **b–d**, Phenotypic analysis of non-follicular splenic B cell subsets during mid- and late-gestation. Marginal zone (MZ) and CD93<sup>+</sup> transitional splenic B cells were gated as DUMP<sup>neg</sup>CD19<sup>+</sup>IgM<sup>hi</sup>CD21/35<sup>hi</sup> and DUMP<sup>neg</sup>CD19<sup>+</sup>CD21/35<sup>lo/neg</sup>CD93<sup>+</sup> cells, respectively, and then assessed for CD95 and MHCII expression. **b**, Representative flow plots (E17.5) are from  $n = 13$  xB6 and  $n = 17$  xmOVA late gestation pregnancies. **c, d**, Quantification of CD95/MHCII expression from xB6 and xmOVA pregnancies. For OVA-specific (OVA-Tet<sup>+</sup>) cells, there was variable but significant t-mOVA-driven CD95/MHCII upregulation at late gestation (E14.5–18.5) in CD93<sup>+</sup> but not MZ B cells (left graphs). This upregulation was much less dramatic than that seen for follicular B cells (Fig. [2b](#)). A similar frequency of CD95<sup>hi</sup>MHCII<sup>hi</sup> cells was seen for the total population of CD93<sup>+</sup> and MZ B cells in xB6 and xmOVA pregnancies (right graphs). *P*-values were determined by two-tailed, unpaired *t*-test. Bars show mean $\pm$ s.d. Data were accumulated over 22 independent experiments and all mice are displayed. **e, f**, OVA-specific (OVA-Tet<sup>+</sup>) B cells slightly expand in late gestation in xmOVA pregnancies. An increase in the absolute number (**e**) and frequency (**f**) of OVA-specific B cells becomes evident in late gestation (E14.5–18.5) only in xmOVA pregnancies. *P*-values were determined by two-tailed, unpaired *t*-test. Data were accumulated over 22 independent experiments and all

mice are displayed. **g, h**, Antigen-specific B cell suppression by t-mOVA. Absolute number of total (**g**) or CD95<sup>+</sup>GL7<sup>+</sup> GC phenotype (**h**) OVA-specific B cells 6 days after i.v. vaccination with 5x10<sup>4</sup> OT-II cells ± c-OVA ± poly(I:C) (Adj) on E11.5 or E12.5.  $P_{adj}$  Adjusted *P*-values were determined by ordinary one-way ANOVA with Šídák's multiple comparisons test. Each group was compared to the xB6 control group, and the xB6+c-OVA/Adj group was compared to the xmOVA+c-OVA/Adj group. Bars show mean±s.d. The dashed line (**h**) indicates the limit of detection. This analysis employed the same samples used for Fig. [2c, d](#). Data were accumulated over 26 independent experiments and all mice are displayed.

#### Source data

#### Extended Data Fig. 6 Although reduced in comparison to the response seen in virgin mice, c-OVA elicits similar levels of OT-II cell expansion and OT-II Tfh differentiation in xB6 and xmOVA pregnant mice, and does not induce OT-II Tfr differentiation.

**a**, Representative flow plots ( $n = 5/\text{group}$ ) showing the frequency of Bcl-6<sup>hi</sup>PD-1<sup>hi</sup> and CXCR5<sup>hi</sup>PD-1<sup>hi</sup> OT-II cells among total OT-II cells five days after adoptive transfer into virgin mice or on E11.5–12.5 into WT pregnant mice mated as indicated. All groups received the vaccination protocol used in Fig. [2c, d](#) that generated a strong OVA-specific B cell response in control pregnancies (i.v. adjuvant [poly(I:C)] + c-OVA) at the time of OT-II transfer. Analysis on day 5 post-transfer was chosen for this experiment as it was the peak of OT-II expansion. **b**, Quantification of results from **(a)** showing that both the absolute number of OT-II cells and those with Tfh phenotype are diminished in pregnant mice compared to virgins. Presumably, this reflects an antigen non-specific effect of pregnancy. Adjusted *P*-values were determined by ordinary one-way ANOVA with Šídák's multiple comparisons test. Each group was compared to the virgin+c-OVA/Adj group, and the xB6+c-OVA/Adj group was compared to the xmOVA+c-OVA/Adj group. Bars show mean±s.d. Data were accumulated over 4 independent experiments and all mice are shown. **c**,

Representative flow plots showing the frequency of FOXP3<sup>+</sup> cells among Bcl-6<sup>hi</sup>PD-1<sup>hi</sup> Tfh phenotype OT-II cells for the three groups shown in (a, b).

[Source data](#)

**Extended Data Fig. 7 Additional characteristics of t-mOVA and demonstration of sialic acid at trophoblast membranes in contact with maternal blood.**

**a, b**, Shed t-mOVA is present within the non-pelletable, non-exosomal fraction of maternal plasma. Equal volumes of plasma from xB6- and xmOVA-mated mice (E18.5) were subjected to differential centrifugation (see [Methods](#)) followed by either anti-OVA immunoprecipitation and anti-OVA immunoblotting (**a**) or immunoblotting for the exosome-specific marker CD9 (**b**). We analyzed three different fractions: “clarified plasma” (i.e., the plasma after an initial low-speed (10,000xg) centrifugation), and the “sup” and “pellet” fractions from 110,000xg ultracentrifugation. Anti-OVA immunoblotting identified mOVA in both the clarified plasma and 110,000xg “sup” fractions, but not in the 110,000xg “pellet” fraction where exosomes reside, as demonstrated by the anti-CD9 immunoblotting. Data are representative of 3 (**a**) and 2 (**b**) separate experiments. **c**, Shed t-mOVA contains *N*-glycolylneuraminic acid (Neu5Gc), the sialic acid variant required for strong  $\alpha$ (2,6)-Sia binding to mouse CD22<sup>52</sup>. Equal volumes of plasma respectively pooled from 3 xB6-mated and xmOVA-mated mice (E16.5–18.5) were subjected to differential centrifugation, anti-OVA immunoprecipitation on the 110,000xg supernatant, sialidase or mock digestion as indicated, and then anti-Neu5Gc immunoblotting. c-OVA was similarly immunoprecipitated and sialidase/mock-treated, or loaded directly onto the gel. As expected, Neu5Gc is absent from c-OVA. The non-specific (n.s.) band is likely free IgG heavy chain. Data are representative of 2 separate experiments. For gel source data, see Supplementary Fig. 1. **d–i**, Distribution of  $\alpha$ (2,6)-Sia and Neu5Gc in the mouse placental labyrinth. Placental sections prepared from mice on E12.5 were stained with SNA-I (**d–f**) or anti-Neu5Gc antibodies (**g–i**). For the SNA-I control (**f**), the adjacent section was pretreated with neuraminidase A; for the anti-Neu5Gc

control (**i**), free Neu5Gc was added in with the primary antibody. Fetal and maternal blood spaces were respectively identified by the presence of DAPI<sup>+</sup> (blue counterstain) nucleated or DAPI<sup>neg</sup> enucleated RBCs, both of which are autofluorescent on the green channel. Note the SNA-I staining on trophoblast membranes in direct contact with maternal blood (arrowheads). This staining was not as continuous as the staining for Neu5Gc, which was present in all trophoblast membranes in contact with maternal blood. The small round structures showing strong Neu5Gc staining are morphologically consistent with platelets. Images are representative of 3 independent experiments.

### **Extended Data Fig. 8 CD22 and LYN mediate B cell suppression to t-mOVA.**

**a–d**, Absolute number of total (**a, c**) and GL7<sup>+</sup>CD95<sup>+</sup> GC phenotype (**b, d**) OVA-specific B cells in pregnant *Cd22*<sup>-/-</sup> and *Lyn*<sup>-/-</sup> mice 6 days after i.v. vaccination with 5x10<sup>4</sup> OT-II cells ± c-OVA ± poly(I:C) (Adj) on E11.5 or E12.5. Adjusted *P*-values were determined by ordinary one-way ANOVA with Šídák's multiple comparisons test. Each group was compared to the xB6 control group, and the xB6+c-OVA/Adj group was compared to the xmOVA+c-OVA/Adj group. Bars show mean±s.d. This analysis employed the same samples used for Fig. [3c, d](#). The *Lyn*<sup>-/-</sup> and *Cd22*<sup>-/-</sup> data were accumulated over 14 and 11 independent experiments, respectively, and all mice are displayed. See Extended Data Fig. [5g, h](#) for corresponding WT data. Note that in *Lyn*<sup>-/-</sup> but not WT nor *Cd22*<sup>-/-</sup> mice, the number of GC phenotype OVA-specific B cells in xmOVA pregnancies approaches that seen in xB6 pregnancies following c-OVA/Adj immunization. The dashed line (**b, d**) indicates the limit of detection. **e–i**, Antigen-specific B cell suppression to t-mOVA in pregnant *Fcgr2b*<sup>-/-</sup> mice. Representative flow plots (**e**), frequencies of total (**f**) or GL7<sup>+</sup>CD95<sup>+</sup> GC phenotype (**g**) OVA-specific B cells, and absolute number of total (**h**) or GL7<sup>+</sup>CD95<sup>+</sup> GC phenotype (**i**) OVA-specific B cells in pregnant WT or *Fcgr2b*<sup>-/-</sup> mice 6 days after i.v. vaccination with 5x10<sup>4</sup> OT-II cells ± c-OVA ± poly(I:C) (Adj) on E11.5 or E12.5. WT data are from Fig. [2c, d](#) and Extended Data Fig. [5g, h](#) and were accumulated over 26 independent experiments. *Fcgr2b*<sup>-/-</sup> data were accumulated over 4 independent experiments. All mice are shown.

The flow plots show OVA-tetramer<sup>+</sup> B cells gated from a fixed number of total B cells across all groups. *P*-values were determined by two-tailed, unpaired *t*-test. Bars show mean±s.d.

[Source data](#)

[Extended Data Fig. 9 Analysis of sialylated glycopeptides in pregnant mouse and human plasma.](#)

**a**, Workflow. Plasma (*n* = 3/group) from non-pregnant or pregnant mice (E15.5–17.5) and humans (collected at delivery) was first run over species-specific Multiple Affinity Removal System (MARS) columns to remove high abundance proteins, whose presence would obscure the detection of other glycopeptides. In one analysis (top), the MARS-depleted plasma was analyzed by mass spectrometry to identify endogenous proteins of potential placental origin bearing *N*-linked glycans with terminal α(2,6)-Sia. Specifically, the depleted plasma was first subjected to trypsin digestion followed by SNA-I lectin chromatography to isolate α(2,6)-Sia containing glycopeptides. These peptides were then deglycosylated using PNGase F and then subjected to LC-MS/MS. Detected peptides representing true *N*-glycosites were identified as those bearing the consensus N-X-S/T/C motif with an asparagine → aspartic acid substitution, which occurs as a result of the PNGase F digestion. In addition (bottom), the depleted plasma was subjected to SNA-I and MAL-I lectin blotting, to reveal the overall pattern and prevalence of α(2,6)- and α(2,3)-linked sialylated glycoproteins, respectively. **b**, Identification of proteins unique to pregnant plasma that contain *N*-linked glycans with α(2,6)-Sia. Peptides with *N*-glycosites present in at least one pregnant specimen but absent from all three non-pregnant specimens were tallied. For mice, there were 30 such peptides, corresponding to 26 proteins, and for humans, there were 68 such peptides, corresponding to 53 proteins. For the mouse proteins, database queries (biogps.org)<sup>53</sup> identified 7 (27% of total) of the encoding genes to be likely expressed primarily if not exclusively by the conceptus. These genes and corresponding representative peptide species are shown in the table (with the *N*-glycosites colored red), and include *Lepr* (encoding the Leptin receptor), a recently identified marker of sinusoidal trophoblast giant cells<sup>54</sup>, *Lifr* (LIF receptor alpha subunit), which is expressed by a number of

trophoblast subtypes<sup>54</sup>, *Ceacam11* and *Ceacam12*, which are expressed by spongiotrophoblasts<sup>54</sup>, and *Afp* ( $\alpha$ -fetoprotein), which is expressed by the yolk sac and fetal liver<sup>55,56</sup>. The pregnant mice were from xmOVA matings but OVA sequences were not identified, suggesting that this mass spectrometry experiment identified only a subset of the shed proteins, perhaps only those with high abundance or with favorable ionization properties. For the human proteins, we found that 7 of the encoding genes were likely expressed primarily if not exclusively by the placenta. These included *PSG1/2/7/5/9/11*, i.e., members of the *PSG* gene family encoding Pregnancy Specific Glycoproteins, which are known to be the most abundant protein species released from the trophoblasts into maternal blood<sup>57</sup>. Peptide assignments were redundant due to shared sequences. *PSG1* has recently been shown to carry primarily  $\alpha$ (2,3)-Sia with a small amount of  $\alpha$ (2,6)-Sia<sup>58</sup>. Both mouse and human pregnant plasma contained unique peptides derived from fibronectin (*Fn1* and *FNI*). While fibronectin is abundant in non-pregnant plasma, fibronectin isolated from the human placenta is more heavily glycosylated, carrying polylactosamine chains<sup>59</sup>. Of note, many of the sialoglycopeptides unique to pregnancy but not obviously derived from the conceptus appeared instead to be produced by the liver (not shown), suggesting that the endocrine state of pregnancy might systemically alter protein sialylation. Complete data is available at the ProteomeXchange under identifier PXD029966

(<http://proteomecentral.proteomexchange.org>). **c–i**, Lectin blotting.

Volumes of MARS column eluates corresponding to equivalent volumes of starting plasma were subjected to SDS-PAGE followed by lectin blotting, Neu5Gc immunoblotting, and silver staining. Although sialylated glycoproteins were more abundant in pregnant plasma specimens, it is important to emphasize that these blotting experiments alone do not demonstrate that the corresponding proteins were derived from the placenta. Many were likely derived from the liver, in accord with our mass spectrometry data. For gel source data, see Supplementary Fig. 1. The experiment was performed once.

## **Extended Data Fig. 10 LYN deficiency allows for partial OT-II cell priming in response to t-mOVA.**

**a, b**, Analysis of CD44 and CD62L expression. Representative flow plots ( $n \geq 5$  mice/group) (**a**) showing the frequency of  $\text{CD44}^{\text{hi}}\text{CD62L}^{\text{lo}}$  OT-II cells among total OT-II cells 6 days after adoptive transfer into WT or  $\text{Lyn}^{-/-}$  pregnant mice mated as indicated; the  $\text{CD44}^{\text{hi}}\text{CD62L}^{\text{lo}}$  gate was set based upon host CD4 cells for each cytometry run, which is shown in (**b**). Some groups received i.v. adjuvants (poly(I:C)+anti-CD40 antibodies)  $\pm$  c-OVA at the time of OT-II transfer. See Fig. 4c for summary data. **c–f**, Assessment of the immunogenicity of Endoglycosidase H- (Endo H-) deglycosylated c-OVA. The extent of deglycosylation in Endo H- or mock (-)-treated c-OVA was determined via Concanavalin A (Con A) lectin blotting (**c**). Proliferation index (**d**), activation marker expression (**e**), and IFN- $\gamma$  production (**f**) of CFSE-labeled OT-II cells were determined 6 days after transfer into virgin WT and  $\mu\text{MT}$  females, with 300  $\mu\text{g}$  mock-treated or deglycosylated c-OVA given i.p. together with adjuvants (poly(I:C)+anti-CD40 antibodies) on the same day as the OT-II transfer. Bars show mean $\pm$ s.d.  $P$ -values were determined by two-tailed, unpaired  $t$ -test. Data were accumulated from 3 independent experiments and all mice are shown. **g–j**, Upregulation of CD80 and CD86 by antigen-specific B cells is suppressed by t-mOVA in WT and  $\text{Lyn}^{-/-}$  pregnant mice. Frequency of  $\text{CD80}^{\text{hi}}$  or  $\text{CD86}^{\text{hi}}$  OVA-Tetramer $^+$  (**g, i**) or total (**h, j**) B cells 6 days after i.v. vaccination with  $5 \times 10^4$  OT-II cells  $\pm$  c-OVA  $\pm$  poly(I:C) (Adj) on E11.5 or E12.5. Adjusted  $P$ -values were determined by ordinary one-way ANOVA with Šídák's multiple comparisons test. Each group was compared to the control xB6 group, and the xB6+c-OVA/Adj group was compared to the xmOVA+c-OVA/Adj group. Bars show mean $\pm$ s.d. Data were accumulated over 39 independent experiments and all mice are shown.

[Source data](#)

## Supplementary information

### [Supplementary Figure 1](#)

Full (uncropped) gels for Fig. 3a–c, and Extended Data Figs. 7a–c, 9c–i and 10c.

## Reporting Summary

## Peer Review File

## **Source data**

[Source Data Fig. 1](#)

[Source Data Fig. 2](#)

[Source Data Fig. 3](#)

[Source Data Fig. 4](#)

[Source Data Extended Data Fig. 2](#)

[Source Data Extended Data Fig. 3](#)

[Source Data Extended Data Fig. 5](#)

[Source Data Extended Data Fig. 6](#)

[Source Data Extended Data Fig. 8](#)

[Source Data Extended Data Fig. 10](#)

## **Rights and permissions**

[Reprints and Permissions](#)

## **About this article**

**Cite this article**

Rizzuto, G., Brooks, J.F., Tuomivaara, S.T. *et al.* Establishment of fetomaternal tolerance through glycan-mediated B cell suppression. *Nature* **603**, 497–502 (2022). <https://doi.org/10.1038/s41586-022-04471-0>

- Received: 22 June 2021
- Accepted: 25 January 2022
- Published: 02 March 2022
- Issue Date: 17 March 2022
- DOI: <https://doi.org/10.1038/s41586-022-04471-0>

## Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-022-04471-0>

- Article
- [Published: 09 March 2022](#)

# Ribosome collisions induce mRNA cleavage and ribosome rescue in bacteria

- [Kazuki Saito](#)<sup>1 na1</sup>,
- [Hanna Kratzat](#) [ORCID: orcid.org/0000-0001-5849-6131](#)<sup>2 na1</sup>,
- [Annabelle Campbell](#)<sup>1</sup>,
- [Robert Buschauer](#) [ORCID: orcid.org/0000-0002-5794-4841](#)<sup>2</sup>,
- [A. Maxwell Burroughs](#) [ORCID: orcid.org/0000-0002-2229-8771](#)<sup>3</sup>,
- [Otto Berninghausen](#) [ORCID: orcid.org/0000-0002-9255-0522](#)<sup>2</sup>,
- [L. Aravind](#)<sup>3</sup>,
- [Rachel Green](#) [ORCID: orcid.org/0000-0001-9337-2003](#)<sup>1,4</sup>,
- [Roland Beckmann](#) [ORCID: orcid.org/0000-0003-4291-3898](#)<sup>2</sup> &
- [Allen R. Buskirk](#) [ORCID: orcid.org/0000-0003-2720-6896](#)<sup>1</sup>

*Nature* volume 603, pages 503–508 (2022)

- 6647 Accesses
- 170 Altmetric
- [Metrics details](#)

## Subjects

- [Cryoelectron microscopy](#)
- [Ribosome](#)

# Abstract

Ribosome rescue pathways recycle stalled ribosomes and target problematic mRNAs and aborted proteins for degradation<sup>1,2</sup>. In bacteria, it remains unclear how rescue pathways distinguish ribosomes stalled in the middle of a transcript from actively translating ribosomes<sup>3,4,5,6</sup>. Here, using a genetic screen in *Escherichia coli*, we discovered a new rescue factor that has endonuclease activity. SmrB cleaves mRNAs upstream of stalled ribosomes, allowing the ribosome rescue factor tmRNA (which acts on truncated mRNAs<sup>3</sup>) to rescue upstream ribosomes. SmrB is recruited to ribosomes and is activated by collisions. Cryo-electron microscopy structures of collided disomes from *E. coli* and *Bacillus subtilis* show distinct and conserved arrangements of individual ribosomes and the composite SmrB-binding site. These findings reveal the underlying mechanisms by which ribosome collisions trigger ribosome rescue in bacteria.

This is a preview of subscription content

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

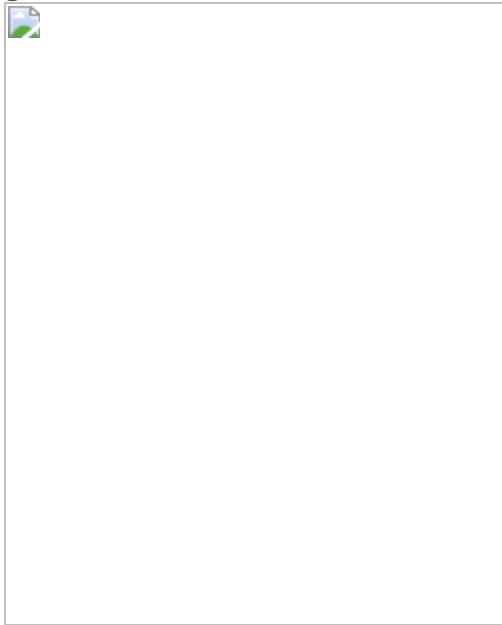
[Buy](#)

All prices are NET prices.

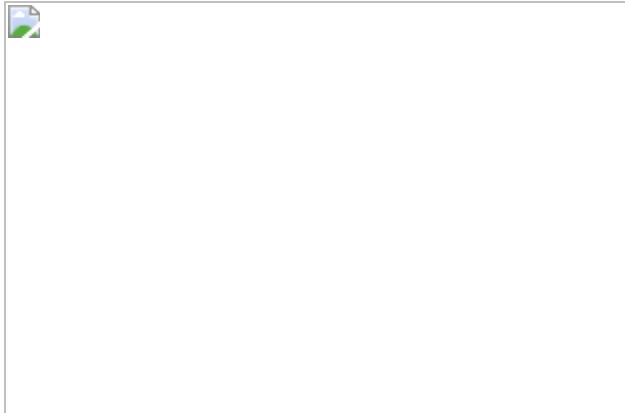
**Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

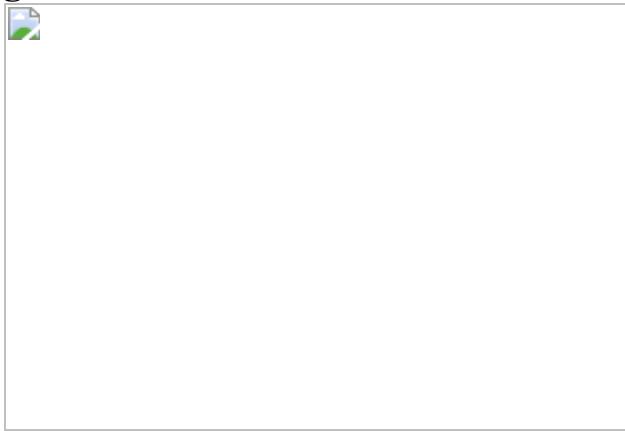
**Fig. 1: SmrB is a ribosome rescue factor.**



**Fig. 2: SmrB cleavage at the 5' boundary of stalled ribosomes promotes ribosome rescue.**



**Fig. 3: SmrB acts on collided ribosomes.**



**Fig. 4: Cryo-EM structure of the *E. coli* disome.**



**Fig. 5: Cryo-EM structure of the SmrB-bound *E. coli* disome.**



## Data availability

Cryo-EM volumes and molecular models have been deposited at the Electron Microscopy Data Bank (EMDB) and Protein Data Bank (PDB) with the following respective accession codes: for the *E. coli* disome, EMD-[13952](#) and [7QG8](#) (stalled 70S) and EMD-[13955](#) and [7QGH](#) (collided 70S); for the *E. coli* trisome, EMD-[13964](#); for the *B. subtilis* disome, EMD-[13959](#) and [7QGU](#) (stalled 70S) and EMD-[13961](#) and [7QH4](#) (collided 70S); for the *E. coli* disome–SmrB complex, EMD-[13956](#) and [7QGN](#) (stalled 70S) and EMD-[13958](#) and [7QGR](#) (collided 70S). Gel source images are provided in Supplementary Figs. [1](#) and [2](#).

## Code availability

Custom Python scripts used to analyse the Tn-seq and RACE data are freely available at [https://github.com/greenlabjhmi/2021\\_SmrB](https://github.com/greenlabjhmi/2021_SmrB).

## References

1. Müller, C., Crowe-McAuliffe, C. & Wilson, D. N. Ribosome rescue pathways in bacteria. *Front. Microbiol.* **12**, 652980 (2021).

2. Keiler, K. C., Waller, P. R. & Sauer, R. T. Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science* **271**, 990–993 (1996).
3. Ivanova, N., Pavlov, M. Y., Felden, B. & Ehrenberg, M. Ribosome rescue by tmRNA requires truncated mRNAs. *J. Mol. Biol.* **338**, 33–41 (2004).
4. Hayes, C. S. & Sauer, R. T. Cleavage of the A site mRNA codon during ribosome pausing provides a mechanism for translational quality control. *Mol. Cell* **12**, 903–911 (2003).
5. Subramaniam, A. R., Zid, B. M. & O’Shea, E. K. An integrated approach reveals regulatory controls on bacterial translation elongation. *Cell* **159**, 1200–1211 (2014).
6. Janssen, B. D., Garza-Sánchez, F. & Hayes, C. S. A-site mRNA cleavage is not required for tmRNA-mediated ssrA-peptide tagging. *PLoS ONE* **8**, e81319 (2013).
7. Moore, S. D. & Sauer, R. T. Ribosome rescue: tmRNA tagging activity and capacity in *Escherichia coli*. *Mol. Microbiol.* **58**, 456–466 (2005).
8. Yan, L. L. & Zaher, H. S. How do cells cope with RNA damage and its consequences? *J. Biol. Chem.* **294**, 15158–15171 (2019).
9. Thomas, E. N., Kim, K. Q., McHugh, E. P., Marcinkiewicz, T. & Zaher, H. S. Alkylative damage of mRNA leads to ribosome stalling and rescue by *trans* translation in bacteria. *eLife* **9**, e61984 (2020).
10. Roche, E. D. & Sauer, R. T. SsrA-mediated peptide tagging caused by rare codons and tRNA scarcity. *EMBO J.* **18**, 4579–4589 (1999).
11. Sunohara, T., Jojima, K., Tagami, H., Inada, T. & Aiba, H. Ribosome stalling during translation elongation induces cleavage of mRNA being translated in *Escherichia coli*. *J. Biol. Chem.* **279**, 15368–15375 (2004).

