

nature



ONE YEAR. TEN STORIES.

Ten people who helped
shape science in 2021

Imaginary friends
Putting complex
numbers in quantum
theory to the test

Smoke-free signals
Weight gain after
giving up smoking
linked to microbiota

Events directory
The *Nature* guide to
global science events
and courses in 2022

Vol. 595, No. 7390
23/30 December 2021

Nature.2022.01.01

[Sat, 01 Jan 2022]

- [This Week](#)
- [News in Focus](#)
- [Opinion](#)
- [Work](#)
- [Research](#)

This Week

- **[Sustainability at the crossroads](#)** [21 December 2021]
Editorial • A look back at 2021 through the Sustainable Development Goals.
- **[Build a registry of results that students can replicate](#)** [20 December 2021]
World View • To speed research, express conclusions as testable statements, and incorporate testing into training.
- **[Silk Road tech transfer: this ancient lyre went global](#)** [15 December 2021]
Research Highlight • A stringed instrument more than 1,000 years old found in what is now Kazakhstan is nearly identical to one from the Sutton Hoo 'ship burial' in Britain.
- **[A fatty molecule shows promise for staving off HIV](#)** [16 December 2021]
Research Highlight • When tested in monkeys, a lipid-rich molecule drove the concentration of an HIV-like virus below detectable levels.
- **[Young star's streamer of scorching-hot gas gives astronomers a fiery first](#)** [14 December 2021]
Research Highlight • Eruption from a star called EK Draconis was observed along with a gargantuan stellar explosion.
- **[Lift off! The biggest known flying creature had an explosive launch](#)** [14 December 2021]
Research Highlight • The gigantic flying reptile Quetzalcoatlus northropi, which lived in the age of the dinosaurs, could also walk and even run, with the help of 'ski pole' front limbs.
- **[Heatwaves afflict even the far north's icy seas](#)** [10 December 2021]
Research Highlight • Arctic waters have notched a growing number of extreme events called marine heatwaves, raising fears for the region's more heat-sensitive sea creatures.
- **[Why the 'wobbly bridge' did its infamous shimmy](#)** [10 December 2021]
Research Highlight • Model further challenges the idea that pedestrians walking in step caused the oscillation of London's Millennium Bridge.

- **How the little finches that amazed Darwin battle a vampire fly** [15 December 2021]

Research Highlight • Mother birds on the Galapagos Islands that hunker down in their nests are the best at warding off the bloodthirsty parasite.

| [Next section](#) | [Main menu](#) |

- EDITORIAL
- 21 December 2021

Sustainability at the crossroads

A look back at 2021 through the Sustainable Development Goals.



A medical worker observes people with COVID-19 inside a makeshift care facility at the Commonwealth Games Village in New Delhi in May 2021. Credit: Getty

There were high hopes for 2021. The year promised progress on the push for sustainable development after months of pandemic-induced delays and uncertainty. We heard ambitious talk of a ‘green recovery’, and world leaders were due to gather for meetings of the United Nations conventions on biological diversity and on climate to set future agendas.

How did the year's sustainability debates evolve? We take a look through *Nature*'s science lens.

2021: a year of multiple crises

As 2021 draws to a close, the world is facing numerous crises. The COVID-19 pandemic is far from over. A year after the first vaccines began to clear regulatory hurdles, the emergence of the SARS-CoV-2 Omicron variant is challenging the fragile and unequal gains in bringing the virus under control. Progress is slow on mitigating and adapting to climate change, protecting biodiversity and ending hunger — parts of the Sustainable Development Goals (SDGs), the United Nations' flagship plan to end poverty and promote a healthier planet by 2030. The plan, already off track before the pandemic, has been all but derailed by COVID-19.

Nature has argued¹ that the setback requires a more rapid response by the researchers who are writing the latest UN Global Sustainable Development Report — the scientific input to the SDGs, which runs on a four-year cycle. But attempts to feed science into policy have come up against strong barriers. Democracy and multilateralism are in retreat, undermining the commitment needed to make progress on sustainability goals. Still, this should not be a reason to disengage. On the contrary, researchers generally need to redouble their efforts.

Fighting the climate crisis

Early November was marked by a momentous climate summit, the 26th UN Climate Change Conference of the Parties (COP26) in Glasgow, UK. For the first time, the final agreement included mention of a phase down of coal-fired power, although phase out was the original aim. It also called for the ending of some public subsidies for other fossil fuels — one of the biggest financial barriers to the shift to renewable energy. More than 100 countries pledged to cut methane emissions, flagged for their role in global warming in the latest report from the Intergovernmental Panel on Climate Change (IPCC)². Richer nations committed to doubling their funding by 2025 to help low- and middle-income countries (LMICs) deal with the damage

already caused by climate change, and they agreed to set up an office to research a long-proposed fund to compensate LMICs for that damage.

But even if the pledges announced are implemented, temperatures are still projected to rise to a catastrophic 2.4 °C by 2100. And below the surface lay disagreements on definitions and the detail of implementation. And this is where research must continue to offer essential input. ‘Net-zero’ is one example. There is no agreed definition or measure of it, and without this, it’s impossible to know whether pledges will actually stop global warming. There is also no agreed definition of climate finance for LIMCs. This means that richer countries can make up their quotas with loans or official development aid that links to climate change only indirectly. Arguments have persisted for years over the funding promised more than a decade ago — what has been disbursed and who owes what — and this has undermined trust and has cast a shadow over negotiations, including those in the lead-up to the Glasgow meeting.



Protesters hold a ‘Biodiversity Emergency’ banner during the demonstration outside the Bank of England in London in November 2021. Credit: Vuk Valcic/SOPA Images/LightRocket/Getty

Elusive biodiversity protection

Just days before COP26, at a separate COP hosted by China in Kunming in Yunnan province, governments debated measures to protect the diversity and richness of plant and animal species. In the first sessions of a two-part UN summit on biological diversity, due to conclude in May 2022, discussions centred on a widely supported target to protect 30% of the world's land and sea areas by 2030 — up from the previous 'Aichi target' of 17%. Among other targets under debate was the provision of greater financial support to low-income countries to preserve biodiversity.

Progress on biodiversity protection has proved elusive since the first 'Earth Summit' in Rio de Janeiro in 1992. The Kunming summit ended with a modest boost in funding for projects that help to preserve biodiversity — unlike climate change, funding for biodiversity comes mostly from the public sector. We argued that these contributions should be given as grants, rather than loans that saddle poor countries with debt³. This is now more important than ever, as the pandemic piles perilous debt on the developing world.

Protecting biodiversity goes hand in hand with managing land and water resources sustainably, and in this way aligns with tackling climate change. And if nature continues to degrade, sooner or later economic output will suffer. This link is captured by debates over assigning monetary and other values to ecosystems, an idea no longer theoretical or controversial. In March, we welcomed a move by members of the UN Statistical Commission to finalize a set of principles that will help national statisticians record ecosystem health and work out payments for ecosystem services⁴.



Icebergs that calved from the Sermeq Kujalleq glacier in Greenland this year help mark one of Greenland's biggest ice-melt years in recorded history. Credit: Mario Tama/Getty

Revamping food systems

Like biodiversity protection, the world's food system needs fixing. One in ten people is undernourished and one in four is overweight. The number of people going hungry is rising fast, a trend fuelled by the pandemic. *Nature*'s coverage emphasizes the fact that science needs to guide the transformation of the food system. The task is challenging, because food spans many disciplines. We have yet to pin down what diets that are both healthy and sustainable should look like. And an IPCC-like system of scientific advice to inform policymaking has so far been missing from food and agriculture.

That changed in September, when António Guterres, the UN secretary-general, convened a controversial but historic Food Systems Summit. A group of scientists was tasked with ensuring that the science underpinning the summit was robust, broad and independent. Writing in *Nature*, this

scientific group issued seven priorities for research, among them a greater focus on sustainable aquatic foods⁵. Soil-based agriculture tends to dominate discussions on food, with ‘blue foods’ — organisms such as fish, shellfish and seaweeds — rarely considered.

Nature joined the scientific group’s call to argue that it’s time to change that (see go.nature.com/3e3ss6r). We published the Blue Food Assessment — the first systematic evaluation of how aquatic food contributes to food security — which explores how research can help transform the global food system. This work also shows some pitfalls of a greater reliance on blue foods without sustainable management, as a rapidly increasing demand for fish adds to risks for coastal ecosystems and the people of coastal communities.



Volunteers prepare meals for distribution in the Paraisópolis favela in São Paulo, Brazil, in March 2021 Credit: Jonne Roriz/Bloomberg/Getty

Strong moves from the UN’s centre

The year 2021 also saw various arms of the UN consider how their own governance needs to respond and adapt to changing times. Guterres is set to appoint a new board of scientific advisers to his office, a decision that *Nature* welcomed⁶. The decision is part of the organization's 25-year vision, laid out in the secretary-general's report, *Our Common Agenda* (see go.nature.com/3egrudq), in September. Specialized agencies also needed to stocktake. Over the fifty years since its founding, the UN Environment Programme has pushed important initiatives that bring science into 'green' policy — co-founding the IPCC, for one — and we urged it to do more to bring together researchers from across environmental sciences to tackle interconnected challenges⁷. *Nature* also urged the International Monetary Fund's shareholders to lend money to strengthen universities, so that science can better work towards global goals⁸.

The right moves at the top echelons of global governance matter – but support for science and collaboration within and between countries matter just as much. In some ways, LMICs are leading the way. A 700-page report by the UN science and cultural organization UNESCO is a first attempt to understand the impact of the SDGs on research priorities⁹. It found that, unlike richer nations, lower-income countries' share of research publications jumped in areas such as photovoltaics and climate-resilient crops. Individual countries need to do better to boost innovation, but collaboration will prove crucial. We need look no further than the pandemic for examples of how researchers working across borders, cultures and disciplines can benefit science and society.

Collaboration and inclusion

We need — and can — do better on collaboration. Global problems need diverse teams to help navigate social and geopolitical challenges. Our COVID-19 coverage comes with a host of inspiring stories of scientists joining forces to tackle the crisis. It serves as a reminder of what can be done. But it's not easy. Collaboration means spending less time achieving metrics of performance and more time nurturing relationships. Link-ups between science and industry suffer without rules around data ownership and intellectual property. And mounting geopolitical tensions, particularly

between the United States and China, are limiting exchanges of people and knowledge¹⁰.

The benefits of international research are worth the effort for both LIMCs and wealthy nations. But collaborations often come with concerns over equity and who benefits. Concerns over inclusion extend to policy forums too. At COP26, *Nature* found that researchers were frequently prevented by the organizers from accessing negotiations. Representatives from civil society and the global south also complained of exclusion. That experience must not be repeated. We've also argued that forums such as the G7 group of wealthy nations and the World Health Organization should regard emerging economies as equals. And UN bodies that solicit scientific input need to reach out beyond their usual expert networks to involve under-represented communities. The Food Systems Dialogues (see go.nature.com/3ykm2ye) could be a model: this initiative has engaged hundreds of participants across six continents since 2018, becoming an official mechanism to build international consensus at the UN food summit.

An eye on the future

Looking ahead to 2022, we're keeping our finger on the pulse. *Nature* will maintain a focus on climate, global health and sustainability. We expect more attention to the food crisis and climate-related migration, and more debate on solutions and trade-offs tied up with the energy transition.

The fallout from the pandemic will be a key focus. It includes the burden of disability from long COVID, lost ground in the fight against polio, malaria, tuberculosis and HIV, the lifelong impact of the loss of education for millions of children, and rising violence against women and girls. As economies struggle to get back on their feet, the financing of sustainability goals is an urgent issue that needs resolving. Researchers should also work towards resolving some of the long-standing tensions between climate, biodiversity conservation and food provision.

The SDGs remain a holistic framework for guiding priorities for sustainable development. In the shorter term, we look to next year's conclusion of the biodiversity summit, and the climate summit in Cairo. And we stand ready

to support science as it responds to global challenges by engaging with policy and the public.

Nature **600**, 569–570 (2021)

doi: <https://doi.org/10.1038/d41586-021-03781-z>

References

1. 1.

Nature **589**, 329–330 (2021).

2. 2.

Intergovernmental Panel on Climate Change. *Sixth Assessment Report — Climate Change 2021: The Physical Science Basis* (IPCC, 2021).

3. 3.

Nature **598**, 539–540 (2021).

4. 4.

Nature **591**, 178 (2021).

5. 5.

von Braun, J., Afsana, K., Fresco, L. O. & Hassan, M. *Nature* **597**, 28–30 (2021).

6. 6.

Nature **600**, 189–190 (2021).

7. 7.

Nature **591**, 8 (2021).

8. 8.

Nature **592**, 325–326 (2021).

9. 9.

UNESCO. *The Race Against Time for Smarter Development* (UNESCO, 2021).

10. 10.

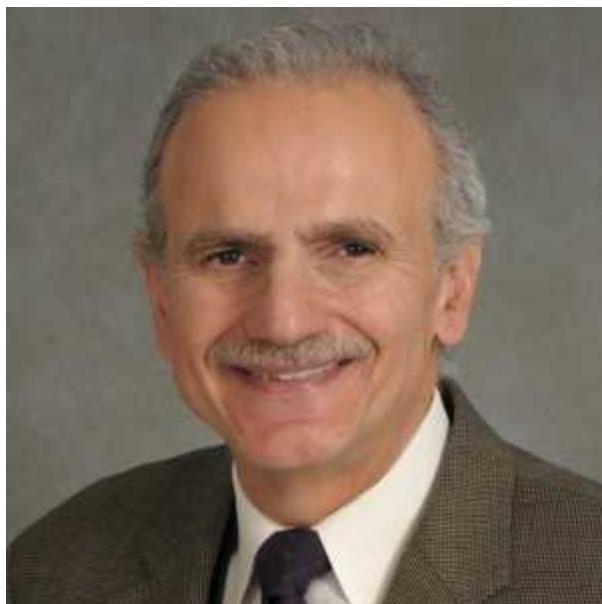
Nature **593**, 477 (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03781-z>

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

- WORLD VIEW
- 20 December 2021

Build a registry of results that students can replicate



To speed research, express conclusions as testable statements, and incorporate testing into training.

- [Yusuf A. Hannun](#) 0

Earlier this month, a large-scale replication project underscored [just how hard it is to repeat results in published papers](#). Over my three decades running a basic-biology laboratory, I have experienced the same. One of my biggest frustrations as a scientist is that it is so hard to know which exciting results are sturdy enough to build on. As I mentor early-career scientists, I try to sharpen their skill at spotting what is unlikely to be replicable, for instance published papers with oddly cropped images or protocols that mention no replicates. Still, my students have cumulatively wasted decades

of research pursuing results that were impossible to confirm. The virtuous cycle of progress is broken.

There have been countless moves to ameliorate this problem: better reporting, better career incentives, separating exploratory from confirmatory work and developing infrastructure for large, collaborative confirmatory experiments ([O. B. Amaral and K. Neves *Nature* 597, 329–331; 2021](#)).

As the year comes to a close, it's natural to consider how to improve in future. One step would be to explicitly restructure scientific publications to fulfil their functions as building blocks of knowledge. Past suggestions include requiring authors to include statements of generalizability or a numerical confidence level. Here I propose two new strategies.

First, every published study should articulate specific testable conclusions.

In my field — cancer biology — an overall conclusion might be that enzyme Y regulates cell migration in cancer. This could be built from a series of experimental results, each laid out in a quantitative way, with the relevant metrics. It's easy to imagine a series like this: (a) compound X inhibits Y *in vitro*, with a K_i (a biochemical measure of inhibition) of 300 nM; (b) compound X inhibits Y in cells with an IC_{50} (concentration giving 50% inhibition) of 1 μM ; (c) compound X inhibits cell migration with an IC_{50} of 1 μM ; (d) deletion of the gene encoding Y inhibits cell migration by >50%.

Each statement includes a ‘testable unit’. These units can be assembled at the end of the article in a ‘compendium of discrete authenticable observations’, or CODA-O. Authors can append a section in which they list specific testable units in work from other groups that they validated in the course of their current work.

The main goal is for authors to articulate and ‘own’ explicit testable statements by expressing extremely high confidence in them. That should, in turn, prompt them to clearly articulate the experimental conditions required, for instance by stating whether results were obtained in one cell

line or several cancer cell lines. That will also clarify how work might be extended (for instance, testing a result in other cell types).

Second, I propose that scientists extract these discrete observations from the literature and compile them in a registry that can seed experiments to be conducted in undergraduate or beginning graduate lab courses.

My work focuses on biologically active lipids, their roles in cell signalling pathways and how those pathways go awry in cancer. Besides working with many postdocs and graduate students, I've trained and mentored about 150 undergraduate and high-school students. Assigning a team of these trainees a claim from the registry to reproduce would be meaningful to them. Such a training programme would produce researchers skilled in troubleshooting and results that contribute to science. An approach with similar goals has been implemented in the field of psychology (K. Button *Nature* 561, 287; 2018): graduate students craft a basic protocol and supervise groups of undergraduates running that protocol.

A registry of claims for replication would bring multiple benefits. First, requirements to specify which components of an experiment will be replicable would reduce temptations to overstate and overgeneralize results. And it might head off some of the bickering that occurs when one set of researchers says it cannot reproduce another's 'work'. Second, researchers might be encouraged to describe their experiments more completely if others are more likely to attempt to formally replicate the work. Third, trainees would learn to appreciate the experimental nature of the biological sciences (too often, undergraduates are taught that all that matters is the existing body of knowledge, not how it is built). Fourth, the registry could let others check whether claims in a paper have been reproduced, and perhaps uncover subtle requirements to make an experiment work. Graduate students and trainees working to update the registry would gain practical insights on what makes work reproducible.

When too many data in the system are irreproducible, the noise becomes overwhelming — 'garbage in, garbage out'. However, with reliable building blocks of information, synthesis and bioinformatics techniques will be much more productive and predictive.

This idea has limitations. Some techniques are too specialized and some experiments too expensive or resource-intensive to be replicated. In these cases, a registry would still be useful because it would articulate the specific units of the work, and might prompt a focused replication effort if the results are deemed crucial for the community. It would also establish precedence — simply requiring claims to be registered in a testable form would prompt researchers to report what would be replicable under which conditions. That alone would spare my lab and others much fruitless work.

Nature **600**, 571 (2021)

doi: <https://doi.org/10.1038/d41586-021-03707-9>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03707-9>

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

- RESEARCH HIGHLIGHT
- 15 December 2021

Silk Road tech transfer: this ancient lyre went global

A stringed instrument more than 1,000 years old found in what is now Kazakhstan is nearly identical to one from the Sutton Hoo ‘ship burial’ in Britain.



The 1,600-year-old lyre (artist’s impression) excavated in Dzhetyasar, Kazakhstan, has the same design as those found in cemeteries thousands of kilometres away in Western Europe. Credit: Gjermund Kolltveit

A 1,600-year-old lyre found in modern Kazakhstan matches musical instruments seen in Anglo-Saxon burials of the first millennium ad, suggesting that technology transfer occurred across thousands of kilometres in antiquity¹.

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Nature **600**, 572 (2021)

doi: <https://doi.org/10.1038/d41586-021-03685-y>

References

1. 1.

Kolltveit, G. *Antiquity* <https://doi.org/10.15184/aqy.2021.164> (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03685-y>.

- RESEARCH HIGHLIGHT
- 16 December 2021

A fatty molecule shows promise for staving off HIV

When tested in monkeys, a lipid-rich molecule drove the concentration of an HIV-like virus below detectable levels.

 Coloured transmission electron micrograph of Simian HIV particles

A designer molecule prevents monkeys from being infected with simian immunodeficiency virus (artificially coloured particles shown), which is closely related to HIV. Credit: Klaus Boller/Science Photo Library

A fat-laden designer molecule keeps monkeys from becoming infected with an HIV-like virus, offering the potential that it could be used to treat AIDS and prevent HIV infection¹.

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Nature **600**, 572 (2021)

doi: <https://doi.org/10.1038/d41586-021-03750-6>

References

1. 1.

Xue, J. *et al.* *Cell* <https://doi.org/10.1016/j.cell.2021.11.032> (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03750-6>

- RESEARCH HIGHLIGHT
- 14 December 2021

Young star's streamer of scorching-hot gas gives astronomers a fiery first

Eruption from a star called EK Draconis was observed along with a gargantuan stellar explosion.

 Artist's impression of a supermassive filament released by a superflare on EK Draconis

A filament of superheated, charged gas (artist's impression) writhes across the young star EK Draconis. Credit: NAOJ

For the first time, astronomers have detected a massive stream of charged gas erupting from the surface of a young, Sun-like star — offering a glimpse of the harsh conditions faced by any surrounding planets¹.

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Nature **600**, 572 (2021)

doi: <https://doi.org/10.1038/d41586-021-03683-0>

References

1. 1.

Namekata, K. *et al.* *Nature Astron.* <https://doi.org/10.1038/s41550-021-01532-8> (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03683-0>

- RESEARCH HIGHLIGHT
- 14 December 2021

Lift off! The biggest known flying creature had an explosive launch

The gigantic flying reptile *Quetzalcoatlus northropi*, which lived in the age of the dinosaurs, could also walk and even run, with the help of ‘ski pole’ front limbs.

 An artist’s interpretation of *Quetzalcoatlus northropi* wading in the water

Even walking, the pterosaur *Quetzalcoatlus northropi* (artist’s impression) cut an impressive figure, standing about 3 metres tall. Credit: James Kuether

With an 11–12-metre wingspan — comparable to that of a small plane — the winged reptile *Quetzalcoatlus northropi* must have cast a fearsome shadow as it flew above what is now Texas, 70 million years ago. It remains the largest known flying animal, but relatively few bones of it have been found and almost nothing about it has been published.

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Nature **600**, 572 (2021)

doi: <https://doi.org/10.1038/d41586-021-03677-y>

References

1. 1.

Padian, K., Cunningham, J. R., Langston, W. Jr & Conway, J. J. *Vertebr. Paleontol.* **41**, 218–251 (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03677-y>.

- RESEARCH HIGHLIGHT
- 10 December 2021

Heatwaves afflict even the far north's icy seas

Arctic waters have notched a growing number of extreme events called marine heatwaves, raising fears for the region's more heat-sensitive sea creatures.



An Arctic iceberg. Rising air temperatures are contributing to the growing incidence of marine heatwaves in the region. Credit: Felipe Dana/AP/Shutterstock

Ocean 'heatwaves' are becoming increasingly common even in the frigid Arctic, where these events are more extreme, and possibly more damaging, than those in warmer seas farther south¹.

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Nature **600**, 573 (2021)

doi: <https://doi.org/10.1038/d41586-021-03676-z>

References

1. 1.

Huang, B. *et al.* *Geo. Res. Lett.* <https://doi.org/10.1029/2021GL095590> (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03676-z>

- RESEARCH HIGHLIGHT
- 10 December 2021

Why the ‘wobbly bridge’ did its infamous shimmy

Model further challenges the idea that pedestrians walking in step caused the oscillation of London’s Millennium Bridge.



The sheer number of people on the Millennium Bridge in London set the structure swaying in 2000, modelling reveals. Credit: Grant Rooney Premium/Alamy

Crossing the Thames was never more rhythmic than it was on 10 June 2000, when London’s Millennium Bridge opened and then visibly rocked as pedestrians ambled across. The bridge’s motion has been cited as evidence of the dangers of synchronized walking — everyone seemed to move in unison, supposedly tapping into the resonant frequency of the bridge and causing it to sway.

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Nature **600**, 573 (2021)

doi: <https://doi.org/10.1038/d41586-021-03675-0>

References

1. 1.

Belykh, I. *et al.* *Nature Commun.* <https://doi.org/10.1038/s41467-021-27568-y> (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03675-0>

- RESEARCH HIGHLIGHT
- 15 December 2021

How the little finches that amazed Darwin battle a vampire fly

Mother birds on the Galapagos Islands that hunker down in their nests are the best at warding off the bloodthirsty parasite.



Attentive parenting by birds (such as *Geospiza fuliginosa*; pictured) of the Galapagos Islands helps to keep the avian vampire fly away from their nests. Credit: Tui De Roy/Minden Pictures/Alamy

Some of the small, plain songbirds known as Darwin's finches might protect their young from flesh-eating parasites just by staying in the nest¹.

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Nature **600**, 573 (2021)

doi: <https://doi.org/10.1038/d41586-021-03684-z>

References

1. 1.

Kleindorfer, S. *et al. Proc. R. Soc. B*
<https://doi.org/10.1098/rspb.2021.1668> (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03684-z>

News in Focus

- **Imperilled glacier, COVID immunity and ‘touching’ the Sun** [22 December 2021]
News Round-Up • The latest science news, in brief.
- **How severe are Omicron infections?** [17 December 2021]
News • As cases spread and countries plan their response, researchers await crucial data on the severity of the disease caused by the coronavirus variant.
- **Evidence of racism found at prestigious London university**
[15 December 2021]
News • London School of Hygiene & Tropical Medicine has pledged to revise equity and diversity policies after a review highlighted racial inequality among students and staff.
- **Omicron blindspots: why it’s hard to track coronavirus variants** [16 December 2021]
News • Researchers are rapidly sequencing the genomes of virus samples worldwide, but shortcomings in the global surveillance system make the task a challenge.
- **How COVID vaccines shaped 2021 in eight powerful charts**
[16 December 2021]
News • The extraordinary vaccination of more than four billion people, and the lack of access for many others, were major forces this year — while Omicron’s arrival complicated things further.
- **The best science images of 2021** [13 December 2021]
News • A helicopter on Mars, human–monkey embryos, volcanic ash and more.
- **Nature’s 10: ten people who helped shape science in 2021**
[15 December 2021]
News Feature • An Omicron investigator, a Mars explorer and an AI ethics pioneer are some of the people behind the year’s big research stories.

- NEWS ROUND-UP
- 22 December 2021

Imperilled glacier, COVID immunity and ‘touching’ the Sun

The latest science news, in brief.



The Thwaites Glacier’s fractures were identified in satellite imagery.Credit: NASA

Antarctic Glacier edges closer to collapse

Giant fractures in the floating ice of Antarctica’s massive Thwaites Glacier — a fast-melting formation that has become an icon of climate change — [could shatter part of the shelf within five years](#), research suggests. If that happens, in what had been considered a relatively stable part of Thwaites, the glacier could begin flowing much faster into the ocean, funnelling ice

that had been resting on land into the sea, where it would contribute to sea-level rise.

For decades, scientists have carefully tracked changes in Thwaites Glacier, which already loses around 50 billion tonnes of ice each year and causes 4% of global sea-level rise. The latest fractures are deep, fast-moving cracks in its eastern ice shelf. They have appeared in satellite images over the past few years and their growth seems to be accelerating.

“I visualize it somewhat similar to that car window where you have a few cracks that are slowly propagating, and then suddenly you go over a bump in your car and the whole thing just starts to shatter in every direction,” said Erin Pettit, a glaciologist at Oregon State University in Corvallis, on 13 December at the American Geophysical Union meeting. If Thwaites’s eastern ice shelf collapses, ice in this region could flow into the sea up to three times faster, Pettit says.



A man recovering from COVID-19 exercises at a rehabilitation centre in Genoa, Italy.Credit: Marco Di Lauro/Getty

Waning COVID super-immunity raises concerns

Research suggests that [protection against the SARS-CoV-2 virus slides over time](#) even in people who have experienced both an infection and had vaccines against it, a combination that initially provides hyper-charged immunity. The research, which has not yet been peer reviewed, was conducted before the Omicron variant emerged (Y. Goldberg *et al.* Preprint at medRxiv <https://doi.org/g9rq>; 2021). But it sharpens questions about how well ‘super-immunity’ — also known as hybrid immunity — will fare against the latest iteration of the coronavirus.

Scientists are now racing to learn by how much Omicron eludes immunity conferred by infection or vaccination. Preliminary results from laboratory studies suggest that vaccination and super-immunity offer some protection against the variant, and that super-immunity might offer more than jabs alone.

However, previous evidence shows that the available vaccines’ potency against other coronavirus variants gradually dwindles, and public-health officials say that Omicron’s spread has heightened the need for people to receive a booster dose. Whether hybrid immunity will prove more durable than vaccines against Omicron is not yet known.



The Parker Solar Probe will travel past the Sun 24 times, looping ever closer to its surface.Credit: NASA/Johns Hopkins APL

NASA spacecraft ‘touches’ the Sun for the first time ever

A NASA spacecraft has [entered a previously unexplored region of the Solar System](#) — the Sun’s outer atmosphere, or corona. The long-awaited milestone, which was reached in April but announced on 14 December, is a major accomplishment for the Parker Solar Probe, a craft that is flying closer to the Sun than any mission in history.

“We have finally arrived,” said Nicola Fox, director of NASA’s heliophysics division, at the agency’s headquarters in Washington DC. “Humanity has touched the Sun.”

She and other team members spoke during a press conference at the American Geophysical Union meeting in New Orleans, Louisiana, held on 13–17 December. A paper describing the findings appeared that week in *Physical Review Letters* ([J. C. Kasper et al. Phys. Rev. Lett. 127, 255101; 2021](#)).

In many ways, the Parker Solar Probe is a counterpoint to NASA’s twin Voyager spacecraft. In 2012, Voyager 1 travelled so far from the Sun that it became the first mission to leave the region of space dominated by the solar wind — the energetic flood of particles coming from the Sun. By contrast, the Parker probe is travelling ever closer to the heart of the Solar System, flying head-on into the solar wind and our star’s atmosphere. With this new front-row seat, scientists can explore some of the biggest unanswered questions about the Sun, such as how it generates the solar wind and how its corona gets heated to temperatures more extreme than those on the Sun’s surface.

“This is a huge milestone,” says Craig DeForest, a solar physicist at the Southwest Research Institute in Boulder, Colorado, who is not involved in the mission. Flying into the solar corona represents “one of the last great unknowns”, he says.

The Parker probe crossed the boundary, called the Alfvén surface, into the Sun's atmosphere at 09:33 universal time on 28 April. As it crossed the Alfvén surface, the probe flew through a ‘pseudostreamer’ of electrically charged material, inside which conditions were quieter than the roiling environment outside of it.

Since its launch in 2018, the spacecraft has been orbiting the Sun: with each pass, it loops ever closer to the solar surface. It ultimately aims to make 24 close passes of the Sun. It crossed the Alfvén surface on the eighth of those fly-bys.

Nature **600**, 575 (2021)

doi: <https://doi.org/10.1038/d41586-021-03708-8>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03708-8>

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

- NEWS
- 17 December 2021

How severe are Omicron infections?

As cases spread and countries plan their response, researchers await crucial data on the severity of the disease caused by the coronavirus variant.

- [Heidi Ledford](#)



The Omicron variant is causing hospitalizations in many nations. Credit: Laurent Gillieron/EPA-EFE/Shutterstock

It has been less than four weeks since the announcement that [a mutation-laden coronavirus variant](#) had been discovered in southern Africa. Since then, dozens of countries around the world have reported Omicron cases — including [a worrying number of infections](#) in people who have either been vaccinated or experienced previous SARS-CoV-2 infections.

But as political leaders and public-health officials try to chart a course through oncoming Omicron surges, they must do so without a firm answer to a key question: how severe will those Omicron infections be?

So far, the data are scarce and incomplete. “There is inevitably a lag between infection and hospitalization,” says infectious-disease epidemiologist Mark Woolhouse at the University of Edinburgh, UK. “In the meantime, policy decisions have to be made and that’s not straightforward.”

Hospitalization rate

Early results suggest a glimmer of hope. Reports from South Africa have consistently noted a lower rate of hospitalization as a result of Omicron infections compared with infections caused by the Delta variant, which is currently responsible for most SARS-CoV-2 infections globally. On 14 December, the South African private health insurer Discovery Health in Johannesburg announced that hospitalization risk has been 29% lower among people infected with Omicron, compared with people infected with a previous variant.

This has fuelled suggestions that Omicron causes milder disease than previous variants. But researchers say it is too early to be sure, and key methodological details of that study have not yet been published. Such details are crucial when interpreting data on disease severity, which can be confounded by factors such as hospital capacity, the age and overall health of those initially infected, and the extent of previous exposure to coronavirus.

But the results from Discovery Health are in keeping with other studies in the country, says Waasila Jassat, a clinician and public-health specialist at the National Institute for Communicable Diseases in Johannesburg. “There

are many caveats and disclaimers around early severity data,” she says. “But the picture is very consistent.”

It will take time for a consistent picture to emerge from countries that currently have fewer Omicron infections. On 13 December, Denmark released data showing that hospitalization rates for people infected with Omicron seemed to be on a par with those for people infected with other variants. But this comparison was based on only about 3,400 cases of Omicron infection and 37 hospitalizations.

Similarly, [a 16 December report](#) from Imperial College London found no evidence of diminished hospitalizations from Omicron infections compared with Delta in England, although this was again based on relatively few cases. Overall, the numbers are still too small to draw firm conclusions about the severity of disease caused by Omicron, says Troels Lillebæk, an infectious-disease specialist at the University of Copenhagen.

And a rapidly spreading variant could dangerously strain health-care systems, even if the risk of severe disease or death is relatively low for any individual. “A small fraction of a very large number is still a large number,” says Woolhouse. “So the population-level threat is very real.”

South Africa’s optimistic data might not be a sign that Omicron itself is more benign than previous variants. More than 70% of the population in regions heavily infected with Omicron have had previous exposure to SARS-CoV-2, and about 40% have received at least one dose of a COVID-19 vaccine, says Jassat. This makes it difficult to disentangle the effects of pre-existing immunity from inherent properties of the variant itself.

Vaccine protection

Laboratory studies have suggested that Omicron [might be able to evade some COVID vaccine-induced immunity](#), and early data from the UK Health Security Agency suggest that the vaccines are not as protective against Omicron infections as they have been against other variants, although the number of cases studied was too small to be sure about how much protection has decreased.

Even so, vaccines could continue to protect many recipients from severe disease and death from COVID-19. In addition to antibodies, the immune system in previously infected and vaccinated people deploys cells called T cells that can recognize fragments of viral proteins and destroy virus-infected cells, potentially [limiting the scope of an infection](#).

Researchers have mapped Omicron's panoply of mutations onto the menu of SARS-CoV-2 protein fragments recognized by T cells following natural infection and vaccination, and found no mutations in most of these fragments. In the case of vaccination, more than 70% of the fragments are fully intact, according to immunologist Alessandro Sette at the La Jolla Institute for Immunology in California.

There is more work to be done — scientists are already conducting laboratory assays to determine how well T cells generated in response to vaccines and infection with other variants react to Omicron, with results expected in the coming weeks. "I'm optimistic that the reactivity is going to be preserved, at least in part," Sette says. "How much of it will be preserved remains to be seen."

At the moment, there is no way to draw a direct line between the degree of T-cell reactivity and protection against severe disease. Previous studies have found that robust T-cell reactions to SARS-CoV-2 are correlated with lower viral loads and less-severe disease, but do not establish a threshold at which that protection might begin to wane, Sette says. Ultimately, it will again come down to waiting for data on hospitalizations and deaths from Omicron.

Infections in children

As those data emerge, researchers will be looking particularly at the effects of Omicron on children. Results from South Africa have suggested that hospitalization rates for children infected with Omicron are higher than were seen in previous waves. But researchers again caution that this does not necessarily mean that children are more vulnerable to Omicron than they were to Delta or other variants. Jassat notes that children have lower rates of previous coronavirus infection and vaccination than adults, meaning that their levels of pre-existing immunity are not as high.

Higher rates of hospitalization in children during the early stages of an outbreak could also reflect more hospital capacity, affording the luxury of keeping a child in for observation who might otherwise be sent home, she adds.

And the setting in which children are exposed can also play a part: extended exposures at home from an infected parent could mean a higher initial exposure to the virus than a transient exposure at school, says David Dowdy, an infectious-disease epidemiologist at Johns Hopkins Bloomberg School of Public Health in Baltimore, Maryland. “Everyone is focused on the pathogen here,” he says. “But it’s not just about the variant, it’s also about the host and the environment.”

Nature **600**, 577-578 (2021)

doi: <https://doi.org/10.1038/d41586-021-03794-8>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03794-8>

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

- NEWS
- 15 December 2021

Evidence of racism found at prestigious London university

London School of Hygiene & Tropical Medicine has pledged to revise equity and diversity policies after a review highlighted racial inequality among students and staff.

- [Linda Nordling](#)



The London School of Hygiene & Tropical Medicine has promised to revise its efforts to promote equity and inclusion. Credit: David A Eastley/Alamy

Evidence of racism and inequality at the London School of Hygiene & Tropical Medicine (LSHTM) — a prestigious 120-year-old research university — has been reported in an independent review that was commissioned by the institution last year.

The LSHTM says it is “determined to do better” after the review found that staff members of colour remain under-represented and face barriers to career advancement.

The [70-page report](#), published on 13 December, notes that “stakeholders” perceive that the LSHTM itself has not addressed and acknowledged the ways in which it has benefited from and perpetuated European colonialism. Staff and students also told its authors that they don’t trust that complaints about racist behaviour will be dealt with fairly or sensitively.

The findings are “difficult to confront”, said LSHTM director Liam Smeeth in a [statement](#). “Facing up to them is an essential step towards creating an environment where everyone’s contributions and perspectives are valued,” he added. “We deeply regret this and apologise sincerely to everyone affected.”

Research at LSHTM focuses on tropical diseases and public health, and includes several projects related to the COVID-19 pandemic.

“So many global-health institutions have either looked away or been lukewarm in response to growing calls to address inequities and injustices in their operations,” says Sèye Abimbólá, a global-health specialist based at the University of Sydney in Australia. On whether the review will result in meaningful change at the LSHTM, he adds: “We’ll have to wait and see.”

‘Blatant racism’

The LSHTM commissioned the review in 2020, after the Black Lives Matter movement inspired discussions about racism worldwide. The review’s publication follows [calls to “decolonize” the funding structures of global-health research](#), and [tensions around inequality in high-profile international health-research programmes](#). The LSHTM is not the first university to undertake such a review: Imperial College London has published a number

of reports, including one from a focus group in 2019 that [highlighted experiences of racism and inequality](#).

The LSHTM report was produced by a group of six independent reviewers, supported by three experts in racial equality. They reviewed data and documents from the LSHTM, surveyed current and former staff members and students, and had targeted discussions with some participants.

Fifty-two per cent of survey respondents who were people of colour said that they had witnessed or experienced racism at the university, and the review heard of several instances in which the LSHTM had failed to act on complaints about racist behaviour. Staff and students of colour also told the reviewers that they were reluctant to challenge racist acts because they feared causing offence, or were concerned about repercussions.

“Reports are brushed under the carpet,” one former staff member told the reviewers. A member of the professional support staff said that senior staff members who are the subject of complaints “get a slap on the wrist behind a closed door”. A current staff member said the “blatant racism” in their research unit was “out of this world”.

The review’s authors point out that the survey is not representative of all LSHTM staff, because participants were self-selected. But the report noted that institutional data shows that staff members of colour have faced greater barriers to career progression than their white counterparts. Two-thirds of white staff members who applied for promotion from assistant to associate professor between 2017 and 2020 were successful, compared with one-third of staff members of colour.

Fifty-eight per cent of people of colour who responded to the survey ‘disagreed’ or ‘strongly disagreed’ that people of colour had equal opportunities to make progress in their careers at the LSHTM. Only 27% of white respondents gave the same answers.

Structural issues

“These structural issues are not specific to LSHTM. They are issues that will have gone on in other places too. It’s really important for institutions to dig

in and address them,” says Mishal Khan, a social epidemiologist at the LSHTM who sits on the institution’s council.

Ngozi Erondu, an epidemiologist and public-health adviser at Chatham House, a policy institute in London, notes that the LSHTM was founded before most colonized African countries became independent between the 1950s and the 1970s. “Pretty much any institution established before then really needs to review and reflect not only on its historical links to colonialism but also its current practices that may have just continued since that time”, she says.

The review makes a number of recommendations for the LSHTM, including mandating anti-racism training for senior leaders, and ensuring that procedures for complaints and reporting are visible and protect whistle-blowers.

Smeeth said in his statement that the university will revise its equity, diversity and inclusion plan by the end of January 2022, in line with the review’s recommendations.

“It will not be a quick or easy journey, but work is already under way and this review accelerates and strengthens the change that is needed,” says Mohamed Osman, an independent member of the LSHTM’s council and chair of its diversity and inclusion committee. “If, by releasing this report, LSHTM is taking one of many steps to help to create a more inclusive and equitable sector, then that is one positive outcome,” he adds.

Nature **600**, 578-579 (2021)

doi: <https://doi.org/10.1038/d41586-021-03755-1>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03755-1>

- NEWS
- 16 December 2021

Omicron blindspots: why it's hard to track coronavirus variants

Researchers are rapidly sequencing the genomes of virus samples worldwide, but shortcomings in the global surveillance system make the task a challenge.

- [Amy Maxmen](#)



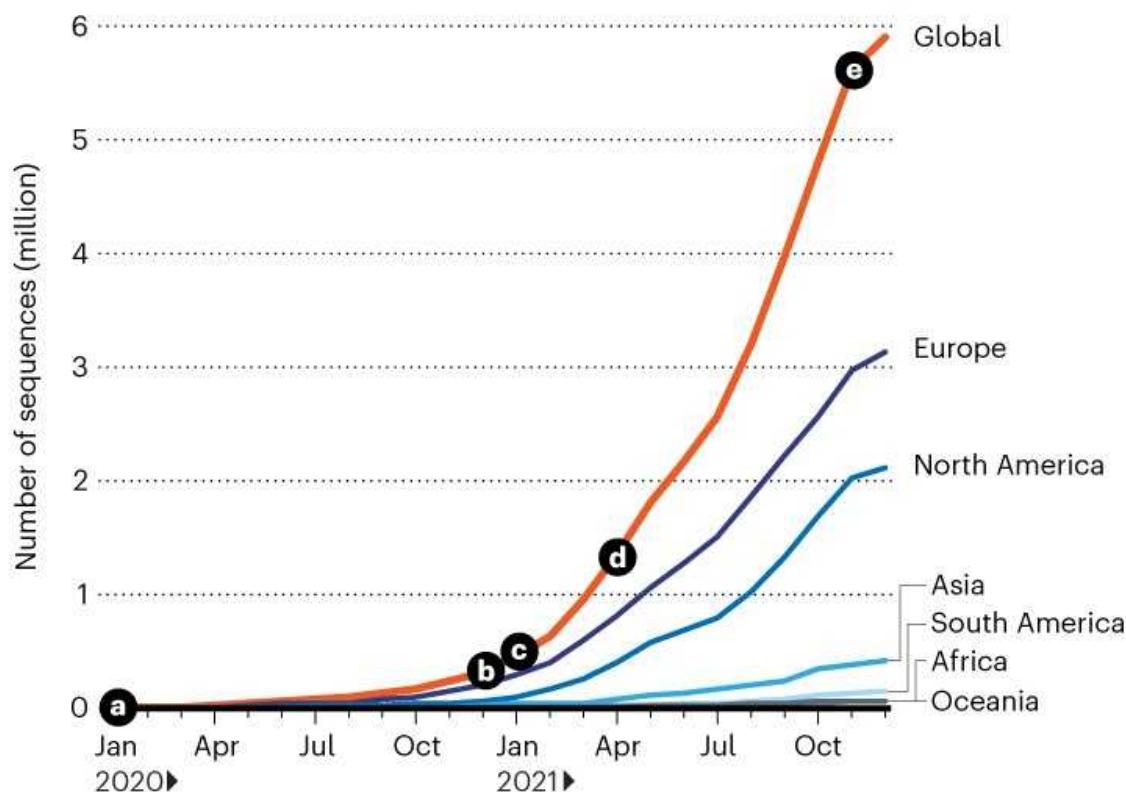
Laboratories such as this one in New Delhi, India, are sequencing the genomes of coronavirus samples, racing to detect the Omicron variant. Credit: T. Narayan/Bloomberg via Getty

Researchers are racing to detect Omicron, the latest SARS-CoV-2 variant of concern, by sequencing the genomes of coronaviruses infecting people. But surveillance through genomic sequencing can be slow and patchy, complicating the picture of how and where Omicron spreads.

One positive development is that researchers are sequencing more SARS-CoV-2 genomes than ever before. This is what enabled them to notice Omicron relatively swiftly. Last April — about 16 months into the pandemic — an online database belonging to the GISAID data-science initiative [contained one million SARS-CoV-2 genomic sequences](#). Since then, researchers have submitted another five million sequences to GISAID in about eight months — a nearly tenfold rate increase (see ‘Genome explosion’). “We are in much better shape to find Omicron or any other emerging variant now,” says Kelly Wroblewski, director of infectious diseases at the Association of Public Health Laboratories in Silver Spring, Maryland.

GENOME EXPLOSION

Scientists have shared about six million SARS-CoV-2 genome sequences on the GISAID data-sharing platform since January 2020, many of them in just the past eight months. Those deposited from Africa recently proved invaluable in sounding the alarm on the Omicron variant.



- **a January 2020:** First genome of the SARS-CoV-2 coronavirus reported
- **b December 2020:** Alpha and Beta variants named
- **c January 2021:** Gamma variant named
- **d April 2021:** Delta variant named
- **e November 2021:** Omicron variant named

©nature

Source: GISAID

Yet researchers warn that there are still troubling gaps in sequencing data that make any interpretation of a variant's movement fraught. "The numbers are complex, and there are so many caveats," Wroblewski says. For one, some countries don't have the laboratory capacity to sequence pathogen

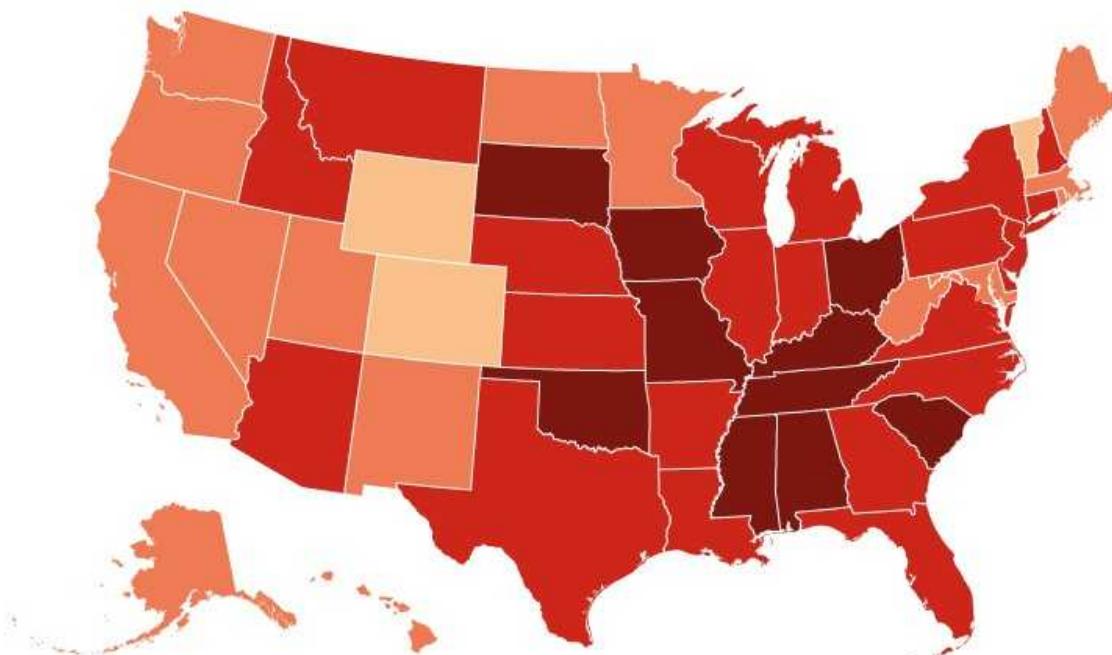
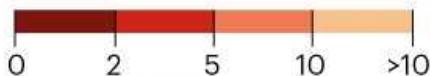
genomes, so it might look like those places have no variants, when in fact the mutated viruses are spreading under the radar.

Sequencing rates vary within countries, as well, yielding an uneven picture of how a variant is spreading within a nation's borders. For instance, 10 US states have sequenced less than 2% of the coronaviruses infecting people who tested positive for COVID-19 in those states in the past month, according to sequences posted at GISAID. By contrast, Wyoming, Colorado and Vermont sequenced more than 10% of their positive cases over the same time frame (see 'States of surveillance').

STATES OF SURVEILLANCE

The percentage of laboratory-sequenced coronavirus genomes varies widely across the United States. Researchers recommend sampling at least 5% of samples in order to detect Omicron or other emerging variants.

% of detected COVID-19 cases sequenced in past 30 days*



*Data as of 13 December

©nature

Sources: GISAID and the Pandemic Prevention Institute at the Rockefeller Foundation

But even if a location is sequencing many of its positive cases, variants could still slip by if testing is poor or biased. “It’s easy to sequence 100% of your cases if you only test a few people to begin with,” explains Jennifer Nuzzo, an epidemiologist at Johns Hopkins University in Baltimore, Maryland. For example, some countries mainly test international travelers. Even if they sequence all of those samples, they might miss a concerning variant that is circulating domestically.

Minding the data gap

Faced with such surveillance challenges, epidemiologist Sam Scarpino and his colleagues at the Pandemic Prevention Institute at the Rockefeller Foundation in Washington DC have been seeking new ways to understand the spread of variants. One method is to use a model they’ve developed to estimate how prevalent Omicron would need to be in a given locality before it would be detected by public-health officials, given the state of testing and sequencing in that particular area. Omicron would need to be relatively common for researchers to identify it in a place with little surveillance, for example.

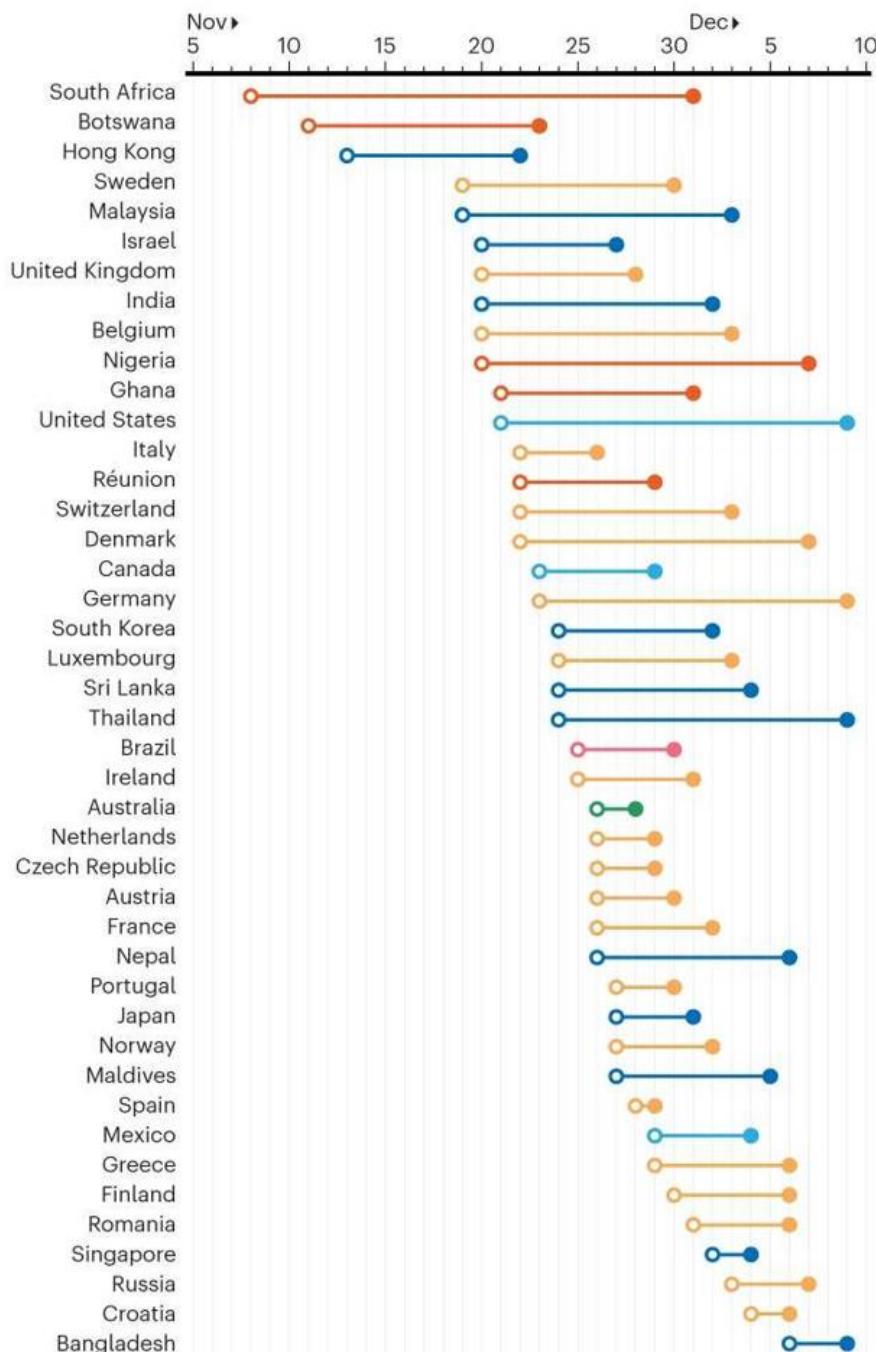
The team is also constructing timelines using Omicron reports that are uploaded to GISAID each day, to draw a clearer picture of detection. They order sequences on the basis of the dates that the samples were collected — rather than when they appear online in the database. Timing can be confusing because weeks might pass between when a person tests positive for the coronavirus and when a sample is shipped to a genomics lab, sequenced and then reported online and to authorities. For example, according to data that was on GISAID as of 9 December, the first person known to have been infected with Omicron was sampled in South Africa on 8 November, about three weeks before the viral sequence for that particular sample was posted online — and nearly two weeks before South Africa’s first report of Omicron. Since then, more data have streamed in, and a new sequence of Omicron dates back to a sample that had been collected in South Africa on 5 November. In contrast, hardly two days passed between sampling the first person known to have been infected in Spain and sequencing (see ‘Sequence of events’).

SEQUENCE OF EVENTS

Weeks can pass between when a SARS-CoV-2 coronavirus sample is collected from a person and when the genome sequence of the virus is submitted to a database, such as the one at the GISAID initiative. This makes tracking the spread of the Omicron variant tricky. Here, the earliest cases of Omicron in 43 locations are plotted on the basis of sample collection date, as of 9 December.

■ Africa ■ Asia ■ Europe ■ North America ■ Oceania ■ South America

○ Earliest sample collected ● Sequence submitted to GISAID



Sources: GISAID and the Pandemic Prevention Institute at the Rockefeller Foundation

Dave Luo, a data scientist who advises Rockefeller's pandemic institute, warns that this type of timeline can't alone determine how Omicron is spreading. To do that, scientists must compare the genetic codes of different SARS-CoV-2 sequences, building an evolutionary tree that shows how closely related one virus is to another. Genomic epidemiologists, such as those working on the [Nextstrain project](#), are currently conducting these sorts of analyses.

All of these studies are evolving daily as new Omicron sequences pour in from around the world. A hint of how fast this field is moving can be seen in the rapid rise in genomes reported after the World Health Organization named Omicron a variant of concern on 26 November. Soon after the agency's announcement, 15 countries submitted 187 genomic sequences belonging to Omicron to GISAID. By 14 December, 55 countries had shared 4,265 Omicron sequences. The figures are on course to balloon further — but Luo warns that's not necessarily representative of how fast the variant is spreading. Many testing centres are preferentially sequencing samples after a simple, fast genotyping test picks up a possible signal for Omicron — a particular amino acid in the gene for its spike protein. As a result, Omicron might be overrepresented among SARS-CoV-2 genome sequences right now.

Genomic information is biased and messy in so many ways, Luo says. “We have to be careful about what we take away from any one source of data.”

Nature **600**, 579 (2021)

doi: <https://doi.org/10.1038/d41586-021-03698-7>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03698-7>

- NEWS
- 16 December 2021

How COVID vaccines shaped 2021 in eight powerful charts

The extraordinary vaccination of more than four billion people, and the lack of access for many others, were major forces this year — while Omicron's arrival complicated things further.

- [Smriti Mallapaty](#),
- [Ewen Callaway](#),
- [Max Kozlov](#),
- [Heidi Ledford](#),
- [John Pickrell](#) &
- [Richard Van Noorden](#)



A mobile vaccination team at work during a house call in a remote region of Turkey. Credit: Chris McGrath/Getty

A year ago, vaccine drives against COVID-19 were just beginning. Now, more than 4.4 billion people have had one or more dose — about 56% of the world population. The vaccination of so many in such a short space of time, so soon after the unparalleled rapid development of the vaccines, has saved huge numbers of lives and is a triumph for science and research.

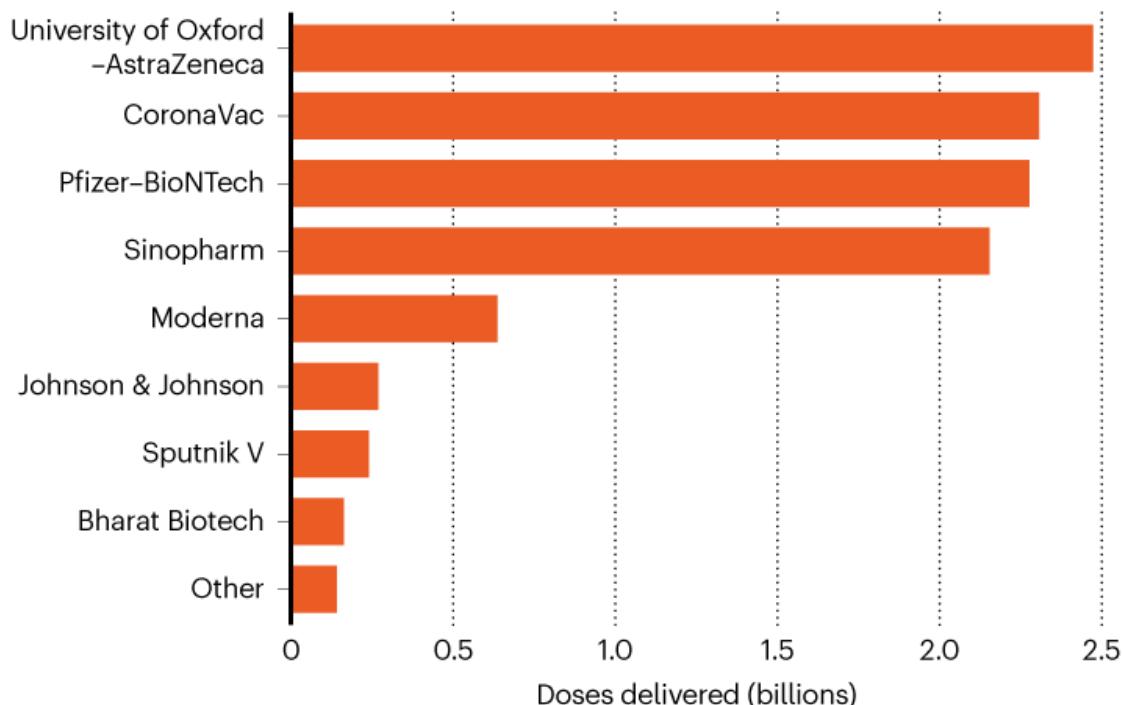
Sadly, the vaccines have not been shared or taken up equitably across the world, nor even, sometimes, within nations. But the extraordinary roll-out of a plethora of COVID-19 vaccines — or the lack thereof — has been a major force shaping politics, science and everyday human experience in 2021. In this graphic-led story, *Nature* offers a guide to the successes, failures and impact of COVID-19 vaccines in 2021.

Winning the vaccine race

More than eight billion doses, mostly of eight front-runner vaccines, have now been administered around the world, the vast majority in 2021 (see ‘The race to vaccinate’). “Just making that much vaccine has been the standout success,” says Gagandeep Kang, a virologist at the Christian Medical College in Vellore, India.

THE RACE TO VACCINATE

Nearly 10 billion doses of COVID-19 vaccine have been delivered around the world since mid-2020, 8.5 billion of which had been administered by late 2021. Eight different vaccines make up the vast majority of doses*.



©nature

*Data as of 14 December 2021.

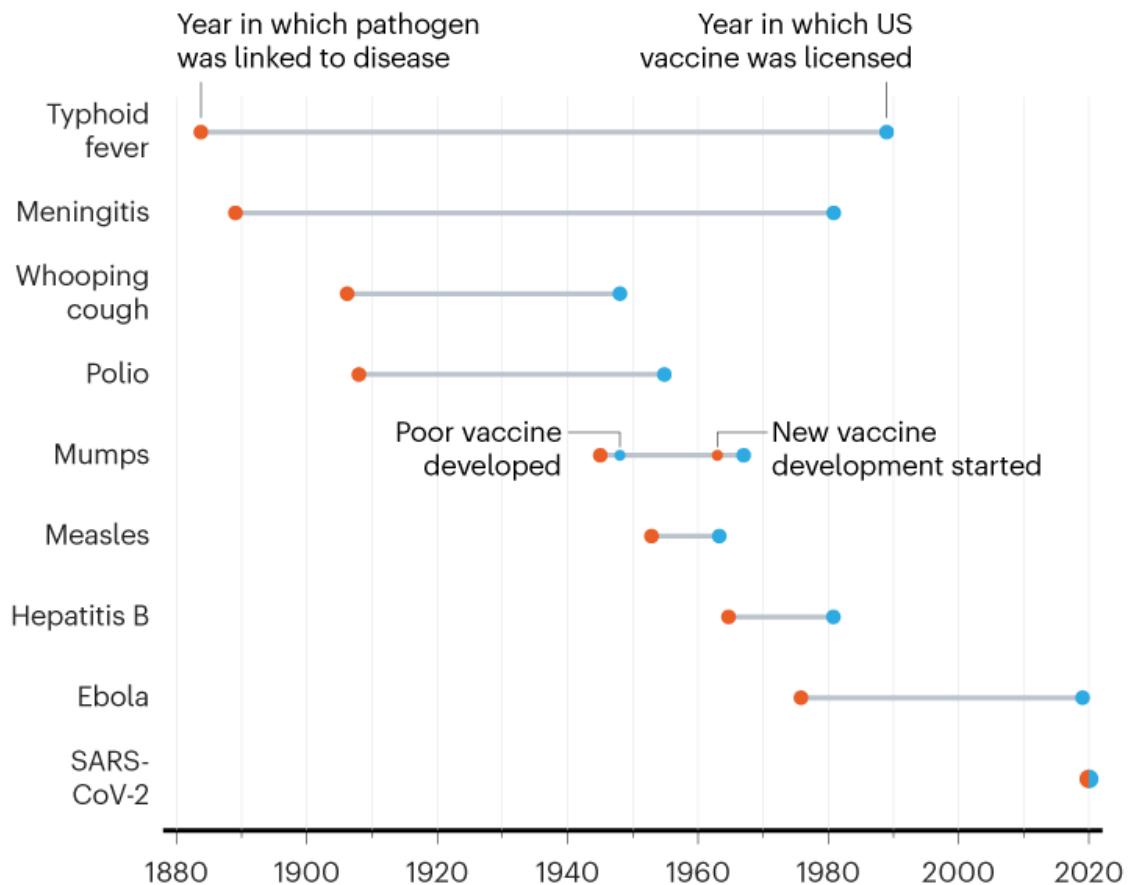
Source: Data from Airfinity

“The vaccines have had a huge impact on averting deaths and helping country’s economies return to normal,” says Soumya Swaminathan, chief scientist at the World Health Organization (WHO) in Geneva, Switzerland. “In countries with high coverage, infections have been uncoupled from deaths, so that even with new surges of infection, deaths have stayed low.”

Also noteworthy is the speed of the development of the vaccines (see ‘Vaccine innovation’). No vaccines in history have been developed so fast, yet 23 different vaccines against SARS-CoV-2 have already been approved for use around the world — and hundreds more are in development.

VACCINE INNOVATION

Most vaccines take years to develop, but scientists created multiple vaccines for SARS-CoV-2 within a year.



©nature

Source: Our World in Data; *Nature* analysis

It is estimated that this astonishingly rapid development and deployment has saved at least 750,000 lives in the United States and Europe alone — and probably many more globally, although researchers are as yet unwilling to commit to a number. A study from the WHO and the European Centre for Disease Prevention and Control in Solna, Sweden, published last month¹

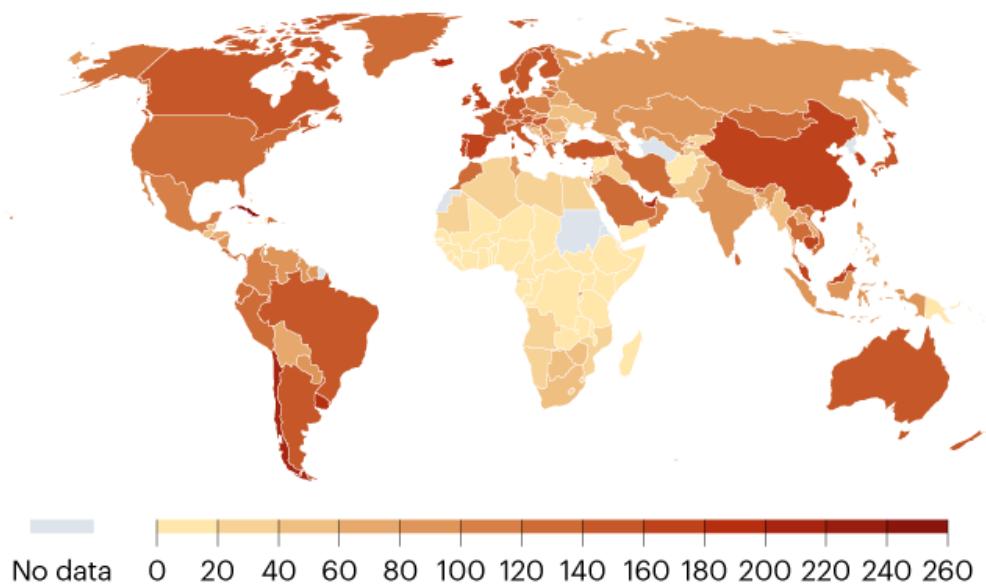
estimated that 470,000 deaths had been averted across 33 European countries in those aged 60 and over alone. Another [modelling study](#), which is yet to be peer reviewed, from epidemiologists at Yale University in New Haven, Connecticut, estimated that 279,000 lives had been saved by late June by the vaccination drive in the United States.

Vaccine haves and have-nots

But despite the astonishing success of the vaccines, it's a story of haves and have-nots and the roll-out has been anything but equitable. "We were so together and so divided," says Kang. "Very together on the science, very divided on the access."

GLOBAL DOSES

Vaccines have been rolled out unevenly across the world, as shown by the number of COVID-19 vaccine doses administered per 100 people in the total population*.



*Data as of 29 November 2021. Data don't reflect the number of people who have been vaccinated because some people have received two doses of a vaccine. Nature publications remain neutral with regard to contested jurisdictional claims in published maps.

©nature

Source: Our World in Data

In the world's most-vaccinated nations, such as the United Arab Emirates, Chile and Cuba, more than 200 doses have been administered per 100 people — but at the opposite end of the scale, in places such as Tanzania,

Afghanistan and Papua New Guinea, fewer than 20 people per 100 have received at least one dose (see ‘Global doses’).

“Vaccine inequity has been one of the most painful experiences of the pandemic,” says Swaminathan, who notes that there now exist two parallel worlds. In some regions, infections have been uncoupled from deaths and life is normalizing. But in others, there is “fear in opening up, schools remain shut, long-term plans cannot be made, and surges in infections translate soon into higher deaths”, she says.

On average, in high-income countries, 83% of the eligible populations have had at least one shot, but in low-income countries that falls to 21%. These figures “never cease to amaze”, says Andrew Azman, an infectious-disease epidemiologist at Johns Hopkins University in Baltimore, Maryland, who co-authored an analysis on the inequities in doses, posted as a preprint² in October.

It was expected that poorer nations would get increased supplies once demand began to fall in wealthy nations, but most of them are now administering boosters. This, combined with the fact that many countries are stockpiling doses, could be contributing to a lack of access to those who really need them, says Kang.

Disparities exist not just between countries, but also within them. One study in the United States found lower vaccination coverage in areas that had larger numbers of people on low incomes, or who were single parents, or who had disabilities³. Other studies show disparities in vaccination coverage along racial or ethnic lines⁴.

Waning immunity and variants

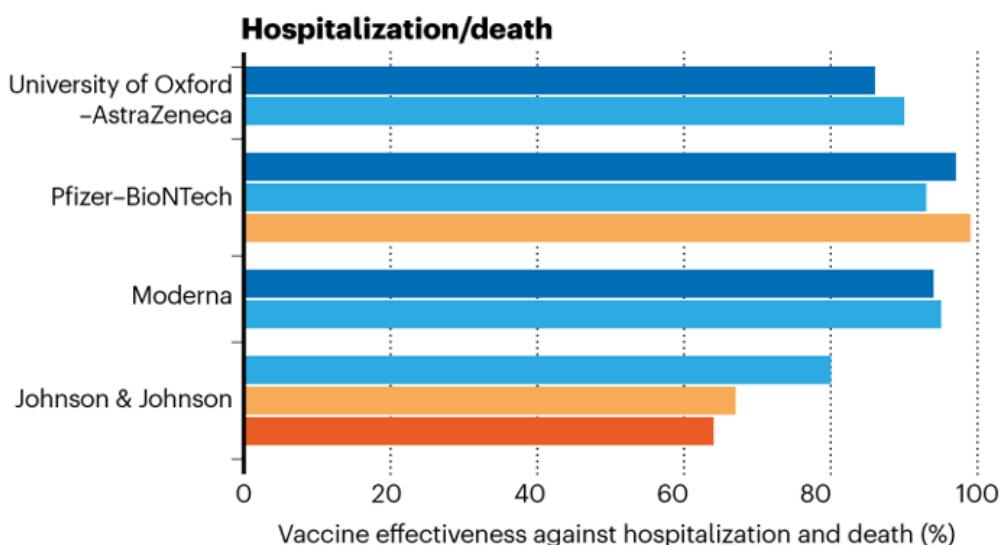
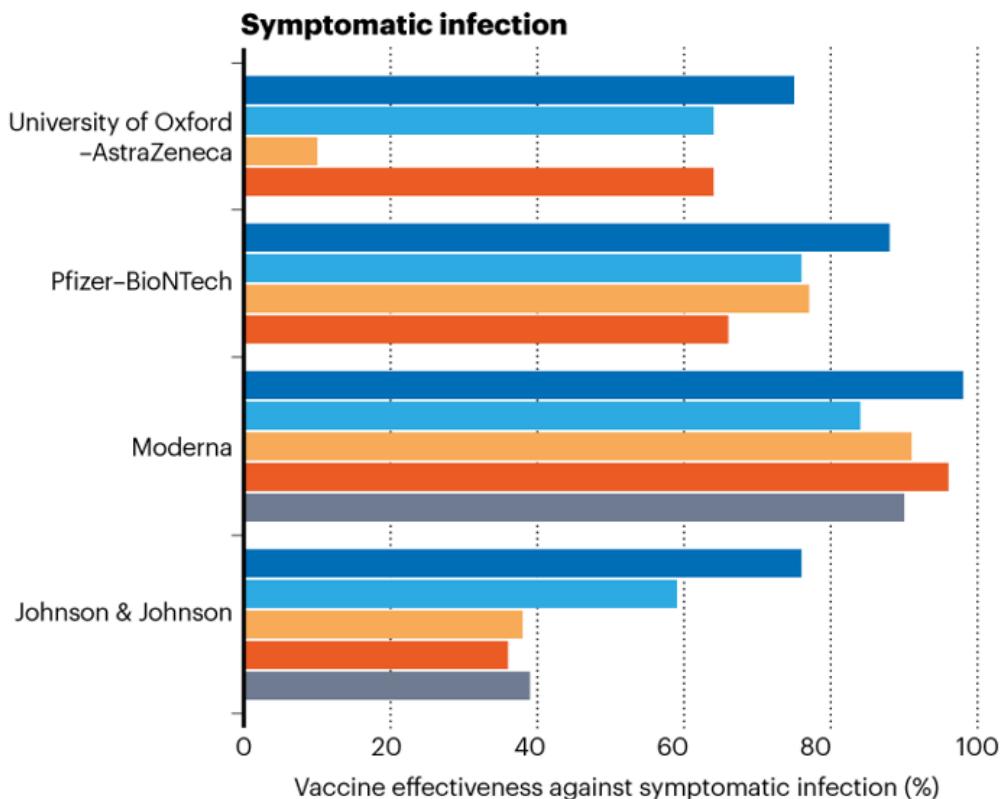
2021 was the year of COVID-19 vaccines, but it was also the year of variants. Researchers identified a trio of SARS-CoV-2 ‘variants of concern’ in late 2020 and early 2021, now called Alpha, Beta and Gamma. They seemed to spread faster than earlier circulating viral lineages, and scientists worried that these variants might also blunt the effectiveness of vaccines.

Laboratory studies and real-world epidemiology confirmed that vaccines remained highly effective against the most widespread of the three, Alpha, which was identified in the United Kingdom. But Beta and Gamma — first spotted in South Africa and Brazil, respectively — were linked to reduced effectiveness of some vaccines, particularly those based on viral vectors, such as the Oxford–AstraZeneca vaccine, or on inactivated viruses, such as those developed in China and India (see ‘Variants and vaccines’).

VARIANTS AND VACCINES

Over the course of the past year, the emergence of SARS-CoV-2 variants Alpha, Beta, Gamma, Delta and Mu has challenged the effectiveness of vaccines, although most have held their ground. How vaccines will fare in the face of highly mutated Omicron is yet to be determined*.

■ Alpha ■ Delta ■ Beta ■ Gamma ■ Mu



Source: Data from Airfinity

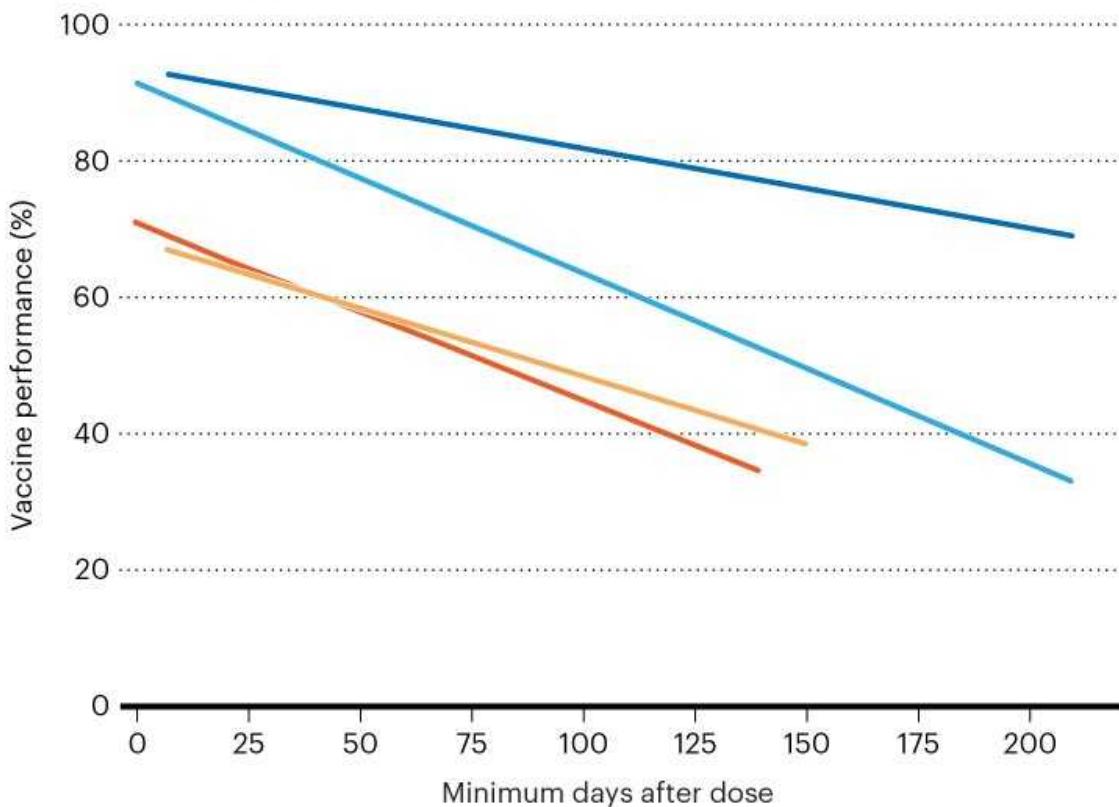
Delta, designated as a variant of concern in May, is currently responsible for most new infections globally and has further challenged vaccines. Countries such as Israel, the United States and the United Kingdom that began their campaigns early are now seeing signs that vaccines lose their potency over time (see ‘Waning immunity’).

Despite these challenges, the vaccines are still doing a good job at protecting against the most severe forms of COVID-19, says Laith Jamal Abu-Raddad, an infectious-disease epidemiologist at Weill Cornell Medicine–Qatar in Doha. “We now have lots of data and we see a very clear pattern that the vaccines are working very well against severity.”

WANING IMMUNITY

The immunity conferred by COVID-19 vaccines, particularly to prevent infections, falls over time — as shown in these estimates of vaccine efficacy against Delta in the months following a second dose*

— AstraZeneca — Johnson & Johnson — Moderna — Pfizer-BioNTech



©nature

*Data as of 25 November 2021. Estimates of vaccine effectiveness modelled by Airfinity, based on available data.

Source: Data from Airfinity

However, researchers are racing to determine how different vaccines [will hold up against the fast-spreading Omicron](#), designated a variant of concern in late November. A preliminary study from the United Kingdom found that two vaccine doses offer little protection against becoming infected with Omicron (a third booster dose restored vaccine effectiveness to above 70%). Researchers expect that vaccines will continue to prevent severe disease caused by the variant — but to what extent is not yet clear.

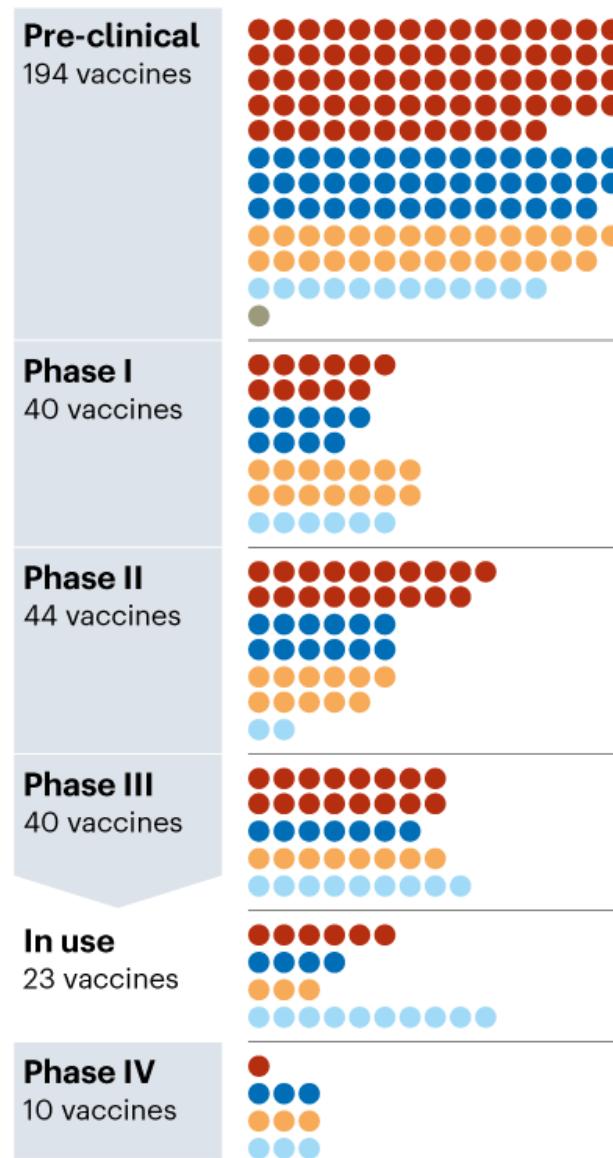
New vaccines on the horizon

While a little less than half the world's population still awaits a first dose of a COVID-19 vaccine, researchers are developing more than 300 fresh options. Of these, nearly 200 are still being tested in lab and animal studies rather than in people, but 40 are in large international clinical trials (see 'Under development').

UNDER DEVELOPMENT

Researchers are developing more than 300 COVID-19 vaccines in addition to the 23 already in use around the world; 84 are in early-stage clinical trials and 40 are at much later stages of development*.

- Protein based
- Viral vector
- Nucleic acid
- Whole virus
- Bacterial antigen-spore expression vector



*Data as of 1 December 2021

©nature

Source: GAVI

Some of these next-generation vaccines could have key advantages over those currently available. For example, [protein vaccines](#) use SARS-CoV-2

proteins to rouse the immune system against the virus, and promise to be easier to produce and transport than some existing vaccines.

In particular, two protein vaccines made by Novavax, in Gaithersburg, Maryland, and Clover Biopharmaceuticals in Chengdu, China, will be pivotal to hitting the COVID-19 Vaccines Global Access (COVAX) initiative's goal of distributing two billion doses to low-income nations next year, says Nicholas Jackson, head of programmes and innovative technology at the Coalition for Epidemic Preparedness Innovations (CEPI) in Oslo, which is a co-leader of COVAX.

Other upcoming COVID-19 vaccines are being formulated so that they can be administered by mouth or inhaled through the nose, such as nasally administered vaccines being developed by CanSino and AstraZeneca. Because these vaccines would be administered directly into the tissues that SARS-CoV-2 first infiltrates when it enters the body, it is hoped that oral or nasal vaccines could work well to prevent infection. They would also require fewer trained health-care personnel to administer injections.

Some COVID-19 vaccines are being developed to tackle specific SARS-CoV-2 variants — such as Omicron — or even a variety of coronaviruses. Three diseases caused by novel coronaviruses have already emerged in less than 20 years, says Jackson — severe acute respiratory syndrome (SARS) in 2002, Middle East respiratory syndrome (MERS) in 2012 and COVID-19 in late 2019. “A broadly protective coronavirus vaccine could revolutionize our response to future infectious-disease outbreaks,” he says.

Vaccinating children

How the pandemic unfolds from now on might not only be driven by novel variants, but also by how quickly vaccines reach another large part of the global population that is yet to be vaccinated — children.

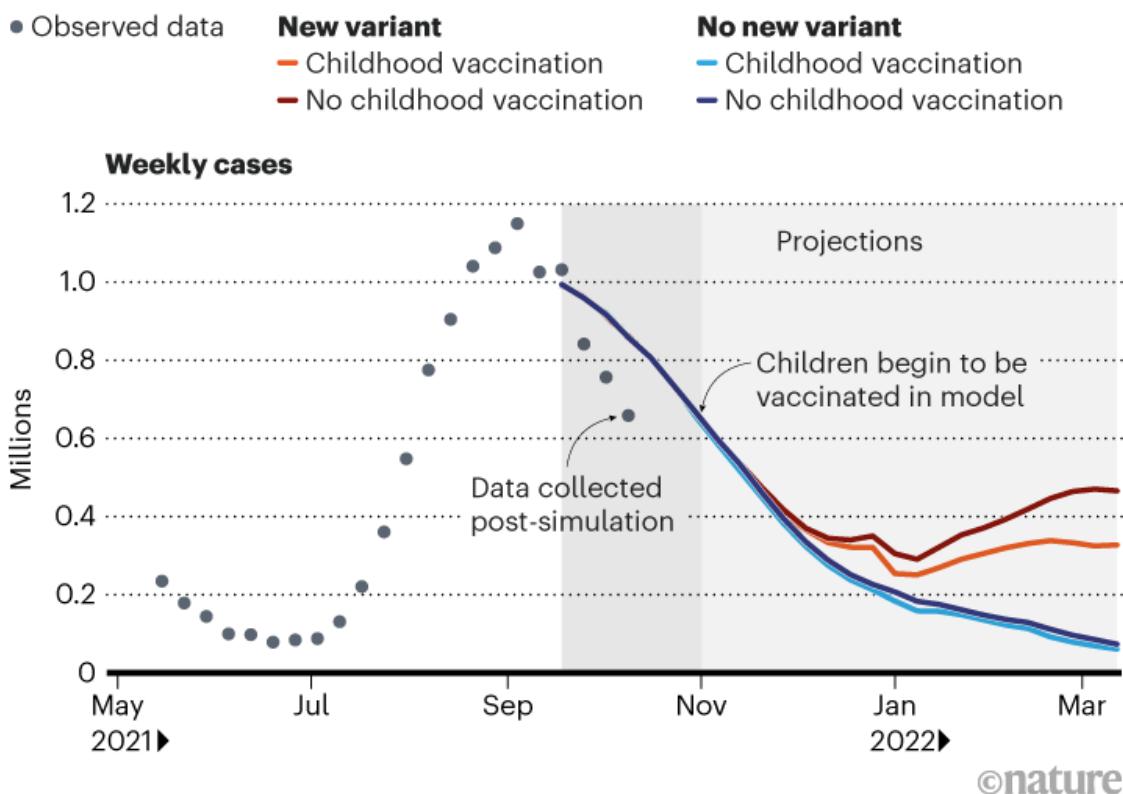
During 2021, the highly transmissible Delta variant caused a sharp rise in paediatric cases worldwide. Although only a relatively small proportion of kids develop severe disease, that still translates to huge numbers of severe cases globally, says Andrew Pavia, a paediatric infectious-disease researcher

at the University of Utah Health in Salt Lake City. Widespread vaccination of children will limit severe cases in that age group and help to control the spread of the virus, he says.

In the United States — where children have accounted for the largest numbers of COVID-19 cases of any age group since late October — the Food and Drug Administration (FDA) approved Pfizer–BioNTech's vaccine [for the nation's roughly 28 million kids](#) aged 5 to 11 in early November. Since then, more than five million children there have received a dose — and [modelling studies](#) in September that looked at the impact both in a scenario where there were no new variants and where there were, show that the benefits could be significant — particularly now, as we face the impacts of Omicron (see ‘The kid effect’). The same researchers are now starting to model the possible impacts of Omicron on case numbers in the United States.

THE KID EFFECT

A simulation of the US pandemic, run in September and averaging multiple models, found that starting to vaccinate children aged 5 to 11 would not only lower COVID-19's toll, but would also have a large impact if a new, more transmissible coronavirus variant emerged.



Source: [COVID-19 Scenario Modelling Hub](#)

Elsewhere, vaccinations for younger children have slowly been taking hold, too. Regulators in Canada, Israel and the European Medicines Agency, for example, all provisionally approved the Pfizer vaccine for children in late November, followed by Australia in early December. Colombia, Chile, Argentina and Venezuela are all now offering China's Sinopharm vaccine to children.

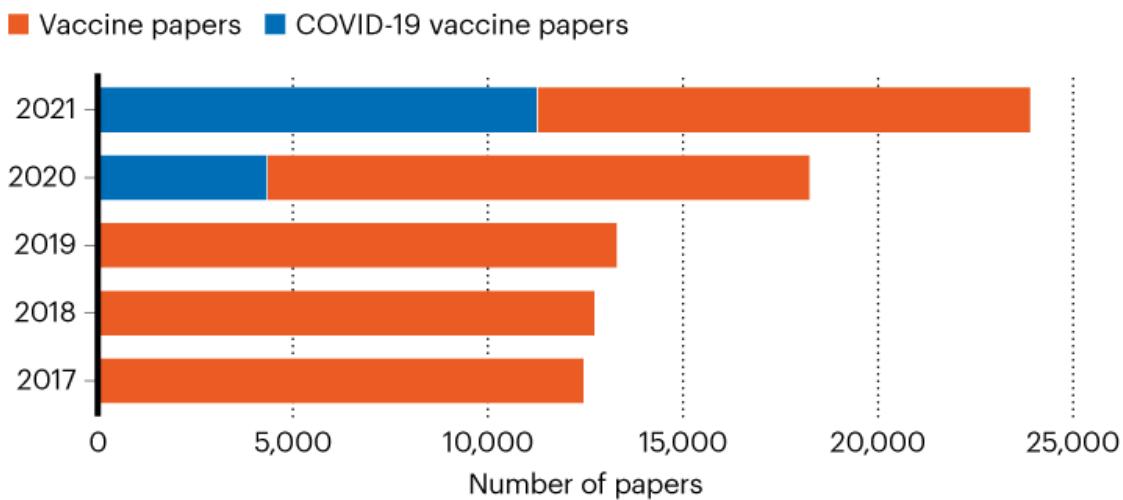
Vaccine papers soar

The development and deployment of COVID-19 vaccines has seen an extraordinary research effort over the past year. According to *Nature's*

calculations, at least 15,000 papers on vaccines mentioning COVID-19 or SARS-CoV-2 have been published since early last year, with more than 11,000 of those during 2021 (see ‘Explosion of knowledge’). These made up more than 47% of all papers on vaccines published in 2021 — and made it a record-breaking year for vaccine-related publications.

EXPLOSION OF KNOWLEDGE

More than 15,000 vaccine-related papers that mention COVID-19 or SARS-CoV-2 have been published since early last year; 11,000 were published in 2021 alone, making up an astonishing 47% of all vaccine-related publications this year*.



*Journal articles, preprints, and clinical trial reports indexed on the PubMed database. Data as of 24 November 2021.

©nature

Source: Data from PubMed; *Nature* analysis

The benefits of that research extends beyond just COVID-19, to vaccines more generally, say researchers. “Humanity coming together to develop and deploy vaccines has opened up a lot of doors for vaccines and understanding what they are, how they work and why we might want to use them in the future,” says Azman.

Vaccines will continue to save lives and help some semblance of normal life to return and energize researchers. But the extent to which the world curtails the pandemic in 2022 and beyond will depend on how quickly it provides access in low-income nations, administers boosters in populations with

waning immunity, and provides doses to children — as well as the nature and extent of new variants, such as Omicron.

Nature **600**, 580–583 (2021)

doi: <https://doi.org/10.1038/d41586-021-03686-x>

References

1. 1.

Meslé, M. M. I. *et al.* *Euro Surveill.* **26**, pii=2101021 (2021).

2. 2.

Chen, Z. *et al.* Preprint at <https://doi.org/10.1101/2021.10.25.21265504> (2021).

3. 3.

Barry, V. *et al.* *Morb. Mortal. Wkly Rep.* **70**, 818–824 (2021).

4. 4.

Wrigley-Field, E. *et al.* Preprint at
<https://doi.org/10.1101/2021.11.19.21266612> (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03686-x>

The best science images of 2021

A helicopter on Mars, human–monkey embryos, volcanic ash and more.

Emma Stoye

13 December 2021



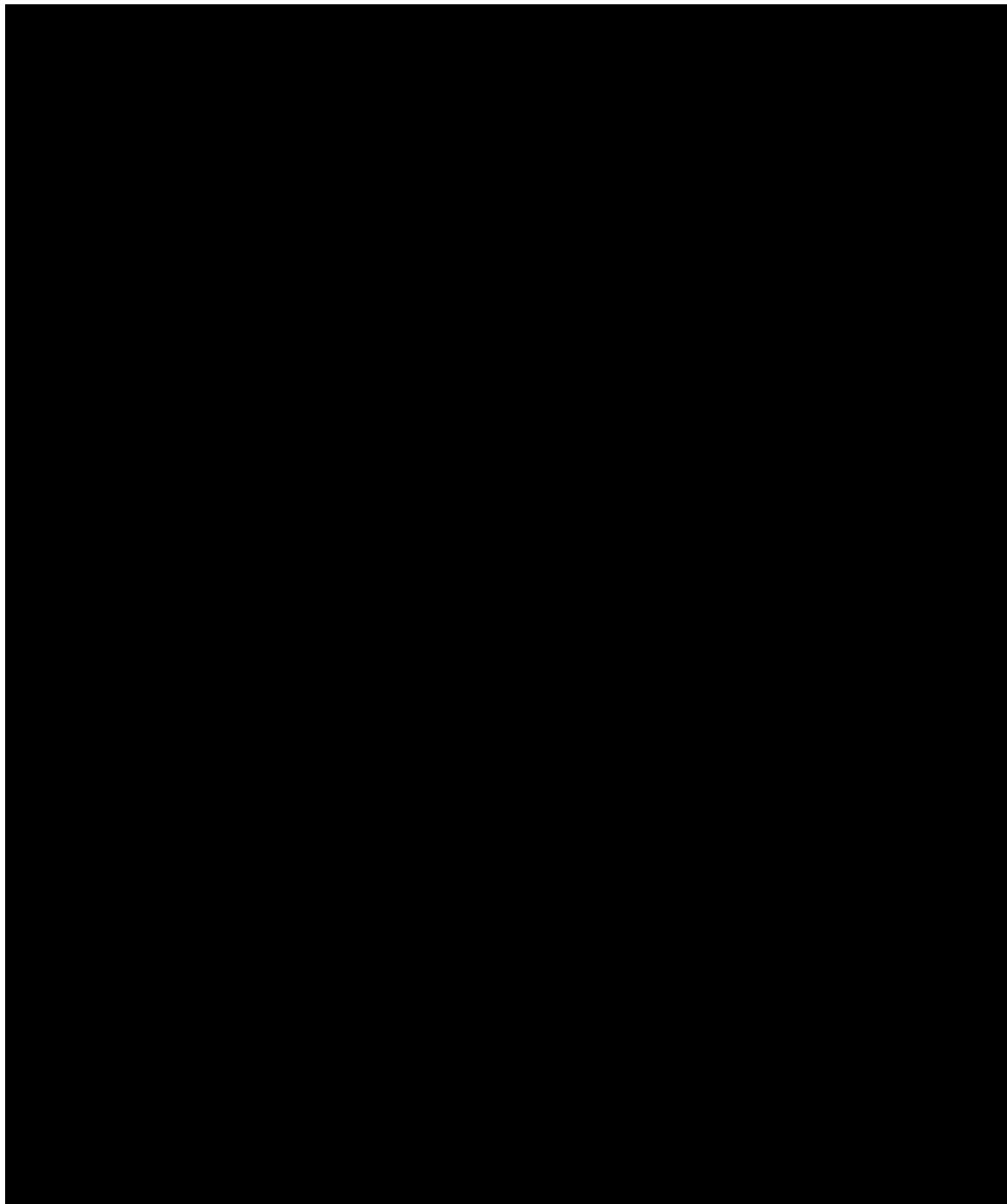
COVID-19 continued to dominate lives in 2021. But the year also produced many stunning science images unrelated to the pandemic. From human–monkey embryos to volcanic ash, here are the most striking shots that caught the attention of Nature's news and art teams.

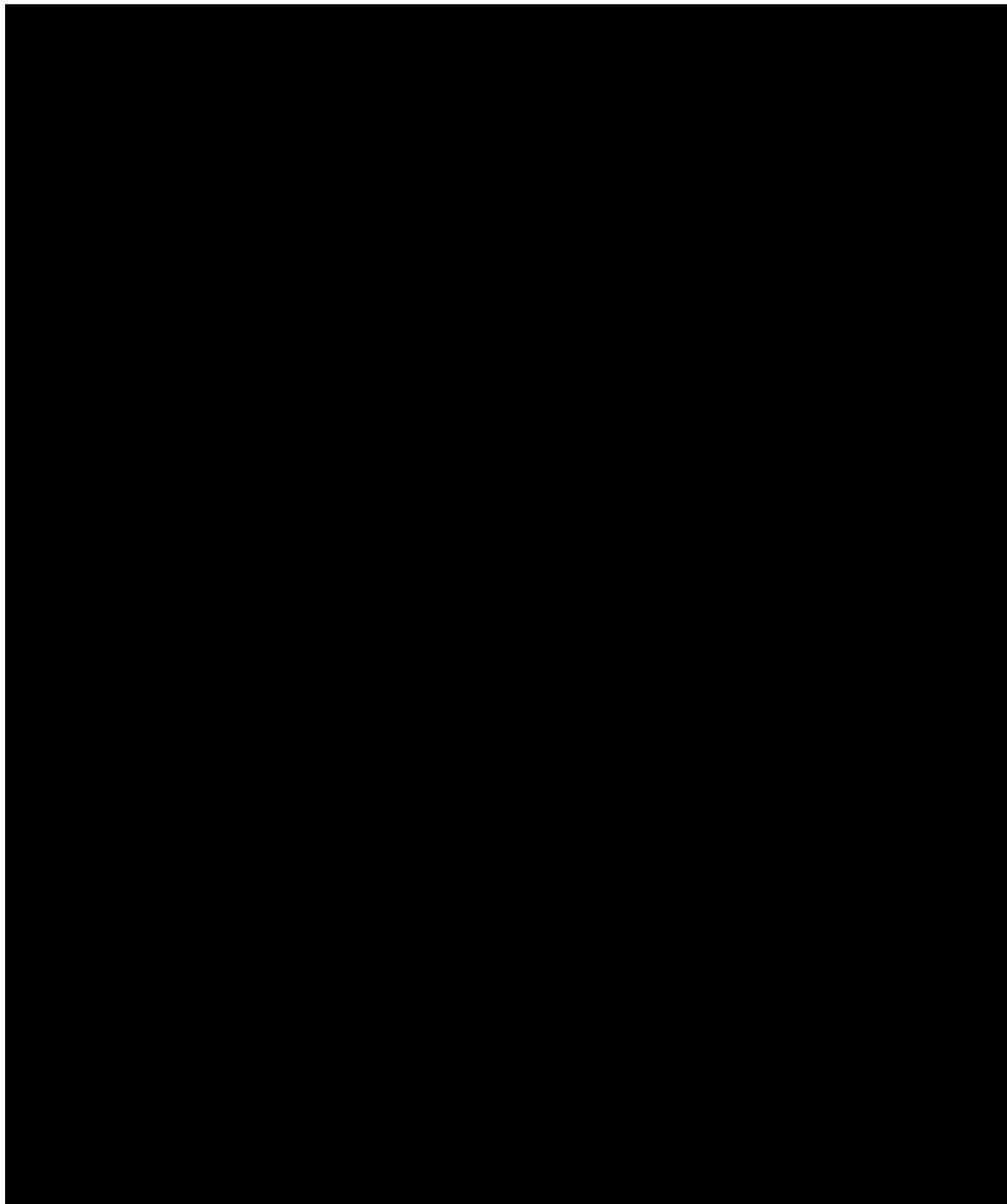


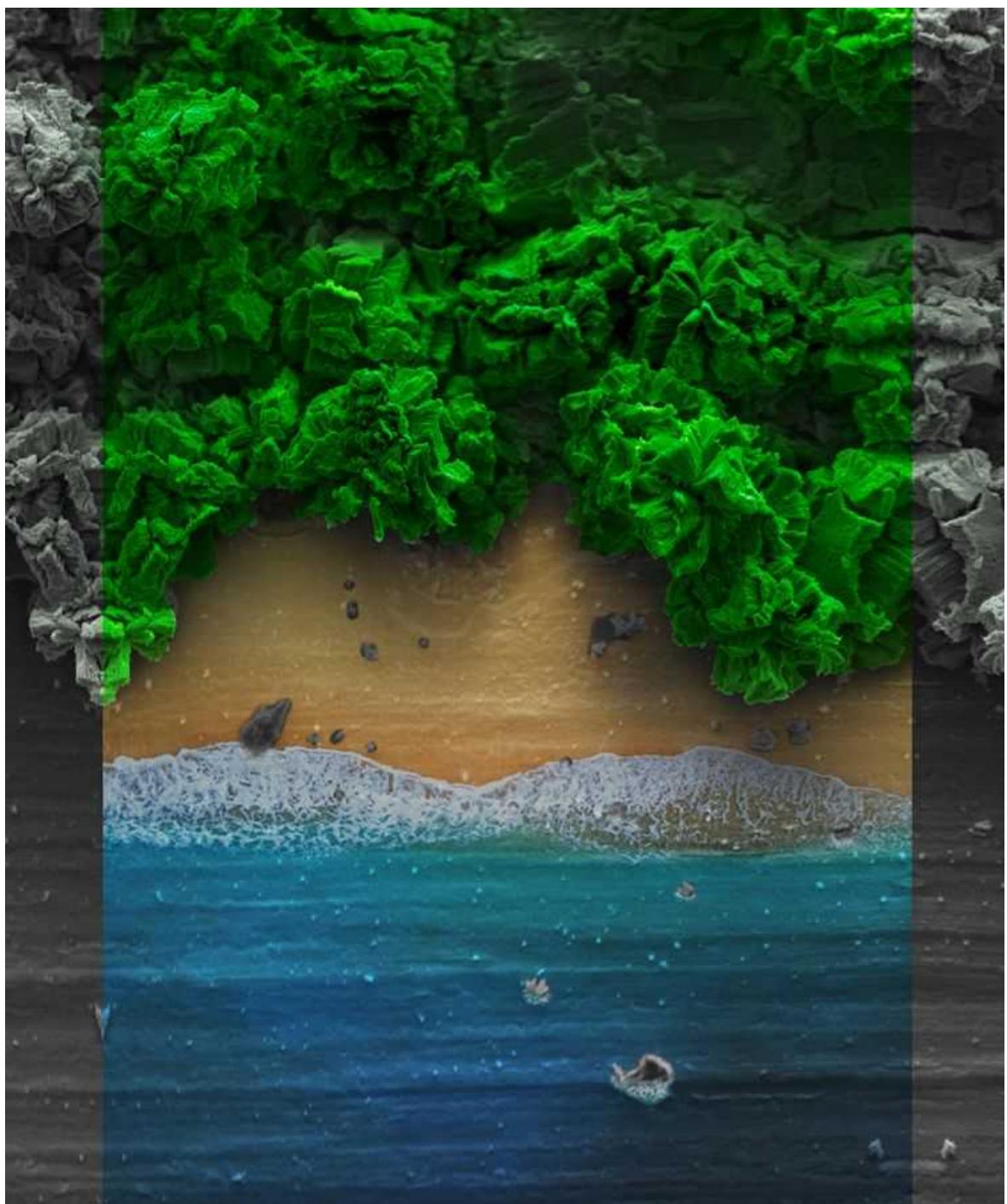
Credit: Emilio Morenatti/AP/Shutterstock

Credit: Emilio Morenatti/AP/Shutterstock

Volcanic ash. The [eruption of the Cumbre Vieja volcano](#) on La Palma in the Spanish Canary Islands in September left vast areas of land buried by ash. The volcano also spewed out a huge volume of lava, which destroyed hundreds of homes as it flowed all the way to the island's west coast, eventually spilling over the cliffs and into the Atlantic Ocean.



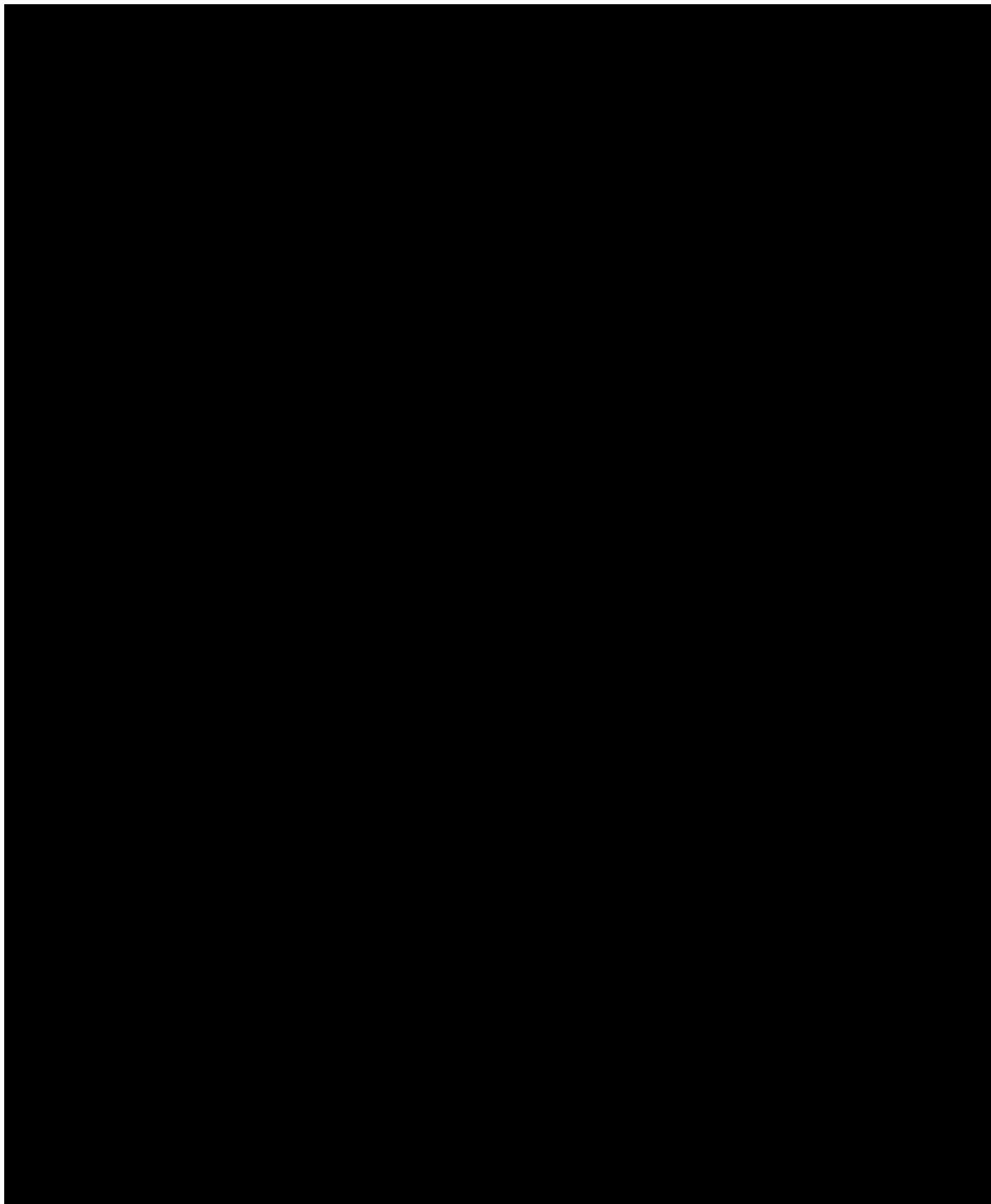


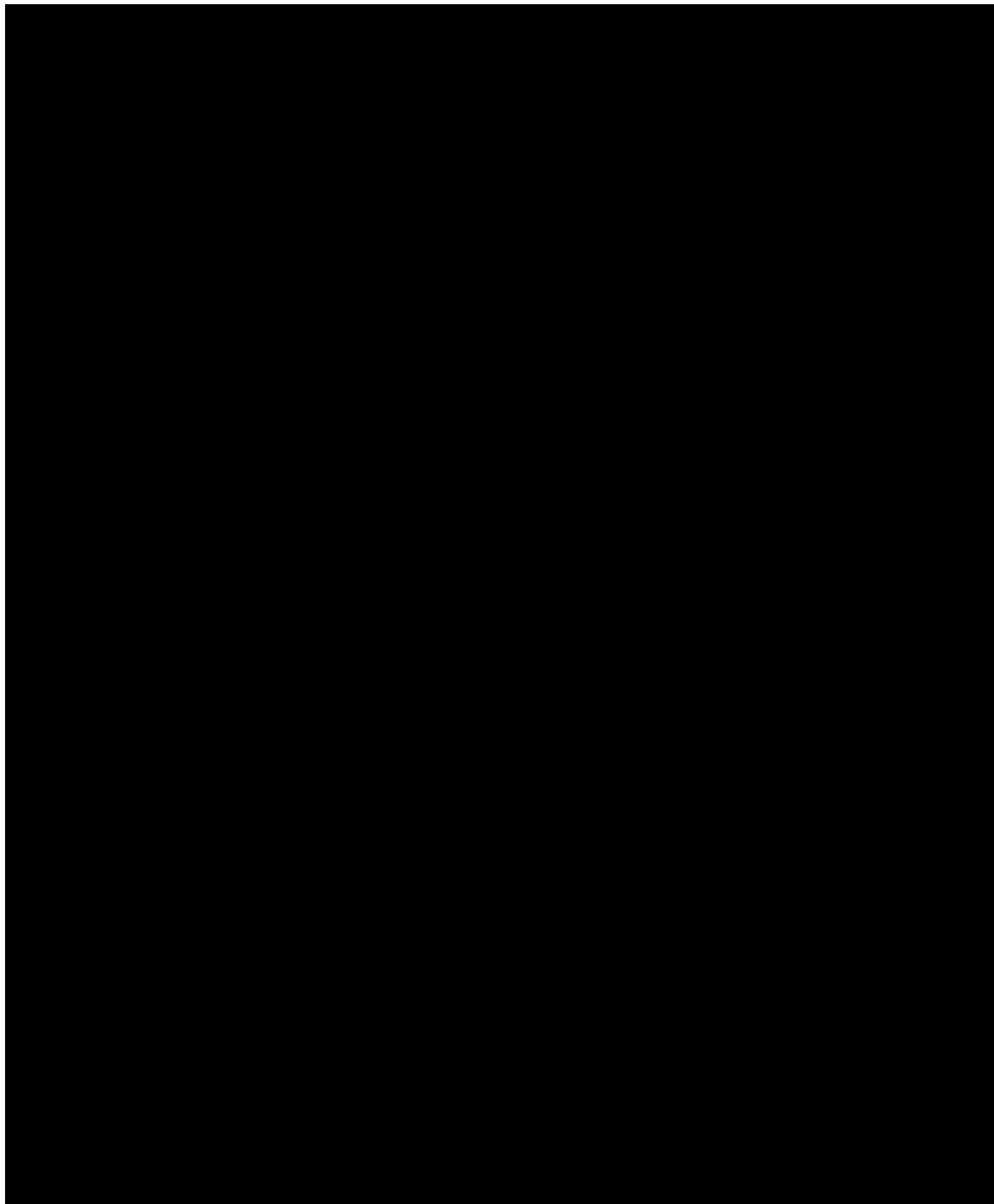


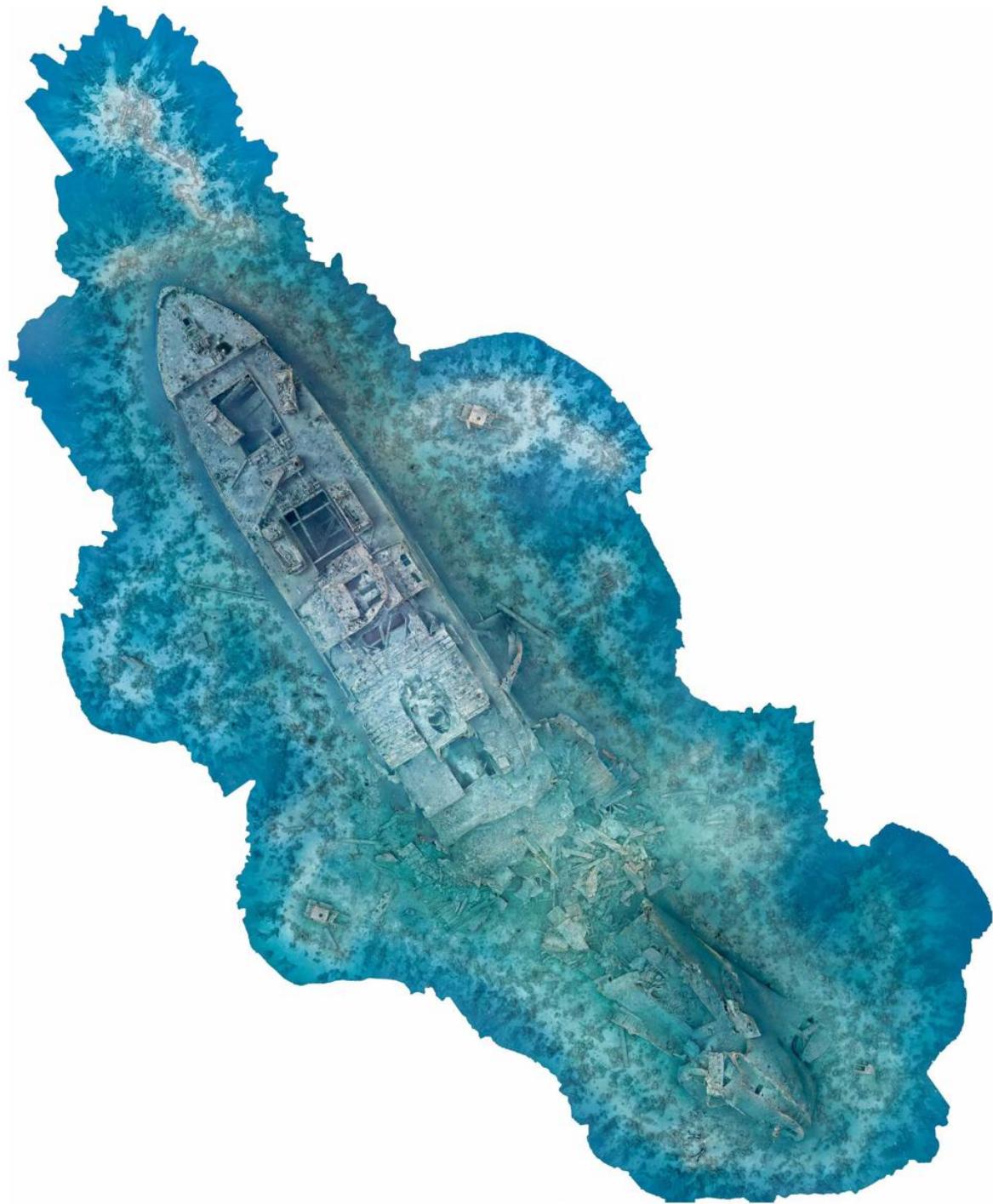
Credit: Parvin Fathi-Hafshejani and Seyed Adib Taba, Auburn University/NanoArtography

Credit: Parvin Fathi-Hafshejani and Seyed Adib Taba, Auburn University/NanoArtography

Nano beach. This tropical scene is just one-quarter of a millimetre tall. It is made from molybdenum disulfide nanostructures, shaped using a laser and imaged with a scanning electron microscope. Artificial colours complete the minuscule masterpiece, which [won third place](#) in the 2021 Nanoartography competition.





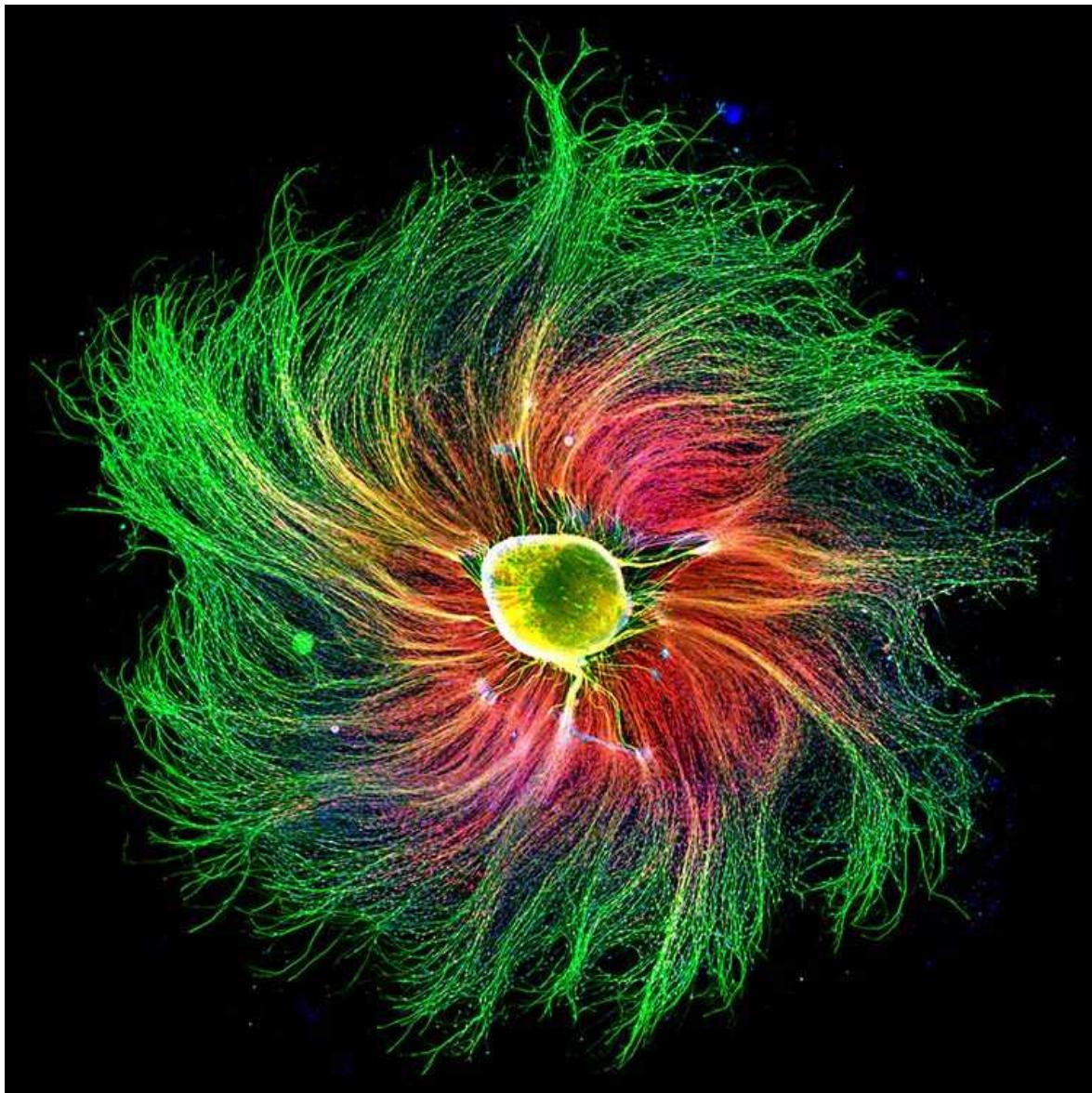


Credit: Simon Brown/RPS Science Photographer of the Year

Credit: Simon Brown/RPS Science Photographer of the Year

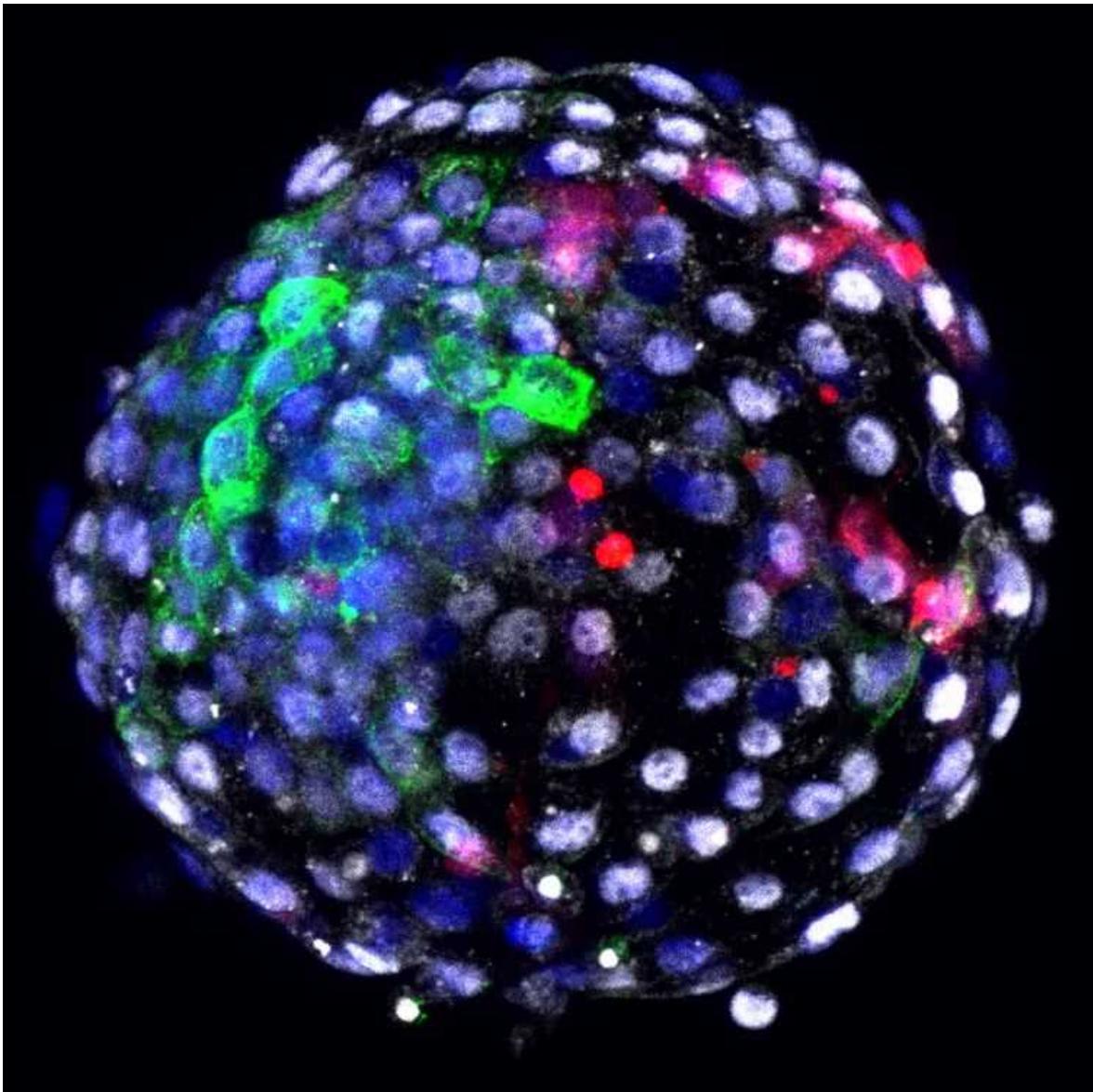
Shipwreck shot. SS *Thistlegorm* was a British Merchant Navy ship that sank in the Red Sea during the Second World War. The wreck is slowly

becoming part of the local coral reef. Photographer Simon Brown [captured this image of the ship](#) in 15,005 frames, each adjusted to give a ‘straight down’ view before being tagged with GPS data and merged with the others.



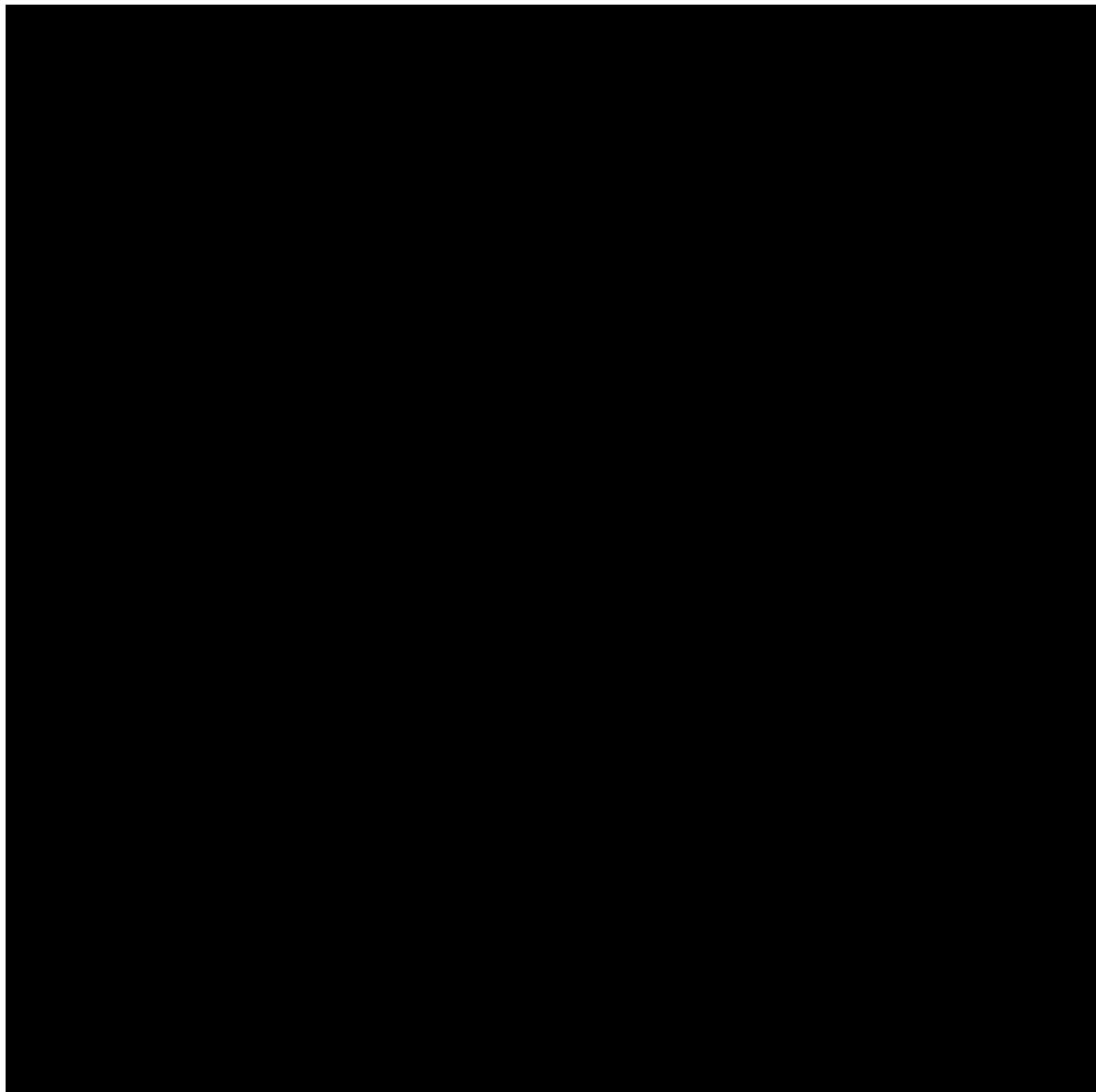
Credit: Paula Diaz, Pontificia Universidad Católica de Chile/Nikon Small World

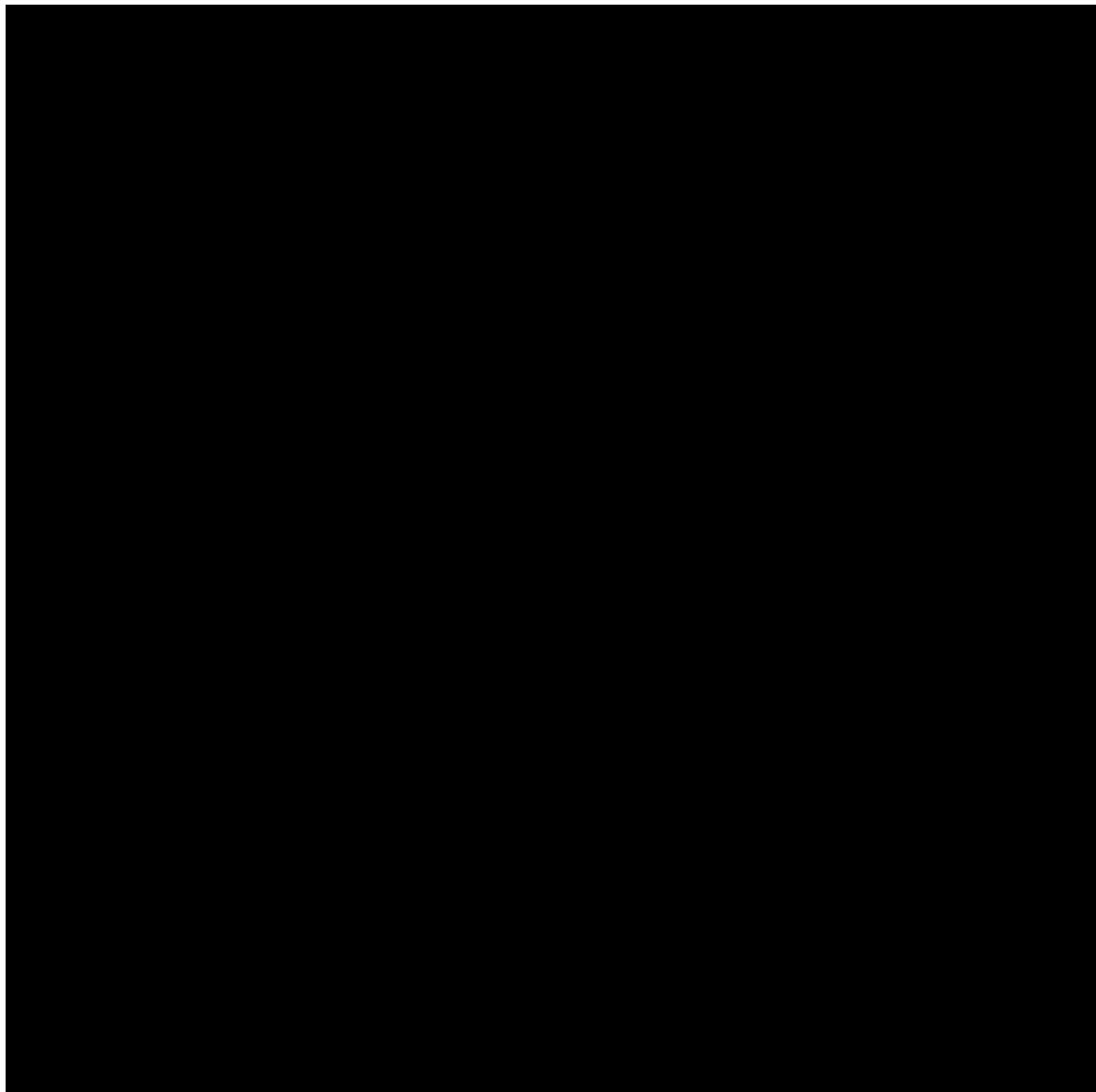
Cell close-up. PhD student Paula Diaz used fluorescence microscopy to capture this [colourful image of a rat embryo's dorsal root ganglion](#), a group of sensory nerve cells that detect pain, among other things. The shot won fourth place in the 2021 Nikon Small World Photomicrography Competition.

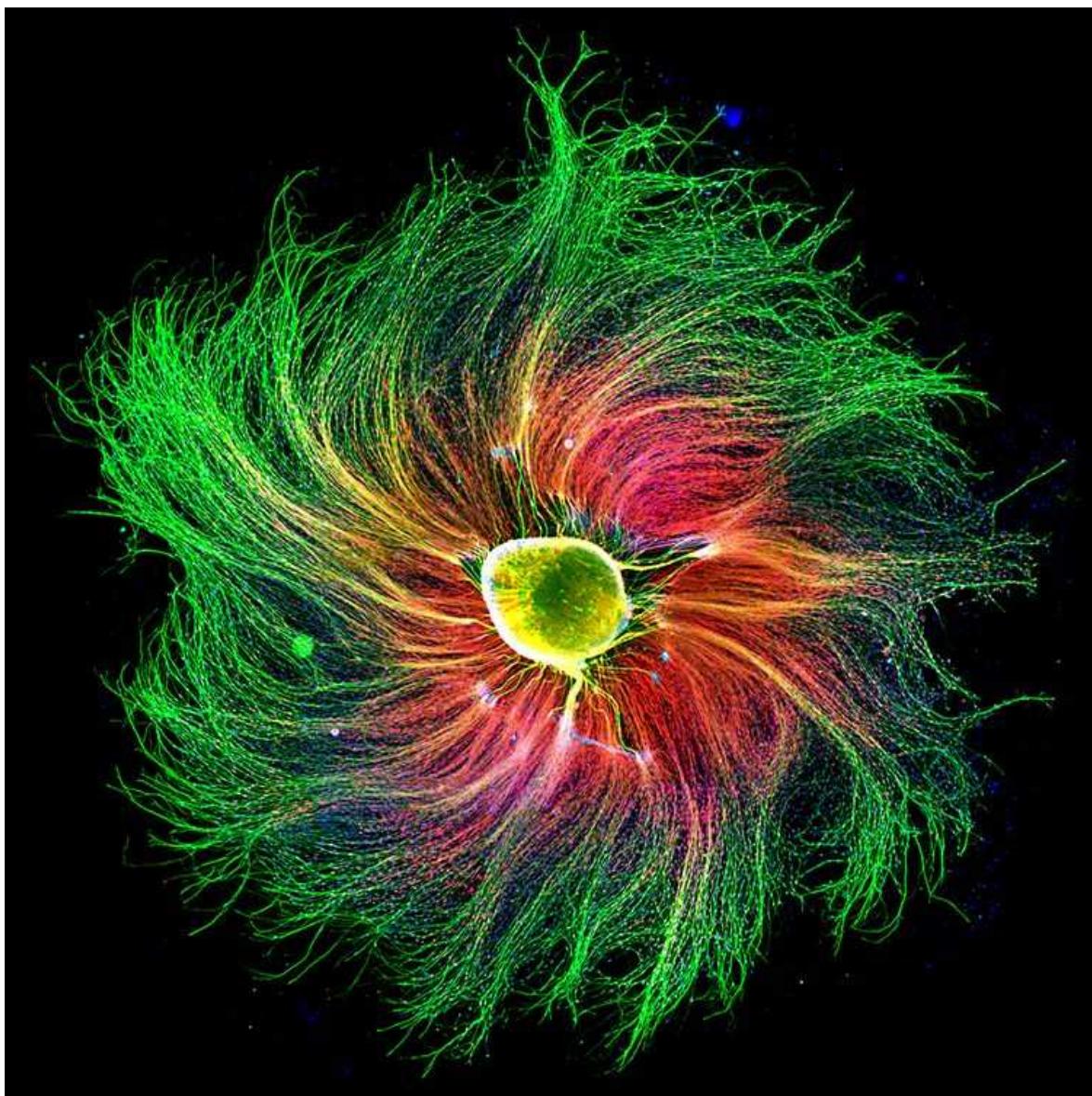


Credit: Weizhi Ji, Kunming University of Science and Technology

Hybrid embryo. In April, scientists reported successfully growing the [first monkey embryos containing human cells](#). Human–animal hybrids might one day provide better models for testing drugs, or be used to grow human organs for transplants. But the field has also raised ethical questions.

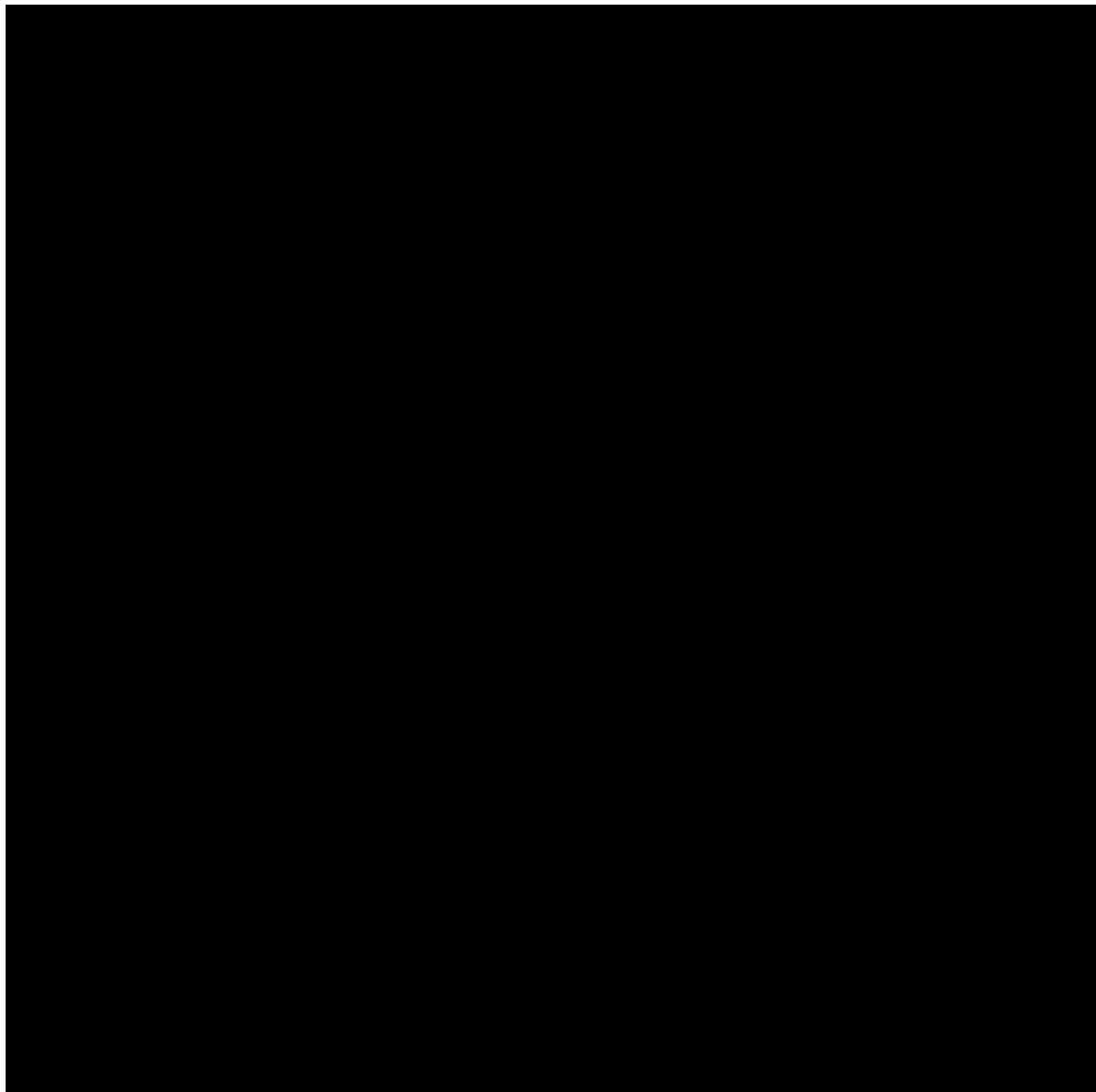


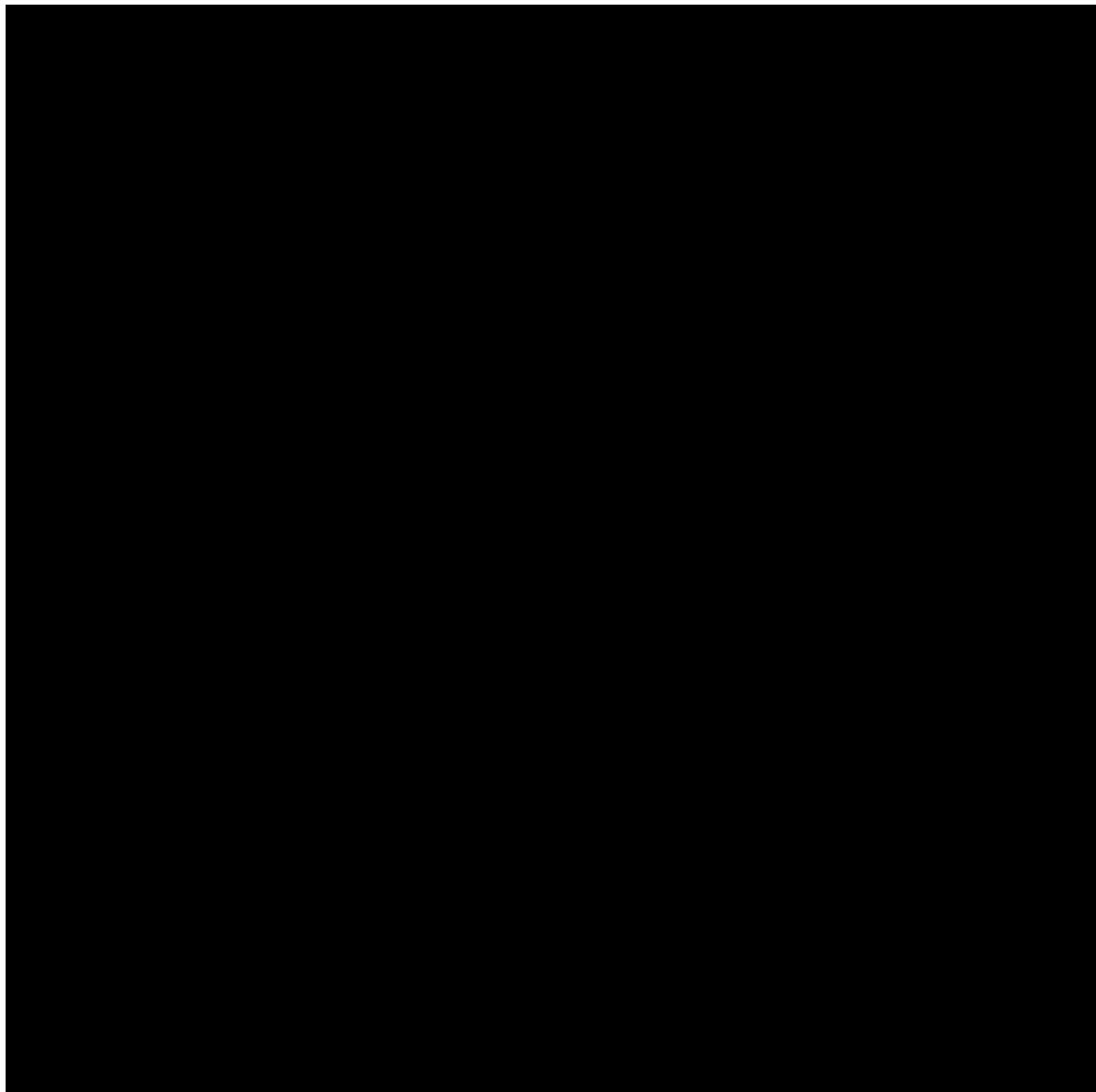


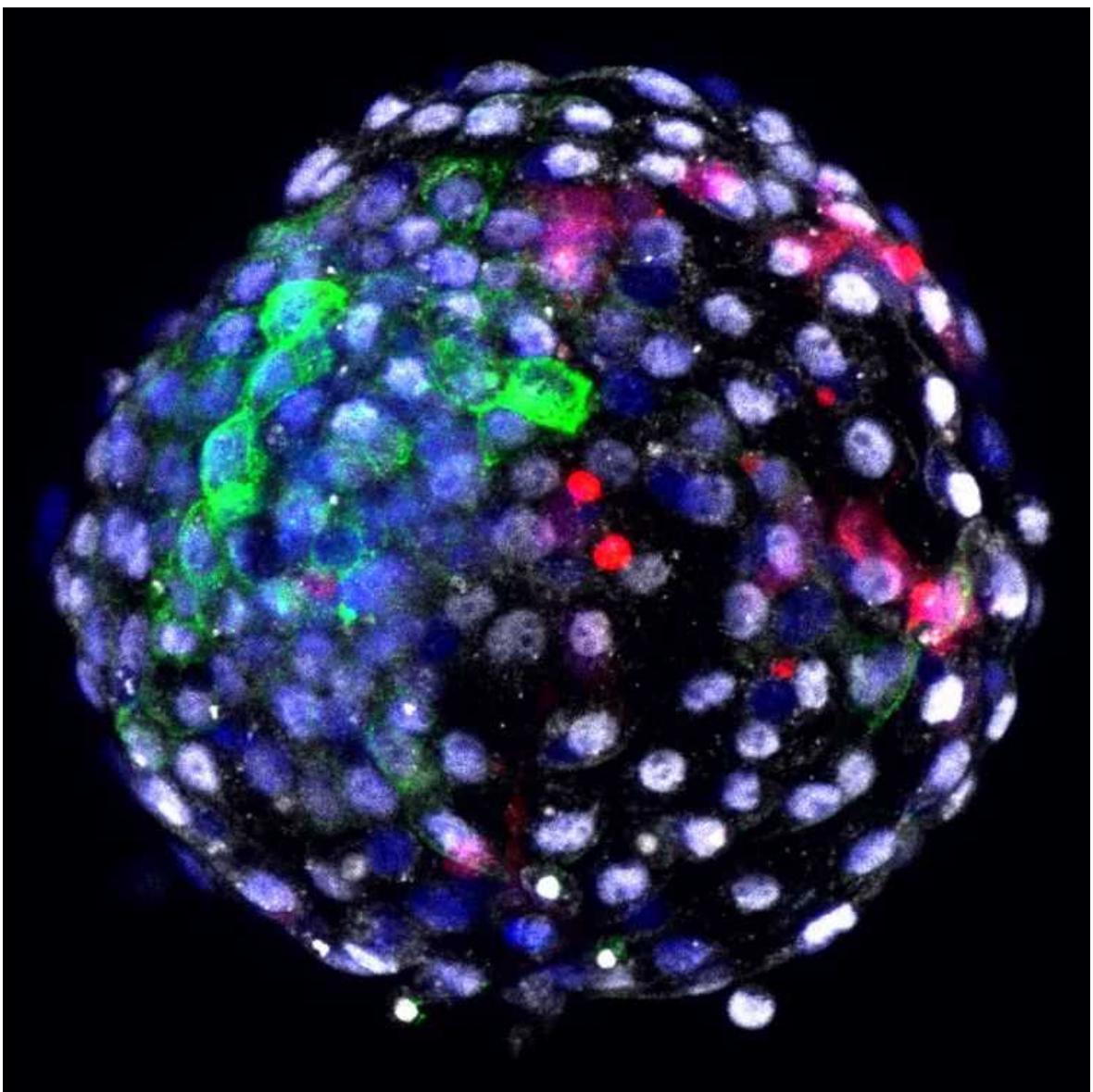


Credit: Paula Diaz, Pontificia Universidad Católica de Chile/Nikon Small World

Credit: Paula Diaz, Pontificia Universidad Católica de Chile/Nikon Small World







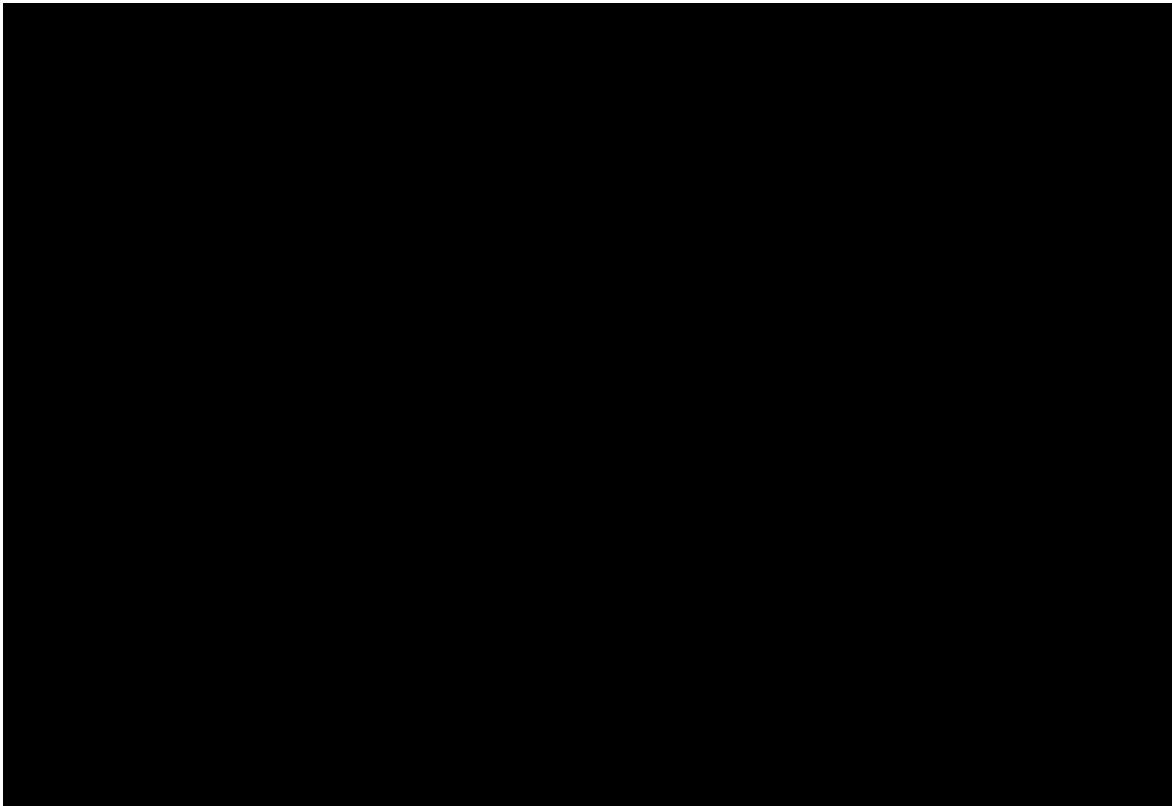
Credit: Weizhi Ji, Kunming University of Science and Technology

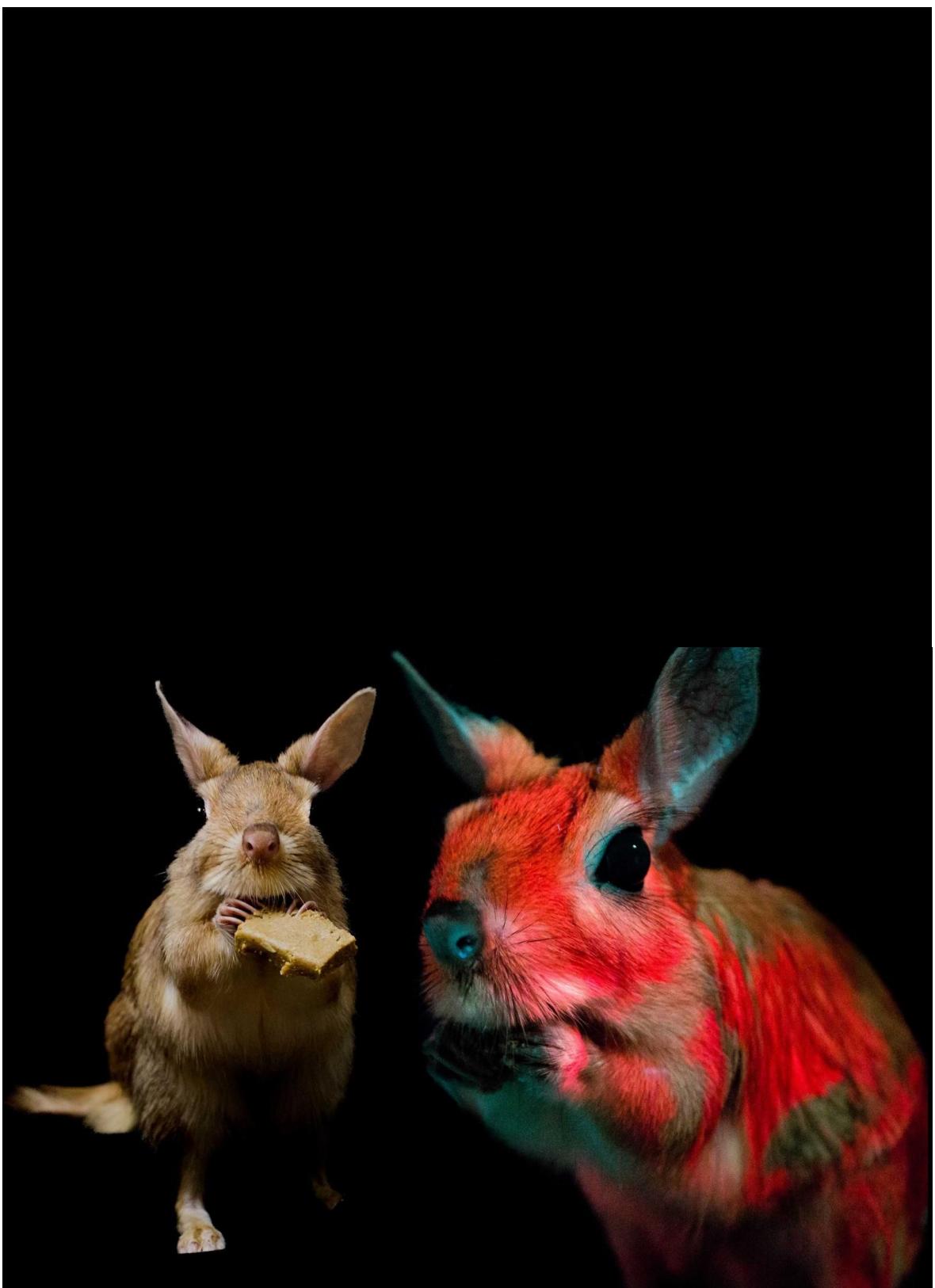
Credit: Weizhi Ji, Kunming University of Science and Technology

This video has no sound. Credit: NASA's Goddard Space Flight Center/Jeremy Schnittman and Brian P. Powell

This video has no sound. Credit: NASA's Goddard Space Flight Center/Jeremy Schnittman and Brian P. Powell

Black holes. This visualization shows how [two supermassive black holes would distort light emanating from the hot gases around them](#). To create it, astrophysicist Jeremy Schnittman at NASA's Goddard Space Flight Center in Greenbelt, Maryland, used a supercomputer to calculate the path taken by light rays through the warped space-time around the simulated black holes.