12. Hayes, C. S., Bose, B. & Sauer, R. T. Proline residues at the C terminus of nascent chains induce SsrA tagging during translation termination. *J. Biol. Chem.* **277**, 33825–33832 (2002).
13. Neubauer, C., Gillet, R., Kelley, A. C. & Ramakrishnan, V. Decoding in the absence of a codon by tmRNA and SmpB in the ribosome. *Science* **335**, 1366–1369 (2012).
14. Ikeuchi, K. et al. Collided ribosomes form a unique structural interface to induce Hel2-driven quality control pathways. *EMBO J.* **38**, e100276 (2019).
15. Juszkiewicz, S. et al. ZNF598 is a quality control sensor of collided ribosomes. *Mol. Cell* **72**, 469–481 (2018).
16. Matsuo, Y. et al. Ubiquitination of stalled ribosome triggers ribosome-associated quality control. *Nat. Commun.* **8**, 159 (2017).
17. Simms, C. L., Yan, L. L. & Zaher, H. S. Ribosome collision is critical for quality control during no-go decay. *Mol. Cell* **68**, 361–373 (2017).
18. Langridge, G. C. et al. Simultaneous assay of every *Salmonella Typhi* gene using one million transposon mutants. *Genome Res.* **19**, 2308–2316 (2009).
19. Kannan, K. et al. The general mode of translation inhibition by macrolide antibiotics. *Proc. Natl Acad. Sci. USA* **111**, 15958–15963 (2014).
20. Beckert, B. et al. Structural and mechanistic basis for translation inhibition by macrolide and ketolide antibiotics. *Nat. Commun.* **12**, 4466 (2021).
21. Zhou, W. et al. PPR-SMR protein SOT1 has RNA endonuclease activity. *Proc. Natl Acad. Sci. USA* **114**, E1554–E1563 (2017).
22. Wu, W. J. et al. SOT1, a pentatricopeptide repeat protein with a small MutS-related domain, is required for correct processing of plastid 23S-

- 4.5S rRNA precursors in *Arabidopsis thaliana*. *Plant J.* **85**, 607–621 (2016).
23. D’Orazio, K. N. et al. The endonuclease Cue2 cleaves mRNAs at stalled ribosomes during no go decay. *eLife* **8**, e49117 (2019).
  24. Liu, S., Melonek, J., Boykin, L. M., Small, I. & Howell, K. A. PPR-SMRs: ancient proteins with enigmatic functions. *RNA Biol.* **10**, 1501–1510 (2013).
  25. Glover, M. L. et al. NONU-1 encodes a conserved endonuclease required for mRNA translation surveillance. *Cell Rep.* **30**, 4321–4331 (2020).
  26. Mohammad, F., Green, R. & Buskirk, A. R. A systematically-revised ribosome profiling method for bacteria reveals pauses at single-codon resolution. *eLife* **8**, e42591 (2019).
  27. Chiba, S., Lamsa, A. & Pogliano, K. A ribosome-nascent chain sensor of membrane protein biogenesis in *Bacillus subtilis*. *EMBO J.* **28**, 3461–3475 (2009).
  28. Ishii, E. et al. Nascent chain-monitored remodeling of the Sec machinery for salinity adaptation of marine bacteria. *Proc. Natl Acad. Sci. USA* **112**, E5513–E5522 (2015).
  29. Su, T. et al. The force-sensing peptide VemP employs extreme compaction and secondary structure formation to induce ribosomal stalling. *eLife* **6**, e25642 (2017).
  30. Schuwirth, B. S. et al. Structures of the bacterial ribosome at 3.5 Å resolution. *Science* **310**, 827–834 (2005).
  31. Selmer, M., Gao, Y. G., Weixlbaumer, A. & Ramakrishnan, V. Ribosome engineering to promote new crystal forms. *Acta Crystallogr. D* **68**, 578–583 (2012).

32. Atkins, J. F. & Bjork, G. R. A gripping tale of ribosomal frameshifting: extragenic suppressors of frameshift mutations spotlight P-site realignment. *Microbiol. Mol. Biol. R.* **73**, 178 (2009).
33. Beckert, B. et al. Structure of the *Bacillus subtilis* hibernating 100S ribosome reveals the basis for 70S dimerization. *EMBO J.* **36**, 2061–2072 (2017).
34. Beckert, B. et al. Structure of a hibernating 100S ribosome reveals an inactive conformation of the ribosomal protein S1. *Nat. Microbiol.* **3**, 1115 (2018).
35. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
36. Ferrin, M. A. & Subramaniam, A. R. Kinetic modeling predicts a stimulatory role for ribosome collisions at elongation stall sites in bacteria. *eLife* **6**, e23629 (2017).
37. Simms, C. L., Yan, L. L., Qiu, J. K. & Zaher, H. S. Ribosome collisions result in +1 frameshifting in the absence of no-go decay. *Cell Rep.* **28**, 1679–1689 (2019).
38. Smith, A. M., Costello, M. S., Kettring, A. H., Wingo, R. J. & Moore, S. D. Ribosome collisions alter frameshifting at translational reprogramming motifs in bacterial mRNAs. *Proc. Natl Acad. Sci. USA* **116**, 21769–21779 (2019).
39. Chai, Q. et al. Organization of ribosomes and nucleoids in *Escherichia coli* cells during growth and in quiescence. *J. Biol. Chem.* **289**, 11342–11352 (2014).
40. Datsenko, K. A. & Wanner, B. L. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl Acad. Sci. USA* **97**, 6640–6645 (2000).
41. Crameri, A., Whitehorn, E. A., Tate, E. & Stemmer, W. P. C. Improved green fluorescent protein by molecular evolution using DNA shuffling.

*Nat. Biotechnol.* **14**, 315–319 (1996).

42. Jiang, H. S., Lei, R., Ding, S. W. & Zhu, S. F. Skewer: a fast and accurate adapter trimmer for next-generation sequencing paired-end reads. *BMC Bioinformatics* **15**, 182 (2014).
43. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet J.* **17**, 10–12 (2011).
44. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
45. Altschul, S. F. et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
46. Eddy, S. R. A new generation of homology search tools based on probabilistic inference. *Genome Inform.* **23**, 205–211 (2009).
47. Finn, R. D. et al. The Pfam protein families database: towards a more sustainable future. *Nucleic Acids Res.* **44**, D279–D285 (2016).
48. Lassmann, T., Frings, O. & Sonnhammer, E. L. L. Kalign2: high-performance multiple alignment of protein and nucleotide sequences allowing external features. *Nucleic Acids Res.* **37**, 858–865 (2009).
49. Cole, C., Barber, J. D. & Barton, G. J. The Jpred 3 secondary structure prediction server. *Nucleic Acids Res.* **36**, W197–W201 (2008).
50. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5**, e9490 (2010).
51. Shevchenko, A., Wilm, M., Vorm, O. & Mann, M. Mass spectrometric sequencing of proteins from silver stained polyacrylamide gels. *Anal. Chem.* **68**, 850–858 (1996).

52. Sohmen, D. et al. Structure of the *Bacillus subtilis* 70S ribosome reveals the basis for species-specific stalling. *Nat. Commun.* **6**, 6941 (2015).
53. Schafer, H. et al. The alarmones (p)ppGpp are part of the heat shock response of *Bacillus subtilis*. *PLoS Genet.* **16**, e1008275 (2020).
54. Wells, J. N. et al. Structure and function of yeast Ls02 and human CCDC124 bound to hibernating ribosomes. *PLoS Biol.* **18**, e3000780 (2020).
55. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
56. Rohou, A. & Grigorieff, N. CTFFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
57. Zivanov, J., Nakane, T. & Scheres, S. H. W. Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in RELION-3.1. *IUCrJ* **7**, 253–267 (2020).
58. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
59. Loveland, A. B., Demo, G. & Korostelev, A. A. Cryo-EM of elongating ribosome with EFTu-GTP elucidates tRNA proofreading. *Nature* **584**, 640–645 (2020).
60. Byrne, R. T., Konevega, A. L., Rodnina, M. V. & Antson, A. A. The crystal structure of unmodified tRNA<sup>Phe</sup> from *Escherichia coli*. *Nucleic Acids Res.* **38**, 4154–4162 (2010).
61. Loveland, A. B. & Korostelev, A. A. Structural dynamics of protein S1 on the 70S ribosome visualized by ensemble cryo-EM. *Methods* **137**, 55–66 (2018).

62. Mirdita, M., Steinegger, M. & Soding, J. MMseqs2 desktop and local web server app for fast, interactive sequence searches. *Bioinformatics* **35**, 2856–2858 (2019).
63. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
64. Goddard, T. D. et al. UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).

## Acknowledgements

We thank S. Sanyal for sharing *E. coli* strains QC101 and QC901, H. Hao at the JHMI Transcriptomics and Deep Sequencing Core for providing assistance with high-throughput sequencing, B. Cole and T. Boronina at JHMI in the Mass Spectrometry and Proteomics Facility, J. Musial for providing assistance during protein purification, T. Mackens-Kiani for helping with the nuclease assay data analysis, C. Ungewickell and S. Rieder for providing technical assistance, and L. Kater and K. Best for offering support with the preprocessing pipeline for the cryo-EM data. This work was supported by NIH grant GM136960 (A.R.B.), HHMI (R.G.), the Intramural Research Program of the National Library of Medicine at the NIH (A.M.B. and L.A.) and the German Research Council (TRR174) (R. Beckmann). H.K. is supported by a DFG fellowship through the Graduate School of Quantitative Bioscience Munich (QBM).

## Author information

### Author notes

1. These authors contributed equally: Kazuki Saito, Hanna Kratzat

## Affiliations

1. Department of Molecular Biology and Genetics, Johns Hopkins University School of Medicine, Baltimore, USA

Kazuki Saito, Annabelle Campbell, Rachel Green & Allen R. Buskirk

2. Gene Center and Department of Biochemistry, University of Munich,  
Munich, Germany

Hanna Kratzat, Robert Buschauer, Otto Berninghausen & Roland  
Beckmann

3. Computational Biology Branch, National Center for Biotechnology  
Information, National Library of Medicine, National Institutes of  
Health, Bethesda, USA

A. Maxwell Burroughs & L. Aravind

4. Howard Hughes Medical Institute, Johns Hopkins University School  
of Medicine, Baltimore, USA

Rachel Green

## Contributions

K.S. performed the genetic screening, analysis of the NanoLuc–ble reporters and the sucrose gradients. A.C. analysed the CRP reporters and prepared samples for the MS experiments. A.M.B. and L.A. performed the phylogenetic analyses. H.K. performed the in vitro translation and in vitro nuclease assays, prepared samples for cryo-EM analysis and processed the cryo-EM data. O.B. and R. Buschauer collected the cryo-EM data. H.K. and R. Buschauer prepared the molecular models. H.K., R. Buschauer and R. Beckmann analysed and interpreted the structures, and R. Buschauer prepared the structural figures. R.G., R. Beckmann and A.R.B. supervised the project.

## Corresponding authors

Correspondence to [Roland Beckmann](#) or [Allen R. Buskirk](#).

## Ethics declarations

## Competing interests

The authors declare no competing interests.

## Peer review

### Peer review information

*Nature* thanks Vasili Hauryliuk, Yury Polikanov and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

## Additional information

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Extended data figures and tables

### [Extended Data Fig. 1 SmrB is a general ribosome rescue factor.](#)

**a**, Reporter protein from wild-type and  $\Delta ssrA$  strains was detected by antibodies against the N-terminal Strep-tag. Arrows indicate the full-length fusion protein (FL) and shorter NanoLuc protein (N). The RpoB protein serves as a loading control. **b**, Additional reporters to study ribosome rescue in *E. coli* with various stalling motifs. **c**, The expression of full-length NanoLuc-Ble protein was monitored with an anti-Flag antibody; anti-Strep antibodies reveal both full-length NanoLuc-Ble and truncated NanoLuc protein. RpoB serves as a loading control.

### [Extended Data Fig. 2 Phylogenetic tree of SMR domain proteins.](#)

Stylized phylogenetic tree depicting relationships between SMR domain clades. Clades with indicated bootstrap support are marked with circles.

Clade names are given to the right of the tree. Dotted lines indicate positions with little or no bootstrap support.

### Extended Data Fig. 3 Distribution and architectures of SMR-domain proteins.

**a**, Heat map demonstrating the conservation and distribution of SMR-domain proteins and other related translational quality control factors. Smr-all includes all types of SMR-domain proteins; Smr-euk includes only the eukaryotic branch. **b**, Domain organization of three representative bacterial proteins containing an SMR domain. **c**, Multiple alignment of the conserved regions in the N-terminal extension of SMR proteins from proteobacteria. Columns in the alignment are shaded and labeled according to biochemical character: -, negatively charged; h, hydrophobic in yellow; a, aromatic; p, polar in blue; l, aliphatic in yellow; s, small in green; u, tiny in green. Residue positions in the—xxxa motif are colored in white and shaded in black, marked by asterisks above the alignment. Residue positions forming part of the active site of the core SMR domain are colored in white and shaded in red. Sequences are labeled with NCBI accession number and organism abbreviation; abbreviations are provided below alignment. Secondary structure provided at top of alignment. Numbers to left and right of alignment denote positioning of the region. Internal numbers give the size of excised variable insert regions. **d**, Sequence alignment of SMR domains of representative proteins. Identical residues are shown in white with a red background; conserved residues are shown in red. The identity of each sequence is represented by the gene name, species name, and numbers indicating the beginning and the end of the residues used for the alignment. Ecol, *Escherichia coli*; Scer, *Saccharomyces cerevisiae*; Cele, *Caenorhabditis elegans*; Hsap, *Homo sapiens*; Atha, *Arabidopsis thaliana*; Bsub, *Bacillus subtilis*.

### Extended Data Fig. 4 SmrB cleavage, tmRNA tagging, and ribosome collisions.

**a**, The results of 5'-RACE showing the 5'-ends of downstream fragments in reads per million on the EP\* reporter. The first nt in the A site codon in the

stall motif is designated as zero. **b**, The results of 3'-RACE showing the 3'-ends of upstream fragments. **c**, tmRNA tagging sites on the EP\* reporter in the wild-type and  $\Delta smrB$  strains, corresponding to the residue immediately preceding the tmRNA tag in peptide sequences detected by targeted LC-MS-MS. The relative spectrum count is normalized by the count at the EP\* stall site (red) where tmRNA tagging was expected to occur in both the wild-type and  $\Delta smrB$  strains. The spectrum count corresponds to the mean and the standard deviation of three replicates. The arrow indicates the SmrB cleavage site demonstrated by 5'-RACE. **d**, 5'-RACE data on the SecM reporter reveal the SmrB cleavage sites as in Fig. 2b, zoomed in to show smaller peaks upstream. **e**, The distribution of Flag-SmrB in cells treated with 5, 50, or 500  $\mu\text{g}/\text{mL}$  erythromycin (ERY) was determined by fractionation over a sucrose gradient and detected with an anti-Flag antibody.

### Extended Data Fig. 5 Cryo-EM data processing for the *E. coli* disome sample.

Shown are the classification scheme, representative micrographs (the scale bar is 500 Å), 2D class averages and the Gold standard Fourier Shell Correlation (GSFSC) curve for the volume containing the 70S stalled ribosome and the 30S of the collided ribosome, as well as the full disome.

### Extended Data Fig. 6 Analysis of the *E. coli* disome structure and comparison of different disome structures.

**a**, The architecture of the *E. coli* disome is not compatible with bS1 remaining bound to the stalled ribosome. Aligned models of the 30S subunits of the collided (left) and the stalled (middle) ribosomes are shown in surface representation. The position of bS1 as observed in the collided ribosome is shown in purple and the same position of bS1 in the stalled ribosome is indicated by a dashed line. The clash between bS1 of the stalled ribosome and the 30S subunit of the collided ribosome that would occur upon disome formation is shown on the right. **b**, Cartoon representation of the individual interactions as they occur at the *E. coli* disome interface. **c**, 2D class averages and cryo-EM structure model of an *E. coli* trisome. **d**, **e**,

Comparison of the *E. coli* (*E.c.*) and *B. subtilis* (*B.s.*) disomes displaying full and cut views. Note the smaller space between stalled and collided ribosomes in the *B.s.* disome interface as illustrated by comparing the positions of uS2 proteins in the zoomed view in **c**. **f**, Surface representation of the structural model of the *S. cerevisiae* disome. **g&h**, Surface representation of the *E. coli* and *B. subtilis* hibernation disomes.

**Extended Data Fig. 7 Cryo-EM data processing for the *B. subtilis* disome and *E. coli* trisome sample.**

- a**, Shown are the classification scheme, and the Gold standard Fourier Shell Correlation (GSFSC) curves for the final volumes of the *B. subtilis* disome containing the 70S stalled ribosome and the 70S of the collided ribosome. **b**, Shown are the 2D class averages, classification scheme, and the Gold standard Fourier Shell Correlation (GSFSC) of the *E. coli* trisome.

**Extended Data Fig. 8 Production of collided and non-collided disomes and relative peak areas of monosomes and disomes in the SmrB nuclease assay.**

- a**, mRNA construct to create the collided *E. coli* disomes and trisomes and below the sucrose density gradient after *in vitro* translation. The ribosome stalling site is indicated by an asterisk. **b**, mRNA construct to create the non-collided disomes that were used in the nuclease assay and below the sucrose density gradient after *in vitro* translation. **c**, Relative monosome and disome peak area calculated from the sucrose gradient profiles of the SmrB nuclease assay, showing the mean and standard deviation of three replicates. **d**, The relative decrease of the area of the disome peak upon addition of SmrB is shown as the mean and standard deviation of three replicates. (The mean difference of the relative disome peak area of collided ribosomes between control and SmrB reaction was set to 1).

**Extended Data Fig. 9 Cryo-EM data processing for the *E. coli* disome sample.**

Shown are the classification scheme, representative micrographs (the scale bar is 500 Å), 2D class averages and the Gold standard Fourier Shell Correlation (GSFSC) curve for the respective 3D reconstructions. The segmented density for SmrB is colored according to local resolution.

## Extended Data Fig. 10 Structural model of SmrB.

**a**, Secondary structure of SmrB. The DLH to ALA mutation is indicated. **b**, AF2 prediction models 1-5 as predicted through the API from the Söding lab. The SMR domain is predicted with high confidence, while the linker to the N-terminal helix appears flexible. **c**, AF2 prediction of the interaction between SmrB and uS2. For this prediction uS2 was fused to the C-terminus of SmrB with a glycine serine linker (39 copies of GS). The prediction shows the N-terminal helix of SmrB folded back onto uS2. **d**, Adjustment of the AF2 predicted model of SmrB-uS2. Without adjustment according to the cryo-EM density (as shown in D) the SMR domain would clash with the ribosome. **e**, Top: Cryo-EM density and adjusted model of the SmrB. Middle: Cryo-EM density and rigid body docked model of the N-terminus of SmrB from the collided 30S onto the stalled 30S. A second copy of SmrB was found anchored to uS2 of the stalled ribosome. However, there was no density for the SMR domain of the second SmrB, indicating a high degree of flexibility due to the lack of another ribosome in front of the stalled one. Bottom: in the control disome without SmrB, there is no density for the N-terminus of SmrB. **f**, Comparison of the AF2 prediction, the homology model, and the adjusted model of SmrB. Compared to the AF2 prediction, the homology model is missing the two N-terminal helices and most of the loops are slightly different (top). The AF2 prediction almost perfectly matched the cryo-EM density map and the corresponding adjusted model (middle and bottom). Only the catalytic loop (carrying the active site mutations) had to be slightly adjusted to prevent clashes with the mRNA. The N-terminus was adjusted as discussed above. **g**, During the preparation of this manuscript the AF2 prediction for SmrB (YfcN) became available at the alphafold database at EMBL-EBI. The deposited model resembles our final adjusted model very well including the position of the N-terminus. The confidence of the prediction (pLDDT) is indicated.

## Extended Data Fig. 11 Testing the importance of structural interactions for SmrB activity.

**a**, Examples of operons containing both uS21 and SMR-domain proteins. **b**, The distribution of Flag-tagged full-length SmrB and a construct with only the SMR domain (residues 88–183) was determined by fractionation over sucrose gradient and detection with an anti-Flag antibody. A non-specific band is marked with \*. **c**, Northern blots using the 3'-probe against the CRP reporters with the short SecM stalling motif in wild-type cells, bL9-deletion strain ( $\Delta rplI$ ), and a strain where mCherry is fused to the C-terminus of bL9 (bL9-mCherry). Ethidium bromide staining of 16S rRNA serves as a loading control. **d**, Northern blots using the 3'-probe against the CRP reporters with the short SecM stalling motif in wild-type cells, a strain where MBP is fused to the N-terminus of uS21, and a strain where GFP is fused to the C-terminus of uS6.

### **Extended Data Table 1 Data collection and refinement statistics**

## **Supplementary information**

### Supplementary Information

This file contains Supplementary Figs. 1 and 2, which show the uncropped western blots and northern blots corresponding to the main and Extended Data Figures.

### Reporting Summary

### Supplementary Table 1

This Excel file contains a list and description of oligonucleotides, plasmids and strains used in this study.

## **Rights and permissions**

## [Reprints and Permissions](#)

## About this article

### Cite this article

Saito, K., Kratzat, H., Campbell, A. *et al.* Ribosome collisions induce mRNA cleavage and ribosome rescue in bacteria. *Nature* **603**, 503–508 (2022). <https://doi.org/10.1038/s41586-022-04416-7>

- Received: 03 September 2021
- Accepted: 07 January 2022
- Published: 09 March 2022
- Issue Date: 17 March 2022
- DOI: <https://doi.org/10.1038/s41586-022-04416-7>

### Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-022-04416-7>

- Article
- [Published: 09 March 2022](#)

# Bacterial ribosome collision sensing by a MutS DNA repair ATPase parologue

- [Federico Cerullo](#) [ORCID: orcid.org/0000-0002-6042-7810](#)<sup>1</sup> nat,
- [Sebastian Filbeck](#) [ORCID: orcid.org/0000-0001-5729-4538](#)<sup>1</sup> nat,
- [Pratik Rajendra Patil](#)<sup>1</sup>,
- [Hao-Chih Hung](#)<sup>1</sup>,
- [Haifei Xu](#)<sup>2</sup>,
- [Julia Vornberger](#)<sup>1</sup>,
- [Florian W. Hofer](#) [ORCID: orcid.org/0000-0002-6531-3820](#)<sup>1</sup>,
- [Jaro Schmitt](#) [ORCID: orcid.org/0000-0001-6701-7112](#)<sup>1</sup>,
- [Guenter Kramer](#) [ORCID: orcid.org/0000-0001-7552-8393](#)<sup>1</sup>,
- [Bernd Bukau](#) [ORCID: orcid.org/0000-0003-0521-7199](#)<sup>1</sup>,
- [Kay Hofmann](#) [ORCID: orcid.org/0000-0002-2289-9083](#)<sup>3</sup>,
- [Stefan Pfeffer](#) [ORCID: orcid.org/0000-0001-5843-2795](#)<sup>1</sup> &
- [Claudio A. P. Joazeiro](#) [ORCID: orcid.org/0000-0003-1433-8013](#)<sup>1,2</sup>

*Nature* volume 603, pages 509–514 (2022)

- 4453 Accesses
- 125 Altmetric
- [Metrics details](#)

## Subjects

- [Cryoelectron microscopy](#)
- [Ribosome](#)
- [RNA quality control](#)

## Abstract

Ribosome stalling during translation is detrimental to cellular fitness, but how this is sensed and elicits recycling of ribosomal subunits and quality control of associated mRNA and incomplete nascent chains is poorly understood<sup>1,2</sup>. Here we uncover *Bacillus subtilis* MutS2, a member of the conserved MutS family of ATPases that function in DNA mismatch repair<sup>3</sup>, as an unexpected ribosome-binding protein with an essential function in translational quality control. Cryo-electron microscopy analysis of affinity-purified native complexes shows that MutS2 functions in sensing collisions between stalled and translating ribosomes and suggests how ribosome collisions can serve as platforms to deploy downstream processes: MutS2 has an RNA endonuclease small MutS-related (SMR) domain, as well as an ATPase/clamp domain that is properly positioned to promote ribosomal subunit dissociation, which is a requirement both for ribosome recycling and for initiation of ribosome-associated protein quality control (RQC). Accordingly, MutS2 promotes nascent chain modification with alanine-tail degrons—an early step in RQC—in an ATPase domain-dependent manner. The relevance of these observations is underscored by evidence of strong co-occurrence of MutS2 and RQC genes across bacterial phyla. Overall, the findings demonstrate a deeply conserved role for ribosome collisions in mounting a complex response to the interruption of translation within open reading frames.

This is a preview of subscription content

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

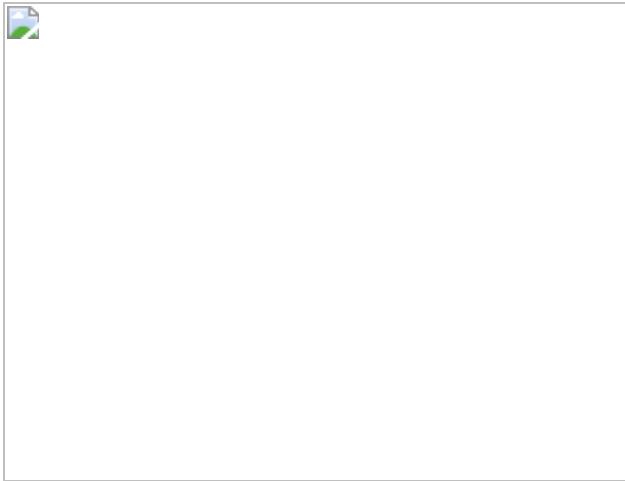
[Buy](#)

All prices are NET prices.

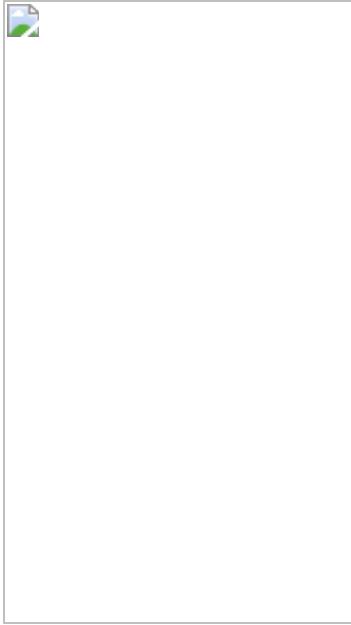
### **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

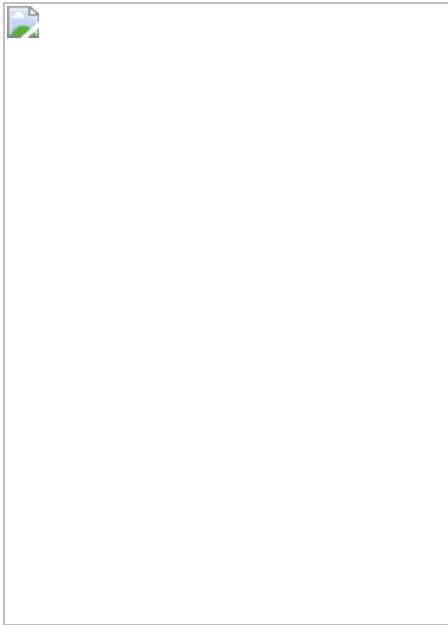
**Fig. 1: MutS2 binds to ribosomes and protects cells from translational stress.**



**Fig. 2: Cryo-EM structure of the bacterial collided disome.**



**Fig. 3: Structural basis for ribosome collision sensing by MutS2.**



**Fig. 4: MutS2 promotes ribosome splitting upstream of RQC.**



## Data availability

Cryo-EM densities have been deposited in the Electron Microscopy Data Bank under accession codes EMD-[14163](#) ('collided disomes': 30S leading ribosome), EMD-[14162](#) ('collided disomes': 70S leading ribosome), EMD-[14165](#) ('collided disomes': 30S collided ribosome), EMD-[14164](#) ('collided disomes': 70S collided ribosome), EMD-[14157](#) ('collided disomes':

composite density), EMD-[14166](#) (dataset 1: 50S ribosome with nascent chain-linked P-site tRNA), EMD-[14160](#) ('+MutS2': 70S leading ribosome), EMD-[14161](#) ('+MutS2': 70S collided ribosome), EMD-[14156](#) ('+MutS2 conformation 1': 70S leading ribosome) and EMD-[14159](#) ('+MutS2 conformation 2': 70S leading ribosome). Atomic coordinates have been deposited in the Protein Data Bank under accession codes [7QV1](#) ('collided disomes': 70S leading ribosome), [7QV2](#) ('collided disomes': 70S collided ribosome) and [7QV3](#) ('+MutS2 conformation 2': 70S leading ribosome bound to MutS2 in conformation 2). Published structural data used in this article were obtained from Protein Data Bank under codes [6HA1](#), [5NGM](#), [1DIV](#), [3J9W](#), [7N2V](#), [6WDB](#), [6WDG](#), [5X9W](#) and [1EWQ](#) and from the AlphaFold Protein Structure Database under code [Q2FZD3](#). [Source data](#) are provided with this paper.