Credit: J. Martin and E. Olson, Northland College/Olson et al. 2021, Scientific Reports

Credit: J. Martin and E. Olson, Northland College/Olson et al. 2021, Scientific Reports

Raving rodent. The South African springhare (*Pedetes capensis*) [fluoresces hot pink under ultraviolet light](#), researchers reported in February. The function of the striking patterns and intense colour — which are unique among known biofluorescent mammals — is a mystery.



Credit: Simone Tramonte/WaterBear and CIWEM

Credit: Simone Tramonte/WaterBear and CIWEM

Plant life. Inside these stacks of tubes, microalgae exposed to ultraviolet light produce astaxanthin — a red pigment that is extracted and sold as a dietary supplement. The system is part of a [bioreactor facility run by Icelandic company Algalif](#). It runs on geothermal electricity and uses an ultra-energy-efficient lighting system.





Credit: NASA/JPL-Caltech/MSSS

Credit: NASA/JPL-Caltech/MSSS

Martian Ingenuity. This year, missions from China, the United States and the United Arab Emirates arrived at Mars. This photo, [snapped by a camera on NASA's Perseverance rover](#), shows its helicopter Ingenuity being lowered onto the red planet's surface. In April, Ingenuity became the [first machine to achieve powered flight on another planet](#).

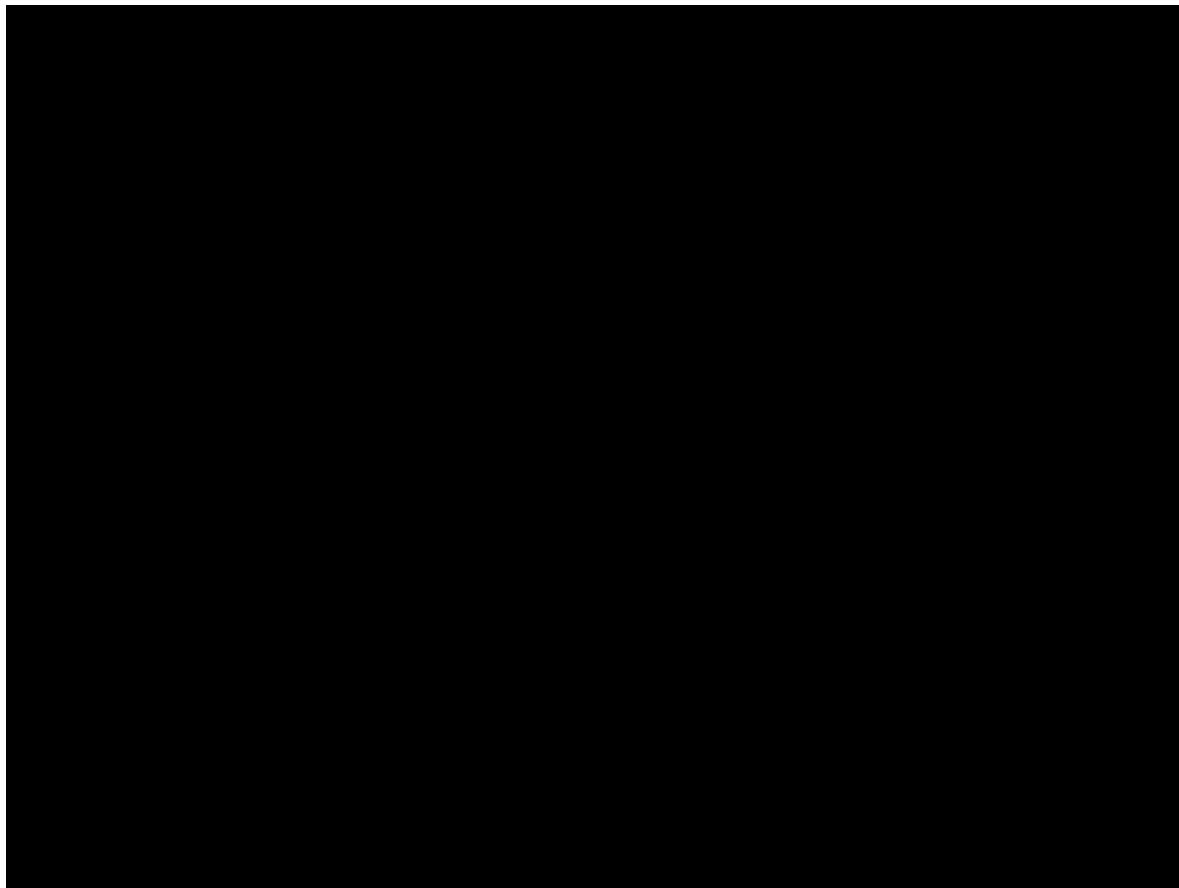


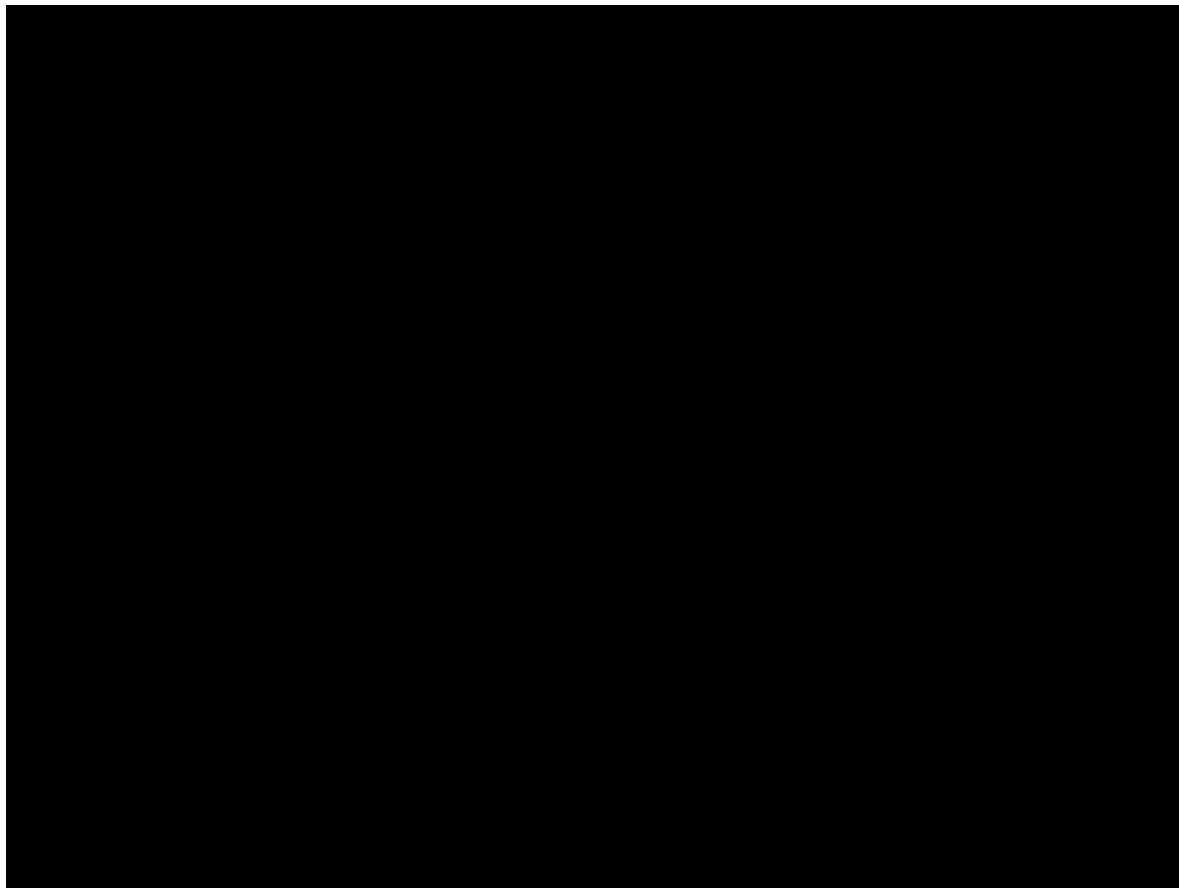


Credit: Max Whittaker/NYT/Redux/eyevine

Credit: Max Whittaker/NYT/Redux/eyevine

Burning up. This year the United States was once again [ravaged by extreme wildfires](#), exacerbated by heatwaves linked to climate change. Here, firefighters in California battle a blaze near the southern shore of Lake Tahoe, where thousands of people were evacuated in late summer.







Credit: Lindsey Swierk

Credit: Lindsey Swierk

Scuba divers. Evolutionary biologists reported that some species of *Anolis* lizard have adapted to re-breathe exhaled air underwater, using a bubble on their snouts. The trick allows them to stay submerged for more than 15 minutes, hiding from predators.





Credit: Levon Biss/Nikon Small World

Credit: Levon Biss/Nikon Small World

Ancient insect. This portrait of a [40-million-year-old gnat in amber](#), by photographer Levon Biss, got an honourable mention in the 2021 Nikon Small World Photomicrography Competition.

A personal view of the news

In compiling this year's collection of striking science images, *Nature*'s media and news editors identified a photograph that said something special to them. Here is their take on the past 12 months.

Flames and fury. Tom Houghton (Media editor). Despite the abundance of incredible photography that has crossed my desk this year, I struggle to think of a moment more moving than the one captured by Bloomberg photographer Konstantinos Tsakalidis on the island of Evia, Greece. This woman's anguish as the raging wildfire draws closer to her house is heartbreakingly powerful. The colour and composition reminds me of the painting *The Scream* by Edvard Munch. No matter how many times I look at it, it remains

both beautiful and upsetting. Credit: Konstantinos Tsakalidis/Bloomberg via Getty

Twin pandas. Jessica Hallett (Associate media editor). Earlier this year, giant panda Huan Huan [gave birth to twin female cubs](#) at Beauval ZooParc in France. It was such a joy to see the safe delivery of these newborns, and even more so to see them in this picture, healthy and thriving at one month old. This story stands out to me as a real glimmer of hope, a welcome break from the doom and gloom of the past couple of years. I can't wait to follow the twins' progress and see even more gorgeous images of them growing up. Credit: Eric Baccega/Nature Picture Library

Goma evacuation. Lizzy Brown (Managing media editor). When the volcano Mount Nyiragongo suddenly started spewing lava towards the crowded city of Goma in the Democratic Republic of the Congo, photojournalist Guerchom Ndebo was there to document the [mass evacuation of the city in the middle of the night](#), under the glow of light from the crater. Scientists fear that further volcanic activity in the area could lead to a '[limnic eruption](#)' from nearby [Lake Kivu](#), which would release huge amounts of toxic gas into the surrounding valley. Luckily, Nyiragongo's eruption in May this year did not trigger a such an event, and life has slowly returned to normal in Goma. Credit: Guerchom Ndebo for Fondation Carmignac

Liquid Rothko. Agnese Abrusci (Media editor). A narrow strip of road divides the ochre-coloured toxic waters of a mine reservoir from fresh green waters near the Spanish city of Huelva. This image — reminiscent of a painting by Mark Rothko — instantly captures the viewer's attention, with its abstract simplicity evoking tranquillity and peace. But this eventually gives way to the realization that, once again, human activity is tainting natural beauty. Credit: Roberto Bueno/WaterBear and CIWEM

Fast fashion. Amelia Hennighausen (US media editor). The fashion industry's carbon footprint is coming under increasing scrutiny. People currently accumulate more garments, use them less often and discard them faster than at any time in history — a trend that is predicted to worsen. This huge dump in Accra, Ghana, is filled with second-hand clothing imported from Australia, Europe and the United States that cannot be resold. People

pick through the remnants while cows scavenge for food. The scene is a reminder that humanity needs to consider all of its wasteful behaviour if it wants to protect Earth. Credit: Andrew Esiebo/Panos Pictures



- [Privacy Policy](#)
- [Use of cookies](#)
- [Legal notice](#)
- [Terms & Conditions](#)
- [Accessibility statement](#)

© 2021 Springer Nature Limited. All rights reserved.

[Top](#)



[Built with Shorthand](#)

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03521-3>

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

Nature's 10

Ten people who helped shape science in 2021

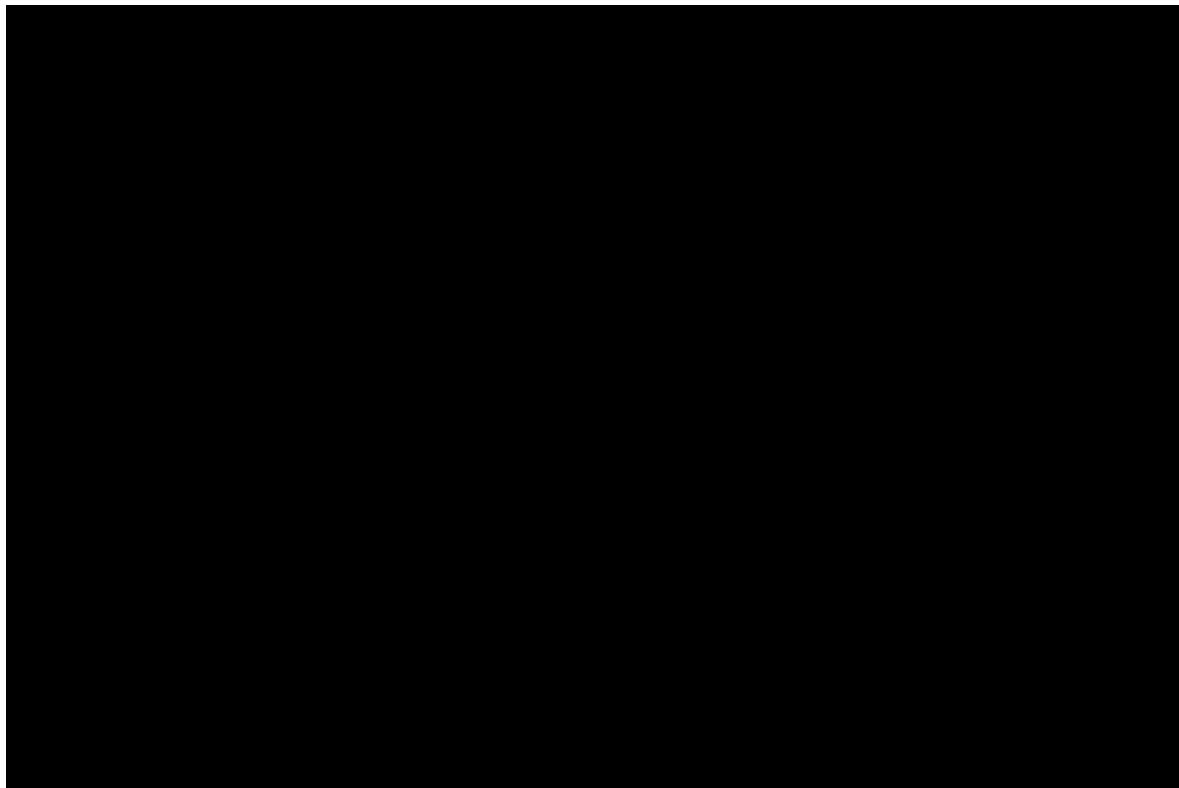
15 DECEMBER 2021

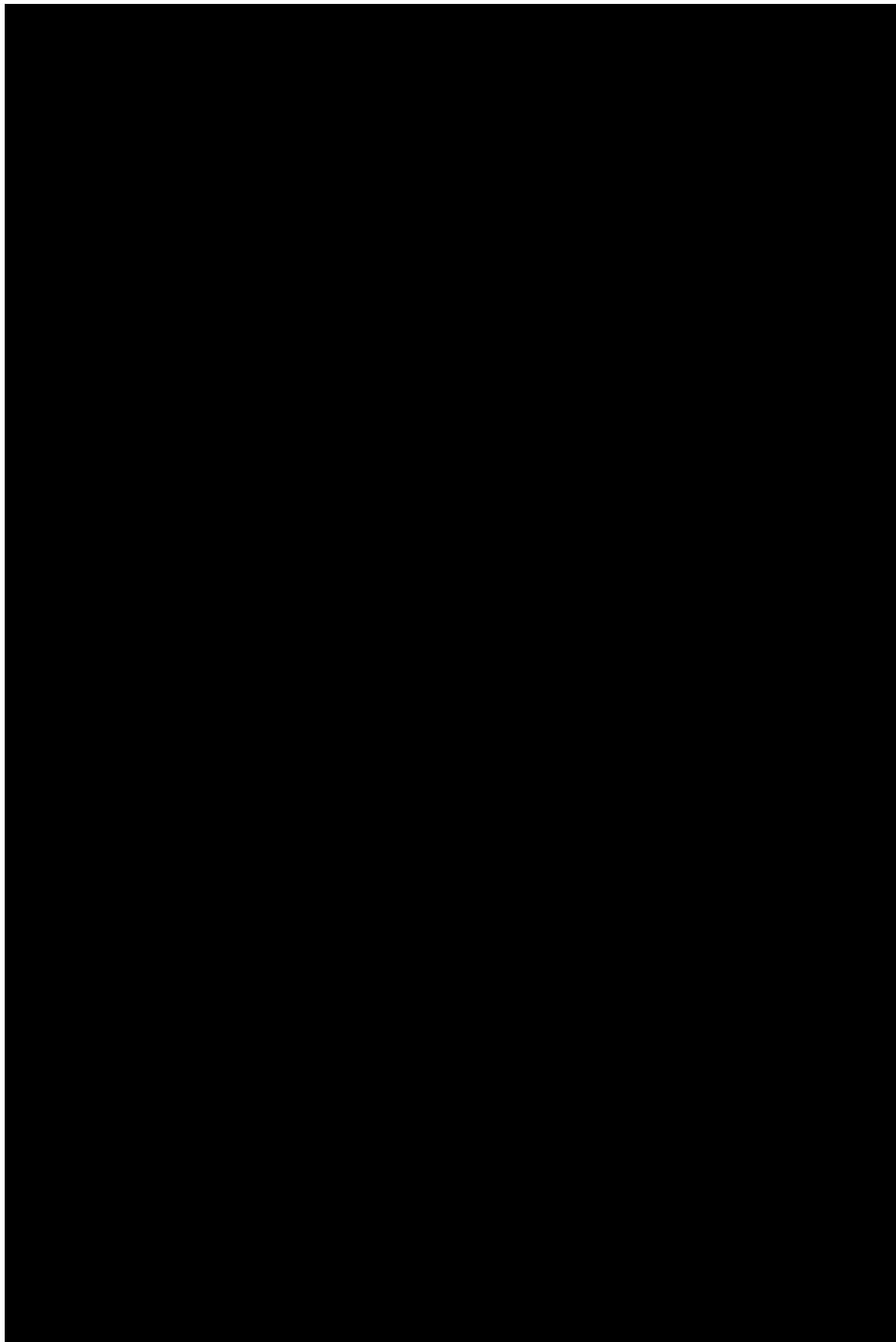
An Omicron investigator, a Mars explorer and an AI ethics pioneer are some of the people behind this year's big research stories.

The [Nature's 10](#) explores key developments in science this year and some of the people who played important parts in these milestones. Along with their colleagues, these individuals helped to make amazing discoveries and brought attention to crucial issues. Nature's 10 is not an award or a ranking. The selection is compiled by Nature's editors to highlight key events in science through the compelling stories of those involved.

Nature's 10

- [**Winnie Byanyima**: Vaccine warrior](#)
- [**Friederike Otto**: Weather detective](#)
- [**Zhang Rongqiao**: Mars explorer](#)
- [**Timnit Gebru**: AI ethics leader](#)
- [**Tulio de Oliveira**: Variant tracker](#)
- [**John Jumper**: Protein predictor](#)
- [**Victoria Tauli-Corpuz**: Indigenous defender](#)
- [**Guillaume Cabanac**: Deception sleuth](#)
- [**Meaghan Kall**: COVID communicator](#)
- [**Janet Woodcock**: Drug chief](#)
- [**Ones to watch in 2022**](#)







Credit: Erich Bartlebaugh/BuzzFeed News/eyevine

Credit: Erich Bartlebaugh/BuzzFeed News/eyevine

Winnie Byanyima: Vaccine warrior

This UN leader knew that vaccine equity wouldn't happen without a fight.

By Amy Maxmen

Before vaccines for COVID-19 even existed, Winnie Byanyima knew that distributing them equitably would be a challenge. In early 2020, she was one of the few voices warning that low- and middle-income countries could not rely on donations alone to vaccinate their people. The only way to get life-saving shots to everyone, she argued, would be by helping as many companies as possible to manufacture them and by setting up systems of distribution to get them where they're needed.

That hasn't happened. Companies that developed coronavirus vaccines, such as Pfizer–BioNTech and Moderna, have held on tightly to the intellectual-property (IP) rights, and wealthy countries have snapped up most doses. Several nations are now distributing booster shots, whereas only about 6% of people in low-income countries have received a single dose. Byanyima expected this disparity because of her experience with the AIDS epidemic. In the early 2000s, life-saving drugs existed but were largely unavailable in her home country of Uganda.

She finds the profit models for some drugs and vaccines infuriating. "This idea that you can sell a life-saving health technology the way you sell a luxury handbag is not normal," says Byanyima, who leads UNAIDS, the United Nations agency heading the effort to end AIDS around the world. "We shouldn't normalize it, we shouldn't respect it and we should call it what it is: immoral, greedy and wrong." Byanyima co-founded the advocacy group the People's Vaccine Alliance to change that way of thinking. Its strategy has been to enlist powerful leaders by presenting them with carefully tailored arguments that highlight how supporting vaccine equity will further their own goals. "Governments are not saints, but they respond to people's demands."

This May, Byanyima and her colleagues celebrated an unexpected victory when the United States — historically a strident patent defender — threw its weight behind a proposal from South Africa and India [to waive the IP protections surrounding COVID-19 vaccines](#) in the hope of bolstering manufacturing capacity.

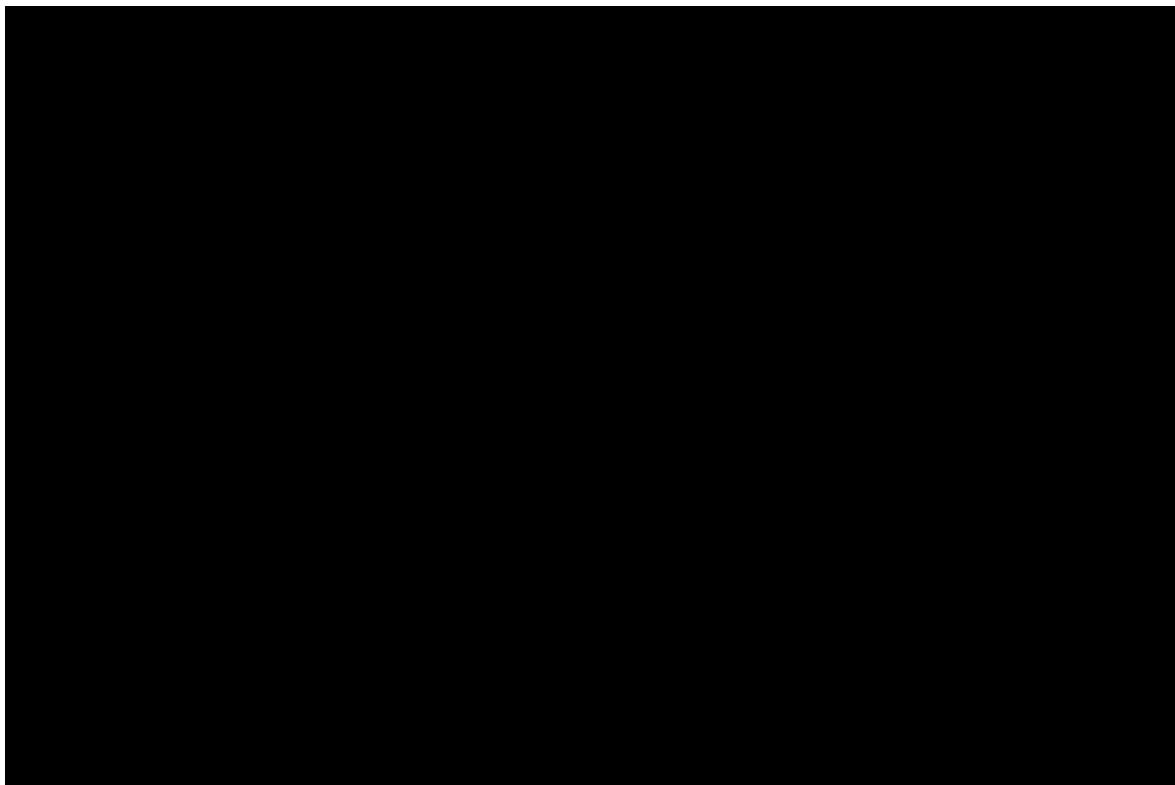
There is still much work to be done. Several countries and the European Union remain opposed to such a waiver, and the companies that own the IP have rejected requests to license their vaccine technologies and share the knowledge needed to produce them. The corporations defend their actions by saying that waivers would not change the situation and would undermine innovation.

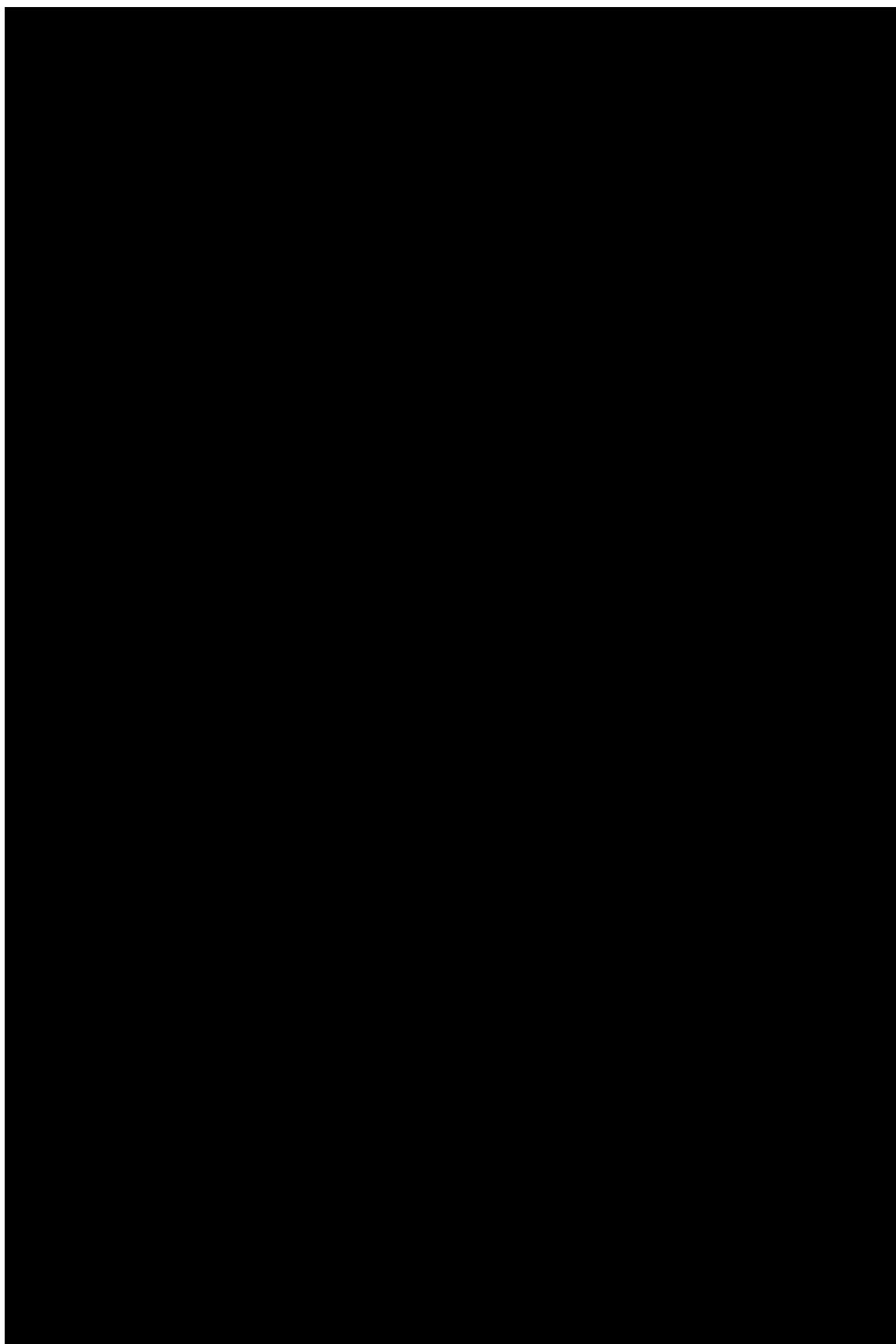
Byanyima disagrees, and her frank statements on this and other matters of inequity have won her many admirers. "Some people will say inequality is bad," says John Nkengasong, director of the Africa Centres for Disease

Control and Prevention. “But when Winnie talks about it, you feel there is a fire in her that comes from within.”

Byanyima says it was this thirst for justice that caused her to leave her career in aeronautical engineering soon after the overthrow of Uganda’s former authoritarian president Idi Amin. In 1981, she joined a guerilla movement fighting to restore democracy and human rights to Uganda. They prevailed, and by 1994, she was elected to Uganda’s parliament. She was appointed head of UNAIDS in 2019, where she is putting equity at the centre of the programme’s work around the world. Global-health-policy researcher Matthew Kavanagh took leave from a position at Georgetown University in Washington DC, to work for Byanyima because of the way she targets underlying inequalities that foster the spread of HIV. The same goes for COVID-19: “Winnie drove the conversation on vaccine equity, starting way before vaccines existed, and others have raced to catch up,” Kavanagh says.

Byanyima is working to ensure that the fruits of science change lives. “Without political decisions to reduce inequality,” she says, “we can’t get anywhere.”







Credit: Joakim Stahl/SvD/TT/Alamy

Credit: Joakim Stahl/SvD/TT/Alamy

Friederike Otto: Weather detective

As heatwaves, floods and droughts multiply, this researcher assesses whether humans bear some blame.

By Quirin Schiermeier

Friederike Otto has spent the past seven years studying extreme weather events, but even she was shocked when an extraordinary heatwave hit Canada and the US Pacific Northwest in July, shattering temperature records and killing hundreds of people.

Whenever extreme weather strikes these days, people immediately wonder whether climate change is to blame. This is exactly the kind of question that Otto and her collaborators in the [World Weather Attribution \(WWA\) group](#)

[try to answer quickly](#). Otto set up a video call with the WWA team and they planned a speedy research study. The team pored over meteorological data to gauge how big the heatwave was, studied climate records for the region and ran computer models to find out how much more likely this kind of heatwave has become, relative to a hypothetical world without climate change. The result: it would have been all but impossible for a heatwave of that magnitude to have happened in the region without human-induced climate change.

“Temperature records were broken by 5 °C in some places,” Otto says. “That’s immense.”

Otto, a climate researcher at the Grantham Institute for Climate Change and the Environment in London, helped to set up the WWA in 2015 with the aim of rapidly analysing whether climate change plays a part in extreme heat, cold, downpours, drought and wildfire activity. She chairs the ad hoc group, which includes about a dozen climate modellers and statisticians.

Aside from the American Northwest heatwave this year, she and the group analysed the role of climate change in the devastating floods in July in Germany and Belgium, an April ‘cold wave’ in France, and the persistent drought in Madagascar.

Otto earnt her PhD in the philosophy of science before turning to physics, and eventually to climate science. Like many researchers, she is sincerely worried about the impacts of climate change. “I’m into justice,” she says. “And climate change is one of the biggest threats to justice.”

Until a few years ago, scientists would have been hard pressed to answer with certainty whether climate change is to blame for specific extremes, and how much more (or less) likely they have become. Many scientists viewed attribution studies critically when the WWA made its first attempts to analyse extreme events — using just one or two climate models without evaluating whether these were able to reliably simulate the extreme in question.

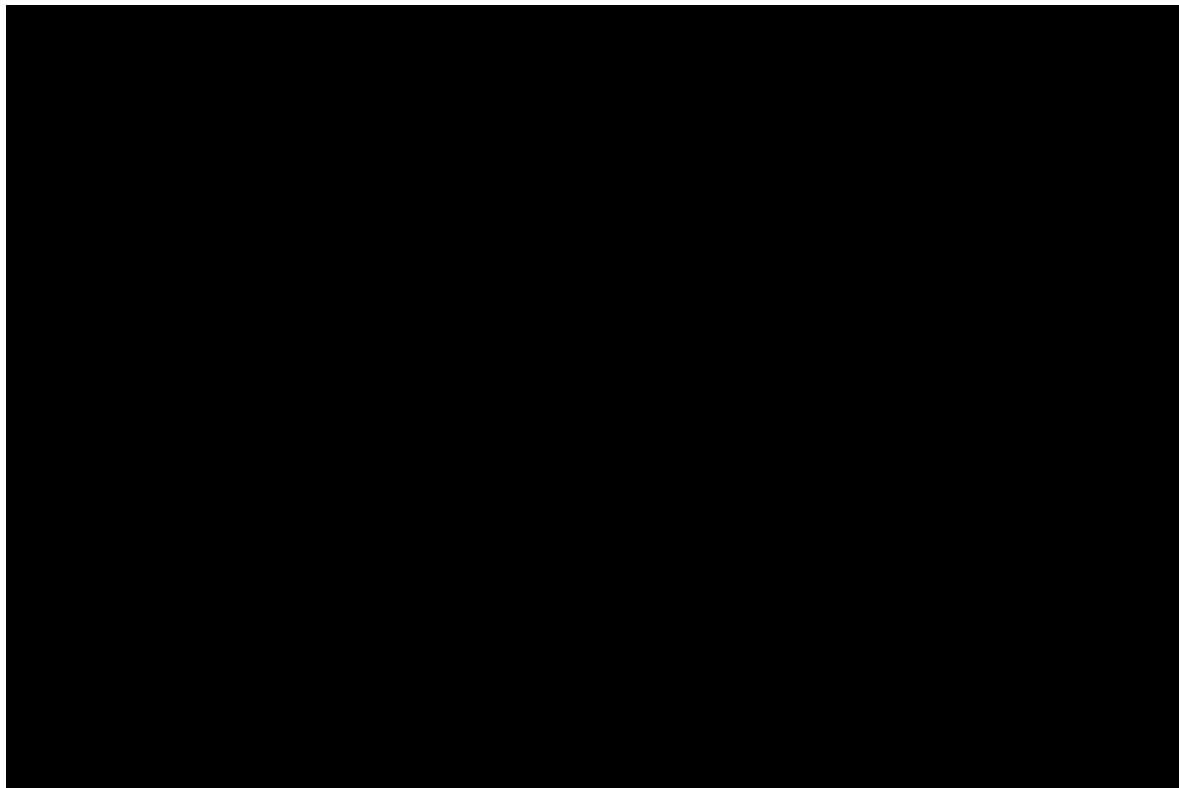
This has changed entirely. Otto and her team — including her former co-chair, the Dutch climate modeller Geert Jan van Oldenborgh, who died this

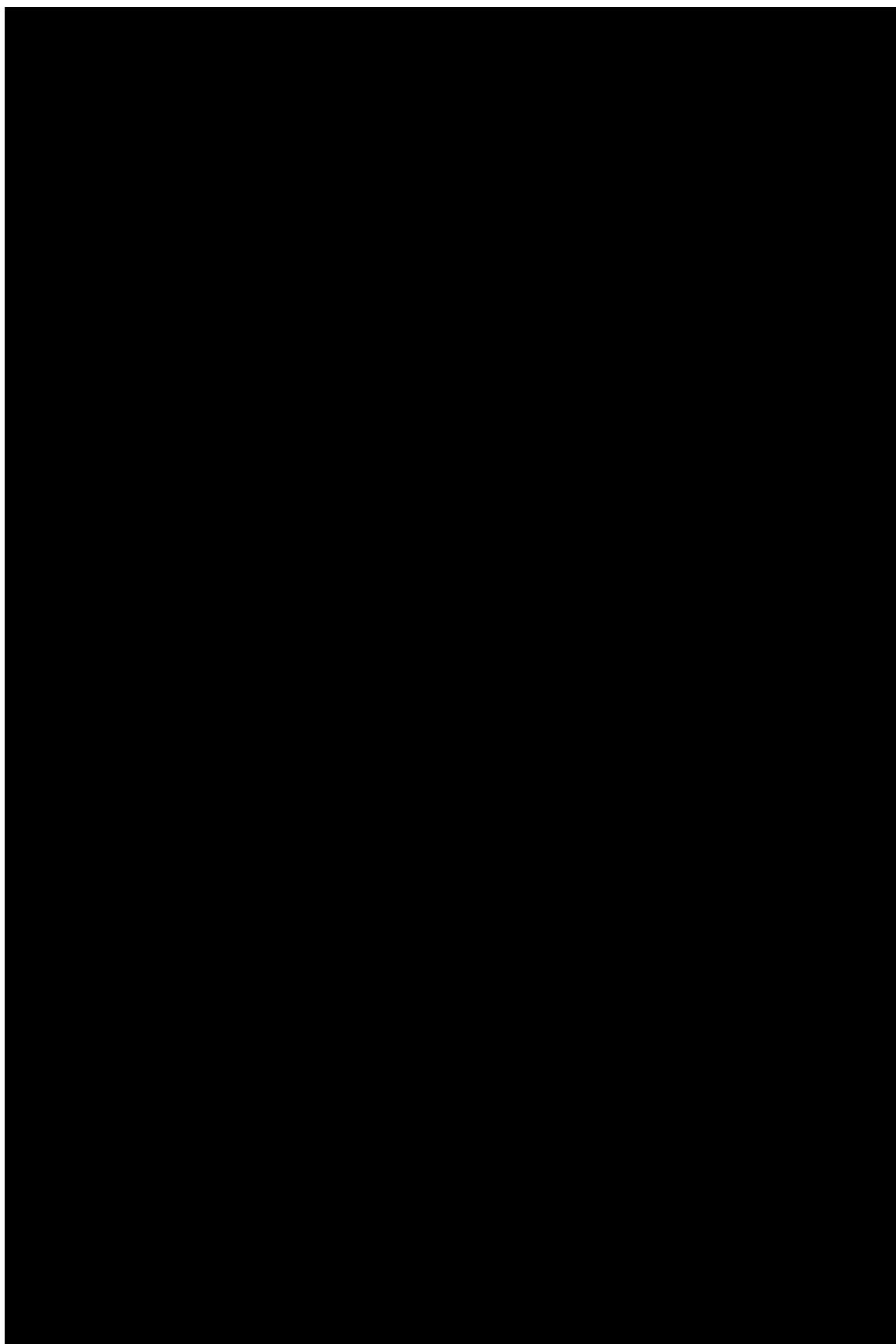
year after a long illness — have developed a strategy that uses climate simulations from as many as 50 models. This approach and the studies generated are now widely viewed as highly robust; they feature prominently in a report issued in August by the Intergovernmental Panel on Climate Change, a scientific committee established by the United Nations. It is now an “established fact”, the panel said, that rising greenhouse-gas emissions have made some weather extremes — in particular, extreme heat — [more frequent and more intense](#). Its report came out shortly before the [26th Conference of the Parties \(COP26\) to the global climate-change treaty](#), held last month in Glasgow, UK.

Attribution studies are more difficult in the global south, [where reliable climate data are often lacking](#), and where local research capacities are limited. But these are some of the places most at risk from climate change and the extreme weather that it can spark. Otto hopes that lower-income countries will be able to strengthen their research in these areas in coming years, with support from wealthier countries.

“Attribution studies are really essential in terms of understanding human impacts of climate change,” says Emily Boyd, a social scientist at Lund University in Sweden who studies climate adaptation and governance. “The science is shifting our mindsets — it allows us to think about the relation between climate and vulnerability in a completely new way.”

Together with Boyd and legal scholars, Otto will study how vulnerable groups and countries might be able to capitalize on attribution studies. “The science”, says Boyd, “has every potential to drive government action and promote climate justice.”







Credit: Sun Zifa/CNSphoto

Credit: Sun Zifa/CNSphoto

Zhang Rongqiao: Mars explorer

This engineer leads China's first successful Mars mission, which reached the planet this year and landed a rover on its surface.

By Smriti Mallapaty

On 15 May, Zhang Rongqiao wiped tears from his eyes as [China's Mars rover landed safely on the planet's sandy, auburn plains](#). "I was so overwhelmed," says Zhang, who coordinated the mission.

The touchdown marked the conclusion of a 475-million-kilometre journey full of peril for Zhang and the China National Space Administration, which had never before sent a successful mission to Mars.

The landing, says Zhang, gave him a taste of the old Chinese saying — it takes ten years to sharpen a good sword. China is only the second nation, after the United States, to place a rover on Mars, which is notorious for crushing the hopes of space agencies; nearly half of all missions to the planet have ended in failure.

China's team faced many unknowns in what Zhang calls “such a strange and complex environment”. As chief designer, he is responsible for coordinating a team of tens of thousands who built and operate the Mars mission, named Tianwen-1. The project consists of an orbiter, a lander and the rover, called Zhurong. “The buck stops with him,” says David Flannery, an astrobiologist at Queensland University of Technology in Brisbane, Australia.

The mission was one of three to arrive at Mars in 2021 — the others were [NASA's Perseverance rover](#) and an [orbiter delivered by the United Arab Emirates](#). The success of China's mission has made a national hero of Zhang, who has appeared numerous times on state media, but rarely talks to the press outside China. He responded to *Nature*'s questions by e-mail.

Science took a back seat to Tianwen-1's primary goal, which was to develop and demonstrate China's prowess in deep-space missions that travel beyond the Moon. But Zhang says that getting rich and high-quality information from Mars was a key consideration of the design. And researchers say that the data generated by the rover's six scientific instruments, and another seven on the orbiter, will contribute to a better understanding of a previously unexplored patch of the planet.

Born in 1966 in the town of Anling, eastern China, Zhang studied engineering at Xidian University in Xi'an. He later completed a master's degree at the Chinese Academy of Space Technology in Beijing, and has worked on Earth-observation satellites.

Lu Pan, a planetary scientist at the University of Copenhagen, says that Zhang probably played a key part in the CNSA's decision to send an orbiter, lander and rover to Mars in one shipment — making China the first country to do so. Researchers also say that Zhang considered their input on the choice of instruments and landing site, which will help to ensure that the mission generates as much research as possible.

“He encouraged scientists to participate in the mission to get more scientific output,” says Wenzhe Fa, a planetary scientist at Peking University, Beijing, who is analysing radar data from the Mars mission.

Launched on 23 July 2020, Tianwen-1 arrived at the red planet in February, and dropped the lander and rover in May. The spacecraft settled on a vast impact crater named Utopia Planitia — selected mainly because it is flat and a relatively safe place to land. Since then, the rover has travelled more than 1,200 metres south, taking panoramic images as well as selfies that have been widely shared online.

In mid-September, Zhurong went into hibernation because the Sun got in the way of communications between Mars and the Earth, but it returned to work in late October. It is now heading towards a region that might once have been the coastline of an ancient ocean, where researchers will search for clues about the evolution of Mars.

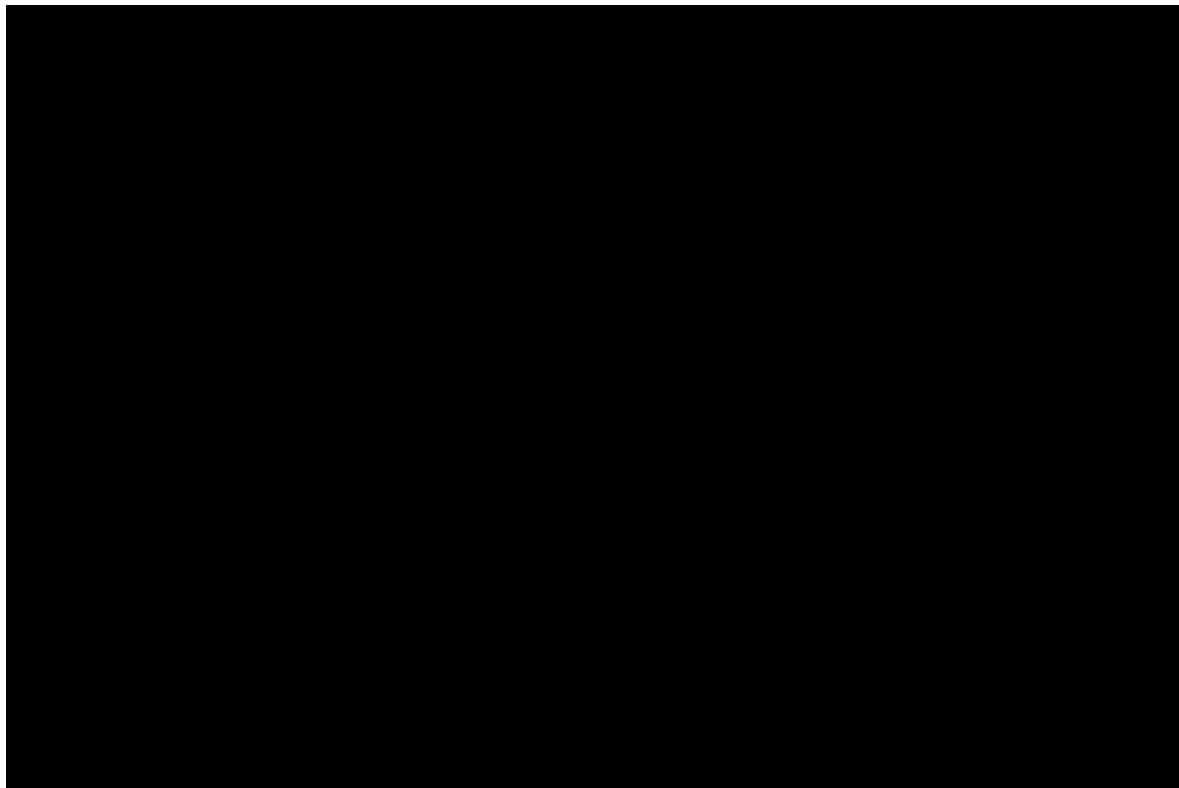
The mission has produced limited science so far, but data collected by some instruments on the rover and orbiter have been shared with more than two dozen teams across the Chinese mainland, Hong Kong and Macau, says Fa, and results are seeping out. They expect to learn insights about the geology of the Utopia Planitia region and the fate of water on the planet.

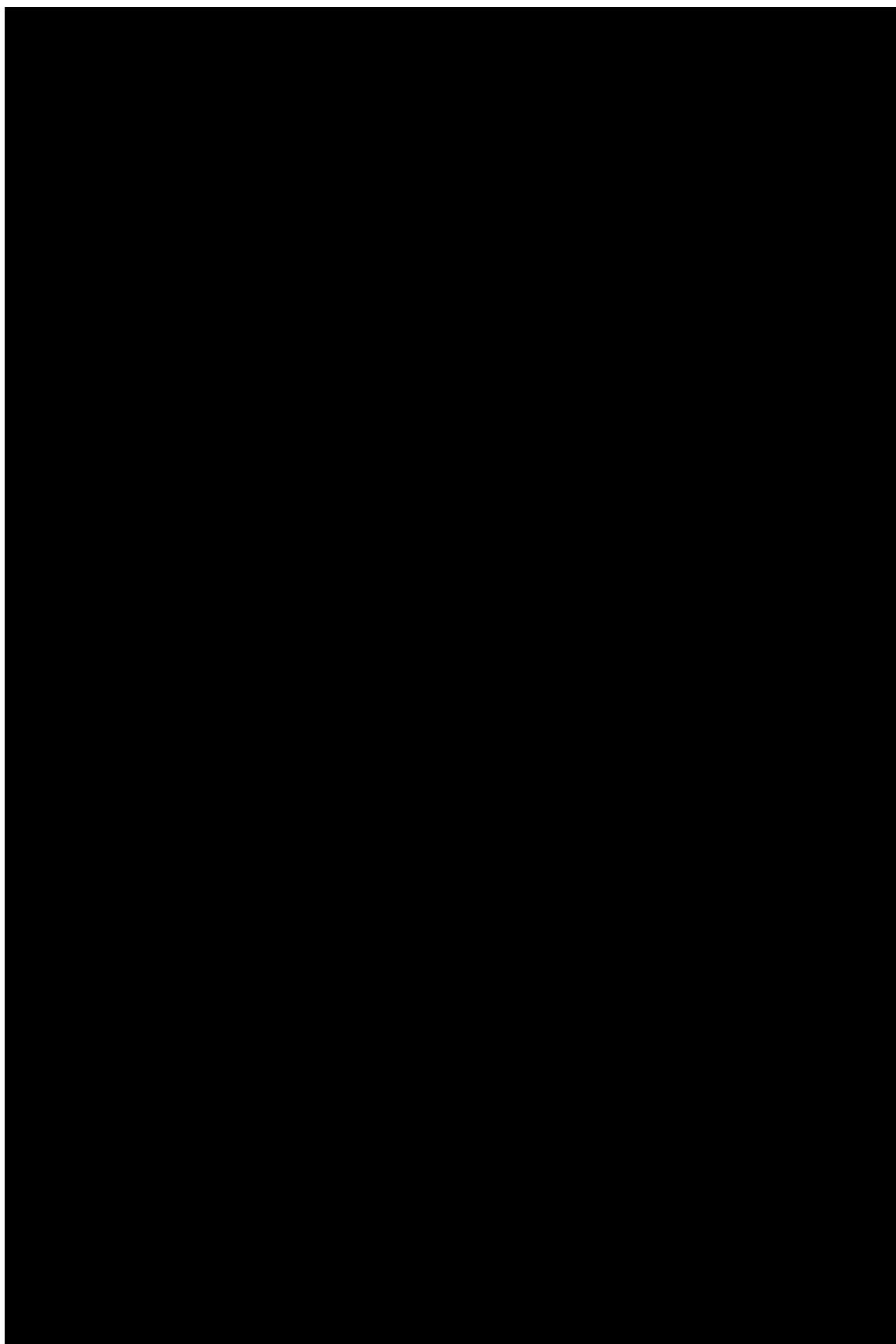
For China’s deep-space missions to take a big leap scientifically, the country will need to refocus towards advancing research rather than chiefly demonstrating engineering. That switch has already happened with China’s lunar missions, says Pan. “These processes take time.”

The real research riches for China, says Flannery, will come later — with the next round of planetary missions. China plans to launch sample-return missions to the asteroid Kamo’oalewa in 2024, and to Mars before 2030. And it has its sights set on Jupiter, too.

Tianwen-1 has also given China’s nascent field of planetary science a boost, say researchers.

“A new generation of scientists is being created right now with this mission,” says Flannery.







Credit: Djeneba Aduayom

Credit: Djeneba Aduayom

Timnit Gebru: AI ethics leader

After losing her job at Google, an artificial-intelligence pioneer founded an independent institute to raise questions about ethics in technology.

By Holly Else

Timnit Gebru, a researcher who studies the ethics of artificial intelligence (AI), says her past year has been — in a word — horrible. Last December, she lost her job at Google after a row over the tech giant's vetting of her work. The highly publicized ousting shocked scientists, including some in the firm, and thousands of researchers rallied to support her, amplifying her concerns around anti-Black discrimination in AI, and around the harms that the technology can cause to marginalized groups in society.

Now, Gebru has forged her own path. On 2 December this year, exactly 12 months after her split with Google, she launched a research institute to study AI independently of big tech companies. The events of the past year, she says, reflect a growing realization that the faults of AI should not be framed as technical problems: they are a symptom of the flawed environment in which the technology is developed.

Born in Ethiopia to parents from Eritrea, Gebru fled the region during a time of war as a teenager and eventually arrived in the United States as a refugee. During her PhD at Stanford University in California, she co-founded a ‘Black in AI’ group with computer scientist Rediet Abebe. And while working at Microsoft, she and computer scientist Joy Buolamwini reported that [facial-recognition software performed less well at identifying the gender of people who were not white men](#) — a finding that drew more attention to bias in AI.

Gebru joined Google in 2018, where she co-led the firm’s ethical AI team with Margaret Mitchell. The pair had a reputation for creating a supportive environment for Black and brown researchers at the firm, where 1.6% of researchers (now 1.8%) were Black women. Their team studied the potential harms of AI, helped Google product teams to think through societal risks of their technologies and supported workforce diversity and inclusion.

But in late 2020, a dispute flared about a paper Gebru had authored with Mitchell and external academics. It [critiqued the environmental impacts and potential biases of large language models](#) — AI software that generates fluent prose and that Google uses in search engines. Gebru was told that internal reviewers at the firm wanted her to withdraw the paper from a conference that she’d submitted it to, or to remove Google-affiliated authors; when she asked for details about who had suggested this — and e-mailed colleagues saying that Google was silencing marginalized voices — she found herself without a job. Google says it accepted her resignation; Gebru says that she was fired, as she had only threatened to resign. (Google didn’t respond to *Nature*’s request for comment for this article.)

Gebru tweeted about the split, and the company faced a storm of protest. Almost 7,000 researchers and engineers, including more than 2,600 from Google, signed a petition in December 2020 calling for an overhaul of the

company's research integrity. In February, the firm fired Gebru's colleague Mitchell after she searched for incidents of discrimination against Gebru in her company e-mail. Two other Google employees left the company outraged at Gebru's treatment, and several groups that support minority researchers in AI ended sponsorship agreements with Google.

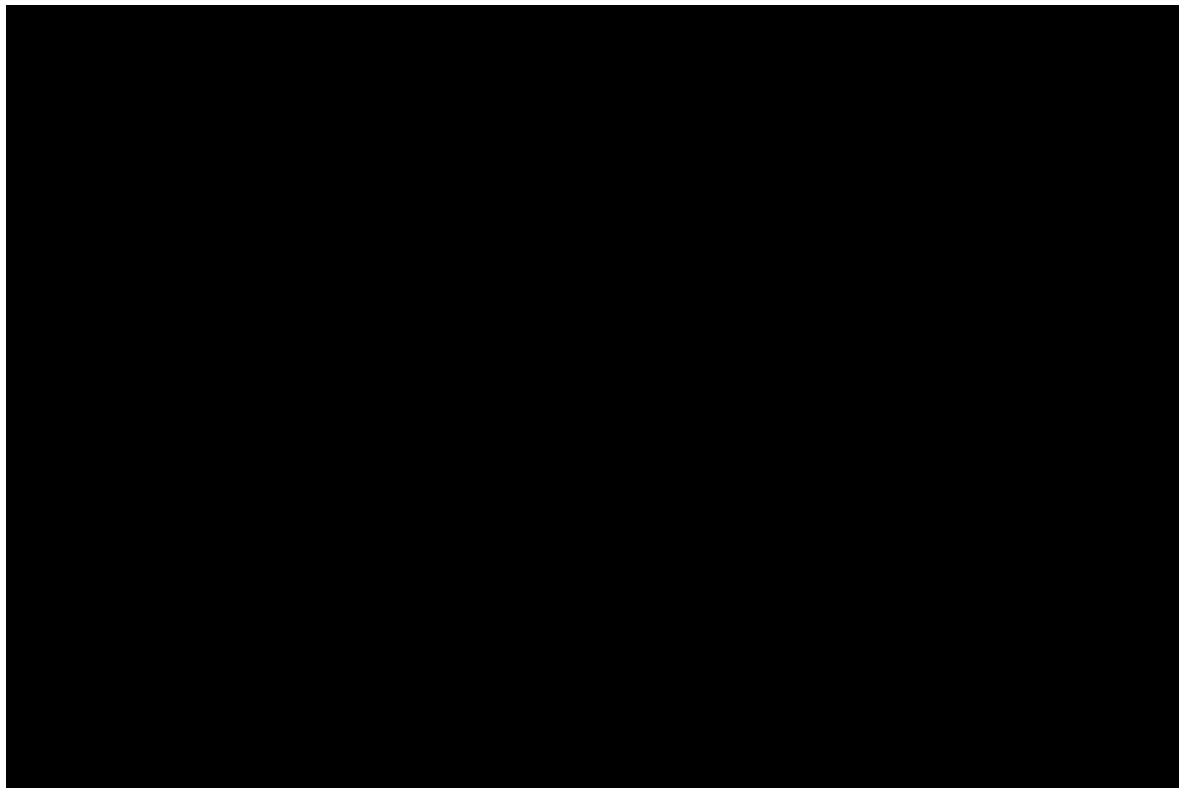
Gebru says that what happened to her was a display of disrespect to her and her work that amounts to misogynoir — anti-Black sexism. "They would never do what they did to me to someone else," she says. "Google had a problem with me speaking up about discrimination."

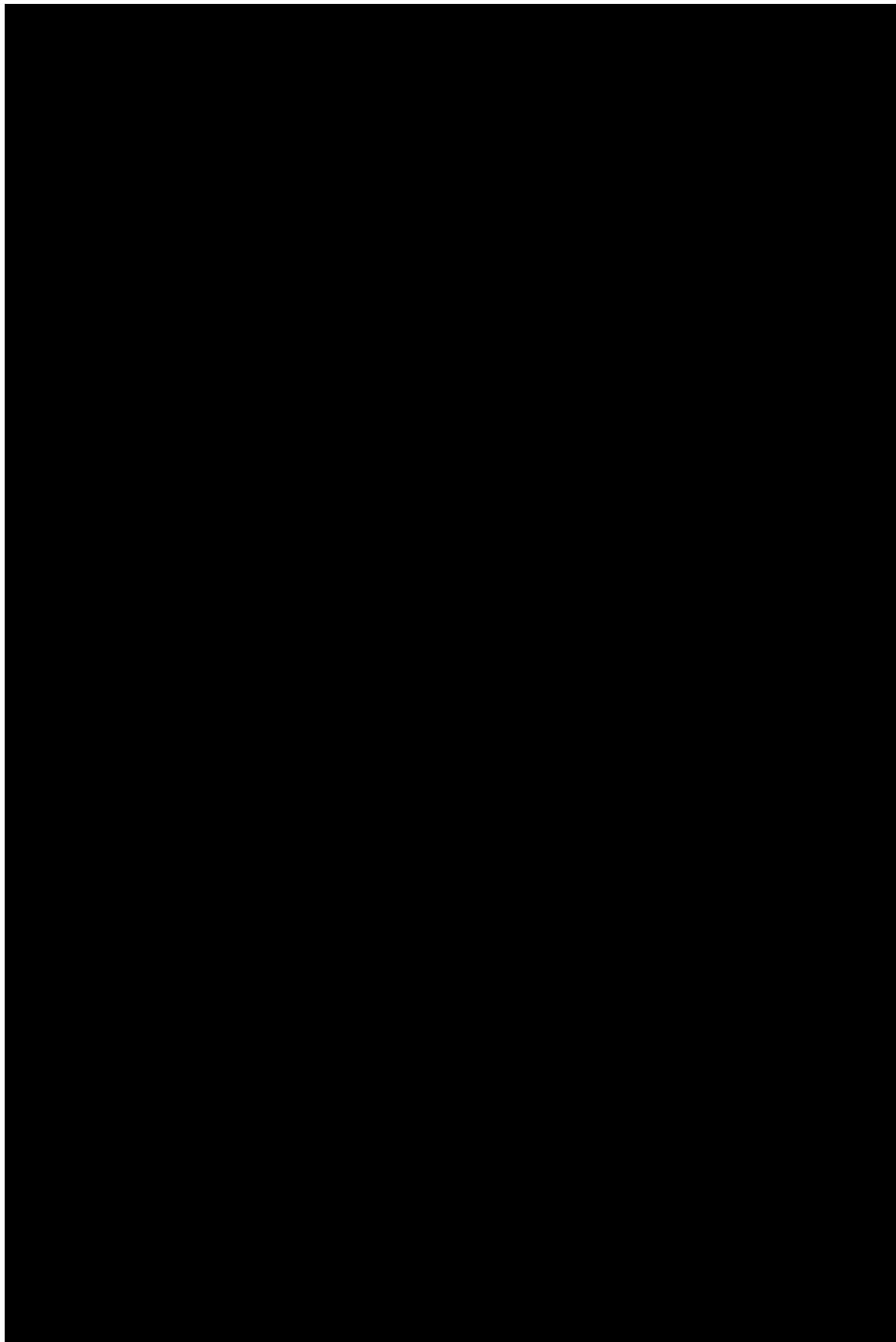
Meredith Whittaker, who researches the social implications of AI at New York University, says that it suddenly became clear to those on the outside that Google's commitment to ethics "was only paper thin".

Reverberations from the firing were powerful because Gebru has such a high-profile voice, says Luke Stark, a researcher who studies the social impacts of AI at Western University in Ontario, Canada. "It was immediately obvious that this was censorship," he says. The firing also brought into focus AI research's dependency on corporate money, and how academics have become so entangled in it, Whittaker adds.

Gebru says that she has long had ideas about creating her own institute to build a positive model of how AI work should be done. Those dreams came to fruition after she applied for grants and won US\$3.7 million in funding from several philanthropic organizations. The [Distributed Artificial Intelligence Research Institute](#) — a remote interdisciplinary centre — will develop AI models and applications that do not depend on the large data sets and computing power that the big tech companies hold.

The organization has two advisory board members and two research fellows (including Gebru), and expects to hire more people soon. "Her institute is not going to shy away from the reality of things," says Deborah Raji, who has collaborated with Gebru on AI ethics and works at the non-profit Internet foundation Mozilla. "It's going to be a great landing spot for people who want to ask these questions and don't have somewhere to do it. She's going to define the field in ways people don't appreciate. It is going to be amazing."







Credit: Rogan Ward for *Nature*

Credit: Rogan Ward for *Nature*

Tulio de Oliveira: Variant tracker

A bioinformatician in South Africa helped to identify troubling variants of the coronavirus SARS-CoV-2.

By Linda Nordling

On 25 November, Tulio de Oliveira announced the [discovery of a new variant of SARS-CoV-2](#). Omicron, detected in samples from Botswana, South Africa and Hong Kong, had a Swiss Army knife of mutations that de Oliveira and other leading scientists feared might help it to evade immunity from previous infection or vaccinations.

For de Oliveira, director of South Africa's KwaZulu-Natal Research Innovation and Sequencing Platform (KRISP), it was eerily reminiscent of

the previous year, when his team had discovered another SARS-CoV-2 variant of concern in South African samples. Beta, as that variant became known, led foreign governments to curb travel to and from South Africa many months after its discovery. Both variants were spotted after doctors and laboratory workers flagged unexpected rises in infections in areas that had already been hit hard by COVID-19.

De Oliveira knew that by reporting yet another concerning variant, he ran the risk of incurring fresh sanctions, which would economically penalize countries in southern Africa. But he also knew it was the right thing to do. “The way that one stops a pandemic is by quick action,” says the Brazilian-born bioinformatician. “Wait and see has not been a good option.”

The rapid identification of both Beta and Omicron in southern Africa reinforces the importance of having disease surveillance spread evenly around the world, says Jeremy Farrar, director of the biomedical research charity Wellcome, based in London. “If an imbalance continues, then where disease surveillance is limited, we risk new variants of COVID-19 — or even new diseases entirely — cropping up and spreading unchecked,” he says.

The COVID-19 pandemic isn’t the first time that genomic sequencing has been used to trace outbreaks in Africa; scientists used it in the Ebola outbreak in West Africa from 2014 to 2016. KRISP, created in 2017 with de Oliveira at the helm, has tracked pathogens behind diseases including dengue and Zika, and more common scourges such as AIDS and tuberculosis. But never before have so many different samples of the same virus been sequenced in such a short period of time — both in Africa and around the world.

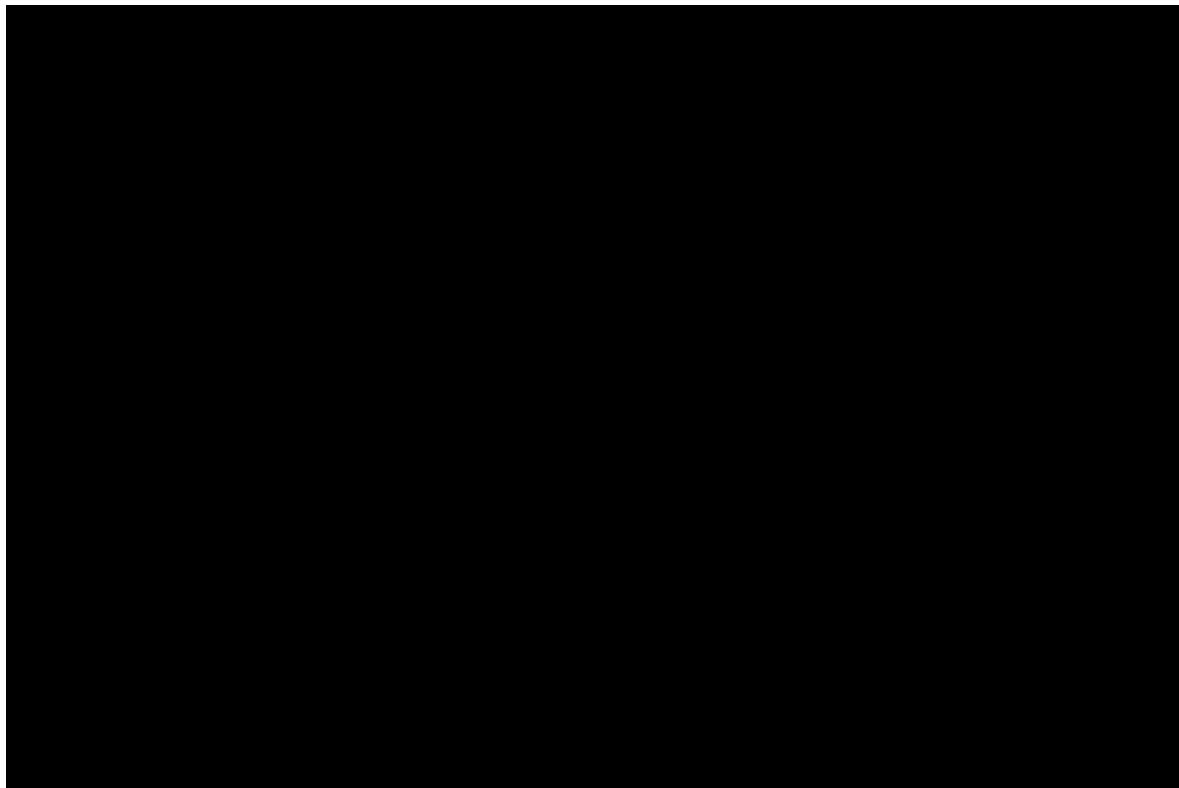
De Oliveira’s work has also influenced policymaking. KRISP’s way of working combines cutting-edge molecular technology with close links to doctors and nurses on the front line, to inform policy in real time. For example, their mapping of an early hospital outbreak of COVID-19 resulted in guidelines for ward layouts to prevent the virus from spreading in hospitals. “Tulio has done an incredible job pioneering a new way of science responding to epidemics,” says Christian Happi, a molecular biologist who

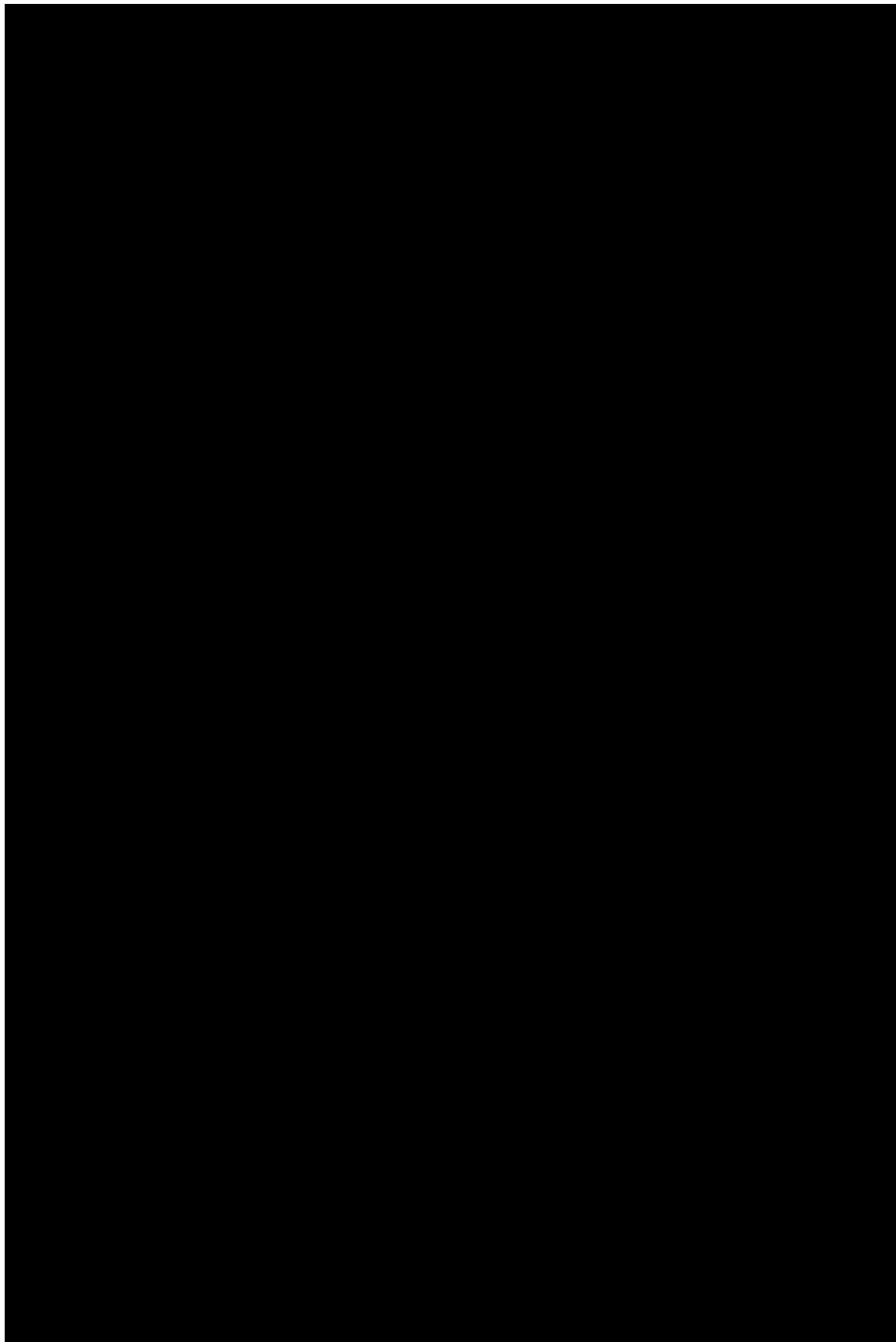
heads the African Centre of Excellence for Genomics of Infectious Diseases at Redeemer's University in Ede, Nigeria.

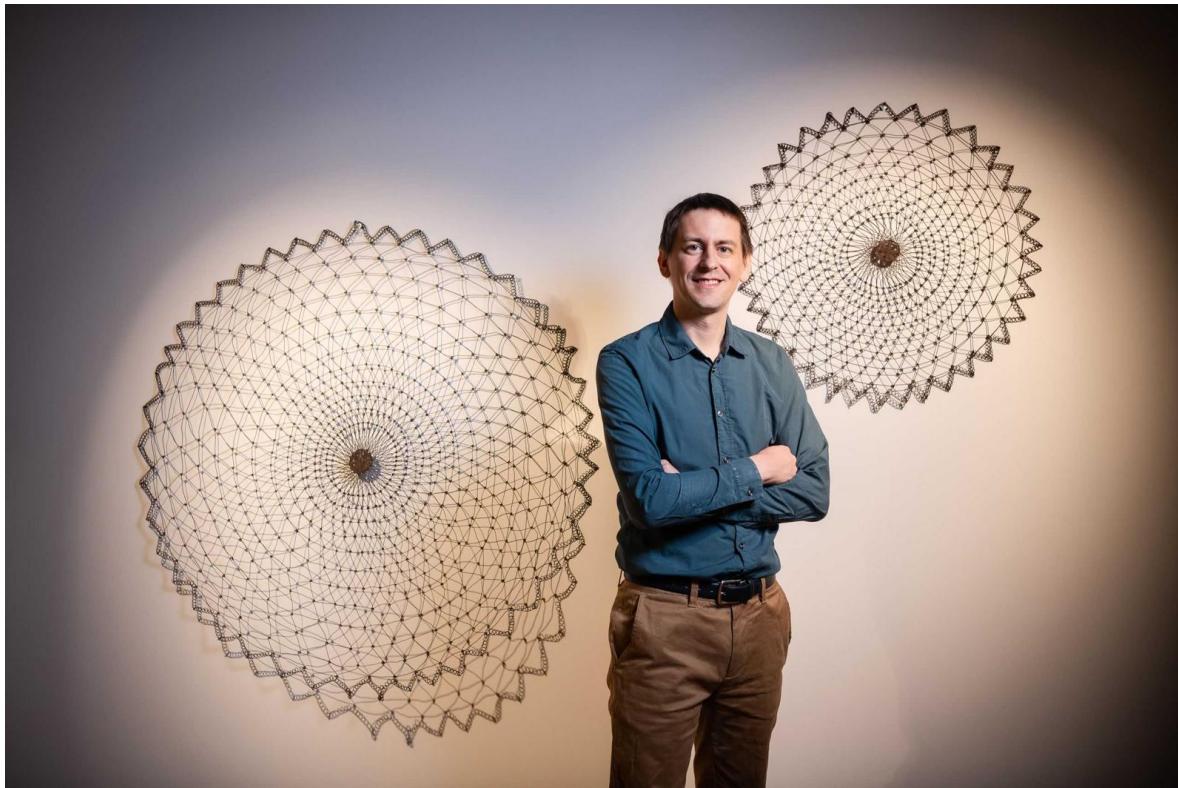
In December, de Oliveira moved permanently to Stellenbosch, outside Cape Town, South Africa, where he has been setting up the Centre for Epidemic Research, Response and Innovation (CERI) since July (he will keep his position at KRISP). The centre will work to control epidemics in Africa and the global south, and will house Africa's largest sequencing facility. The coronavirus pandemic has fuelled these investments, but the momentum is already spilling over into surveillance on other diseases, says de Oliveira. “The main thing we have shown the world is that these things can be done in developing countries.”

Not that those countries have been rewarded for it — quite the opposite. De Oliveira says he was extremely disappointed when rich countries imposed travel bans on southern Africa simply because the country had the scientific skill to discover new variants. The scapegoating of South Africa “was almost a smokescreen for the vaccine hoarding, and for rich countries losing control of the pandemic”, says de Oliveira. “Of course I expected more.”

De Oliveira’s role in announcing two variants of concern has given him a reputation for delivering bad news. When the Omicron announcement brought fresh travel bans, some South Africans, including politicians, queried de Oliveira’s right to make such pronouncements. Some people even view the genomic-surveillance community as the enemy. But, he says: “We are not the enemies, we are the opposite.”







Credit: Alecsandra Dragoi for *Nature*

Credit: Alecsandra Dragoi for *Nature*

John Jumper: Protein predictor

A team led by this AI researcher released a tool that is transforming biology.

By Ewen Callaway

What would it mean if determining the structure of almost any protein — all of its intricate folds, pockets and catalytic surfaces — were as simple as performing a web search? John Jumper and his colleagues at DeepMind in London attempted to answer that question earlier this year with the public release of AlphaFold, which uses artificial intelligence (AI) to predict protein structures with stunning accuracy.

“It is going to change the face of modern biology,” says Tobin Sosnick, a biophysicist at the University of Chicago, Illinois, and one of Jumper’s

former PhD advisers. “There’s going to be a before and after 2021, when people ask, ‘What is structural biology?’”

AlphaFold’s development was a process of destruction and rebirth. In 2018, the team working on it jolted the small community of scientists interested in protein-structure prediction. At a biennial competition called CASP (short for Critical Assessment of Structure Prediction), an early version of AlphaFold [outperformed all other computational tools for determining a protein’s shape from its sequence.](#)

But despite the win, AlphaFold didn’t generate predictions with fine enough detail, says Jumper, who was co-leader of the team at the time. Efforts to improve its performance hit a wall. So the researchers decided to start from scratch. “You had to throw away everything,” says Pushmeet Kohli, head of AI for Science at DeepMind, a subsidiary of Alphabet, Google’s parent company. Jumper had a key role in the decision to start afresh, says Kohli, and stuck with it, even when early versions of ‘AlphaFold2’ performed much worse than its predecessor. “He’s not afraid of taking on new directions.”

Jumper has a history of changing tack. He started a PhD in condensed-matter physics at the University of Cambridge, UK, but decided it wasn’t the right research topic. So he left with a master’s degree and wound up working on computer simulations of proteins at a private research group run by a physicist-turned-hedge-fund-manager. “I didn’t know what a protein was when I showed up,” says Jumper.

He next embarked on a chemistry PhD programme, in which he used machine learning to study protein dynamics. Applying AI to a scientific problem was “really magical”, says Jumper. But he wanted stability. So, he applied for jobs in finance as well as at commercial AI laboratories. During an interview with DeepMind, the company revealed its plans to tackle protein-structure prediction, and Jumper was intrigued. “I probably would have left science, if not for DeepMind,” Jumper notes.

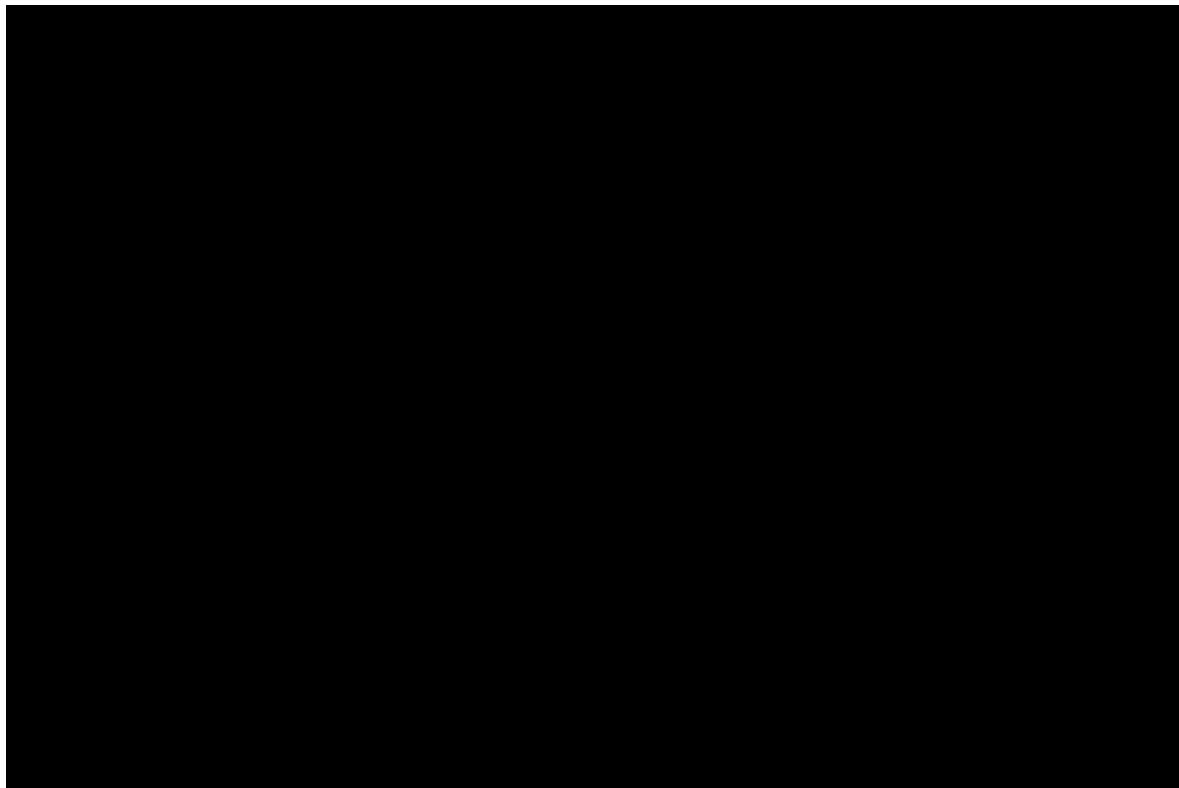
The first iteration of AlphaFold was based on a neural network that predicted the distance between parts of a target protein, an approach that other teams were also taking. Jumper wanted AlphaFold to deliver

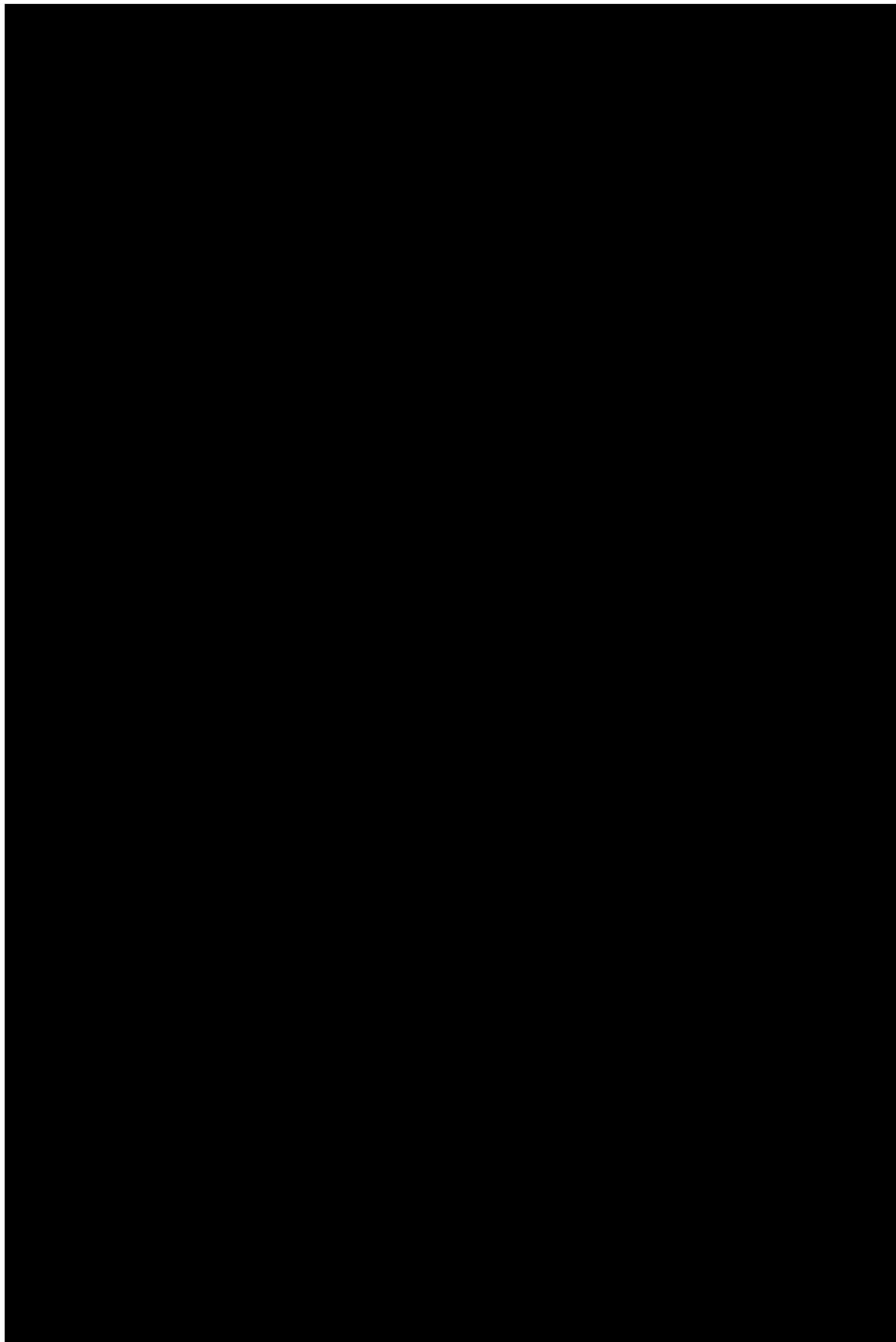
predictions that scientists could have confidence in, necessitating a complete overhaul of the underlying neural network.

The [second version of AlphaFold dominated CASP again in late 2020](#), this time by an even wider margin. Furthermore, nearly two-thirds of its predictions were on a par with experimentally determined structures. For Jumper, however, the most rewarding chapter in AlphaFold's story came in July. He and his team released the network's underlying code, as well as [predicted structures for almost all proteins in humans](#) and 20 other model organisms — 250,000 structures in total — together with the European Molecular Biology Laboratory's European Bioinformatics Institute in Hinxton, UK. They plan to release the structures of nearly half of all known proteins — totalling 130 million structures — next year.

Jumper's team regularly hears from other researchers who now use AlphaFold. One of the projects he was most excited about was a map of the nuclear pore complex, a gargantuan molecular machine that is a gatekeeper to the genomes of eukaryotic cells. The work combined AlphaFold's structures and other predictions with experimental structures comprising the complex, which consists of more than 1,000 individual protein chains.

For Jumper, such applications are the highest compliment. “To see the amount to which AlphaFold has changed the work of experimentalists has been really, really incredible,” he says. “The dream is to do something really useful.”







Credit: Annie Ling/The New York Times/ Redux/eyevine

Credit: Annie Ling/The New York Times/ Redux/eyevine

Victoria Tauli-Corpuz: Indigenous defender

A former revolutionary helped Indigenous peoples to gain international recognition for protecting biodiversity and the climate.

By Jeff Tollefson

As the [United Nations climate summit COP26 kicked off](#) in Glasgow, UK, several wealthy nations and more than a dozen philanthropic organizations stepped up with an unprecedented commitment. They pledged to provide US\$1.7 billion to help Indigenous peoples around the world to preserve forests, protect biodiversity and prevent global warming by keeping carbon locked up in plants and soils.

It was a watershed moment for Indigenous groups, and much of the credit goes to decades of work by Victoria Tauli-Corpuz, an Indigenous leader from the Philippines who served for six years as the UN special rapporteur on the rights of Indigenous peoples.

Tauli-Corpuz has spent years criss-crossing the globe to convince governments, environmentalists and philanthropic foundations that Indigenous peoples are the best stewards of forests and other hotspots of biodiversity — something that has recently been backed up by scientific literature.

“The world caught up with her, and also science caught up with her,” says David Kaimowitz, an economist at the Food and Agriculture Organization of the UN in Rome. Kaimowitz says that it has really only been in the past five or ten years that peer-reviewed literature has provided data showing that Indigenous lands serve as protective buffers against environmentally harmful activities such as mining, dams and deforestation (see [A. Blackman et al. Proc. Natl Acad. Sci. USA 114, 4123–4128; 2017](#)). Having grown up in an Igorot village without electricity in the mountains of the Philippines, Tauli-Corpuz learnt that lesson earlier than most. When the regime of former president Ferdinand Marcos sought to log her people’s forests and install a hydroelectric dam in their river in the 1970s, she joined the opposition.

“We defeated the dam, and the logging stopped,” says Tauli-Corpuz, who is the founder and executive director of the Tebtebba Foundation in Baguio City, the Philippines.

She realized that the fates of Indigenous peoples and the world’s forests are inextricably intertwined. As she worked her way through the UN bureaucracy over the course of 35 years, Tauli-Corpuz became a forceful critic of what she calls “fortress conservation” — a model that presumes that nature can be preserved only if walled off from humanity. “The conservation mindset has to be changed,” she says, because the forests that hold much of Earth’s biodiversity and carbon are also home to the world’s Indigenous populations. “People live in these forests, and we should be working with them.” In recent years, with the help of satellite imagery, scientists have come to the same conclusion. Like national parks and other protected areas,

Indigenous territories are less prone to deforestation, mining and dams than neighbouring lands.

That idea gained traction this year. Indigenous rights were recognized during the virtual UN Convention on Biological Diversity in October, and received unprecedented attention at the world conservation congress in Marseilles, France, in September. Indigenous groups attended the latter as members for the first time and successfully pushed for a motion that called on governments to protect 80% of the Amazon by 2025. And for the first time, governments and donors stepped up at COP26 with real funding commitments.

Activism came naturally for Tauli-Corpuz. In 1966, she earned a spot in an elite high school in Manila, and joined protests against her country's involvement in the Vietnam War. She was at university when the Marcos regime declared martial law in 1972, and studied to be a nurse in preparation for potentially violent uprisings. After university, she returned home to help organize community health programmes, while also promoting resistance against the regime and the proposed dam and logging — activities that resulted in a military raid on her home in 1980.

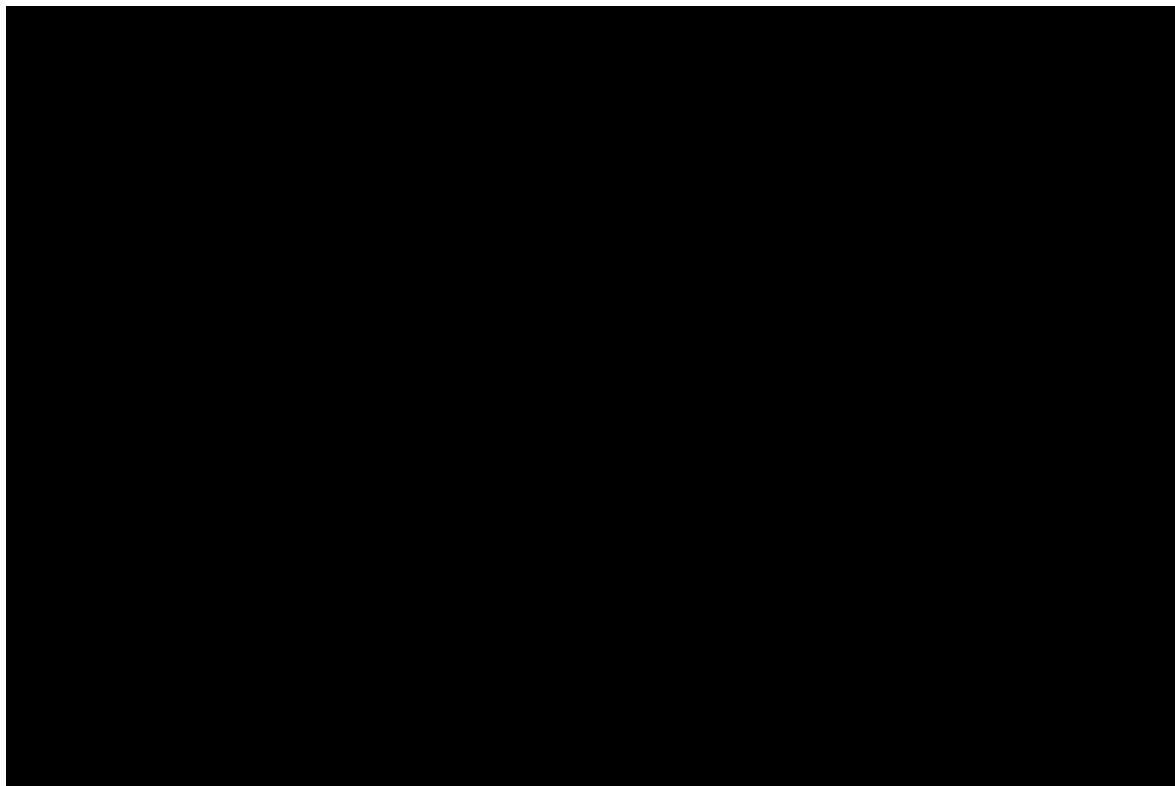
In 1985, Tauli-Corpuz received an invitation to join a UN panel in Geneva, Switzerland, investigating Indigenous rights. It took nearly a quarter of a century and plenty of travel, but Tauli-Corpuz saw the effort through. In September 2007, the UN General Assembly in New York City adopted a landmark declaration that recognized, for the first time, the collective rights of Indigenous peoples. Tauli-Corpuz then carried that effort forwards under the UN climate convention, ultimately helping to garner recognition of Indigenous rights in the 2015 Paris agreement. At the COP26 summit, Indigenous rights were once again recognized in an agreement governing international partnerships and carbon markets.

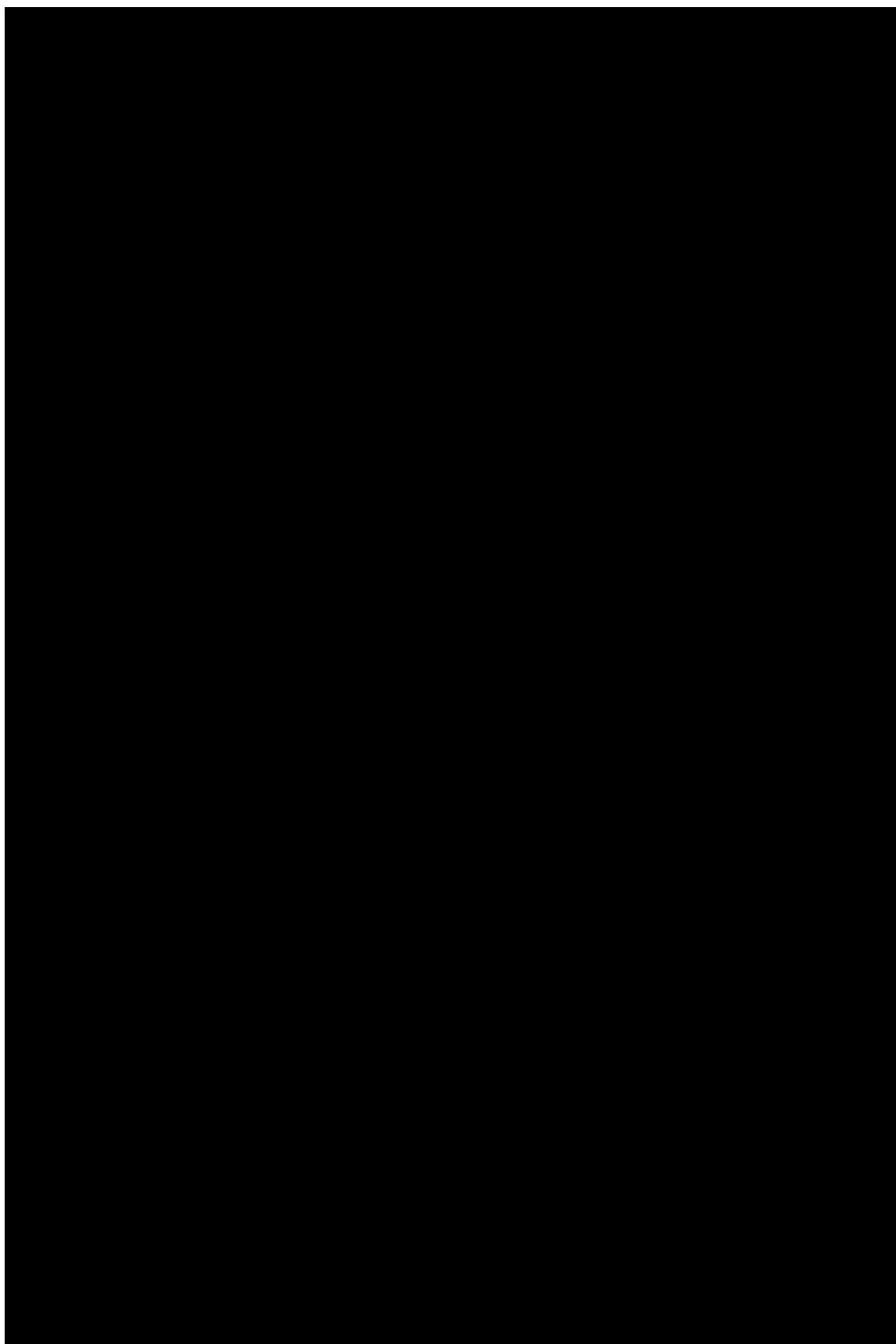
“The attention on Indigenous peoples was unprecedented,” says Jennifer Corpuz, one of Tauli-Corpuz’s four children, who has followed in her mother’s footsteps in the Indigenous-rights movement. An attorney with the US-based non-profit advocacy group Nia Tero, Corpuz acknowledges that she sometimes suffers from the weight of expectation, but says she never felt any pressure from her mother.

From 2014 to 2020, as the UN special rapporteur, Tauli-Corpuz travelled the world, holding meetings with Indigenous communities to talk about the challenges they are facing on the ground. In one [influential 2016 report](#) to the UN, she shone a light on how the creation and enforcement of conventional protected areas such as national parks and nature reserves has often impinged on the rights and land claims of Indigenous communities.

Now back at her foundation, Tauli-Corpuz continues to work with Indigenous communities around the world, helping them to understand their rights and gain the title to their traditional lands. She is also helping Indigenous communities to bolster their own governance systems, which will be crucial as they seek to propose projects and access the newly committed international funds.

“It’s really about helping the Indigenous peoples empower themselves,” she says. “Hopefully, we can strengthen their capacity to do what they need to do.”







Credit: Frédéric Scheiber for *Nature*

Credit: Frédéric Scheiber for *Nature*

Guillaume Cabanac: Deception sleuth

This computer scientist helped to uncover a new kind of fabricated paper.

By Diana Kwon

Underground creepy crawly state. Bosom malignancy. Sun oriented force. These might sound like expressions from a work of fiction, but they are actually strange translations, pulled from the scholarly literature, of scientific terms — ant colony, breast cancer and solar energy, respectively. Guillaume Cabanac, a computer scientist at the University of Toulouse, France, spots such bizarre phrases in academic papers every day.

This year, Cabanac and his colleagues found these [tortured phrases, as they call them, in thousands of papers](#). A handful have been retracted; publishers

are investigating many more. Cabanac has built a website to keep track of the mushrooming problem. “They found this whole new hornet’s nest of articles that appear to be completely fake,” says Elisabeth Bik, a research-integrity analyst in California.

Weeding out these problems is related to Cabanac’s day job: he specializes in analysing the scholarly literature, and now devotes around two hours a day to finding tortured phrases. Some people might find them funny, but Cabanac takes the problem seriously. “This shouldn’t be happening,” he says.

Cabanac’s hunt for gibberish papers began in 2015, when he started collaborating with Cyril Labbé, a computer scientist at the University of Grenoble Alpes in France. Labbé had developed a program to spot gibberish computer-science papers automatically generated using SCIGen, a piece of software created initially as a joke. Labbé’s work led [journals to withdraw more than 120 manuscripts](#). Cabanac helped to update Labbé’s program to find papers only partially written by SCIGen, and to locate them using Dimensions, a search engine for scholarly literature. This year, they [reported finding hundreds more papers containing nonsense text](#), published in journals and conference proceedings and as preprints.

To raise awareness, Cabanac and his colleagues e-mailed publishers and posted their findings on social media and on PubPeer, a post-publication peer-review site. Cabanac also created the [Problematic Paper Screener](#), a website for flagging and reporting questionable manuscripts. “He gets frustrated about fake papers,” Labbé says. “He’s really willing to do whatever it takes to prevent these things from happening.”

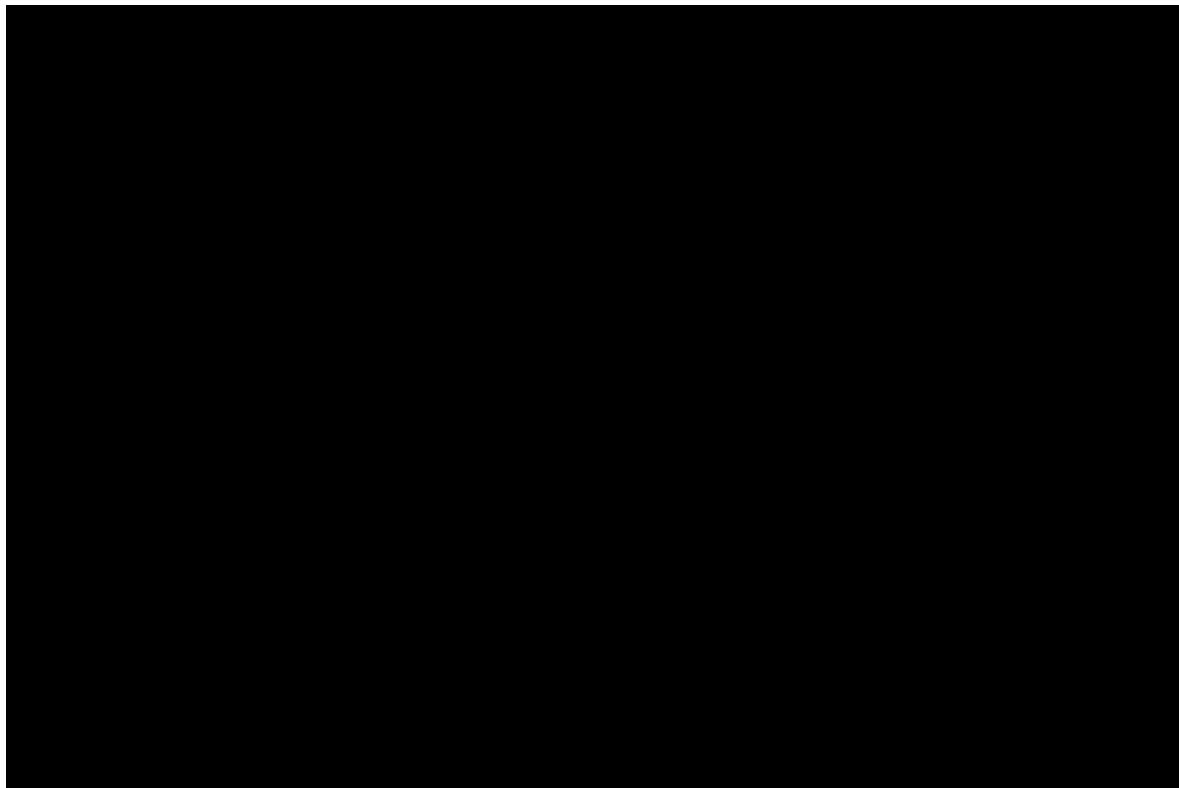
The SCIGen work led Alexander Magazinov, a software engineer at the multinational technology firm Yandex, headquartered in Moscow, to contact Labbé and Cabanac. Magazinov asked whether SCIGen might be behind oddly paraphrased versions of scientific concepts he’d noticed in papers, such as “colossal information” for “big data”. Together, the three [located the terms in hundreds of papers](#), which they reported in July. Digging deeper, they suggested that machine-paraphrasing tools might have been used to create them.

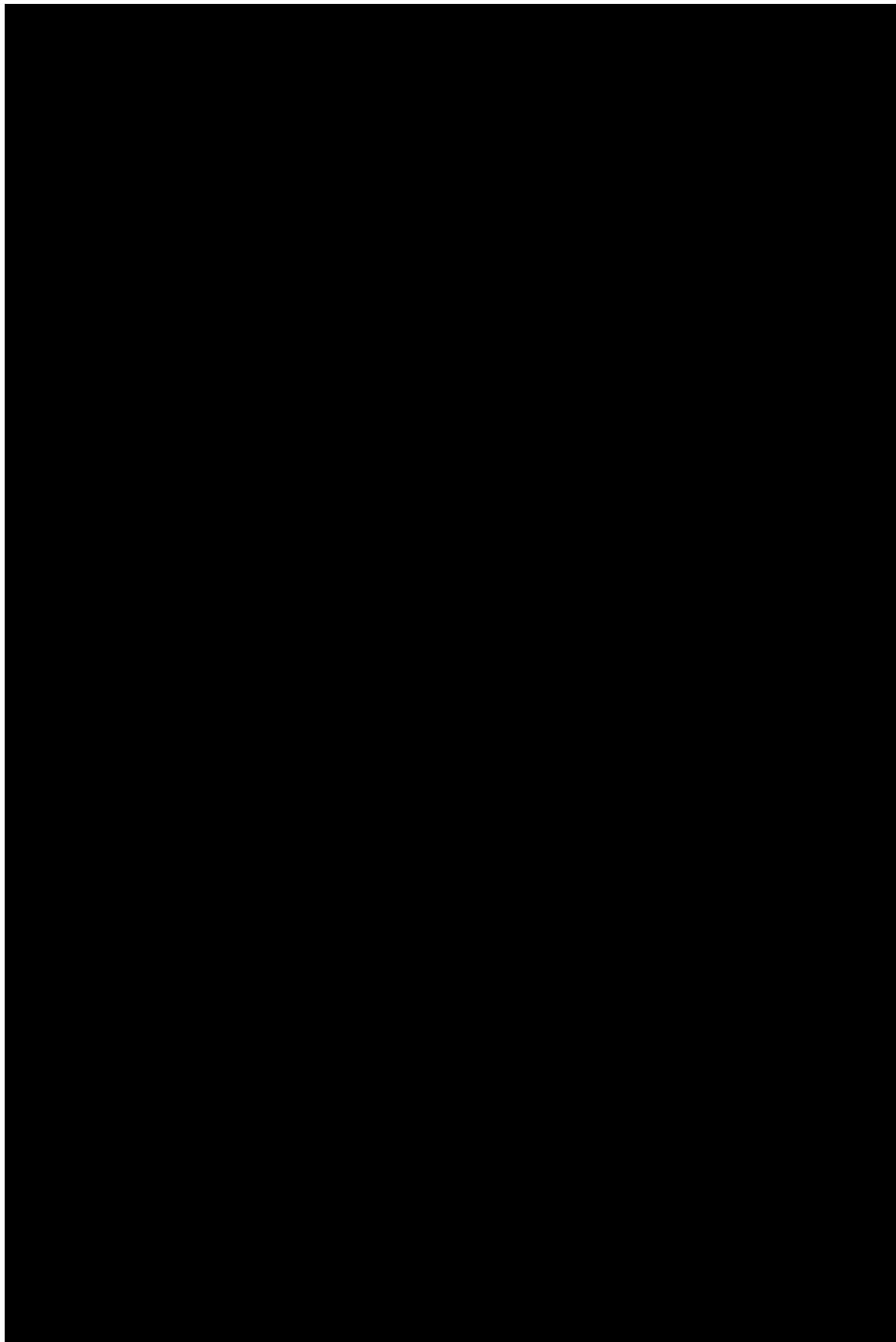
“I think these tortured phrases indicate a failure of peer review,” says Jennifer Byrne, a cancer researcher at the University of Sydney in Australia, with whom Cabanac has worked on other scientific-integrity projects.

“Surely, somebody who was conscious during the peer-review process would have seen that that’s not really right.”

By now, Cabanac and colleagues — along with volunteers from the PubPeer community — have pinpointed nearly 400 tortured phrases in more than 2,000 papers, including ones in journals from well-known publishers such as Elsevier and Springer Nature (*Nature*’s news team is editorially independent of its publisher). Each such phrase first has to be spotted by a person; a search algorithm then runs on Dimensions’ index to find papers that include it. Cabanac and a host of helpers manually scan each of these articles to weed out false positives. Eventually, Cabanac would like to develop a program that can identify tortured phrases automatically.

Cabanac hopes that his work will help to decontaminate the scientific literature. But he knows that will not be easy. “I’m afraid of new techniques that would help scammers publish papers containing errors that would be less detectable,” he says. “It’s a whack-a-mole game. We need to be prepared.”







Credit: Jessica Hallett/*Nature*

Credit: Jessica Hallett/*Nature*

Meaghan Kall: COVID communicator

A government epidemiologist went against norms to tweet explanations of UK coronavirus data.

By Richard Van Noorden

On the afternoon of 8 January, Meaghan Kall, a UK government epidemiologist, helped to put the finishing touches to a technical briefing document about a concerning SARS-CoV-2 variant spreading in southeast England. Then, about half an hour after the report was published, she tried something new: she posted a Twitter thread breaking down its key points.

Kall had seen increasing disquiet and confusion about the coronavirus variants online, and wanted to explain publicly what the government data

showed. She hadn't asked permission from her bosses at Public Health England (PHE) — an agency tasked with responding to health threats, now succeeded by the UK Health Security Agency. "I just did it," she says. But she soon got an audience. UK researchers chimed in with questions; so did a US philosopher and an Argentinian programmer.

It was the first of a series of accessible, rapid explainers from Kall on dozens of the agency's coronavirus briefings. [Through her tweets](#) — prepared and posted around her day job — she became a human face for a government team that has provided many early answers to burning questions about COVID-19 in 2021.

Thanks to its early roll-out of vaccines, well-equipped genomics laboratories and unified National Health Service (NHS), the United Kingdom was quick to produce high-quality data on the coronavirus, from the spread of new variants to the effectiveness of vaccines. "PHE's data have been absolutely invaluable and have been used by the whole world to understand many key aspects of COVID-19," says Marm Kilpatrick, an infectious-disease researcher at the University of California, Santa Cruz.

UK researchers did a better job communicating these data than other countries with early vaccine roll-outs. Scientists from Israel, for instance, often shared initial data on social media only as pictures, making it hard to extract data, and in Hebrew, making the results difficult for an international audience to understand, says Ben Cowling, an epidemiologist at the University of Hong Kong.

With so much misinformation and confusion about what data such as these can mean, public-health agencies struggled to get their messages across. Kall stepped into this breach. "Meaghan has been a wonderful source of hot-off-the-presses information and has done a fantastic job in breaking down the reports," says Kilpatrick.

Dozens of scientists have emerged as communication stars on Twitter during the pandemic, but Kall's position is unusual. "As a civil servant, I have limited scope to speak my truth," she tweeted in September. In more than 6,000 tweets this year, she also shared photos of her rapid coronavirus tests when her family had to isolate, and invited people who were worried about

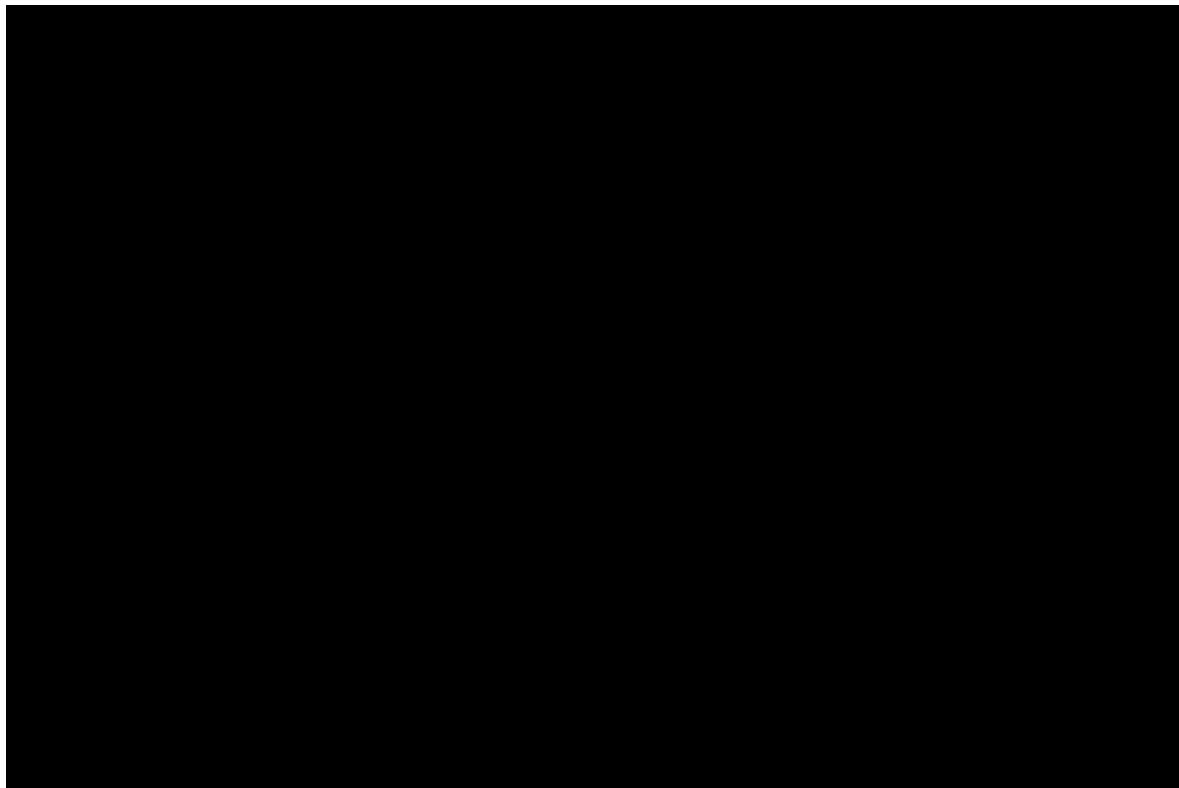
COVID-19 vaccines to message her privately. Her tweeting was neither forbidden nor encouraged by her bosses, she says, although some colleagues have been personally supportive.

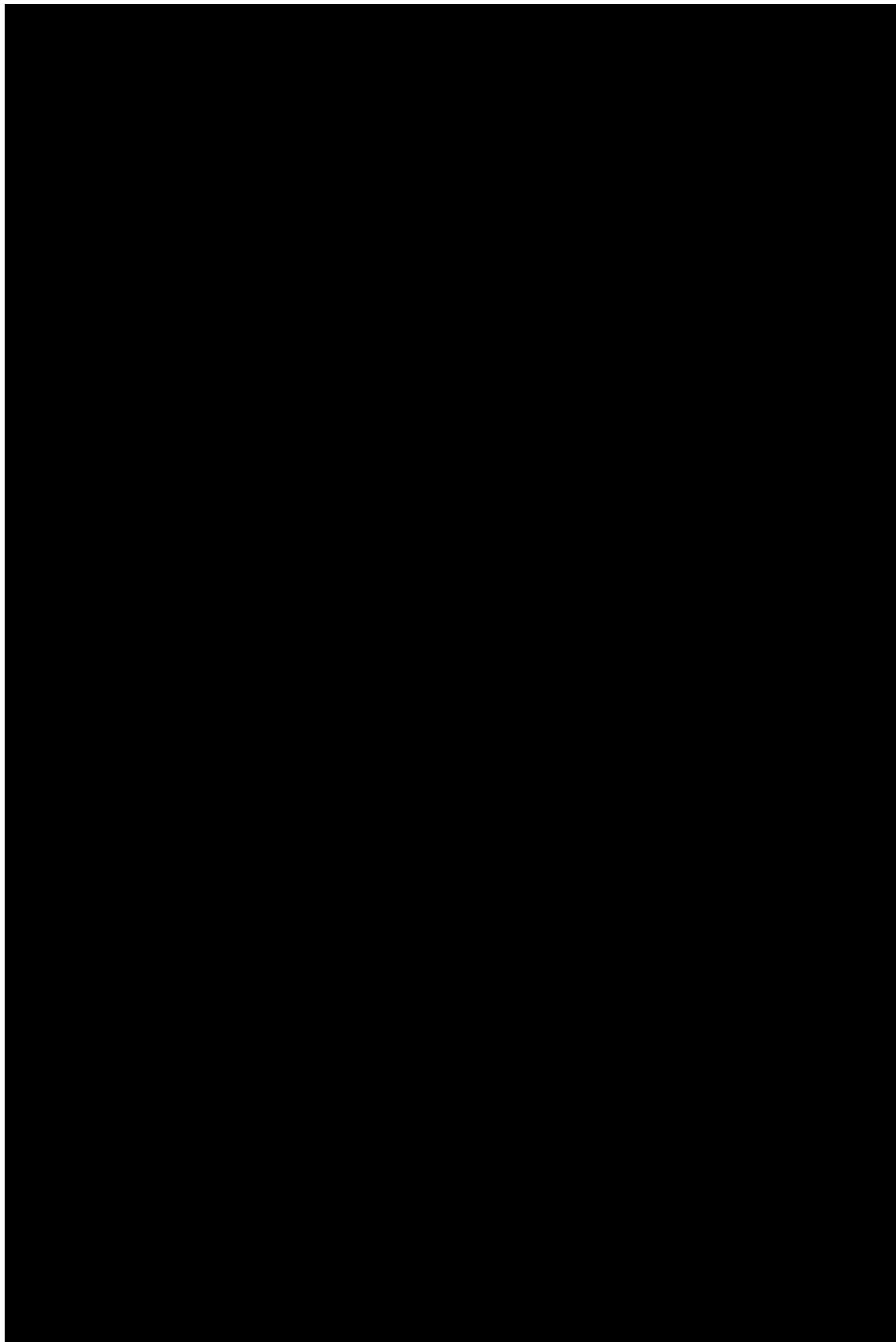
Kall thinks her public communication has aided trust in UK government data. “My favourite responses are the people who say, ‘I didn’t really have a lot of faith in PHE, I didn’t really trust it, until I started following your Twitter feed,’” she says.

Before COVID-19, Kall spent a decade monitoring HIV infections for UK government agencies. She says that her work with people with HIV — she was writing up her PhD on a national survey of those living with the virus when COVID-19 struck — gave her experience in understanding people’s sometimes opposing views, as did her upbringing in a small, conservative town in Michigan.

Kall’s most popular tweets have criticized the UK coronavirus response, which included the controversial privatization of most of its testing and contact tracing. In November, she wrote: “I still think it was a huge oversight that our NHS sexual-health advisers, who are professional contact tracers, were never drafted in or consulted on Test and Trace.” Asked — with a press officer watching — about how much freedom she has to tweet, she says that she’s never been told what to say or been reprimanded, but she carefully chooses the issues that she feels strongly about.

Kall says putting together her Twitter threads cuts into her personal life — it takes her an hour or so to prepare each briefing breakdown, in addition to responding to questions — and colleagues ask her how she keeps going. “One of my main aims”, she says, “is really just to try and make sure people are empowered, and have agency to understand the data to make their own decisions, from a reliable source.”







Credit: Stefani Reynolds/The New York Times/eyevine

Credit: Stefani Reynolds/The New York Times/eyevine

Janet Woodcock: Drug chief

This career administrator led the US's premier drug agency through a challenging year.

By Heidi Ledford

Only days after Joe Biden became US president this January, he appointed Janet Woodcock as acting commissioner of the US Food and Drug Administration (FDA). Soon, the letters came flooding in. Some were in her favour: one signed by 82 rare-disease patient-advocacy organizations praised her leadership and her focus on integrating patient voices in drug-approval decisions.

Others were less laudatory: 31 advocacy organizations urged the secretary of the US Department of Health and Human Services to keep Woodcock's time at the helm of the FDA brief. "Dr Woodcock presided over one of the worst regulatory agency failures in US history," they wrote, laying partial responsibility for the country's raging opioid crisis at her feet. The debates over her would continue through a tumultuous year at the top of the agency.

Woodcock, a former medical doctor, spent most of her 35-year career at the FDA running the Center for Drug Evaluation and Research, which is responsible for ensuring that drugs are safe and effective before they are approved for the US market. Woodcock helped to modernize the centre's drug-evaluation process, ushering in advanced clinical-trial designs and pathways to drug approval that are coupled with sophisticated diagnostic tests. She also oversaw the development of a bigger role for patients and their advocates in the approval process. In January, there was speculation that President Biden might nominate her to stay on as commissioner.

By June, however, that prospect dimmed after a controversial FDA decision to [approve the drug aducanumab for treatment of Alzheimer's disease](#). The drug, developed by Biogen in Cambridge, Massachusetts, had been shown to reduce tangled amyloid- β proteins in the brains of people with the disease. But it did not seem to improve cognitive function or symptoms.

A panel of external advisers to the FDA voted against the approval, but the agency made the unusual decision to ignore the recommendation. Michael Carome, director of health research at the consumer advocacy group Public Citizen in Washington DC, says it was a pivotal moment: the agency had approved a drug that could be taken by millions, without solid evidence that it helps people. "It was one of the worst decisions the agency has ever made," says Carome.

Woodcock declined to comment for this article, but an FDA spokesperson said that she was not involved in the aducanumab approval. Even so, as acting head of the agency, she bore some responsibility, says Aaron Kesselheim, a physician who also studies drug regulation at Harvard Medical School in Boston, Massachusetts. Kesselheim served on the FDA advisory panel and, along with two others, resigned in protest over the decision. Woodcock has long advocated a close relationship between

industry and the FDA, raising concerns among some consumer groups and academic scientists.

Kesselheim quit to call attention to what he worried might be an emerging trend at the agency: “My hope is that this kind of bad decision-making doesn’t metastasize.”

The agency faced more controversy in the summer, when the White House announced that the country would soon administer booster vaccines for COVID-19. The FDA had not yet issued a decision on the boosters, and Biden’s announcement was seen as not only premature, but also antithetical to his promise to put science and evidence at the centre of decision-making. Woodcock, along with several other US public-health officials, endorsed the plan but said that it was subject to regulatory approval.

The episode caused a disturbance at the FDA, and two key vaccine evaluators resigned in protest.

In November, Biden nominated cardiologist [Robert Califf, a former FDA commissioner, to return to lead the agency](#). But Woodcock’s legacy will go beyond her year at the FDA’s helm, says Kesselheim.

Despite his critiques, he says, “I think the FDA makes the right decision most of the time, and continues to serve as a gold standard for drug regulation around the world.” Woodcock, he adds, “deserves some recognition for the way the FDA has evolved with the times”.

Ones to watch in 2022

- **Chikwe Ihekweazu**

WHO Hub for Pandemic and Epidemic Intelligence

This epidemiologist will direct the [surveillance hub and gather data on the COVID-19 pandemic](#) and other outbreaks.

- **Jane Rigby**

NASA Goddard Space Flight Center

As project scientist for operations for the [soon-to-be-launched James Webb Space Telescope](#), this astrophysicist will help to orchestrate what discoveries are made.

- **Love Dalén**

Swedish Museum of Natural History

This geneticist has [sequenced the oldest DNA on record](#) — 1.65 million years old, from a mammoth — and is now going after more genetic remains.

- **Xie Zhenhua**

China's special envoy for climate change

This politician will have a central role in ongoing talks to [strengthen international efforts aimed at combating climate change](#).

- **Graziano Venanzoni**

Italian National Institute for Nuclear Physics

This physicist is co-spokesperson for the [Muon g-2 experiment](#). The effort could produce strong evidence of discrepancies between experiments and the standard model of particle physics, hinting at new fundamental discoveries.

This article is also available as a [pdf version](#).

Correction

The profile on Timnit Gebru incorrectly referred to Ethiopia's 1998 conflict with Eritrea as a civil war.



- [Privacy Policy](#)
- [Use of cookies](#)
- [Legal notice](#)
- [Terms & Conditions](#)
- [Accessibility statement](#)

© 2021 Springer Nature Limited. All rights reserved.

[Top](#)



[Built with Shorthand](#)

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03621-0>

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

Opinion

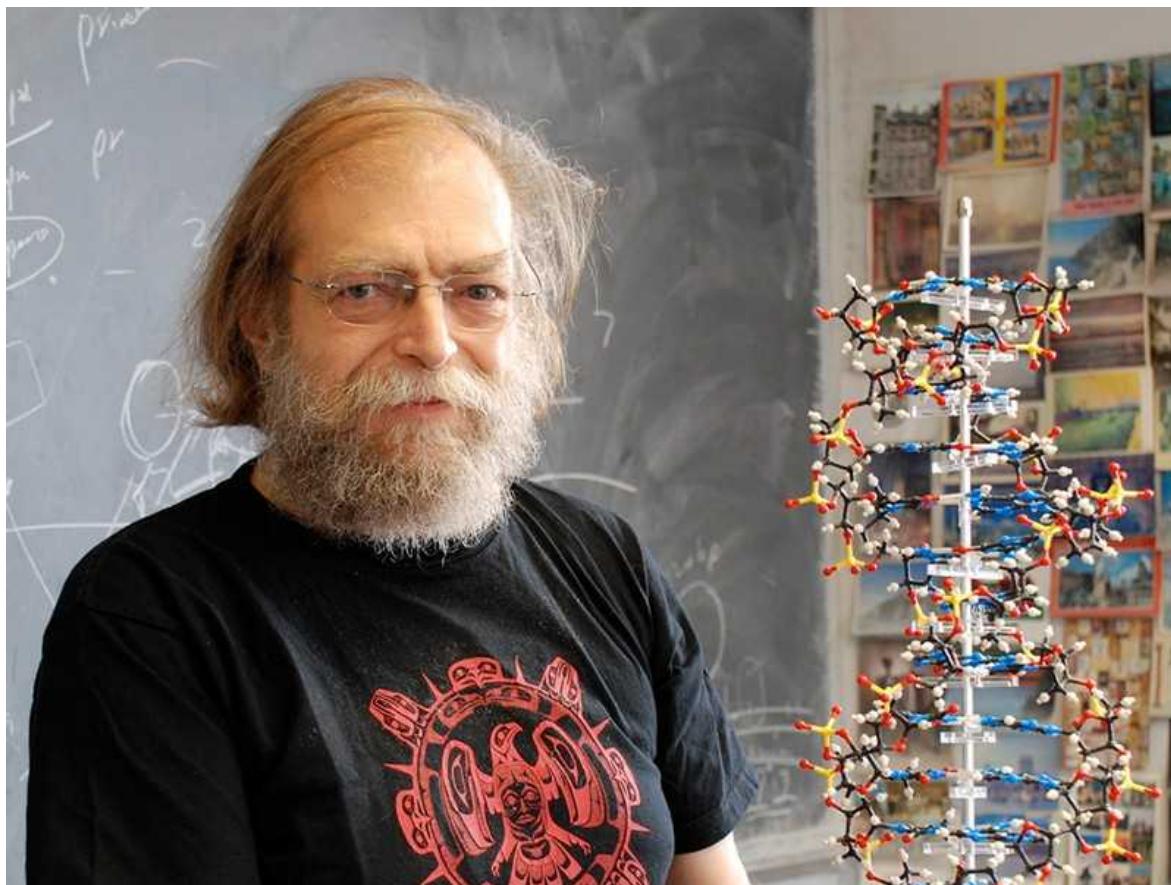
- **[Ned Seeman \(1945–2021\)](#)** [16 December 2021]
Obituary • Nanotechnologist who built the first self-assembling DNA structures.
- **[Upgrade the science and technology policy system the US already has](#)** [21 December 2021]
Correspondence •
- **[Early-career researchers help Wellcome funding panel](#)** [21 December 2021]
Correspondence •
- **[Half measures in One Health fail people and the environment](#)** [21 December 2021]
Correspondence •
- **[Secret of Hellenic Foundation's success? Its staff](#)** [21 December 2021]
Correspondence •

- OBITUARY
- 16 December 2021

Ned Seeman (1945–2021)

Nanotechnologist who built the first self-assembling DNA structures.

- [Philip Ball](#) ⁰



Credit: Michael Summers

Nadrian (Ned) Seeman is widely credited with being the first person to recognize that DNA can be used to design and build programmable nanostructures and nanomachines. With this new way of thinking about self-

assembly, he helped to turn chemistry into an information science. He has died aged 75.

Seeman was born in Chicago, Illinois, in 1945. A self-professed ‘Sputnik kid’, his passion for science was ignited by his high-school biology teacher. As a medical undergraduate at the University of Chicago, he was delighted to discover that professors did more than just teach. “I had no idea that I could spend most of my day having fun by doing research and also get paid to do it,” he wrote in an autobiographical note on receiving a Kavli Prize in Nanoscience in 2010.

Although by his own admission not a diligent student, he stayed at Chicago to take a doctorate in biochemistry, before switching to crystallography at the University of Pittsburgh in Pennsylvania. As a postdoctoral researcher, he worked with pioneer of DNA chemistry Alexander Rich at the Massachusetts Institute of Technology in Cambridge. He then joined the biology department of the State University of New York at Albany in 1977. He was disparaging about his time there. “The only thing worse than looking for a job was finding this one,” he wrote.

Ostensibly still a crystallographer, he bent his attentions in a new direction while at Albany. The turning point came when a postdoc asked him to build a model of a forked, double-stranded DNA structure called a Holliday junction, which can occur naturally.

The possibilities opened up by these structures dawned on him one day as he sat in the Albany campus bar recalling the image *Depth* by Dutch graphic artist M. C. Escher. In the 1955 print, rocket-like fish with horizontal and vertical fins touch ends in a 3D array. Seeman’s epiphany was that the branched units of the Holliday junction could assemble themselves into an orderly structure in a similar way, through interactions at the tips. If these tips were short single-stranded sections of DNA, their ends would stick to other complementary DNA sequences.

Seeman’s 1982 paper¹ outlining the idea of making lattices from DNA junctions got, he recalled, zero reaction. Because he was using DNA, a common response to his work was ‘where’s the biology?’ To which Seeman would reply that he was exploiting an aspect of DNA that to the best of his

knowledge is not used by biology. People took more notice when, in 1991, he and his graduate student Junghuei Chen synthesized a DNA molecule shaped like the edges of a cube². Working in the chemistry department at New York University in New York City, where Seeman spent the rest of his career, they used three-armed Holliday junctions as the vertices.

The resulting report² was initially regarded as something of a curiosity, but Seeman went on to produce more ambitious architectures and functions. In collaboration with Erik Winfree at the California Institute of Technology (Caltech) in Pasadena, and others, he began to think about assembly in algorithmic terms: programming DNA components through their sequences to link up in a manner governed by specific interaction rules. After Winfree, Seeman and their collaborators reported 2D DNA arrays³ in 1998, they and others created DNA ‘tiles’ that self-assemble into a connected mosaic following an algorithm, thereby conducting a kind of computation⁴.

From the late 1990s onwards, many other laboratories took up the challenge of what came to be called DNA nanotechnology. This pleased Seeman. “We don’t have to have all the ideas any more, and we don’t have to make all the mistakes,” he said in an interview.

One goal of DNA nanotechnology is to use the nanostructures as a scaffold, for example to arrange into regular arrays molecules that are hard to crystallize alone. Another goal is to use these biocompatible materials for drug delivery and tissue engineering.

Paul Rothemund at Caltech developed this programmable approach to make ‘DNA origami’, in which instructions for more or less any shape could be built into the folding strands⁵. Seeman and others made assembly reversible using short, single strands that can unzip, allowing DNA nanostructures to be fashioned into dynamic nanomachines.

Seeman was renowned for his blunt way with words and dry, self-deprecating wit. He has been described as a singular character — at once gruff and caring, vulgar and articulate, stubborn and visionary. Winfree told me, “I feel it’s no accident that DNA nanotechnology sprang forth from someone who took almost a perverse pleasure in thinking differently.”

Seeman's many honours and awards included the American Chemical Society's Nichols Medal in 2008. His real legacy, however, is to have launched a field of startling originality, the full potential of which has yet to be felt. "DNA nanotechnology has taken on a life of its own since Ned's original vision," Rothemund wrote⁶ in 2006. "We look forward to a new 'DNA world' in which an all-DNA 'bacterium' wriggles, reproduces and computes."

Nature **600**, 605 (2021)

doi: <https://doi.org/10.1038/d41586-021-03709-7>

References

1. 1.

Seeman, N. C. *J. Theor. Biol.* **99**, 237–247 (1982).

2. 2.

Chen, J. & Seeman, N. C. *Nature* **350**, 631–633 (1991).

3. 3.

Winfree, E., Liu, F., Wenzler, L. A. & Seeman, N. C. *Nature* **394**, 539–544 (1998).

4. 4.

Mao, C., LaBean, T. H., Reif, J. H. & Seeman, N. C. *Nature* **407**, 493–496 (2000).

5. 5.

Rothemund, P. K. *Nature* **440**, 297–302 (2006).

6. 6.

Rothenmund, P. K. in *Nanotechnology: Science and Computation* (eds Chen, J., Jonoska, N. & Rozenberg, G.) 3–21 (Springer, 2006).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03709-7>

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

- CORRESPONDENCE
- 21 December 2021

Upgrade the science and technology policy system the US already has

- [Valerie J. Karplus](#) ORCID: <http://orcid.org/0000-0001-5493-9150>⁰
&
- [M. Granger Morgan](#)¹

We disagree that the United States needs a cabinet-level Department of Technology and Science Policy ([H. Varmus and E. Zerhouni *Nature* 600, 30–32; 2021](#)). Instead, it should strengthen the existing — and highly effective — institution of the White House’s Office of Science and Technology Policy (OSTP).

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Nature **600**, 606 (2021)

doi: <https://doi.org/10.1038/d41586-021-03778-8>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03778-8>

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

- CORRESPONDENCE
- 21 December 2021

Early-career researchers help Wellcome funding panel

- [Rachel V. Guest](#) ORCID: <http://orcid.org/0000-0003-3213-7688>⁰,
- [Laura Benjamin](#) ORCID: <http://orcid.org/0000-0002-9685-1664>¹ &
- [Daniel R. Morales](#)²

We are among the first early-career researchers who were recruited by Wellcome, one of the world's biggest bioscience funding bodies, to sit on a grant panel this year. We provided a voice that represented trainees in the clinical academic community. In our view, this improved the programme experience and raised standards. We urge other funders to follow Wellcome's lead in helping to transform research culture.

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Nature **600**, 606 (2021)

doi: <https://doi.org/10.1038/d41586-021-03777-9>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03777-9>

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

- CORRESPONDENCE
- 21 December 2021

Half measures in One Health fail people and the environment

- [Andrew Peters](#) ORCID: <http://orcid.org/0000-0003-4938-2823>⁰ &
- [Carlos das Neves](#) ORCID: <http://orcid.org/0000-0003-0348-4808>¹

We applaud this month's joint statement on a shared definition of One Health from the United Nations Environment Programme and the Tripartite collaboration of agencies (see go.nature.com/3dnnpb5). It aims for integrated recognition of the interdependence of human, animal and environmental health. Too many international, regional and national bodies still persist in compartmentalizing their responses to public-health crises in ways that run counter to the One Health goal.

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Nature **600**, 606 (2021)

doi: <https://doi.org/10.1038/d41586-021-03780-0>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03780-0>

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

- CORRESPONDENCE
- 21 December 2021

Secret of Hellenic Foundation's success? Its staff

- [Nektarios K. Nasikas](#) 9

The Hellenic Foundation for Research and Innovation (HFRI) was established five years ago to fund the best research proposals from Greek universities and research centres ([Nature 535, 333; 2016](#)). In its relatively short existence, its contribution to advancing blue-skies research in Greece has been remarkable — particularly given the country's persistent economic challenges and the impact of the COVID-19 pandemic on the functioning of our society.

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Nature **600**, 606 (2021)

doi: <https://doi.org/10.1038/d41586-021-03779-7>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03779-7>

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

Work

- **The scientific workplace in 2021** [17 December 2021]
Career Feature • The impact on careers of a lingering pandemic and industrial unrest were among the challenges faced by working scientists in the past 12 months
- **Embryo-like models shed fresh light on early human development** [16 December 2021]
Technology Feature • A wave of stem-cell systems are enabling researchers to unpick what happens after an embryo implants in the uterus.
- **A partridge in hand on the Spanish steppe** [17 December 2021]
Where I Work • As part of his PhD research into the effects of farming and hunting on endangered bird populations in Spain, Xabier Cabodevilla tracked birds and collected faecal samples from roosting sites.

- CAREER FEATURE
- 17 December 2021

The scientific workplace in 2021

The impact on careers of a lingering pandemic and industrial unrest were among the challenges faced by working scientists in the past 12 months

- [Chris Woolston](#) ⁰

[Find a new job](#)



Academics in London join a three-day nationwide strike over pay and pensions in December.Credit: UCU

The past year delivered challenges as well as opportunities to working scientists around the world. Every week brought new developments: changes in funding (for better or worse), disruptions from a lingering pandemic, fresh warnings about inequity and toxicity in the workplace, and other reminders that careers in science remain dynamic and rewarding, but can also be daunting.

COVID-19 continued to shape science. Surveys of researchers conducted in the first few months of 2021 confirmed that the pandemic had hampered the productivity and [strained the mental health of researchers](#) in the United States and the United Kingdom. The [disruptions were especially harmful to female scientists](#), according to a report by the US National Academies of Sciences, Engineering, and Medicine. A separate survey [broadened that to anyone with childcare duties](#). Shortages of reagents, pipette tips and other supplies [forced researchers to scramble and adapt](#). More than 40% of those who responded to *Nature*'s global salary and job satisfaction survey in June and July reported that the [pandemic had negatively affected their career prospects](#). Despite all the downsides, some researchers still saw a pandemic silver lining. The results of another *Nature* poll, published in March, found [wide support for the continuation of virtual conferences](#).

Prospects

Job security remained a leading concern, especially for scientists in academia. In *Nature*'s 2021 salary and job satisfaction survey, [just over 40% of respondents in academia felt positively about their career prospects](#), compared with more than 60% of respondents in industry. “Academia is turning into a milling machine with no regard for life–work balance and fair compensation,” said one respondent.

There is reason to worry. In April, a survey of faculty members by the American Association of University Professors documented [falling salaries and job losses at US universities](#). A survey published in July found [widespread anger among UK academics](#) over university job losses and funding cuts. In December, lecturers at 58 universities [launched a 3-day strike over pensions and pay](#). In May, the Organisation for Economic Co-operation and Development, an international coalition of 38 nations, warned that universities and research institutions around the world must redouble efforts to expand training for PhD students and postdoctoral researchers to [prepare them for jobs outside academia](#). Job insecurity and pressure to publish papers are deep-rooted issues in China, so much so that they [might have been involved in a high-profile killing](#) of a member of the mathematics faculty at Fudan University, Shanghai, in June.

Funding and evaluation

Government support (or otherwise) of science made headlines throughout the year. Researchers in the United States anticipated billions of extra dollars in federal funding from President Joe Biden's administration — a cash injection that [raised concerns about a glut of new PhD students and postdocs](#), as well as [hopes of a resurgence of US research](#) — but scientists elsewhere were coping with disappointment. In October, [researchers in Brazil decried crippling cutbacks in government funding](#), and researchers in the United Kingdom [lamented that funding was growing much more slowly than hoped](#).

The year saw increased scrutiny of the process of researcher evaluation. Citations are often seen as the currency of success, but a study published in February found that [a few elite researchers rack up an outsize share of citations](#). Utrecht University in the Netherlands announced in June that it was doing away with the journal impact factor when [assessing a candidate's publication record for hiring and promotion](#), three months after the University of Liverpool, UK, drew criticism for its use of [metrics to identify academics at risk of losing their positions](#) through job cuts.

The Declaration on Research Assessment, (DORA), a global initiative to reform the evaluation of researchers, announced a new project to develop an [online dashboard that will track hiring and promotion metrics](#) used at institutions around the world. It is due to launch in mid-to-late 2022.

Deciding which researchers deserve grants proved to be complicated. A study of grants awarded by the European Research Council, posted as a preprint in March, suggests that simply sharing an institutional affiliation with certain panellists [gives an applicant a better chance of success](#). A plan by the Australian Research Council to ban any mention of preprints or other non-peer-reviewed publications in grant applications was [quickly abandoned after an outcry](#). A Swiss funder took a simpler approach: [determine grant winners by the luck of the draw](#).