## References

1. Joazeiro, C. A. P. Ribosomal stalling during translation: providing substrates for ribosome-associated protein quality control. *Annu. Rev. Cell Dev. Biol.* **33**, 343–368 (2017).
2. Meydan, S. & Guydosh, N. R. A cellular handbook for collided ribosomes: surveillance pathways and collision types. *Curr. Genet.* **67**, 19–26 (2021).
3. Modrich, P. Mechanisms in *E. coli* and human mismatch repair (Nobel Lecture). *Angew. Chem. Int. Ed. Engl.* **55**, 8490–8501 (2016).
4. Moore, S. D. & Sauer, R. T. The tmRNA system for translational surveillance and ribosome rescue. *Annu. Rev. Biochem.* **76**, 101–124 (2007).
5. Joazeiro, C. A. P. Mechanisms and functions of ribosome-associated protein quality control. *Nat. Rev. Mol. Cell Biol.* **20**, 368–383 (2019).
6. Venkataraman, K., Guja, K. E., Garcia-Diaz, M. & Karzai, A. W. Non-stop mRNA decay: a special attribute of *trans*-translation mediated ribosome rescue. *Front. Microbiol.* **5**, 93 (2014).

7. Howard, C. J. & Frost, A. Ribosome-associated quality control and CAT tailing. *Crit. Rev. Biochem. Mol. Biol.* **56**, 603–620 (2021).
8. Lytvynenko, I. et al. Alanine tails signal proteolysis in bacterial ribosome-associated quality control. *Cell* **178**, 76–90 (2019).
9. D’Orazio, K. N. & Green, R. Ribosome states signal RNA quality control. *Mol. Cell* **81**, 1372–1383 (2021).
10. Vind, A. C., Genzor, A. V. & Bekker-Jensen, S. Ribosomal stress-surveillance: three pathways is a magic number. *Nucleic Acids Res.* **48**, 10648–10661 (2020).
11. Ikeuchi, K. et al. Collided ribosomes form a unique structural interface to induce Hel2-driven quality control pathways. *EMBO J.* **38**, 1–40 (2019).
12. Simms, C. L., Yan, L. L. & Zaher, H. S. Ribosome collision is critical for quality control during no-go decay. *Mol. Cell* **68**, 361–373 (2017).
13. Juszkiewicz, S. et al. ZNF598 is a quality control sensor of collided ribosomes. *Mol. Cell* **72**, 469–481 (2018).
14. Garzia, A. et al. The E3 ubiquitin ligase and RNA-binding protein ZNF598 orchestrates ribosome quality control of premature polyadenylated mRNAs. *Nat. Commun.* **8**, 16056 (2017).
15. Sundaramoorthy, E. et al. ZNF598 and RACK1 regulate mammalian ribosome-associated quality control function by mediating regulatory 40S ribosomal ubiquitylation. *Mol. Cell* **65**, 751–760 (2017).
16. Juszkiewicz, S., Speldewinde, S. H., Wan, L., Svejstrup, J. Q. & Hegde, R. S. The ASC-1 complex disassembles collided ribosomes. *Mol. Cell* **79**, 603–614 (2020).
17. Matsuo, Y. et al. RQT complex dissociates ribosomes collided on endogenous RQC substrate SDD1. *Nat. Struct. Mol. Biol.* **27**, 323–332 (2020).

18. Matsuo, Y. et al. Ubiquitination of stalled ribosome triggers ribosome-associated quality control. *Nat. Commun.* **8**, 159 (2017).
19. Glover, M. L. et al. NONU-1 encodes a conserved endonuclease required for mRNA translation surveillance. *Cell Rep.* **30**, 4321–4331 (2020).
20. D’Orazio, K. N. et al. The endonuclease Cue2 cleaves mRNAs at stalled ribosomes during no go decay. *eLife* **8**, e49117 (2019).
21. Nürenberg-Goloub, E. & Tampé, R. Ribosome recycling in mRNA translation, quality control, and homeostasis. *Biol. Chem.* **401**, 47–61 (2019).
22. Donaldson, K. M., Yin, H., Gekakis, N., Supek, F. & Joazeiro, C. A. P. Ubiquitin signals protein trafficking via interaction with a novel ubiquitin binding domain in the membrane fusion regulator, Vps9p. *Curr. Biol.* **13**, 258–262 (2003).
23. Burby, P. E. & Simmons, L. A. MutS2 promotes homologous recombination in *Bacillus subtilis*. *J. Bacteriol.* **199**, e00682-16 (2017).
24. Pinto, A. V. et al. Suppression of homologous and homeologous recombination by the bacterial MutS2 protein. *Mol. Cell* **17**, 113–120 (2005).
25. Hingorani, M. M. Mismatch binding, ADP–ATP exchange and intramolecular signaling during mismatch repair. *DNA Repair* **38**, 24–31 (2016).
26. Groothuizen, F. S. & Sixma, T. K. The conserved molecular machinery in DNA mismatch repair enzyme structures. *DNA Repair* **38**, 14–23 (2016).
27. Kyrpides, N. C., Woese, C. R. & Ouzounis, C. A. KOW: a novel motif linking a bacterial transcription factor with ribosomal proteins. *Trends Biochem. Sci.* **21**, 425–426 (1996).

28. Fukui, K. & Kuramitsu, S. Structure and function of the small MutS-related domain. *Mol. Biol. Int.* **2011**, 691735 (2011).
29. Sachadyn, P. Conservation and diversity of MutS proteins. *Mutat. Res.* **694**, 20–30 (2010).
30. Pochopien, A. A. et al. Structure of Gcn1 bound to stalled and colliding 80S ribosomes. *Proc. Natl Acad. Sci. USA* **118**, e2022756118 (2021).
31. Sohmen, D. et al. Structure of the *Bacillus subtilis* 70S ribosome reveals the basis for species-specific stalling. *Nat. Commun.* **6**, 6941 (2015).
32. Smith, A. M., Costello, M. S., Kettring, A. H., Wingo, R. J. & Moore, S. D. Ribosome collisions alter frameshifting at translational reprogramming motifs in bacterial mRNAs. *Proc. Natl Acad. Sci. USA* **116**, 21769–21779 (2019).
33. Borovinskaya, M. A., Shoji, S., Holton, J. M., Fredrick, K. & Cate, J. H. D. A steric block in translation caused by the antibiotic spectinomycin. *ACS Chem. Biol.* **2**, 545–552 (2007).
34. Brodersen, D. E. et al. The structural basis for the action of the antibiotics tetracycline, pactamycin, and hygromycin B on the 30S ribosomal subunit. *Cell* **103**, 1143–1154 (2000).
35. Svetlov, M. S. et al. High-resolution crystal structures of ribosome-bound chloramphenicol and erythromycin provide the ultimate basis for their competition. *RNA* **25**, 600–606 (2019).
36. Filbeck, S. et al. Mimicry of canonical translation elongation underlies alanine tail synthesis in RQC. *Mol. Cell* **81**, 104–114 (2021).
37. Crowe-McAuliffe, C. et al. Structural basis for bacterial ribosome-associated quality control by RqcH and RqcP. *Mol. Cell* **81**, 115–126 (2021).

38. Takada, H. et al. RqcH and RqcP catalyze processive poly-alanine synthesis in a reconstituted ribosome-associated quality control system. *Nucleic Acids Res.* **49**, 8355–8369 (2021).
39. Thrun, A. et al. Convergence of mammalian RQC and C-end rule proteolytic pathways via alanine tailing. *Mol. Cell* **81**, 2112–2122 (2021).
40. Korostelev, A., Trakhanov, S., Laurberg, M. & Noller, H. F. Crystal structure of a 70S ribosome–tRNA complex reveals functional interactions and rearrangements. *Cell* **126**, 1065–1077 (2006).
41. Tejada-Arranz, A., de Crécy-Lagard, V. & de Reuse, H. Bacterial RNA degradosomes: molecular machines under tight control. *Trends Biochem. Sci.* **45**, 42–57 (2020).
42. Arnaud, M., Chastanet, A. & Débarbouillé, M. New vector for efficient allelic replacement in naturally nontransformable, low-GC-content, gram-positive bacteria. *Appl. Environ. Microbiol.* **70**, 6887–6891 (2004).
43. Koo, B.-M. et al. Construction and analysis of two genome-scale deletion libraries for *Bacillus subtilis*. *Cell Syst.* **4**, 291–305 (2017).
44. Mende, D. R. et al. proGenomes2: an improved database for accurate and consistent habitat, taxonomic and functional annotations of prokaryotic genomes. *Nucleic Acids Res.* **48**, D621–D625 (2020).
45. Bucher, P., Karplus, K., Moeri, N. & Hofmann, K. A flexible motif search technique based on generalized profiles. *Comput. Chem.* **20**, 3–23 (1996).
46. Levy, J. A., LaFlamme, C. W., Tsaprailis, G., Crynen, G. & Page, D. T. *Dyrk1a* mutations cause undergrowth of cortical pyramidal neurons via dysregulated growth factor signaling. *Biol. Psychiatry* **90**, 295–306 (2021).

47. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).
48. Zivanov, J., Nakane, T. & Scheres, S. H. W. Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in RELION-3.1. *IUCrJ* **7**, 253–267 (2020).
49. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
50. Crowe-McAuliffe, C. et al. Structural basis for antibiotic resistance mediated by the *Bacillus subtilis* ABCF ATPase VmlR. *Proc. Natl Acad. Sci. USA* **115**, 8978–8983 (2018).
51. Matzov, D. et al. The cryo-EM structure of hibernating 100S ribosome dimer from pathogenic *Staphylococcus aureus*. *Nat. Commun.* **8**, 723 (2017).
52. Hoffman, D. W. et al. Crystal structure of prokaryotic ribosomal protein L9: a bi-lobed RNA-binding protein. *EMBO J.* **13**, 205–212 (1994).
53. Zimmermann, L. et al. A completely reimplemented MPI Bioinformatics Toolkit with a new HHpred server at its core. *J. Mol. Biol.* **430**, 2237–2243 (2018).
54. Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D* **75**, 861–877 (2019).
55. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
56. Croll, T. I. ISOLDE: a physically realistic environment for model building into low-resolution electron-density maps. *Acta Crystallogr. D* **74**, 519–530 (2018).

57. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
58. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).
59. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
60. Ribeiro, J. V. et al. QwikMD—integrative molecular dynamics toolkit for novices and experts. *Sci. Rep.* **6**, 26536 (2016).
61. Trabuco, L. G., Villa, E., Mitra, K., Frank, J. & Schulten, K. Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics. *Structure* **16**, 673–683 (2008).
62. Phillips, J. C. et al. Scalable molecular dynamics on CPU and GPU architectures with NAMD. *J. Chem. Phys.* **153**, 044130 (2020).
63. Rundlet, E. J. et al. Structural basis of early translocation events on the ribosome. *Nature* **595**, 741–745 (2021).
64. Loveland, A. B., Demo, G. & Korostelev, A. A. Cryo-EM of elongating ribosome with EF-Tu•GTP elucidates tRNA proofreading. *Nature* **584**, 640–645 (2020).
65. Goddard, T. D. et al. UCSF ChimeraX: meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).
66. Amiri, H. & Noller, H. F. Structural evidence for product stabilization by the ribosomal mRNA helicase. *RNA* **25**, 364–375 (2019).

## Acknowledgements

The authors thank G. Tsaprailis for MS analyses, G. Crynen for bioinformatic and statistical analyses of proteomics data, E. Zupa for help with atomic modelling, R. Tampé and E. Nürenberg-Goloub for comments

on the manuscript, R. Manna and O. Canterbury for help with molecular biology experiments and S. Kreger for technical assistance. S.P. acknowledges access to the infrastructure of the Cryo-EM Network at Heidelberg University (HDcryoNET) and support by G. Hofhaus and D. Flemming. S.P. also acknowledges the services SDS@hd and bwHPC supported by the Ministry of Science, Research and the Arts Baden-Württemberg. F.C., S.F., H.-C.H. and J.S. are members of the Heidelberg Biosciences International Graduate School (HBIGS). We acknowledge financial support by a fellowship of the Deutsche Studienstiftung to J.S. and by grants of the Deutsche Forschungsgemeinschaft (DFG) to B.B. (SFB1036), C.A.P.J. (SFB1036) and G.K. (KR 3593/2-1), of the European Research Council to B.B. (ERC Advanced Grant 743118) and of the National Institute of Neurological Disorders and Stroke (NINDS) of the NIH (R01 NS102414) to C.A.P.J. S.P. acknowledges funding by the Aventis Foundation and the Chica and Heinz Schaller Foundation.

## Author information

### Author notes

1. These authors contributed equally: Federico Cerullo, Sebastian Filbeck

## Affiliations

1. Zentrum für Molekulare Biologie der Universität Heidelberg, DKFZ-ZMBH Alliance, Heidelberg, Germany

Federico Cerullo, Sebastian Filbeck, Pratik Rajendra Patil, Hao-Chih Hung, Julia Vornberger, Florian W. Hofer, Jaro Schmitt, Guenter Kramer, Bernd Bukau, Stefan Pfeffer & Claudio A. P. Joazeiro

2. Department of Molecular Medicine, Scripps Florida, Jupiter, FL, USA

Haifei Xu & Claudio A. P. Joazeiro

3. Institute for Genetics, University of Cologne, Cologne, Germany

Kay Hofmann

## Contributions

F.C. performed most biochemical and molecular genetic work, together with H.X., H.-C.H. and J.V. S.F. performed all cryo-EM and structure analysis work, with help from F.W.H. P.R.P. developed the Ala-PIRHification method and performed analyses with help from F.C. J.S., G.K. and B.B. assisted with sucrose gradient fractionation and disome analyses. K.H. performed bioinformatic analyses of bacterial genomes and protein domains. S.P. and C.A.P.J. designed studies and conceived and supervised the project. C.A.P.J., S.P., S.F. and F.C. wrote the manuscript. C.A.P.J., G.K., B.B. and S.P. acquired funding. All authors edited and commented on the manuscript.

## Corresponding authors

Correspondence to [Stefan Pfeffer](#) or [Claudio A. P. Joazeiro](#).

## Ethics declarations

## Competing interests

The authors declare no competing interests.

## Peer review

## Peer review information

*Nature* thanks Vasili Hauryliuk, Yury Polikanov and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

## Additional information

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Extended data figures and tables

### Extended Data Fig. 1 Cryo-EM data processing scheme.

**a**, Two datasets were acquired and processed separately using Relion. To reject false-positive particles, auto-picked particles were subjected to a 3D classification. A subset of 50S depicting particles from dataset 1 was subjected to 3D auto-refinement. **b**, To enrich for collided disomes, 70S depicting particles were subsequently subjected to multiple rounds of 3D classification using a spherical mask with a diameter of 435 Å (dataset 1) or focused on the 30S subunit of the collided disome (dataset 2). After the first round of classification particles were merged. **c**, Particles of collided disomes were re-extracted centered on the leading and the colliding ribosome. 3D auto-refinements were focused on the 30S subunit or the full 70S of the respective ribosome. **d**, MutS2 conformations were sorted by two rounds of 3D classification using a mask encompassing the MutS2-homodimer. To increase local resolution for the MustS2-A clamp domain, the two separate conformations were merged and subjected to 3D auto-refinements based on either the leading or the colliding ribosome. **e**, To enrich for particles that depict collided trisomes, particles from collided disomes centered on the collided ribosome were subjected to a 3D classification focused on the third ribosome. Retained particles were re-extracted centered on the third ribosome (Collided 2) and subjected to 3D auto-refinement.

### Extended Data Fig. 2 Structural similarity between the predicted MutS2 structure and the X-ray structure of a bacterial MutS protein.

**a**, Overall structures of *B. subtilis* MutS2 (this study) and *Neisseria gonorrhoeae* MutS (PDB 5X9W). Segments of MutS2 that were not resolved by cryo-EM, including the KOW and SMR domains, are shown in grey for one monomer as predicted by AlphaFold. **b–f**, Superposition of the

two models with respect to the ATPase domain (**b**), the individual MutS clamp and lever domains (**c**, **d**) and all conserved segments of the individual monomer structures (**e**, **f**). Dashed boxes indicate the zoomed area for panels **b–f**.

### Extended Data Fig. 3 Global and local resolution estimation.

For each 3D auto-refinement run, densities have been colored according to local resolution. Mask-corrected Fourier shell correlation (FSC) has been calculated between two independently refined half-sets of the data ('gold standard'; FSC threshold at 0.143). Shown particle sets are "Collided Disomes" (**a**, **b**, **c**, **d**), "+MutS2" (**g**, **h**), MutS2 conformation 1 and 2 (respectively **e**, **f**) and 50S subunits obstructed with a nascent chain-linked P-site tRNA (**i**). Particles are either centered on the leading (**a**, **b**, **e**, **f**, **g**) or on the colliding ribosome (**c**, **d**, **h**) and refined using either a 30S (**a**, **c**) or a 70S mask of the respective ribosome (**b**, **d**, **e**, **f**, **g**, **h**)

### Source data

### Extended Data Fig. 4 Molecular details of contacts stabilizing the disome interface.

**a**, The 30S subunit heads interact via complementary charged patches on uS9 of the leading ribosome and uS10 of the collided ribosome ('head contact'). Surface representations for uS9 and uS10 are shown and colored according to electrostatic potential (blue: positive, red: negative, white: neutral). The interacting patches are indicated. **b**, On the opposing side of the inter-ribosomal mRNA trajectory, helix 25 of the leading ribosome 30S subunit rRNA is accommodated in a groove on the 30S subunit of the collided ribosome formed by uS2 and uS8 ('body contact'). In particular, helix 25 of the leading ribosome directly interacts with a surface exposed  $\alpha$ -helix (Leu43 - Glu63) and the partially unordered C-terminus of uS2 of the collided ribosome. **c**, In close proximity, two additional contacts between the two collided ribosomes ('platform contacts') complete the network of interactions clustered around the mRNA entry and exit sites. First, the L1-stalk adopting an extreme out-conformation on the leading ribosome directly contacts the 30S subunit rRNA of the collided ribosome. Second,

uS11 of the leading ribosome contacts uS4 of the collided ribosome, mainly via hydrophobic interactions (Val14 and Ile18 of S11; Val157 and Gly23 of S4) and aromatic stacking (His40 of S11; Phe160 of S4). **d**, A more peripheral interaction is mediated by ribosomal protein bL9 of the leading ribosome ('bL9 contact'). **e**, The binding site of bL9 on the *B. subtilis* 50S subunit has been significantly remodeled compared to the *E. coli* ribosome<sup>66</sup> (PDB 6BY1). While the interaction area between bL28 and the N-terminal half of bL9 is reduced in *B. subtilis*, this is compensated by an expansion of the 50S subunit rRNA (helix 15), which together with the L1-stalk forms an extended binding groove for bL9.

### **Extended Data Fig. 5 Structural and compositional remodeling of the mRNA exit site on the leading ribosome.**

**a**, Upper panel: Trajectory of unstrained mRNA exiting the mRNA channel in a defined direction to interact with the anti-Shine-Dalgarno (SD) rRNA sequence of the 16S rRNA 3'end, thereby forming an RNA duplex reminiscent of the SD helix during translation initiation<sup>31</sup> (PDB 3J9W, EMDB 6306). Middle panel: The unstrained mRNA exiting the collided ribosome interacts with the anti-SD rRNA to form an RNA duplex. Bottom panel: In the collided disome, the mRNA under strain follows a vastly different trajectory, which is accompanied by remarkable structural remodeling of the mRNA exit site of the leading ribosome. In particular, the rRNA anti-SD sequence of the leading ribosome can no longer interact with the mRNA and becomes disordered, which renders the binding site for the bS21 C-terminus on the leading ribosome accessible and at the same time reduces the interaction surface for uS2 on the 30S subunit. Superposition: Superposition of mRNAs exiting the ribosomes and the respective SD-helices in the translating and collided ribosomes, as well as bS21 in the leading ribosome. **b**, Comparison of uS2 density in the leading and collided ribosomes at the same density threshold level. In all panels, local resolution filtered densities based on 3D auto-refinements focused on either the 30S or the 70S of the respective ribosome are shown (See Extended Data Fig. 1).

### **Extended Data Fig. 6 The C-terminal half of bL9 sterically excludes binding of EF-G on the collided ribosome.**

**a**, Binding site of bL9 on the 30S subunit of the collided ribosome. **b**, Atomic model and simulated density of EF-G<sup>63</sup> (PDB 7N2V) mapped onto the 30S subunit of the collided ribosome by fitting the 30S-EF-G complex as a rigid body. Overlapping segments of bL9 and EF-G are shown in transparent grey. **c**, As in ‘b’, but not showing the EF-G atomic model and with overlapping segments of bL9 and EF-G colored in purple.

### Extended Data Fig. 7 Conformational plasticity of the MutS2-B clamp region.

**a**, Two views on the local resolution-filtered density of the MutS2 dimer after 3D auto-refinement of all MutS2-containing particles. Highly fragmented density for the MutS2-B clamp and lever domains (left panel) indicated conformational heterogeneity. **b**, **c**, Computational particle sorting focused on MutS2-B produced two structurally distinct subpopulations slightly differing in the positioning of the MutS-domains III and IV, in which the clamp region either binds to ribosomal protein L5 (**b**) or the nascent chain-associated P-site tRNA of the leading ribosome (**c**). **d**, Overlay of the two conformations from ‘b’ and ‘c’. In all panels, local resolution-filtered cryo-EM densities are shown as obtained after 3D autorefinement centered on the leading ribosome using either all MutS2-containing particles (ribosome densities in all panels, MutS2 density in panel A) or using only particles representing one of the two different MutS2 conformations (MutS2 density in panels **b**, **c**).

### Extended Data Fig. 8 Conformational plasticity of the A-site finger during the translational elongation cycle.

**a-c**, Atomic models for rRNA, tRNAs and ribosomal protein uL5. **a**, The leading ribosome of the *B. subtilis* collided disome. **b**, The *E. coli* ribosome in accommodation state IV-A<sup>64</sup> (PDB 6WDB). **c**, The pre-translocation state VI-B<sup>64</sup> (PDB 6WDG). **d**, Superposition of the structures shown in (a-c) according to ribosomal protein uL5, demonstrating remodeling of the MutS2-A binding site during the translation elongation cycle.

## Extended Data Fig. 9 Structure of the 50S ribosomal subunit obstructed with a nascent chain-linked P-site tRNA.

**a**, Local resolution-filtered cryo-EM density of the 50S ribosomal subunit obstructed with the nascent chain-linked P-site tRNA. Due to conformational flexibility, cryo-EM density for peripheral segments of the P-site tRNA is fragmented. **b**, Density was sliced open to allow for an unobstructed view on the P-site tRNA and the associated incomplete nascent chain. **c**, Model of a nascent chain-linked P-site tRNA<sup>36</sup> (PDB 7AQC) superposed to the cryo-EM density. Zoom on the CCA-tail of the P-site-tRNA with linked nascent chain.

## Extended Data Fig. 10 Evidence from genomic analyses link *mutSB* and *rqcH*.

**a**, *mutSB* and *rqcH* strongly co-occur. Distribution of *rqcH* and *mutSB* quantitated separately for the indicated bacterial phyla. Both absolute numbers and frequencies are presented. For the latter, the higher the frequency the darker the background red color. **b**, *mutSB* localizes in the vicinity of *rqcH* in diverse bacteria. Genes are represented by arrows reflecting the direction of transcription, with *mutSB* and *rqcH* indicated in blue and red, respectively. In the instances where *mutSB* and *rqcH* are separated by genes represented as grey arrows, those genes are highly diverse, have no obvious relationship to translational quality control, and are generally unrelated between different species.

## Extended Data Fig. 11 Model for MutS2 function in sensing ribosome collisions and eliciting downstream responses.

The model depicts, from top to bottom: ribosomes translating an mRNA with a stalling site within the open reading frame (“Translation”); the leading ribosome becoming stalled (“Internal stalling”); a trailing ribosome colliding with the stalled ribosome (“Ribosomal collision”); MutS2 sensing the collision and promoting both separation of the ribosomal subunits (“Ribosome splitting”) and endonucleolytic cleavage (“mRNA cleavage”). *Left side*: Ribosomal splitting generates a 50S subunit still obstructed with a

nascent chain-tRNA conjugate, which is sensed by RqcH and RqcP, resulting in elongation of the nascent chain with a C-terminal Ala tail (“Ala tailing”). Nascent-chain release is accompanied by 50S recycling (“Ribosome recycling”, dotted line). *Right side*: mRNA cleavage can result in mRNA decay (dotted line) or in trailing ribosomes becoming stalled at the mRNA 3’end, which are sensed by SsrA/tmRNA. The SsrA reaction leads to ribosome recycling (“Ribosome recycling”, dotted line) and to nascent-chain modification with a C-terminal SsrA tag (“SsrA tagging”). Both Ala-tails and the SsrA tag act as degrons, recognized by ClpXP and other proteases (“Proteolysis”). See the main text for additional details. *Objects*: the mRNA is shown in red, with a stalling site within the open reading frame represented by ‘!’ within a triangle and the mRNA stop codon shown as a ‘Stop’ traffic sign; the direction of translation is indicated by arrows; the stalled ribosome is shown in orange (light, 50S subunit; dark, 30S subunit); trailing and collided ribosomes are shown in green (dark, 50S subunit; light, 30S subunit). Quality control factors are indicated by their names.

## **Extended Data Table 1 Cryo-EM data collection, refinement and validation statistics.**

## **Supplementary information**

### **Supplementary Information**

This file contains Supplementary Figs. 1–8 and Supplementary Tables 1–3.