Publishing and politics

Publishing took steps towards improved access and transparency, sometimes with a nudge from the outside. UK Research and Innovation, the country's largest public funder, announced that, starting in April 2022, all grant recipients must [make papers immediately available on open-access platforms after publication](#). To further improve access to research findings, US technologist Carl Malamud created a free, searchable online database of words and phrases from more than 100 million journal articles, [including many currently locked behind paywalls](#).

The forces of global politics continued to affect daily life in the laboratory. Senior research officers in the United States noted that they spend large chunks of time completing [tasks intended to protect research integrity and to prevent espionage](#) by foreign countries, particularly China. Against that backdrop, polls of researchers of Chinese descent based in the United States found [concerns about racial bias](#) that deterred many from collaborating with institutions in China. [Thousands of Chinese scientists have returned to their home country](#), driven partly by unease about China in the countries they worked in. In February, scientists in Russia protested about a new [law that they feared could impede international collaboration](#).

Research integrity remained a hot topic. In September, a coalition of publishers announced an [initiative to crack down on doctored images](#). There's clearly work to do: a poll published in July found that 8% of researchers in the Netherlands [admitted to fabricating data](#). China took steps to [crack down on fake publishers](#), but some scammers still found ways to fool academics, often by impersonating guest editors.

Diversity, equity and inclusion

The move towards greater equality in science picked up momentum in 2021 with renewed calls for action (by [UK Members of Parliament, among others](#)) and [fresh evidence of disparities](#). *Nature*'s salary and satisfaction survey found that only 40% of respondents felt that their [employers were doing enough to promote diversity](#), down from 51% in 2018. One respondent said: “Academics like to think of their community as free spirited and innovative, but there is massive systemic discrimination and power hierarchies that ruin people and careers.”

Gender inequity was a particular focus. An analysis of top science prizes awarded over the past decade found that women were less lauded than men. The pay gap between male and female researchers in North America has widened, and men won an outsize proportion of medical awards in Australia. Last month, researchers at Germany's Max Planck Institute raised concerns about the treatment of women research leaders there. A survey of astronomers found that reports of discrimination and harassment were common, with women and people from minority ethnic groups the most frequent targets.

The issue of racial equality continued to make headlines. A US survey of nearly 20 million workers found only modest gains in the representation of minority ethnic groups in STEM fields (science, technology, engineering and mathematics). An analysis of grants awarded by Cancer Research UK found that members of minority ethnic groups were less likely to receive funding. The American Physics Society announced that it would no longer hold conferences in cities with a history of racist police practices.

Historically Black colleges and universities in the United States, major incubators of scientific talent, anticipated substantial funding increases thanks to a series of legal settlements and promises of increased federal support.

Mental health

In these stressful times, mental health is a priority. The US National Academies of Sciences, Engineering and Medicine urged universities to invest more in graduate mental health. The chair of the committee that produced the report told *Nature* that the problem is “greater than it’s ever been”. *Nature*’s salary and satisfaction survey found that more than 40% of respondents had accessed professional help for anxiety or depression caused by their work, or had wanted to do so. That’s a jump from 2016, when just over 30% had sought or wanted help.

Next year will undoubtedly bring new developments that will change the way science is done. Stay tuned.

Nature **600**, 765-766 (2021)

doi: <https://doi.org/10.1038/d41586-021-03710-0>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03710-0>

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

- TECHNOLOGY FEATURE
- 16 December 2021

Embryo-like models shed fresh light on early human development

A wave of stem-cell systems are enabling researchers to unpick what happens after an embryo implants in the uterus.

- [Sandeep Ravindran](#) ⁰



Embryo models of the blastocyst stage of early development, reprogrammed from skin cells. Credit: Monash University

Cells in culture typically form continuous sheets, like skin. But in 2017, bioengineer Jianping Fu realized that if he cultured human stem cells in a 3D scaffold, they would spontaneously organize into structures that looked, under a microscope, a bit like an embryo¹. Gene-expression analyses suggested the cells were similar to those in an embryo immediately after it implants in the uterus, meaning they could serve as experimental models for a previously opaque point in early development.

“Once the human embryo implants into the maternal uterus, it becomes invisible,” says Fu, at the University of Michigan in Ann Arbor. “That time frame is really a black box.” But it’s crucially important. This stage, about 7 to 10 days after fertilization in humans, is marked by formation of the amniotic sac and the first signs of the primitive streak. That structure marks the point at which the embryo sets up the body axes and begins to distinguish head from tail and left from right.

Researchers have long sought to observe and study these developmental stages. But working with human embryos has always been technically and

ethically fraught. Animal models go only so far in mimicking humans. Natural human embryos, donated by people undergoing fertility treatment, are hard to come by. And until May this year, scientists were barred from culturing such embryos in the laboratory for longer than two weeks after fertilization. The International Society for Stem Cell Research (ISSCR) relaxed this 14-day rule in May, allowing research groups in countries where such work is legal to apply for permission to continue studies beyond 14 days.

Given these limitations, it is no surprise that researchers have sought alternatives to natural human embryos in the lab. Stem cells from species such as mice have long provided a replacement — either embryonic stem cells (taken directly from early-stage embryos) or induced pluripotent stem cells (reprogrammed from adult cells). Now, a new wave of human embryo-like models is catching up. “If we want to understand human-specific features, we really need to look at a human system,” says Naomi Morris, a developmental biologist at the Francis Crick Institute in London.

Various embryo-like models exist. They include blastoids, which mimic the stages before the embryo implants into the uterus, and gastruloids, which model formation of the body plan and organ precursors. Others include tissues that surround the embryo, such as the yolk sac and amniotic cavity (the fluid-filled sac that encloses the embryo), and some recapitulate the early steps of brain, spinal-cord and heart formation, which occurs around day 22 of human development.

Researchers are still benchmarking how similar these models are to human embryos — the greater the similarity, the greater the technical and ethical challenges. But they open up new experimental approaches. “For the first time, we are able to do genetic screens and drug screens on structures that are remarkably similar to the embryo,” says Nicolas Rivron, a developmental biologist at the Institute of Molecular Biotechnology of the Austrian Academy of Sciences in Vienna. Such screens could help to identify treatments for developmental disorders or infertility. The models could also help to tease apart how organs form, potentially leading to advances in artificial transplants, for instance.

“We are using these models to study embryology from a bottom-up perspective,” says Susanne van den Brink, a developmental biologist at Pompeu Fabra University in Barcelona, Spain. “It’s like studying how a car works by starting at the smallest pieces and putting it together from scratch,” she says.

Pre-implantation

Among the earliest developmental stages is a hollow sphere of cells called a blastocyst, which develops before implantation, when the human embryo comprises about 100 cells.

Before 2018, models of the blastocyst stage were still rudimentary. So, by combining two types of mouse stem cell in microwell arrays, Rivron and his team developed the first mouse blastoids — self-aggregating structures that resemble 3.5-day-old blastocysts². Formed from embryonic stem cells and trophoblast stem cells, the blastoids revealed that signals from embryonic stem cells regulate and induce the development of trophoblast cells — the progenitor of the placenta. (This latter group of cells has a crucial role in mediating uterine implantation.) “This is a little counterintuitive, because during later development, it’s the placenta that takes care of the embryo,” says Rivron. “But actually, in these early stages, it’s the opposite.”

Several teams translated the blastoid model to human cells by growing stem cells in a 3D inverted-pyramid culture system called AggreWell, available from Stemcell Technologies in Vancouver, Canada. They reported their findings this year. For instance, stem-cell biologist Jun Wu at the University of Texas Southwestern Medical Center in Dallas and his team built blastoids out of human embryonic stem cells by treating them with chemical factors that induce the signalling needed for blastocyst development. They found that the same method worked on human induced pluripotent stem cells generated from reprogrammed skin cells³. Stem-cell biologist Xiaodong Liu and colleagues at Monash University in Melbourne, Australia, likewise came up with a method to form human blastoids from human skin cells⁴. And Rivron and his team have developed a human blastoid model⁵.

Unlike some models of human development such as Fu's, blastoids contain all three cell lineages found in natural human embryos — not just the cells that form the embryo itself but also the 'extra-embryonic' ones that lead to the yolk sac and placenta. "Human blastoids represent the first complete or integrated embryo models established from cultured stem cells," says Wu.

Using these, Wu has begun exploring the molecular signals that drive human development. By treating his blastoids with inhibitors of different versions of an enzyme called protein kinase C, for instance, he has identified which specific isoforms of the enzyme drive formation of the blastocyst cavity, the fluid-filled space at its centre³. "Questions concerning the molecular and cellular crosstalk between embryonic and extra-embryonic tissues during early development can only be answered with integrated models such as blastoids," he says.

Implantation and beyond

Multiple groups — including those of Magdalena Zernicka-Goetz, a developmental and stem-cell biologist at the University of Cambridge, UK, and developmental biologist Yang Yu of Peking University Third Hospital in Beijing — have used human blastocyst models to look at the stages before and immediately beyond implantation (around 7–10 days post-fertilization)^{6,7}.

These models use extended or expanded pluripotent stem cells, which are derived by reprogramming stem cells so that they can generate both embryonic and extra-embryonic cell lineages. Growing these cells in 3D inverted pyramid microwell plates with a combination of growth and other chemical factors allowed the cells to aggregate into blastocyst-like structures.

Both groups then extended these models into the early post-implantation stage by implanting them in a 3D scaffold called Matrigel that mimics the protein and chemical environment surrounding the cells. "[Matrigel] provides the signals to allow those stem cells to organize, and it allows us to mimic the transition that happens when an embryo implants in the body of the mother," says Zernicka-Goetz.

Fu extended his post-implantation embryo models by using bioengineered microfluidics systems made of three channels — a central gel mimicking the uterine wall, and one channel each for delivering stem cells and chemical signals in a precisely controlled manner. The system was sufficiently scalable to produce hundreds or even thousands of embryo-like structures, he and his team reported in 2019⁸. This microfluidics approach also increased efficiency and reproducibility, he says, making the models particularly compelling for translational applications such as drug and toxicity screening.

Fu used his models to demonstrate that certain aspects of human embryonic development, such as differentiation of the cells that form the innermost layer of the placenta (the amnion) from stem cells, are particularly sensitive to the mechanical rigidity and 3D nature of the matrix on which they are grown⁸. Changing the thickness of the gel matrix or growing it in 2D instead of 3D culture prevented proper amnion differentiation. “Such knowledge can only be generated, as of this moment, from *in vitro* models, because *in vivo*, it’s very hard to conduct such mechanistic studies,” he says.

Going past day 14

In the third week of human development (14–22 days post-fertilization), the main body pattern emerges in a process called gastrulation. So do the three ‘germ layers’ — the endoderm that gives rise to the gut and internal organs, the mesoderm that forms muscle and connective tissue, and the ectoderm, which creates the skin and nervous system. “These germ layers are the precursors of all organs and every cell type in the body, so it is really crucial to understand how the cellular differentiation occurs,” says Berna Sozen, a developmental biologist at Yale School of Medicine in New Haven, Connecticut.

Developmental biologists Ali Brivanlou and Eric Siggia at the Rockefeller University in New York City developed the earliest gastrulation models in 2014⁹. The models “were mimicking a period of our development that we had never ever seen before, and this was one of the most beautiful things”, says Brivanlou.

But those were 2D structures. Van den Brink developed the first mouse 3D models of gastrulation, called gastruloids, while working with developmental biologist Alfonso Martinez Arias at the University of Cambridge. The researchers discovered that growing small aggregates of mouse stem cells with the correct signals, such as activators of the Wnt signalling pathway, caused them to differentiate into embryo-like structures resembling the gastrulation stage¹⁰. “They form a head–tail axis, the back–belly axis and the left–right axis, and they form all organ progenitors in the right location,” van den Brink says.



Gastruloid embryo-like models are used to study formation of the body and organs. Credit: Naomi Moris

Researchers had previously thought that, to form the head–tail axis, the embryo had to be surrounded by and receive signals from extra-embryonic tissues, which gastruloids lack, she says. “That’s a very nice example of how these models are starting to challenge our textbooks on embryology.”

In 2020, while working as a graduate student with Martinez Arias, Moris translated that technology to human gastruloids, finding that their patterns of

gene expression roughly “followed the head-to-tail axis in these structures”, she says.

By comparing gastruloid morphology to existing curated collections of human embryos, Moris was able to place her model system at about day 20–21 days post-fertilization, when structures called somites, which give rise to the vertebral column and some muscles, begin to appear¹¹. “Not only were these structures recapitulating something that we knew happened in mammalian development, but it also let us kind of put a pin really where they were in the timeline of human development,” she says. These models could therefore be useful in replicating these later stages of human development.

Researchers have now pushed gastruloid models even further. Brivanlou, for instance, has developed 3D models of early human-brain tissue, dubbed neuruloids¹². By embedding gastruloids in Matrigel, van den Brink generated mouse-embryo-like models that form somites¹³. And Jesse Veenvliet, a developmental biologist at the Max Planck Institute of Molecular Cell Biology and Genetics in Dresden, Germany, generated mouse embryo-like models with somites and a neural tube — the structure that eventually forms the spinal cord¹⁴.

“We can use the original gastruloid that doesn’t have a neural tube as a blank slate and then start to add cues on top of it, so we can really start to properly dissect the minimal inputs that are needed to obtain mammalian embryo-like architecture,” says Veenvliet. Researchers can even capture embryonic development as it unfolds, live on camera, he says.

Varied applications

These tools could help researchers to understand how development goes awry. Creating human embryo-like models from a person’s own skin cells could allow scientists to explore the impacts of that individual’s genetics on development, for instance.

Researchers such as Rivron are using organoid models of the womb lining to develop detailed implantation models, which could help to reveal causes of

infertility and early pregnancy loss, and could lead to new therapies. “We can recapitulate now, in a dish, those processes that normally happen very well hidden into the womb,” says Rivron.

Human embryo models could also be used to screen for drug toxicity. In a proof-of-concept study, Moris has shown that the drug thalidomide — infamous for causing birth defects in humans, but not particularly toxic in mice — had a stronger effect on human gastruloids than on mouse gastruloids, suggesting that these models could be useful for human-specific toxicological screens¹⁵.

Human embryo-like models could even one day lead to advances in regenerative medicine, such as the ability to generate artificial organs and tissues for transplantation, says Fu. Models based on patient-derived pluripotent stem cells could grow artificial organs that the body recognizes as its own. “That’s really the holy grail of regenerative medicine,” says Fu.

Technical and ethical challenges

Embryo-like models can generally be made in 3–9 days using off-the-shelf consumables and protocols, and standard cell-culture skills. “You have to plate cells and then give them the right signals, but other than that, it’s mostly just waiting,” says van den Brink. Efficiency is often low, with only 5–20% of cell colonies developing into fully formed models, but some strategies do better. Fu’s microfluidics-based models, for instance, can exceed 90% efficiency, although they require specialized bioengineering expertise.

Start with high-quality cells, advises Rivron. “Having pristine stem cells to start with is absolutely crucial.” That means using cells that haven’t been cultured for too long, and keeping them as healthy as possible. “Culturing them is kind of an art,” he says.

Existing protocols provide a good starting point, but be prepared to fine-tune experimental conditions for optimal results. “If you’re experienced in cell culture, then you can learn how to make gastruloids in about two weeks,” says van den Brink.

Once they're grown, be sure to compare your models to natural embryos using both microscopy and single-cell gene-expression analysis, to ensure they are what they appear to be. "Looks can be deceiving," says Rivron. "You have to be able to understand the cells that you form and benchmark them to the cells of the embryo," he says.

Such comparisons become more difficult past day 14, because such experiments were completely banned until earlier this year, and still remain technically challenging. But that could change as techniques improve and labs apply for exceptions to the day-14 rule. "It would be very useful and would definitely allow us to validate the human gastruloid system," says van den Brink.

But ethical considerations also become more complicated past day 14. The latest ISSCR guidelines differentiate between models that contain only a part of the embryo — such as gastruloids — and those that contain all three cell lineages, such as blastoids. The latter models are subject to the same intensive review process as natural embryos before they can be studied past day 14¹⁶. "There's also a consensus that none of the models should be used for human reproduction purposes, and none of the models, whether they are complete or not, should be used for implantation," says Fu, who helped to draft the guidelines.

As human models of embryogenesis improve, further ethical discussions are only going to become more important. "This is moving forward now with a speed that I could have not even imagined a few years ago," says Brivanlou.

Nature **600**, 767-769 (2021)

doi: <https://doi.org/10.1038/d41586-021-03711-z>

References

1. 1.

Shao, Y. *et al.* *Nature Commun.* **8**, 208 (2017).

2. 2.

Rivron, N. C. *et al.* *Nature* **557**, 106–111 (2018).

3. 3.

Yu, L. *et al.* *Nature* **591**, 620–626 (2021).

4. 4.

Liu, X. *et al.* *Nature* **591**, 627–632 (2021).

5. 5.

Kagawa, H. *et al.* *Nature* <https://doi.org/10.1038/s41586-021-04267-8> (2021).

6. 6.

Fan, Y. *et al.* *Cell Discov.* **7**, 81 (2021).

7. 7.

Sozen, B. *et al.* *Nature Commun.* **12**, 5550 (2021).

8. 8.

Zheng, Y. *et al.* *Nature* **573**, 421–425 (2019).

9. 9.

Warmflash, A., Sorre, B., Etoc, F., Siggia, E. D. & Brivanlou, A. H. *Nature Meth.* **11**, 847–854 (2014).

10. 10.

van den Brink, S. C. *et al.* *Development* **141**, 4231–4242 (2014).

11. 11.

Moris, N. *et al.* *Nature* **582**, 410–415 (2020).

12. 12.

Haremaki, T. *et al.* *Nature Biotechnol.* **37**, 1198–1208 (2019).

13. 13.

van den Brink, S. C. *et al.* *Nature* **582**, 405–409 (2020).

14. 14.

Veenvliet, J. V. *et al.* *Science* **370**, eaba4937 (2020).

15. 15.

Mantziou, V. *et al.* *Reprod. Toxicol.* **105**, 72–90 (2021).

16. 16.

Lovell-Badge, R. *et al.* *Stem Cell Rep.* **16**, 1398–1408 (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03711-z>

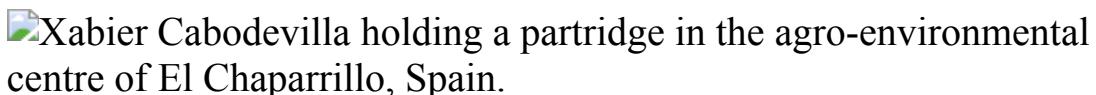
| [Section menu](#) | [Main menu](#) |

- WHERE I WORK
- 17 December 2021

A partridge in hand on the Spanish steppe

As part of his PhD research into the effects of farming and hunting on endangered bird populations in Spain, Xabier Cabodevilla tracked birds and collected faecal samples from roosting sites.

- [Jack Leeming](#)



Xabier Cabodevilla recently completed his PhD at the University of the Basque Country in Vitoria-Gasteiz, Spain, and the Spanish Institute for Game and Wildlife Research in Ciudad Real. Credit: Guillermo Gutierrez Carrasca for *Nature*

In this photograph, taken on a late summer's day in the region of Campo de Calatrava in central Spain, I'm tagging a red-legged partridge (*Alectoris rufa*) with a GPS tracker to help me to monitor the impact of human activity on these and other steppe birds' lives. I also study two sandgrouse species (*Pterocles alchata* and *Pterocles orientalis*) and two types of bustard (*Tetrax tetrax* and *Otis tarda*).

The outlook isn't great for the birds I research. Red-legged partridges are native to the Iberian Peninsula, as well as to some parts of France and Italy, but their populations are falling rapidly owing to activities such as intensive agriculture and hunting. In central Spain, populations of the birds have declined by around 50% in less than a decade, according to my PhD research.

Partridge fieldwork isn't as easy as this picture suggests. Generally, I prefer to track birds from a distance, following them into the fields with a telescope or a pair of binoculars, staying far enough away so as not to spook them, and waiting until they roost on the ground in the evening. This can mean long days tracking the birds under a blazing hot Spanish sun.

In the morning I'll come back to collect faecal samples from where we spotted the birds roosting. Finding partridge poo is also hard: sometimes I start at 6 a.m. and don't find anything until the afternoon, even if I have a pretty good sense of where they were the previous day.

I completed my PhD in September and am looking for a postdoctoral research position where I can continue studying. It's a little stressful, but I'm trying to keep things in perspective and stay relaxed about my future: it's been a difficult year finalizing my degree and publishing papers. Often, among researchers, our work is also our hobby; it can be difficult to switch off. We'll see where science takes me next year.

Nature **600**, 772 (2021)

doi: <https://doi.org/10.1038/d41586-021-03712-y>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03712-y>.

Research

- **[Alternatives to standard quantum theory ruled out](#)** [15 December 2021]
News & Views • Standard quantum theory contains square roots of negative numbers. But how essential are these ‘imaginary’ numbers? A way of disproving analogous theories that omit them has been proposed — and confirmed experimentally.
- **[Stomach cancer gets a triple punch of therapy](#)** [15 December 2021]
News & Views • Harnessing immune cells to target tumours is a growing trend. The results of a clinical trial combining such treatment with other standard therapies for gastric cancer have altered medical practice — and more changes are to come.
- **[Low-power light modifies electron microscopy](#)** [22 December 2021]
News & Views • An optical device designed to control the properties of electron waves inside an electron microscope demonstrates that clever platforms for integrated photonics need not be powered by expensive laser systems.
- **[Gut clues to weight gain after quitting smoking](#)** [08 December 2021]
News & Views Forum • Research has uncovered factors that underlie the weight gain associated with cessation of smoking. Here, scientists consider the implications of this finding from the perspectives of gut biology and of smoking.
- **[Clever substitutions reveal magnetism in zigzag graphene nanoribbons](#)** [22 December 2021]
News & Views • The inclusion of nitrogen atoms stabilizes the zigzag edges of carbon-based nanoribbons, enabling the ribbons to be decoupled from a substrate and providing a probe for their unconventional magnetism.
- **[What surveys really say](#)** [08 December 2021]
News & Views • Increasing the sample size of a survey is often thought to increase the accuracy of the results. However, an analysis of big surveys on the uptake of COVID-19 vaccines shows that larger sample sizes do not protect against bias.
- **[Gravity, AlphaFold and neural interfaces: a year of remarkable science](#)** [17 December 2021]
News & Views • Highlights from News & Views published this year.

- **Very-high-frequency oscillations in the main peak of a magnetar giant flare** [22 December 2021]

Article • Two very-high-frequency quasi-periodic oscillations (at 2,132 Hz and 4,250 Hz) are detected within the initial hard spike of a magnetar giant flare originating from the galaxy NGC 253, and detailed temporal and spectral analyses are performed.

- **Quantum theory based on real numbers can be experimentally falsified** [15 December 2021]

Article • A Bell-like experiment that discriminates between real-number and complex-number multipartite quantum systems could disprove real quantum theory.

- **Programmable interactions and emergent geometry in an array of atom clouds** [22 December 2021]

Article • The reported network of connectivity between atomic ensembles is entirely programmable and independent of its geometrical arrangement, because of the interaction with an optical cavity.

- **Resonance from antiferromagnetic spin fluctuations for superconductivity in UTe2** [22 December 2021]

Article • Inelastic neutron scattering measurements show that superconductivity in UTe2 is associated with a resonance near antiferromagnetic order that suggests an unexpected spin-singlet component to the electron pairing.

- **Quantum anomalous Hall effect from intertwined moiré bands** [22 December 2021]

Article • An electric-field-induced topological phase transition from a Mott insulator to a quantum anomalous Hall insulator in near-60-degree-twisted (or AB-stacked) MoTe2/WSe2 heterobilayers is reported.

- **Spin splitting of dopant edge state in magnetic zigzag graphene nanoribbons** [22 December 2021]

Article • Decoupling spin-polarized edge states using substitutional N-atom dopants along the edges of a zigzag graphene nanoribbon (ZGNR) reveals giant spin splitting of a N-dopant edge state, and supports the predicted emergent magnetic order in ZGNRs.

- **Integrated photonics enables continuous-beam electron phase modulation** [22 December 2021]

Article • A silicon nitride microresonator is used for coherent phase modulation of a transmission electron microscope beam, with future applications in combining high-resolution microscopy with spectroscopy, holography and metrology.

- **Dynamic spatial progression of isolated lithium during battery operations** [22 December 2021]

Article • An electrochemical process stimulates the progression toward the electrode of isolated or ‘dead’ lithium in a battery, recovering its electrical connection, and the effect is demonstrated by increased cycle life.

- **High-entropy polymer produces a giant electrocaloric effect at low fields** [22 December 2021]

Article • A study reports and characterizes a high-entropy electrocaloric polymer that switches under low fields, and discusses its potential suitability for use in caloric heat pumps.

- **Rapid microbial methanogenesis during CO₂ storage in hydrocarbon reservoirs** [22 December 2021]

Article • Microbial methanogenesis converts up to 19% of the carbon dioxide injected into an oil field to methane, suggesting that microbial methanogenesis may be a globally important subsurface process.

- **The power of genetic diversity in genome-wide association studies of lipids** [09 December 2021]

Article • A genome-wide association meta-analysis study of blood lipid levels in roughly 1.6 million individuals demonstrates the gain of power attained when diverse ancestries are included to improve fine-mapping and polygenic score generation, with gains in locus discovery related to sample size.

- **Cortical responses to touch reflect subcortical integration of LTMR signals** [17 November 2021]

Article • Genetic manipulation of skin peripheral sensory neurons in mice shows that cortical neuron responses to touch reflect subcortical mixing of signals from both rapidly adapting and slowly adapting low-threshold mechanoreceptors.

- **Mechanical actions of dendritic-spine enlargement on presynaptic exocytosis** [24 November 2021]

Article • A mechanism of mechanosensation and transduction in the presynaptic boutons is identified, in which sensing of fine pressure leads to enhanced neurotransmitter release.

- **Collective durotaxis along a self-generated stiffness gradient in vivo** [08 December 2021]

Article • The neural crest of *Xenopus laevis* self-generates a stiffness gradient in the adjacent placodal tissue and follows this gradient by durotaxis.

- **Unrepresentative big surveys significantly overestimated US vaccine uptake** [08 December 2021]

Article • An analysis of three surveys of COVID-19 vaccine behaviour shows that larger surveys overconfidently overestimated vaccine uptake, a demonstration of how larger sample sizes can paradoxically lead to less accurate estimates.

- **Immunogenicity and efficacy of heterologous ChAdOx1–BNT162b2 vaccination** [21 October 2021]

Article • An observational study of more than 13,000 healthcare workers shows that a prime–boost vaccine regimen of AstraZeneca ChAdOx1-S-nCoV-19 with Pfizer BNT162b2 provides enhanced protection against SARS-CoV-2 infection compared with two doses of BNT162b2.

- **Antigen-presenting innate lymphoid cells orchestrate neuroinflammation** [01 December 2021]

Article • A subset of inflammatory group 3 innate lymphoid cells, here termed iILC3s, infiltrate the central nervous system and promote neuroinflammation and disease progression in a mouse model of multiple sclerosis.

- **Gut microbiota modulates weight gain in mice after discontinued smoke exposure** [08 December 2021]

Article • A study of mice exposed to cigarette smoke suggests that smoking-cessation-induced weight gain is associated with a dysbiotic state that is driven by smoking-related metabolites.

- **A hormone complex of FABP4 and nucleoside kinases regulates islet function** [08 December 2021]

Article • Hormonal FABP4 is discovered to be a pivotal regulator of an adipose–beta-cell endocrine axis that coordinates energy status and metabolic organ function, and targeting this axis improved metabolic outcomes.

- **The KEYNOTE-811 trial of dual PD-1 and HER2 blockade in HER2-positive gastric cancer** [15 December 2021]

Article • Interim analysis of a phase III clinical trial of HER2-positive gastric adenocarcinoma shows pembrolizumab plus trastuzumab and chemotherapy improves response rates compared with trastuzumab and chemotherapy alone.

- **ecDNA hubs drive cooperative intermolecular oncogene expression** [24 November 2021]

Article • Extrachromosomal DNA (ecDNA) congregates in clusters called ecDNA hubs that promote intermolecular interactions between gene-regulatory regions and thereby amplify the expression of oncogenes such as MYC in cancer cell lines.

- **Sex-specific chromatin remodelling safeguards transcription in germ cells** [08 December 2021]

Article • Following global DNA demethylation, mouse gonadal primordial germ cells undergo remodelling of repressive chromatin modifications, resulting in a sex-specific signature that is required to safeguard the transcriptional program.

- **A conserved mechanism for regulating replisome disassembly in eukaryotes** [26 October 2021]

Article • A conserved mechanism for the regulation of replisome disassembly in eukaryotes is shown using cryo-electron microscopy, revealing a role for DNA in the preservation of replisome integrity.

- **Activation of homologous recombination in G1 preserves centromeric integrity** [01 December 2021]

Article • Centromeres are able to recruit the homologous recombination machinery during G1 via CENP-A and HJURP, thereby preserving centromeric integrity even in the absence of a sister chromatid.

- **β -NAD as a building block in natural product biosynthesis**

[08 December 2021]

Article • β -Nicotinamide adenine dinucleotide is used as a building block in secondary metabolite biosynthetic pathways for a novel class of natural products.

- **Structures of the σ 2 receptor enable docking for bioactive ligand discovery** [08 December 2021]

Article • Crystal structures of the σ 2 receptor are determined and used to perform a docking screen of nearly 500 million molecules, identifying σ 2-selective ligands and providing insight into the role of σ 2 in neuropathic pain.

- NEWS AND VIEWS
- 15 December 2021

Alternatives to standard quantum theory ruled out

Standard quantum theory contains square roots of negative numbers. But how essential are these ‘imaginary’ numbers? A way of disproving analogous theories that omit them has been proposed — and confirmed experimentally.

- [William K. Wootters](#) 0

The square root of a negative number could easily be regarded as a curious mathematical construction that has little bearing on real life. Indeed, the French polymath René Descartes gave such numbers the derogatory name ‘imaginary’, reserving the term ‘real numbers’ for those that occur on an ordinary number line. However, since the development of quantum theory nearly a century ago, it has seemed likely that complex numbers — which combine imaginary and real numbers — are deeply embedded in the structure of the physical world. Indeed, the imaginary unit that represents the square root of minus one occurs prominently in the basic equations of quantum theory. [Writing in Nature](#), Renou *et al.*¹ show how experiments could strengthen the evidence that complex numbers are key to describing the quantum world, by eliminating a class of theory that contains only real numbers. And two versions of the experiment they propose have already been implemented^{2,3}.

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Nature **600**, 607-608 (2021)

doi: <https://doi.org/10.1038/d41586-021-03678-x>

References

1. 1.

Renou, M.-O. *et al. Nature* <https://doi.org/10.1038/s41586-021-04160-4> (2021).

2. 2.

Chen, M.-C. *et al.* *Phys. Rev. Lett.* (in the press); preprint at
<https://arxiv.org/abs/2103.08123>

3. 3.

Li, Z.-D. *et al.* *Phys. Rev. Lett.* (in the press).

4. 4.

Schrödinger, E. *Math. Proc. Camb. Phil. Soc.* **31**, 555–563 (1935).

5. 5.

Pál, K. F. & Vértesi, T. *Phys. Rev. A* **77**, 042105 (2008).

6. 6.

McKague, M., Mosca, M. & Gisin, N. *Phys. Rev. Lett.* **102**, 020505 (2009).

7. 7.

Moroder, T., Bancal, J.-D., Liang, Y.-C., Hofmann, M. & Gühne, O. *Phys. Rev. Lett.* **111**, 030501 (2013).

8. 8.

Dyson, F. J. *J. Math. Phys.* **3**, 1199–1215 (1962).

9. 9.

Boaz, J. C. *Found. Phys.* **42**, 819–855 (2012).

10. 10.

Stueckelberg, E. C. G. *Helv. Phys. Acta* **33**, 727–752 (1960).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03678-x>

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

- NEWS AND VIEWS
- 15 December 2021

Stomach cancer gets a triple punch of therapy

Harnessing immune cells to target tumours is a growing trend. The results of a clinical trial combining such treatment with other standard therapies for gastric cancer have altered medical practice — and more changes are to come.

- [Myriam Chalabi](#) 0

Gastric cancer is a highly aggressive disease. [Writing in Nature](#), Janjigian *et al.*¹ present clinical data that point to a way to improve outcomes for people who have this type of tumour.

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Nature **600**, 608-609 (2021)

doi: <https://doi.org/10.1038/d41586-021-03458-7>

References

1. 1.

Janjigian, Y. Y. *et al.* *Nature* <https://doi.org/10.1038/s41586-021-04161-3> (2021).

2. 2.

Janjigian, Y. Y. *et al.* *Lancet* **398**, 27–40 (2021).

3. 3.

Baselga, J. *et al.* *J. Clin. Oncol.* **14**, 737–744 (1996).

4. 4.

Slamon, D. J. *et al.* *N. Engl. J. Med.* **344**, 783–792 (2001).

5. 5.

Bang, Y.-J. *et al.* *Lancet* **376**, 687–697 (2010).

6. 6.

Gall, V. A. *et al.* *Cancer Res.* **77**, 5374–5383 (2017).

7. 7.

Shitara, K. *et al.* *N. Engl. J. Med.* **382**, 2419–2430 (2020).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03458-7>

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

- NEWS AND VIEWS
- 22 December 2021

Low-power light modifies electron microscopy

An optical device designed to control the properties of electron waves inside an electron microscope demonstrates that clever platforms for integrated photonics need not be powered by expensive laser systems.

- [Martin Kozák](#) ⁰

The wavelength of an electron can be up to 100,000 times shorter than that of a photon, which means that microscopes that use electrons to illuminate a sample are able to resolve much smaller structures than can those that use light. But electrons and photons also work together: interactions with photons can be used to modulate the wave-like nature of electrons, changing the energy spectrum of the electron beam in a way that could be useful for tailored microscopy¹. However, these electron–photon interactions are weak, and usually require high-power laser sources. [Writing in Nature](#), Henke *et al.*² report an optical platform for manipulating the properties of electron waves using a light source that is not much more powerful than an average laser pointer. The low light power required of this device, together with its integrated design, make it readily applicable to many existing electron microscopes.

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Nature **600**, 610–611 (2021)

doi: <https://doi.org/10.1038/d41586-021-03767-x>

References

1. 1.

Feist, A. *et al.* *Nature* **521**, 200–203 (2015).

2. 2.

Henke, J.-W. *et al.* *Nature* **600**, 653–658 (2021).

3. 3.

Kapitza, P. L. & Dirac, P. A. M. *Math. Proc. Camb. Phil. Soc.* **29**, 297–300 (1933).

4. 4.

Freimund, D. L., Aflatooni, K. & Batelaan, H. *Nature* **413**, 142–143 (2001).

5. 5.

Barwick, B., Flannigan, D. J. & Zewail, A. H. *Nature* **462**, 902–906 (2009).

6. 6.

Ryabov, A., Thurner, J. W., Nabben, D., Tsarev, M. V. & Baum, P. *Sci. Adv.* **6**, eabb1393 (2020).

7. 7.

Breuer, J. & Hommelhoff, P. *Phys. Rev. Lett.* **111**, 134803 (2013).

8. 8.

Dahan, R. *et al.* *Science* **373**, eabj7128 (2021).

9. 9.

Wang, K. *et al.* *Nature* **582**, 50–54 (2020).

10. 10.

Gover, A. & Yariv, A. *Phys. Rev. Lett.* **124**, 064801 (2020).

11. 11.

Kfir, O., Di Giulio, V., García de Abajo, F. J. & Ropers, C. *Sci. Adv.* **7**, eabf6380 (2021).

12. 12.

Krivanek, O. *et al.* *Nature* **514**, 209–212 (2014).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03767-x>

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

- NEWS & VIEWS FORUM
- 08 December 2021

Gut clues to weight gain after quitting smoking

Research has uncovered factors that underlie the weight gain associated with cessation of smoking. Here, scientists consider the implications of this finding from the perspectives of gut biology and of smoking.

- [Matthew P. Spindler](#)⁰,
- [Jeremiah J. Faith](#)¹,
- [Junshi Wang](#)² &
- [Paul J. Kenny](#)³

THE PAPER IN BRIEF

- People who stop smoking usually go on to gain weight, which can cause some individuals to start smoking again.
- The factors underlying this type of weight gain are not fully understood.
- Microorganisms in the gut (termed the microbiota) can influence aspects of human health, and have been implicated in obesity.
- [Writing in Nature](#), Fluhr *et al.*¹ report evidence from studies in mice and humans that point to the microbiota as having a role in weight gain associated with the cessation of smoking.
- These findings might aid efforts to enable people to avoid side effects associated with quitting smoking.

MATTHEW P. SPINDLER & JEREMIAH J. FAITH: Gut microbial mischief

Cigarette smoking is the leading preventable cause of death worldwide, and yet many smokers never attempt to quit^{2,3}. The weight gain associated with cessation of smoking is cited as a major reason why more people don't try to stop^{2,3}. This weight gain is broadly attributed to smoking-associated effects on energy intake, metabolic rate and physical activity, but the specific underlying molecular mechanisms are not understood. Fluhr and colleagues now provide evidence that implicates gut microorganisms in this phenomenon.

The authors established a mouse model that replicates features of the weight change that occurs after smoking cessation in humans. These animals gained less weight during smoke exposure, and their weight returned to the non-smoking baseline after exposure ceased. The authors demonstrate that microbiota-dependent factors affect how much weight is regained. They found that administering antibiotics reduced the amount of weight regained, which suggests that a bacterial component of the microbiota targeted by antibiotics contributes to the process. This effect on the animals' weight was maintained for weeks after antibiotic administration was halted, and the results were unaffected by changes in diet or differences in the original microbiota of mice obtained from various vendors.

Fluhr and colleagues then carried out experiments to pinpoint the effect of smoke-associated microbiotas on weight change. Mice that lacked their natural microbiota (germ-free animals) and that received a transplant of faecal microbiota from mice exposed to smoke gained more weight than did animals transplanted with control microbiotas from mice not exposed to smoke. This type of weight gain was observed consistently for two recipient strains of mice lacking their natural microbiota and on two different diets, suggesting that smoke-associated microbiotas directly affect weight gain.

However, several aspects remain to be addressed before these observations can be generalized to humans. Studies in humans that compare the microbiota of smokers and non-smokers have yielded disparate results —

some research indicates that the microbiota is perturbed, whereas other studies find no difference^{4–6}. Fluhr and colleagues show that the microbial composition of faeces differs between mice exposed to smoke and unexposed mice, but such a distinction is less clear in human studies. Furthermore, the community of bacteria that make up the microbiota varies tremendously between individuals. Experiments establishing whether the microbiota-dependent modulation of this weight-loss phenomenon is robust, despite the interpersonal variability of the human microbiota, would provide much-needed insight into the generalizability of these observations. To gather such data, one could perform a retrospective analysis of clinical data to assess whether weight changes in recipients who received faecal-microbiota transplants from smoking donors differed from weight changes in those who received transplants from non-smoking donors.

Groups of people in the population who are most at risk of greater weight gain after quitting smoking include women; those on a low income or with a poor diet; people whose physical activity is limited; and those who have been heavy tobacco users^{3,4}. It is therefore worth noting that although insights using simple model systems can revolutionize how we think about the weight gain associated with the cessation of smoking, efforts to extend these insights from models to clinical practice should take into consideration the complex interplay of factors that influence individual behaviour linked to lifestyle modifications.

JUNSHI WANG & PAUL J. KENNY: Smoke clears on weight gain

The saying, ‘you are what you eat’ implies that what we consume influences our health. Fluhr and colleagues’ results go one step further by suggesting that cigarette smoke can also have an effect, acting on the gut microbiota to facilitate the extraction of energy from our food.

Tobacco smokers tend to have lower body weights than non-smokers, a situation that is reversed when smokers quit the habit⁷. Concerns about weight gain discourage many smokers from quitting, which in turn increases their risk of developing smoking-related diseases.

Nicotine is the major reinforcing component of tobacco that drives the smoking habit⁸. It stimulates nicotinic acetylcholine receptors (nAChRs) located on cells that function in reward circuits in the brain to motivate tobacco use⁹. Nicotine also stimulates nAChRs on pro-opiomelanocortin neurons and other appetite-suppressing neurons in a brain region called the hypothalamus to suppress food intake¹⁰. The general assumption was that the weight gain in ex-smokers is a consequence of the opposite process — decreased activity of these hypothalamic neurons, and perhaps increased activity of pro-feeding circuits in the brain.

However, Fluhr and colleagues show instead that, compared with microbiotas not exposed to tobacco smoke, a microbiota exposed to smoke produces a greater abundance of metabolite molecules that aid the extraction of energy from food by the gut and that promote weight gain on cessation of exposure. Most notably, the molecules associated with this weight gain are derivatives of the amino acid glycine, such as dimethylglycine (DMG). When mice were put on a diet deficient in choline sulfate, a molecule needed for the synthesis of DMG and related compounds, they were less prone to weight gain associated with smoking cessation.

Unexpectedly, this ability of tobacco smoke to modify the microbiota occurred independently of the nicotine content of the smoke, and the effects on the microbiota were not reproduced if mice were given nicotine instead of being exposed to cigarette smoke. This finding suggests that non-nicotine components of tobacco are responsible for remodelling the gut microbiota to enhance the extraction of energy from food and promote weight gain. It challenges the idea that nAChRs in the hypothalamus and other brain sites involved in maintaining energy homeostasis are solely responsible for the actions of cigarettes on appetite, metabolism and the regulation of body weight.

Fluhr and colleagues' findings have major implications for our understanding of the mechanisms that underlie smoking-related diseases. For example, glycine and certain other amino acids can function as neurotransmitter molecules that aid communication between neurons, or that serve as key precursors in the synthesis of neurotransmitters. Given that

cigarette smoke influences the microbiota-mediated production of glycine derivatives, this raises the possibility that these molecules can enter the bloodstream and subsequently gain access to sites in the central nervous system that are involved in appetite regulation and energy metabolism. It is also possible that microbiota-generated inhibitory or excitatory neurotransmitters act in the gut itself to modify the activity of local sensory neurons. This could therefore influence the transmission of information to brain centres involved in energy homeostasis and other processes relevant to tobacco use, such as the desire to smoke.

It is possible that the gut microbiotas of smokers more efficiently harvest not only nutrients, but other substances, too. For example, it will be interesting to discover whether the amounts of nicotine extracted from tobacco smoke, acting in the brain to perpetuate the tobacco habit, or nicotine-derived bioactive molecules (metabolites), are influenced by the gut microbiota or by the microbiota of organs such as the lungs. By extension, the absorption, metabolism and distribution of substances that are often taken together with tobacco products, particularly alcohol, might be affected by the actions of cigarette smoke on the microbiota.

These findings offer possibilities to consider from a therapeutic perspective. Dietary supplements that act in the gut, such as the choline sulfate needed to generate DMG and related glycine derivatives, or precursors of other microbiota-generated bioactive metabolites, might limit weight gain in abstinent smokers and encourage smokers to attempt to quit. Identifying the non-nicotine components of tobacco smoke responsible for remodelling the gut microbiota might permit their removal from cigarettes to lessen weight gain during abstinence. Alternatively, these components might be harnessed for use in individuals who would benefit from enhanced harvesting of nutrients from food, such as those with the cancer-associated weight loss known as cachexia.

This fresh perspective on the weight gain of ex-smokers should encourage closer consideration of whether other smoking-related physiological and behavioural changes involve processes that are not initiated in the brain.

doi: <https://doi.org/10.1038/d41586-021-03548-6>

References

1. 1.

Fluhr, L. *et al.* *Nature* <https://doi.org/10.1038/s41586-021-04194-8> (2021).

2. 2.

Filozof, C., Fernández Pinilla, M. C. & Fernández-Cruz, A. *Obes. Rev.* **5**, 95–103 (2004).

3. 3.

Bush, T., Lovejoy, J. C., Deprey, M. & Carpenter, K. M. *Obesity* **24**, 1834–1841 (2016).

4. 4.

Biedermann, L. *et al.* *Inflamm. Bowel Dis.* **20**, 1496–1501 (2014).

5. 5.

Lee, S. H. *et al.* *J. Clin. Med.* **7**, 282 (2018).

6. 6.

Lim, M. Y. *et al.* *Sci. Rep.* **6**, 23745 (2016).

7. 7.

Flegal, K. M. *et al.* *Am. J. Clin. Nutr.* **89**, 500–508 (2009).

8. 8.

Stolerman, I. P. & Jarvis, M. J. *Psychopharmacology* **117**, 2–10 (1995).

9. 9.

Picciotto, M. R. & Kenny, P. J. *Cold Spring Harb. Perspect. Med.*
<https://doi.org/10.1101/cshperspect.a039610> (2021).

10. 10.

Jessen, A., Buemann, B., Toubro, S., Skovgaard, I. M. & Astrup, A. *Diabetes Obes. Metab.* **7**, 327–333 (2005).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03548-6>

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

- NEWS AND VIEWS
- 22 December 2021

Clever substitutions reveal magnetism in zigzag graphene nanoribbons

The inclusion of nitrogen atoms stabilizes the zigzag edges of carbon-based nanoribbons, enabling the ribbons to be decoupled from a substrate and providing a probe for their unconventional magnetism.

- [Aran Garcia-Lekue](#) ⁰ &
- [Daniel Sánchez-Portal](#) ¹

Graphene is a single layer of carbon atoms arranged in a honeycomb lattice. Thin, flexible, transparent and metallic, it therefore forms an ideal material for many applications, especially for a type of electronics known as spintronics. In spintronic devices, the magnetic moment (spin) of an electron can be just as useful as its charge for storing information and performing logic operations. It has been predicted that when graphene is shaped into nanoribbons, with zigzag edges that are stabilized by carbon–hydrogen bonds, it should exhibit magnetic states that show particular promise for carbon-based electronics¹. However, a clear experimental demonstration of this magnetism in nanoribbons that are long enough to be technologically relevant has not been possible. [Writing in Nature](#), Blackwell *et al.*² overcome this hurdle — reporting the synthesis and characterization of zigzag graphene nanoribbons in which carbon atoms spaced at regular intervals along the edges have been replaced by nitrogen atoms.

Access options

[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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Nature **600**, 613-614 (2021)

doi: <https://doi.org/10.1038/d41586-021-03768-w>

References

1. 1.

Wang, H. *et al.* *Nature Rev. Phys.* **3**, 791–802 (2021).

2. 2.

Blackwell, R. E. *et al.* *Nature* **600**, 647–652 (2021).

3. 3.

Son, Y.-W., Cohen, M. L. & Louie, S. G. *Nature* **444**, 347–349 (2006).

4. 4.

Mishra, S. *et al.* *Nature Nanotechnol.* **15**, 22–28 (2020).

5. 5.

Li, J. *et al.* *Nature Commun.* **10**, 200 (2019).

6. 6.

Lawrence, J. *et al.* *ACS Nano* **14**, 4499–4508 (2020).

7. 7.

Mishra, S. *et al.* *Nature* **598**, 287–292 (2021).

8. 8.

Cai, J. *et al.* *Nature* **466**, 470–473 (2010).

9. 9.

Ruffieux, P. *et al.* *Nature* **531**, 489–492 (2016).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03768-w>

- NEWS AND VIEWS
- 08 December 2021

What surveys really say

Increasing the sample size of a survey is often thought to increase the accuracy of the results. However, an analysis of big surveys on the uptake of COVID-19 vaccines shows that larger sample sizes do not protect against bias.

- [Frauke Kreuter](#) 0

The accuracy of survey results is often thought to increase with sample size. However, [writing in *Nature*](#), Bradley *et al.*¹ show that this is not always the case. Although ‘big’ surveys can, under certain conditions, be useful for tracking changes in a population measure over time and across space, their estimates of population variables can be considerably biased.

Early in the COVID-19 pandemic, many nations lacked essential epidemiological data — even those with well-developed public-health monitoring infrastructures. There was a scarcity of timely information on regional increases in SARS-CoV-2 infections, on adherence to physical-distancing measures and on the social and economic effects of the pandemic. The state-sponsored data collections that existed at the time were often too slow to meet the demands generated by the pandemic.

As a result, some private companies jumped in to offer data; for example, Google, in Mountain View, California, provided anonymized, aggregated data on people’s mobility (go.nature.com/3htjccv), and Facebook in Menlo Park, California, presented anonymized and aggregated data about the development of connections between different geographical regions (go.nature.com/3lwknax). The London-based lifestyle company ZOE built the ZOE COVID Study app in collaboration with academic partners (go.nature.com/3i7ypxj). The app surveyed participants who downloaded it,

to identify infection hotspots and track the effect of mitigation measures. And when vaccination programmes were rolled out, it was used to record COVID-19 vaccine side effects. In addition, various private-sector surveys — many of which were archived by the US-based Societal Experts Action Network (go.nature.com/3rcmkwh) — produced data on changes in the public's response to the pandemic.

The US Census Bureau, in partnership with various federal agencies, and the Delphi group at Carnegie Mellon University, based in Pittsburgh, Pennsylvania, in partnership with Facebook, designed and performed massive surveys to forecast the spread of COVID-19 and measure its effects; questions about vaccination were added in early 2021. With more than 3 million and 25 million responses collected, respectively (as of November 2021; see go.nature.com/3dg0qvy and go.nature.com/3y2r1bk), these are now probably the largest US surveys relating to the pandemic. However, using a subset of responses, Bradley and colleagues demonstrate that the US Census Bureau–federal agencies survey (dubbed the Census Household Pulse survey) and the Delphi–Facebook survey overestimated the vaccination uptake compared with the benchmark data from the US Centers for Disease Control and Prevention (CDC) (Fig. 1).

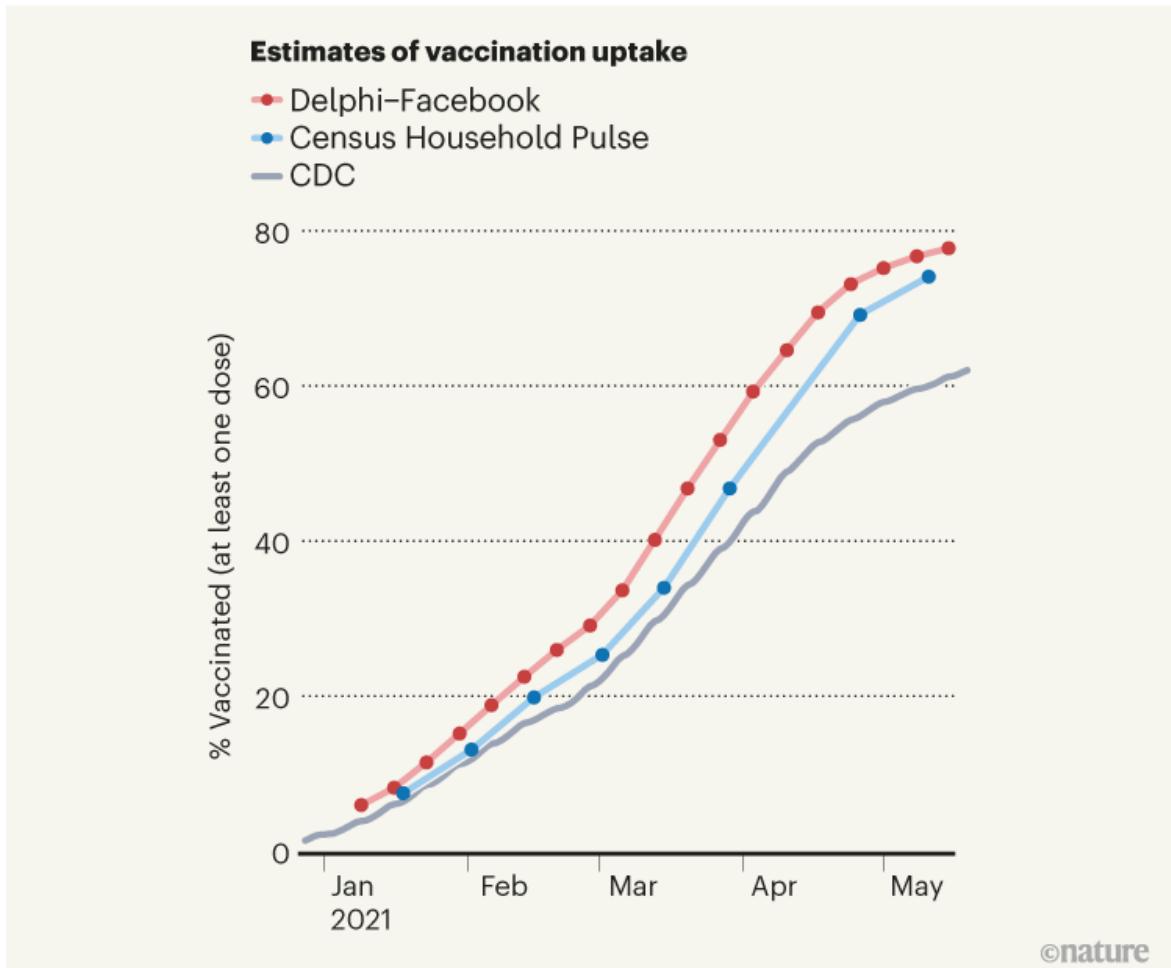


Figure 1 | Big surveys can give biased estimates of population variables. Bradley *et al.*¹ compared estimates of the uptake of SARS-CoV-2 vaccines among US adults, as reported by large surveys, with numbers of administered vaccine doses reported by the US Centers for Disease Control and Prevention (CDC) on 26 May 2021. Results from a survey carried out by the US Census Bureau in partnership with various federal agencies (Census Household Pulse), and another survey by the Delphi group at Carnegie Mellon University in Pittsburgh, Pennsylvania, in partnership with Facebook (Delphi–Facebook), overestimated vaccine uptake, but were useful in tracking the increase in vaccination over time in the first half of 2021. Bradley and colleagues explain how design choices in these surveys could account for the bias in the surveys' absolute estimates of vaccine uptake.

The authors conclude that having more data does not necessarily lead to better estimates. They discuss how design choices in survey-data collection can lead to error — in this case, the overestimation of vaccination uptake. Their findings are a reminder to researchers that statistical precision does not equate to unbiased population estimates.

Bradley and co-workers focus on three elements that can contribute to the size of the error — that is, the difference between estimates from big surveys and actual population values. These elements are data quantity (the fraction of a population that is captured in the sample), problem difficulty (how much variation in the outcome of interest there is in the population) and data quality. The quality is very difficult to assess, because there is usually no independently verified ‘ground truth’ or ‘gold standard’ with which to compare survey data. In this case, the CDC’s reports of the numbers of vaccines administered provide benchmark data with which the estimates reported in the surveys could be compared. Under the strong assumption that these reports are indeed the gold standard and reflect the correct vaccination rates, the survey estimates can be compared with these official numbers (which the CDC frequently updates; state-level estimates updated more recently than those used by Bradley *et al.* can be found at go.nature.com/3dtrdit). Using this approach, Bradley *et al.* evaluated estimates from several surveys and found that they did not match the CDC’s reported rates of vaccination uptake.

However, what the metric used by Bradley and colleagues does not enable us to answer — at least, not quantitatively — is the cause of the differences in data quality. To address this issue, the authors used a conceptual framework from survey methodology² called the total survey error (TSE) framework³, which can help to optimize survey-data quality in three key ways.

First, the TSE framework seeks to ensure that the population of interest and the members included in the ‘frame’ from which the sample is drawn are aligned. Facebook’s active user base is an example of a population that is not aligned with the entire population of the United States. Therefore, if Facebook users have different vaccination habits from those who do not use Facebook, estimates from a survey of Facebook users will be biased. Second, the framework aims to minimize the extent to which those who are

sampled and respond differ from the sample members who do not respond. For example, some people who don't trust the government might be less likely to respond to a government survey. Third, the accordance between the survey measure and the construct of interest should be maximized, and the respondents need to answer in the way intended. For example, questions about vaccination are at risk of being answered positively if respondents feel that they need to present themselves in a favourable light.

For certain inferential tasks, surveys with deficiencies can be useful⁴. The usefulness of a data set can be evaluated only in the context of a specific research question. For example, data from samples that are known to be biased have provided useful information for monitoring inflation rates, as exemplified by the Billion Prices Project (go.nature.com/3i6qock)⁵ — which, for years, used prices of online goods and services to estimate alternative indices of inflation. The project was able to do this because, even though not all goods and services were online, online and offline price changes tracked each other. Similarly, the data produced by the US Census Bureau and its partner agencies, and by the Delphi–Facebook partnership, can help to create early-warning systems when administrative data are lacking, as well as help to track cases⁶ and evaluate the effectiveness of measures designed to mitigate the spread of SARS-CoV-2 infections, if the errors of these surveys stay constant over time.

Large sample sizes can also reveal relationships between variables — such as reasons for vaccine hesitancy in subgroups of the population, and changes in these reasons over time — unless these relationships for survey respondents differ from those for people who do not respond. Samples collected at a high frequency over time and across relatively small geographical areas, such as some of the samples discussed by Bradley and colleagues, can also be used to evaluate the need for and effectiveness of policy interventions, such as mask-wearing mandates, lockdowns and school-based measures to limit COVID-19 spread^{7–9}.

The world is moving towards making decisions on the basis of data — as reflected, for example, in the US Foundations for Evidence-Based Policymaking Act of 2018 and the European Data Strategy (go.nature.com/3cp1f7o). In response to these changes, we will probably see

more data from all kinds of sources, not just surveys. Strong hopes rest on having more available administrative data, such as those from the CDC, that can in some instances replace survey data¹⁰ and, in others, improve survey estimates¹¹.

However, as with survey data, we will need robust frameworks and metrics to assess the quality of the data provided by governments, academic institutions and the private sector, and to guide us in using such data. The work by Bradley and colleagues reminds us that, alongside the studies themselves, research is needed on how best to use data — and on their quality and relevance to the question being asked.

Nature **600**, 614-615 (2021)

doi: <https://doi.org/10.1038/d41586-021-03604-1>

References

1. 1.

Bradley, V. C. *et al.* *Nature* <https://doi.org/10.1038/s41586-021-04198-4> (2021).

2. 2.

Groves, R. M. *et al.* *Survey Methodology* Vol. 56, 1 Ch. 2 (Wiley, 2011).

3. 3.

Biemer, P. P. *et al.* (eds) *Total Survey Error in Practice* (Wiley, 2017).

4. 4.

Kohler, U. *Surv. Methods Insights Field* <https://doi.org/10.13094/SMIF-2019-00014> (2019).

5. 5.

Cavallo, A. & Rigobon, R. *J. Econ. Perspect.* **30**, 151–178 (2016).

6. 6.

Reinhart, A. *et al.* Preprint at medRxiv
<https://doi.org/10.1101/2021.07.12.21259660> (2021).

7. 7.

Lessler, J. *et al.* *Science* **372**, 1092–1097 (2021).

8. 8.

Bilinski, A., Emanuel, E., Salomon, J. A. & Venkataramani, A. *J. Gen. Intern. Med.* **36**, 1825–1828 (2021).

9. 9.

Adjodah, D. *et al.* *PLoS ONE* **16**, e0252315 (2021).

10. 10.

Lane, J. *Democratizing our Data* (MIT Press, 2020).

11. 11.

National Academies of Sciences, Engineering, and Medicine. *Federal Statistics, Multiple Data Sources, and Privacy Protection: Next Steps* (National Academies, 2017).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03604-1>

- NEWS AND VIEWS
- 17 December 2021

Gravity, AlphaFold and neural interfaces: a year of remarkable science

Highlights from News & Views published this year.

January: Cosmic electromagnetic bombs laid bare

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Nature **600**, 617-620 (2021)

doi: <https://doi.org/10.1038/d41586-021-03730-w>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03730-w>

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

- Article
- [Published: 22 December 2021](#)

Very-high-frequency oscillations in the main peak of a magnetar giant flare

- [A. J. Castro-Tirado](#) ORCID: [orcid.org/0000-0003-2999-3563^{1,2}](https://orcid.org/0000-0003-2999-3563),
- [N. Østgaard](#) ORCID: [orcid.org/0000-0002-2572-7033³](https://orcid.org/0000-0002-2572-7033),
- [E. Göğüş](#) ⁴,
- [C. Sánchez-Gil](#) ORCID: [orcid.org/0000-0003-0829-4157⁵](https://orcid.org/0000-0003-0829-4157),
- [J. Pascual-Granado](#) ORCID: [orcid.org/0000-0003-0139-6951¹](https://orcid.org/0000-0003-0139-6951),
- [V. Reglero](#) ORCID: [orcid.org/0000-0002-0437-5286^{6,7}](https://orcid.org/0000-0002-0437-5286),
- [A. Mezentsev](#) ³,
- [M. Gabler](#) ORCID: [orcid.org/0000-0002-1663-4513⁶](https://orcid.org/0000-0002-1663-4513),
- [M. Marisaldi](#) ^{3,8},
- [T. Neubert](#) ORCID: [orcid.org/0000-0001-7851-7788⁹](https://orcid.org/0000-0001-7851-7788),
- [C. Budtz-Jørgensen](#) ORCID: [orcid.org/0000-0002-1618-2270⁹](https://orcid.org/0000-0002-1618-2270),
- [A. Lindanger](#) ORCID: [orcid.org/0000-0002-0723-0920³](https://orcid.org/0000-0002-0723-0920),
- [D. Sarria](#)³,
- [I. Kuvvetli](#)⁹,
- [P. Cerdá-Durán](#) ORCID: [orcid.org/0000-0003-4293-340X⁶](https://orcid.org/0000-0003-4293-340X),
- [J. Navarro-González](#) ORCID: [orcid.org/0000-0002-3799-1237⁷](https://orcid.org/0000-0002-3799-1237),
- [J. A. Font](#) ORCID: [orcid.org/0000-0001-6650-2634^{6,10}](https://orcid.org/0000-0001-6650-2634),
- [B.-B. Zhang](#) ORCID: [orcid.org/0000-0003-4111-5958^{11,12,13}](https://orcid.org/0000-0003-4111-5958),
- [N. Lund](#) ORCID: [orcid.org/0000-0001-5620-0886²](https://orcid.org/0000-0001-5620-0886),
- [C. A. Oxborrow](#) ORCID: [orcid.org/0000-0002-7555-6623⁹](https://orcid.org/0000-0002-7555-6623),
- [S. Brandt](#)⁹,
- [M. D. Caballero-García](#)¹,
- [I. M. Carrasco-García](#)¹⁴,

- [A. Castellón](#) ORCID: [orcid.org/0000-0003-4646-0451^{2,15}](https://orcid.org/0000-0003-4646-0451),
- [M. A. Castro Tirado^{1,16}](#),
- [F. Christiansen](#) ORCID: [orcid.org/0000-0001-9415-0621⁹](https://orcid.org/0000-0001-9415-0621),
- [C. J. Eyles⁷](#),
- [E. Fernández-García¹](#),
- [G. Genov³](#),
- [S. Guziy^{17,18}](#),
- [Y.-D. Hu](#) ORCID: [orcid.org/0000-0002-7400-4608^{1,19}](https://orcid.org/0000-0002-7400-4608),
- [A. Nicuesa Guelbenzu](#) ORCID: [orcid.org/0000-0002-6856-9813²⁰](https://orcid.org/0000-0002-6856-9813),
- [S. B. Pandey²¹](#),
- [Z.-K. Peng^{11,12}](#),
- [C. Pérez del Pulgar](#) ORCID: [orcid.org/0000-0001-5819-8310²](https://orcid.org/0000-0001-5819-8310),
- [A. J. Reina Terol²](#),
- [E. Rodríguez](#) ORCID: [orcid.org/0000-0001-6827-9077¹](https://orcid.org/0000-0001-6827-9077),
- [R. Sánchez-Ramírez²²](#),
- [T. Sun^{1,23,24}](#),
- [K. Ullaland³](#) &
- [S. Yang³](#)

Nature volume **600**, pages 621–624 (2021)

- 1477 Accesses
- 370 Altmetric
- [Metrics details](#)

Subjects

- [Astronomical instrumentation](#)
- [Compact astrophysical objects](#)
- [Transient astrophysical phenomena](#)

Abstract

Magnetars are strongly magnetized, isolated neutron stars^{1,2,3} with magnetic fields up to around 10^{15} gauss, luminosities of approximately 10^{31} – 10^{36} ergs per second and rotation periods of about 0.3–12.0 s. Very energetic giant flares from galactic magnetars (peak luminosities of 10^{44} – 10^{47} ergs per second, lasting approximately 0.1 s) have been detected in hard X-rays and soft γ -rays⁴, and only one has been detected from outside our galaxy⁵. During such giant flares, quasi-periodic oscillations (QPOs) with low (less than 150 hertz) and high (greater than 500 hertz) frequencies have been observed^{6,7,8,9}, but their statistical significance has been questioned¹⁰. High-frequency QPOs have been seen only during the tail phase of the flare⁹. Here we report the observation of two broad QPOs at approximately 2,132 hertz and 4,250 hertz in the main peak of a giant γ -ray flare¹¹ in the direction of the NGC 253 galaxy^{12,13,14,15,16,17}, disappearing after 3.5 milliseconds. The flare was detected on 15 April 2020 by the Atmosphere–Space Interactions Monitor instrument^{18,19} aboard the International Space Station, which was the only instrument that recorded the main burst phase (0.8–3.2 milliseconds) in the full energy range (50×10^3 to 40×10^6 electronvolts) without suffering from saturation effects such as deadtime and pile-up. Along with sudden spectral variations, these extremely high-frequency oscillations in the burst peak are a crucial component that will aid our understanding of magnetar giant flares.

This is a preview of subscription content

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Fig. 1: Temporal variability of GRB 200415A.

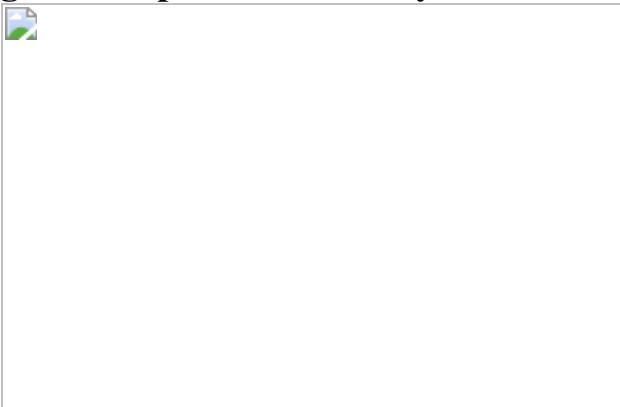


Fig. 2: Time-resolved spectroscopy.

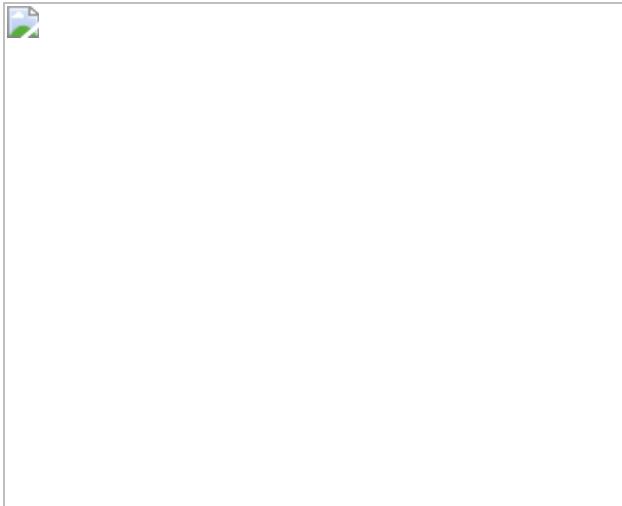
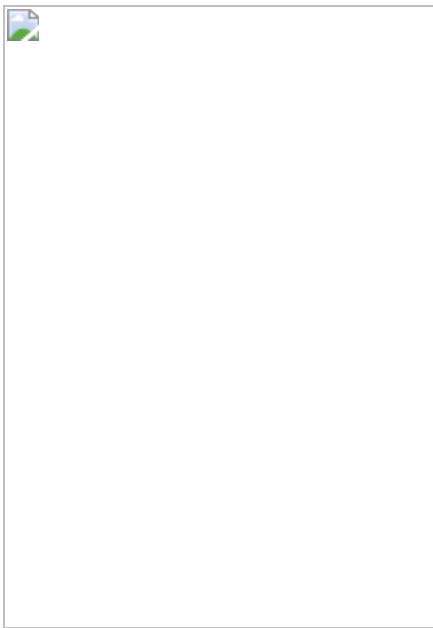


Fig. 3: Periodogram and fits for quasi-periodicities search for the time interval 0–5 ms.



Data availability

The ASIM data used for this study are available on the Zenodo repository, <https://doi.org/10.5281/zenodo.5242975>. The Swift 100- μ s-resolution data are publicly available at the Swift BAT GUANO repository: <https://www.swift.psu.edu/guano>. The Fermi NaI and BGO data are publicly available at <https://heasarc.gsfc.nasa.gov/W3Browse/fermi/fermigbrst.html>.

Code availability

For the search for QPOs within the Bayesian framework, we use the Stingray Python Package developed by Huppenkothen et al.³⁰ (<https://docs.stingray.science>). The Z^2 search and related calculations of confidence level and chance probabilities are performed using Interactive Data Language (IDL) version 8.7.3. The IDL codes developed can be provided upon request. For spectral analysis we used the software XSPEC, version v12.10.1f, (including the XANADU package) available at <https://heasarc.gsfc.nasa.gov/xanadu/xspec/>.

References

1. 1.

Kaspi, V. M. & Beloborodov, A. M. Magnetars. *Annu. Rev. Astron. Astrophys.* **55**, 261–301 (2017).

2. 2.

Mereghetti, S., Pons, J. A. & Melatos, A. Magnetars: properties, origin and evolution. *Space Sci. Rev.* **191**, 315–338 (2015).

3. 3.

Duncan, R. C. & Thompson, C. Formation of very strongly magnetized neutron stars: implications for gamma-ray bursts. *Astrophys. J.* **392**, L9–L13 (1992).

4. 4.

Mazets, E. P., Golenetskii, S. V., Il’inskii, V. N., Aptekar’, R. L. & Guryan, Yu. A. Observations of a flaring X-ray pulsar in Dorado. *Nature* **282**, 587–589 (1979).

5. 5.

Ofek, E. O. et al. The short-hard GRB 051103: observations and implications for its nature. *Astrophys. J.* **652**, 507–511 (2006).

6. 6.

Barat, C. et al. Fine time structure in the 1979 March 5 gamma ray burst. *Astron. Astrophys.* **126**, 400–402 (1983).

7. 7.

Strohmayer, T. E. & Watts, A. L. Discovery of fast X-ray oscillations during the 1998 giant flare from SGR 1900+14. *Astrophys. J. Lett.* **632**, L111–L114 (2005).

8. 8.

Israel, G. L. et al. The discovery of rapid X-ray oscillations in the tail of the SGR 1806-20 hyperflare. *Astrophys. J.* **628**, L53–L56 (2005).

9. 9.

Strohmayer, T. E. & Watts, A. L. The 2004 hyper flare from SGR 1806-20: further evidence for global torsional vibrations. *Astrophys. J.* **653**, 593–601 (2006).

10. 10.

Pumpe, D., Gabler, M., Steininger, T. & Enßlin, T. A. Search for quasi-periodic signals in magnetar giant flares – Bayesian inspection of SGR 1806-20 and SGR 1900+14. *Astron. Astrophys.* **610**, A61 (2018).

11. 11.

Marisaldi, M., Mezentsev, A., Østgaard, N., Reglero, V. & Neubert, T. GRB 200415A: ASIM observation. *GCN Circular* 27622 (2020).

12. 12.

Svinkin, D. et al. A bright γ -ray flare interpreted as a giant magnetar flare in NGC 253. *Nature* **589**, 211–213 (2021).

13. 13.

Roberts, O. R. et al. Rapid variability of a giant flare from a magnetar in NGC 253. *Nature* **589**, 207–210 (2021).

14. 14.

The Fermi-LAT Collaboration. High-energy emission from a magnetar giant flare in the Sculptor galaxy. *Nat. Astron.* **5**, 385–391 (2021).

15. 15.

Yang, J. et al. GRB 200415A: a short gamma-ray burst from a magnetar giant flare? *Astrophys. J.* **899**, 106 (2020).

16. 16.

Zhang, H. M., Liu, R.-Y., Zhong, S.-Q. & Wang, X.-Y. Magnetar giant flare origin for GRB 200415A inferred from a new scaling relation. *Astrophys. J. Lett.* **903**, 32 (2020).

17. 17.

Burns, E. et al. Identification of a local sample of gamma-ray bursts consistent with a magnetar giant flare origin. *Astrophys. J. Lett.* **907**, 28 (2021).

18. 18.

Neubert, T. et al. The ASIM mission on the International Space Station. *Space Sci. Rev.* **215**, 26 (2019).

19. 19.

Østgaard, N. et al. The Modular X- and Gamma-Ray Sensor (MXGS) of the ASIM payload on the International Space Station. *Space Sci.*

Rev. **215**, 23 (2019).

20. 20.

Leahy, D. A. et al. On searches for pulsed emission with application to four globular cluster X-ray sources: NGC 1851, 6441, 6624 and 6712. *Astrophys. J.* **266**, L160–L170 (1983).

21. 21.

Buccheri, R. et al. Search for pulsed γ -ray emission from radio pulsars in the COS-B data. *Astron. Astrophys.* **128**, 245–251 (1983).

22. 22.

Gehrels, N. et al. The Swift gamma-ray burst mission. *AAstrophys. J.* **611**, 1005 (2004).

23. 23.

Barthelmy, S. D. et al. The Burst Alert Telescope on the Swift Midex mission. *Space Sci. Rev.* **120**, 143–164 (2005).

24. 24.

Tohuvavohu, A. et al. Gamma-Ray Urgent Archiver for Novel Opportunities (GUANO): Swift/BAT event data dumps on demand to enable sensitive subthreshold GRB searches. *AAstrophys. J.* **900**, 35 (2020).

25. 25.

Mahlmann, J. F., Akgün, T., Pons, J. A., Aloy, M. A. & Cerdá-Durán, P. Instability of twisted magnetar magnetospheres. *Mon. Not. R. Astron. Soc.* **490**, 4858–4876 (2019).

26. 26.

Parfrey, K., Beloborodov, A. M. & Hui, L. Twisting, reconnecting magnetospheres and magnetar spindown. *Astrophys. J. Lett.* **754**, 12 (2012).

27. 27.

Li, X., Zrake, J. & Beloborodov, A. M. Dissipation of Alfvén waves in relativistic magnetospheres of magnetars. *Astrophys. J.* **881**, 13 (2019).

28. 28.

Duncan, R. C. Global seismic oscillations in soft gamma repeaters. *Astrophys. J.* **498**, L45–L49 (1998).

29. 29.

Gabler, M., Cerdá-Durán, P., Stergioulas, N., Font, J. A. & Müller, E. Constraining properties of high-density matter in neutron stars with magneto-elastic oscillations. *Mon. Not. R. Astron. Soc.* **476**, 4199–4212 (2018).

30. 30.

Chanrion, O. et al. The Modular Multispectral Imaging Array (MMIA) of the ASIM payload on the International Space Station. *Space Sci. Rev.* **215**, 28 (2019).

31. 31.

Vaughan, S. A Bayesian test for periodic signals in red noise. *Mon. Not. R. Astron. Soc.* **402**, 307–320 (2010).

32. 32.

Huppenkothen, D. et al. Quasi-periodic oscillations and broadband variability in short magnetar bursts. *Astrophys. J.* **768**, 87 (2013).

33. 33.

Huppenkothen, D. et al. Stingray: a modern Python library for spectral timing. *Astrophys. J.* **881**, 39 (2019).

34. 34.

McDermott, P. N., van Horn, H. M. & Hansen, C. J. Nonradial oscillations of neutron stars. *Astrophys. J.* **325**, 725–748 (1988).

35. 35.

Vaughan, S. A simple test for periodic signals in red noise. *Astron. Astrophys.* **431**, 391–403 (2005).

36. 36.

Sotani, H., Kokkotas, K. D. & Stergioulas, N. Torsional oscillations of relativistic stars with dipole magnetic fields. *Mon. Not. R. Astron. Soc.* **375**, 261–277 (2007).

37. 37.

Brazier, K. T. S. Confidence intervals from the Rayleigh test. *Mon. Not. R. Astron. Soc.* **268**, 709–712 (1994).

38. 38.

Timokhin, A. N., Eichler, D. & Lubarsky, Yu. On the nature of quasi-periodic oscillations in the tail of soft gamma repeater giant flares. *Astrophys. J.* **680**, 1398–1404 (2008).

39. 39.

Gabler, M., Cerdá-Durán, P., Stergioulas, N., Font, J. A. & Müller, E. Modulating the magnetosphere of magnetars by internal magneto-elastic oscillations. *Mon. Not. R. Astron. Soc.* **443**, 1416–1424 (2014).

40. 40.

Gabler, M. et al. Magneto-elastic oscillations modulating the emission of magnetars. *Astron. Nachr.* **338**, 1105–1108 (2017).

41. 41.

Watts, A. L. & Strohmayer, T. E. Detection with RHESSI of high-frequency X-ray oscillations in the tail of the 2004 hyperflare from SGR 1806-20. *Astrophys. J.* **637**, L117–L120 (2006).

42. 42.

Huppenkothen, D., Heil, L. M., Watts, A. L. & Göğüş, E. Quasi-periodic oscillations in short recurring bursts of magnetars SGR 1806-20 and SGR 1900+14 observed with RXTE. *Astrophys. J.* **795**, 114 (2014).

43. 43.

Miller, M. C., Chirenti, C. & Strohmayer, T. E. On the persistence of QPOs during the SGR 1806-20 giant flare. *Astrophys. J.* **871**, 95 (2019).

44. 44.

Steiner, A. W. & Watts, A. L. Constraints on neutron star crusts from oscillations in giant flares. *Phys. Rev. Lett.* **103**, 181101 (2009).

45. 45.

Carrasco, F., Viganò, D., Palenzuela, C. & Pons, J. A. Triggering magnetar outbursts in 3D force-free simulations. *Mon. Not. R. Astron. Soc.* **484**, L124–L129 (2019).

46. 46.

Li, X. & Beloborodov, A. M. Plastic damping of Alfvén waves in magnetar flares and delayed afterglow emission. *Astrophys. J.* **815**, 25 (2015).

Thompson, C. & Duncan, R. C. The giant flare of 1998 August 27 from SGR 1900+14. II. Radiative mechanism and physical constraints on the source. *Astrophys. J.* **561**, 980–1005 (2001).

Acknowledgements

This work made use of data from ASIM. ASIM is an international experiment aboard the International Space Station. ASIM is a mission of the European Space Agency (ESA), funded by ESA and national grants from Denmark, Norway and Spain. ESA PRODEX contracts C 4000115884 (DTU) and 4000123438 (Bergen) support the ASIM Science Data Centre. The science analysis is supported by ESA Topical Team contract 4200019920/06/NL/VJ; European Research Council grant AdG-FP7/2007-2013: n 320839; Research Council of Norway contract 223252/F50 (CoE/BCSS); and Ministerio de Ciencia, Innovación y Universidades grants ESP 2017-86263-C4 and ESP2017-87676-C5-5-R. J.A.F. and P.C.-D. acknowledge support by the Spanish Agencia Estatal de Investigación (grant PGC2018-095984-B-I00) and by the Generalitat Valenciana (PROMETEO/2019/071). P.C.-D. acknowledges the Ramon y Cajal funding (RYC-2015-19074) for supporting his research. M.G. acknowledges support through the Generalitat Valenciana via the grant CIDEVENT/2019/031. A.J.C.-T., J.P.-G., M.D.C.-G. and E.R. acknowledge financial support from the State Agency for Research of the Spanish MCIU through the “Center of Excellence Severo Ochoa” award to the Instituto de Astrofísica de Andalucía (SEV-2017-0709). This work is partly based on observations made with the Gran Telescopio Canarias (GTC), installed at the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias, on the island of La Palma. We thank J. D. Scargle for his help with the preliminary time-series analysis and his comments. J.P.-G. also acknowledges funding support from Spanish public funds for research from project PID2019-107061GB-C63 from the ‘Programas Estatales de Generación de Conocimiento y Fortalecimiento Científico y Tecnológico del Sistema de I+D+i y de I+D+i Orientada a los Retos de la Sociedad’ from the Spanish Ministry of Science, Innovation and Universities (MCIU).

Author information

Affiliations

1. Instituto de Astrofísica de Andalucía del Consejo Superior de Investigaciones Científicas (IAA-CSIC), Granada, Spain
A. J. Castro-Tirado, J. Pascual-Granado, M. D. Caballero-García, M. A. Castro Tirado, E. Fernández-García, Y.-D. Hu, E. Rodríguez & T. Sun
2. Unidad Asociada al CSIC Departamento de Ingeniería de Sistemas y Automática, Escuela de Ingeniería Industrial, Universidad de Málaga, Málaga, Spain
A. J. Castro-Tirado, A. Castellón, C. Pérez del Pulgar & A. J. Reina Terol
3. Birkeland Centre for Space Science, Department of Physics and Technology, University of Bergen, Bergen, Norway
N. Østgaard, A. Mezentsev, M. Marisaldi, A. Lindanger, D. Sarria, G. Genov, K. Ullaland & S. Yang
4. Faculty of Engineering and Natural Sciences, Sabancı University, Istanbul, Turkey
E. Göğüş
5. Departamento de Estadística e Investigación Operativa, Universidad de Cádiz, Puerto Real, Spain
C. Sánchez-Gil
6. Departamento de Astronomía y Astrofísica, Universitat de València, Valencia, Spain
V. Reglero, M. Gabler, P. Cerdá-Durán & J. A. Font

7. Imaging Processing Laboratory, Universitat de València, Valencia, Spain

V. Reglero, J. Navarro-González & C. J. Eyles

8. National Institute for Astrophysics, Osservatorio di Astrofisica e Scienza dello Spazio, Bologna, Italy

M. Marisaldi

9. DTU Space, Technical University of Denmark, Kongens Lyngby, Denmark

T. Neubert, C. Budtz-Jørgensen, I. Kuvvetli, N. Lund, C. A. Oxborrow, S. Brandt & F. Christiansen

10. Observatori Astronòmic, Universitat de València, Valencia, Spain

J. A. Font

11. School of Astronomy and Space Science, Nanjing University, Nanjing, China

B.-B. Zhang & Z.-K. Peng

12. Key Laboratory of Modern Astronomy and Astrophysics, Nanjing University, Ministry of Education, Nanjing, China

B.-B. Zhang & Z.-K. Peng

13. Department of Physics and Astronomy, University of Nevada, Las Vegas, Las Vegas, NV, USA

B.-B. Zhang

14. Departamento de Química Analítica, Universidad de Málaga, Málaga, Spain

I. M. Carrasco-García

15. Departamento de Álgebra, Geometría y Topología, Facultad de Ciencias, Universidad de Málaga, Málaga, Spain
A. Castellón
16. Escuela Técnica Superior de Arquitectura, Universidad de Málaga, Málaga, Spain
M. A. Castro Tirado
17. Astronomical Observatory, Mykolaiv National University, Mykolaiv, Ukraine
S. Guziy
18. Research Institute, Mykolaiv Astronomical Observatory, Mykolaiv, Ukraine
S. Guziy
19. Facultad de Ciencias, Universidad de Granada, Granada, Spain
Y.-D. Hu
20. Thüringer Landessternwarte Tautenburg, Tautenburg, Germany
A. Nicuesa Guelbenzu
21. Aryabhatta Research Institute of Observational Sciences (ARIES), Nainital, India
S. B. Pandey
22. Istituto di Astrofisica e Planetologia Spaziali, INAF, Rome, Italy
R. Sánchez-Ramírez
23. Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing, China

T. Sun

24. School of Astronomy and Space Science, University of Science and Technology of China, Hefei, China

T. Sun

Contributions

A.J.C.-T., N.Ø. and V.R. led the research team for this event. V.R., J.N.-G., N.Ø. and T.N. contributed to ASIM and provided the raw ASIM data. A.M., M.M., N.Ø., A.L. and V.R. provided the data for the time analysis. E.G., C.S.-G., J.P.-G. and E.R. analysed the time variability in the data. A.L. and C.A.O. provided the spectral accepted data. B.-B.Z. and Z.-K.P. performed the preliminary spectral analysis and M.M. performed the detailed spectral analysis. I.K. and C.B.-J. contributed to all ASIM phases for the MXGS development. A.L. and D.S. verified the LED and HED energy-response matrices. D.S. provided the calibration of the HED instrument and also performed the calculations of column density and potential photon absorption in the atmosphere. F.C., C.J.E., G.G., K.U. and S.Y. contributed substantially to the design, building and operation of the ASIM MXGS instrument. I.M.C.-G., M.A.C.T., E.F.-G., C.P.d.P., S. G. and A.J.R.T. provided the optical data. P.C.-D., M.G. and J.A.F. contributed to the theoretical interpretation of the observational results. A.J.C.-T., N.Ø., E.G., C.S.-G., M.G. and J.P.-G. contributed to the writing of the manuscript. All the co-authors, including S.B., M.D.C.-G., A.C., M.A.C.T., P.C.-D., J.A.F., M.G., Y.-D.H., I.K., N.L., A.N.G., S.B.P., R.S.-R. and T.S., have read the manuscript and made useful suggestions.

Corresponding authors

Correspondence to [N. Østgaard](#), [E. Göğüş](#), [A. Mezentsev](#), [M. Gabler](#) or [M. Marisaldi](#).

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.

Extended data figures and tables

Extended Data Fig. 1 Optical content of the final Inter-Planetary Network (IPN) error box¹² (red parallelogram) partly including the NGC 253 galaxy.

a, A 120-s exposure (clear filter) taken by the BOOTES-1A wide field camera in Huelva (South Spain) on 23 July 2020 (04:00 ut). The field of view is $0.5^\circ \times 0.5^\circ$. This field was just 1.84° above the Earth's limb as seen by ASIM. **b**, A magnification fo the IPN error box showing a colour image resulting from the combination of g (10 s), r (5 s) and i (5 s) -band exposures taken by the 10.4-m Gran Telescopio Canarias on 16 August 2020, 04:30 ut.

Extended Data Fig. 2 GRB 200415A lightcurve as measured by Fermi/GBM, ASIM and Swift/BAT.

a, ASIM HED (all 12 BGOs) and Fermi/GBM BGO 0 (as shown in ref. ¹³) in the same energy ranges. **b**, ASIM LED and Fermi/GBM NaI 0,1,2,3,5 (as in ref. ¹³) in the same energy range. **c**, ASIM LED and Swift/BAT in comparable energy ranges.

Extended Data Fig. 3 Time-resolved spectroscopy (including residuals).

Spectra for four time intervals during the magnetar giant flare, defined in the main text. **a**, Precursor: 0.0–0.8 ms. **b**, Peak: 0.8–3.2 ms. **c**, Decay: 3.2–8 ms. **d**, Tail: 8–160 ms. A total number of 2,148 counts in the LED and 2,141 counts in the HED have been used. Left panels: count spectra (blue: LED; red: HED) and best-fit models (green). Middle panels: residuals. Best-fit models are those highlighted in bold in Table 1. Right panels: energy flux density for the best-fit models.

Extended Data Fig. 4 Comparison between ASIM MXGS and Fermi GBM.

Spectral energy density plots for the four time intervals considered in ref. 13. Data points: ASIM MXGS unfolded spectra; LED (blue) and HED (red). Green curves: 1σ confidence region for the Fermi GBM spectra, from figure 1d from ref. 13.

Extended Data Fig. 5 Periodograms and residuals of the LED light curves for time periods from 0-5, 0-10, 0-50 and 0-100 ms after the burst ($t = 0$ is start of the burst), with a time resolution of 50 μ s (frequency bins of 200 Hz).

a, Unsmoothed (black) and smoothed (orange; Wiener filter, $3\Delta\nu$) periodograms. Power-law fit in blue, and power law plus two Lorentzians (centred around the 2,100 and 4,200 Hz features) fit in red. **b**, The residuals of the fits around the 2,100 Hz feature. **c**, The residuals of the fits around the 4,200 Hz feature.

Extended Data Fig. 6 Periodograms and residuals of the HED light curves for time periods from 0-5, 0-10, 0-50 and 0-100 ms after the burst ($t = 0$ is start of the burst), with a time resolution of 50 μ s (frequency bins of 200 Hz).

a, Unsmoothed (black) and smoothed (orange; Wiener filter, $3\Delta\nu$) periodograms. Power-law fit in blue, and power law plus two Lorentzians (centred around the 2,100 and 4,200 Hz features) fit in red. **b**, The residuals

of the fits around the 2,100 Hz feature. **c**, The residuals of the fits around the 4,200 Hz feature.

**Extended Data Fig. 7 **

({\{boldsymbol{Z}\}}_{\{{\bf 1}\}}^{\{{\bf 2}\}}).power spectra in the 300-5,000 Hz frequency range, obtained with the events within the first 5 ms time interval of the burst.

a, Using LED data. **b**, Using HED data. The horizontal dotted, dashed and solid lines represent 95%, 99% and 99.9% confidence levels, respectively.

Extended Data Fig. 8 A spectrogram combined of consecutive periodograms for ASIM-LED, 50-400 keV, with 200 ms long sliding window, and 1 ms step.

Time axis corresponds to the left edge of the sliding window. **a**, Covering the interval between −300 ms and +300 ms relative to the GRB onset. Colour bar represents the Leahy-normalized PSD, where 2 corresponds to the white noise level. Areas of interest are highlighted with the red rectangles and are shown in the next panels for frequency range 1 kHz to 5 kHz. **b**, Zoomed-in segment of the spectrogram when the sliding window just enters the GRB region. A higher time resolution of 10 μ s steps is used. For convenience time axis corresponds to the right edge of the 200-ms long sliding window, relative to the GRB onset. **c**, Zoomed-in segment of the spectrogram when the sliding window starts to leave the GRB region. Time axis corresponds to the left edge of the 200-ms long sliding window, relative to the GRB onset.

Extended Data Table 1 Spectral parameters and fluxes for ASIM and Fermi

Extended Data Table 2 Results for the search for periodicities and quasi-periodicities: posterior summary

Extended Data Table 3 Z^2 search results and corresponding chance probabilities

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Castro-Tirado, A.J., Østgaard, N., Göğüş, E. *et al.* Very-high-frequency oscillations in the main peak of a magnetar giant flare. *Nature* **600**, 621–624 (2021). <https://doi.org/10.1038/s41586-021-04101-1>

- Received: 17 August 2020
- Accepted: 06 October 2021
- Published: 22 December 2021
- Issue Date: 23 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04101-1>

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

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04101-1>

- Article
- Open Access
- [Published: 15 December 2021](#)

Quantum theory based on real numbers can be experimentally falsified

- [Marc-Olivier Renou](#) ORCID: [orcid.org/0000-0002-4358-2384¹](https://orcid.org/0000-0002-4358-2384),
- [David Trillo](#) ORCID: [orcid.org/0000-0002-9836-7366²](https://orcid.org/0000-0002-9836-7366),
- [Mirjam Weilenmann](#) ORCID: [orcid.org/0000-0003-0852-6763²](https://orcid.org/0000-0003-0852-6763),
- [Thinh P. Le²](#),
- [Armin Tavakoli^{2,3}](#),
- [Nicolas Gisin](#) ORCID: [orcid.org/0000-0003-0788-3494^{4,5}](https://orcid.org/0000-0003-0788-3494),
- [Antonio Acín^{1,6}](#) &
- [Miguel Navascués](#) ²

[Nature](#) volume 600, pages 625–629 (2021)

- 33k Accesses
- 1 Citations
- 477 Altmetric
- [Metrics details](#)

Subjects

- [Quantum information](#)
- [Quantum mechanics](#)

Abstract

Although complex numbers are essential in mathematics, they are not needed to describe physical experiments, as those are expressed in terms of probabilities, hence real numbers. Physics, however, aims to explain, rather than describe, experiments through theories. Although most theories of physics are based on real numbers,

quantum theory was the first to be formulated in terms of operators acting on complex Hilbert spaces^{1,2}. This has puzzled countless physicists, including the fathers of the theory, for whom a real version of quantum theory, in terms of real operators, seemed much more natural³. In fact, previous studies have shown that such a ‘real quantum theory’ can reproduce the outcomes of any multipartite experiment, as long as the parts share arbitrary real quantum states⁴. Here we investigate whether complex numbers are actually needed in the quantum formalism. We show this to be case by proving that real and complex Hilbert-space formulations of quantum theory make different predictions in network scenarios comprising independent states and measurements. This allows us to devise a Bell-like experiment, the successful realization of which would disprove real quantum theory, in the same way as standard Bell experiments disproved local physics.

[Download PDF](#)

Main

Without qualification, the question of whether complex numbers are necessary for natural sciences, and, more concretely, for physics, must be answered in the negative: physics experiments are described by the statistics they generate, that is, by probabilities, and hence real numbers, thus, there is no need for complex numbers. The question becomes meaningful, however, when considering a specific theoretical framework, designed to explain existing experiments and make predictions about future ones. Whether complex numbers are needed within a theory to correctly explain experiments, or whether real numbers only are sufficient, is not straightforward. Complex numbers are sometimes introduced in electromagnetism to simplify calculations: one might, for instance, regard the electric and magnetic fields as complex vector fields to describe electromagnetic waves. However, this is just a computational trick. We wonder whether the same can be said about complex numbers in quantum theory.

In its Hilbert space formulation, quantum theory is defined in terms of the following postulates^{5,6}. (1) For every physical system S , there corresponds a Hilbert space \mathcal{H}_S and its state is represented by a normalized vector ϕ in \mathcal{H}_S , that is, $\langle \varphi | \varphi \rangle = 1$. (2) A measurement Π in S corresponds to an ensemble $\{\Pi_r\}_r$ of projection operators, indexed by the measurement result r and acting on \mathcal{H}_S , with $\sum_r \Pi_r = \mathbb{I}$. (3) Born rule: if we measure Π when system S is in state ϕ , the probability of obtaining result r is given by $\Pr(r) = \langle \varphi | \Pi_r | \varphi \rangle$. (4) The Hilbert space \mathcal{H}_{ST} corresponding to the composition of two systems S and T is $\mathcal{H}_S \otimes \mathcal{H}_T$. The operators

used to describe measurements or transformations in system S act trivially on \mathcal{H}_T and vice versa. Similarly, the state representing two independent preparations of the two systems is the tensor product of the two preparations.

This last postulate has a key role in our discussions: we remark that it even holds beyond quantum theory, specifically for space-like separated systems in some axiomatizations of quantum field theory^{7,8,9,10} ([Supplementary Information](#)).

As originally introduced by Dirac and von Neumann^{1,2}, the Hilbert spaces \mathcal{H}_S in postulate (1) are traditionally taken to be complex. We call the resulting postulate (1_c) . The theory specified by postulates (1_c) and (2)–(4) is the standard formulation of quantum theory in terms of complex Hilbert spaces and tensor products. For brevity, we will refer to it simply as ‘complex quantum theory’. Contrary to classical physics, complex numbers (in particular, complex Hilbert spaces) are thus an essential element of the very definition of complex quantum theory.

Owing to the controversy surrounding their irruption in mathematics and their almost total absence in classical physics, the occurrence of complex numbers in quantum theory worried some of its founders, for whom a formulation in terms of real operators seemed much more natural (“What is unpleasant here, and indeed directly to be objected to, is the use of complex numbers. Ψ is surely fundamentally a real function.” (Letter from Schrödinger to Lorentz, 6 June 1926; ref. ³)). This is precisely the question we address in this work: whether complex numbers can be replaced by real numbers in the Hilbert space formulation of quantum theory without limiting its predictions. The resulting ‘real quantum theory’, which has appeared in the literature under various names^{11,12}, obeys the same postulates (2)–(4) but assumes real Hilbert spaces \mathcal{H}_S in postulate (1), a modified postulate that we denote by \((\{\{1\}_{\{\{\mathbb{R}\}\}}})\).

If real quantum theory led to the same predictions as complex quantum theory, then complex numbers would just be, as in classical physics, a convenient tool to simplify computations but not an essential part of the theory. However, we show that this is not the case: the measurement statistics generated in certain finite-dimensional quantum experiments involving causally independent measurements and state preparations do not admit a real quantum representation, even if we allow the corresponding real Hilbert spaces to be infinite dimensional.

Our main result applies to the standard Hilbert space formulation of quantum theory, through axioms (1)–(4). It is noted, though, that there are alternative formulations able to recover the predictions of complex quantum theory, for example, in terms of path integrals¹³, ordinary probabilities¹⁴, Wigner functions¹⁵ or Bohmian mechanics¹⁶. For

some formulations, for example, refs. [17,18](#), real vectors and real operators play the role of physical states and physical measurements respectively, but the Hilbert space of a composed system is not a tensor product. Although we briefly discuss some of these formulations in [Supplementary Information](#), we do not consider them here because they all violate at least one of the postulates $\langle\{1\}_{\{\{\mathbb{R}\}\}}\rangle$ and (2)–(4). Our results imply that this violation is in fact necessary for any such model.

It is instructive to address our main question as a game between two players—the ‘real’ quantum physicist Regina and the ‘complex’ quantum physicist Conan. Regina is convinced that our world is governed by real quantum theory, whereas Conan believes that only complex quantum theory can describe it. Through a well chosen quantum experiment, Conan aims to prove Regina wrong; that is, to falsify real quantum theory by exhibiting an experiment that this theory cannot explain.

At first, Conan thinks of conducting simple experiments involving a single quantum system. Unfortunately, for any such quantum experiment, Regina can find a real quantum explanation. For instance, if ρ is the complex density matrix that Conan uses to model his experiment, Regina could propose the state

$$\begin{aligned} \text{\$\$} &\mathop{\rho} \limits^{\sim} = \text{\rm R} \{ e \} (\rho) \otimes \\ &\frac{1}{\sqrt{2}} \{ \text{\rm I} \} \{ 2 \} + \{ \text{\rm I} \} \{ \text{\rm m} \} (\rho) \otimes \frac{1}{\sqrt{2}} \begin{array}{l} \{ \text{\rm r} \} 0 \& 1 \\ -1 \& 0 \end{array} \end{aligned} \quad (1)$$

where $\langle \pm i \rangle = \frac{1}{\sqrt{2}} (\langle 0 | \pm i | 1 \rangle)$, and the asterisk denotes complex conjugation. The operator $\langle \tilde{\rho} \rangle$ is real and positive semidefinite: it is thus a real quantum state. Fig. [1](#) (left) explains how Regina can analogously define real measurement operators that, acting on $\langle \tilde{\rho} \rangle$, reproduce the statistics of any (complex) measurement conducted by Conan on ρ . This construction is just one of the infinitely many ways that Regina has to explain the measurement statistics of any single-particle experiment using real operators, but it already implies that real quantum theory cannot be falsified in this scenario. It does not imply, however, that states in real quantum theory are restricted to have this form: they remain arbitrary, as in complex quantum theory.

Fig. 1: Simulating single-site and multipartite quantum experiments through real quantum theory.



Left: a single-site quantum experiment. A complex quantum system in state ρ is probed via the measurement $\{\Pi_r\}_r$. One way to reproduce the measurement statistics of this experiment using real quantum theory requires adding an extra real qubit: the state ρ is then replaced by the real state $\langle \tilde{\rho} \rangle$ in equation (1), while every measurement operator is replaced by the real measurement operator $\langle \tilde{\Pi}_r \rangle = \langle \{ \Pi_i \}_r \rangle \otimes | \{ \rm{rm}{ i } \} \rangle \langle \{ \rm{rm}{ i } \} | + \langle \{ \Pi_i \}_r \rangle^{\ast} \otimes | - \{ \rm{rm}{ i } \} \rangle \langle - \{ \rm{rm}{ i } \} |$. Using the fact that probabilities are real, and thus $\langle P(r) = P(r_i) \rangle^{\ast} = \langle \{ \rm{rm}{ t } \}_r \{ \rm{rm}{ r } \} \langle \{ \rm{rm}{ r } \} | \{ \rm{rho} \}^{\ast} \{ \Pi_i \}_r \rangle$, it is straightforward that $\langle P(r) = \langle \{ \rm{rm}{ tr } \} \rangle \langle \{ \rm{rho} \} \{ \{ \Pi_i \}_r \} \rangle = \langle \{ \rm{rm}{ tr } \} \rangle \langle \langle \tilde{\rho} \{ \{ \Pi_i \} \}_r \rangle \rangle$. It is noted that this construction doubles the Hilbert space dimension of the original complex quantum system (when finite). This is not surprising, as a complex number is defined by two real numbers, and may just be seen as yet another example of how complex numbers simplify the calculation of experimental predictions, as in electromagnetism. Right: a multipartite quantum experiment. A complex Bell scenario consists of two particles (or systems) distributed between Alice and Bob, who perform local measurements, labelled by x and y , and get results a and b . By postulates (1 ϵ)–(4), a complex Hilbert space is assigned to each particle, and the Hilbert space describing the overall bipartite system is defined by the tensor product of these. The state of the two particles is thus described by an operator ρ_{AB} acting on the joint space, whereas operators $\langle A \rangle_a|x\rangle$ and $\langle B \rangle_b|y\rangle$ acting on each local Hilbert space describe the local measurements. The observed measurement statistics or correlations are described by the conditional probability distribution $\langle P(ab|xy) = \langle \{ \rm{rm}{ tr } \} \langle \{ \rm{rho} \}_{AB} \{ A \}_a \{ a|x \} \otimes \{ B \}_b \{ b|y \} \rangle \rangle$. One way to reproduce these statistics using real

quantum theory consists of assigning an extra real qubit to each particle. The quantum state is replaced by the real state $\langle \tilde{\rho}_{AB} \rangle = \frac{1}{2} (\rho_{AB} + i(\rho_{AB}^* - \rho_{AB}))$, and the local measurements are replaced by the same transformation as before for a single system. The observed statistics are again recovered, that is, $\langle P(ab|xy) \rangle = \langle \tilde{\rho}_{AB} \rangle \langle \tilde{\rho}_A \rangle \langle \tilde{\rho}_B \rangle$.

It is noted that, assuming a fixed Hilbert space dimension, Conan could come up with single-site experiments where real and complex quantum theory differ, for instance, because the former does not satisfy local tomography, or even leads to different experimental predictions (see, for example, ref. [19](#)). However, as dimension cannot be upper bounded experimentally^{[20](#)}, Regina would be right not to interpret any such experiment as a disproof of real quantum theory. In practice, any experimental system has infinite degrees of freedom: a finite dimension may just be an approximation made to simplify its theoretical description. Hence, to defeat Regina, Conan has to design an experiment in which no explanation using real Hilbert spaces is valid, no matter their dimension.

Conan may next consider experiments involving several distant labs, where phenomena such as entanglement^{[21](#)} and Bell non-locality^{[22](#)} can manifest. For simplicity, we focus on the case of two separate labs. A source emits two particles (for example, a crystal pumped by a laser emitting two photons) in a state ρ_{AB} , each being measured by different observers, called Alice (A) and Bob (B) (Fig. 1, right). Alice (Bob) conducts measurement x (y) on her (his) particle, obtaining the outcome a (b). As pointed out by Bell^{[22](#)}, there exist quantum experiments where the observed correlations, encapsulated by the measured probabilities $\langle P(a,b|x,y) \rangle$, are such that they cannot be reproduced by any local deterministic model. An experimental realization of such correlations disproves the universal validity of local classical physics.

Next we consider whether Conan could similarly refute real quantum theory via a (complex) quantum Bell experiment. Such an experiment should necessarily violate some Bell inequality; otherwise, one could reproduce the measured probabilities with diagonal (and hence real) density matrices and measurement operators. The mere observation of a Bell violation is, however, insufficient to disprove real quantum theory, as already exemplified by the famous Clauser–Horne–Shimony–Holt (CHSH) Bell inequality^{[23](#)} $\langle \langle x_1, x_2; y_1, y_2 \rangle \rangle = \langle \langle A_{x_1} B_{y_1} \rangle \rangle + \langle \langle A_{x_1} B_{y_2} \rangle \rangle - \langle \langle A_{x_2} B_{y_1} \rangle \rangle - \langle \langle A_{x_2} B_{y_2} \rangle \rangle$. The inequality is derived for a Bell

experiment where Alice and Bob perform two measurements with outcomes ± 1 , and where $A_x (B_y)$ denotes the results by Alice (Bob) when performing measurement $x (y)$.

The maximal quantum violation of this inequality is $\sqrt{2}$ and Alice and Bob can attain it using real measurements on a real two-qubit state.

To find a gap between the predictions of real and complex quantum theory, Conan shall explore more complicated Bell inequalities. A priori, promising candidates are the elegant inequality of ref. [24](#) or the combination of three CHSH inequalities introduced in refs. [25,26](#)

$$\text{CHSH}_3 := \text{CHSH}(1,2;1,2) + \text{CHSH}(1,3;3,4) + \text{CHSH}(2,3;5,6) \leq 6 \quad (2)$$

designed for a scenario in which Alice and Bob perform three and six measurements, respectively. The maximal violation of inequality (2) is $6\sqrt{2}$, which can be attained with complex measurements on qubits [26](#).

However, none of these Bell inequalities will work: as shown in refs. [4,27,28](#), real quantum Bell experiments can reproduce the statistics of any quantum Bell experiment, even if conducted by more than two separate parties. Indeed, the construction of equation (1) for single complex quantum systems can be adapted to the multipartite case if we allow the source to distribute an extra qubit for each observer (see Fig. 1, right, for details).

To defeat Regina, Conan may also look for inspiration to other no-go theorems in quantum theory, such as the Pusey–Barrett–Rudolph construction [29](#) involving states prepared in independent labs subject to joint measurements. Unfortunately, Regina is again able to provide an explanation to such scenarios using real quantum theory ([Supplementary Information](#)). At this point, Conan might give up and accept that he will never change Regina’s mind. He would not be alone. For years, it was generally accepted that real quantum theory was experimentally indistinguishable from complex quantum theory. In other words, in quantum theory, complex numbers would only be convenient, but not necessary, to make sense of quantum experiments. Next we prove this conclusion wrong.

All it takes for Conan to win the discussion is to go beyond the previous constructions and consider experimental scenarios where independent sources prepare and send entangled states to several parties, who in turn conduct independent measurements [30,31,32,33,34](#). Such general network scenarios correspond to the future

quantum internet, which will connect many quantum computers and guarantee quantum confidentiality over continental distances. Our results demonstrate how these networks, beyond their practical relevance, open radically new perspectives to solve open questions in the foundations of quantum theory when exploiting the causal constraints associated with their geometries.

To disprove real quantum theory, Conan proposes the network corresponding to a standard entanglement-swapping scenario, depicted in Fig. 2, consisting of two independent sources and three observers: Alice, Bob and Charlie. The two sources prepare two maximally entangled states of two qubits, the first one $\langle \bar{\sigma}_A | \bar{\sigma}_B \rangle$ distributed to Alice and Bob; and the second $\langle \bar{\sigma}_B | \bar{\sigma}_C \rangle$, to Bob and Charlie. Bob performs a standard Bell-state measurement on the two particles that he receives from the two sources. This measurement has the effect of swapping the entanglement from Alice and Bob and Bob and Charlie to Alice and Charlie: namely, for each of Bob's four possible outcomes, Alice and Charlie share a two-qubit entangled state. Note that the actual state depends on Bob's outcome, but not on its degree of entanglement, which is always maximal. Alice and Charlie implement the measurements leading to the maximal violation of the CHSH₃ inequality (2). For these measurements, the state shared by Alice and Charlie, conditioned on Bob's result, maximally violates the inequality or a variant thereof produced by simple relabellings of the measurement outcomes.

Fig. 2: Network scenario separating real and complex quantum theory.

 **figure2**

In complex quantum theory (top), two independent sources distribute the two-qubit states $\langle \bar{\sigma}_A | \bar{\sigma}_B \rangle$ and $\langle \bar{\sigma}_B | \bar{\sigma}_C \rangle$ to, respectively, Alice and Bob, and Bob and Charlie. At Bob's location, a Bell measurement, of four outputs, is implemented. Alice and Charlie apply the complex measurements leading to the maximal violation of the CHSH₃ inequality: three and six measurements with two possible outputs, labelled by ± 1 . According to quantum physics, the observed correlations read $\langle P(abc|xz) = \text{Tr}(\bar{\sigma}_A \otimes \bar{\sigma}_B \otimes \bar{\sigma}_C)$. These correlations cannot be reproduced, or even well approximated, when all the states and measurements in the network are constrained to be real operators of arbitrary dimension (bottom). The impossibility still holds if the two preparations are correlated through shared randomness (dashed arrows), resulting in correlations of the form $\langle P(abc|xz) = \sum_{\lambda} P(\lambda) \langle \tilde{\sigma}_A | \tilde{\sigma}_B | \tilde{\sigma}_C \rangle^{\lambda}$, where all operators are real.

Regina takes up Conan's challenge and seeks to reproduce the statistics predicted by Conan. As she works under the postulates $\langle 1 \rangle$ and $\langle 2 \rangle$ – $\langle 4 \rangle$, she models the experiment of Fig. 2 as follows: each subsystem is represented by a real

Hilbert space \mathcal{H}_S for $S = A, B_1, B_2, C$, the states of the two sources are arbitrary real density matrices acting on $(\{\mathcal{H}\}_{A} \otimes \{\mathcal{H}\}_{B_1})$ and $(\{\mathcal{H}\}_{B_2} \otimes \{\mathcal{H}\}_{C})$, respectively, and the arbitrary real measurements act on $(\{\mathcal{H}\}_{A} \otimes \{\mathcal{H}\}_{B_1})$ and \mathcal{H}_C respectively. For each choice of states and measurements, she computes the probabilities via the Born rule. Regina's goal is to search over all states and measurements of the aforementioned form, acting on real Hilbert spaces of arbitrary dimension, until she can match Conan's predictions.

However, no construction by Regina is able to reproduce the measurement probabilities $\langle P(a,b,c|x,z) \rangle$ observed in the experiment. The proof, given in [Supplementary Information](#), exploits the results of ref. [26](#), where all quantum realizations leading to the maximal quantum value of inequality (2) were characterized. From this characterization, we show that the marginal state shared by Alice and Charlie at the beginning of the experiment cannot be decomposed as a convex combination of real product states^{[35](#)}, as the network depicted in Fig. [2](#) requires, and thus the statement follows. We moreover show the result to be robust, in the sense that the impossibility of real simulation also holds for non-maximal violations of the inequality (2) between Alice and Charlie. This result settles the argument between Conan and Regina: as she cannot accommodate Conan's experimental observations within the real quantum framework, she must admit her defeat.

A different question now is whether it is experimentally feasible to disprove real quantum theory. To assess this, it is convenient to rephrase our impossibility result in terms of a Bell-type parameter, that is, $\langle \mathcal{T} \rangle$, a linear function of the observed correlations. To this aim, we propose the Bell-type functional, defined by the sum of the violations of (the variants of) the CHSH₃ inequality for each of Bob's measurement outputs, weighted by the probability of the output. In the ideal entanglement-swapping realization with two-qubit maximally entangled states, the maximal quantum value of CHSH₃, equal to $6\sqrt{2}$, is obtained for each of the four outputs by Bob, so $\langle \mathcal{T} \rangle$ also attains its maximum quantum value, $\langle \mathcal{T} \rangle = 6\sqrt{2} \approx 8.49$. In [Supplementary Information](#), we explain how to reduce the problem of upper bounding $\langle \mathcal{T} \rangle$ to a convex optimization problem, making use of the hierarchies^{[28,36,37,38](#)}, that we solve numerically^{[39,40](#)}, for real quantum systems, to give $\langle \mathcal{T} \rangle \leq 7.66$. It remains open whether this upper bound is tight. As the map $\langle \mathcal{T} \rangle$ is a linear function of the observed probabilities, the impossibility result holds even when the real simulation is assisted by shared randomness (Fig. [2](#), bottom). As shown in refs. [41,42](#), this feature allows one to drop the assumption of independent and identical realizations in multiple-round hypothesis tests.

The setup needed to experimentally falsify real quantum theory is very similar to the bilocality scenario described in ref. ³⁰, for which several experimental implementations have been reported^{[43,44,45,46](#)}. Beating the real bound on $\langle\langle \mathscr{T} \rangle\rangle$ requires the two distributed states to have each a visibility beyond $\sqrt{7.66/6}/\sqrt{2} \approx 0.95$, a value attained in several experimental labs worldwide. The experiment similarly relies on the implementation of a challenging^{[47](#)} but feasible^{[48](#)} two-qubit entangled measurement. All things considered, we believe that an experimental disproof of real quantum physics based on the inequality $\langle\langle \mathscr{T} \rangle\rangle$ is within reach of current quantum technology (see [Supplementary Information](#) for more details).

Since the birth of modern science four centuries ago, abstract mathematical entities have played a big role in formalizing physical concepts. Our current understanding of velocity was only possible through the introduction of derivatives. The modern conception of gravity is attributable to the invention of non-Euclidean geometry. Basic notions from representation theory made it possible to formalize the notion of a fundamental particle. Here we considered whether the same holds for the complex numbers. Somewhat surprisingly, we found that there do exist natural scenarios that require the use of complex numbers to account for experimental observations within the standard Hilbert space formulation of quantum theory. As it turns out, some such experiments are within reach of current experimental capabilities, so it is not unreasonable to expect a convincing experimental disproof of real quantum theory in the near future.

From a broader point of view, our results advance the research programme, started in ref. ⁴⁹, of singling out quantum correlations by demanding maximal performance in a device-independent information-theoretic task. In this regard, our work shows that complex quantum theory outperforms real quantum theory when the non-local game $\langle\langle \mathscr{T} \rangle\rangle$ is played in the entanglement-swapping scenario. This game can be interpreted as an extension of the adaptive CHSH game proposed in ref. ⁴⁹, which was recently shown to rule out a number of alternative physical theories in favour of quantum theory^{[50](#)}. Whether the average score of $\langle\langle \mathscr{T} \rangle\rangle$ is maximized by complex quantum theory, or whether any physical theory other than complex quantum theory must necessarily produce a lower score are intriguing questions that we leave open.

Data availability

There are no data to be shared.

Code availability

The MATLAB codes used to prove Theorem 3 in the [Supplementary Information](#) can be found in the [Supplementary Data](#).

References

1. 1.

Dirac, P. *The Principles of Quantum Mechanics* (International Series of Monographs on Physics, Clarendon Press, 1958).

2. 2.

von Neumann, J. & Beyer, R. *Mathematical Foundations of Quantum Mechanics* (Princeton Univ. Press, 1955).

3. 3.

Einstein, A., Przibram, K. & Klein, M. *Letters on Wave Mechanics: Correspondence with H. A. Lorentz, Max Planck, and Erwin Schrödinger* (Philosophical Library/Open Road, 2011).

4. 4.

McKague, M., Mosca, M. & Gisin, N. Simulating quantum systems using real Hilbert spaces. *Phys. Rev. Lett.* **102**, 020505 (2009).

5. 5.

Ismael, J. in *The Stanford Encyclopedia of Philosophy Fall 2021 Edition* (ed. Zalta, E. N.) (Metaphysics Research Lab, Stanford University, 2021);
<https://plato.stanford.edu/archives/fall2021/entries/qm/>

6. 6.

Piron, C. Axiomatique quantique. *Helv. Phys. Acta* **37**, 439–468 (1964).

7. 7.

Werner, R. Local preparability of states and the split property in quantum field theory. *Lett. Math. Phys.* **13**, 325–329 (1987).

8. 8.

Buchholz, D. & Jacobi, P. On the nuclearity condition for massless fields. *Lett. Math. Phys.* **13**, 313–323 (1987).

9. 9.

Buchholz, D. Product states for local algebras. *Commun. Math. Phys.* **36**, 287–304 (1974).

10. 10.

Summers, S. J. Normal product states for fermions and twisted duality for CCR- and CAR-type algebras with application to the Yukawa₂ quantum field model. *Commun. Math. Phys.* **86**, 111–141 (1982).

11. 11.

Wootters, W. K. in *Complexity, Entropy, and the Physics of Information* (ed. Zurek, W. H.) 39–46 (Addison-Wesley, 1990).

12. 12.

Caves, C. M., Fuchs, C. A. & Schack, R. Unknown quantum states: the quantum de finetti representation. *J. Math. Phys.* **43**, 4537 (2002).

13. 13.

Shankar, R. *Principles of Quantum Mechanics* (Plenum, 1980).

14. 14.

Fuchs, C. A. & Schack, R. Quantum-Bayesian coherence. *Rev. Mod. Phys.* **85**, 1693 (2013).

15. 15.

Wigner, E. On the quantum correction for thermodynamic equilibrium. *Phys. Rev.* **40**, 749 (1932).

16. 16.

Bohm, D. A suggested interpretation of the quantum theory in terms of “hidden” variables. I. *Phys. Rev.* **85**, 166 (1952).

17. 17.

Stueckelberg, E. C. Quantum theory in real Hilbert space. *Helv. Phys. Acta* **33**, 727–752 (1960).

18. 18.

Aleksandrova, A., Borish, V. & Wootters, W. K. Real-vector-space quantum theory with a universal quantum bit. *Phys. Rev. A* <https://doi.org/10.1103/physreva.87.052106> (2013).

19. 19.

Wu, K.-D. et al. Operational resource theory of imaginarity. *Phys. Rev. Lett.* **126**, 090401 (2021).

20. 20.

Brunner, N. et al. Testing the dimension of Hilbert spaces. *Phys. Rev. Lett.* **100**, 210503 (2008).

21. 21.

Schrödinger, E. Discussion of probability relations between separated systems. *Math. Proc. Cambridge Phil. Soc.* **31**, 555–563 (1935).

22. 22.

Bell, J. S. On the Einstein, Podolsky, Rosen paradox. *Physics* **1**, 195 (1964).

23. 23.

Clauser, J. F., Horne, M. A., Shimony, A. & Holt, R. A. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.* **23**, 880 (1969).

24. 24.

Gisin, N. in *Quantum Reality, Relativistic Causality, and Closing the Epistemic Circle: Essays in Honour of Abner Shimony* (eds Myrvold, W. C. & Christian, J.) 125–138 (Springer, 2009).

25. 25.

Aćin, A., Pironio, S., Vértesi, T. & Wittek, P. Optimal randomness certification from one entangled bit. *Phys. Rev. A* **93**, 040102 (2016).

26. 26.

Bowles, J., Šupić, I., Cavalcanti, I. & Acin, A. Self-testing of Pauli observables for device-independent entanglement certification. *Phys. Rev. A* **98**, 042336 (2018).

27. 27.

Pál, K. F. & Vértesi, T. Efficiency of higher-dimensional Hilbert spaces for the violation of Bell inequalities. *Phys. Rev. A* **77**, 042105 (2008).

28. 28.

Moroder, T., Bancal, J.-D., Liang, Y.-C., Hofmann, M. & Gühne, O. Device-independent entanglement quantification and related applications. *Phys. Rev. Lett.* **111**, 030501 (2013).

29. 29.

Pusey, M. F., Barrett, J. & Rudolph, T. On the reality of the quantum state. *Nat. Phys.* **8**, 475–478 (2012).

30. 30.

Branciard, C., Gisin, N. & Pironio, S. Characterizing the nonlocal correlations created via entanglement swapping. *Phys. Rev. Lett.* **104**, 170401 (2010).

31. 31.

Fritz, T. Beyond Bell’s theorem: correlation scenarios. *New J. Phys.* **14**, 103001 (2012).

32. 32.

Fritz, T. Beyond Bell’s theorem II: scenarios with arbitrary causal structure. *Commun. Math. Phys.* **341**, 391–434 (2016).

33. 33.

Renou, M. O., Kaniewski, J. M. K. & Brunner, N. Self-testing entangled measurements in quantum networks. *Phys. Rev. Lett.* **121**, 250507 (2018).

34. 34.

Bancal, J.-D., Sangouard, N. & Sekatski, P. Noise-resistant device-independent certification of Bell state measurements. *Phys. Rev. Lett.* **121**, 250506 (2018).

35. 35.

Caves, C. M., Fuchs, C. A. & Rungta, P. Entanglement of formation of an arbitrary state of two rebits. *Found. Phys. Lett.* **14**, 199–212 (2001).

36. 36.

Navascués, M., Pironio, S. & Acín, A. Bounding the set of quantum correlations. *Phys. Rev. Lett.* **98**, 010401 (2007).

37. 37.

Navascués, M., Pironio, S. & Acín, A. A convergent hierarchy of semidefinite programs characterizing the set of quantum correlations. *New J. Phys.* **10**, 073013 (2008).

38. 38.

Pironio, S., Navascués, M. & Acín, A. Convergent relaxations of polynomial optimization problems with noncommuting variables. *SIAM J. Optim.* **20**, 2157 (2010).

39. 39.

Löfberg, J. YALMIP: a toolbox for modeling and optimization in MATLAB. In *2004 IEEE International Conference on Robotics and Automation* 284–289 (IEEE, 2004).

40. 40.

Vandenberghe, L. & Boyd, S. *The MOSEK Optimization Toolbox for MATLAB Manual Version 9.3.10* (MOSEK ApS, 2021).

41. 41.

Elkouss, D. & Wehner, S. (Nearly) optimal P values for all Bell inequalities. *npj Quantum Inf.* **2**, 16026 (2016).

42. 42.

Aráujo, M., Hirsch, F. & Quintino, M. T. Bell nonlocality with a single shot. *Quantum* **4**, 353 (2020).

43. 43.

Carvacho, G. et al. Experimental violation of local causality in a quantum network. *Nat. Commun.* **8**, 14775 (2017).

44. 44.

Saunders, D. J., Bennet, A. J., Branciard, C. & Pryde, G. J. Experimental demonstration of nonbilocal quantum correlations. *Sci. Adv.* **3**, e1602743 (2017).

45. 45.

Sun, Q.-C. et al. Experimental violation of local causality in a quantum network. *Nat. Photon.* **13**, 687–691 (2019).

46. 46.

Bäumer, E., Gisin, N. & Tavakoli, A. Demonstrating the power of quantum computers, certification of highly entangled measurements and scalable quantum nonlocality. *npj Quant. Inf.* **7**, 117 (2021).

47. 47.

Calsamiglia, J. & Lütkenhaus, N. Maximum efficiency of a linear-optical Bell-state analyzer. *Appl. Phys. B* **72**, 67–71 (2001).

48. 48.

Grice, W. P. Arbitrarily complete Bell-state measurement using only linear optical elements. *Phys. Rev. A* **84**, 042331 (2011).

49. 49.

Weilenmann, M. & Colbeck, R. Self-testing of physical theories, or, is quantum theory optimal with respect to some information-processing task? *Phys. Rev. Lett.* **125**, 060406 (2020).

50. 50.

Weilenmann, M. & Colbeck, R. Toward correlation self-testing of quantum theory in the adaptive Clauser–Horne–Shimony–Holt game. *Phys. Rev. A* **102**,

022203 (2020).

Acknowledgements

D.T. is a recipient of the DOC Fellowship OÖAS000047 of the Austrian Academy of Sciences at the Institute of Quantum Optics and Quantum Information (IQOQI), Vienna. M.W. and T.P.L. are each supported by a Lise Meitner Fellowship awarded by Fonds zur Förderung der wissenschaftlichen Forschung (FWF), with project numbers M 3109-N and M 2812-N, respectively. M.-O.R. and A.T. are supported by the Swiss National Fund Early Mobility Grants P2GEP2_191444 and P2GEP2 194800, respectively. We acknowledge support from the Government of Spain (FIS2020-TRANQI and Severo Ochoa CEX2019-000910-S), Fundacio Cellex, Fundacio Mir-Puig, Generalitat de Catalunya (CERCA, AGAUR SGR 1381 and QuantumCAT), the ERC AdG CERQUTE, the AXA Chair in Quantum Information Science and the Swiss NCCR SwissMap.

Author information

Affiliations

1. ICFO-Institut de Ciencies Fotoniques, The Barcelona Institute of Science and Technology, Castelldefels (Barcelona), Spain

Marc-Olivier Renou & Antonio Acín

2. Institute for Quantum Optics and Quantum Information (IQOQI) Vienna, Austrian Academy of Sciences, Vienna, Austria

David Trillo, Mirjam Weilenmann, Thinh P. Le, Armin Tavakoli & Miguel Navascués

3. Institute for Atomic and Subatomic Physics, Vienna University of Technology, Vienna, Austria

Armin Tavakoli

4. Group of Applied Physics, University of Geneva, Geneva, Switzerland

Nicolas Gisin

5. Schaffhausen Institute of Technology–SIT, Geneva, Switzerland

Nicolas Gisin

6. ICREA-Institució Catalana de Recerca i Estudis Avançats, Barcelona, Spain

Antonio Acín

Contributions

All authors contributed equally to this work.

Corresponding author

Correspondence to [Miguel Navascués](#).

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 and Data Sections 1–9.

[Supplementary Data](#)

MATLAB codes.

[Peer Review File](#)

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

Renou, MO., Trillo, D., Weilenmann, M. *et al.* Quantum theory based on real numbers can be experimentally falsified. *Nature* **600**, 625–629 (2021).
<https://doi.org/10.1038/s41586-021-04160-4>

- Received: 28 January 2021
- Accepted: 18 October 2021
- Published: 15 December 2021
- Issue Date: 23 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04160-4>

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

Further reading

- [Alternatives to standard quantum theory ruled out](#)

◦ William K. Wootters

Nature (2021)

[Alternatives to standard quantum theory ruled out](#)

- William K. Wootters

News & Views 15 Dec 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04160-4>

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

- Article
- [Published: 22 December 2021](#)

Programmable interactions and emergent geometry in an array of atom clouds

- [Avikar Periwal](#) [ORCID: orcid.org/0000-0001-7601-0932^{1 na1}](#),
- [Eric S. Cooper^{1 na1}](#),
- [Philipp Kunkel](#) [ORCID: orcid.org/0000-0002-8145-9184^{1,2 na1}](#),
- [Julian F. Wienand](#) [ORCID: orcid.org/0000-0001-9347-0850^{1,3}](#),
- [Emily J. Davis¹](#) &
- [Monika Schleier-Smith](#) [ORCID: orcid.org/0000-0002-4686-3528^{1,2}](#)

[Nature](#) volume **600**, pages 630–635 (2021)

- 1532 Accesses
- 18 Altmetric
- [Metrics details](#)

Subjects

- [Quantum optics](#)
- [Quantum simulation](#)
- [Ultracold gases](#)

Abstract

Interactions govern the flow of information and the formation of correlations between constituents of many-body quantum systems, dictating phases of matter found in nature and forms of entanglement generated in the laboratory. Typical interactions decay with distance and thus produce a network of connectivity governed by geometry—such as the crystalline structure of a material or the trapping sites of atoms in a quantum simulator^{1,2}. However, many envisioned applications in quantum simulation and computation require more complex coupling graphs including non-local interactions, which feature in models of information scrambling in black holes^{3,4,5,6} and mappings of hard optimization problems onto frustrated classical magnets^{7,8,9,10,11}. Here we describe the realization of programmable non-local interactions in an array of atomic ensembles within an optical cavity, in which photons carry information between atomic spins^{12,13,14,15,16,17,18,19}. By programming the distance dependence of the interactions, we access effective geometries for which the dimensionality, topology and metric are entirely distinct from the physical geometry of the array. As examples, we engineer an antiferromagnetic triangular ladder, a Möbius strip with sign-changing interactions and a treelike geometry inspired by concepts of quantum gravity^{5,20,21,22}. The tree graph constitutes a toy model of holographic duality^{21,22}, in which the quantum system lies on the boundary of a higher-dimensional geometry that emerges from measured correlations²³. Our work provides broader prospects for simulating frustrated magnets and topological phases²⁴, investigating quantum optimization paradigms^{10,11,25,26} and engineering entangled resource states for sensing and computation^{27,28}.

This is a preview of subscription content

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Fig. 1: Engineering distance-dependent interactions.



Fig. 2: Pair creation at programmable distance.

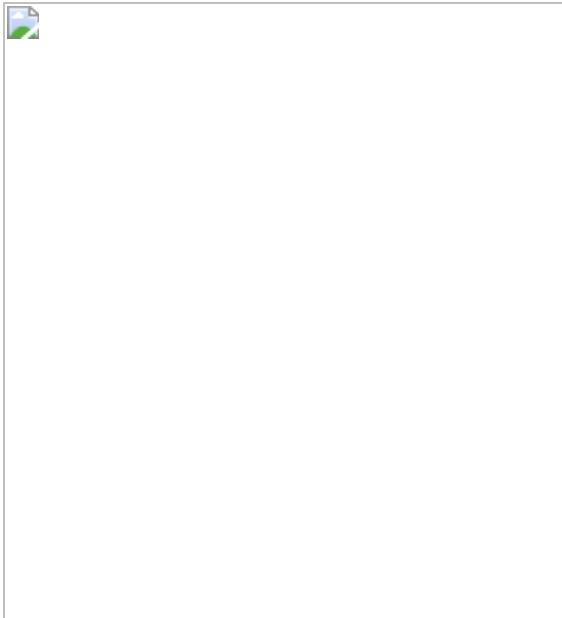


Fig. 3: Geometry extracted from correlations.

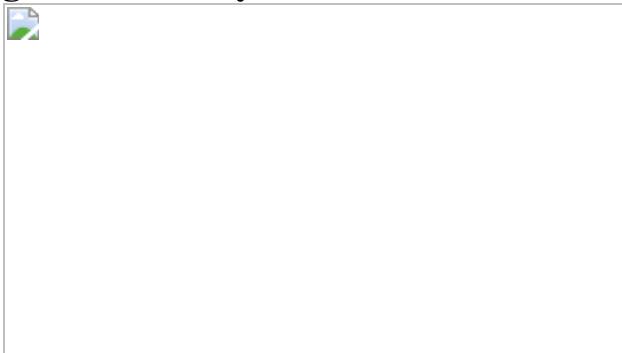
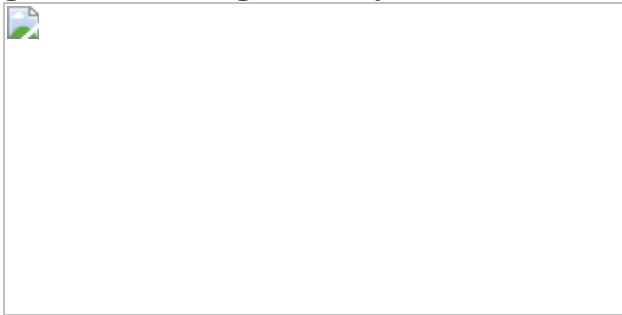


Fig. 4: Treelike geometry.



Data availability

All data displayed in Figs. 1–4 and Extended Data Figs. 1–4 are available from the corresponding author upon reasonable request.

Code availability

All code used for simulation and analysis is available from the corresponding author upon reasonable request.

References

1. 1.

Bloch, I., Dalibard, J. & Nascimbene, S. Quantum simulations with ultracold quantum gases. *Nat. Phys.* **8**, 267–276 (2012).

2. 2.

Browaeys, A. & Lahaye, T. Many-body physics with individually controlled Rydberg atoms. *Nat. Phys.* **16**, 132–142 (2020).

3. 3.

Hayden, P. & Preskill, J. Black holes as mirrors: quantum information in random subsystems. *J. High Energy Phys.* **2007**, 120–120 (2007).

4. 4.

Maldacena, J. & Stanford, D. Remarks on the Sachdev-Ye-Kitaev model. *Phys. Rev. D* **94**, 106002 (2016).

5. 5.

Bentsen, G. et al. Treelike interactions and fast scrambling with cold atoms. *Phys. Rev. Lett.* **123**, 130601 (2019).

6. 6.

Belyansky, R., Bienias, P., Kharkov, Y. A., Gorshkov, A. V. & Swingle, B. Minimal model for fast scrambling. *Phys. Rev. Lett.* **125**, 130601 (2020).

7. 7.

Das, A. & Chakrabarti, B. K. Colloquium: quantum annealing and analog quantum computation. *Rev. Mod. Phys.* **80**, 1061–1081 (2008).

8. 8.

Gopalakrishnan, S., Lev, B. L. & Goldbart, P. M. Frustration and glassiness in spin models with cavity-mediated interactions. *Phys. Rev. Lett.* **107**, 277201 (2011).

9. 9.

Strack, P. & Sachdev, S. Dicke quantum spin glass of atoms and photons. *Phys. Rev. Lett.* **107**, 277202 (2011).

10. 10.

McMahon, P. L. et al. A fully programmable 100-spin coherent Ising machine with all-to-all connections. *Science* **354**, 614–617 (2016).

11. 11.

Berloff, N. G. et al. Realizing the classical XY Hamiltonian in polariton simulators. *Nat. Mater.* **16**, 1120–1126 (2017).

12. 12.

Leroux, I. D., Schleier-Smith, M. H. & Vuletić, V. Implementation of cavity squeezing of a collective atomic spin. *Phys. Rev. Lett.* **104**, 073602 (2010).

13. 13.

Barontini, G., Hohmann, L., Haas, F., Estève, J. & Reichel, J. Deterministic generation of multiparticle entanglement by quantum Zeno dynamics. *Science* **349**, 1317–1321 (2015).

14. 14.

Hosten, O., Krishnakumar, R., Engelsen, N. J. & Kasevich, M. A. Quantum phase magnification. *Science* **352**, 1552–1555 (2016).

15. 15.

Pedrozo-Peñaflor, E. et al. Entanglement on an optical atomic-clock transition. *Nature* **588**, 414–418 (2020).

16. 16.

Welte, S., Hacker, B., Daiss, S., Ritter, S. & Rempe, G. Photon-mediated quantum gate between two neutral atoms in an optical cavity. *Phys. Rev. X* **8**, 011018 (2018).

17. 17.

Davis, E. J., Bentsen, G., Homeier, L., Li, T. & Schleier-Smith, M. H. Photon-mediated spin-exchange dynamics of spin-1 atoms. *Phys. Rev. Lett.* **122**, 010405 (2019).

18. 18.

Davis, E. J. et al. Protecting spin coherence in a tunable Heisenberg model. *Phys. Rev. Lett.* **125**, 060402 (2020).

19. 19.

Muniz, J. A. et al. Exploring dynamical phase transitions with cold atoms in an optical cavity. *Nature* **580**, 602–607 (2020).

20. 20.

Barbón, J. L. & Magán, J. M. Fast scramblers and ultrametric black hole horizons. *J. High Energy Phys.* **2013**, 163 (2013).

21. 21.

Gubser, S. S., Knaute, J., Parikh, S., Samberg, A. & Witaszczyk, P. p -adic AdS/CFT. *Commun. Math. Phys.* **352**, 1019–1059 (2017).

22. 22.

Heydeman, M., Marcolli, M., Saberi, I. & Stoica, B. Tensor networks, p-adic fields, and algebraic curves: arithmetic and the AdS3/CFT2 correspondence. Preprint at <https://arxiv.org/abs/1605.07639> (2016).

23. 23.

Qi, X.-L. Does gravity come from quantum information? *Nat. Phys.* **14**, 984–987 (2018).

24. 24.

Hung, C.-L., González-Tudela, A., Cirac, J. I. & Kimble, H. J. Quantum spin dynamics with pairwise-tunable, long-range interactions. *Proc. Natl Acad. Sci. USA* **113**, E4946–E4955 (2016).

25. 25.

Marsh, B. P. et al. Enhancing associative memory recall and storage capacity using confocal cavity QED. *Phys. Rev. X* **11**, 021048 (2021).

26. 26.

Anikeeva, G. et al. Number partitioning with Grover’s algorithm in central spin systems. *PRX Quantum* **2**, 020319 (2021).

27. 27.

Masson, S. J., Barrett, M. & Parkins, S. Cavity QED engineering of spin dynamics and squeezing in a spinor gas. *Phys. Rev. Lett.* **119**, 213601 (2017).

28. 28.

Qin, Z. et al. Characterizing the multipartite continuous-variable entanglement structure from squeezing coefficients and the Fisher information. *npj Quantum Inf.* **5**, 3 (2019).

29. 29.

Pezzè, L., Smerzi, A., Oberthaler, M. K., Schmied, R. & Treutlein, P. Quantum metrology with nonclassical states of atomic ensembles. *Rev. Mod. Phys.* **90**, 035005 (2018).

30. 30.

Omran, A. et al. Generation and manipulation of Schrödinger cat states in Rydberg atom arrays. *Science* **365**, 570–574 (2019).

31. 31.

Hamley, C. D., Gerving, C., Hoang, T., Bookjans, E. & Chapman, M. S. Spin-nematic squeezed vacuum in a quantum gas. *Nat. Phys.* **8**, 305–308 (2012).

32. 32.

Fadel, M., Zibold, T., Décamps, B. & Treutlein, P. Spatial entanglement patterns and Einstein-Podolsky-Rosen steering in Bose-Einstein condensates. *Science* **360**, 409–413 (2018).

33. 33.

Kunkel, P. et al. Spatially distributed multipartite entanglement enables EPR steering of atomic clouds. *Science* **360**, 413–416 (2018).

34. 34.

Lange, K. et al. Entanglement between two spatially separated atomic modes. *Science* **360**, 416–418 (2018).

35. 35.

Islam, R. et al. Emergence and frustration of magnetism with variable-range interactions in a quantum simulator. *Science* **340**, 583–587 (2013).

36. 36.

Jurcevic, P. et al. Quasiparticle engineering and entanglement propagation in a quantum many-body system. *Nature* **511**, 202–205 (2014).

37. 37.

Vaidya, V. D. et al. Tunable-range, photon-mediated atomic interactions in multimode cavity QED. *Phys. Rev. X* **8**, 011002 (2018).

38. 38.

Manovitz, T., Shapira, Y., Akerman, N., Stern, A. & Ozeri, R. Quantum simulations with complex geometries and synthetic gauge fields in a trapped ion chain. *PRX Quantum* **1**, 020303 (2020).

39. 39.

Kollár, A. J., Fitzpatrick, M. & Houck, A. A. Hyperbolic lattices in circuit quantum electrodynamics. *Nature* **571**, 45–50 (2019).

40. 40.

Lücke, B. et al. Detecting multiparticle entanglement of Dicke states. *Phys. Rev. Lett.* **112**, 155304 (2014).

41. 41.

Lee, S. & Lee, K.-C. Phase transitions in the fully frustrated triangular XY model. *Phys. Rev. B* **57**, 8472 (1998).

42. 42.

Gubser, S. S., Jepsen, C., Ji, Z. & Trundy, B. Continuum limits of sparse coupling patterns. *Phys. Rev. D* **98**, 045009 (2018).

43. 43.

Shi, Y.-Y., Duan, L.-M. & Vidal, G. Classical simulation of quantum many-body systems with a tree tensor network. *Phys. Rev. A* **74**, 022320 (2006).

44. 44.

Murg, V., Verstraete, F., Legeza, Ö. & Noack, R. M. Simulating strongly correlated quantum systems with tree tensor networks. *Phys. Rev. B* **82**, 205105 (2010).

45. 45.

Pastawski, F., Yoshida, B., Harlow, D. & Preskill, J. Holographic quantum error-correcting codes: toy models for the bulk/boundary correspondence. *J. High Energ. Phys.* **2015**, 149 (2015).

46. 46.

Amit, D. J., Gutfreund, H. & Sompolinsky, H. Spin-glass models of neural networks. *Phys. Rev. A* **32**, 1007–1018 (1985).

47. 47.

Aidelsburger, M. et al. Measuring the Chern number of Hofstadter bands with ultracold bosonic atoms. *Nat. Phys.* **11**, 162–166 (2015).

48. 48.

Kennedy, C. J., Burton, W. C., Chung, W. C. & Ketterle, W. Observation of Bose–Einstein condensation in a strong synthetic magnetic field. *Nat. Phys.* **11**, 859–864 (2015).

49. 49.

Zeiher, J. et al. Microscopic characterization of scalable coherent Rydberg superatoms. *Phys. Rev. X* **5**, 031015 (2015).

50. 50.

Evans, R. E. et al. Photon-mediated interactions between quantum emitters in a diamond nanocavity. *Science* **362**, 662–665 (2018).

51. 51.

Lee, J., Vrijen, G., Teper, I., Hosten, O. & Kasevich, M. A. Many-atom–cavity QED system with homogeneous atom–cavity coupling. *Opt. Lett.* **39**, 4005–4008 (2014).

52. 52.

Xu, V. et al. Probing gravity by holding atoms for 20 seconds. *Science* **366**, 745–749 (2019).

53. 53.

Sørensen, A. S. & Mølmer, K. Entangling atoms in bad cavities. *Phys. Rev. A* **66**, 022314 (2002).

54. 54.

Stamper-Kurn, D. M. & Ueda, M. Spinor Bose gases: symmetries, magnetism, and quantum dynamics. *Rev. Mod. Phys.* **85**, 1191 (2013).

55. 55.

Sinatra, A., Lobo, C. & Castin, Y. The truncated Wigner method for Bose-condensed gases: limits of validity and applications. *J. Phys. B* **35**, 3599–3631 (2002).

56. 56.

Blakie, P. B., Bradley, A. S., Davis, M. J., Ballagh, R. J. & Gardiner, C. W. Dynamics and statistical mechanics of ultra-cold Bose gases using c-field techniques. *Adv. Phys.* **57**, 363–455 (2008).

57. 57.

Lewis-Swan, R. J., Olsen, M. K. & Kheruntsyan, K. V. Approximate particle number distribution from direct stochastic sampling of the Wigner function. *Phys. Rev. A* **94**, 033814 (2016).

58. 58.

Lauritzen, S. L. *Graphical Models* Vol. 17 (Clarendon Press, 1996).

Acknowledgements

We thank S. Gubser for discussions that inspired our exploration of non-Archimedean geometry. We also acknowledge discussions with G. Bentsen, A. Daley, I. Bloch, B. Lev, N. Berloff, A. Deshpande, B. Swingle and P. Hayden. This work was supported by the DOE Office of Science, Office of High Energy Physics and Office of Basic Energy Sciences under grant no. DE-SC0019174. A.P. and E.S.C. acknowledge support from the NSF under grant no. PHY-1753021. We additionally acknowledge support from the National Defense Science and Engineering Graduate Fellowship (A.P.), the NSF Graduate Research Fellowship Program (E.J.D. and E.S.C.), the Hertz Foundation (E.J.D.) and the German Academic Scholarship Foundation (J.F.W.).

Author information

Author notes

1. These authors contributed equally: Avikar Periwal, Eric S. Cooper, Philipp Kunkel

Affiliations

1. Department of Physics, Stanford University, Stanford, CA, USA

Avikar Periwal, Eric S. Cooper, Philipp Kunkel, Julian F. Wienand, Emily J. Davis & Monika Schleier-Smith

2. SLAC National Accelerator Laboratory, Menlo Park, CA, USA

Philipp Kunkel & Monika Schleier-Smith

3. Fakultät für Physik, Ludwig-Maximilians-Universität, Munich,
Germany

Julian F. Wienand

Contributions

A.P., E.S.C., P.K., J.F.W. and E.J.D. performed the experiments. A.P., E.S.C., P.K. and M.S.-S. analysed the experimental data and developed supporting theoretical models. A.P., E.S.C., P.K. and M.S.-S. wrote the manuscript. All authors contributed to the discussion and interpretation of results.

Corresponding author

Correspondence to [Monika Schleier-Smith](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Robert Lewis-Swan and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

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 Coupling graphs.

Sketch of couplings $\langle J(i-j) \rangle$ for the model in equation (4) with local interactions ($s = -1$, left) or treelike interactions ($s = 1$, right). The strengths of the interactions are indicated by the thickness and transparency of the red lines. For $s = 1$, reordering the sites according to the Monna map makes the couplings more local, corroborating the treelike geometry.

Extended Data Fig. 2 Experimental sequence and imaging.

a Schematic of experimental sequence for measurements of $\langle F_i^x \rangle$. After driving the cavity to induce interactions, we apply spin rotations sequentially to the M sites of the array and subsequently perform state-sensitive readout via fluorescence imaging. **b** Fluorescence images after spin rotation, showing the signal for the $F=2$ manifold and the three magnetic substates for the case of interactions at distance $r=3$ with periodic boundary conditions. **c** Transverse magnetization $\langle F_i^x \rangle$ and structure factor $\langle \tilde{F}_k^x \rangle$ extracted from the image in **b**.

Extended Data Fig. 3 Effect of finite statistics.

Left, correlation plot reproduced from Fig. 1, showing $\langle C_{\text{pm}} \rangle$ obtained from 50 realizations of the experiment with interactions at distance $r=10$. Right, simulation results obtained from a truncated Wigner approximation, where we either choose the same number of realisations as in the experiment or increase the number of realisations by a factor of 10 to reduce statistical uncertainty. The simulations indicate that residual correlations in the experimental data are mainly due to the finite sample size.

Extended Data Fig. 4 Comparison between measured structure factor and simulation results.

The left graph shows the measured structure factor after $T=3$ Bloch periods of evolution, which is also shown in Fig. 2c. The two plots at right show results of a truncated Wigner simulation with and without periodic

boundary conditions. For the simulated data we used 100 realizations of the TWA simulation, which is four times higher than the number of experimental realizations to reduce statistical fluctuations. For open boundary conditions, we find that the simulation has an offset with respect to the theoretical prediction (blue line). We attribute this offset to the finite system size, as the model is exact only for an infinite system or a system with periodic boundary conditions. Repeating the same simulation with a pulsed drive shown on the right shows that in this case the TWA simulation is consistent with the analytical model. The error bars indicate the standard error of the mean.