### **Reporting Summary**

## **Source data**

### **Source Data Fig. 1**

### **Source Data Fig. 4**

## [Source Data Extended Data Fig. 3](#)

## Rights and permissions

### [Reprints and Permissions](#)

## About this article

### Cite this article

Cerullo, F., Filbeck, S., Patil, P.R. *et al.* Bacterial ribosome collision sensing by a MutS DNA repair ATPase paralogue. *Nature* **603**, 509–514 (2022).  
<https://doi.org/10.1038/s41586-022-04487-6>

- Received: 28 September 2021
- Accepted: 28 January 2022
- Published: 09 March 2022
- Issue Date: 17 March 2022
- DOI: <https://doi.org/10.1038/s41586-022-04487-6>

## Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-022-04487-6>

| [Section menu](#) | [Main menu](#) |

- Article
- [Published: 09 March 2022](#)

# Discovery of a Ni<sup>2+</sup>-dependent guanidine hydrolase in bacteria

- [D. Funck ORCID: orcid.org/0000-0002-9855-0419](#)<sup>1</sup> na1,
- [M. Sinn ORCID: orcid.org/0000-0002-3790-9686](#)<sup>1</sup> na1,
- [J. R. Fleming ORCID: orcid.org/0000-0003-4016-8740](#)<sup>2</sup>,
- [M. Stanoppi ORCID: orcid.org/0000-0003-0644-3091](#)<sup>1</sup>,
- [J. Dietrich](#)<sup>1</sup>,
- [R. López-Igual ORCID: orcid.org/0000-0002-2369-1583](#)<sup>3</sup>,
- [O. Mayans](#)<sup>2,4</sup> &
- [J. S. Hartig ORCID: orcid.org/0000-0001-6601-7217](#)<sup>1,4</sup>

*Nature* volume 603, pages 515–521 (2022)

- 2719 Accesses
- 96 Altmetric
- [Metrics details](#)

## Subjects

- [Bacterial physiology](#)
- [Hydrolases](#)
- [Metabolic pathways](#)
- [Microbial ecology](#)
- [X-ray crystallography](#)

# Abstract

Nitrogen availability is a growth-limiting factor in many habitats<sup>1</sup>, and the global nitrogen cycle involves prokaryotes and eukaryotes competing for this precious resource. Only some bacteria and archaea can fix elementary nitrogen; all other organisms depend on the assimilation of mineral or organic nitrogen. The nitrogen-rich compound guanidine occurs widely in nature<sup>2,3,4</sup>, but its utilization is impeded by pronounced resonance stabilization<sup>5</sup>, and enzymes catalysing hydrolysis of free guanidine have not been identified. Here we describe the arginase family protein GdmH (Sll1077) from *Synechocystis* sp. PCC 6803 as a Ni<sup>2+</sup>-dependent guanidine hydrolase. GdmH is highly specific for free guanidine. Its activity depends on two accessory proteins that load Ni<sup>2+</sup> instead of the typical Mn<sup>2+</sup> ions into the active site. Crystal structures of GdmH show coordination of the dinuclear metal cluster in a geometry typical for arginase family enzymes and allow modelling of the bound substrate. A unique amino-terminal extension and a tryptophan residue narrow the substrate-binding pocket and identify homologous proteins in further cyanobacteria, several other bacterial taxa and heterokont algae as probable guanidine hydrolases. This broad distribution suggests notable ecological relevance of guanidine hydrolysis in aquatic habitats.

This is a preview of subscription content

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

[Subscribe to Journal](#)

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

[Buy article](#)

Get time limited or full article access on ReadCube.

\$32.00

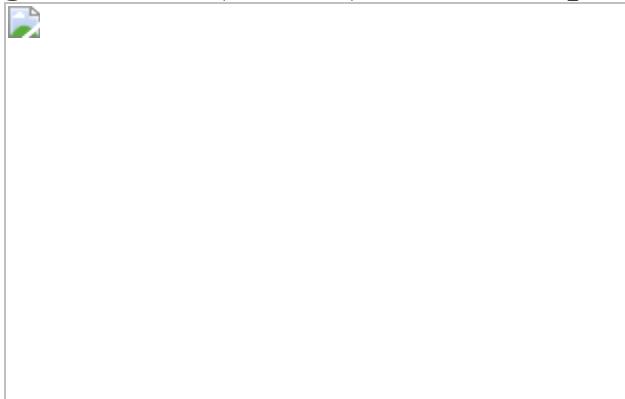
[Buy](#)

All prices are NET prices.

### **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

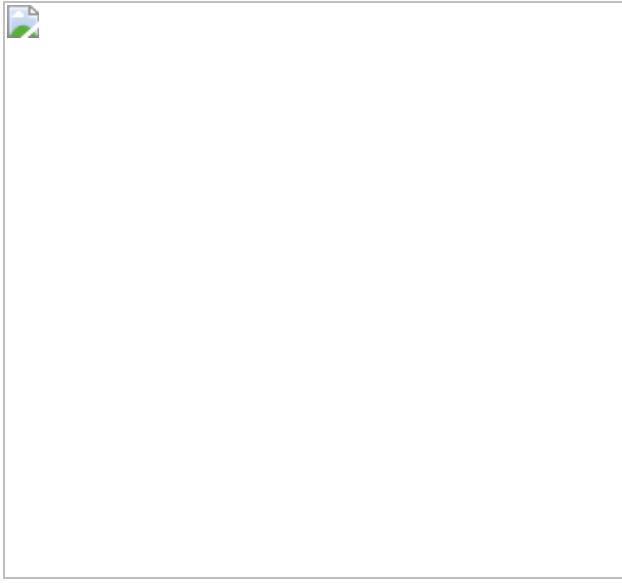
**Fig. 1: GdmH (Sll1077) is a Ni<sup>2+</sup>-dependent guanidine hydrolase.**



**Fig. 2: Kinetic and energetic properties of GdmH.**

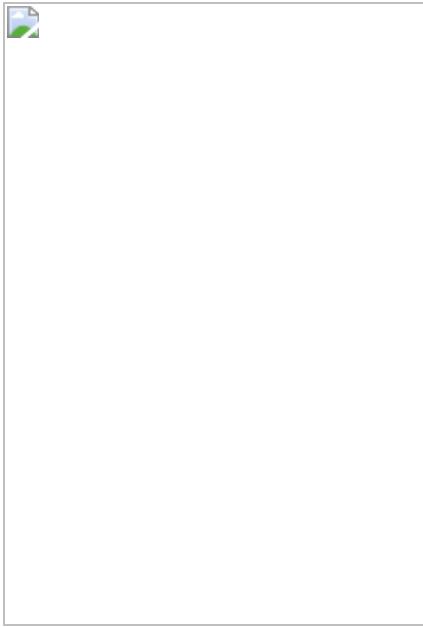


**Fig. 3: Structural characterization of GdmH.**



**Fig. 4: Growth of *Synechocystis* sp. PCC 6803 with guanidine as the sole source of nitrogen.**





## Data availability

Structure coordinates of GdmH from *Synechocystis* sp. PCC6803 and experimental structure factor amplitudes have been deposited with the Protein Data Bank as entries [7OI1](#) and [7ESR](#), for space groups *C*2 and *R*32, respectively. X-ray diffraction images have been deposited with Zenodo, at which <https://doi.org/10.5281/zenodo.4750963> corresponds to PDB entry [7OI1](#) and <https://doi.org/10.5281/zenodo.4750940> corresponds to [7ESR](#). The raw data are presented in the Article and are available from the corresponding authors upon reasonable request.

## References

1. Du, E. et al. Global patterns of terrestrial nitrogen and phosphorus limitation. *Nat. Geosci.* **13**, 221–226 (2020).
2. Schulze, E. Ueber einige stickstoffhaltige Bestandtheile der Keimlinge von *Vicia sativa*. *Z. Phys. Chem.* **17**, 193–216 (1893).
3. Wishart, D. S. et al. HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Res.* **46**, D608–D617 (2018).

4. Kato, T., Yamagata, M. & Tsukahara, S. Guanidine compounds in fruit trees and their seasonal variations in citrus (*Citrus unshiu* Marc.). *J. Jpn. Soc. Hortic. Sci.* **55**, 169–173 (1986).
5. Gund, P. Guanidine, trimethylenemethane, and "Y-delocalization." Can acyclic compounds have "aromatic" stability? *J. Chem. Educ.* **49**, 100 (1972).
6. GÜTHNER, T., MERTSCHENK, B. & SCHULZ, B. In *Ullmann's Fine Chemicals* vol. 2, 657–672 (Wiley-VCH, 2014).
7. Strecker, A. Untersuchungen über die chemischen Beziehungen zwischen Guanin, Xanthin, Theobromin, Caffein und Kreatinin. *Justus Liebigs Ann. Chem.* **118**, 151–177 (1861).
8. Iwanoff, N. N. & Awetissowa, A. N. The fermentative conversion of guanidine in urea. *Biochem. Z.* **231**, 67–78 (1931).
9. Lenkeit, F., Eckert, I., Hartig, J. S. & Weinberg, Z. Discovery and characterization of a fourth class of guanidine riboswitches. *Nucleic Acids Res.* **48**, 12889–12899 (2020).
10. Salvail, H., Balaji, A., Yu, D., Roth, A. & Breaker, R. R. Biochemical validation of a fourth guanidine riboswitch class in bacteria. *Biochemistry* **59**, 4654–4662 (2020).
11. Nelson, J. W., Atilho, R. M., Sherlock, M. E., Stockbridge, R. B. & Breaker, R. R. Metabolism of free guanidine in bacteria is regulated by a widespread riboswitch class. *Mol. Cell* **65**, 220–230 (2017).
12. Sherlock, M. E. & Breaker, R. R. Biochemical validation of a third guanidine riboswitch class in bacteria. *Biochemistry* **56**, 359–363 (2016).
13. Sherlock, M. E., Malkowski, S. N. & Breaker, R. R. Biochemical validation of a second guanidine riboswitch class in bacteria. *Biochemistry* **56**, 352–358 (2016).

14. Kermani, A. A., Macdonald, C. B., Gundepudi, R. & Stockbridge, R. B. Guanidinium export is the primal function of SMR family transporters. *Proc. Natl Acad. Sci. USA* **115**, 3060–3065 (2018).
15. Sinn, M., Hauth, F., Lenkeit, F., Weinberg, Z. & Hartig, J. S. Widespread bacterial utilization of guanidine as nitrogen source. *Mol. Microbiol.* **116**, 200–210 (2021).
16. Schneider, N. O. et al. Solving the conundrum: widespread proteins annotated for urea metabolism in bacteria are carboxyguanidine deiminases mediating nitrogen assimilation from guanidine. *Biochemistry* **59**, 3258–3270 (2020).
17. Zhao, J., Zhu, L., Fan, C., Wu, Y. & Xiang, S. Structure and function of urea amidolyase. *Biosci. Rep.* **38**, BSR20171617 (2018).
18. Mobley, H. L., Island, M. D. & Hausinger, R. P. Molecular biology of microbial ureases. *Microbiol. Rev.* **59**, 451–480 (1995).
19. Mazzei, L., Musiani, F. & Ciurli, S. The structure-based reaction mechanism of urease, a nickel dependent enzyme: tale of a long debate. *J. Biol. Inorg. Chem.* **25**, 829–845 (2020).
20. Uribe, E. et al. Functional analysis of the Mn<sup>2+</sup> requirement in the catalysis of ureohydrolases arginase and agmatinase - a historical perspective. *J. Inorg. Biochem.* **202**, 110812 (2020).
21. Perozich, J., Hempel, J. & Morris, S. M. Jr Roles of conserved residues in the arginase family. *Biochim. Biophys. Acta* **1382**, 23–37 (1998).
22. Sekowska, A., Danchin, A. & Risler, J. L. Phylogeny of related functions: the case of polyamine biosynthetic enzymes. *Microbiology* **146**, 1815–1828 (2000).
23. Sekula, B. The neighboring subunit is engaged to stabilize the substrate in the active site of plant arginases. *Front. Plant Sci.* **11**, 987 (2020).

24. Quintero, M. J., Muro-Pastor, A. M., Herrero, A. & Flores, E. Arginine catabolism in the cyanobacterium *Synechocystis* sp. strain PCC 6803 involves the urea cycle and arginase pathway. *J. Bacteriol.* **182**, 1008–1015 (2000).
25. Lacasse, M. J., Summers, K. L., Khorasani-Motlagh, M., George, G. N. & Zamble, D. B. Bimodal nickel-binding site on *Escherichia coli* [NiFe]-hydrogenase metallochaperone HypA. *Inorg. Chem.* **58**, 13604–13618 (2019).
26. Hoffmann, D., Gutekunst, K., Klissenbauer, M., Schulz-Friedrich, R. & Appel, J. Mutagenesis of hydrogenase accessory genes of *Synechocystis* sp. PCC 6803. *FEBS J.* **273**, 4516–4527 (2006).
27. Dowling, D. P., Di Costanzo, L., Gennadios, H. A. & Christianson, D. W. Evolution of the arginase fold and functional diversity. *Cell. Mol. Life Sci.* **65**, 2039–2055 (2008).
28. Dutta, A., Mazumder, M., Alam, M., Gourinath, S. & Sau, A. K. Metal-induced change in catalytic loop positioning in *Helicobacter pylori* arginase alters catalytic function. *Biochem. J.* **476**, 3595–3614 (2019).
29. Di Costanzo, L. et al. Crystal structure of human arginase I at 1.29-Å resolution and exploration of inhibition in the immune response. *Proc. Natl Acad. Sci. USA* **102**, 13058–13063 (2005).
30. Suzek, B. E. et al. UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. *Bioinformatics* **31**, 926–932 (2014).
31. Alfano, M. & Cavazza, C. Structure, function, and biosynthesis of nickel-dependent enzymes. *Protein Sci.* **29**, 1071–1089 (2020).
32. Wang, B. et al. A guanidine-degrading enzyme controls genomic stability of ethylene-producing cyanobacteria. *Nat. Commun.* **12**, 5150 (2021).

33. McGee, D. J. et al. Purification and characterization of *Helicobacter pylori* arginase, RocF: unique features among the arginase superfamily. *Eur. J. Biochem.* **271**, 1952–1962 (2004).
34. Arakawa, N., Igarashi, M., Kazuoka, T., Oikawa, T. & Soda, K. d-Arginase of *Arthrobacter* sp. KUJ 8602: characterization and its identity with Zn<sup>2+</sup>-guanidinobutyrase. *J. Biochem.* **133**, 33–42 (2003).
35. Saragadam, T., Kumar, S. & Punekar, N. S. Characterization of 4-guanidinobutyrase from *Aspergillus niger*. *Microbiology* **165**, 396–410 (2019).
36. Viator, R. J., Rest, R. F., Hildebrandt, E. & McGee, D. J. Characterization of *Bacillus anthracis* arginase: effects of pH, temperature, and cell viability on metal preference. *BMC Biochem.* **9**, 15 (2008).
37. D'Antonio, E. L., Hai, Y. & Christianson, D. W. Structure and function of non-native metal clusters in human arginase I. *Biochemistry* **51**, 8399–8409 (2012).
38. Andresen, E., Peiter, E. & Küpper, H. Trace metal metabolism in plants. *J. Exp. Bot.* **69**, 909–954 (2018).
39. Eisenhut, M. Manganese homeostasis in cyanobacteria. *Plants* **9**, 18 (2019).
40. Burnat, M. & Flores, E. Inactivation of agmatinase expressed in vegetative cells alters arginine catabolism and prevents diazotrophic growth in the heterocyst-forming cyanobacterium *Anabaena*. *MicrobiologyOpen* **3**, 777–792 (2014).
41. Callahan, B. P., Yuan, Y. & Wolfenden, R. The burden borne by urease. *J. Am. Chem. Soc.* **127**, 10828–10829 (2005).
42. Lewis, C. A. Jr & Wolfenden, R. The nonenzymatic decomposition of guanidines and amidines. *J. Am. Chem. Soc.* **136**, 130–136 (2014).

43. Grobben, Y. et al. Structural insights into human Arginase-1 pH dependence and its inhibition by the small molecule inhibitor CB-1158. *J. Struct. Biol.* **X4**, 100014 (2020).
44. Mills, L. A., McCormick, A. J. & Lea-Smith, D. J. Current knowledge and recent advances in understanding metabolism of the model cyanobacterium *Synechocystis* sp. PCC 6803. *Biosci. Rep.* **40**, BSR20193325 (2020).
45. Giner-Lamia, J. et al. Identification of the direct regulon of NtcA during early acclimation to nitrogen starvation in the cyanobacterium *Synechocystis* sp PCC 6803. *Nucleic Acids Res.* **45**, 11800–11820 (2017).
46. Martinez, S. & Hausinger, R. P. Biochemical and spectroscopic characterization of the non-heme Fe(II)- and 2-oxoglutarate-dependent ethylene-forming enzyme from *Pseudomonas syringae* pv. *phaseolicola* PK2. *Biochemistry* **55**, 5989–5999 (2016).
47. Copeland, R. A. et al. An iron(IV)-oxo intermediate initiating l-arginine oxidation but not ethylene production by the 2-oxoglutarate-dependent oxygenase, ethylene-forming enzyme. *J. Am. Chem. Soc.* **143**, 2293–2303 (2021).
48. Rippka, R., Deruelles, J., Waterbury, J. B., Herdman, M. & Stanier, R. Y. Generic assignments, strain histories and properties of pure cultures of cyanobacteria. *Microbiology* **111**, 1–61 (1979).
49. Geyer, J. W. & Dabich, D. Rapid method for determination of arginase activity in tissue homogenates. *Anal. Biochem.* **39**, 412–417 (1971).
50. van Anken, H. C. & Schiphorst, M. E. A kinetic determination of ammonia in plasma. *Clin. Chim. Acta* **56**, 151–157 (1974).
51. Kabsch, W. XDS. *Acta Crystallogr. D* **66**, 125–132 (2010).
52. McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).

53. Lamzin, V. S. P. A., Wilson, K. S. In International Tables for Crystallography Vol. F (eds Arnold, E. et al.) 525–528 (Kluwer, 2012).
54. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
55. Adams, P. D. et al. The Phenix software for automated determination of macromolecular structures. *Methods* **55**, 94–106 (2011).
56. Williams, C. J. et al. MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2018).
57. Wang, J., Wang, W., Kollman, P. A. & Case, D. A. Automatic atom type and bond type perception in molecular mechanical calculations. *J. Mol. Graph. Model.* **25**, 247–260 (2006).
58. Maier, J. A. et al. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).
59. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A. & Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174 (2004).
60. Trott, O. & Olson, A. J. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–461 (2010).
61. Ashkenazy, H. et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res.* **44**, W344–W350 (2016).
62. Lemoine, F. et al. Renewing Felsenstein’s phylogenetic bootstrap in the era of big data. *Nature* **556**, 452–456 (2018).

63. Lemoine, F. et al. NGPhylogeny.fr: new generation phylogenetic services for non-specialists. *Nucleic Acids Res.* **47**, W260–W265 (2019).
64. Letunic, I. & Bork, P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* **47**, W256–W259 (2019).
65. Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).

## Acknowledgements

J.S.H. acknowledges the ERC CoG project 681777 “RiboDisc” for financial support. R.L.-I. was supported by grant PID2019-104784RJ-100 MCIN/AEI/10.13039/501100011033 Spain, and grant BFU2017-88202-P MCIN/AEI/10.13039/501100011033 Spain, cofinanced by ERDF A way of making Europe. We thank R. Winter and E. Flores for helpful discussions and A. Joachimi and D. Galetskiy for technical assistance. We acknowledge the Paul Scherrer Institut, Villigen, Switzerland, for provision of synchrotron radiation beamtime at beamline X06SA-PXI of the Swiss Light Source, and thank K. M. L. Smith for assistance. We thank K. Forchhammer for providing the *Synechocystis* PCC 6803 GT cells and for helpful advice regarding their cultivation.

## Author information

### Author notes

1. These authors contributed equally: D. Funck, M. Sinn

## Affiliations

1. Department of Chemistry, University of Konstanz, Konstanz, Germany  
D. Funck, M. Sinn, M. Stanoppi, J. Dietrich & J. S. Hartig

2. Department of Biology, University of Konstanz, Konstanz, Germany

J. R. Fleming & O. Mayans

3. Instituto de Bioquímica Vegetal y Fotosíntesis, Universidad de Sevilla  
and C.S.I.C, Seville, Spain

R. López-Igual

4. Konstanz Graduate School Chemical Biology (KoRS-CB), University  
of Konstanz, Konstanz, Germany

O. Mayans & J. S. Hartig

## Contributions

D.F., M. Sinn and J.S.H. conceived the project. D.F. and M. Sinn performed protein expression, purification, activity assays and sequence analyses. R.L.-I. generated the  $\Delta sll1077$  mutant and J.D., R.L.-I. and D.F. performed growth assays. M. Stanoppi performed NMR analyses. J.R.F. and O.M. performed protein crystallization, structure determination, analysis and modelling. D.F., M. Sinn and J.S.H. wrote the manuscript with input from all authors. D.F., M. Sinn, J.D., M. Stanoppi and J.R.F. prepared figures. The manuscript was reviewed and approved by all coauthors.

## Corresponding author

Correspondence to [J. S. Hartig](#).

## Ethics declarations

## Competing interests

The authors declare no competing interests.

## Peer review

## Peer review information

*Nature* thanks David Richardson and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

## Additional information

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Extended data figures and tables

### [Extended Data Fig. 1 Additional characterization GdmH.](#)

**a:** Activation of GdmH expressed in the absence of added metal ions and either co-expressed with GhaA and GhaB or with either of the accessory proteins alone by Ni<sup>2+</sup> or Fe<sup>2+</sup>. The orange columns represent the activity in a mixture of the latter two extracts. Where indicated, 1 mM β-mercaptoethanol (β-MSH) was additionally included. Columns represent the average of n = 3 technical replicates, error bars indicate s.d. The entire experiment was repeated with independent bacterial extracts, which yielded slightly different absolute values for specific activities but virtually identical results for the relative values. **b:** <sup>13</sup>C-NMR spectra of <sup>13</sup>C<sup>15</sup>N-guanidine after incubation over night with purified GdmH (overexpressed together with GhaA and GhaB). Exclusively the triplet signal of <sup>13</sup>C<sup>15</sup>N-urea was detected after incubation with purified GdmH, whereas after incubation with GdmH partially inactivated by heat treatment, both the quadruplet of <sup>13</sup>C<sup>15</sup>N-guanidine and the triplet signal of <sup>13</sup>C<sup>15</sup>N-urea were detected. **c:** Coupled enzymatic assay of GdmH with glutamate dehydrogenase (GDH). NADH gets oxidized by GDH during reductive amination of α-ketoglutarate with ammonia released from guanidine by GdmH. Michaelis-constant  $K_M$ , maximal specific activity  $A_{max}$ , and catalytic constant  $k_{cat}$  were determined using different guanidine concentrations. Data points represent the average of n = 3 technical

replicates and the black line represents the least-square fit to the Michaelis-Menten equation. Error bars represent s.d. The whole experiment was repeated with an independent enzyme preparation, which gave consistent results. **d:** Influence of Na-cacodylate on the activity of GdmH with 10 mM guanidine as substrate. A cacodylate molecule occupied the active site in one of the crystal structures but even a tenfold excess of cacodylate had no influence on GdmH activity. Columns represent the average of  $n = 3$  independent enzyme preparations, error bars indicate s.d. **e:** Effect of point mutations on the activity and  $K_M$  of purified, recombinant GdmH. The black lines represent the least square fits to the Michaelis-Menten equation of single experiments. The entire analysis was repeated with independent enzyme preparations with consistent results. The inset shows a Coomassie-stained gel with the purified GdmH variants and the positions of molecular weight markers. For gel source data, see Supplementary Fig. 1b. **f:** Time-dependence of urea production by GdmH in the presence of 10 mM guanidine. Note that almost half of the substrate was hydrolyzed after 3 days of incubation.

### [Extended Data Fig. 2 Additional images of the GdmH structure.](#)

**a:** Top view of the GdmH hexamer with one subunit in surface display (bright yellow) and two subunits as ribbons below a transparent surface (orange and sky blue). The extended N-terminus of the orange subunit is highlighted by saturated color. **b:** Comparison of the GdmH capping helix (blue) with the same helix from the most similar protein with a high resolution structure, guanidinobutyrase from *Pseudomonas aeruginosa* (grey with capping helix in yellow, PDB entry 3NIO). Although the overall fold is well conserved between both enzymes, the position of the highlighted  $\alpha$ -helix is shifted towards the active site.

### [Extended Data Fig. 3 Relation of GdmH to other proteins from the arginase superfamily.](#)

Unrooted neighbor-joining tree of 509 representatives of >15000 UniRef90 sequences with at least 27% identity to residues 60–390 of GdmH, selected

and aligned with the ConSurf webserver<sup>61</sup>. Accession numbers of all sequences are given in Supplementary Data 2. The branches are colored according to the taxonomic group: Black: bacteria; cyan: archaea; bright green: fungi; green: plants; brown: stramenopile; magenta: animals. Branches containing selected enzymes with biochemically confirmed function are highlighted and demonstrate that very dissimilar sequences can have the same enzymatic activity in different taxa.

**Extended Data Table 1 Kinetic properties of *Synechocystis* guanidine hydrolase GdmH**

**Extended Data Table 2 X-ray diffraction data collection and refinement statistics**

## Supplementary information

**Supplementary Fig. 1**

Source data for SDS gel (Fig. 1e); source data for SDS gel (Extended Data Fig. 1e) as pdf file.

**Reporting Summary**

**Peer Review File**

**Supplementary Data 1**

Multiple-sequence alignment of GdmH with similar sequences as fasta formatted plain text file.

**Supplementary Data 2**

Manually optimized sequence alignment of GdmH with further sequences of characterized enzymes of the arginase family (source data for Extended Data Fig. 3) as fasta formatted plain text file.

**Supplementary Data 3**

Multiple-sequence alignment of the 194 closest homologues of GdmH (source data for Fig. 3g) as fasta formatted plain text file.

## Rights and permissions

### [Reprints and Permissions](#)

## About this article

### Cite this article

Funck, D., Sinn, M., Fleming, J.R. *et al.* Discovery of a Ni<sup>2+</sup>-dependent guanidine hydrolase in bacteria. *Nature* **603**, 515–521 (2022).  
<https://doi.org/10.1038/s41586-022-04490-x>

- Received: 09 June 2021
- Accepted: 31 January 2022
- Published: 09 March 2022
- Issue Date: 17 March 2022
- DOI: <https://doi.org/10.1038/s41586-022-04490-x>

## Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-022-04490-x>

| [Section menu](#) | [Main menu](#) |

- Article
- [Published: 02 March 2022](#)

# Overcoming universal restrictions on metal selectivity by protein design

- [Tae Su Choi ORCID: orcid.org/0000-0002-1515-3275<sup>1</sup>](#) &
- [F. Akif Tezcan ORCID: orcid.org/0000-0002-4733-6500<sup>1</sup>](#)

*Nature* volume **603**, pages 522–527 (2022)

- 4920 Accesses
- 102 Altmetric
- [Metrics details](#)

## Subjects

- [Biophysical chemistry](#)
- [Metalloproteins](#)
- [Protein design](#)
- [Structural biology](#)

## Abstract

Selective metal coordination is central to the functions of metalloproteins:<sup>1,2</sup> each metalloprotein must pair with its cognate metallocofactor to fulfil its biological role<sup>3</sup>. However, achieving metal selectivity solely through a three-dimensional protein structure is a great

challenge, because there is a limited set of metal-coordinating amino acid functionalities and proteins are inherently flexible, which impedes steric selection of metals<sup>3,4</sup>. Metal-binding affinities of natural proteins are primarily dictated by the electronic properties of metal ions and follow the Irving–Williams series<sup>5</sup> ( $Mn^{2+} < Fe^{2+} < Co^{2+} < Ni^{2+} < Cu^{2+} > Zn^{2+}$ ) with few exceptions<sup>6,7</sup>. Accordingly, metalloproteins overwhelmingly bind  $Cu^{2+}$  and  $Zn^{2+}$  in isolation, regardless of the nature of their active sites and their cognate metal ions<sup>1,3,8</sup>. This led organisms to evolve complex homeostatic machinery and non-equilibrium strategies to achieve correct metal speciation<sup>1,3,8,9,10</sup>. Here we report an artificial dimeric protein,  $(AB)_2$ , that thermodynamically overcomes the Irving–Williams restrictions in vitro and in cells, favouring the binding of lower-Irving–Williams transition metals over  $Cu^{2+}$ , the most dominant ion in the Irving–Williams series. Counter to the convention in molecular design of achieving specificity through structural preorganization,  $(AB)_2$  was deliberately designed to be flexible. This flexibility enabled  $(AB)_2$  to adopt mutually exclusive, metal-dependent conformational states, which led to the discovery of structurally coupled coordination sites that disfavour  $Cu^{2+}$  ions by enforcing an unfavourable coordination geometry. Aside from highlighting flexibility as a valuable element in protein design, our results illustrate design principles for constructing selective metal sequestration agents.

This is a preview of subscription content

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

[Subscribe to Journal](#)

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

[Buy article](#)

Get time limited or full article access on ReadCube.

\$32.00

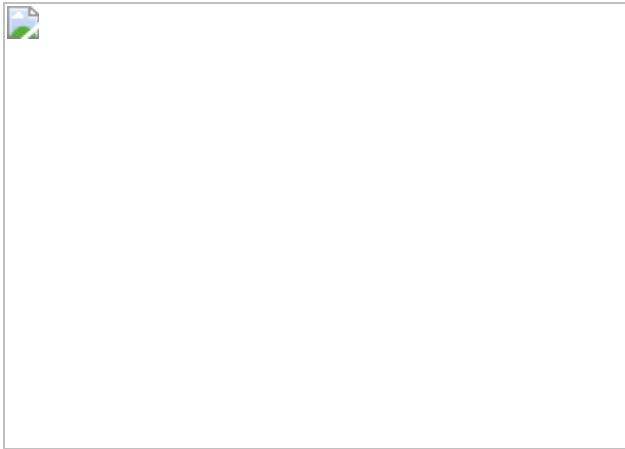
[Buy](#)

All prices are NET prices.

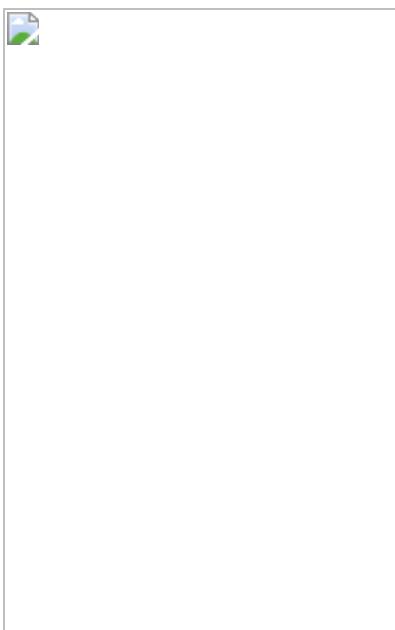
### **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

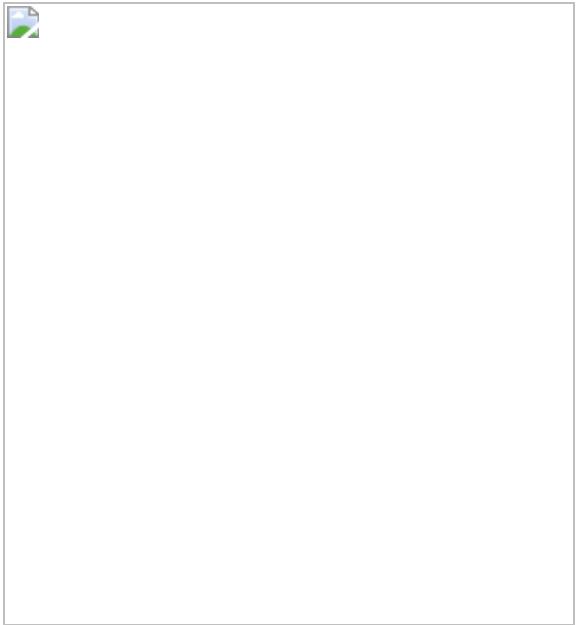
**Fig. 1: Design and characterization of the (AB)<sub>2</sub> scaffold.**



**Fig. 2:  $\text{Co}^{2+}$  and  $\text{Ni}^{2+}$  selectivity of  $(\text{AB})_2$ .**



**Fig. 3: Primary metal coordination spheres of central and peripheral sites in  $(\text{AB})_2$ .**



**Fig. 4: Selective  $\text{Co}^{2+}$  and  $\text{Ni}^{2+}$  binding of  $(\text{AB})_2$  in the *E. coli* periplasmic space.**



## Data availability

The principal data supporting the findings of this work are available within the figures and the [Supplementary Information](#). Additional data that support the findings of this study are available from the corresponding author on request. Crystallographic data for protein structures (coordinates and structure factors) have been deposited into the RCSB PDB under the following accession codes: [7MK4](#), [7LRV](#), [7LV1](#), [7N4G](#), [7N4F](#), [7LR5](#), [7LRA](#), [7LRB](#) and [7LRR](#). The model structure used for molecular replacement is available in the PDB ([2BC5](#)).

## References

1. Waldron, K. J., Rutherford, J. C., Ford, D. & Robinson, N. J. Metalloproteins and metal sensing. *Nature* **460**, 823–830 (2009).
2. Gray, H. B., Stiefel, E. I., Valentine, J. S. & Bertini, I. Biological Inorganic Chemistry: Structure and Reactivity (University Science Books, 2007).
3. Waldron, K. J. & Robinson, N. J. How do bacterial cells ensure that metalloproteins get the correct metal? *Nat. Rev. Microbiol.* **7**, 25–35 (2009).
4. Dudev, T. & Lim, C. Competition among metal ions for protein binding sites: determinants of metal ion selectivity in proteins. *Chem. Rev.* **114**, 538–556 (2014).
5. Frausto da Silva, J. J. R. & Williams, R. J. P. The Biological Chemistry of the Elements (Oxford University Press, 2001).
6. Kisgeropoulos, E. C. et al. Key structural motifs balance metal binding and oxidative reactivity in a heterobimetallic Mn/Fe protein. *J. Am. Chem. Soc.* **142**, 5338–5354 (2020).
7. Grāve, K., Griese, J. J., Berggren, G., Bennett, M. D. & Högbom, M. The *Bacillus anthracis* class Ib ribonucleotide reductase subunit NrdF intrinsically selects manganese over iron. *J. Biol. Inorg. Chem.* **25**, 571–582 (2020).
8. Reyes-Caballero, H., Campanello, G. C. & Giedroc, D. P. Metalloregulatory proteins: metal selectivity and allosteric switching. *Biophys. Chem.* **156**, 103–114 (2011).
9. O'Halloran, T. V. & Culotta, V. C. Metallochaperones, an intracellular shuttle service for metal ions. *J. Biol. Chem.* **275**, 25057–25060 (2000).
10. Tottey, S. et al. Protein-folding location can regulate manganese-binding versus copper- or zinc-binding. *Nature* **455**, 1138–1142 (2008).

11. Lombardi, A., Pirro, F., Maglio, O., Chino, M. & DeGrado, W. F. De novo design of four-helix bundle metalloproteins: one scaffold, diverse reactivities. *Acc. Chem. Res.* **52**, 1148–1159 (2019).
12. Lu, Y., Yeung, N., Sieracki, N. & Marshall, N. M. Design of functional metalloproteins. *Nature* **460**, 855–862 (2009).
13. Yu, F. et al. Protein design: toward functional metalloenzymes. *Chem. Rev.* **114**, 3495–3578 (2014).
14. Schwizer, F. et al. Artificial metalloenzymes: reaction scope and optimization strategies. *Chem. Rev.* **118**, 142–231 (2018).
15. Churchfield, L. A. & Tezcan, F. A. Design and construction of functional supramolecular metalloprotein assemblies. *Acc. Chem. Res.* **52**, 345–355 (2019).
16. Faiella, M. et al. An artificial di-iron oxo-protein with phenol oxidase activity. *Nat. Chem. Biol.* **5**, 882–884 (2009).
17. Zastrow, M. L., Peacock, F. A., Stuckey, J. A. & Pecoraro, V. L. Hydrolytic catalysis and structural stabilization in a designed metalloprotein. *Nat. Chem.* **4**, 118–123 (2012).
18. Studer, S. et al. Evolution of a highly active and enantiospecific metalloenzyme from short peptides. *Science* **362**, 1285–1288 (2018).
19. Khare, S. D. et al. Computational redesign of a mononuclear zinc metalloenzyme for organophosphate hydrolysis. *Nat. Chem. Biol.* **8**, 294–300 (2012).
20. Yeung, N. et al. Rational design of a structural and functional nitric oxide reductase. *Nature* **462**, 1079–1082 (2009).
21. Song, W. J. & Tezcan, F. A. A designed supramolecular protein assembly with in vivo enzymatic activity. *Science* **346**, 1525–1528 (2014).

22. Churchfield, L. A., Medina-Morales, A., Brodin, J. D., Perez, A. & Tezcan, F. A. De novo design of an allosteric metalloprotein assembly with strained disulfide bonds. *J. Am. Chem. Soc.* **138**, 13163–13166 (2016).
23. Zhou, L. et al. A protein engineered to bind uranyl selectively and with femtomolar affinity. *Nat. Chem.* **6**, 236–241 (2014).
24. Wegner, S. V., Boyaci, H., Chen, H., Jensen, M. P. & He, C. Engineering a uranyl-specific binding protein from NikR. *Angew. Chem. Int. Ed. Engl.* **48**, 2339–2341 (2009).
25. Brodin, J. D. et al. Evolution of metal selectivity in templated protein interfaces. *J. Am. Chem. Soc.* **132**, 8610–8617 (2010).
26. Guffy, S. L., Der, B. S. & Kuhlman, B. Probing the minimal determinants of zinc binding with computational protein design. *Protein Eng. Des. Sel.* **29**, 327–338 (2016).
27. Akcapinar, G. B. & Sezerman, O. U. Computational approaches for de novo design and redesign of metal-binding sites on proteins. *Biosci. Rep.* **37**, BSR20160179 (2017).
28. Byrd, J. & Winge, D. R. Cooperative cluster formation in metallothionein. *Arch. Biochem. Biophys.* **250**, 233–237 (1986).
29. Halling, D. B., Liebeskind, B. J., Hall, A. W. & Aldrich, R. W. Conserved properties of individual  $\text{Ca}^{2+}$ -binding sites in calmodulin. *Proc. Natl Acad. Sci. USA* **113**, E1216–E1225 (2016).
30. Zygiel, E. M. & Nolan, E. M. Transition metal sequestration by the host-defense protein calprotectin. *Annu. Rev. Biochem.* **87**, 621–643 (2018).
31. Rittle, J., Field, M. J., Green, M. T. & Tezcan, F. A. An efficient, step-economical strategy for the design of functional metalloproteins. *Nat. Chem.* **11**, 434–441 (2019).

32. Faraone-Mennella, J., Tezcan, F. A., Gray, H. B. & Winkler, J. R. Stability and folding kinetics of structurally characterized cytochrome *c*-*b*<sub>562</sub>. *Biochemistry* **45**, 10504–10511 (2006).
33. Choi, T. S., Lee, H. J., Han, J. Y., Lim, M. H. & Kim, H. I. Molecular insights into human serum albumin as a receptor of amyloid- $\beta$  in the extracellular region. *J. Am. Chem. Soc.* **139**, 15437–15445 (2017).
34. Burgot, J.-L. Ionic Equilibria in Analytical Chemistry (Springer, 2012).
35. Osman, D. et al. Bacterial sensors define intracellular free energies for correct enzyme metalation. *Nat. Chem. Biol.* **15**, 241–249 (2019).
36. Young, T. R. et al. Calculating metalation in cells reveals CobW acquires Co<sup>II</sup> for vitamin B<sub>12</sub> biosynthesis while related proteins prefer Zn<sup>II</sup>. *Nat. Commun.* **12**, 1195 (2021).
37. Jeschek, M. et al. Directed evolution of artificial metalloenzymes for in vivo metathesis. *Nature* **537**, 661–665 (2016).
38. Thompson, A. N. et al. Mechanism of potassium-channel selectivity revealed by Na<sup>+</sup> and Li<sup>+</sup> binding sites within the KcsA pore. *Nat. Struct. Mol. Biol.* **16**, 1317–1324 (2009).
39. Capdevila, D. A., Braymer, J. J., Edmonds, K. A., Wu, H. & Giedroc, D. P. Entropy redistribution controls allosteric control in a metalloregulatory protein. *Proc. Natl Acad. Sci. USA* **114**, 4424–4429 (2017).
40. Tokuriki, N. & Tawfik, D. S. Protein dynamism and evolvability. *Science* **324**, 203–207 (2009).
41. Motlagh, H. N., Wrabl, J. O., Li, J. & Hilser, V. J. The ensemble nature of allosteric control. *Nature* **508**, 331–339 (2014).
42. Papaleo, E. et al. The role of protein loops and linkers in conformational dynamics and allosteric control. *Chem. Rev.* **116**, 6391–6423 (2016).

43. Arslan, E., Schulz, H., Zufferey, R., Künzler, P. & Thöny-Meyer, L. Overproduction of the *Bradyrhizobium japonicum* c-type cytochrome subunits of the *cbb<sub>3</sub>* oxidase in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* **251**, 744–747 (1998).
44. Bailey, J. B., Subramanian, R. H., Churchfield, L. A. & Tezcan, F. A. in Peptide, Protein and Enzyme Design: Methods in Enzymology Vol. 580 (ed. Pecoraro, V. L.) 223–250 (Academic Press, 2016).
45. Martel, A., Liu, P., Weiss, T. M., Niebuhr, M. & Tsuruta, H. An integrated high-throughput data acquisition system for biological solution X-ray scattering studies. *J. Synchrotron Radiat.* **19**, 431–434 (2012).
46. Manalastas-Cantos, K. et al. ATSAS 3.0: expanded functionality and new tools for small-angle scattering data analysis. *J. Appl. Crystallogr.* **54**, 343–355 (2021).
47. Svergun, D., Barberato, C. & Koch, M. H. J. CRYSTOL—a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* **28**, 768–773 (1995).
48. Collaborative Computational Project. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
49. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).
50. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).
51. The PyMOL Molecular Graphics System v.1.8 (Schrödinger, 2015).
52. Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and Lamm equation modeling. *Biophys. J.* **78**, 1606–1619 (2000).

53. Manoil, C. & Beckwith, J. A genetic approach to analyzing membrane protein topology. *Science* **233**, 1403–1408 (1986).
54. Dapprich, S., Komáromi, I., Byun, K. S., Morokuma, K. & Frisch, M. J. A new ONIOM implementation in Gaussian98. Part I. The calculation of energies, gradients, vibrational frequencies and electric field derivatives. *J. Mol. Struct. THEOCHEM* **461–462**, 1–21 (1999).
55. Vreven, T., Morokuma, K., Farkas, Ö., Schlegel, H. B. & Frisch, M. J. Geometry optimization with QM/MM, ONIOM, and other combined methods. I. Microiterations and constraints. *J. Comput. Chem.* **24**, 760–769 (2003).
56. Tao, P. et al. Matrix metalloproteinase 2 inhibition: combined quantum mechanics and molecular mechanics studies of the inhibition mechanism of (4-phenoxyphenylsulfonyl)methylthiirane and its oxirane analogue. *Biochemistry* **48**, 9839–9847 (2009).
57. Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **98**, 5648–5652 (1993).
58. Lee, C., Yang, W. & Parr, R. G. Development of the Colle–Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **37**, 785–789 (1988).
59. Hariharan, P. C. & Pople, J. A. The effect of d-functions on molecular orbital energies for hydrocarbons. *Chem. Phys. Lett.* **16**, 217–219 (1972).
60. Rassolov, V. A., Pople, J. A., Ratner, M. A. & Windus, T. L. 6-31G\* basis set for atoms K through Zn. *J. Chem. Phys.* **109**, 1223–1229 (1998).
61. Rassolov, V. A., Ratner, M. A., Pople, J. A., Redfern, P. C. & Curtiss, L. A. 6-31G\* basis set for third-row atoms. *J. Comput. Chem.* **22**, 976–984 (2001).

62. Freindorf, M., Shao, Y., Furlani, T. R. & Kong, J. Lennard-Jones parameters for the combined QM/MM method using the B3LYP/6-31G\*/AMBER potential. *J. Comput. Chem.* **26**, 1270–1278 (2005).
63. Case, D. A. et al. The Amber biomolecular simulation programs. *J. Comput. Chem.* **26**, 1668–1688 (2005).
64. Bakowies, D. & Thiel, W. Hybrid models for combined quantum mechanical and molecular mechanical approaches. *J. Phys. Chem.* **100**, 10580–10594 (1996).
65. Weiner, S. J., Singh, U. C. & Kollman, P. A. Simulation of formamide hydrolysis by hydroxide ion in the gas phase and in aqueous solution. *J. Am. Chem. Soc.* **107**, 2219–2229 (1985).
66. Kakkis, A., Gagnon, D., Esselborn, J., Britt, R. D. & Tezcan, F. A. Metal-templated design of chemically switchable protein assemblies with high-affinity coordination sites. *Angew. Chem. Int. Ed. Engl.* **59**, 21940–21944 (2020).
67. Kocyła, A., Pomorski, A. & Krężel, A. Molar absorption coefficients and stability constants of metal complexes of 4-(2-pyridylazo)resorcinol (PAR): revisiting common chelating probe for the study of metalloproteins. *J. Inorg. Biochem.* **152**, 82–92 (2015).
68. Kuzmič, P. Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase. *Anal. Biochem.* **237**, 260–273 (1996).
69. Stoll, S. & Schweiger, A. EasySpin, a comprehensive software package for spectral simulation and analysis in EPR. *J. Magn. Reson.* **178**, 42–55 (2006).
70. Smilgies, D.-M. & Folta-Stogniew, E. Molecular weight-gyration radius relation of globular proteins: a comparison of light scattering, small-angle X-ray scattering and structure-based data. *J. Appl. Crystallogr.* **48**, 1604–1606 (2015).

## Acknowledgements

We thank members of the Tezcan group, J. Rittle, H. Gray, S. Cohen and M. Green for discussions. We also thank A. Kakkis and C.-J. Yu for AUC and EPR measurements, respectively. This work was funded primarily by the National Institutes of Health (NIH; R01-GM138884) and by NASA (80NSSC18M0093; ENIGMA: Evolution of Nanomachines in Geospheres and Microbial Ancestors (NASA Astrobiology Institute Cycle 8)). Parts of this research were carried out at the Stanford Synchrotron Radiation Lightsource (supported by the DOE, Office of Basic Energy Sciences contract DE-AC02-76SF00515 and NIH P30-GM133894) and the Advanced Light Source (supported by the DOE, Office of Basic Energy Sciences contract DE-AC02-05CH11231 and NIH P30-GM124169-01). EPR data was acquired on an instrument obtained through a Major Research Instrumentation fund supported by the National Science Foundation (NSF; CHE-2019066).

## Author information

### Affiliations

1. Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, CA, USA

Tae Su Choi & F. Akif Tezcan

### Contributions

T.S.C. conceived the project, designed and performed all experiments, calculations and analyses and co-wrote the paper. F.A.T. conceived and directed the project and co-wrote the manuscript.

### Corresponding author

Correspondence to [F. Akif Tezcan](#).

# Ethics declarations

## Competing interests

The authors declare no competing interests.

## Peer review

### Peer review information

*Nature* thanks Jan Lipfert and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

## Additional information

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Extended data figures and tables

### [Extended Data Fig. 1 Analysis of the oligomerization states of \(AB\)<sub>2</sub>, \(AC\)<sub>2</sub> and \(BC\)<sub>2</sub> variants.](#)

**a**, Sedimentation velocity/AUC analysis of the oligomerization state of (AB)<sub>2</sub> (black), (AC)<sub>2</sub> (red), and (BC)<sub>2</sub> (blue) at 25 μM dimer concentration. The inset shows a higher-order oligomer population of (BC)<sub>2</sub>. **b–d**, Guinier analysis of (b) AB (black), (c) AC (red), and (d) BC (blue) constructs measured by solution small-angle X-ray scattering (SAXS) at 0.8 mM dimer concentration. Larger I(0) values in (AC)<sub>2</sub> and (BC)<sub>2</sub> indicate higher oligomeric states compared to (AB)<sub>2</sub>. Radius of gyration ( $R_g$ ) value of (AB)<sub>2</sub> corresponds to approximately 25 kDa in  $M_w - R_g$  relationship of globular proteins<sup>70</sup>, indicating a dimeric state for (AB)<sub>2</sub>.

## Extended Data Fig. 2 Structural characterization and metal-binding analysis of M<sup>II</sup>-(AB)<sub>2</sub> complexes.

**a**, X-band EPR spectrum of Cu<sup>II</sup>-(AB)<sub>2</sub> (orange) and simulated pattern (black line), along with the fit parameters consistent with a rhombic coordination geometry. Conditions: 2.5 mM Cu<sup>II</sup>, 20 mM MOPS (pH 7.4) and 150 mM NaCl, 298 K. **b**, Crystal structure of 2Co<sup>II</sup>-(AB)<sub>2</sub>. Co<sup>II</sup> ions are represented as magenta spheres. **c**, Solution characterization of 2Co<sup>II</sup>-(AB)<sub>2</sub> complex by competitive Fura-2 titration (left), SAXS (middle), and ESI-MS (right). Circles (magenta) in ESI-MS spectrum represent the number of Co<sup>II</sup> ions bound to (AB)<sub>2</sub>. Experimental data points and error bars in the Fura-2 titration are presented as mean and standard deviation of three independent measurements. **d**, Competitive metal-binding titrations (AB)<sub>2</sub> for Mn<sup>II</sup> (pink) and Zn<sup>II</sup> (grey) binding. Mag-Fura-2 and Fura-2 were used for Mn<sup>II</sup> and Zn<sup>II</sup> titrations, respectively. **e**, Investigation of M<sup>II</sup>-(AB)<sub>2</sub> complexation using ESI-MS: Mn<sup>II</sup> (top), Fe<sup>II</sup> (middle), and Zn<sup>II</sup> (bottom). Mn<sup>II</sup>- and Fe<sup>II</sup>-(AB)<sub>2</sub> complexes were not observed owing to the low metal-binding affinities. **f**, ESI-MS spectra of (AB)<sub>2</sub> collected under Cu<sup>II</sup>/Zn<sup>II</sup> and Zn<sup>II</sup>/Cu<sup>II</sup> competition conditions. Cu<sup>II</sup> outcompetes Zn<sup>II</sup> or forms heterometallic complexes of (AB)<sub>2</sub>. **g**, Theoretical SAXS scattering profiles of M<sup>II</sup>-(AB)<sub>2</sub> generated by CRYSTAL. **h**, **i**, Experimental SAXS profiles of (**h**) 2Co<sup>II</sup>- and 2Ni<sup>II</sup>-(AB)<sub>2</sub> complexes compared with theoretical scattering profile of 1Cu<sup>II</sup>-(AB)<sub>2</sub> (cyan), and (**i**) 1Cu<sup>II</sup>-(AB)<sub>2</sub> complex compared with theoretical scattering profiles of 2Co<sup>II</sup>- (magenta) and 2Ni<sup>II</sup>-(AB)<sub>2</sub> (green). **j**, **k**, Log-scale plots of Fig. [1e–f](#) to compare (**j**) experimental scattering profiles between 2Co<sup>II</sup>/2Ni<sup>II</sup>-(AB)<sub>2</sub> and 1Cu<sup>II</sup>-(AB)<sub>2</sub> and (**k**) experimental scattering profiles of 2Co<sup>II</sup>- (left), 2Ni<sup>II</sup>- (middle), and 1Cu<sup>II</sup>-(AB)<sub>2</sub> (right) with theoretical scattering profiles. **l**, **m**, Expanded low q-ranges of the scattering plots in (**l**) (panel **j**) and (**m**) (**k**).

## Extended Data Fig. 3 Structural characterization and metal-binding analysis of M<sup>II</sup>-(AB)<sub>2</sub> complexes in competitive conditions with Cu<sup>II</sup>.

**a**, Crystal structures (ribbon models) of (AB)<sub>2</sub> formed under Co<sup>II</sup>/Cu<sup>II</sup> (PDB:7LRB), Cu<sup>II</sup>/Co<sup>II</sup> (PDB:7LRR), Ni<sup>II</sup>/Cu<sup>II</sup> (PDB:7LR5), and Cu<sup>II</sup>/Ni<sup>II</sup> (PDB:7LRA) conditions. Root mean square deviation (RMSD) values were determined in comparison with the crystal structures of 2Co<sup>II</sup>-(AB)<sub>2</sub> and 2Ni<sup>II</sup>-(AB)<sub>2</sub> (grey cartoon models). **b**, Experimental SAXS profile of (AB)<sub>2</sub> (0.8 mM) in Co<sup>II</sup>/Cu<sup>II</sup>, Cu<sup>II</sup>/Co<sup>II</sup>, Ni<sup>II</sup>/Cu<sup>II</sup>, and Cu<sup>II</sup>/Ni<sup>II</sup> conditions ([M<sup>II</sup>] = 1.6 mM) compared with the theoretical SAXS profile of 1Cu<sup>II</sup>-(AB)<sub>2</sub>. **c**, ESI-MS spectrum of (AB)<sub>2</sub> (5 μM) without metal ions. **d**, ESI-MS spectrum of (AB)<sub>2</sub> (5 μM) with sub-stoichiometric amounts of Cu<sup>II</sup> (5 μM) and Co<sup>II</sup> or Ni<sup>II</sup> (5 μM). Circles in ESI-MS spectra represent the number of Co<sup>II</sup> (magenta), Ni<sup>II</sup> (green), and Cu<sup>II</sup> (cyan) ions bound to (AB)<sub>2</sub>. The number of circles indicates the equivalents of bound metal ions. **e**, Competitive metal-binding titration of (AB)<sub>2</sub> with Co<sup>II</sup> (magenta), Ni<sup>II</sup> (green), and Cu<sup>II</sup> (cyan) in 20 mM NH<sub>4</sub>HCO<sub>3</sub> (pH 7.8). Mag-Fura-2 (5 μM) was used for competitive Co<sup>II</sup> and Ni<sup>II</sup> titration, and Newport green DCF (5 μM) was used for competitive Cu<sup>II</sup> titration. Experimental data points and error bars are presented as mean and standard deviation of three independent measurements. **f**, **g**, Log-scale plots (left) of Fig. [2c](#) to compare experimental scattering profiles of (**f**) Co<sup>II</sup>/Cu<sup>II</sup> and Cu<sup>II</sup>/Co<sup>II</sup>, and (**g**) Ni<sup>II</sup>/Cu<sup>II</sup> and Cu<sup>II</sup>/Ni<sup>II</sup> with theoretical scattering profiles. Right panels present expanded low q-ranges of the scattering plots (left).

## Extended Data Fig. 4 Peripheral 5His coordination site of 2Co<sup>II</sup>-(AB)<sub>2</sub> and rotamer analysis of the H100 side chain.