Extended Data Table 1 Experimental parameters

Supplementary information

Supplementary Figs. 1–4 and references.

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Periwal, A., Cooper, E.S., Kunkel, P. *et al.* Programmable interactions and emergent geometry in an array of atom clouds. *Nature* **600**, 630–635 (2021). <https://doi.org/10.1038/s41586-021-04156-0>

- Received: 07 June 2021
- Accepted: 15 October 2021
- Published: 22 December 2021
- Issue Date: 23 December 2021

- DOI: <https://doi.org/10.1038/s41586-021-04156-0>

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/s41586-021-04156-0>

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

- Article
- [Published: 22 December 2021](#)

Resonance from antiferromagnetic spin fluctuations for superconductivity in UTe₂

- [Chunruo Duan](#)¹,
- [R. E. Baumbach](#)^{2,3},
- [Andrey Podlesnyak](#) ORCID: orcid.org/0000-0001-9366-6319⁴,
- [Yuhang Deng](#)⁵,
- [Camilla Moir](#)⁵,
- [Alexander J. Breindel](#)⁵,
- [M. Brian Maple](#)⁵,
- [E. M. Nica](#)⁶,
- [Qimiao Si](#) ORCID: orcid.org/0000-0003-1357-2705¹ &
- [Pengcheng Dai](#) ORCID: orcid.org/0000-0002-6088-3170¹

Nature volume 600, pages 636–640 (2021)

- 2368 Accesses
- 41 Altmetric
- [Metrics details](#)

Subjects

- [Magnetic properties and materials](#)
- [Superconducting properties and materials](#)

Abstract

Superconductivity originates from the formation of bound (Cooper) pairs of electrons that can move through the lattice without resistance below the superconducting transition temperature T_c (ref. 1). Electron Cooper pairs in most superconductors form anti-parallel spin singlets with total spin $S=0$ (ref. 2), although they can also form parallel spin-triplet Cooper pairs with $S=1$ and an odd parity wavefunction³. Spin-triplet pairing is important because it can host topological states and Majorana fermions relevant for quantum computation^{4,5}. Because spin-triplet pairing is usually mediated by ferromagnetic (FM) spin fluctuations³, uranium-based materials near an FM instability are considered to be ideal candidates for realizing spin-triplet superconductivity⁶. Indeed, UTe₂, which has a $T_c \approx 1.6$ K (refs. 7,8), has been identified as a candidate for a chiral spin-triplet topological superconductor near an FM instability^{7,8,9,10,11,12,13,14}, although it also has antiferromagnetic (AF) spin fluctuations^{15,16}. Here we use inelastic neutron scattering (INS) to show that superconductivity in UTe₂ is coupled to a sharp magnetic excitation, termed resonance^{17,18,19,20,21,22,23}, at the Brillouin zone boundary near AF order. Because the resonance has only been found in spin-singlet unconventional superconductors near an AF instability^{17,18,19,20,21,22,23}, its observation in UTe₂ suggests that AF spin fluctuations may also induce spin-triplet pairing²⁴ or that electron pairing in UTe₂ has a spin-singlet component.

This is a preview of subscription content

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

Rent or Buy

All prices are NET prices.

Additional access options:

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

Fig. 1: Crystal structure, heat capacity and a summary of INS results.

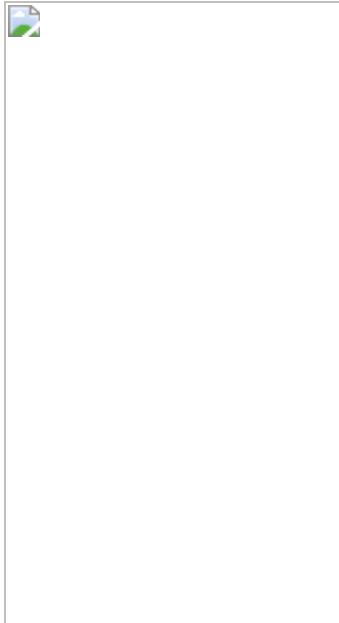


Fig. 2: The wavevector, energy and temperature dependence of the scattering function $S(Q, E)$ in the $[H, K, O]$ plane.

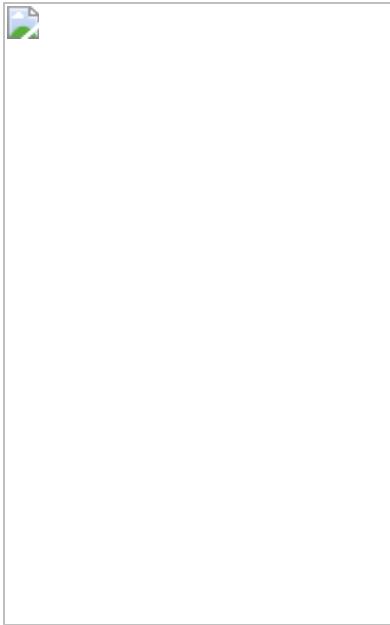


Fig. 3: The wavevector and energy dependence of the scattering below and above T_c .

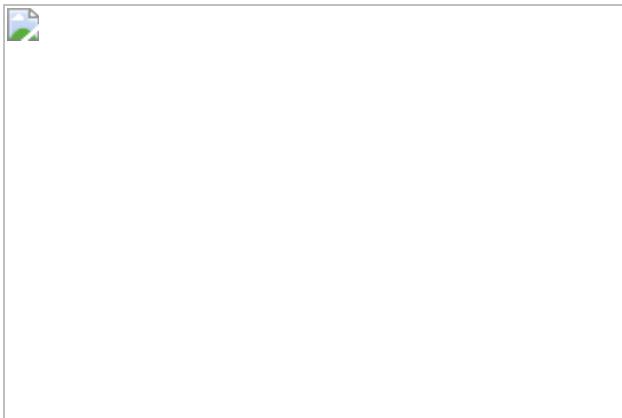
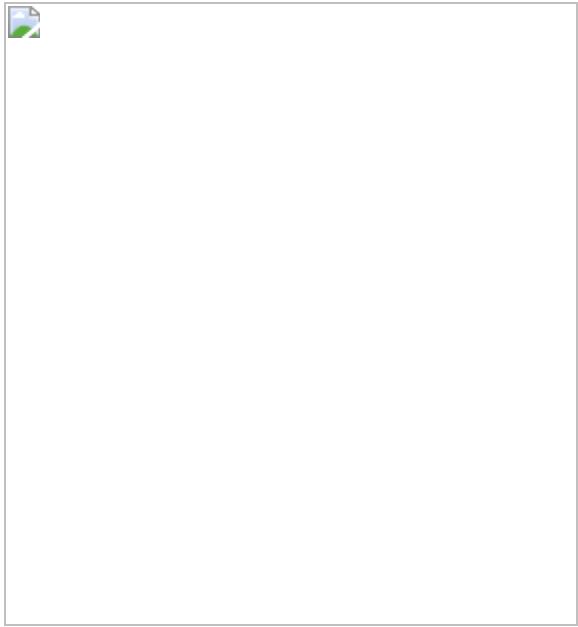


Fig. 4: The wavevector, energy and temperature dependence of the scattering at Brillouin zone boundary points.



Data availability

The data that support the plots within this paper and other findings of this study are available from the corresponding authors upon reasonable request.

References

1. 1.

Bardeen, J., Cooper, L. N. & Schrieffer, J. R. Theory of superconductivity. *Phys. Rev.* **108**, 1175–1204 (1957).

2. 2.

Scalapino, D. J. A common thread: the pairing interaction for unconventional superconductors. *Rev. Mod. Phys.* **84**, 1383–1417 (2012).

3. 3.

Mackenzie, A. P. & Maeno, Y. The superconductivity of Sr_2RuO_2 and the physics of spin-triplet pairing. *Rev. Mod. Phys.* **75**, 657–712

(2003).

4. 4.

Sato, M. & Ando, Y. Topological superconductors: a review. *Rep. Prog. Phys.* **80**, 076501 (2017).

5. 5.

Kitaev, A. Y. Unpaired Majorana fermions in quantum wires. *Phys.-Usp.* **44**, 131–136 (2001).

6. 6.

Aoki, D., Ishida, K. & Flouquet, J. Review of U-based ferromagnetic superconductors: comparison between UGe_2 , URhGe , and UCoGe . *J. Phys. Soc. Jpn.* **88**, 022001 (2019).

7. 7.

Ran, S. et al. Nearly ferromagnetic spin-triplet superconductivity. *Science* **365**, 684–687 (2019).

8. 8.

Aoki, D. et al. Unconventional superconductivity in heavy fermion UTe_2 . *J. Phys. Soc. Jpn.* **88**, 043702 (2019).

9. 9.

Ran, S. et al. Extreme magnetic field-boosted superconductivity. *Nat. Phys.* **15**, 1250–1254 (2019).

10. 10.

Knebel, G. et al. Field-reentrant superconductivity close to a metamagnetic transition in the heavy-fermion superconductor UTe_2 . *J. Phys. Soc. Jpn.* **88**, 063707 (2019).

11. 11.

Sundar, S. et al. Coexistence of ferromagnetic fluctuations and superconductivity in the actinide superconductor UTe₂. *Phys. Rev. B* **100**, 140502 (2019).

12. 12.

Jiao, L. et al. Chiral superconductivity in heavy-fermion metal UTe₂. *Nature* **579**, 523–527 (2020).

13. 13.

Nakamine, G. et al. Anisotropic response of spin susceptibility in the superconducting state of UTe₂ probed with ¹²⁵Te–NMR measurement. *Phys. Rev. B* **103**, L100503 (2021).

14. 14.

Hayes, I. M. et al. Multicomponent superconducting order parameter in UTe₂. *Science* **373**, 797–801 (2021).

15. 15.

Thomas, S. M. et al. Evidence for a pressure-induced antiferromagnetic quantum critical point in intermediate-valence UTe₂. *Sci. Adv.* **6**, eabc8709 (2020).

16. 16.

Duan, C. et al. Incommensurate spin fluctuations in the spin-triplet superconductor candidate UTe₂. *Phys. Rev. Lett.* **125**, 237003 (2020).

17. 17.

Rossat-Mignod, J. et al. Neutron scattering study of the YBa₂Cu₃O_{6+x} system. *Physica C* **185–189**, 86–92 (1991).

18. 18.

Wilson, S. D. et al. Resonance in the electron-doped high-transition-temperature superconductor $\text{Pr}_{0.88}\text{LaCe}_{0.12}\text{CuO}_{4-\delta}$. *Nature* **442**, 59–62 (2006).

19. 19.

Dai, P. Antiferromagnetic order and spin dynamics in iron-based superconductors. *Rev. Mod. Phys.* **87**, 855–896 (2015).

20. 20.

Sato, N. K. et al. Strong coupling between local moments and superconducting ‘heavy’ electrons in UPd_2Al_3 . *Nature* **410**, 340–343 (2001).

21. 21.

Bernhoeft, N. Superconductor order parameter symmetry in UPd_2Al_3 . *Eur. Phys. J. B* **13**, 685–694 (2000).

22. 22.

Stock, C., Broholm, C., Hudis, J., Kang, H. J. & Petrovic, C. Spin resonance in the *d*-wave superconductor CeCoIn_5 . *Phys. Rev. Lett.* **100**, 087001 (2008).

23. 23.

Stockert, O. et al. Magnetically driven superconductivity in CeCu_2Si_2 . *Nat. Phys.* **7**, 119–124 (2011).

24. 24.

Kuwabara, T. & Ogata, M. Spin-triplet superconductivity due to antiferromagnetic spin-fluctuation in Sr_2RuO_4 . *Phys. Rev. Lett.* **85**, 4586–4589 (2000).

25. 25.

Eschrig, M. The effect of collective spin-1 excitations on electronic spectra in high- T_c superconductors. *Adv. Phys.* **55**, 47–183 (2006).

26. 26.

Yu, G., Li, Y., Motoyama, E. M. & Greven, M. A universal relationship between magnetic resonance and superconducting gap in unconventional superconductors. *Nat. Phys.* **5**, 873–875 (2009).

27. 27.

Huxley, A. D., Raymond, S. & Ressouche, E. Magnetic excitations in the ferromagnetic superconductor UGe₂. *Phys. Rev. Lett.* **91**, 207201 (2003).

28. 28.

Stock, C. et al. Anisotropic critical magnetic fluctuations in the ferromagnetic superconductor UCoGe. *Phys. Rev. Lett.* **107**, 187202 (2011).

29. 29.

Kunkemöller, S. et al. Absence of a large superconductivity-induced gap in magnetic fluctuations of Sr₂RuO₄. *Phys. Rev. Lett.* **118**, 147002 (2017).

30. 30.

Steffens, P. et al. Spin fluctuations in Sr₂RuO₄ from polarized neutron scattering: implications for superconductivity. *Phys. Rev. Lett.* **122**, 047004 (2019).

31. 31.

Pustogow, A. et al. Constraints on the superconducting order parameter in Sr_2RuO_4 from oxygen-17 nuclear magnetic resonance. *Nature* **574**, 72–75 (2019).

32. 32.

Joynt, R. & Taillefer, L. The superconducting phases of UPt_3 . *Rev. Mod. Phys.* **74**, 235–294 (2002).

33. 33.

Aeppli, G. et al. Magnetic order and fluctuations in superconducting UPt_3 . *Phys. Rev. Lett.* **60**, 615–618 (1988).

34. 34.

Gannon, W. J. et al. Spin susceptibility of the topological superconductor UPt_3 from polarized neutron diffraction. *Phys. Rev. B* **96**, 041111(R) (2017).

35. 35.

Song, Y. et al. Nature of the spin resonance mode in CeCoIn_5 . *Commun. Phys.* **3**, 98 (2020).

36. 36.

Zwicknagl, G. & Fulde, P. The dual nature of $5f$ electrons and the origin of heavy fermions in U compounds. *J. Phys. Condens. Matter* **15**, S1911–S1916 (2003).

37. 37.

Fujimori, S. et al. Core-level photoelectron spectroscopy study of UTe_2 . *J. Phys. Soc. Jpn.* **90**, 015002 (2021).

38. 38.

Ishizuka, J. & Yanase, Y. Periodic Anderson model for magnetism and superconductivity in UTe₂. *Phys. Rev. B* **103**, 094504 (2021).

39. 39.

Wang, M. et al. Doping dependence of spin excitations and its correlations with high-temperature superconductivity in iron pnictides. *Nat. Commun.* **4**, 2874 (2013).

40. 40.

Miao, L. et al. Low energy band structure and symmetries of UTe₂ from angle-resolved photoemission spectroscopy. *Phys. Rev. Lett.* **124**, 076401 (2020).

41. 41.

Kim, J. S., Tam, G. N. & Stewart, G. R. Universal scaling law for the condensation energy across a broad range of superconductor classes. *Phys. Rev. B* **92**, 224509 (2015).

42. 42.

Ehlers, G., Podlesnyak, A. A., Niedziela, J. L., Iverson, E. B. & Sokol, P. E. The new Cold Neutron Chopper Spectrometer at the Spallation Neutron Source: design and performance, *Rev. Sci. Instrum.* **82**, 085108 (2011).

43. 43.

Ewings, R. A. et al. Horace: software for the analysis of data from single crystal spectroscopy experiments at time-of-flight neutron instruments. *Nucl. Instrum. Meth. Phys. Res. A* **834**, 132–142 (2016).

44. 44.

Nica, E. N. & Si, Q. Multiorbital singlet pairing and $d + d$ superconductivity. *npj Quantum Mater.* **6**, 3 (2021).

45. 45.

Nica, E. M., Yu, R. & Si, Q. Orbital-selective pairing and superconductivity in iron selenides. *npj Quantum Mater.* **2**, 24 (2017).

46. 46.

Pang, G. M. et al. Fully gapped *d*-wave superconductivity in CeCu₂Si₂. *Proc. Natl Acad. Sci. USA* **115**, 5343–5347 (2018).

47. 47.

Amorese, A. et al. Possible multiorbital ground state in CeCu₂Si₂. *Phys. Rev. B* **102**, 245146 (2020).

48. 48.

Shick, A. B. & Pickett, W. E. Spin-orbit coupling induced degeneracy in the anisotropic unconventional superconductor UTe₂. *Phys. Rev. B* **100**, 134502 (2019).

49. 49.

Shick, A. B., Fujimori, S. & Pickett, W. E. UTe₂: a nearly insulating half-filled $\frac{5}{2}$ heavy-fermion metal. *Phys. Rev. B* **103**, 125136 (2021).

50. 50.

Koster, G. F. et al. *The Properties of the Thirty-Two Point Groups* (MIT Press, 1963).

51. 51.

Pixley, J. H., Deng, L., Ingersent, K. & Si, Q. Pairing correlations near a Kondo-destruction quantum critical point. *Phys. Rev. B* **91**, 201109(R) (2015).

52. 52.

Nguyen, D. H. et al. Superconductivity in an extreme strange metal.
Nat. Commun. **12**, 4341 (2021).

Acknowledgements

P.D. thanks D. Natelson, W. P. Halperin, N. Butch and J. Paglione for discussions. E.M.N. and Q.S. acknowledge discussions with H. Hu, S. Paschen and J.-X. Zhu. The INS work at Rice is supported by the US DOE, BES under grant no. DE-SC0012311 (P.D.). Part of the material characterization efforts at Rice is supported by the Robert A. Welch Foundation grant nos C-1839 (P.D.). Work performed by R.E.B. at the National High Magnetic Field Laboratory was supported by National Science Foundation Cooperative Agreement No. DMR-1644779 and the State of Florida. Synthesis of crystalline materials and measurements by R.E.B. were supported by the Center for Actinide Science and Technology (CAST), an Energy Frontier Research Center (EFRC) funded by the US DOE, BES, under grant no. DE-SC0016568. Research at UC San Diego was supported by the US DOE, BES under grant no. DEFG02-04-ER46105 (single-crystal growth) and US NSF under grant no. DMR-1810310 (characterization of physical properties). The theory work at Rice has primarily been supported by the US DOE, BES under award no. DE-SC0018197, with travel support provided by the Robert A. Welch Foundation grant no. C-1411. Q.S. acknowledges the hospitality of the Aspen Center for Physics, which is supported by NSF grant no. PHY-1607611. E.M.N. was supported by an ASU startup grant. A portion of this research used resources at the Spallation Neutron Source, a DOE Office of Science User Facility operated by ORNL.

Author information

Affiliations

1. Department of Physics and Astronomy, Rice Center for Quantum Materials, Rice University, Houston, TX, USA

Chunruo Duan, Qimiao Si & Pengcheng Dai

2. National High Magnetic Field Laboratory, Florida State University,
Tallahassee, FL, USA

R. E. Baumbach

3. Department of Physics, Florida State University, Tallahassee, FL, USA

R. E. Baumbach

4. Neutron Scattering Sciences Division, Oak Ridge National Laboratory,
Oak Ridge, TN, USA

Andrey Podlesnyak

5. Department of Physics, University of California, San Diego, San
Diego, CA, USA

Yuhang Deng, Camilla Moir, Alexander J. Breindel & M. Brian Maple

6. Department of Physics, Arizona State University, Tempe, AZ, USA

E. M. Nica

Contributions

P.D. and M.B.M. conceived the project. R.E.B. grew the single crystals and made specific-heat measurements on the crystals. The single crystals of UTe₂ were aligned using Laue X-ray diffraction by C.D., Y.D., C.M. and A.J.B. and characterized by means of powder X-ray diffraction by C.M., A.J.B. and Y.D. at UCSD. The INS experiments were carried out by A.P. in remote discussion with C.D. and P.D. The data analysis was carried out by C.D. and P.D. E.M.N. and Q.S. contributed to the theoretical idea that AF spin fluctuations may facilitate spin-triplet superconductivity. The paper was written by P.D., C.D., R.E.B., E.M.N and Q.S., and all coauthors made comments on the paper.

Corresponding author

Correspondence to [Pengcheng Dai](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

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 Pictures of the UTe₂ single crystals used in the INS experiment.](#)

a, A typical piece of UTe₂ single crystal of 10 mm by 3 mm by 3 mm in size. The direction of the longest edge is the intersection of [1, 1, 0] plane and [0, 0, 1] plane. **b, c**, 27 pieces of UTe₂ single crystals co-aligned on two oxygen-free Cu sample plates. The total mass is 0.9 grams.

[Extended Data Fig. 2 Summary of temperature-dependent heat capacity C/T for single-crystal specimens of UTe₂.](#)

a, Comparison of C/T versus T for two representative crystals of UTe₂. One crystal shows a single superconducting phase transition whereas the other

shows two features. Several other crystals were measured, which all show similar behavior. **b**, The electronic component of the heat capacity C_e/T , which was obtained by subtracting the low temperature phonon heat capacity βT^2 , which was obtained by fitting the data for $T > T_c$ using the expression $C/T = \gamma + \beta T^2$. The normal state electronic coefficient of the heat capacity γ is indicated by the horizontal dotted blue line. An equal entropy construction is also indicated by dotted blue lines to determine T_c and the ideal size of the heat-capacity jump $\Delta C/T_c$.

Extended Data Fig. 3 X-ray Laue pattern of the [0,0,1] plane of UTe₂.

Pattern is shown for one of the samples used in the experiment.

Extended Data Fig. 4 Unsymmetrized raw data in the [H,K,0] plane with $E_i = 3.32$ meV.

a–l, Constant-energy cuts of the unsymmetrized $S(\mathbf{Q}, E)$ with $E_i = 3.32$ meV at **(a)** 0.0 ± 0.1 meV and BT, **(b)** 0.0 ± 0.1 meV and 2 K, **(c)** 0.4 ± 0.1 meV and BT, **(d)** 0.4 ± 0.1 meV and 2 K, **(e)** 0.7 ± 0.1 meV and BT, **(f)** 0.7 ± 0.1 meV and 2 K, **(g)** 1.0 ± 0.1 meV and BT, **(h)** 1.0 ± 0.1 meV and 2 K, **(i)** 1.5 ± 0.1 meV and BT, **(j)** 1.5 ± 0.1 meV and 2 K, **(k)** 2.0 ± 0.1 meV and BT, **(l)** 2.0 ± 0.1 meV and 2 K. The bin size is 0.035 r.l.u. along both H and K . The integration range is ± 0.2 r.l.u. in L , and ± 0.1 meV in E . The unit of the colour bars in Extended Data Figs. 4, 5, 6 is the same as that of Fig. 2b.

Extended Data Fig. 5 Unsymmetrized raw data in the [H,K,0] plane with $E_i = 12$ meV.

a–d, Constant energy cuts of the unsymmetrized $S(\mathbf{Q}, E)$ with $E_i = 12$ meV and BT at **(a)** 0.0 ± 0.5 meV, **(b)** 3.25 ± 0.25 meV, **(c)** 5.25 ± 0.25 meV, **(d)** 7.25 ± 0.25 meV. The bin size is 0.04 r.l.u. along both H and K . The integration range is ± 0.2 r.l.u. in L , and ± 0.25 meV in E . The rings of

scattering in **a** are from the nuclear (1, 1, 1) and (2, 0, 0) Bragg peaks of the Cu sample holder.

Extended Data Fig. 6 Unsymmetrized raw data, E - \mathbf{Q} plots and one-dimensional energy cuts with $E_i = 2.5$ meV.

a–d, Constant energy cuts of the unsymmetrized $S(\mathbf{Q}, E)$ with $E_i = 2.5$ meV at (**a**) 0.25 to 0.3 meV and BT, (**b**) 0.25 to 0.3 meV and 3.5 K, (**c**) 1.05 to 1.1 meV and BT, (**d**) 1.05 to 1.1 meV and 3.5 K. The bin size is 0.02 r.l.u. along H and 0.03 r.l.u. along K . The integration range is ± 0.3 r.l.u. in L . **e, f**, E - \mathbf{Q} plots of the scattering function $S(\mathbf{Q}, E)$ with $E_i = 2.5$ meV at BT (**e**) and 3.5 K (**f**), respectively. The integration range is ± 0.08 r.l.u. in H and ± 0.3 r.l.u. in L , the bin size along K is 0.03 r.l.u., and the E step is 0.03 meV. **g**, One-dimensional cuts of the scattering function $S(\mathbf{Q})$ with high temperature data ($S^{HT}(\mathbf{Q})$) subtracted. The cuts are taken at Y1 along E taken at BT (blue), 0.4 K (red), 0.8 K (yellow), 1.2 K (purple), and 1.5 K (green) with $E_i = 2.5$ meV. The high-temperature data are taken at 3.5 K. At low energy the excitation at Y1 is not fully covered with this E_i , which causes the gap feature between 0.2 to 0.7 meV to be hard to observe in the subtracted one-dimensional data. Different temperature data in **g** are artificially shifted, with the dashed black line representing the base line for each temperature. The integration ranges in **g** are: ± 0.08 r.l.u. in H , ± 0.15 r.l.u. in K , and ± 0.3 r.l.u. in L . The bin size in E is 0.04 meV.

Extended Data Fig. 7 Temperature dependence of the excitations at different \mathbf{Q} positions.

a, b, One-dimensional cuts of $S(E)$ with $E_i = 3.32$ meV at Bragg peak (1, -1, 0) along E at BT and 2 K, respectively. Incoherent background scattering integrated at \mathbf{Q}_{bkg} is plotted in green triangles. There are no FM spin fluctuation signals observed above the background. The broad peak around $E = 0.7$ meV is powder ring of scattering not associated with UTe₂ (see Extended Data Fig. 4e,g). (**c**) One-dimensional cuts of $S(E)$ with $E_i = 3.32$ meV at Y1 along E at 1.5, 1.8, and 2 K. There is no significant change in the quasielastic energy range for temperature close to and above T_c . **d, e**,

One-dimensional cuts of $S(E)$ with $E_i = 3.32$ meV (**d**) and 2.5 meV (**e**), respectively. The subtle increase of $S(E)$ above T_c near 1.4 meV with $E_i = 3.32$ meV is just above one standard deviation, and is not observed with $E_i = 2.5$ meV. The integration ranges of the one-dimensional data in **d**, **e** are: ± 0.1 r.l.u. in H , ± 0.15 r.l.u. in K , and ± 0.3 r.l.u. in L . The bin size in E is 0.04 meV.

Supplementary information

[Peer Review File](#)

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Duan, C., Baumbach, R.E., Podlesnyak, A. *et al.* Resonance from antiferromagnetic spin fluctuations for superconductivity in UTe₂. *Nature* **600**, 636–640 (2021). <https://doi.org/10.1038/s41586-021-04151-5>

- Received: 07 June 2021
- Accepted: 15 October 2021
- Published: 22 December 2021
- Issue Date: 23 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04151-5>

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/s41586-021-04151-5>

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

- Article
- [Published: 22 December 2021](#)

Quantum anomalous Hall effect from intertwined moiré bands

- [Tingxin Li^{1,2} na1](#),
- [Shengwei Jiang](#) [ORCID: orcid.org/0000-0001-5213-0139^{1,2} na1](#),
- [Bowen Shen¹ na1](#),
- [Yang Zhang](#) [ORCID: orcid.org/0000-0003-4630-5056³ na1](#),
- [Lizhong Li¹](#),
- [Zui Tao¹](#),
- [Trithip Devakul³](#),
- [Kenji Watanabe](#) [ORCID: orcid.org/0000-0003-3701-8119⁴](#),
- [Takashi Taniguchi](#) [ORCID: orcid.org/0000-0002-1467-3105⁴](#),
- [Liang Fu³](#),
- [Jie Shan](#) [ORCID: orcid.org/0000-0003-1270-9386^{1,5}](#) &
- [Kin Fai Mak](#) [ORCID: orcid.org/0000-0002-5768-199X^{1,5}](#)

[Nature](#) volume 600, pages 641–646 (2021)

- 3247 Accesses
- 9 Altmetric
- [Metrics details](#)

Subjects

- [Phase transitions and critical phenomena](#)
- [Topological insulators](#)
- [Two-dimensional materials](#)

Abstract

Electron correlation and topology are two central threads of modern condensed matter physics. Semiconductor moiré materials provide a highly tuneable platform for studies of electron correlation^{1,2,3,4,5,6,7,8,9,10,11,12}. Correlation-driven phenomena, including the Mott insulator^{2,3,4,5}, generalized Wigner crystals^{2,6,9}, stripe phases¹⁰ and continuous Mott transition^{11,12}, have been demonstrated. However, non-trivial band topology has remained unclear. Here we report the observation of a quantum anomalous Hall effect in AB-stacked MoTe₂/WSe₂ moiré heterobilayers.

Unlike in the AA-stacked heterobilayers¹¹, an out-of-plane electric field not only controls the bandwidth but also the band topology by intertwining moiré bands centred at different layers. At half band filling, corresponding to one particle per moiré unit cell, we observe quantized Hall resistance, h/e^2 (with h and e denoting the Planck's constant and electron charge, respectively), and vanishing longitudinal resistance at zero magnetic field. The electric-field-induced topological phase transition from a Mott insulator to a quantum anomalous Hall insulator precedes an insulator-to-metal transition. Contrary to most known topological phase transitions¹³, it is not accompanied by a bulk charge gap closure. Our study paves the way for discovery of emergent phenomena arising from the combined influence of strong correlation and topology in semiconductor moiré materials.

This is a preview of subscription content

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Figure 1: AB-stacked MoTe₂/WSe₂ heterobilayer.

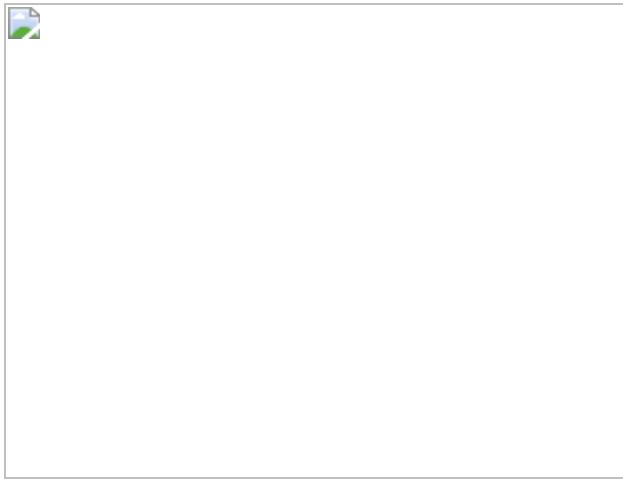


Figure 2: Quantized anomalous Hall effect at $\nu = 1$.

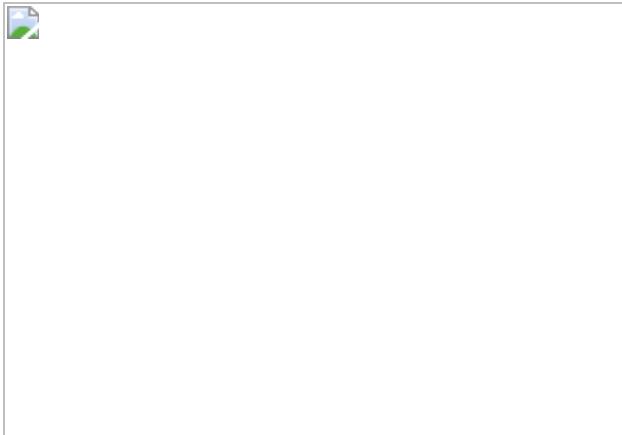


Figure 3: Mott–QAH insulator transition at $\nu = 1$.

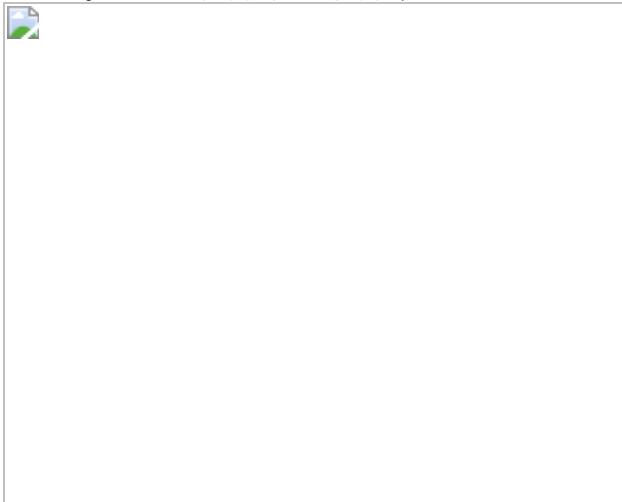
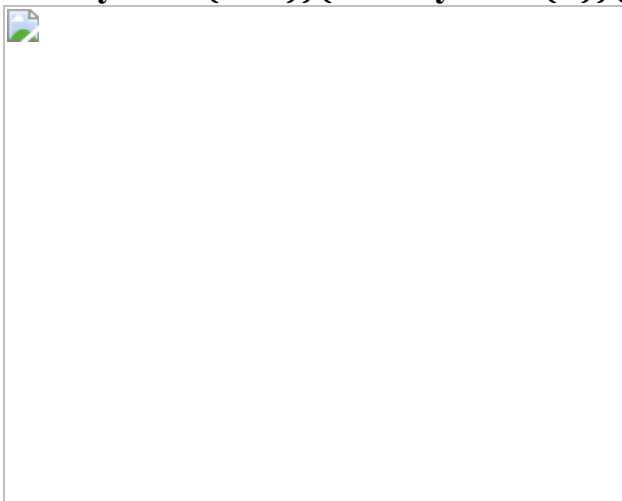


Figure 4: Evidence of a quantum valley-spin Hall insulator at $\nu = 2$.



Data availability

[Source data](#) are provided with this paper. All other data are available from the corresponding authors.

References

1. 1.

Wu, F., Lovorn, T., Tutuc, E. & MacDonald, A. H. Hubbard model physics in transition metal dichalcogenide moiré bands. *Phys. Rev. Lett.* **121**, 026402 (2018).

2. 2.

Regan, E. C. et al. Mott and generalized Wigner crystal states in WSe₂/WS₂ moiré superlattices. *Nature* **579**, 359–363 (2020).

3. 3.

Tang, Y. et al. Simulation of Hubbard model physics in WSe₂/WS₂ moiré superlattices. *Nature* **579**, 353–358 (2020).

4. 4.

Wang, L. et al. Correlated electronic phases in twisted bilayer transition metal dichalcogenides. *Nat. Mater.* **19**, 861–866 (2020).

5. 5.

Shimazaki, Y. et al. Strongly correlated electrons and hybrid excitons in a moiré heterostructure. *Nature* **580**, 472–477 (2020).

6. 6.

Xu, Y. et al. Correlated insulating states at fractional fillings of moiré superlattices. *Nature* **587**, 214–218 (2020).

7. 7.

Zhang, Y., Yuan, N. F. Q. & Fu, L. Moiré quantum chemistry: charge transfer in transition metal dichalcogenide superlattices. *Phys. Rev. B* **102**, 201115 (2020).

8. 8.

Andrei, E. Y. et al. The marvels of moiré materials. *Nat. Rev. Mater.* **6**, 201–206 (2021).

9. 9.

Huang, X. et al. Correlated insulating states at fractional fillings of the WS₂/WSe₂ moiré lattice. *Nat. Phys.* **17**, 715–719 (2021).

10. 10.

Jin, C. et al. Stripe phases in WSe₂/WS₂ moiré superlattices. *Nat. Mater.* **20**, 940–944 (2021).

11. 11.

Li, T. et al. Continuous Mott transition in semiconductor moiré superlattices. *Nature* **597**, 350–354 (2021).

12. 12.

Ghiotto, A. et al. Quantum criticality in twisted transition metal dichalcogenides. *Nature* **597**, 345–349 (2021).

13. 13.

Hasan, M. Z. & Kane, C. L. Colloquium: topological insulators. *Rev. Mod. Phys.* **82**, 3045–3067 (2010).

14. 14.

Andrei, E. Y. & MacDonald, A. H. Graphene bilayers with a twist. *Nat. Mater.* **19**, 1265–1275 (2020).

15. 15.

Balents, L., Dean, C. R., Efetov, D. K. & Young, A. F. Superconductivity and strong correlations in moiré flat bands. *Nat. Phys.* **16**, 725–733 (2020).

16. 16.

Kennes, D. M. et al. Moiré heterostructures as a condensed-matter quantum simulator. *Nat. Phys.* **17**, 155–163 (2021).

17. 17.

Serlin, M. et al. Intrinsic quantized anomalous Hall effect in a moiré heterostructure. *Science* **367**, 900–903 (2020).

18. 18.

Sharpe, A. L. et al. Emergent ferromagnetism near three-quarters filling in twisted bilayer graphene. *Science* **365**, 605–608 (2019).

19. 19.

Chen, G. et al. Tunable correlated Chern insulator and ferromagnetism in a moiré superlattice. *Nature* **579**, 56–61 (2020).

20. 20.

Cao, Y. et al. Unconventional superconductivity in magic-angle graphene superlattices. *Nature* **556**, 43–50 (2018).

21. 21.

Wu, F., Lovorn, T., Tutuc, E., Martin, I. & MacDonald, A. H. Topological insulators in twisted transition metal dichalcogenide homobilayers. *Phys. Rev. Lett.* **122**, 086402 (2019).

22. 22.

Devakul, T., Crépel, V., Zhang, Y. and Fu, L. Magic in twisted transition metal dichalcogenide bilayers. *Nat. Commun.* <https://doi.org/10.1038/s41467-021-27042-9> (2021).

23. 23.

Hohenadler, M. & Assaad, F. F. Correlation effects in two-dimensional topological insulators. *J. Phys. Condens. Matter* **25**, 143201 (2013).

24. 24.

Witczak-Krempa, W., Chen, G., Kim, Y. B. & Balents, L. Correlated quantum phenomena in the strong spin-orbit regime. *Annu. Rev. Condens. Matter Phys.* **5**, 57–82 (2014).

25. 25.

Pesin, D. & Balents, L. Mott physics and band topology in materials with strong spin–orbit interaction. *Nat. Phys.* **6**, 376–381 (2010).

26. 26.

Raghu, S., Qi, X.-L., Honerkamp, C. & Zhang, S.-C. Topological Mott insulators. *Phys. Rev. Lett.* **100**, 156401 (2008).

27. 27.

Zhang, Y., Devakul, T. & Fu, L. Spin-textured Chern bands in AB-stacked transition metal dichalcogenide bilayers. *Proc. Natl Acad. Sci. USA* **118**, e2112673118 (2021).

28. 28.

Chang, C.-Z. et al. Experimental observation of the quantum anomalous Hall effect in a magnetic topological insulator. *Science* **340**, 167–170 (2013).

29. 29.

Kou, X. et al. Scale-invariant quantum anomalous Hall effect in magnetic topological insulators beyond the two-dimensional limit. *Phys. Rev. Lett.* **113**, 137201 (2014).

30. 30.

Mogi, M. et al. Magnetic modulation doping in topological insulators toward higher-temperature quantum anomalous Hall effect. *Appl. Phys. Lett.* **107**, 182401 (2015).

31. 31.

Liu, C.-X., Zhang, S.-C. & Qi, X.-L. The quantum anomalous Hall effect: theory and experiment. *Annu. Rev. Condens. Matter Phys.* **7**, 301–321 (2016).

32. 32.

MacDonald, A. H. Introduction to the physics of the quantum Hall regime. Preprint at arXiv:cond-mat/9410047 (1994).

33. 33.

Young, A. F. et al. Tunable symmetry breaking and helical edge transport in a graphene quantum spin Hall state. *Nature* **505**, 528–532 (2014).

34. 34.

Li, T. et al. Charge-order-enhanced capacitance in semiconductor moiré superlattices. *Nat. Nanotechnol.* **16**, 1068–1072 (2021).

35. 35.

Y. M. Xie, C. P. Zhang, J. X. Hu, K. F. Mak & Law, K. T. Theory of valley polarized quantum anomalous Hall state in moiré MoTe₂/WSe₂ heterobilayers. Preprint at arXiv:2106.13991 (2021).

36. 36.

Fei, Z. et al. Edge conduction in monolayer WTe₂. *Nat. Phys.* **13**, 677–682 (2017).

37. 37.

Amaricci, A., Budich, J. C., Capone, M., Trauzettel, B. & Sangiovanni, G. First-order character and observable signatures of topological quantum phase transitions. *Phys. Rev. Lett.* **114**, 185701 (2015).

38. 38.

Ezawa, M., Tanaka, Y. & Nagaosa, N. Topological phase transition without gap closing. *Sci. Rep.* **3**, 2790 (2013).

39. 39.

Senthil, T. Theory of a continuous Mott transition in two dimensions. *Phys. Rev. B* **78**, 045109 (2008).

40. 40.

Wang, L. et al. One-dimensional electrical contact to a two-dimensional material. *Science* **342**, 614–617 (2013).

41. 41.

Zibrov, A. A. et al. Tunable interacting composite fermion phases in a half-filled bilayer-graphene Landau level. *Nature* **549**, 360–364 (2017).

42. 42.

Ashoori, R. C. et al. Single-electron capacitance spectroscopy of discrete quantum levels. *Phys. Rev. Lett.* **68**, 3088–3091 (1992).

43. 43.

Xie, Y.-M., Zhang, C.-P., Hu, J.-X., Mak, K. F. & Law, K. Theory of valley polarized quantum anomalous Hall state in moiré MoTe₂/WSe₂ heterobilayers. Preprint at arXiv:2106.13991 (2021).

44. 44.

Perdew, J. P., Burke, K. & Ernzerhof, M. generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).

45. 45.

Peng, H., Yang, Z.-H., Perdew, J. P. & Sun, J. Versatile van der Waals density functional based on a meta-generalized gradient approximation. *Phys. Rev.* **6**, 041005 (2016).

46. 46.

Kresse, G. & Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mater. Sci.* **6**, 15–50 (1996).

47. 47.

Liu, G.-B., Xiao, D., Yao, Y., Xu, X. & Yao, W. Electronic structures and theoretical modelling of two-dimensional group-VIB transition metal dichalcogenides. *Chem. Soc. Rev.* **44**, 2643–2663 (2015).

Acknowledgements

This research was primarily supported by the US Department of Energy, Office of Science, Basic Energy Sciences, under award no. DE-SC0019481 (electrical measurements). The experimental study was partially supported by the Cornell Center for Materials Research with funding from the NSF MRSEC program (DMR-1719875) for device fabrication, the Air Force Office of Scientific Research under award number FA9550-19-1-0390 (capacitance measurements) and FA9550-20-1-0219 (optical characterizations), and the US Army Research Office under grant no.

W911NF-17-1-0605 (analysis). The theoretical work at Massachusetts Institute of Technology was supported by the Simons Foundation through a Simons Investigator Award (theoretical analysis) and by the US Department of Energy, Office of Science, Basic Energy Sciences, Division of Materials Sciences and Engineering Awards no. DE-SC0020149 (band structure calculation). Growth of the hBN crystals was supported by the Elemental Strategy Initiative of MEXT, Japan and CREST (no. JPMJCR15F3), JST. This work made use of the Cornell NanoScale Facility, an NNCI member supported by NSF Grant no. NNCI-1542081. L.F. is partially supported by the David and Lucile Packard Foundation.

Author information

Author notes

1. These authors contributed equally: Tingxin Li, Shengwei Jiang, Bowen Shen, Yang Zhang

Affiliations

1. School of Applied and Engineering Physics and Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, NY, USA

Tingxin Li, Shengwei Jiang, Bowen Shen, Lizhong Li, Zui Tao, Jie Shan & Kin Fai Mak

2. Key Laboratory of Artificial Structures and Quantum Control (Ministry of Education), School of Physics and Astronomy, Shanghai Jiao Tong University, Shanghai, China

Tingxin Li & Shengwei Jiang

3. Department of Physics, Massachusetts Institute of Technology, Cambridge, MA, USA

Yang Zhang, Trithep Devakul & Liang Fu

4. National Institute for Materials Science, Tsukuba, Japan

Kenji Watanabe & Takashi Taniguchi

5. Kavli Institute at Cornell for Nanoscale Science, Ithaca, NY, USA

Jie Shan & Kin Fai Mak

Contributions

T.L., S.J., B.S. and L.L. fabricated the devices, performed the measurements and analysed the data. Z.T. provided assistance with the optical characterization. Y.Z. performed the DFT calculations. Y.Z., T.D. and L.F. performed theoretical analysis. K.W. and T.T. grew the bulk hBN crystals. T.L., S.J., J.S. and K.F.M. designed the scientific objectives and oversaw the project. All authors discussed the results and commented on the manuscript.

Corresponding authors

Correspondence to [Jie Shan](#) or [Kin Fai Mak](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Lede Xian, You Zhou and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

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 Figure 1 Optical micrograph of device 1.

MoTe₂/WSe₂ heterobilayer is outlined by the black line. Electrode 1-9 are labeled. For the results presented in the main text, electrode 5 and 6 are grounded; electrode 1 is used as a source electrode. The longitudinal voltage drop is measured between 3 and 4; and the transverse voltage drop is measured between 3 and 8. The scale bar is 10 μm .

Extended Data Figure 2 2D maps in filling factor and electric field.

2D map of $\langle R_{xx} \rangle$ (a), $\langle R_{xy} \rangle$ (b) and $\langle C/C_g \rangle$ (c) as a function of filling factor and electric field. The data are converted from the 2D maps in Fig. 1c, d and 3d. The QAH region is circled by green dashed lines.

[Source data](#)

Extended Data Figure 3 Other devices exhibiting the QAH effect at $\nu = 1$.

a,b, Magnetic-field dependence of $\langle R_{xy} \rangle$ and $\langle R_{xx} \rangle$ from device 2 and 3 at 1.6 K. Quantized Hall resistance at zero magnetic field is observed in device 2; $\langle R_{xx} \rangle$ is not measured. The resistance quantization in device 3 is nearly perfect at a small magnetic field, accompanied by a significant drop in $\langle R_{xx} \rangle$. **c,d**, The same as **a,b** for device 1 measured from different pairs of probes (than the result shown in the main text) at 300 mK. The labeling of the electrodes is shown in Extended Data Fig. 1. Nearly quantized resistance with similar magnetic-field dependence is observed. The coercive field varies slightly from device to device.

[Source data](#)

Extended Data Figure 4 Temperature and electric-field dependences of $\langle\langle \boldsymbol{R} \rangle\rangle_{\langle\langle \boldsymbol{xx} \rangle\rangle}$ and $\langle\langle \boldsymbol{R} \rangle\rangle_{\langle\langle \boldsymbol{xy} \rangle\rangle}$ at $\langle\langle \boldsymbol{\nu} \rangle\rangle = \langle\langle \boldsymbol{\Xi} \rangle\rangle$.

a, Hall resistance at $\langle\langle \boldsymbol{\nu} = 1 \rangle\rangle$ as a function of electric field under both forward and backward scans at 300 mK (device 1). The out-of-plane magnetic field is fixed at 0 or ± 0.1 T. No hysteresis is observed. Typical fluctuation of the Hall plateau from the ideal value $\langle\langle h/e^2 \rangle\rangle$ is about 5%. **b, c**, Electric-field dependence of $\langle\langle R_{xy} \rangle\rangle$ (**b**) and $\langle\langle R_{xx} \rangle\rangle$ (**c**) at varying temperatures in the zero-magnetic-field limit. A Hall plateau $\langle\langle R_{xy} \rangle\rangle \approx h/e^2$ in electric field, correlated with a substantial resistance dip $\langle\langle R_{xx} \rangle\rangle \ll \langle\langle R_{xy} \rangle\rangle$, is observed at low temperatures.

[Source data](#)

Extended Data Figure 5 Magnetic-field scans of $\langle\langle \boldsymbol{R} \rangle\rangle_{\langle\langle \boldsymbol{xx} \rangle\rangle}$ and $\langle\langle \boldsymbol{R} \rangle\rangle_{\langle\langle \boldsymbol{xy} \rangle\rangle}$ at $\langle\langle \boldsymbol{\nu} \rangle\rangle = \langle\langle \boldsymbol{\Xi} \rangle\rangle$.

Magnetic-field dependence of $\langle\langle R_{xy} \rangle\rangle$ (**a**) and $\langle\langle R_{xx} \rangle\rangle$ (**b**) under varying electric fields at 1.6 K. Each curve for $\langle\langle R_{xy} \rangle\rangle$ is vertically shifted by 60 k Ω for clarity. No vertical shift is applied to $\langle\langle R_{xx} \rangle\rangle$. The green and purple curves are inside the QAH region; they display nearly quantized $\langle\langle R_{xy} \rangle\rangle$ and small $\langle\langle R_{xx} \rangle\rangle$. A remnant anomalous Hall response (clear magnetic hysteresis but not quantized) is observed near the QAH-metal boundary (yellow curve). A nearly quantized Hall response with large $\langle\langle R_{xx} \rangle\rangle$ and small $\langle\langle R_{xy} \rangle\rangle$ is observed near the Mott-QAH boundary (blue curve). It is a quantized Hall insulator. Further away from the boundary and into the Mott region, $\langle\langle R_{xx} \rangle\rangle$ diverges rapidly and $\langle\langle R_{xy} \rangle\rangle$ diminishes (red and black curves).

[Source data](#)

Extended Data Figure 6 Charge gap from thermal activation transport.

a, Arrhenius plot of $\langle R \rangle_{xx}$ at varying electric fields for the Mott insulator. The high-temperature data is well described by the thermal activation model (dashed lines), from which the activation gap is extracted. **b**, Same as **a** for the QAH insulator. Here thermal activation behavior (dashed line) is observed only at low temperature, when the QAH state develops. The extracted charge gaps are shown as empty red circles (Mott insulator) and filled circle (QAH insulator) in Fig. 3c of the main text.

[Source data](#)

Extended Data Figure 7 Identification of the QAH region in the capacitance device (device 6).

a–d, Normalized differential top-gate capacitance, (C/C_g) , versus magnetic field and filling factor at 300 mK. Results for four different electric fields are shown. Most strongly dispersive incompressible states shown arise from the Landau levels of the graphite top gate and are irrelevant in this study. The black dashed lines, originated from $(\nu=1)$, are the theoretical magnetic-field dispersion of a QAH state with Chern number -1. We use this to determine the Mott-QAH insulator boundary in Fig. 3c of the main text. The system is a Mott insulator in **a** since there is no magnetic field dispersion of the incompressible state; it is near the Mott-QAH insulator boundary in **b**; and it is a QAH insulator in **c, d**. The QAH insulator-metal boundary is determined as capacitance reaches $(C/C_g) \approx 1$ in the metallic state. **e**, Electric-field dependence of (C/C_g) at $(\nu=1)$, 300 mK and zero magnetic field. The QAH region determined from **a–d** is shaded in blue. The normalized capacitance (C/C_g) equals to 1 for a metallic state at large electric fields. It is below 1 for incompressible states, which include both the Mott and QAH insulating states. Charge gap closure (i.e. $(C/C_g)=1$) is not observed at the Mott-QAH phase boundary.

[Source data](#)

Extended Data Figure 8 Kohn-Sham wave functions.

a, Interlayer distance (top) and stacking alignment (bottom) of relaxed AB-stacked MoTe₂/WSe₂ heterobilayers. MM, MX, and XX (M = Mo or W, X = Se or Te) are the high-symmetry stacking sites. **b-d**, Real-space wave function at the Γ_m -point of the MBZ for the flat band (**b**), the first (**c**) and second (**d**) dispersive valence moiré bands. The envelope wave function for the flat band is located at the MX site. The atomic-scale orbitals consist of both the *d*-orbitals of the M atom and the *p*-orbitals of the X atom, forming a honeycomb network. The wave functions for the first and second dispersive moiré bands are located at the MM and XX sites, respectively. The atomic-scale orbitals are dominated by the *d*-orbitals of the M atom, forming a triangular network. The results show that the dispersive bands are originated from the K or K' valleys of the TMD monolayers.

Extended Data Figure 9 Moiré band structure at varying displacement fields.

Moiré valence band structure of AB-stacked MoTe₂/WSe₂ heterobilayers from DFT with displacement field 0 V/nm (**a**), 0.2 V/nm (**b**), 0.25 V/nm (**c**) and 0.3 V/nm (**d**). The first moiré band is shown in red (The two branches are from the different folding scheme of the K and –K valley states, which are degenerate in energy and shifted in momentum); the flat band is in blue; and the rest of the bands are in black. The energy gap at the K_m -point (marked by a red circle) between the two dispersive moiré bands closes and reopens between 0.25 V/nm and 0.3 V/nm. The insets in **c** and **d** show the envelope wave function at the K_m -point of the MBZ for the first (left inset) and second (right inset) dispersive bands. Before gap closure, the K_m -point wave function for the two dispersive bands is centered at the MM and XX site, respectively. They are centered at the same sites as the Γ_m -point wave function (see Extended Data Fig. 8). The envelope wave function switches moiré sites after gap reopening, showing the emergence of topological bands (see [Methods](#) for the interaction effects).

[Source data](#)

Extended Data Figure 10 Magnetoresistance and nonlocal transport at $\nu = 2$.

a, Magnetoresistance, $R_{xx}(B)/R_{xx}(B = 0)$ at 300 mK under an out-of-plane magnetic field at varying electric fields. The effect is much weaker than magnetoresistance under an in-plane magnetic field (Fig. 4c, d of the main text). **b**, Nonlocal resistance versus filling factor around $(\nu = 2)$ after gap reopening. The magnetic field is zero. The arrow for I in the inset shows the direction of the bias electric field between the source and drain electrodes. Voltage drops between electrode 3 and 4 (V_{3-4}) and between 7 and 6 (V_{7-6}) are measured. A change in sign for the nonlocal resistance is consistent with the presence of helical edge state transport for a quantum valley-spin Hall insulator.

[Source data](#)

Supplementary information

[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. 4**](#)

[**Source Data Extended Data Fig. 5**](#)

[**Source Data Extended Data Fig. 6**](#)

[**Source Data Extended Data Fig. 7**](#)

[**Source Data Extended Data Fig. 9**](#)

[**Source Data Extended Data Fig. 10**](#)

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Li, T., Jiang, S., Shen, B. *et al.* Quantum anomalous Hall effect from intertwined moiré bands. *Nature* **600**, 641–646 (2021).
<https://doi.org/10.1038/s41586-021-04171-1>

- Received: 05 July 2021
- Accepted: 20 October 2021
- Published: 22 December 2021
- Issue Date: 23 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04171-1>

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

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04171-1>

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

- Article
- [Published: 22 December 2021](#)

Spin splitting of dopant edge state in magnetic zigzag graphene nanoribbons

- [Raymond E. Blackwell](#)¹ na1,
- [Fangzhou Zhao](#)^{2,3} na1,
- [Erin Brooks](#) [ORCID: orcid.org/0000-0002-5834-0679](#)¹,
- [Junmian Zhu](#)¹,
- [Ilya Piskun](#)¹,
- [Shenkai Wang](#)¹,
- [Aidan Delgado](#)¹,
- [Yea-Lee Lee](#)²,
- [Steven G. Louie](#) [ORCID: orcid.org/0000-0003-0622-0170](#)^{2,3} &
- [Felix R. Fischer](#) [ORCID: orcid.org/0000-0003-4723-3111](#)^{1,3,4}

[Nature](#) volume **600**, pages 647–652 (2021)

- 3126 Accesses
- 102 Altmetric
- [Metrics details](#)

Subjects

- [Electronic properties and devices](#)
- [Magnetic materials](#)
- [Magnetic properties and materials](#)

- [Scanning probe microscopy](#)

Abstract

Spin-ordered electronic states in hydrogen-terminated zigzag nanographene give rise to magnetic quantum phenomena^{1,2} that have sparked renewed interest in carbon-based spintronics^{3,4}. Zigzag graphene nanoribbons (ZGNRs)—quasi one-dimensional semiconducting strips of graphene bounded by parallel zigzag edges—host intrinsic electronic edge states that are ferromagnetically ordered along the edges of the ribbon and antiferromagnetically coupled across its width^{1,2,5}. Despite recent advances in the bottom-up synthesis of GNRs featuring symmetry protected topological phases^{6,7,8} and even metallic zero mode bands⁹, the unique magnetic edge structure of ZGNRs has long been obscured from direct observation by a strong hybridization of the zigzag edge states with the surface states of the underlying support^{10,11,12,13,14,15}. Here, we present a general technique to thermodynamically stabilize and electronically decouple the highly reactive spin-polarized edge states by introducing a superlattice of substitutional N-atom dopants along the edges of a ZGNR. First-principles GW calculations and scanning tunnelling spectroscopy reveal a giant spin splitting of low-lying nitrogen lone-pair flat bands by an exchange field (\sim 850 tesla) induced by the ferromagnetically ordered edge states of ZGNRs. Our findings directly corroborate the nature of the predicted emergent magnetic order in ZGNRs and provide a robust platform for their exploration and functional integration into nanoscale sensing and logic devices^{15,16,17,18,19,20,21}.

This is a preview of subscription content

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Fig. 1: Bottom-up synthesis of N-doped N-6-ZGNRs.

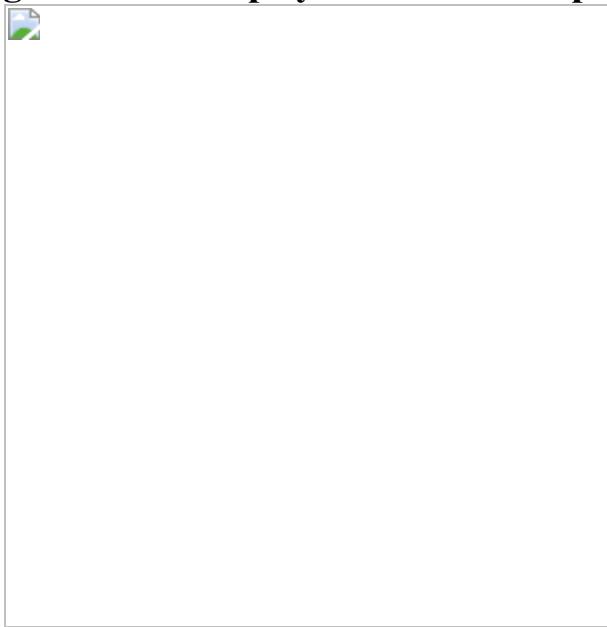


Fig. 2: Tip-induced decoupling of magnetic edge states in N-6-ZGNRs from Au surface.

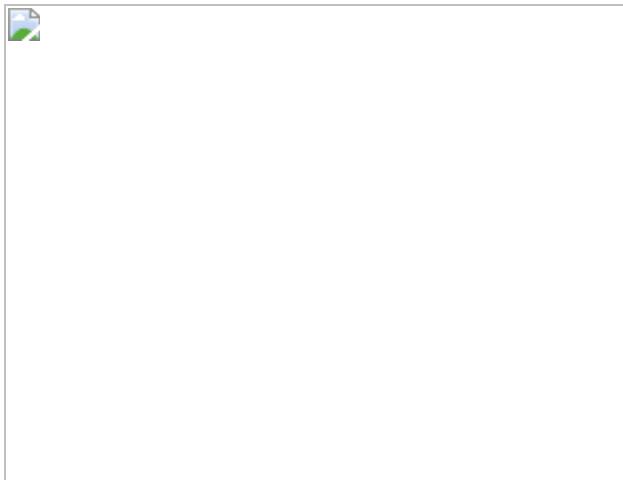


Fig. 3: Electronic structure of N-6-ZGNR.



Fig. 4: Band structure and spatial distribution of spin-ordered edge states in N-6-ZGNRs.

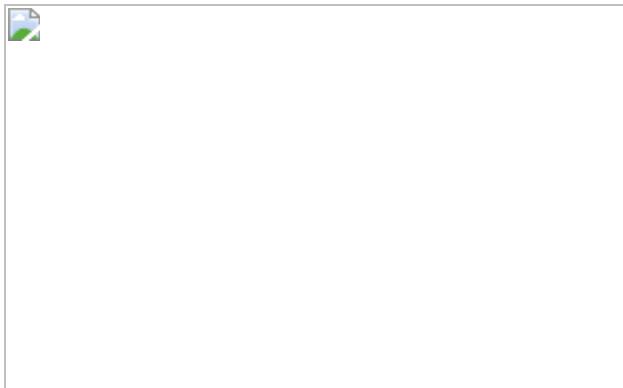
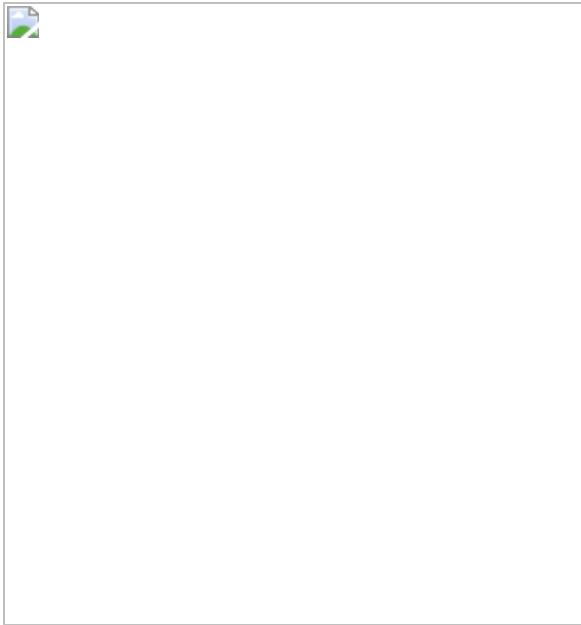


Fig. 5: Spin splitting of nitrogen flat band states (NFB) of sp^2 lone-pair orbitals in N-6-ZGNRs.



Data availability

DFT code with pseudopotentials and GW code can be downloaded from the Quantum Espresso (<https://www.quantum-espresso.org>) and the BerkeleyGW (<https://www.berkeleygw.org>) websites, respectively. We used Quantum Espresso version 6.4.1 and BerkeleyGW version 2.1 for the theoretical calculations. All data presented in the main text and the [supplementary information](#) are available from the corresponding authors upon reasonable request.

References

1. 1.

Fujita, M., Wakabayashi, K., Nakada, K. & Kusakabe, K. Peculiar localized state at zigzag graphite edge. *J. Phys. Soc. Japan* **65**, 1920–1923 (1996).

2. 2.

Son, Y.-W., Cohen, M. L. & Louie, S. G. Half-metallic graphene nanoribbons. *Nature* **444**, 347–349 (2006).

3. 3.

Han, W., Kawakami, R. K., Gmitra, M. & Fabian, J. Graphene spintronics. *Nat. Nanotechnol.* **9**, 794–807 (2014).

4. 4.

Bullard, Z., Girao, E. C., Owens, J. R., Shelton, W. A. & Meunier, V. Improved all-carbon spintronic device design. *Sci. Rep.* **5**, 7634 (2015).

5. 5.

Yang, L., Park, C. -H., Son, Y. -W., Cohen, M. L. & Louie, S. G. Quasiparticle energies and band gaps in graphene nanoribbons. *Phys. Rev. Lett.* **99**, 186801 (2007).

6. 6.

Rizzo, D. J. et al. Topological band engineering of graphene nanoribbons. *Nature* **560**, 204–208 (2018).

7. 7.

Groning, O. et al. Engineering of robust topological quantum phases in graphene nanoribbons. *Nature* **560**, 209–213 (2018).

8. 8.

Cao, T., Zhao, F. Z. & Louie, S. G. Topological phases in graphene nanoribbons: junction states, spin centers, and quantum spin chains. *Phys. Rev. Lett.* **119**, 076401 (2017).

9. 9.

Rizzo, D. J. et al. Inducing metallicity in graphene nanoribbons via zero-mode superlattices. *Science* **369**, 1597–1603, (2020).

10. 10.

Lee, G. & Cho, K. Electronic structures of zigzag graphene nanoribbons with edge hydrogenation and oxidation. *Phys. Rev. B* **79**, 165440 (2009).

11. 11.

Ritter, K. A. & Lyding, J.W. The influence of edge structure on the electronic properties of graphene quantum dots and nanoribbons. *Nat. Mater.* **8**, 235–242 (2009).

12. 12.

Tao, C. et al. Spatially resolving edge states of chiral graphene nanoribbons. *Nat. Phys.* **7**, 616–620 (2011).

13. 13.

Li, Y. Y., Chen, M. X., Weinert, M. & Li, L. Direct experimental determination of onset of electron-electron interactions in gap opening of zigzag graphene nanoribbons. *Nat. Comm.* **5**, 4311 (2014).

14. 14.

Magda, G. Z. et al. Room-temperature magnetic order on zigzag edges of narrow graphene nanoribbons. *Nature* **514**, 608–611 (2014).

15. 15.

Ruffieux, P. et al. On-surface synthesis of graphene nanoribbons with zigzag edge topology. *Nature* **531**, 489–492 (2016).

16. 16.

Beyer, D. et al. Graphene nanoribbons derived from zigzag edge-encased poly(para-2,9-dibenzo[bc,kl]coronenylenes) polymer chains. *J. Am. Chem. Soc.* **141**, 2843–2846 (2019).

17. 17.

Fu, Y. et al. On-surface synthesis of NBN-doped zigzag-edged graphene nanoribbons. *Angew. Chem. Int. Ed.* **59**, 8873–8879 (2020).

18. 18.

Li, Y., Zhou, Z., Cabrera, C. R. & Chen, Z. Preserving the edge magnetism of zigzag graphene nanoribbons by ethylene termination: insight by Clar's rule. *Sci. Rep.* **3**, 2030 (2013).

19. 19.

Shinde, P. P. et al. Stability of edge magnetism in functionalized zigzag graphene nanoribbons. *Carbon* **124**, 123–132 (2017).

20. 20.

Salemi, L., Lherbier, A. & Charlier, J. C. Spin-dependent properties in zigzag graphene nanoribbons with phenyl-edge defects. *Phys. Rev. B* **98**, 214204 (2018).

21. 21.

Lopez-Urias, F et al. Spin-dependent band-gap driven by nitrogen and oxygen functional groups in zigzag graphene nanoribbons. *Appl. Surf. Sci.* **521**, 146435 (2020).

22. 22.

Guo, G. -P. et al. Quantum computation with graphene nanoribbon. *New J. Phys.* **11**, 123005 (2009).

23. 23.

Luis, F. & Coronado, E. Spinning on the edge of graphene. *Nature* **557**, 645–647 (2018).

24. 24.

Mandal, B., Sarkar, S., Pramanik, A. & Sarkar, P. Doped defective graphene nanoribbons: a new class of materials with novel spin filtering properties. *RSC Adv.* **4**, 49946–49952 (2014).

25. 25.

Dery, H., Dalal, P., Cywinski, L. & Sham, L. J. Spin-based logic in semiconductors for reconfigurable large-scale circuits. *Nature* **447**, 573–576 (2007).

26. 26.

Dery, H. et al. Nanospintronics based on magnetologic gates. *IEEE Trans. Electron Devices* **59**, 259–262 (2012).

27. 27.

Lin, X., Yang, W., Wang, K. L. & Zhao, W. Two-dimensional spintronics for low-power electronics. *Nat. Electron.* **2**, 274–283 (2019).

28. 28.

Ahn, E. C., Wong, H. -S. P. & Pop, E. Carbon nanomaterials for non-volatile memories. *Nat. Rev. Mater.* **3**, 18009 (2018).

29. 29.

Zhang, D. -B. & Wei, S. -H. Inhomogeneous strain-induced half-metallicity in bent zigzag graphene nanoribbons. *NPJ Comput. Mater.* **3**, 32 (2017).

30. 30.

Berdonces-Layunta, A. et al. Chemical stability of (3,1)-chiral graphene nanoribbons. *ACS Nano* **15**, 5610–5617 (2021).

31. 31.

Perdew, J. P. & Zunger, A. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B* **23**, 5048–5079 (1981).

32. 32.

Hybertsen, M. S. & Louie, S. G. Electron correlation in semiconductors and insulators — band gaps and quasiparticle energies. *Phys. Rev. B* **34**, 5390–5413 (1986).

33. 33.

Lieb, E. H. Two theorems on the Hubbard model. *Phys. Rev. Lett.* **62**, 1201–1204 (1989).

34. 34.

Wang, S. K., Zhu, J. M., Blackwell, R. & Fischer, F. R. Automated tip conditioning for scanning tunnelling spectroscopy. *J. Phys. Chem. A* **125**, 1384–1390 (2021).

35. 35.

Giannozzi, P et al. QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys. Condens. Mat.* **21**, 395502 (2009).

36. 36.

Deslippe, J et al. BerkeleyGW: a massively parallel computer package for the calculation of the quasiparticle and optical properties of materials and nanostructures. *Comput. Phys. Commun.* **183**, 1269–1289 (2012).

Acknowledgements

This work was primarily funded by the US Department of Energy (DOE), Office of Science, Basic Energy Sciences (BES), Materials Sciences and

Engineering Division under contract no. DE-AC02-05-CH11231 (Nanomachine program KC1203) (molecular design, surface growth, calculations and analyses of surface–GNR interactions). Research was also supported by the Office of Naval Research under award no. N00014-19-1-2503 (STM characterization), the National Science Foundation under grant nos. DMR-1839098 (image analysis) and DMR-1926004 (GW calculations), the Center for Energy Efficient Electronics Science ECCS-0939514 (magnetic modelling), and the Office of Naval Research MURI under award no. N00014-16-1-2921 (molecular synthesis, ab initio DFT calculations). This research used resources of the National Energy Research Scientific Computing Center (NERSC), a US Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05-CH11231. Computational resources were also provided by the National Science Foundation through XSEDE resources at the NICS. R.E.B. acknowledges support through a National Science Foundation Graduate Research Fellowship under grant DGE-11064000. We thank the College of Chemistry for use of resources at their NMR facility, and we thank their staff for assistance. Instruments in CoC-NMR are supported in part by NIH S10OD024998.

Author information

Author notes

1. These authors contributed equally: Raymond E. Blackwell, Fangzhou Zhao

Affiliations

1. Department of Chemistry, University of California, Berkeley, CA, USA

Raymond E. Blackwell, Erin Brooks, Junmian Zhu, Ilya Piskun, Shenkai Wang, Aidan Delgado & Felix R. Fischer

2. Department of Physics, University of California, Berkeley, CA, USA

Fangzhou Zhao, Yea-Lee Lee & Steven G. Louie

3. Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Fangzhou Zhao, Steven G. Louie & Felix R. Fischer

4. Kavli Energy NanoScience Institute at the University of California Berkeley and the Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Felix R. Fischer

Contributions

R.E.B., F.Z., S.G.L and F.R.F. initiated and conceived the research. E.B., I.P. and F.R.F designed, synthesized, and characterized the molecular precursors. R.E.B., S.W., J.Z., A.D. and F.R.F. performed on-surface synthesis and STM characterization and analysis. F.Z., Y.-L.L. and S.G.L. performed DFT and GW calculations as well as theoretical analyses, and assisted with data interpretation. R.E.B., F.Z., S.G.L. and F.R.F. wrote the manuscript. All authors contributed to the scientific discussion.

Corresponding authors

Correspondence to [Steven G. Louie](#) or [Felix R. Fischer](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Rémy Pawlak, Levente Tapaszto and Daniel Sanchez-Portal for their contribution to the peer review of this work.

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 Height profiles of N-6-ZGNRs on Au\(111\).](#)

a, STM topographic image of N-6-ZGNR on Au(111) ($V_s = 50$ mV, $I_t = 20$ pA; CO-functionalized tip). **b**, Height profile recorded along the arrows marked in **a**.

[Extended Data Fig. 2 dI/dV point spectroscopy of N-6-ZGNRs on Au\(111\).](#)

a, dI/dV point spectra collected on bare Au(111) ($V_{ac} = 11$ mV, $f = 455$ Hz). **b**, dI/dV point spectra collected on as grown segments of N-6-ZGNRs. **c**, dI/dV point spectra collected on decoupled segments of N-6-ZGNRs following the SPM tip-induced decoupling protocol.

[Extended Data Fig. 3 Tip induced decoupling of N-6-ZGNRs.](#)

I_t/V_s plot showing in red the positive voltage sweep ($V_s = +1.50$ V to $V_s = +2.50$ V) used during the decoupling procedure. The respective return sweep ($V_s = +2.50$ V to $V_s = +1.50$ V) is depicted in grey and shows the irreversible shift in the tunnelling current I_t .

[Extended Data Fig. 4 Tip-induced decoupling of magnetic edge states in N-6-ZGNRs.](#)

a, Topographic image of a fully cyclized N-6-ZGNR segment recorded with CO-functionalized STM tip. **b**, Constant-current dI/dV map recorded at a voltage bias of $V_s = +0.5$ V of N-6-ZGNR segment following tip-induced decoupling using a positive voltage sweep from $V_s = 0.0$ V to $V_s = +2.5$ V at the position marked by a red cross in **a** ($V_{ac} = 11$ mV, $I_t = 200$ pA, $f = 455$ Hz, CO functionalized tip). **c**, Constant-current dI/dV map recorded at a voltage bias of $V_s = +0.5$ V of N-6-ZGNR segment following tip-induced decoupling using a negative voltage sweep from $V_s = 0.0$ V to $V_s = -2.5$ V at the position marked by a blue cross in **a** ($V_{ac} = 11$ mV, $I_t = 200$ pA, $f = 455$ Hz, CO functionalized tip). Arrows mark the position of selected N-atom along the edge of the N-6-ZGNR.