**a**, Co<sup>II</sup>-coordination in the dimer interface of (AB)<sub>2</sub> with the 2mF<sub>o</sub>-DF<sub>c</sub> electron density map (grey mesh) contoured at 5.0σ (metal) and 1.5σ (ligand). Coordination distances and angles between Co<sup>II</sup> and ligands are

shown on the right. **b**, Possible rotamer orientations of H100 residue in the crystal structure of  $2\text{Ni}^{\text{II}}\text{-}(\text{AB})_2$ . Rotamers were generated with the combination of  $\chi_1$  (torsion angle of  $\text{C}_\alpha\text{-C}_\beta = 60\text{--}300^\circ$ ,  $60^\circ$  interval) and  $\chi_2$  (torsion angle of  $\text{C}_\beta\text{-C}_\gamma = 0\text{--}330^\circ$ ,  $30^\circ$  interval). **c**, Averaged number of atomic contacts of H100 rotamers as a function of van der waals (VdW) overlap. Error bars reflect the standard deviations of the atomic contacts in rotamers with fixed  $\chi_1$  and variable  $\chi_2$ . Compared to the original conformation of H100, all possible rotamers show significant clashes with proximal residues. **d**, Location of proximal residues interacting with H100 rotamers.

### Extended Data Fig. 5 QM/MM calculations of models 1–3.

**a–c**, QM/MM-optimized structures of **(a)** Model 1:  $\text{Cu}^{\text{II}}$  in the peripheral site of  $(\text{AB})_2$  **(b)** Model 2:  $\text{Cu}^{\text{II}}$  in the peripheral site of  ${}^{\text{H100A}}(\text{AB})_2$ , and **(c)** Model 3:  $\text{Cu}^{\text{II}}$  in the central site of  $(\text{AB})_2$ . Conformations of the opposite peripheral binding site (red) from the modelled  $\text{Cu}^{\text{II}}$  coordination site are shown in the panel on the right in **a** and **b**. **d**, RMSDs of  $\text{C}_\alpha$  positions between QM/MM-optimized models and the crystal structure (His:  $\text{C}_\alpha$ 's of 5His or 4His residues in the opposite peripheral site, All: all  $\text{C}_\alpha$ 's in  $(\text{AB})_2$  or  ${}^{\text{H100A}}(\text{AB})_2$ , and Interface:  $\text{C}_\alpha$ 's of residues in the dimer interface).

### Extended Data Fig. 6 Characterization of metal-bound ${}^{\text{H100A}}(\text{AB})_2$ complexes.

**a**, ESI–MS spectra of  ${}^{\text{H100A}}(\text{AB})_2$  ( $5 \mu\text{M}$ ) complexed with metal ions. Circles in ESI–MS spectra represent the number of  $\text{Co}^{\text{II}}$  (magenta),  $\text{Ni}^{\text{II}}$  (green), and  $\text{Cu}^{\text{II}}$  (cyan) ions bound to  ${}^{\text{H100A}}(\text{AB})_2$ . The number of circles indicates the equivalents of bound metal ions. Metal concentrations were  $10 \mu\text{M}$ . **b**, Fura-2 competitive metal titration assay of  ${}^{\text{H100A}}(\text{AB})_2$  with  $\text{Cu}^{\text{II}}$ . Changes in Fura-2 absorbance at  $335 \text{ nm}$  (cyan) are plotted with theoretical metal-binding isotherms in the absence (grey) and the presence (black) of  $(\text{AB})_2$ . Experimental data points and error bars in the Fura-2 titration are

presented as mean and standard deviation of three independent measurements. **c**, **d**, Crystal structure of (**c**)  $\text{Co}^{\text{II}}\text{-H}^{100\text{A}}(\text{AB})_2$  (PDB:7N4G) and (**d**)  $\text{Ni}^{\text{II}}\text{-H}^{100\text{A}}(\text{AB})_2$  (PDB:7N4F).  $\text{Co}^{\text{II}}$  and  $\text{Ni}^{\text{II}}$  ions are represented as magenta and green spheres. Atomic details of each metal coordination site are shown in the right panels, with the  $2\text{mF}_o\text{-DF}_c$  electron density map (grey mesh) contoured at  $5.0\sigma$  (metal) and  $1.5\sigma$  (ligand).

### **Extended Data Fig. 7 Characterization of periplasmic extracts using SDS-PAGE and ESI-MS.**

**a, b**, SDS-PAGE of periplasmic extracts and medium stained with (**a**) Coomassie blue for all proteinaceous contents and (**b**) o-dianisidine for haem-proteins. Regardless of supplemented metal ions, no significant amount of  $(\text{AB})_2$  dimer and  $(\text{AB})$  monomer was observed in the medium. Uncropped gel images are shown in Supplementary Fig. 2. Reproducibility of SDS-PAGE was tested with two independently extracted sample sets. **c**, ESI-MS spectra of  $(\text{AB})_2$  extracted from cells grown without metal supplement (top) and with  $\text{Co}^{\text{II}}$  (middle) or  $\text{Ni}^{\text{II}}$  (bottom) supplement (20  $\mu\text{M}$ ) in LB medium. **d, e**, ESI-MS spectra of  $(\text{AB})_2$  extracted from cells grown with (**d**) 100  $\mu\text{M}$  and (**e**) 5  $\mu\text{M}$  metal ions in LB medium. Inlet spectra represent expanded  $m/z$  ranges for  $2\text{M}^{\text{II}}\text{-}(\text{AB})_2$  complexes with magenta, green, and cyan lines corresponding to theoretical  $m/z$  values of  $2\text{Co}^{\text{II}}$ ,  $2\text{Ni}^{\text{II}}$ , and  $2\text{Cu}^{\text{II}}\text{-}(\text{AB})_2$  complexes. Supplemented metal ions are represented as magenta ( $\text{Co}^{\text{II}}$ ), green ( $\text{Ni}^{\text{II}}$ ), and cyan ( $\text{Cu}^{\text{II}}$ ) circles. The number of circles indicates the equivalents of metal ions bound to  $(\text{AB})_2$ . D and M in gel pictures refer to the protein dimer and monomer, respectively.

### **Extended Data Table 1 Equilibrium dissociation constants ( $K_d$ ) of $\text{M}^{\text{II}}\text{-}(\text{AB})_2$ complexes**

**Extended Data Table 2 X-ray data collection and refinement statistics for the crystals obtained under non-competitive conditions**  
**Extended Data Table 3 X-ray data collection and refinement statistics for the crystals obtained under competitive conditions with  $\text{Cu}^{\text{II}}$**

# Supplementary information

## Supplementary Information

This file contains the Supplementary Methods, Supplementary Figures 1–2 and Supplementary Tables 1–5.

## Reporting Summary

## Rights and permissions

### Reprints and Permissions

## About this article

### Cite this article

Choi, T.S., Tezcan, F.A. Overcoming universal restrictions on metal selectivity by protein design. *Nature* **603**, 522–527 (2022).  
<https://doi.org/10.1038/s41586-022-04469-8>

- Received: 17 June 2021
- Accepted: 21 January 2022
- Published: 02 March 2022
- Issue Date: 17 March 2022
- DOI: <https://doi.org/10.1038/s41586-022-04469-8>

## Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-022-04469-8>

| [Section menu](#) | [Main menu](#) |

- Article
- [Published: 02 March 2022](#)

# Structure determination of high-energy states in a dynamic protein ensemble

- [John B. Stiller](#) ORCID: [orcid.org/0000-0002-6989-2225](https://orcid.org/0000-0002-6989-2225)<sup>1</sup>,
- [Renee Otten](#) ORCID: [orcid.org/0000-0001-7342-6131](https://orcid.org/0000-0001-7342-6131)<sup>1</sup>,
- [Daniel Häussinger](#) ORCID: [orcid.org/0000-0002-4798-0072](https://orcid.org/0000-0002-4798-0072)<sup>2</sup>,
- [Pascal S. Rieder](#) ORCID: [orcid.org/0000-0001-7110-9175](https://orcid.org/0000-0001-7110-9175)<sup>2</sup>,
- [Douglas L. Theobald](#) ORCID: [orcid.org/0000-0002-2695-8343](https://orcid.org/0000-0002-2695-8343)<sup>3</sup> &
- [Dorothee Kern](#) ORCID: [orcid.org/0000-0002-7631-8328](https://orcid.org/0000-0002-7631-8328)<sup>1</sup>

[Nature](#) volume **603**, pages 528–535 (2022)

- 7148 Accesses
- 71 Altmetric
- [Metrics details](#)

## Subjects

- [Enzyme mechanisms](#)
- [Kinases](#)
- [Solution-state NMR](#)

## Abstract

Macromolecular function frequently requires that proteins change conformation into high-energy states<sup>1,2,3,4</sup>. However, methods for solving the structures of these functionally essential, lowly populated states are lacking. Here we develop a method for high-resolution structure determination of minorly populated states by coupling NMR spectroscopy-derived pseudocontact shifts<sup>5</sup> (PCSs) with Carr–Purcell–Meiboom–Gill (CPMG) relaxation dispersion<sup>6</sup> (PCS–CPMG). Our approach additionally defines the corresponding kinetics and thermodynamics of high-energy excursions, thereby characterizing the entire free-energy landscape. Using a large set of simulated data for adenylate kinase (Adk), calmodulin and Src kinase, we find that high-energy PCSs accurately determine high-energy structures (with a root mean squared deviation of less than 3.5 angström). Applying our methodology to Adk during catalysis, we find that the high-energy excursion involves surprisingly small openings of the AMP and ATP lids. This previously unresolved high-energy structure solves a longstanding controversy about conformational interconversions that are rate-limiting for catalysis. Primed for either substrate binding or product release, the high-energy structure of Adk suggests a two-step mechanism combining conformational selection to this state, followed by an induced-fit step into a fully closed state for catalysis of the phosphoryl-transfer reaction. Unlike other methods for resolving high-energy states, such as cryo-electron microscopy and X-ray crystallography, our solution PCS–CPMG approach excels in cases involving domain rearrangements of smaller systems (less than 60 kDa) and populations as low as 0.5%, and enables the simultaneous determination of protein structure, kinetics and thermodynamics while proteins perform their function.

This is a preview of subscription content

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

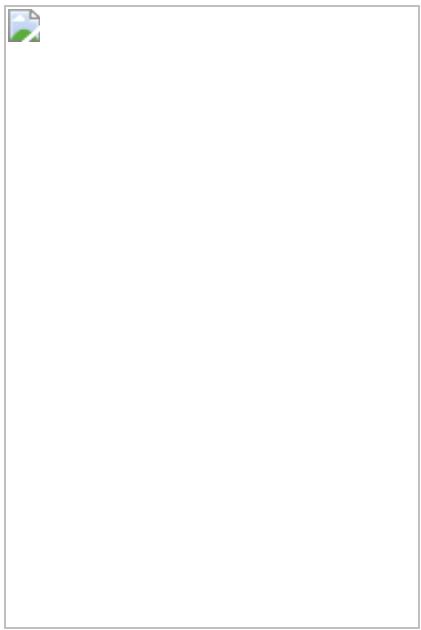
[Buy](#)

All prices are NET prices.

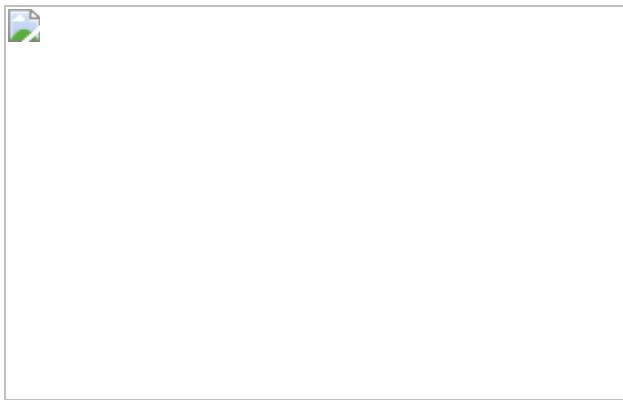
**Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

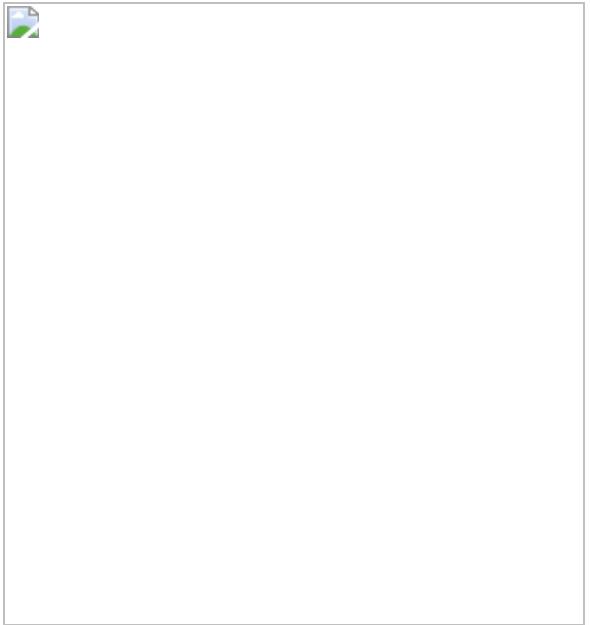
**Fig. 1: Paramagnetic enhanced NMR for structure determination of a high-energy state during catalysis.**



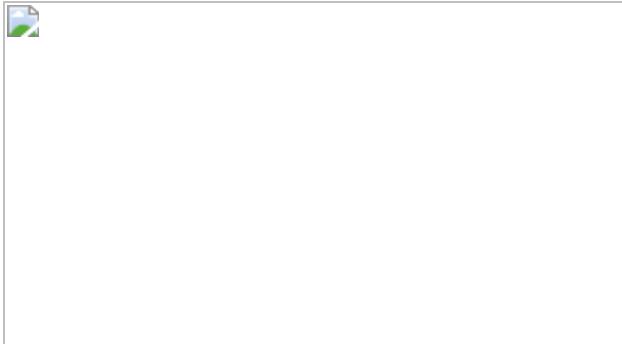
**Fig. 2: Maximum-likelihood classification method for high-energy structure determination.**



**Fig. 3: The high-energy state of Adk during enzyme catalysis.**



**Fig. 4: General applicability of PCS–CPMG methodology.**



## Data availability

The NMR assignments of *G. stearothermophilus* adenylate kinase in its Zn<sup>2+</sup>- and Co<sup>2+</sup>-bound states have been deposited in the BioMagResBank<sup>74</sup> under accession codes [51232](#) and [51233](#), respectively. Peak lists for the tagged ubiquitin variants can be obtained from <https://github.com/kernlab-brandeis/PCS-CPMG>. All relevant data are available from the corresponding author upon request.

## Code availability

Any relevant code is available by request from the corresponding author. A general script for the Expectation Maximization algorithm and test datasets

are made available at <https://github.com/kernlab-brandeis/PCS-CPMG>.

## References

1. Sekhar, A. & Kay, L. E. NMR paves the way for atomic level descriptions of sparsely populated, transiently formed biomolecular conformers. *Proc. Natl Acad. Sci. USA* **110**, 12867–12874 (2013).
2. Orellana, L. Large-scale conformational changes and protein function: breaking the in silico barrier. *Front. Mol. Biosci.* **6**, 117 (2019).
3. Nussinov, R. Introduction to protein ensembles and allostery. *Chem. Rev.* **116**, 6263–6266 (2016).
4. Haliloglu, T. & Bahar, I. Adaptability of protein structures to enable functional interactions and evolutionary implications. *Curr. Opin. Struct. Biol.* **35**, 17–23 (2015).
5. Bertini, I., Luchinat, C. & Parigi, G. Magnetic susceptibility in paramagnetic NMR. *Prog. Nucl. Magn. Reson. Spectrosc.* **40**, 249–273 (2002).
6. Kleckner, I. R. & Foster, M. P. An introduction to NMR-based approaches for measuring protein dynamics. *Biochim. Biophys. Acta* **1814**, 942–968 (2011).
7. Boehr, D. D., McElheny, D., Dyson, H. J. & Wright, P. E. The dynamic energy landscape of dihydrofolate reductase catalysis. *Science* **313**, 1638–1642 (2006).
8. Cianfrocco, M. A. et al. Human TFIID binds to core promoter DNA in a reorganized structural state. *Cell* **152**, 120–131 (2013).
9. Zhao, J., Benlekbir, S. & Rubinstein, J. L. Electron cryomicroscopy observation of rotational states in a eukaryotic V-ATPase. *Nature* **521**, 241–245 (2015).

10. Neudecker, P. et al. Structure of an intermediate state in protein folding and aggregation. *Science* **336**, 362–366 (2012).
11. Dethoff, E. A., Petzold, K., Chugh, J., Casiano-Negróni, A. & Al-Hashimi, H. M. Visualizing transient low-populated structures of RNA. *Nature* **491**, 724–728 (2012).
12. Zhao, B., Guffy, S. L., Williams, B. & Zhang, Q. An excited state underlies gene regulation of a transcriptional riboswitch. *Nat. Chem. Biol.* **13**, 968–974 (2017).
13. Fraser, J. S. et al. Accessing protein conformational ensembles using room-temperature X-ray crystallography. *Proc. Natl Acad. Sci. USA* **108**, 16247–16252 (2011).
14. Bonomi, M. & Vendruscolo, M. Determination of protein structural ensembles using cryo-electron microscopy. *Curr. Opin. Struct. Biol.* **56**, 37–45 (2019).
15. Vogeli, B., Olsson, S., Guntert, P. & Riek, R. The exact NOE as an alternative in ensemble structure determination. *Biophys. J.* **110**, 113–126 (2016).
16. Leung, H. T. et al. A rigorous and efficient method to reweight very large conformational ensembles using average experimental data and to determine their relative information content. *J. Chem. Theory Comput.* **12**, 383–394 (2016).
17. Clore, G. M. & Iwahara, J. Theory, practice, and applications of paramagnetic relaxation enhancement for the characterization of transient low-population states of biological macromolecules and their complexes. *Chem. Rev.* **109**, 4108–4139 (2009).
18. Maltsev, A. S., Grishaev, A., Roche, J., Zasloff, M. & Bax, A. Improved cross validation of a static ubiquitin structure derived from high precision residual dipolar couplings measured in a drug-based liquid crystalline phase. *J. Am. Chem. Soc.* **136**, 3752–3755 (2014).

19. Korzhnev, D. M., Religa, T. L., Banachewicz, W., Fersht, A. R. & Kay, L. E. A transient and low-populated protein-folding intermediate at atomic resolution. *Science* **329**, 1312–1316 (2010).
20. Nerli, S., McShan, A. C. & Sgourakis, N. G. Chemical shift-based methods in NMR structure determination. *Prog. Nucl. Magn. Reson. Spectrosc.* **106-107**, 1–25 (2018).
21. Bertini, I. et al. Experimentally exploring the conformational space sampled by domain reorientation in calmodulin. *Proc. Natl Acad. Sci. USA* **101**, 6841–6846 (2004).
22. Hass, M. A. S. et al. A minor conformation of a lanthanide tag on adenylate kinase characterized by paramagnetic relaxation dispersion NMR spectroscopy. *J. Biomol. NMR* **61**, 123–136 (2015).
23. Xu, D. et al. Ligand proton pseudocontact shifts determined from paramagnetic relaxation dispersion in the limit of NMR intermediate exchange. *J. Phys. Chem. Lett.* **9**, 3361–3367 (2018).
24. Eichmuller, C. & Skrynnikov, N. R. Observation of microsecond time-scale protein dynamics in the presence of  $\text{Ln}^{3+}$  ions: application to the N-terminal domain of cardiac troponin C. *J. Biomol. NMR* **37**, 79–95 (2007).
25. Kerns, S. J. et al. The energy landscape of adenylate kinase during catalysis. *Nat. Struct. Mol. Biol.* **22**, 124–131 (2015).
26. Moon, S., Bannen, R. M., Rutkoski, T. J., Phillips, G. N. Jr & Bae, E. Effectiveness and limitations of local structural entropy optimization in the thermal stabilization of mesophilic and thermophilic adenylate kinases. *Proteins* **82**, 2631–2642 (2014).
27. Hanson, J. A. et al. Illuminating the mechanistic roles of enzyme conformational dynamics. *Proc. Natl Acad. Sci. USA* **104**, 18055–18060 (2007).

28. Aden, J. & Wolf-Watz, M. NMR identification of transient complexes critical to adenylate kinase catalysis. *J. Am. Chem. Soc.* **129**, 14003–14012 (2007).
29. Pelz, B., Zoldak, G., Zeller, F., Zacharias, M. & Rief, M. Subnanometre enzyme mechanics probed by single-molecule force spectroscopy. *Nat. Commun.* **7**, 10848 (2016).
30. Mukhopadhyay, A. et al. Crystal structure of the zinc-, cobalt-, and iron-containing adenylate kinase from *Desulfovibrio gigas*: a novel metal-containing adenylate kinase from Gram-negative bacteria. *J. Biol. Inorg. Chem.* **16**, 51–61 (2011).
31. Carver, J. P. & Richards, R. E. General 2-site solution for chemical exchange produced dependence of T2 upon Carr–Purcell pulse separation. *J. Mag. Res.* **6**, 89–105 (1972).
32. Aviram, H. Y. et al. Direct observation of ultrafast large-scale dynamics of an enzyme under turnover conditions. *Proc. Natl Acad. Sci. USA* **115**, 3243–3248 (2018).
33. Skrynnikov, N. R., Dahlquist, F. W. & Kay, L. E. Reconstructing NMR spectra of “invisible” excited protein states using HSQC and HMQC experiments. *J. Am. Chem. Soc.* **124**, 12352–12360 (2002).
34. Schwieters, C. D., Kuszewski, J. J., Tjandra, N. & Clore, G. M. The Xplor-NIH NMR molecular structure determination package. *J. Mag. Res.* **160**, 65–73 (2003).
35. Fallon, J. L. & Quiocho, F. A. A closed compact structure of native Ca<sup>2+</sup>-calmodulin. *Structure* **11**, 1303–1307 (2003).
36. Cowan-Jacob, S. W. et al. The crystal structure of a c-Src complex in an active conformation suggests possible steps in c-Src activation. *Structure* **13**, 861–871 (2005).
37. Müntener, T., Kottelat, J., Huber, A. & Häussinger, D. New lanthanide chelating tags for PCS NMR spectroscopy with reduction stable, rigid

- linkers for fast and irreversible conjugation to proteins. *Bioconjugate Chem.* **29**, 3344–3351 (2018).
38. Chou, J. J., Li, S., Klee, C. B. & Bax, A. Solution structure of Ca<sup>2+</sup>-calmodulin reveals flexible hand-like properties of its domains. *Nat. Struct. Biol.* **8**, 990–997 (2001).
  39. Russel, D. et al. Putting the pieces together: integrative modeling platform software for structure determination of macromolecular assemblies. *PLoS Biol.* **10**, e1001244 (2012).
  40. Häussinger, D., Huang, J. R. & Grzesiek, S. DOTA-M8: an extremely rigid, high-affinity lanthanide chelating tag for PCS NMR spectroscopy. *J. Am. Chem. Soc.* **131**, 14761–14767 (2009).
  41. Morgado, L., Burmann, B. M., Sharpe, T., Mazur, A. & Hiller, S. The dynamic dimer structure of the chaperone Trigger Factor. *Nat. Commun.* **8**, 1992 (2017).
  42. Kovermann, M., Grundstrom, C., Sauer-Eriksson, A. E., Sauer, U. H. & Wolf-Watz, M. Structural basis for ligand binding to an enzyme by a conformational selection pathway. *Proc. Natl Acad. Sci. USA* **114**, 6298–6303 (2017).
  43. Li, D., Liu, M. S. & Ji, B. Mapping the dynamics landscape of conformational transitions in enzyme: the adenylate kinase case. *Biophys. J.* **109**, 647–660 (2015).
  44. Stiller, J. B. et al. Probing the transition state in enzyme catalysis by high-pressure NMR dynamics. *Nat. Catal.* **2**, 726–734 (2019).
  45. Saio, T. & Ishimori, K. Accelerating structural life science by paramagnetic lanthanide probe methods. *Biochim. Biophys. Acta* **1864**, 129332 (2019).
  46. Nitsche, C. & Otting, G. Pseudocontact shifts in biomolecular NMR using paramagnetic metal tags. *Prog. Nucl. Magn. Reson. Spectrosc.* **98-99**, 20–49 (2017).

47. Ma, R. S. et al. Determination of pseudocontact shifts of low-populated excited states by NMR chemical exchange saturation transfer. *Phys. Chem. Chem. Phys.* **18**, 13794–13798 (2016).
48. Gerstein, M., Lesk, A. M. & Chothia, C. Structural mechanisms for domain movements in proteins. *Biochemistry* **33**, 6739–6749 (1994).
49. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
50. Schmitz, C., Stanton-Cook, M. J., Su, X. C., Otting, G. & Huber, T. Numbat: an interactive software tool for fitting Delta chi-tensors to molecular coordinates using pseudocontact shifts. *J. Biomol. NMR* **41**, 179–189 (2008).
51. Cai, M., Huang, Y., Craigie, R. & Clore, G. M. A simple protocol for expression of isotope-labeled proteins in *Escherichia coli* grown in shaker flasks at high cell density. *J. Biomol. NMR* **73**, 743–748 (2019).
52. Otting, G., Ruckert, M., Levitt, M. H. & Moshref, A. NMR experiments for the sign determination of homonuclear scalar and residual dipolar couplings. *J. Biomol. NMR* **16**, 343–346 (2000).
53. Joss, D., Walliser, R. M., Zimmermann, K. & Häussinger, D. Conformationally locked lanthanide chelating tags for convenient pseudocontact shift protein nuclear magnetic resonance spectroscopy. *J. Biomol. NMR* **72**, 29–38 (2018).
54. Romero, P. R. et al. BioMagResBank (BMRB) as a resource for structural biology. *Methods Mol. Biol.* **2112**, 187–218 (2020).
55. Orton, H. W., Huber, T. & Otting, G. Paramagpy: software for fitting magnetic susceptibility tensors using paramagnetic effects measured in NMR spectra. *Magn. Reson.* **1**, 1–12 (2020).
56. Ishima, R. & Torchia, D. A. Extending the range of amide proton relaxation dispersion experiments in proteins using a constant-time

- relaxation-compensated CPMG approach. *J. Biomol. NMR* **25**, 243–248 (2003).
57. Delaglio, F. et al. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
  58. Vranken, W. F. et al. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696 (2005).
  59. Lee, W., Rahimi, M., Lee, Y. & Chiu, A. POKY: a software suite for multidimensional NMR and 3D structure calculation of biomolecules. *Bioinformatics*. **37**, 3041–3042 (2021).
  60. Niklasson, M. et al. Comprehensive analysis of NMR data using advanced line shape fitting. *J. Biomol. NMR* **69**, 93–99 (2017).
  61. Newville, M., Stensitzki, T., Allen, D. B. & Ingargiola, A. LMFIT: Non-Linear Least-Square Minimization and Curve-Fitting for Python <https://lmfit.github.io/lmfit-py/> (2014).
  62. Counago, R., Chen, S. & Shamu, Y. In vivo molecular evolution reveals biophysical origins of organismal fitness. *Mol. Cell* **22**, 441–449 (2006).
  63. Abele, U. & Schulz, G. E. High-resolution structures of adenylate kinase from yeast ligated with inhibitor Ap5A, showing the pathway of phosphoryl transfer. *Protein Sci.* **4**, 1262–1271 (1995).
  64. Berry, M. B. & Phillips, G. N. Jr. Crystal structures of *Bacillus stearothermophilus* adenylate kinase with bound Ap5A, Mg<sup>2+</sup> Ap5A, and Mn<sup>2+</sup> Ap5A reveal an intermediate lid position and six coordinate octahedral geometry for bound Mg<sup>2+</sup> and Mn<sup>2+</sup>. *Proteins* **32**, 276–288 (1998).
  65. Diederichs, K. & Schulz, G. E. The refined structure of the complex between adenylate kinase from beef heart mitochondrial matrix and its substrate AMP at 1.85 Å resolution. *J. Mol. Biol.* **217**, 541–549 (1991).

66. Schlauderer, G. J., Proba, K. & Schulz, G. E. Structure of a mutant adenylate kinase ligated with an ATP-analogue showing domain closure over ATP. *J. Mol. Biol.* **256**, 223–227 (1996).
67. Henzler-Wildman, K. A. et al. Intrinsic motions along an enzymatic reaction trajectory. *Nature* **450**, 838–844 (2007).
68. Muller, C. W., Schlauderer, G. J., Reinstein, J. & Schulz, G. E. Adenylate kinase motions during catalysis: an energetic counterweight balancing substrate binding. *Structure* **4**, 147–156 (1996).
69. Arnold, K., Bordoli, L., Kopp, J. & Schwede, T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22**, 195–201 (2006).
70. Word, J. M., Lovell, S. C., Richardson, J. S. & Richardson, D. C. Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation. *J. Mol. Biol.* **285**, 1735–1747 (1999).
71. Chattopadhyaya, R., Meador, W. E., Means, A. R. & Quiocho, F. A. Calmodulin structure refined at 1.7 Å resolution. *J. Mol. Biol.* **228**, 1177–1192 (1992).
72. Xu, W., Doshi, A., Lei, M., Eck, M. J. & Harrison, S. C. Crystal structures of c-Src reveal features of its autoinhibitory mechanism. *Mol. Cell* **3**, 629–638 (1999).
73. Bertini, I., Janik, M. B., Lee, Y. M., Luchinat, C. & Rosato, A. Magnetic susceptibility tensor anisotropies for a lanthanide ion series in a fixed protein matrix. *J. Am. Chem. Soc.* **123**, 4181–4188 (2001).
74. Ulrich, E. L. et al. BioMagResBank. *Nucleic Acids Res.* **36**, D402–D408 (2007).
75. Tollinger, M., Skrynnikov, N. R., Mulder, F. A., Forman-Kay, J. D. & Kay, L. E. Slow dynamics in folded and unfolded states of an SH3 domain. *J. Am. Chem. Soc.* **123**, 11341–11352 (2001).

## Acknowledgements

We thank I. Hertel-Hering and M. Rogowski for expression of the K6C and S20C ubiquitin constructs, respectively; and S. Hiller for the trigger factor expression plasmid. This work was supported by the Howard Hughes Medical Institute (HHMI) to D.K., the NIH (R01GM121384 and R01GM132499) to D.L.T., and the Department of Chemistry at University of Basel to D.H. Computational resources were provided by NSF XSEDE computing resources. We acknowledge computational support from the Brandeis HPCC which is partially supported by the NSF through DMR-MRSEC 2011846 and OAC-1920147.

## Author information

### Affiliations

1. Department of Biochemistry and Howard Hughes Medical Institute, Brandeis University, Waltham, MA, USA

John B. Stiller, Renee Otten & Dorothee Kern

2. Department of Chemistry, University of Basel, Basel, Switzerland

Daniel Häussinger & Pascal S. Rieder

3. Department of Biochemistry, Brandeis University, Waltham, MA, USA

Douglas L. Theobald

### Contributions

D.K. conceived the project idea. D.K. and J.B.S. developed the research plan and experimental strategy. J.B.S. purified Adk, performed NMR experiments, and analysed results. D.H. and P.S.R. synthesized lanthanide-binding tags and prepared tagged ubiquitin samples. R.O. purified PPD-SBD proteins and prepared lanthanide-bound samples. R.O. performed the

NMR experiments for ubiquitin and PPD–SBD trigger factor. R.O., J.B.S., D.K., P.S.R. and D.H. analysed results from the NMR experiments. D.L.T. designed the expectation maximization algorithm. J.B.S. implemented the algorithm into the XPLOR-NIH software and performed calculations on simulated and real data. All authors discussed results leading to overall scientific findings. J.B.S. and D.K. wrote the manuscript. All authors reviewed and edited the manuscript.

## Corresponding author

Correspondence to [Dorothee Kern](#).

## Ethics declarations

### Competing interests

The authors declare no competing interests.

### Peer review

### Peer review information

*Nature* thanks Hashim Al-Hashimi and the other, anonymous reviewer(s) for their contribution to the peer review of this work. Peer review reports are available.

## Additional information

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Extended data figures and tables

## Extended Data Fig. 1 Diamagnetic and paramagnetic samples were assigned by triple resonance experiments.

Backbone assignment of Zn<sup>2+</sup> (left) and Co<sup>2+</sup> (right) Adk proteins under saturating nucleotide conditions plotted onto Ap5A bound crystal structure, PDB 4QBH<sup>26</sup>. Orange spheres depict residues where no assignment was determined. Prolines are shown as gray sticks. All other amides are assigned. For the cobalt species, many residues surrounding the metal binding site are lost due to either Curie relaxation or exchange on the intermediate timescale.

## Extended Data Fig. 2 PCSs in different ligand states reflect the conformational state of Adk.

Fits of PCSs extracted from [<sup>1</sup>H-<sup>15</sup>N]-HSQC spectra in either 20 mM Mg<sup>2+</sup>ADP (a), 20 mM ADP (b), and apo (c) to the closed crystal structure (PDB 4QBH<sup>26</sup>). Notably, the Mg<sup>2+</sup>/ADP data fits reasonably well ( $Q = 11.6\%$ ), but reports on PCSs from both the closed crystal structure and the minor state (i.e., population averaged). In the absence of magnesium, the open/closing exchange rate is in the slow time regime, leading to PCSs that better fit the closed state ( $Q = 7.6\%$ ). The *apo* PCSs report on a more open state, and, therefore, fit significantly worse ( $Q = 47.9\%$ ) to the closed state. (d) Calculating tensors using either paramagnetic-induced RDCs (left) or PCSs (right) provide similar tensor values, indicating little ps-ns motion of the paramagnet.

## Extended Data Fig. 3 In the absence of magnesium, the open/closed rate exchange rate is in the slow exchange regime, in agreement with earlier reports<sup>25</sup>.

(a) [<sup>1</sup>H, <sup>15</sup>N]-HSQC spectra for Zn<sup>2+</sup> and Co<sup>2+</sup> samples with either 20 mM ADP or 20 mM Mg<sup>2+</sup>ADP. Residues that were broadened in the Mg<sup>2+</sup>ADP sample show reduced linewidths in the ADP sample, indicating a shift in exchange timescales from intermediate to slow exchange. Black lines trace the PCSs between diamagnetic and paramagnetic samples. For residue 38,

the corresponding diamagnetic peaks are at 8.46 ppm ( $^1\text{H}$ ) and 123.5 ppm ( $^{15}\text{N}$ ) and are not shown for illustration purposes. (b)  $^1\text{H}_\text{N}$  CPMG dispersion profiles for  $\text{Co}^{2+}$  Adk with 20 mM ADP and 0 mM  $\text{Mg}^{2+}$ . Representative traces show slow exchange that is fit to the Tollinger equation<sup>75</sup> ( $F$ -statistics were used to determine whether the description by a slow exchange model compared to a “no-exchange” model was justified at the 95% confidence interval;  $p>0.05$ ). (c) Representative CPMG relaxation dispersion profiles for residues in the presence on 20 mM  $\text{Mg}^{2+}$  ADP. Notably, the paramagnetic chemical shift differences in the absence and presence of magnesium are similar, whereas the timescale is significantly altered ( $k_{\text{open}, \text{ADP}} = 2.6 \pm 0.3 \text{ s}^{-1}$  vs  $k_{\text{open}, \text{ADP}} = 180 \pm 36 \text{ s}^{-1}$ ). Uncertainties (s.d.) in  $R_{2,\text{eff}}$  are determined from the rmsd in the intensities of duplicate points ( $n = 3$ ) according to the definition of pooled relative standard deviation. Uncertainties (s.d.) in chemical shift differences were calculated from the covariance matrix.

#### **Extended Data Fig. 4 Comparison of PCS values determined during *apo* and turnover conditions and correlation of PCSs during turnover conditions with the open structure of Adk.**

(a) Overlay of PCS values obtained for *apo* and  $\text{Mg}^{2+}$ ADP conditions. Values were determined from [ $^1\text{H}$ - $^{15}\text{N}$ ]-HSQC comparison in  $\text{Zn}^{2+}$  and  $\text{Co}^{2+}$  states. Note the sizable loss of PCS in the *apo* state compared to the closed state, indicating a more open structure in the absence of ligand. (b) Difference in PCS values for *apo* and turnover conditions. Large absolute differences of  $> 0.1$  ppm are observed for many residues. (c) Zoom in of [ $^1\text{H}$ - $^{15}\text{N}$ ]-HSQC spectra in either 20 mM  $\text{Mg}^{2+}$ ADP or apo conditions of  $\text{Co}^{2+}$  Adk. Noticeable line broadening is observed for *apo* conditions. (d) Fit of open state (4AKE<sup>68</sup>) to observed PCS shift data during catalytic turnover. (e) Best-fit tensor for PCS to open state structure. (f) Calculated PCSs for open state structure when fit with observed PCSs. A poor fit is found as the observed PCSs do not report on the open state structure. (g) The PCS difference expected between the open and closed state structures. Differences of  $|0.5 \text{ ppm}|$  or greater would be expected for residues in AMP lid and core domain near the ATP lid.

**Extended Data Fig. 5 Expectation-maximization during simulated annealing leads to correct PCS identifications for nearly all residues.**

All 12 structures are shown with incorrectly chosen PCS as gray spheres. Mistakes usually occur near the end of secondary structure units, where local differences between crystal structures are most prevalent. For each structure, the core domain, ATP-lid, and AMP-lid are colored in wheat, red, and blue, respectively. The cobalt metal is shown as a green sphere.

**Extended Data Fig. 6 Starting from an open or closed starting structure of Adk in the PCS-CPMG maximum-likelihood calculations results in equivalent final Adk structures.**

(a) Open (lighter colors) and closed (darker colors) crystal structures, PDBs 4AKE<sup>68</sup> and 4QBH<sup>26</sup>, respectively. (b) RMSD, Co<sup>2+</sup>-Core distance, and AMP-lid angle results for four calculated structures, all started from the open state (PDB 4AKE<sup>68</sup>). Similar to starting from the closed state (Fig. 2f), starting from an open state result in excellent collective variables, and excellent agreement with the experimental structures (shown as black stars). (c) Alignment of inferred structures starting from a closed state (blue) and open state (red) compared to the target structure (gray), highlighting that the new method results in converged and accurate structures independent of the starting model.

**Extended Data Fig. 7 Plots of likelihood versus RMSD to the target structures for the first simulated annealing run for all 12 simulated Adk structures.**

Structures which possess relatively low likelihood also have higher RMSD.

**Extended Data Fig. 8 Eight PCS choices possible in the case where CPMG is performed in the presence of one diamagnetic and two paramagnetic metals.**

As the diamagnetic sign is shared between the two paramagnetic dispersions, choices 1/3 and 2/4 are linked. This reduces the possible PCS choice for a specific residue.

### **Extended Data Fig. 9 Expectation-maximization of ambiguous PCSs are solved during structural calculations for calmodulin and src kinase.**

PCSs of the final structures for calmodulin (a) and src kinase (b) have low  $Q$  values for both metals as well as accurate PCS identifications (blue stars for the correct PCS identification, red spheres are incorrect PCS identifications, dark blue line indicates the calculated PCS).

### **Extended Data Fig. 10 Ubiquitin and the chaperone trigger factor support lanthanide-binding tag coordination and produce substantial paramagnetic tensors.**

(a) [ $^1\text{H}$ - $^{15}\text{N}$ ]-HSQC spectra of ubiquitin mutants S20C (left) and K6C (middle, right) bound with either the DOTA-M7PyThiazole (left, middle) or DOTA-M8-(4R4S)-SSPy(right) lanthanide-binding tags. Each spectrum shows large PCSs induced by  $\text{Tm}^{3+}$  bound tags. (b) Correlation plots between each ubiquitin variant's PCSs and the calculated PCSs. (c) [ $^1\text{H}$ - $^{15}\text{N}$ ]-HSQC spectra of WT PPD-SBD, V270C PPD-SBD, M8-Lu-V270C PPD-SBD, and M8-Tm-V270C PPD-SBD. (d) Zoom-in of center section of the [ $^1\text{H}$ - $^{15}\text{N}$ ]-HSQC spectra of (c) showing nearly identical spectra for each sample. (e) Correlation plots between each M8-Tm-V270C PPD-SBD PCSs and the calculated PCSs.

## **Supplementary information**

### **Supplementary Information**

This file contains the Supplementary Methods, Supplementary Table 1 and Supplementary Figs. 1–5.

## Reporting Summary

## Peer Review File

## Rights and permissions

### Reprints and Permissions

## About this article

### Cite this article

Stiller, J.B., Otten, R., Häussinger, D. *et al.* Structure determination of high-energy states in a dynamic protein ensemble. *Nature* **603**, 528–535 (2022). <https://doi.org/10.1038/s41586-022-04468-9>

- Received: 01 February 2021
- Accepted: 25 January 2022
- Published: 02 March 2022
- Issue Date: 17 March 2022
- DOI: <https://doi.org/10.1038/s41586-022-04468-9>

## Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

# Determining the structure of fleeting protein states

Research Briefing 02 Mar 2022

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-022-04468-9>

| [Section menu](#) | [Main menu](#) |

- Matters Arising
- [Published: 16 March 2022](#)

# The size of the land carbon sink in China

- [Yilong Wang<sup>1</sup>](#),
- [Xuhui Wang](#) [ORCID: orcid.org/0000-0003-0818-9816<sup>2</sup>](#),
- [Kai Wang](#) [ORCID: orcid.org/0000-0001-6269-3517<sup>2</sup>](#),
- [Frédéric Chevallier<sup>3</sup>](#),
- [Dan Zhu](#) [ORCID: orcid.org/0000-0002-5857-1899<sup>2</sup>](#),
- [Jinghui Lian<sup>3,4</sup>](#),
- [Yue He<sup>2</sup>](#),
- [Hanqin Tian](#) [ORCID: orcid.org/0000-0002-1806-4091<sup>5</sup>](#),
- [Junsheng Li<sup>6</sup>](#),
- [Jianxiao Zhu<sup>7</sup>](#),
- [Sujong Jeong](#) [ORCID: orcid.org/0000-0003-4586-4534<sup>8</sup>](#) &
- [Josep G. Canadell](#) [ORCID: orcid.org/0000-0002-8788-3218<sup>9</sup>](#)

[Nature](#) volume 603, pages E7–E9 (2022)

- 2249 Accesses
- 1 Citations
- 26 Altmetric
- [Metrics details](#)

## Subjects

- [Carbon cycle](#)

- [Environmental sciences](#)

[Matters Arising](#) to this article was published on 16 March 2022

The [Original Article](#) was published on 28 October 2020

arising from J. Wang et al. *Nature* <https://doi.org/10.1038/s41586-020-2849-9> (2020)

This is a preview of subscription content

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

[Buy](#)

All prices are NET prices.

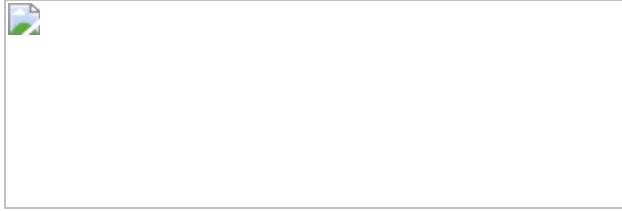
### **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

**Fig. 1: Land–atmosphere CO<sub>2</sub> flux over China.**



**Fig. 2: Representativeness of site Shangri-La in the inversion system.**



## **Data availability**

Atmospheric CO<sub>2</sub> mole fraction data used in the reference, S0 and S1 inversions were collected from the following databases of atmospheric measurements: the National Oceanic and Atmospheric Administration Earth System Research Laboratory archive (Carbon Cycle Greenhouse Gases, <http://www.esrl.noaa.gov/gmd/ccgg/>); the World Data Centre for Greenhouse Gases (<https://gaw.kishou.go.jp/>); the Réseau Atmosphérique de Mesure des Composés à Effet de Serre database (<http://www.lsce.ipsl.fr/>); the Integrated Carbon Observation System–Atmospheric Thematic Center (<https://icos-atc.lsce.ipsl.fr/>); the National Institute for Environmental Studies (<http://db.cger.nies.go.jp>). CO<sub>2</sub> mole fraction data used in the S0 and S1 inversions from the Chinese sites were retrieved from <https://doi.org/10.17632/w3bwmr6rfg.1>. The reference CAMS inversion results are available from <https://ads.atmosphere.copernicus.eu/cdsapp#!/dataset/cams-global-greenhouse-gas-inversion?tab=form>. The results of the high-resolution WRF-Chem simulation for Fig. 2 are available from <https://doi.org/10.6084/m9.figshare.1674667.v1>.

## Code availability

The CAMS inversion system is available on request from F.C. WRF-Chem V3.9.1 is maintained centrally and made available by the National Oceanic and Atmospheric Administration/Earth System Research Laboratories/Global Systems Division (<https://ruc.noaa.gov/wrf/wrf-chem/>).

## References

1. Mallapaty, S. How China could be carbon neutral by mid-century. *Nature* **586**, 482–483 (2020).
2. Wang, J. et al. Large Chinese land carbon sink estimated from atmospheric carbon dioxide data. *Nature* **586**, 720–723 (2020).
3. Jiang, F. et al. A comprehensive estimate of recent carbon sinks in China using both top-down and bottom-up approaches. *Sci. Rep.* **6**,

22130 (2016).

4. Piao, S. et al. The carbon balance of terrestrial ecosystems in China. *Nature* **458**, 1009–1013 (2009).
5. Fang, J., Yu, G., Liu, L., Hu, S. & Chapin, F. S. Climate change, human impacts, and carbon sequestration in China. *Proc. Natl Acad. Sci. USA* **115**, 4015–4020 (2018).
6. Ji, Y. et al. Variation of net primary productivity and its drivers in China’s forests during 2000–2018. *For. Ecosyst.* **7**, 15 (2020).
7. Chapin, F. S., Matson, P. A. & Vitousek, P. M. in *Principles of Terrestrial Ecosystem Ecology* (eds Chapin, F. S. et al.) 183–228 (Springer, 2011).
8. Prentice, I. C. et al. In *Climate Change 2001: The Scientific Basis* (eds Houghton, J. T. et al.) 183–237 (IPCC, Cambridge Univ. Press, 2001).
9. Pan, Y. et al. A large and persistent carbon sink in the world’s forests. *Science* **333**, 988–993 (2011).
10. Tian, H. et al. China’s terrestrial carbon balance: contributions from multiple global change factors. *Global Biogeochem. Cycles* **25**, GB1007 (2011).
11. Ciais, P. et al. Empirical estimates of regional carbon budgets imply reduced global soil heterotrophic respiration. *Natl Sci. Rev.* **8**, nwaa145 (2020).
12. Agustí-Panareda, A. et al. Modelling CO<sub>2</sub> weather – why horizontal resolution matters. *Atmos. Chem. Phys.* **19**, 7347–7376 (2019).
13. Wang, Y. et al. Estimation of observation errors for large-scale atmospheric inversion of CO<sub>2</sub> emissions from fossil fuel combustion. *Tellus B Chem. Phys. Meteorol.* **69**, 1325723 (2017).

14. Chevallier, F. *Validation Report for the Inverted CO<sub>2</sub> Fluxes, v18r1–version 1.0*, [https://atmosphere.copernicus.eu/sites/default/files/2019-01/CAMS73\\_2018SC1\\_D73.1.4.1-2017-v0\\_201812\\_v1\\_final.pdf](https://atmosphere.copernicus.eu/sites/default/files/2019-01/CAMS73_2018SC1_D73.1.4.1-2017-v0_201812_v1_final.pdf) (European Centre for Medium-Range Weather Forecasts (ECMWF), 2018).
15. Canadell, J. et al. (eds) Regional Carbon Cycle Assessment and Processes (RECCAP) *Biogeosciences* [https://bg.copernicus.org/articles/special\\_issue107.html](https://bg.copernicus.org/articles/special_issue107.html) (2012–2015).

## Acknowledgements

This study is supported by the Second Tibetan Plateau Scientific Expedition and Research Program (2019QZKK0405), the National Natural Science Foundation of China (Basic Science Center for Tibetan Plateau Earth System grant number 41988101 and grant number 42171096, 42001104 and 42101090) and the Regional Carbon Cycle Assessment and Processes-2 of the Global Carbon Project. F.C. is financially supported by the Copernicus Atmosphere Monitoring Service, implemented by the European Centre for Medium-Range Weather Forecasts on behalf of the European Commission. H.T. is financially supported by the US National Science Foundation (grant number 1903722). S.J. is supported by the R&D Program for Oceans and Polar Regions of the National Research Foundation (NRF) funded by the Ministry of Science and ICT(NRF-2020M1A5A1110493). We thank J. Zhu for discussions on carbon accumulation in domestic wood products.

## Author information

### Affiliations

1. Key Laboratory of Land Surface Pattern and Simulation, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China

Yilong Wang

2. Sino-French Institute for Earth System Science, College of Urban and Environmental Sciences, Peking University, Beijing, China

Xuhui Wang, Kai Wang, Dan Zhu & Yue He

3. Laboratoire des Sciences du Climat et de l'Environnement, LSCE-IPSL (CEA-CNRS-UVSQ), Université Paris-Saclay, Gif-sur-Yvette, France

Frédéric Chevallier & Jinghui Lian

4. Suez Group, Tour CB21, Paris, France

Jinghui Lian

5. International Center for Climate and Global Change Research, School of Forestry and Wildlife Sciences, Auburn University, Auburn, AL, USA

Hanqin Tian

6. Biodiversity Research Center, Chinese Research Academy of Environmental Sciences, Beijing, China

Junsheng Li

7. State Key Laboratory of Grassland Agro-ecosystems, College of Pastoral Agriculture Science and Technology, Lanzhou University, Lanzhou, China

Jianxiao Zhu

8. Department of Environmental Planning, Graduate School of Environmental Studies, Seoul National University, Seoul, South Korea

Sujong Jeong

9. Global Carbon Project, CSIRO Oceans and Atmosphere, Canberra, Australian Capital Territory, Australia

Josep G. Canadell

## Contributions

Y.W. and X.W. designed the study. Y.W., X.W. and D.Z. coordinated the author team. Y.W., X.W., K.W., F.C. and J. Lian performed the analysis. Y.W., X.W. and D.Z. led the writing of the manuscript with contributions from all authors.

## Corresponding authors

Correspondence to [Xuhui Wang](#) or [Dan Zhu](#).

## Ethics declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Supplementary information

### [Supplementary Information](#)

This file contains Supplementary Text 1–3, Figs. 1 and 2, Supplementary Table 1 and References.

## Rights and permissions

### [Reprints and Permissions](#)

## About this article

### Cite this article

Wang, Y., Wang, X., Wang, K. *et al.* The size of the land carbon sink in China. *Nature* **603**, E7–E9 (2022). <https://doi.org/10.1038/s41586-021-04255-y>

- Received: 10 February 2021
- Accepted: 15 November 2021
- Published: 16 March 2022
- Issue Date: 17 March 2022
- DOI: <https://doi.org/10.1038/s41586-021-04255-y>

### Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

### Further reading

- [\*\*Reply to: The size of the land carbon sink in China\*\*](#)
  - Jing Wang
  - Liang Feng
  - ChaoZong Xia

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04255-y>.

| [Section menu](#) | [Main menu](#) |

- Matters Arising
- [Published: 16 March 2022](#)

# Reply to: The size of the land carbon sink in China

- [Jing Wang](#) ORCID: orcid.org/0000-0002-7149-5157<sup>1,2,3,4</sup>,
- [Liang Feng](#)<sup>3,4</sup>,
- [Paul I. Palmer](#) ORCID: orcid.org/0000-0002-1487-0969<sup>3,4</sup>,
- [Yi Liu](#) ORCID: orcid.org/0000-0001-9305-5358<sup>1,2</sup>,
- [Shuangxi Fang](#) ORCID: orcid.org/0000-0002-2104-5721<sup>5,6</sup>,
- [Hartmut Bösch](#)<sup>7</sup>,
- [Christopher W. O'Dell](#)<sup>8</sup>,
- [Xiaoping Tang](#)<sup>9</sup>,
- [Dongxu Yang](#)<sup>1,2</sup>,
- [Lixin Liu](#)<sup>6</sup> &
- [ChaoZong Xia](#)<sup>9</sup>

[Nature](#) volume 603, pages E10–E12 (2022)

- 914 Accesses
- 1 Altmetric
- [Metrics details](#)

## Subjects

- [Atmospheric science](#)
- [Carbon cycle](#)

The [Original Article](#) was published on 16 March 2022

replying to Y. Wang et al. *Nature* <https://doi.org/10.1038/s41586-021-04255-y> (2022)

This is a preview of subscription content

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

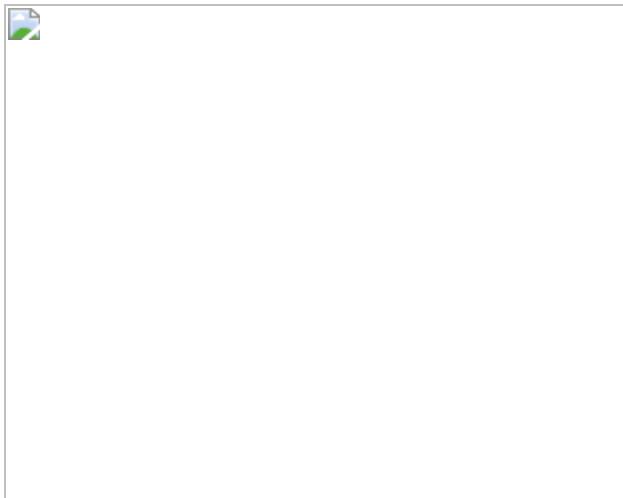
[Buy](#)

All prices are NET prices.

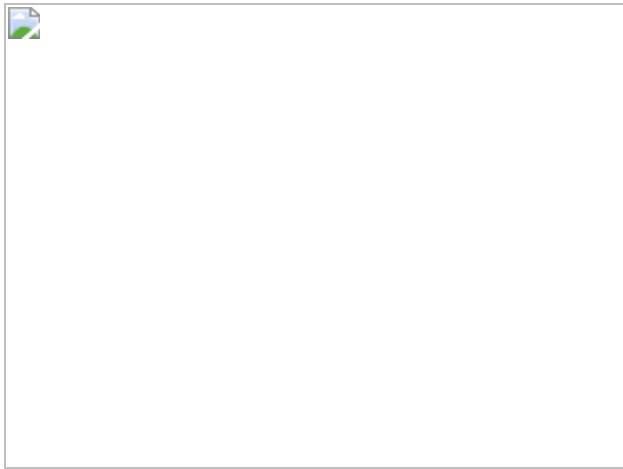
### **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

**Fig. 1: Atmospheric CO<sub>2</sub> mole fractions.**



**Fig. 2: Terrestrial biosphere CO<sub>2</sub> fluxes.**



### **Data availability**

CO<sub>2</sub> mole fraction data from the Chinese sites used in this study are available at <https://doi.org/10.17632/w3bwmr6rfg.1> on <http://data.mendeley.com>.

## Code availability

We used Python Language Reference, version 3.7.7 (Python Software Foundation), available at <http://www.python.org>. We also used Matplotlib (v3.1.3, <https://doi.org/10.5281/zenodo.3984190>). The community-led GEOS-Chem model of atmospheric chemistry and transport is maintained centrally by Harvard University (<http://wiki.seas.harvard.edu/geos-chem>) and is available on request. The ensemble Kalman filter code is publicly available as PyOSSE (<https://www.nceo.ac.uk/data-tools/atmospheric-tools/>).

## References

1. Wang, J. et al. Large Chinese land carbon sink estimated from atmospheric carbon dioxide data. *Nature* **86**, 720–723 (2020).
2. Wang, Y. et al. The size of the land carbon sink in China. *Nature*, <https://doi.org/10.1038/s41586-021-04255-y> (2022).
3. Prentice, I. C. et al. in *Climate Change 2001: The Scientific Basis* (eds Houghton, J. T. et al.) Ch. 3 (IPCC, Cambridge Univ. Press, 2001).
4. Yue, X., Zhang, T. & Shao, C. Afforestation increases ecosystem productivity and carbon storage in China during the 2000s. *Agric. For. Meteorol.* <https://doi.org/10.1016/j.agrformet.2020.108227> (2021).
5. Chapin, F. S. III et al. Reconciling carbon-cycle concepts, terminology, and methods. *Ecosystems* <https://doi.org/10.1007/s10021-005-0105-7> (2006).
6. Noormets, A. et al. Effects of forest management on productivity and carbon sequestration: a review and hypothesis. *For. Ecol. Manag.* **355**, 124–140 (2015).

7. Feng, L., Palmer, P. I., Bösch, H. & Dance, S. Estimating surface CO<sub>2</sub> fluxes from space-borne CO<sub>2</sub> dry air mole fraction observations using an ensemble Kalman Filter. *Atmos. Chem. Phys.* **9**, 2619–2633 (2009).

## Author information

### Affiliations

1. Carbon Neutrality Research Center, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China  
Jing Wang, Yi Liu & Dongxu Yang
2. University of Chinese Academy of Sciences, Beijing, China  
Jing Wang, Yi Liu & Dongxu Yang
3. National Centre for Earth Observation, University of Edinburgh, Edinburgh, UK  
Jing Wang, Liang Feng & Paul I. Palmer
4. School of GeoSciences, University of Edinburgh, Edinburgh, UK  
Jing Wang, Liang Feng & Paul I. Palmer
5. College of Environment, Zhejiang University of Technology, Hangzhou, China  
Shuangxi Fang
6. Meteorological Observation Centre, China Meteorological Administration, Beijing, China  
Shuangxi Fang & Lixin Liu
7. National Centre for Earth Observation, University of Leicester, Leicester, UK

Hartmut Bösch

8. Cooperative Institute for Research in the Atmosphere (CIRA),  
Colorado State University, Fort Collins, CO, USA

Christopher W. O'Dell

9. Academy of Forest Inventory and Planning, State Forestry  
Administration, Beijing, China

Xiaoping Tang & ChaoZong Xia

## Contributions

P.I.P. and L.F. prepared this response. P.I.P. led the writing of the paper, with contributions from L.F., J.W., Y.L., S.F., H.B., C.W.O'D., X.T., D.Y., L.L. and C.X.

## Corresponding authors

Correspondence to [Paul I. Palmer](#), [Yi Liu](#) or [Shuangxi Fang](#).

## Ethics declarations

## Competing interests

The authors declare no competing interests.

## Additional information

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Supplementary information

### [Supplementary Information](#)

This file contains supplementary text regarding estimating the NEP/NPP ratio.

## Rights and permissions

### [Reprints and Permissions](#)

## About this article

### Cite this article

Wang, J., Feng, L., Palmer, P.I. *et al.* Reply to: The size of the land carbon sink in China. *Nature* **603**, E10–E12 (2022).  
<https://doi.org/10.1038/s41586-021-04256-x>

- Published: 16 March 2022
- Issue Date: 17 March 2022
- DOI: <https://doi.org/10.1038/s41586-021-04256-x>

## Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

## [Large Chinese land carbon sink estimated from atmospheric carbon dioxide data](#)

- Jing Wang

- Liang Feng
- ChaoZong Xia

Article 28 Oct 2020

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04256-x>

| [Section menu](#) | [Main menu](#) |

- Matters Arising
- [Published: 16 March 2022](#)

# On the role of atmospheric model transport uncertainty in estimating the Chinese land carbon sink

- [Andrew E. Schuh](#) ORCID: [orcid.org/0000-0002-0445-7708](https://orcid.org/0000-0002-0445-7708)<sup>1</sup>,
- [Brendan Byrne](#) ORCID: [orcid.org/0000-0003-0619-3045](https://orcid.org/0000-0003-0619-3045)<sup>2</sup>,
- [Andrew R. Jacobson](#) ORCID: [orcid.org/0000-0002-8501-3382](https://orcid.org/0000-0002-8501-3382)<sup>3,4</sup>,
- [Sean M. R. Crowell](#) ORCID: [orcid.org/0000-0001-8353-3707](https://orcid.org/0000-0001-8353-3707)<sup>5</sup>,
- [Feng Deng](#)<sup>6</sup>,
- [David F. Baker](#)<sup>1</sup>,
- [Matthew S. Johnson](#)<sup>7</sup>,
- [Sajeev Philip](#)<sup>8</sup> &
- [Brad Weir](#) ORCID: [orcid.org/0000-0001-6160-0577](https://orcid.org/0000-0001-6160-0577)<sup>9,10</sup>

[Nature](#) volume **603**, pages E13–E14 (2022)

- 819 Accesses
- 3 Altmetric
- [Metrics details](#)

## Subjects

- [Atmospheric science](#)
- [Biogeochemistry](#)
- [Climate change](#)

[Matters Arising](#) to this article was published on 16 March 2022

The [Original Article](#) was published on 28 October 2020

arising from J. Wang et al. *Nature* <https://doi.org/10.1038/s41586-020-2849-9> (2020).

This is a preview of subscription content

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

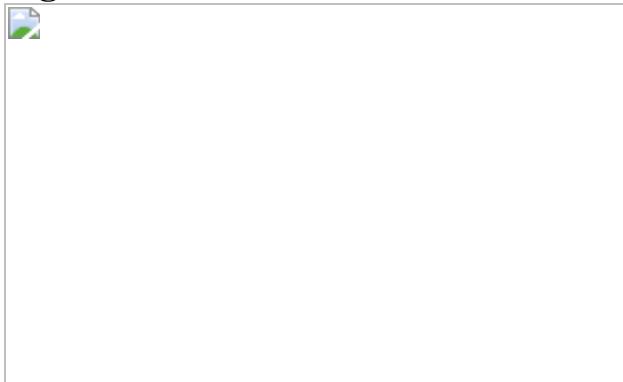
[Buy](#)

All prices are NET prices.

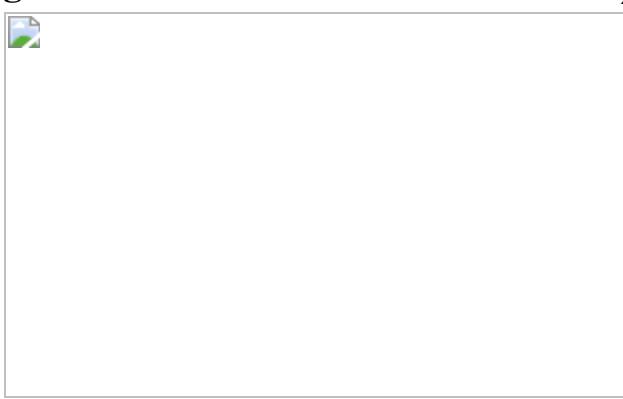
### **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

**Fig. 1: Land biosphere flux estimates over China from inverse models using GEOS-Chem and TM5 for atmospheric transport.**



**Fig. 2: GEOS-Chem minus TM5 XCO<sub>2</sub>.**



## **References**

1. Wang, J. et al. Large Chinese land carbon sink estimated from atmospheric carbon dioxide data. *Nature* **586**, 720–723 (2020).
2. Thompson, R. L. et al. Top-down assessment of the Asian carbon budget since the mid 1990s. *Nat. Commun.* **7**, 10724 (2016).
3. Jiang, F. et al. A comprehensive estimate of recent carbon sinks in China using both top-down and bottom-up approaches. *Sci. Rep.* **6**, 22130 (2016).
4. Zhang, H. F. et al. Net terrestrial CO<sub>2</sub> exchange over China during 2001–2010 estimated with an ensemble data assimilation system for atmospheric CO<sub>2</sub>. *J. Geophys. Res. Atmos.* **119**, 3500–3515 (2014).
5. Bey, I. et al. Global modeling of tropospheric chemistry with assimilated meteorology: model description and evaluation, *J. Geophys. Res.*, **106**, 23073–23096 (2001).
6. Krol, M. et al. The two-way nested global chemistry-transport zoom model TM5: algorithm and applications. *Atmospheric Chem. Phys.* **5**, 417–432 (2005).
7. Peiro, H. et al. Four years of global carbon cycle observed from OCO-2 version 9 and in situ data, and comparison to OCO-2 v7. *Atmos. Chem. Phys. Discuss.* <https://doi.org/10.5194/acp-2021-373> (2021).
8. Schuh, A. E. et al. Quantifying the impact of atmospheric transport uncertainty on CO<sub>2</sub> surface flux estimates. *Glob. Biogeochem. Cycles* **33**, 484–500 (2019).
9. Stanevich, I. et al. Characterizing model errors in chemical transport modeling of methane: impact of model resolution in versions v9-02 of GEOS-Chem and v35j of its adjoint model. *Geosci. Model Dev.* **13**, 3839–3862 (2020).
10. Stephens, B. B. et al. Weak northern and strong tropical land carbon uptake from vertical profiles of atmospheric CO<sub>2</sub>. *Science* **316**, 1732–1735 (2007).

11. Philip, S. et al. Prior biosphere model impact on global terrestrial CO<sub>2</sub> fluxes estimated from OCO-2 retrievals. *Atmos. Chem. Phys.* **19**, 13267–13287 (2019).
12. Byrne, B. et al. Improved constraints on northern extratropical CO<sub>2</sub> fluxes obtained by combining surface-based and space-based atmospheric CO<sub>2</sub> measurements. *J. Geophys. Res. Atmos.* **125**, e2019JD032029 (2020).
13. Liu, J. et al. Comparison between the Local Ensemble Transform Kalman Filter (LETKF) and 4D-Var in atmospheric CO<sub>2</sub> flux inversion with the Goddard Earth Observing System-Chem model and the observation impact diagnostics from the LETKF. *J. Geophys. Res. Atmos.* **121**, 13066–13087 (2017).
14. Peylin, P. et al. Global atmospheric carbon budget: results from an ensemble of atmospheric CO<sub>2</sub> inversions. *Biogeosciences* **10**, 6699–6720 (2013).
15. Friedlingstein, P. et al. Global carbon budget 2019. *Earth Syst. Sci. Data* **11**, 1783–1838 (2019).
16. Canadell, J. G. et al. An international effort to quantify regional carbon fluxes. *Eos Trans. Am. Geophys. Union* **92**, 81–82 (2011).
17. Crowell, S. et al. The 2015–2016 carbon cycle as seen from OCO-2 and the global in situ network. *Atmos. Chem. Phys.* **19**, 9797–9831 (2019).

## Acknowledgements

A.E.S.’s research was performed at Colorado State University under a grant from the National Aeronautics and Space Administration (80NSSCK1312). A.R.J.’s research was carried out at the National Oceanic and Atmospheric Administration Global Monitoring Laboratory and the University of Colorado under grant 80NSSC18K1115 from the National Aeronautics and

Space Administration. B.B.'s research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004).

## Author information

### Affiliations

1. CIRA, Colorado State University, Fort Collins, CO, USA

Andrew E. Schuh & David F. Baker

2. Jet Propulsion Laboratory, California Institute of Technology,  
Pasadena, CA, USA

Brendan Byrne

3. CIRES, University of Colorado, Boulder, CO, USA

Andrew R. Jacobson

4. NOAA Global Monitoring Laboratory, Boulder, CO, USA

Andrew R. Jacobson

5. Department of Meteorology, University of Oklahoma, Norman, OK,  
USA

Sean M. R. Crowell

6. Department of Physics, University of Toronto, Toronto, Ontario,  
Canada

Feng Deng

7. Earth Science Division, NASA Ames Research Center, Moffett Field,  
CA, USA

Matthew S. Johnson

8. NASA Academic Mission Services by Universities Space Research Association at NASA Ames Research Center, Mountain View, CA, USA

Sajeev Philip

9. Universities Space Research Association, Columbia, MD, USA

Brad Weir

10. NASA Goddard Space Flight Center, Greenbelt, MD, USA

Brad Weir

## **Contributions**

A.E.S., A.R.J. and B.B. were the primary authors of this response with comments and feedback from the remaining co-authors, who were participants of the OCO-2 v9 MIP exercise. All authors contributed to the final version.

## **Corresponding author**

Correspondence to [Andrew E. Schuh](#).

## **Ethics declarations**

## **Competing interests**

The authors declare no competing interests.

## **Additional information**

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Rights and permissions

[Reprints and Permissions](#)

## About this article

### Cite this article

Schuh, A.E., Byrne, B., Jacobson, A.R. *et al.* On the role of atmospheric model transport uncertainty in estimating the Chinese land carbon sink. *Nature* **603**, E13–E14 (2022). <https://doi.org/10.1038/s41586-021-04258-9>

- Received: 14 December 2020
- Accepted: 12 November 2021
- Published: 16 March 2022
- Issue Date: 17 March 2022
- DOI: <https://doi.org/10.1038/s41586-021-04258-9>

### Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04258-9>

| [Section menu](#) | [Main menu](#) |

- Matters Arising
- [Published: 16 March 2022](#)

# Reply to: On the role of atmospheric model transport uncertainty in estimating the Chinese land carbon sink

- [Jing Wang](#) ORCID: [orcid.org/0000-0002-7149-5157](https://orcid.org/0000-0002-7149-5157)<sup>1,2,3,4</sup>,
- [Liang Feng](#)<sup>3,4</sup>,
- [Paul I. Palmer](#) ORCID: [orcid.org/0000-0002-1487-0969](https://orcid.org/0000-0002-1487-0969)<sup>3,4</sup>,
- [Yi Liu](#) ORCID: [orcid.org/0000-0001-9305-5358](https://orcid.org/0000-0001-9305-5358)<sup>1,2</sup>,
- [Shuangxi Fang](#) ORCID: [orcid.org/0000-0002-2104-5721](https://orcid.org/0000-0002-2104-5721)<sup>5,6</sup>,
- [Hartmut Bösch](#)<sup>7</sup>,
- [Christopher W. O'Dell](#)<sup>8</sup>,
- [Xiaoping Tang](#)<sup>9</sup>,
- [Dongxu Yang](#)<sup>1,2</sup>,
- [Lixin Liu](#)<sup>6</sup> &
- [ChaoZong Xia](#)<sup>9</sup>

[Nature](#) volume 603, pages E15–E16 (2022)

- 433 Accesses
- 1 Altmetric
- [Metrics details](#)

## Subjects

- [Atmospheric chemistry](#)
- [Carbon cycle](#)

The [Original Article](#) was published on 16 March 2022

replying to A. E. Schuh et al. *Nature* <https://doi.org/10.1038/s41586-021-04258-9> (2022).

This is a preview of subscription content

## Access options

Subscribe to Nature+

Get immediate online access to the entire Nature family of 50+ journals

\$29.99

monthly

[Subscribe](#)

Subscribe to Journal

Get full journal access for 1 year

\$199.00

only \$3.90 per issue

[Subscribe](#)

All prices are NET prices.

VAT will be added later in the checkout.

Tax calculation will be finalised during checkout.

Buy article

Get time limited or full article access on ReadCube.

\$32.00

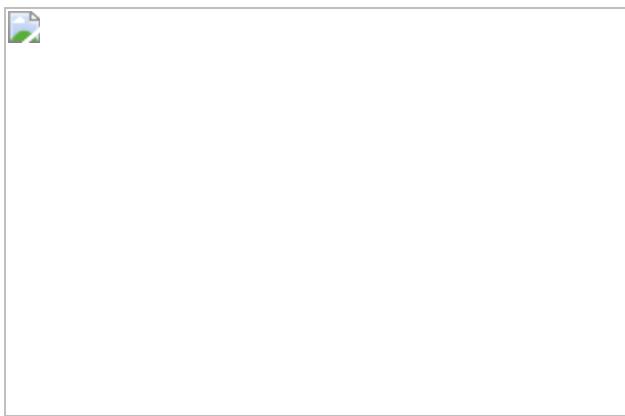
[Buy](#)

All prices are NET prices.

### **Additional access options:**

- [Log in](#)
- [Learn about institutional subscriptions](#)

**Fig. 1: Monthly land CO<sub>2</sub> fluxes for the whole of China during 2015–2018.**



## **Data availability**

The CO<sub>2</sub> mole fraction data from the Chinese sites used in this study are available at <https://doi.org/10.17632/w3bwmr6rfg.1> on <http://data.mendeley.com>.

## **Code availability**

We acknowledge the Python Software Foundation: Python Language Reference, version 3.7.7; available at <http://www.python.org>. We also acknowledge Matplotlib (v3.1.3, <https://doi.org/10.5281/zenodo.3984190>).

The community-led GEOS-Chem model of atmospheric chemistry and transport is maintained centrally by Harvard University (<http://acmg.seas.harvard.edu/geos/>) and is available on request. The ensemble Kalman filter code is publicly available as PyOSSE (<https://www.nceo.ac.uk/data-tools/atmospheric-tools/>).

## References

1. Wang, J. et al. Large Chinese land carbon sink estimated from atmospheric carbon dioxide data. *Nature* **586**, 720–723 (2020).
2. Schuh, A. E. et al. On the role of atmospheric model transport uncertainty in estimating the Chinese land carbon sink. *Nature* <https://doi.org/10.1038/s41586-021-04258-9> (2022).
3. Schuh, A. E. et al. Quantifying the impact of atmospheric transport uncertainty on CO<sub>2</sub> surface flux estimates. *Glob. Biogeochem. Cycles* **33**, 484–500 (2019).

## Acknowledgements

We thank the OCO-2 Model Intercomparison Project for sharing their individual model data to help us to prepare this response.

## Author information

### Affiliations

1. Carbon Neutrality Research Center, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

Jing Wang, Yi Liu & Dongxu Yang

2. University of Chinese Academy of Sciences, Beijing, China

Jing Wang, Yi Liu & Dongxu Yang

3. National Centre for Earth Observation, University of Edinburgh,  
Edinburgh, UK

Jing Wang, Liang Feng & Paul I. Palmer

4. School of GeoSciences, University of Edinburgh, Edinburgh, UK

Jing Wang, Liang Feng & Paul I. Palmer

5. College of Environment, Zhejiang University of Technology,  
Hangzhou, China

Shuangxi Fang

6. Meteorological Observation Centre, China Meteorological  
Administration, Beijing, China

Shuangxi Fang & Lixin Liu

7. National Centre for Earth Observation, University of Leicester,  
Leicester, UK

Hartmut Bösch

8. Cooperative Institute for Research in the Atmosphere (CIRA),  
Colorado State University, Fort Collins, CO, USA

Christopher W. O'Dell

9. Academy of Forest Inventory and Planning, State Forestry  
Administration, Beijing, China

Xiaoping Tang & ChaoZong Xia

## Contributions

P.I.P. and L.F. prepared this response. P.I.P. led the writing of the paper, with contributions from L.F., Y.L., S.F., H.B., C.W.O., X.T., D.Y., L.L. and C.X.

## Corresponding authors

Correspondence to [Paul I. Palmer](#), [Yi Liu](#) or [Shuangxi Fang](#).

## Ethics declarations

### Competing interests

The authors declare no competing interests.

## Additional information

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Rights and permissions

[Reprints and Permissions](#)

## About this article

### Cite this article

Wang, J., Feng, L., Palmer, P.I. *et al.* Reply to: On the role of atmospheric model transport uncertainty in estimating the Chinese land carbon sink. *Nature* **603**, E15–E16 (2022). <https://doi.org/10.1038/s41586-021-04259-8>

- Published: 16 March 2022
- Issue Date: 17 March 2022
- DOI: <https://doi.org/10.1038/s41586-021-04259-8>

### Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

## [Large Chinese land carbon sink estimated from atmospheric carbon dioxide data](#)

- Jing Wang
- Liang Feng
- ChaoZong Xia

Article 28 Oct 2020

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04259-8>

| [Section menu](#) | [Main menu](#) |

# Amendments & Corrections

- **[Author Correction: Inherent mosaicism and extensive mutation of human placentas](#)** [ 28 February 2022]

Author Correction •

- Author Correction
- [Published: 28 February 2022](#)

# Author Correction: Inherent mosaicism and extensive mutation of human placentas

- [Tim H. H. Coorens](#) ORCID: [orcid.org/0000-0002-5826-3554](https://orcid.org/0000-0002-5826-3554)<sup>1 na1</sup>,
- [Thomas R. W. Oliver](#) ORCID: [orcid.org/0000-0003-4306-0102](https://orcid.org/0000-0003-4306-0102)<sup>1,2 na1</sup>,
- [Rashesh Sanghvi](#) ORCID: [orcid.org/0000-0002-7703-9216](https://orcid.org/0000-0002-7703-9216)<sup>1</sup>,
- [Ulla Sovio](#) ORCID: [orcid.org/0000-0002-0799-1105](https://orcid.org/0000-0002-0799-1105)<sup>3</sup>,
- [Emma Cook](#) ORCID: [orcid.org/0000-0002-6142-4443](https://orcid.org/0000-0002-6142-4443)<sup>3</sup>,
- [Roser Vento-Tormo](#) ORCID: [orcid.org/0000-0002-9870-8474](https://orcid.org/0000-0002-9870-8474)<sup>1</sup>,
- [Muzlifah Haniffa](#) ORCID: [orcid.org/0000-0002-3927-2084](https://orcid.org/0000-0002-3927-2084)<sup>1,4,5</sup>,
- [Matthew D. Young](#)<sup>1</sup>,
- [Raheleh Rahbari](#) ORCID: [orcid.org/0000-0002-1839-7785](https://orcid.org/0000-0002-1839-7785)<sup>1</sup>,
- [Neil Sebire](#)<sup>6,7</sup>,
- [Peter J. Campbell](#) ORCID: [orcid.org/0000-0002-3921-0510](https://orcid.org/0000-0002-3921-0510)<sup>1</sup>,
- [D. Stephen Charnock-Jones](#) ORCID: [orcid.org/0000-0002-2936-4890](https://orcid.org/0000-0002-2936-4890)<sup>3,8 na2</sup>,
- [Gordon C. S. Smith](#) ORCID: [orcid.org/0000-0003-2124-0997](https://orcid.org/0000-0003-2124-0997)<sup>2,3,8 na2</sup> &
- [Sam Behjati](#) ORCID: [orcid.org/0000-0002-6600-7665](https://orcid.org/0000-0002-6600-7665)<sup>1,2,9 na2</sup>

[Nature](#) volume 603, page E17 (2022)

- 604 Accesses
- 2 Altmetric
- [Metrics details](#)

## Subjects

- [Cell lineage](#)
- [Development](#)
- [Genetics research](#)
- [Genomics](#)

- [Reproductive biology](#)

The [Original Article](#) was published on 10 March 2021

[Download PDF](#)

Correction to: *Nature* <https://doi.org/10.1038/s41586-021-03345-1> Published online 10 March 2021

In the version of this article initially published, there was a plotting error affecting Fig. 1d,e and Extended Data Fig. 1b. SBS signature extraction was performed for the bulk placental samples on a per-placenta basis, but the results were plotted on a per-sample basis, duplicating the signature proportion estimates where multi-regional sampling was done. The error does not affect the conclusions reported in the paper.

**Fig. 1: Incorrect and corrected Fig.**

 [figure 1](#)

1d,e.

**Fig. 2: Incorrect and corrected Extended Data Fig.**



1b.

In Fig. [1d](#) caption, the text originally reading “SBS signatures in placental biopsies. Each column represents one bulk sample,” has been amended to read “SBS signatures in placentas that underwent bulk sampling. Each column represents the unique substitutions across one placenta ( $n = 36$ , placenta from trisomic rescue case not included).” The Fig. [1e](#) caption, originally beginning “Prevalence of SBS18 mutations in bulk placental biopsies,” has been amended to read “Prevalence of SBS18 mutations across each placenta that underwent bulk sampling.”

In Extended Data Fig. 1b caption, the text originally reading “Mutational signature profiles of each bulk placental sample by their clinical group. Each bar represents a single bulk placental sample,” has been amended to read “Mutational signature profiles of each placenta by their clinical group ( $n = 36$ , placenta from trisomic rescue case not included). Each column represents one placenta (signature extraction performed across all unique substitutions per placenta in view of their low individual mutation burdens).”

The original and corrected Fig. [1d,e](#) and Extended Data Fig. 1b are shown above.

The changes have been made to the online version of the article.

## Author information

### Author notes

1. These authors contributed equally: Tim H. H. Coorens, Thomas R. W. Oliver

2. These authors jointly supervised this work: D. Stephen Charnock-Jones, Gordon C. S. Smith, Sam Behjati

## Affiliations

1. Wellcome Sanger Institute, Hinxton, UK

Tim H. H. Coorens, Thomas R. W. Oliver, Rashesh Sanghvi, Roser Vento-Tormo, Muzlifah Haniffa, Matthew D. Young, Raheleh Rahbari, Peter J. Campbell & Sam Behjati

2. Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK

Thomas R. W. Oliver, Gordon C. S. Smith & Sam Behjati

3. Department of Obstetrics and Gynaecology, University of Cambridge, NIHR Cambridge Biomedical Research Centre, Cambridge, UK

Ulla Sovio, Emma Cook, D. Stephen Charnock-Jones & Gordon C. S. Smith

4. Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne, UK

Muzlifah Haniffa

5. Department of Dermatology, Royal Victoria Infirmary, Newcastle Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK

Muzlifah Haniffa

6. Great Ormond Street Hospital for Children NHS Foundation Trust, NIHR Great Ormond Street Hospital Biomedical Research Centre, London, UK

Neil Sebire

7. UCL Great Ormond Street Institute of Child Health, London, UK

Neil Sebire

8. Centre for Trophoblast Research, Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, UK

D. Stephen Charnock-Jones & Gordon C. S. Smith

9. Department of Paediatrics, University of Cambridge, Cambridge, UK

Sam Behjati

## Corresponding authors

Correspondence to [D. Stephen Charnock-Jones](#), [Gordon C. S. Smith](#) or [Sam Behjati](#).

## Rights and permissions

[Reprints and Permissions](#)

## About this article

### Cite this article

Coorens, T.H.H., Oliver, T.R.W., Sanghvi, R. *et al.* Author Correction: Inherent mosaicism and extensive mutation of human placentas. *Nature* **603**, E17 (2022). <https://doi.org/10.1038/s41586-021-04347-9>

- Published: 28 February 2022
- Issue Date: 17 March 2022
- DOI: <https://doi.org/10.1038/s41586-021-04347-9>

### Share this article

Anyone you share the following link with will be able to read this content:

[Get shareable link](#)

Sorry, a shareable link is not currently available for this article.

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04347-9>