Extended Data Fig. 5 Bond resolved imaging of decoupled N-6-ZGNRs.

a, Constant-height BRSTM image of a N-6-ZGNR segment from ($V_s = 0$ mV, modulation voltage $V_{ac} = 11$ mV, modulation frequency $f = 455$ Hz). **b**, Constant-height BRSTM image of the N-6-ZGNR segment in **a** following tip-induced decoupling using a positive voltage sweep from $V_s = 0.0$ V to $V_s = +2.5$ V at the position marked by a red cross in **a** ($V_s = 0$ mV, $V_{ac} = 11$ mV, $f = 455$ Hz). Arrows mark the position of selected N-atom along the edge of the N-6-ZGNR. **c**, STM topographic image of as-grown N-6-ZGNRs with CO-modified tip. **d**, STM topographic image of N-6-ZGNR after decoupling the GNR on the left. ($V_s = 50$ mV, $I_t = 20$ pA).

Extended Data Fig. 6 Calculated adsorption geometries of N-6-ZGNRs on Au(111).

a, Local minimum Adsorption Geometry I ($E = +0.312$ eV). Four C-atoms per unit cell interact through π -bonding (< 2.5 Å) with the Au(111) surface. The corrugation on opposing zigzag edges is in phase ($\varphi = 0$). **b**, Local minimum Adsorption Geometry II ($E = +0.355$ eV). Five C-atoms and one N-atom per unit cell interact through π -bonding (< 2.5 Å) with the Au(111) surface. The corrugation on opposing zigzag edges is phase shifted by $\varphi = 1/3\pi$. **c**, Global minimum Adsorption Geometry III ($E = +0.000$ eV). Four

C-atoms and two N-atoms per unit cell interact through π -bonding ($< 2.5 \text{ \AA}$) with the Au(111) surface. The corrugation on opposing zigzag edges is phase shifted by $\varphi = \pi$. All calculations performed with ultrasoft pseudopotentials and 40 Ry cut-offs.

Extended Data Fig. 7 Electronic structure of 6-ZGNR and N-6-ZGNRs.

a, GW band structure of freestanding 6-ZGNR (grey) and N-6-ZGNR (red) calculated using the same dimension unit cell. **b**, GW band structure of a freestanding N-6-ZGNR. The colour code shows the normalized contributions from C-atoms, N-atoms, and H-atoms to the wavefunction of each state. The number of $p_z + \sigma$ orbitals for C, N, and H atoms are 280, 8, and 10 per unit cell, respectively. The wavefunction projection of the state in the n^{th} band and at wavevector k to the C, N, and H atoms is described by P_{nk}^C , P_{nk}^N , and P_{nk}^H , respectively. We define the normalized percentage weight P_{nk}^C , P_{nk}^N , and P_{nk}^H as the wavefunction projection on the C, N, and H atoms per atom in the unit cell: $\langle P \rangle_{nk}^C = P_{nk}^C / 280$, $\langle P \rangle_{nk}^N = P_{nk}^N / 8$, $\langle P \rangle_{nk}^H = P_{nk}^H / 10$. The scale bar defines the mapping between the colour scale and the normalized percentage weight. **c**, Spatial distribution of the calculated spin polarized wavefunction for UNFB and LNFB. **d**, Spin unpolarized GW band structure of a freestanding N-6-ZGNR (with DFT within the local density approximation (LDA) as the starting point). In the spin unpolarized calculation the UNFB and LNFB form narrow non-splitting bands with a total band width smaller than 50 meV. **e**, Spatial distribution of the calculated spin unpolarized wavefunction for UNFB and LNFB.

Extended Data Fig. 8 Spatial localization of spin split low-lying nitrogen dopant flat band states.

Waterfall plot of dI/dV point spectra collected along a line marked in the inset long the edge of a N-6-ZGNR ($V_{\text{ac}} = 11 \text{ mV}$, $f = 455 \text{ Hz}$). When the STM tip is located immediately above the position of a nitrogen dopant atom the dI/dV point spectra show two characteristic peaks centred at $V_s = -$

2.60 ± 0.02 V and $V_s = -2.70 \pm 0.02$ V, corresponding to the UNFB and LNFB states, respectively.

Extended Data Fig. 9 dI/dV point spectroscopy of spin split low-lying nitrogen dopant flat band states.

dI/dV point spectroscopy recorded on four different decoupled N-6-ZGNR/Au(111) at the position above the N atoms. Ten dI/dV point spectra were collected at each position (grey). The respective averaged dI/dV point spectra are highlighted in red, orange, yellow, and blue (spectroscopy: $V_{ac} = 11$ mV, $f = 455$ Hz).

Extended Data Fig. 10 Constant-current dI/dV maps of decoupled N-6-ZGNR.

dI/dV maps recorded at voltage biases of **a**, $V_s = -2.000$ V, **b**, $V_s = -2.025$ V, **c**, $V_s = -2.050$ V, **d**, $V_s = -2.100$ V, **e**, $V_s = -2.125$ V, **f**, $V_s = -2.150$ V, **g**, $V_s = -2.200$ V, **h**, $V_s = -2.225$ V, **i**, $V_s = -2.250$ V, **j**, $V_s = -2.400$ V, **k**, $V_s = -2.500$ V, and **l**, $V_s = -2.800$ V ($V_{ac} = 11$ mV, $I_t = 2$ nA, $f = 455$ Hz).

Supplementary information

Supplementary Information

This file contains Supplementary Methods and Supplementary Figures 1–6.

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Blackwell, R.E., Zhao, F., Brooks, E. *et al.* Spin splitting of dopant edge state in magnetic zigzag graphene nanoribbons. *Nature* **600**, 647–652 (2021). <https://doi.org/10.1038/s41586-021-04201-y>

- Received: 18 April 2021
- Accepted: 02 November 2021
- Published: 22 December 2021
- Issue Date: 23 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04201-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

[Clever substitutions reveal magnetism in zigzag graphene nanoribbons](#)

- Aran Garcia-Lekue
- Daniel Sánchez-Portal

News & Views 22 Dec 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04201-y>.

- Article
- Open Access
- [Published: 22 December 2021](#)

Integrated photonics enables continuous-beam electron phase modulation

- [Jan-Wilke Henke](#)^{1,2} na1,
- [Arslan Sajid Raja](#)³ na1,
- [Armin Feist](#) [ORCID: orcid.org/0000-0003-1434-8895](#)^{1,2},
- [Guanhao Huang](#)^{3,4},
- [Germaine Arend](#)^{1,2},
- [Yujia Yang](#) [ORCID: orcid.org/0000-0003-1524-7973](#)^{3,4},
- [F. Jasmin Kappert](#)^{1,2},
- [Rui Ning Wang](#) [ORCID: orcid.org/0000-0002-5704-3971](#)^{3,4},
- [Marcel Möller](#) [ORCID: orcid.org/0000-0002-2518-5764](#)^{1,2},
- [Jiahe Pan](#)^{3,4},
- [Junqiu Liu](#) [ORCID: orcid.org/0000-0003-2405-6028](#)³,
- [Ofer Kfir](#) [ORCID: orcid.org/0000-0003-1253-9372](#)^{1,2},
- [Claus Ropers](#) [ORCID: orcid.org/0000-0002-9539-3817](#)^{1,2} &
- [Tobias J. Kippenberg](#) [ORCID: orcid.org/0000-0002-3408-886X](#)^{3,4}

Nature volume 600, pages 653–658 (2021)

- 6155 Accesses
- 102 Altmetric
- [Metrics details](#)

Subjects

- [Matter waves and particle beams](#)
- [Quantum metrology](#)
- [Quantum optics](#)

- [Sub-wavelength optics](#)

Abstract

Integrated photonics facilitates extensive control over fundamental light–matter interactions in manifold quantum systems including atoms¹, trapped ions^{2,3}, quantum dots⁴ and defect centres⁵. Ultrafast electron microscopy has recently made free-electron beams the subject of laser-based quantum manipulation and characterization^{6,7,8,9,10,11}, enabling the observation of free-electron quantum walks^{12,13,14}, attosecond electron pulses^{10,15,16,17} and holographic electromagnetic imaging¹⁸. Chip-based photonics^{19,20} promises unique applications in nanoscale quantum control and sensing but remains to be realized in electron microscopy. Here we merge integrated photonics with electron microscopy, demonstrating coherent phase modulation of a continuous electron beam using a silicon nitride microresonator. The high-finesse ($Q_0 \approx 10^6$) cavity enhancement and a waveguide designed for phase matching lead to efficient electron–light scattering at extremely low, continuous-wave optical powers. Specifically, we fully deplete the initial electron state at a cavity-coupled power of only 5.35 microwatts and generate >500 electron energy sidebands for several milliwatts. Moreover, we probe unidirectional intracavity fields with microelectronvolt resolution in electron-energy-gain spectroscopy²¹. The fibre-coupled photonic structures feature single-optical-mode electron–light interaction with full control over the input and output light. This approach establishes a versatile and highly efficient framework for enhanced electron beam control in the context of laser phase plates²², beam modulators and continuous-wave attosecond pulse trains²³, resonantly enhanced spectroscopy^{24,25,26} and dielectric laser acceleration^{19,20,27}. Our work introduces a universal platform for exploring free-electron quantum optics^{28,29,30,31}, with potential future developments in strong coupling, local quantum probing and electron–photon entanglement.

[Download PDF](#)

Main

The rapid advancement of electron microscopy epitomizes our growing ability to characterize the structure and function of nanoscale materials, devices and biological systems. Beyond stationary imaging, the advent of *in situ* and time-resolved electron microscopy allows for the observation of transient phenomena and non-equilibrium dynamics^{32,33,34,35}. In the form of photon-induced near-field electron microscopy (PINEM)⁶, ultrafast transmission electron microscopy (UTEM) permits quantitative, high-resolution imaging of nano-optical fields^{7,36,37,38}. The underlying mechanism

involves energy transfer between localized optical excitations and free electrons, modifying the state of the electron by generating discrete photon sidebands in a process termed stimulated inelastic electron-light scattering (IELS). The corresponding longitudinal phase modulation of the electronic wavefunction at optical frequencies is quantum-coherent in nature^{8,12}, and thus can be used for the coherent control of electron quantum states in space^{9,11,39,40} and time^{10,16,17}. Recent theoretical works suggest that such modulated electron beams enable an electron-mediated transfer of optical coherence, predicting the generation of coherent cathodoluminescence, the resonant excitation of two-level systems and superradiance from sequential electrons^{24,25,41,42,43}. These developments create opportunities for a closer integration of electron microscopy with coherent optical spectroscopy, based on local quantum control and enhanced sensing. Harnessing coherent electron–light interactions for scientific and technological applications is, however, hampered by its usual limitation to the ultrafast regime. Recent works implemented IELS^{44,45}, a ponderomotive laser phase plate²² and an attosecond modulator²³ at high continuous-wave powers. Low coupling efficiencies and structural damage have thus far limited inelastic scattering to weak interactions. Despite the use of phase matching and resonant amplification in dielectric laser accelerators¹⁹, prism geometries^{46,47} or free-space coupled whispering-gallery-mode microresonators⁴⁸, achieving strong phase modulation of an electron beam has remained out of reach of regular electron microscopes.

Here, we overcome this challenge and demonstrate highly efficient electron–photon interactions in the continuous-wave regime using an electron microscope and photonic integrated circuits based on Si₃N₄. Our set-up (Fig. 1) allows an electron beam to interact with the copropagating evanescent field of a microresonator waveguide in the object plane of a transmission electron microscope (TEM).

Fig. 1: Principle of continuous-wave photonic-chip-based optical phase modulation of free-electron beams.

 **figure1**

a, Rendering of the experimental set-up, including the electron microscope and a fibre-coupled Si_3N_4 photonic-chip-based microresonator. Inset: magnified interaction region with the electron beam passing the microresonator. CW, continuous wave; PD, photodiode. **b**, Left: during interaction with the laser-driven cavity mode, the initially narrow electron spectrum (black) develops discrete sidebands at integer multiples of the photon energy (red). Right: in a cQED depiction, the cavity photons induce transitions between the free-electron energy ladder states. **c**, Photograph of the fibre-coupled Si_3N_4 photonic chip mounted on a customized TEM holder. The triangular-shaped chip edge minimizes undesired electron–substrate interactions (the inset image shows the optical fibre glued to the input waveguide). **d**, Optical microscope image of the photonic chip showing the bus waveguide and the microresonator. The electron beam (green path, not to scale) traverses the microresonator parallel to the chip surface. **e**, Frequency-dependent effective index (n_{eff}) of the fundamental quasi-TM microresonator mode (orange). The integrated on-chip platform allows for phase matching to be achieved at different electron kinetic energies (blue, 90–145 keV) either by changing the dimensions of the Si_3N_4 waveguide or operating at different optical frequencies. For the waveguide shown in **f**, phase matching is achieved between the optical mode at ~ 193.5 THz (dashed orange line, corresponding to a wavelength of $\sim 1,550$ nm) and the free electrons at ~ 115 keV (dashed blue line). W , width; H , height. **f**, FEM simulation of the E_φ distribution of the fundamental quasi-TM mode of the microresonator (the green dot is an exemplary electron trajectory pointing into the page).

Free-electron–cavity photon interaction

The electron–photon interaction at the photonic chip (schematic in Fig. 1) is described by the Hamiltonian^{28,29} (see derivation in the [Supplementary Information](#))

$$\begin{aligned} \text{\$\$}\{ H \}_{\text{\rm int}} = & \frac{e}{2m} (\hat{p} \cdot \hat{A}) \cdot \hat{p} + \frac{1}{2} \hbar \omega_0^2 \hat{a}^\dagger \hat{a} \\ & + \frac{1}{2} \hbar \omega_0^2 \hat{b}^\dagger \hat{b} + \frac{1}{2} \hbar \omega_0^2 \hat{a}^\dagger \hat{b} + \frac{1}{2} \hbar \omega_0^2 \hat{b}^\dagger \hat{a} \end{aligned} \quad (1)$$

While the length gauge is usually chosen for localized quantum systems in cavity quantum electrodynamics (cQED)⁴⁹, the above velocity gauge is a natural choice for free electrons at finite momentum (\hat{p}) in an electromagnetic field with vector potential (\hat{A}) . The interaction between the optical mode ($(\hat{a}, \hat{a}^\dagger)$: annihilation/creation operator) and an electron ($(\hat{b}, \hat{b}^\dagger)$: electron-energy ladder operators, see [Supplementary Information](#)) is characterized by the vacuum coupling rate g_0 (refs. [28,29](#)), determined by the mode distribution and the matching of electron group and optical phase velocities:

$$\text{\$\$}\{ g \}_0 = \eta \sqrt{\frac{e^2 v_e^2}{2 \varepsilon \omega V}} \quad (2)$$

Here, e is the electron charge, v_e the electron group velocity, ε the optical permittivity, ω the angular frequency and V the effective optical mode volume. The phase matching condition is manifested in the coefficient η defined by the optical mode profile function $u(z)$ describing the electric field component along the electron trajectory, linked to the vector potential by $(\hat{A}(z)) = \hat{a}(z) + \Delta k u(z)$ and the electron wavenumber change $\Delta k = \omega/v_e$ upon photon absorption and emission^{8,50}. The scattering matrix, summarizing the effect of the interaction on both the electron and the cavity, is given by $(S = \alpha \exp(-i g_0 \tau) \hat{a}^\dagger \hat{b} + \text{H.c.})$ ²⁸. An empty cavity, therefore, facilitates electron-driven photon generation^{28,51,52} in a spontaneous Cherenkov or Smith–Purcell-like process⁵³, with coherence properties characteristic of transition radiation and so-called coherent cathodoluminescence⁵⁴. Populated by a coherent state $(|\alpha\rangle)$ from a laser, the free-electron–photon interaction reduces to a dimensionless coupling parameter $g = i\tau\alpha g_0$, where $|\alpha|^2 = n_c$ is the mean intracavity photon number, $\tau = L/v_e$ is the electron’s transit time over the interaction region L , and the scattering operator becomes a displacement operator acting on the electron state¹². The interaction produces electron energy sidebands separated from the initial energy E_0 by integer multiples of the photon energy $(N\hbar\omega)$ ($N \in \mathbb{Z}$, where N is the photon

order), with populations P_N following the Bessel functions of the first kind, $P_N = J_N(2|g|)^2$ (refs. [12,50](#)). In a position representation, the interaction imprints a sinusoidal phase modulation onto the electron wavefunction^{[50](#)}, which upon dispersive propagation will result in a density modulation of the electron beam^{[10,16,17](#)}.

Chip-based integrated photonics platform

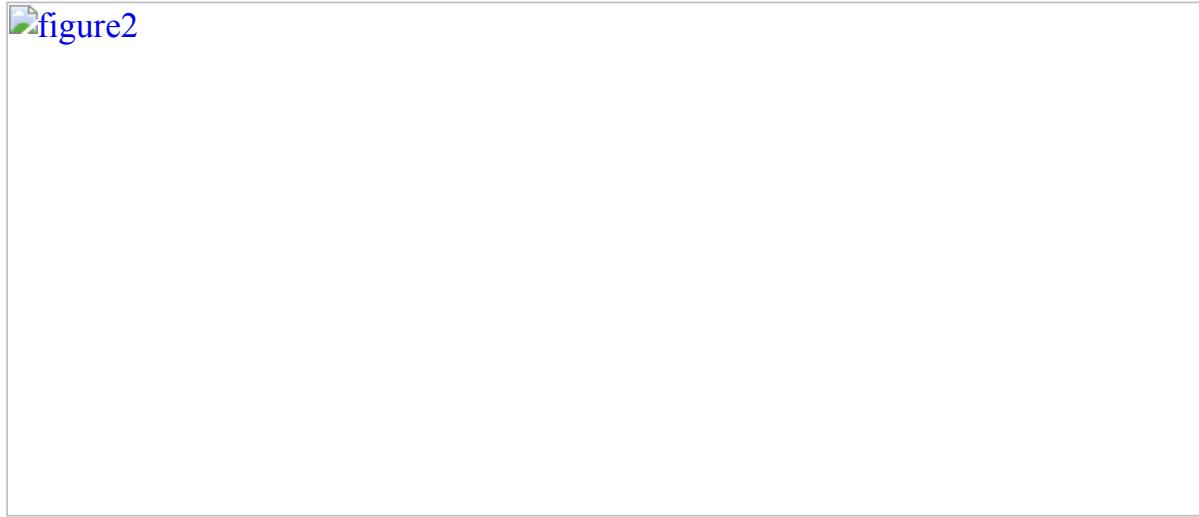
To facilitate high interaction strengths, we use photonic chip-based Si₃N₄ microresonators, a platform with many important features, including radiation hardness, high power handling^{[55](#)}, extremely low propagation losses (below 1 dB m⁻¹ at 1,550 nm)^{[56](#)} and flexibility, to engineer dispersion for phase matching. The chip was fabricated using the photonic Damascene process^{[56,57](#)} ([Methods](#)) without top oxide cladding to allow for efficient free-electron–light interactions with the evanescent field. The ring microresonator of 20 μm radius is coupled to a bus waveguide and placed close to a triangular edge of the chip to minimize undesired electron–substrate interactions (Fig. [1d](#) and [Methods](#)). For operation in the TEM, the photonic structure was packaged via ultrahigh numerical aperture (UHNA) fibres (Fig. [1c](#) and [Methods](#)). The Si₃N₄ microresonators employed here enable phase matching at different electron energies by modifying the waveguide geometry (Fig. [1e](#)); various established integrated platforms could further extend the phase matching range in terms of the electron energy and optical frequency (see [Supplementary Information](#)).

In the current study, we designed a ring microresonator (cross-section: 2 μm × 650 nm) to provide phase matching at an optical frequency of ~193 THz (wavelength $\lambda \approx 1,549$ nm) for a target electron energy of 115 keV. Figure [1f](#) shows a finite-element method (FEM, [Methods](#)) simulation of the quasi-transverse magnetic (quasi-TM) mode profile in terms of its major contributing field component E_φ along the electron propagation direction. Owing to the small mode area and considerable evanescent field component, we predict a vacuum coupling rate of $g_0/2\pi \approx 10^{11}$ Hz over an interaction time of $\tau \approx 10^{-13}$ s ($L \approx 19$ μm). The microresonator’s high Q factor facilitates a unity coupling constant $g \approx 1$ (the coupling constant g is a complex number; for simplicity, we use g in place of $\sqrt{|g|}$ hereafter) at a coupled optical power of $P = n_c \hbar \omega \kappa \approx 1$ μW, where $\kappa/2\pi = 390$ MHz is the measured cavity decay rate inside the electron microscope (the intrinsic Q factor of $Q_0 \approx 0.74 \times 10^6$ was slightly degraded in the microscope, see [Methods](#)).

Combined optical and electron spectroscopy

In the experiments, the fibre-coupled microresonator is driven by a 1,550 nm continuous-wave laser via the bus waveguide. It is mounted on a customized sample holder and placed in the object plane of the field-emission TEM (Fig. 1a and [Methods](#)). Parallel to the surface, the electron beam passes the waveguide and interacts with the confined optical mode (inset in Fig. 1a). After traversing the structure, the electron kinetic energy distribution is characterized with an imaging electron spectrometer in two different ways. Specifically, high-dispersion electron energy spectra are recorded by positioning a sharply focused electron beam in front of the microresonator. Alternatively, energy-filtered TEM using a collimated beam is used to image the interaction across the entire cavity mode in the near field. Whereas the former yields electron spectra for varying experimental parameters (Fig. 2), the latter enables imaging of individual sideband populations with high spatial resolution (Figs. 3 and 4).

Fig. 2: Simultaneous optical and electron spectroscopy of a high- Q microresonator mode.



a, A continuous-wave laser was used to excite the quasi-TM mode of the Si_3N_4 microresonator by using a polarization controller (PC). The relative optical frequency was calibrated by imparting sidebands (± 2 GHz) via an electro-optic phase modulator (EOM). The total transmitted and back-reflected light was detected to calibrate the power coupled into the clockwise propagating mode. CIR, circulator; OSC, oscilloscope; RF, radio-frequency synthesizer. **b**, Normalized transmission scan of the microresonator quasi-TM mode measured outside the TEM with a Q factor of $\sim 0.77 \times 10^6$ ($\kappa_0/2\pi = 112$ MHz, external coupling loss rate $\kappa_{\text{ex}}/2\pi = 139$ MHz) and a free spectral range of ~ 1.090 THz (see [Methods](#) for in situ optical characterization). **c**, Simultaneously measured optical transmission at the output waveguide (top) and $|g|^2$ retrieved from the electron energy spectra (bottom) while the electron beam interacted with the evanescent cavity field. The measured $|g|^2$ trace follows the power coupled to

the clockwise mode; a slight splitting is present due to modal coupling. FWHM, full-width at half-maximum. **d**, Example electron energy spectra for low (top; $g = 0$ (black), $g \approx 3.5$ (red) and $g \approx 6.7$ (blue)) and high (bottom, $g \approx 125$) optical powers. **e**, $|g|^2$ varies linearly with the optical power coupled to the clockwise mode of the cavity (slope: $|g|^2 = P/3.70 \mu\text{W}$). a.u., arbitrary units.

Fig. 3: High-resolution hyperspectral imaging of the quasi-TM microresonator mode.

 figure3

a, Spatial map of g at phase matching (115 keV electron energy) with a 1 nm image resolution. **b**, Exemplary electron spectrum retrieved from the energy-filtered spatial map (position indicated by a cross in **d**). **c**, N -dependent sideband population as a function of the distance $\langle \tilde{z} \rangle$ to the chip surface. **d**, Energy-filtered images of selected N (indicated). The energy window has a width of 800 meV and the position of the microresonator is indicated in grey.

Fig. 4: Phase matching and Ramsey-type interference.

 figure4

a, Geometry (side view) of the electron beam passing the photonic chip and interacting with the quasi-TM microresonator mode. The double interaction and phase shifts between the light and electron wave led to a characteristic spatial distribution of g due to Ramsey-type interference. **b**, Comparison of experimental (red crosses, see **c**) and

numerically simulated (blue line, see **d**) maximum g as a function of the electron kinetic energy. **c**, Quantitative g maps for variable electron energies, showing the spatial distribution and amplitude of the electron–light interaction. The radial position of the Si_3N_4 waveguide is indicated. **d**, Numerical simulation of g as a function of electron energy and lateral position just above the surface. The white dots overlay the experimentally observed minima retrieved from **c** (80–200 kV electron beam energies, the waveguide positions are marked by white lines).

We first investigate the strength of g using a focused electron probe, recording electron spectra while scanning the laser over a resonance employing the set-up depicted in Fig. [2a](#). The transmission spectrum, displayed in Fig. [2b](#), shows the quasi-TM microresonator modes spaced by a free spectral range of 1.090 THz. Figure [2c](#) shows simultaneous optical and in situ electron spectroscopy of the laser-excited mode. An electro-optical modulator (driven at 2.0 GHz) is used to generate sidebands that can be observed in the transmission spectrum to calibrate the optical frequency. The laser is tuned to a single optical mode at \sim 1,549.4 nm ($\kappa/2\pi = 390$ MHz, [Methods](#)) and the focused electron beam (120 keV beam energy, 25 nm focal spot size, 1 mrad convergence semi-angle) is centred just above the surface of the microresonator (Fig. [3](#)) to record electron spectra for a stationary beam. Harnessing the high- Q intracavity enhancement, we observe strong populations P_N in multiple N , reaching the regime of strong phase modulation for a continuous laser light source and electron beam (Fig. [2d](#)). These results relate to sideband generation in a recent experiment using semi-transparent membranes^{[23](#)} and a study employing dielectric gratings posted at the same time as our work^{[14](#)}. The coupling parameter g is retrieved from the spectra ([Methods](#)), whereas the optical power coupled to the clockwise-propagating mode is determined from the recorded optical transmission and reflection data (Fig. [2e](#)). We observe the expected linear dependence $|g|^2 = P/3.70 \mu\text{W}$ of the coupling on the in-coupled clockwise optical power. From this, we find that an optical power of $5.35 \mu\text{W}$ is required to suppress the electron zero-energy-loss peak at $g \approx 1.2$, and a value of $g \approx 40$ for an optical power of about 6 mW. For a power of \sim 38 mW in the bus waveguide, we generate >500 photon sidebands ($g \approx 125$, Fig. [2d](#)). While state-of-the-art dielectric laser accelerators achieve the highest peak acceleration gradients for sub-relativistic electron beams using femtosecond lasers^{[19,58](#)}, our approach enables continuous acceleration of microbunched electron beams^{[10,26](#)} with a gradient of 330 keV m^{-1} per milliwatt of optical power. In terms of the input peak power required, this is an improvement in efficiency of four orders of magnitude over free-space coupled dielectric structures^{[47,48](#)}.

Representing a particular strength of the approach introduced here, integrated photonics permits the controlled interaction with a single optical mode, as described in the following. The spectral line shape of the resonance is analysed by simultaneously recording the transmitted optical power and extracting g for varying frequency

detuning (Fig. 2c). The optical transmission trace, formed by the interference of the input light with the light coupled out from the clockwise-propagating resonator mode, exhibits a full-width at half-maximum of 560 MHz (total line width $\kappa/2\pi = 390$ MHz, [Supplementary Information](#)). Interestingly, the electron spectra, which are sensitive only to the intracavity power stored in the clockwise-rotating, copropagating optical mode, display a double-peaked structure originating from coupling the power to the frequency-degenerate counterclockwise mode⁵⁹. The differences in width and shape between the optical and electron spectroscopic measurements are explained by the interference in the optical transmission channel with the input field, and both curves can be fitted consistently in one model as presented in the [Supplementary Information](#) (Supplementary Figs. 1 and 2). These data demonstrate continuous-wave electron energy gain spectroscopy^{21,60} at an extraordinarily narrow spectral feature of only 3.2 μ eV in width (full-width at half-maximum; 1 μ eV peak separation).

Mode imaging and phase matching

The implementation of chip-based high- Q optical microresonators as efficient electron phase modulators requires insights into the real-space scattering distribution. Employing a collimated electron beam and an imaging energy filter (800 meV energy pass band), we record spatial maps of P_N (Fig. 3d, [Methods](#)). A photon-order resolved spectrum is extracted for each image pixel (Fig. 3b), and g is retrieved with a 1 nm spatial resolution (Fig. 3a). Using a continuous electron beam circumvents previously encountered signal-to-noise ratio limitations in PINEM at about 3–4 orders of magnitude lower flux. The distance-dependent sideband occupations (Fig. 3c) reveal a strong and high-contrast modulation, indicating the absence of spatial and temporal averaging. We note that IELS is typically accompanied by a three-dimensional momentum transfer^{9,11}; however, in the current experiment, these transverse deflections did not limit the resolution for in-focus imaging ([Supplementary Information](#)).

The microresonator waveguide was designed to match the phase velocity of the excited optical mode with the group velocity of electrons at an energy of around 115 keV. Figure 4 shows the spatial pattern of the electron–light coupling near the ring-shaped microresonator (Fig. 4a). The electron-energy-dependent near-field maps reveal an oscillatory modulation of g along the chip surface (Fig. 4c) and an amplitude change with electron energy (Fig. 4b). At phase matching, the spatial distribution of g closely resembles the electric field configuration of the excited quasi-TM mode (Fig. 3a), and the most efficient coupling is achieved for a beam passing the periphery of the ring (tangent line). Trajectories closer to the centre of the microresonator (secant line) are governed by two sequential interactions with the microresonator mode. This results

in Ramsey-type constructive or destructive interference that depends on the relative phase of both individual interactions¹³.

Numerical FEM simulations of the optical mode profile (Fig. 4d, [Methods](#)), considering the electron trajectory through the near field and the microresonator power enhancement, closely reproduce the experimental observations, including the position of the minima in g (white dots). Excellent agreement is achieved for the electron energy dependence of the maximum g (Fig. 4b). In contrast to nanoscale-confined optical fields, a steep cutoff in electron–light coupling is found for beam energies below ~ 100 keV because the optical mode is devoid of higher-momentum electric field components. At high electron energies, the coupling strength reaches a plateau, which enables broadband temporal phase plates for TEM. Besides enhancing the total coupling via a larger interaction time, an increase in microresonator radius will further sharpen the phase matching condition in energy.

Conclusion

In summary, we demonstrate highly efficient, phase-matched and single-optical-mode interaction of free electrons with photonic-chip-based high- Q microresonators, driving transitions at microwatt continuous optical pump powers. Simultaneous in situ optical and electron measurements show that this cQED-type setting yields a quantitative understanding of the interaction and facilitates electron energy gain spectroscopy at the microelectronvolt level. Our approach may enable a seamless integration of temporal phase plates in electron microscopy over a broad range of energies, with various applications in advanced longitudinal and transverse electron beam shaping for high-fidelity attosecond metrology and sub-cycle probing, beam blanking and high-frequency beam modulation. Coherent beam separation in the optical field may enable new variants of inelastic electron holography, and sequential interactions will provide sensitivity enhancements in coherent electron spectroscopy. We envisage various possible architectures featuring anomalous dispersion to directly probe the multimode intracavity field in nonlinear effects such as the formation and dynamics of dissipative Kerr solitons, or coupled microresonators for advanced electron state preparation and readout. The present study provides a missing link in the realization of various theoretical concepts in free-electron cQED by establishing full control over the optical input and output channels of a single confined integrated microresonator mode coupled to an electron beam. Future investigations will address cavity-mediated electron–photon and electron–electron entanglement using coincidence detection, promising enhanced electron imaging and electron-heralded single-photon sources.

Methods

Numerical simulations

The effective refractive index (n_{eff}) and the mode profile of the Si_3N_4 microresonator are calculated via FEM simulations (COMSOL Multiphysics) (Fig. 1f). Mode analysis is performed for the two-dimensional axially symmetric model with the cross-section of the microresonator. The Si_3N_4 core has a rectangular cross-section with width W and height H . The microresonator has a SiO_2 bottom cladding and an air (vacuum) top cladding. The refractive index of Si_3N_4 used in the simulation was obtained from in-house ellipsometry measurements and found to be $n_K = 1.9923$ (J. A. Woollam, unpublished data), whereas the refractive index of SiO_2 was obtained from ref. [61](#) as $n_M = 1.4440$. From the electric field distribution in the cross-section ((r, \tilde{z})), the complex electric field along the electron trajectory (E_{physical}) related to the physical field by $E_{\text{physical}} = \text{real}[E_{\text{traj}}]$, is determined via $(E_{\text{physical}})(r, \tilde{z}) = E_{\text{traj}}(r, \tilde{z}) \cos(z/r) + i E_{\text{traj}}(r, \tilde{z}) \sin(z/r)$. The coupling strength $(g(r, \tilde{z}, E))$ is then found by evaluating $(g = \frac{e}{2\hbar\omega} \int_{-\infty}^{\infty} v \rm{d}z)$ (refs. [12,50](#)) for different electron energies E , where $v = c\sqrt{1 - 1/(1+E/m)^2}$ is the relativistic electron velocity.

Fibre-integrated silicon nitride microresonators

The microresonators used in this study are made from Si_3N_4 , a platform that is CMOS compatible with radiation hardness capability^{[62](#)} and ultra-low propagation losses^{[56,57](#)}. They also allow for dispersion engineering and efficient fibre-to-chip coupling. Si_3N_4 -based integrated photonic circuits have been driving major progress in nonlinear optical devices, in particular soliton microcombs^{[63](#)} that are already used in numerous system-level applications ranging from coherent telecommunication^{[64](#)} to astrophysical spectrometer calibration^{[65,66](#)}. Here, the designed microresonators were fabricated using the photonic Damascene process^{[56,57](#)}. A 4-inch silicon substrate with 4- μm -thick thermal wet SiO_2 cladding was used. The substrate was coated with deep-ultraviolet (DUV) photoresist, and microresonators and bus waveguides were patterned on the substrate via DUV stepper photolithography (248 nm KrF excimer laser). The pattern was then dry-etched into the SiO_2 cladding using $\text{C}_4\text{H}_8/\text{O}_2/\text{He}$ etchants to create waveguide preforms. Stoichiometric Si_3N_4 film was deposited on the patterned substrate via low-pressure chemical vapour deposition (LPCVD), to fill the waveguide preforms and to form the waveguide cores. Afterwards, etchback and chemical-mechanical polishing^{[56](#)} were used to planarize the substrate and remove excess Si_3N_4 . The entire substrate was further annealed at 1,200 °C to remove the residual hydrogen

content in Si_3N_4 , to reduce hydrogen-induced absorption loss in the Si_3N_4 waveguides. This high-temperature annealing is critical to fabricate low-propagation-loss waveguides at telecommunication bands around 1,550 nm. No top SiO_2 cladding was added on top of the Si_3N_4 waveguides, such that the electron beam could interact with the optical mode. Finally, photolithography with alignment was used to precisely position the Si_3N_4 microresonator close to the chip edge (within a distance $<10 \mu\text{m}$, key for aligning the electron beam with the Si_3N_4 microresonator). Deep reactive-ion etching was used to separate the entire substrate into hundreds of individual dies/chips for the following die-level integration with fibres.

The light is coupled into and out of the Si_3N_4 photonic chip on one edge, while the microresonator is placed on the other edge. To avoid cutting the electron beam by the substrate (Fig. 1c), the near edge was clipped, forming a tight supporting triangle around the resonator ring to avoid damage during the chip release. A bus waveguide (dimensions $800 \text{ nm} \times 650 \text{ nm}$) is used to evanescently couple the light into the multimode microresonator (dimensions $2 \mu\text{m} \times 650 \text{ nm}$) to achieve a better coupling ideality⁶⁷. Inverse-taper waveguides⁶⁸ are used to facilitate efficient light coupling via UHNA-7 fibres (mode field diameter $\approx 3.2 \mu\text{m}$ at 1,550 nm). A 2–3-cm-long UHNA-7 fibre with a thermally expandable core was spliced to standard single-mode fibre (SMF-28) with a splicing loss of $<0.2 \text{ dB}$. A chip through-coupling efficiency (fibre–chip–fibre) of $>25\%$ is achieved using the UHNA-7 fibre. The photonic packaging was done by first aligning the UHNA-7 fibre via a custom-built holder to optimize the coupling. Then, a small drop of epoxy ($\sim 150 \mu\text{m}$, Fig. 1b) was dispensed using a precise pneumatic valve and cured using a UV lamp in four small time steps ($\sim 2\text{--}5 \text{ min}$). A long-term ($\sim 1\text{--}2 \text{ days}$) coupling stability test was performed at low optical power by monitoring the transmitted light. The broadband characterization of the microresonators was performed by employing a widely tunable diode laser. The transmission spectrum was calibrated using a self-referenced optical frequency comb and Mach–Zehnder interferometer (MZI)⁶⁹. The resonance is fitted using models explained in the [Supplementary Information](#) (fitting model) to extract $\kappa_0/2\pi$ (intrinsic loss rate), $\kappa_{\text{ex}}/2\pi$ (external coupling loss rate), and $\gamma/2\pi$ (mode coupling rate) for both quasi-TE and quasi-TM mode families. The mean intrinsic linewidth is $\sim 110\text{--}120 \text{ MHz}$ ($Q_0 \approx 1.75 \times 10^6$) for the quasi-TM mode family. The cavity resonance centre at $\sim 1,549 \text{ nm}$ is critically coupled with $\kappa_0/2\pi = 112 \text{ MHz}$ and $\kappa_{\text{ex}}/2\pi = 139 \text{ MHz}$ ($Q \approx 0.77 \times 10^6$).

Optical characterization of microresonators inside the TEM

The packaged sample is transferred into the TEM using a customized holder with vacuum fibre feedthroughs. The fibres connected to the microresonator chip are fed

through the hollowed pipe part of the holder, and the T-shaped base holding the photonic chip is mounted on an adaptor that allows the placement of the entire structure in the sample region of the TEM. The optical set-up used to perform the spectroscopy in the TEM is shown in Fig. 2a. The set-up is driven by a CW-laser (Toptica CTL 1550), with a maximal power of 40 mW, a tunable wavelength from 1,510 nm to 1,630 nm, and a linewidth below 10 kHz. The laser is coupled to an electro-optic modulator (EOM) that allows for frequency calibration of the transmission scan and the dependence of g on the detuning from the resonance frequency. A polarization controller is used to align the input light to either the quasi-TE or TM mode of the microresonator. An optical circulator is used to probe the light reflected from the microresonator due to the bulk and Rayleigh surface scattering responsible for the splitting of clockwise and counterclockwise cavity modes. The optical transmission and reflection are measured simultaneously to calibrate the intracavity power coupled to the clockwise mode (see [Supplementary Information](#)). We note that a degradation of the Q factor from 0.77×10^6 to 0.49×10^6 , resulting in an additional increase in total linewidth (~ 150 MHz), was observed when the sample was transferred into the TEM, which could be related to some charging, strain or contamination of the resonator chip. Theory predicts that $|g|^2$ is proportional to the photon number n_{cw} of the clockwise mode optical cavity mode. The intracavity photon numbers are varied by sweeping the laser frequency through the optical resonance to find the $|g|^2$ dependence on optical power. The $g(t)$ value is extracted for each time step of the sweep, t , by fitting the measured electron spectrum, and the intracavity photon number $n_{\text{cw}}(t)$ is retrieved from the optical transmission signal using the steady-state Langevin equation. By binning the $|g(t)|^2$ trace in 20 ms steps and plotting it against the coupled optical power in the clockwise mode $\langle P \rangle = \kappa \hbar \omega \langle n \rangle \langle \{ \text{rm} \{ \text{cw} \} \} (t) \rangle$, we obtain the power sweep curve shown in Fig. 2e. More detailed data analysis procedures and the analytical expression of various field quantities can be found in the [Supplementary Information](#).

Electron microscope and experimental set-up

The *in situ* experiments were performed in the Göttingen UTEM⁷⁰, which is based on a thermal field-emission TEM (JEM 2100F, JEOL Ltd). Achieving temporal resolutions in the femtosecond regime^{70,71,72}, UTEM has enabled stroboscopic real-space movies of numerous processes, including strain wave dynamics^{34,73,74}, ultrafast demagnetization^{75,76,77} and the evolution of structural phase transformations^{33,35}. In the current experiments, the electron gun is operated in the extended Schottky regime, yielding a continuous electron beam with an initial energy spread of 0.5 eV at variable electron energies from 80 to 200 keV and typical beam currents of 10–50 pA in the sample region. Preventing clipping of the electron beam at the extended microresonator chip (~ 50 μm sample height at the probing position), a low beam

convergence is set by turning off the magnetic objective lens in low magnification (LM) TEM and STEM mode, for high-resolution imaging and spectroscopy, respectively. For analysing the electron beam passing the microresonator structure, a post-column imaging energy filter (CEFID, CEOS) is employed, equipped with a scintillator-coupled CMOS camera (TemCam-XF416ES, TVIPS GmbH; used for all measurements, if not mentioned otherwise) and a hybrid-pixel electron detector (EM CheeTah T3, Amsterdam Scientific Instruments). The sample tilt is adjusted carefully to align the electron beam parallel to the surface of the microresonator chip, probing the local electron–light interaction at the microresonator.

Local spectroscopy of the microresonator mode

Electron spectroscopy (Fig. 2) is implemented in LowMAG STEM mode (indicated magnification: $\times 1,000$) at 120 keV electron energy, achieving an electron focal spot size of about 25 nm (FWHM) with beam semi-convergence angles of 1.1 mrad (100 μm condenser aperture, Fig. 2d) and 0.45 mrad (40 μm condenser aperture, Fig. 2c (lower panel) and e), respectively. The electron beam is either positioned about 30 nm over the microresonator surface, such that no clipping or beam distortion was observed (Fig. 2d), or at a distance of 380 nm from the microresonator surface (Fig. 2c,e), as estimated from the optical power in the bus waveguide and the exponential decay of the optical mode. We observe some charging of the dielectric microresonator chip when using a tightly focused electron beam. Therefore, larger distances to the surface or fast beam sweeps are employed, and future structures will address this point by adding a thin conductive coating on parts of the microresonator chip. The STEM focal plane is set to the middle of the ring microresonator and the lateral beam position relative to the chip is software controlled (Filter Control, CEOS) via an external scan generator (USG, TVIPS). Single spectra are recorded with an integration time of 100 ms (Fig. 2d). Electron spectroscopy with fast sweeps of the laser frequency is captured by an event-based hybrid pixel detector (EM CheeTah T3, Timepix3 ASIC, 1.56 ns timing precision) and discrete binning in time (100 μs windows, Fig. 2c,e). For each spectrum, the electron–light coupling constant g is extracted by fitting the shape and amplitudes of the individual spectral sidebands N with a comb of Voigt-peaks and normalized occupations $P_N = J_N(2|g|)^2$. For details on the fitting procedure, see refs. 12,48. By considering the change of the coupling constant g as a function of frequency detuning, the optical linewidth can be extracted, and the achievable energy resolution in electron energy-gain spectroscopy is ultimately limited by the laser linewidth. The concept can be transferred to arbitrary high- Q optical modes, with previously demonstrated spectral linewidths below $\kappa/2\pi \approx 10$ MHz (ref. 56). In comparison, state-of-the-art monochromated EELS probes material excitations with a zero-loss peak (ZLP) of sub-10-meV energy width^{78,79,80}.

Energy-filtered imaging and quantitative g maps

Quantitative maps of the electron–light coupling constant g are retrieved by energy-filtered transmission electron microscopy (EFTEM) imaging (Figs. 3 and 4), resolving individual photon sidebands with a double-sided energy window of only 800 meV width and typical bi-sided r.m.s. non-isochromaticity (energy window edge sharpness) of about 30 meV (8 mm field of view of the spectrometer entrance aperture plane, for technical details, see ref. 81). Changing the microscope’s high tension (step size of 0.8 V) relative to the centre beam energy (80–200 keV) allows for the recording of images $I_N(x, y)$ that contain inelastically scattered electrons, losing or gaining an integer number N of photon energies $\hbar\omega = 0.8$ eV (Fig. 3c). For each spectral sideband, an image is recorded with 1-s integration time, resulting in energy-filtered TEM datasets containing up to 121 image slices, enabling a quantitative retrieval of the coupling constant g at each image pixel (see Fig. 3d). Accounting for a large number of image positions (up to $4,096 \times 4,096$ pixels in Fig. 3), the data are collocated into 1,024 subsets using the k -means++ clusters algorithm (implemented in MATLAB 2020b, MathWorks Inc.), with the sideband occupations as the input vector. The resulting centroid locations (metric: squared Euclidean distances) give a set of averaged spectra. For each of these, the coupling constant g , relative uncertainty dg/g (with standard deviation dg of Gaussian distribution in g) and spectral amplitudes A are fitted to the recorded image stacks $I_N(x, y)$. The averaged intensity spread of the ZLP to neighbouring energy windows, B , is accounted for by applying a convolution $I_N(x, y) = A(x, y) \cdot P_N(x, y) * B$ (at 115 keV beam energy, the 800 meV energy window contained about 82% of the ZLP intensity). Furthermore, remaining ZLP intensity at stronger coupling ($g > 3$), resulting from limited rejection contrast of the spectrometer slits and scattering at the entrance aperture, is disregarded (about 2(0.8)% diffuse background for $N = 0(\pm 1)$). Finally, the retrieved fitting parameters are assigned to the individual pixels before clustering, resulting in a quantitative map of the electron–light coupling constant with 1 nm spatial resolution (Figs. 3 and 4).

Data availability

The data that support the plots within this paper and other findings of this study are available on Zenodo (<https://doi.org/10.5281/zenodo.5575752>). All other data used in this study are available from the corresponding authors upon reasonable request.

References

1. 1.

Thompson, J. D. et al. Coupling a single trapped atom to a nanoscale optical cavity. *Science* **340**, 1202–1205 (2013).

2. 2.

Mehta, K. K. et al. Integrated optical multi-ion quantum logic. *Nature* **586**, 533–537 (2020).

3. 3.

Niffenegger, R. J. et al. Integrated multi-wavelength control of an ion qubit. *Nature* **586**, 538–542 (2020).

4. 4.

Lodahl, P., Mahmoodian, S. & Stobbe, S. Interfacing single photons and single quantum dots with photonic nanostructures. *Rev. Mod. Phys.* **87**, 347–400 (2015).

5. 5.

Sipahigil, A. et al. An integrated diamond nanophotonics platform for quantum-optical networks. *Science* **354**, 847–850 (2016).

6. 6.

Barwick, B., Flannigan, D. J. & Zewail, A. H. Photon-induced near-field electron microscopy. *Nature* **462**, 902–906 (2009).

7. 7.

Piazza, L. et al. Simultaneous observation of the quantization and the interference pattern of a plasmonic near-field. *Nat. Commun.* **6**, 6407 (2015).

8. 8.

García de Abajo, F. J. Optical excitations in electron microscopy. *Rev. Mod. Phys.* **82**, 209–275 (2010).

9. 9.

Vanacore, G. M. et al. Attosecond coherent control of free-electron wave functions using semi-infinite light fields. *Nat. Commun.* **9**, 2694 (2018).

10. 10.

Priebe, K. E. et al. Attosecond electron pulse trains and quantum state reconstruction in ultrafast transmission electron microscopy. *Nat. Photon.* **11**, 793–797 (2017).

11. 11.

Feist, A., Yalunin, S. V., Schäfer, S. & Ropers, C. High-purity free-electron momentum states prepared by three-dimensional optical phase modulation. *Phys. Rev. Res.* **2**, 043227 (2020).

12. 12.

Feist, A. et al. Quantum coherent optical phase modulation in an ultrafast transmission electron microscope. *Nature* **521**, 200–203 (2015).

13. 13.

Echternkamp, K. E., Feist, A., Schäfer, S. & Ropers, C. Ramsey-type phase control of free-electron beams. *Nat. Phys.* **12**, 1000–1004 (2016).

14. 14.

Dahan, R. et al. Imprinting the quantum statistics of photons on free electrons. *Science* **373**, eabj7128 (2021).

15. 15.

Sears, C. M. S. et al. Production and characterization of attosecond electron bunch trains. *Phys. Rev. Spec. Top. Accel. Beams* **11**, 061301 (2008).

16. 16.

Morimoto, Y. & Baum, P. Diffraction and microscopy with attosecond electron pulse trains. *Nat. Phys.* **14**, 252–256 (2018).

17. 17.

Kozák, M., Schönenberger, N. & Hommelhoff, P. Ponderomotive generation and detection of attosecond free-electron pulse trains. *Phys. Rev. Lett.* **120**, 103203 (2018).

18. 18.

Madan, I. et al. Holographic imaging of electromagnetic fields via electron-light quantum interference. *Sci. Adv.* **5**, eaav8358 (2019).

19. 19.

England, R. J. et al. Dielectric laser accelerators. *Rev. Mod. Phys.* **86**, 1337–1389 (2014).

20. 20.

Sapra, N. V. et al. On-chip integrated laser-driven particle accelerator. *Science* **367**, 79–83 (2020).

21. 21.

García de Abajo, F. J. G. & Kociak, M. Electron energy-gain spectroscopy. *New J. Phys.* **10**, 073035 (2008).

22. 22.

Schwartz, O. et al. Laser phase plate for transmission electron microscopy. *Nat. Methods* **16**, 1016–1020 (2019).

23. 23.

Ryabov, A., Thurner, J. W., Nabben, D., Tsarev, M. V. & Baum, P. Attosecond metrology in a continuous-beam transmission electron microscope. *Sci. Adv.* **6**, eabb1393 (2020).

24. 24.

Gover, A. & Yariv, A. Free-electron–bound-electron resonant interaction. *Phys. Rev. Lett.* **124**, 064801 (2020).

25. 25.

Zhao, Z., Sun, X.-Q. & Fan, S. Quantum entanglement and modulation enhancement of free-electron–bound-electron interaction. *Phys. Rev. Lett.* **126**, 233402 (2021).

26. 26.

Yalunin, S. V., Feist, A. & Ropers, C. Tailored high-contrast attosecond electron pulses for coherent excitation and scattering. *Phys. Rev. Res.* **3**, L032036 (2021).

27. 27.

Shiloh, R. et al. Electron phase-space control in photonic chip-based particle acceleration. *Nature* **597**, 498–502 (2021).

28. 28.

Kfir, O. Entanglements of electrons and cavity photons in the strong-coupling regime. *Phys. Rev. Lett.* **123**, 103602 (2019).

29. 29.

Di Giulio, V., Kociak, M. & de Abajo, F. J. G. Probing quantum optical excitations with fast electrons. *Optica* **6**, 1524–1534 (2019).

30. 30.

Pan, Y. & Gover, A. Spontaneous and stimulated emissions of a preformed quantum free-electron wave function. *Phys. Rev. A* **99**, 052107 (2019).

31. 31.

Ben Hayun, A. et al. Shaping quantum photonic states using free electrons. *Sci. Adv.* **7**, eabe4270 (2021).

32. 32.

Zewail, A. H. Four-dimensional electron microscopy. *Science* **328**, 187–193 (2010).

33. 33.

van der Veen, R. M., Kwon, O.-H., Tissot, A., Hauser, A. & Zewail, A. H. Single-nanoparticle phase transitions visualized by four-dimensional electron microscopy. *Nat. Chem.* **5**, 395–402 (2013).

34. 34.

Cremons, D. R., Plemmons, D. A. & Flannigan, D. J. Femtosecond electron imaging of defect-modulated phonon dynamics. *Nat. Commun.* **7**, 11230 (2016).

35. 35.

Danz, T., Domröse, T. & Ropers, C. Ultrafast nanoimaging of the order parameter in a structural phase transition. *Science* **371**, 371–374 (2021).

36. 36.

Yurtsever, A., van der Veen, R. M. & Zewail, A. H. Subparticle ultrafast spectrum imaging in 4D electron microscopy. *Science* **335**, 59–64 (2012).

37. 37.

Wang, K. et al. Coherent interaction between free electrons and a photonic cavity. *Nature* **582**, 50–54 (2020).

38. 38.

Liebtrau, M. et al. Spontaneous and stimulated electron– photon interactions in nanoscale plasmonic near fields. *Light Sci. Appl.* **10**, 82 (2021).

39. 39.

Vanacore, G. M. et al. Ultrafast generation and control of an electron vortex beam via chiral plasmonic near fields. *Nat. Mater.* **18**, 573–579 (2019).

40. 40.

García de Abajo, F. J. & Konečná, A. Optical modulation of electron beams in free space. *Phys. Rev. Lett.* **126**, 123901 (2021).

41. 41.

Gover, A. et al. Superradiant and stimulated-superradiant emission of bunched electron beams. *Rev. Mod. Phys.* **91**, 035003 (2019).

42. 42.

Di Giulio, V., Kfir, O., Ropers, C. & García de Abajo, F. J. Modulation of cathodoluminescence emission by interference with external light. *ACS Nano* acsnano.1c00549 (2021).

43. 43.

Kfir, O., Di Giulio, V., de Abajo, F. J. G. & Ropers, C. Optical coherence transfer mediated by free electrons. *Sci. Adv.* **7**, eabf6380 (2021).

44. 44.

Das, P. et al. Stimulated electron energy loss and gain in an electron microscope without a pulsed electron gun. *Ultramicroscopy* **203**, 44–51 (2019).

45. 45.

Liu, C. et al. Continuous wave resonant photon stimulated electron energy-gain and electron energy-loss spectroscopy of individual plasmonic nanoparticles. *ACS Photon.* **6**, 2499–2508 (2019).

46. 46.

Kozák, M. et al. Acceleration of sub-relativistic electrons with an evanescent optical wave at a planar interface. *Opt. Express* **25**, 19195–19204 (2017).

47. 47.

Dahan, R. et al. Resonant phase-matching between a light wave and a free-electron wavefunction. *Nat. Phys.* **16**, 1123–1131 (2020).

48. 48.

Kfir, O. et al. Controlling free electrons with optical whispering-gallery modes. *Nature* **582**, 46–49 (2020).

49. 49.

Kimble, H. J. Strong interactions of single atoms and photons in cavity QED. *Phys. Scripta* **T76**, 127 (1998).

50. 50.

Park, S. T., Lin, M. & Zewail, A. H. Photon-induced near-field electron microscopy (PINEM): theoretical and experimental. *New J. Phys.* **12**, 123028 (2010).

51. 51.

Bendaña, X., Polman, A. & García de Abajo, F. J. Single-photon generation by electron beams. *Nano Lett.* **11**, 5099–5103 (2011).

52. 52.

Müller, N. et al. Broadband coupling of fast electrons to high-Q whispering-gallery mode resonators. *ACS Photon.* acsphotonics.1c00456 (2021).

53. 53.

Christopher, J. et al. Electron-driven photon sources for correlative electron-photon spectroscopy with electron microscopes. *Nanophotonics* **9**, 4381–4406 (2020).

54. 54.

Polman, A., Kociak, M. & García de Abajo, F. J. Electron-beam spectroscopy for nanophotonics. *Nat. Mater.* **18**, 1158–1171 (2019).

55. 55.

Brasch, V. et al. Photonic chip– based optical frequency comb using soliton Cherenkov radiation. *Science* **351**, 357–360 (2016).

56. 56.

Liu, J. et al. High-yield, wafer-scale fabrication of ultralow-loss, dispersion-engineered silicon nitride photonic circuits. *Nat. Commun.* **12**, 2236 (2021).

57. 57.

Pfeiffer, M. H. P. et al. Photonic Damascene process for integrated high-Q microresonator based nonlinear photonics. *Optica* **3**, 20–25 (2016).

58. 58.

Leedle, K. J. et al. Dielectric laser acceleration of sub-100 keV electrons with silicon dual-pillar grating structures. *Opt. Lett.* **40**, 4344 (2015).

59. 59.

Kippenberg, T. J., Spillane, S. M. & Vahala, K. J. Modal coupling in traveling-wave resonators. *Opt. Lett.* **27**, 1669–1671 (2002).

60. 60.

Pomarico, E. et al. meV resolution in laser-assisted energy-filtered transmission electron microscopy. *ACS Photon.* **5**, 759–764 (2018).

61. 61.
Malitson, I. H. Interspecimen comparison of the refractive index of fused silica. *J. Opt. Soc. Am.* **55**, 1205–1209 (1965).
62. 62.
Brasch, V., Chen, Q.-F., Schiller, S. & Kippenberg, T. J. Radiation hardness of high-Q silicon nitride microresonators for space compatible integrated optics. *Opt. Express* **22**, 30786–30794 (2014).
63. 63.
Kippenberg, T. J., Gaeta, A. L., Lipson, M. & Gorodetsky, M. L. Dissipative Kerr solitons in optical microresonators. *Science* **361**, eaan8083 (2018).
64. 64.
Marin-Palomo, P. et al. Microresonator-based solitons for massively parallel coherent optical communications. *Nature* **546**, 274–279 (2017).
65. 65.
Obrzud, E. et al. A microphotonic astrocomb. *Nat. Photon.* **13**, 31–35 (2019).
66. 66.
Suh, M.-G. et al. Searching for exoplanets using a microresonator astrocomb. *Nat. Photon.* **13**, 25–30 (2019).
67. 67.
Pfeiffer, M. H. P., Liu, J., Geiselmann, M. & Kippenberg, T. J. Coupling ideality of integrated planar high-Q microresonators. *Phys. Rev. Appl.* **7**, 024026 (2017).
68. 68.
Liu, J. et al. Double inverse nanotapers for efficient light coupling to integrated photonic devices. *Opt. Lett.* **43**, 3200–3203 (2018).
69. 69.
Liu, J. et al. Frequency-comb-assisted broadband precision spectroscopy with cascaded diode lasers. *Opt. Lett.* **41**, 3134–3137 (2016).

70. 70.

Feist, A. et al. Ultrafast transmission electron microscopy using a laser-driven field emitter: femtosecond resolution with a high coherence electron beam. *Ultramicroscopy* **176**, 63–73 (2017).

71. 71.

Piazza, L. et al. Design and implementation of a fs-resolved transmission electron microscope based on thermionic gun technology. *Chem. Phys.* **423**, 79–84 (2013).

72. 72.

Houdellier, F., Caruso, G., Weber, S., Kociak, M. & Arbouet, A. Development of a high brightness ultrafast transmission electron microscope based on a laser-driven cold field emission source. *Ultramicroscopy* **186**, 128–138 (2018).

73. 73.

Yurtsever, A. & Zewail, A. H. 4D nanoscale diffraction observed by convergent-beam ultrafast electron microscopy. *Science* **326**, 708–712 (2009).

74. 74.

Feist, A., Rubiano da Silva, N., Liang, W., Ropers, C. & Schäfer, S. Nanoscale diffractive probing of strain dynamics in ultrafast transmission electron microscopy. *Struct. Dynam.* **5**, 014302 (2018).

75. 75.

Schliep, K. B., Quarterman, P., Wang, J.-P. & Flannigan, D. J. Picosecond Fresnel transmission electron microscopy. *Appl. Phys. Lett.* **110**, 222404 (2017).

76. 76.

Rubiano da Silva, N. et al. Nanoscale mapping of ultrafast magnetization dynamics with femtosecond Lorentz microscopy. *Phys. Rev. X* **8**, 031052 (2018).

77. 77.

Cao, G., Jiang, S., Åkerman, J. & Weissenrieder, J. Femtosecond laser driven precessing magnetic gratings. *Nanoscale* **13**, 3746–3756 (2021).

78. 78.

Lagos, M. J., Trügler, A., Hohenester, U. & Batson, P. E. Mapping vibrational surface and bulk modes in a single nanocube. *Nature* **543**, 529–532 (2017).

79. 79.

Hachtel, J. A. et al. Identification of site-specific isotopic labels by vibrational spectroscopy in the electron microscope. *Science* **363**, 525–528 (2019).

80. 80.

Li, X. et al. Three-dimensional vectorial imaging of surface phonon polaritons. *Science* **371**, 1364–1367 (2021).

81. 81.

Kahl, F. et al. in *Advances in Imaging and Electron Physics* (eds Hawkes, P. & Hytch, M.) Vol. 212, 35–70 (Elsevier Inc., 2019).

Acknowledgements

All samples were fabricated in the Center of MicroNanoTechnology (CMi) at EPFL. This material is based upon work supported by the Air Force Office of Scientific Research under award number FA9550-19-1-0250. This work was further supported by the Swiss National Science Foundation under grant agreement numbers 185870 (Ambizione), 182103 and 176563 (BRIDGE). The experiments were conducted at the Göttingen UTEM Lab, funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through grant number 432680300/SFB 1456 (project C01) and the Gottfried Wilhelm Leibniz programme, and the European Union’s Horizon 2020 research and innovation programme under grant agreement number 101017720 (FET-Proactive EBEAM). Y.Y. acknowledges support from the EU H2020 research and innovation programme under the Marie Skłodowska-Curie IF grant agreement number 101033593 (SEPhIM). O.K. acknowledges the Max Planck Society for funding from the Manfred Eigen Fellowship for postdoctoral fellows from abroad. We thank H. Lourenço-Martins, I. Shomroni, J. Riemsberger and N. Johan Engelsen for useful discussions. We thank J. He for taking photographs of the photonic chips, J. Gaida for support with the TEM holder design and M. Sivis for making the render graphics.

Funding

Open access funding provided by Max Planck Society.

Author information

Author notes

1. These authors contributed equally: Jan-Wilke Henke, Arslan Sajid Raja

Affiliations

1. Georg-August-Universität Göttingen, Göttingen, Germany

Jan-Wilke Henke, Armin Feist, Germaine Arend, F. Jasmin Kappert, Marcel Möller, Ofer Kfir & Claus Ropers

2. Max Planck Institute of Multidisciplinary Sciences, Göttingen, Germany

Jan-Wilke Henke, Armin Feist, Germaine Arend, F. Jasmin Kappert, Marcel Möller, Ofer Kfir & Claus Ropers

3. Swiss Federal Institute of Technology Lausanne (EPFL), Lausanne, Switzerland

Arslan Sajid Raja, Guanhao Huang, Yujia Yang, Rui Ning Wang, Jiahe Pan, Junqiu Liu & Tobias J. Kippenberg

4. Center for Quantum Science and Engineering, EPFL, Lausanne, Switzerland

Guanhao Huang, Yujia Yang, Rui Ning Wang, Jiahe Pan & Tobias J. Kippenberg

Contributions

A.S.R. and J.L. designed the photonic chip devices. J.L. and R.N.W. developed the fabrication process and fabricated the devices. A.S.R. characterized and packaged the devices. J.-W.H., G.H., J.P. and Y.Y. performed numerical simulations. G.H., O.K. and C.R. devised the theory section. J.-W.H. designed the TEM sample mount, supported by A.F., and together with G.A. and F.J.K. built the optical set-up. J.-W.H. and A.F. carried out the TEM experiments and data analysis, supported by G.A., F.J.K. and M.M. The study was planned and directed by O.K., C.R. and T.J.K. The manuscript was written by J.-W.H., A.S.R., A.F., G.H., Y.Y., C.R. and T.J.K. after discussions with, and input from, all authors.

Corresponding authors

Correspondence to [Claus Ropers](#) or [Tobias J. Kippenberg](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Olav Solgaard and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

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

Supplementary information

Supplementary Information

This Supplementary Information file contains Sections 1–4, including Supplementary Figures 1–3 and additional references. Section 1: Analytical equations used for fitting optical resonance; Section 2: Quantum optical description of electron–photon interaction; Section 3: Estimation of transverse beam detection; Section 4: Integrated photonics platforms for electron beam modulation.

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

Henke, JW., Raja, A.S., Feist, A. *et al.* Integrated photonics enables continuous-beam electron phase modulation. *Nature* **600**, 653–658 (2021).
<https://doi.org/10.1038/s41586-021-04197-5>

- Received: 07 May 2021
- Accepted: 01 November 2021
- Published: 22 December 2021
- Issue Date: 23 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04197-5>

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

[Low-power light modifies electron microscopy](#)

- Martin Kozák

News & Views 22 Dec 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04197-5>

- Article
- [Published: 22 December 2021](#)

Dynamic spatial progression of isolated lithium during battery operations

- [Fang Liu](#) ORCID: [orcid.org/0000-0003-0885-5604¹](https://orcid.org/0000-0003-0885-5604),
- [Rong Xu¹](#),
- [Yecun Wu](#) ORCID: [orcid.org/0000-0001-6011-4489²](https://orcid.org/0000-0001-6011-4489),
- [David Thomas Boyle](#) ORCID: [orcid.org/0000-0002-0452-275X³](https://orcid.org/0000-0002-0452-275X),
- [Ankun Yang¹](#),
- [Jinwei Xu¹](#),
- [Yangying Zhu¹](#),
- [Yusheng Ye](#) ORCID: [orcid.org/0000-0001-9832-2478¹](https://orcid.org/0000-0001-9832-2478),
- [Zhiao Yu](#) ORCID: [orcid.org/0000-0001-8746-1640⁴](https://orcid.org/0000-0001-8746-1640),
- [Zewen Zhang¹](#),
- [Xin Xiao](#) ORCID: [orcid.org/0000-0003-1098-9484¹](https://orcid.org/0000-0003-1098-9484),
- [Wenxiao Huang¹](#),
- [Hansen Wang](#) ORCID: [orcid.org/0000-0002-6738-1659¹](https://orcid.org/0000-0002-6738-1659),
- [Hao Chen](#) ORCID: [orcid.org/0000-0002-2852-0070¹](https://orcid.org/0000-0002-2852-0070) &
- [Yi Cui](#) ORCID: [orcid.org/0000-0002-6103-6352^{1,5}](https://orcid.org/0000-0002-6103-6352)

[Nature](#) volume 600, pages 659–663 (2021)

- 5513 Accesses
- 129 Altmetric
- [Metrics details](#)

Subjects

- [Energy science and technology](#)
- [Materials science](#)
- [Nanoscience and technology](#)

Abstract

The increasing demand for next-generation energy storage systems necessitates the development of high-performance lithium batteries^{1,2,3}. Unfortunately, current Li anodes exhibit rapid capacity decay and a short cycle life^{4,5,6}, owing to the continuous generation of solid electrolyte interface^{7,8} and isolated Li (i-Li)^{9,10,11}. The formation of i-Li during the nonuniform dissolution of Li dendrites¹² leads to a substantial capacity loss in lithium batteries under most testing conditions¹³. Because i-Li loses electrical connection with the current collector, it has been considered electrochemically inactive or ‘dead’ in batteries^{14,15}. Contradicting this commonly accepted presumption, here we show that i-Li is highly responsive to battery operations, owing to its dynamic polarization to the electric field in the electrolyte. Simultaneous Li deposition and dissolution occurs on two ends of the i-Li, leading to its spatial progression toward the cathode (anode) during charge (discharge). Revealed by our simulation results, the progression rate of i-Li is mainly affected by its length, orientation and the applied current density. Moreover, we successfully demonstrate the recovery of i-Li in Cu–Li cells with >100% Coulombic efficiency and realize $\text{LiNi}_{0.5}\text{Mn}_{0.3}\text{Co}_{0.2}\text{O}_2$ (NMC)–Li full cells with extended cycle life.

This is a preview of subscription content

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Fig. 1: Dynamic polarization of i-Li under an electric field.

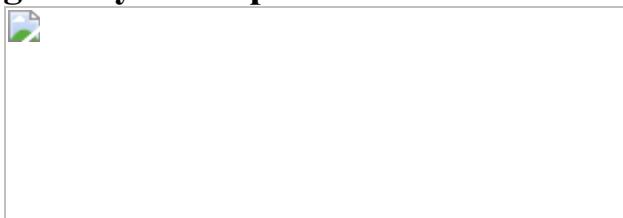


Fig. 2: Morphological evolution of the i-Li island.

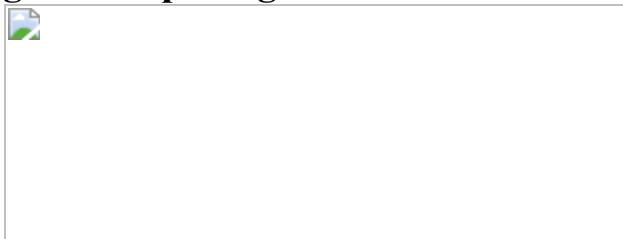


Fig. 3: Quantification of the overpotentials on i-Li in coin cells.

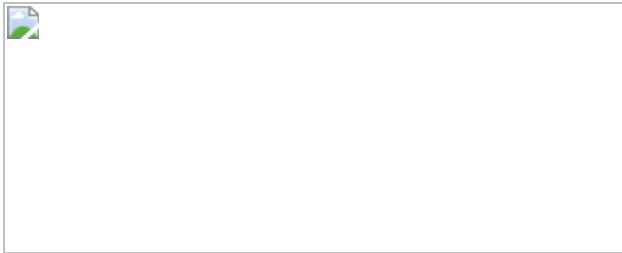
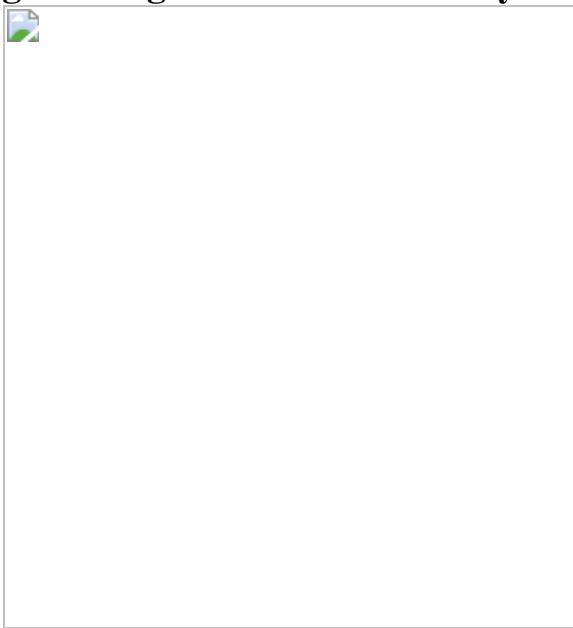


Fig. 4: Progression and recovery of i-Li in coin cells during discharge.



Data availability

All the data that support the findings of this study are available from the corresponding author upon request. [Source data](#) are provided with this paper.

References

1. 1.

Bruce, P. G., Freunberger, S. A., Hardwick, L. J. & Tarascon, J.-M. Li–O₂ and Li–S batteries with high energy storage. *Nat. Mater.* **11**, 19–29 (2012); erratum *Nat. Mater.* **11**, 172 (2012).

2. 2.

Lin, D., Liu, Y. & Cui, Y. Reviving the lithium metal anode for high-energy batteries. *Nat. Nanotechnol.* **12**, 194–206 (2017).

3. 3.

Albertus, P., Babinec, S., Litzelman, S. & Newman, A. Status and challenges in enabling the lithium metal electrode for high-energy and low-cost rechargeable batteries. *Nat. Energy* **3**, 16–21 (2018).

4. 4.

Liu, J. et al. Pathways for practical high-energy long-cycling lithium metal batteries. *Nat. Energy* **4**, 180–186 (2019).

5. 5.

Cheng, X.-B., Zhang, R., Zhao, C.-Z. & Zhang, Q. Toward safe lithium metal anode in rechargeable batteries: a review. *Chem. Rev.* **117**, 10403–10473 (2017).

6. 6.

Xu, W. et al. Lithium metal anodes for rechargeable batteries. *Energy Environ. Sci.* **7**, 513–537 (2014).

7. 7.

Lu, D. et al. Failure mechanism for fast-charged lithium metal batteries with liquid electrolytes. *Adv. Energy Mater.* **5**, 1400993 (2015).

8. 8.

Niu, C. et al. High-energy lithium metal pouch cells with limited anode swelling and long stable cycles. *Nat. Energy* **4**, 551–559 (2019)

9. 9.

Yoshimatsu, I., Hirai, T. & Yamaki, J.-i. Lithium electrode morphology during cycling in lithium cells. *J. Electrochem. Soc.* **135**, 2422–2427

(1988).

10. 10.

Sanchez, A. J. et al. Plan-view operando video microscopy of Li metal anodes: identifying the coupled relationships among nucleation, morphology, and reversibility. *ACS Energy Lett.* **5**, 994–1004 (2020).

11. 11.

Gunnarsdóttir, A. B., Amanchukwu, C. V., Menkin, S. & Grey, C. P. Noninvasive in situ NMR study of “dead lithium” formation and lithium corrosion in full-cell lithium metal batteries. *J. Am. Chem. Soc.* **142**, 20814–20827 (2020).

12. 12.

Li, Y. et al. Correlating structure and function of battery interphases at atomic resolution using cryoelectron microscopy. *Joule* **2**, 2167–2177 (2018).

13. 13.

Fang, C. et al. Quantifying inactive lithium in lithium metal batteries. *Nature* **572**, 511–515 (2019).

14. 14.

Chen, K.-H. et al. Dead lithium: mass transport effects on voltage, capacity, and failure of lithium metal anodes. *J. Mater. Chem. A* **5**, 11671–11681 (2017).

15. 15.

Xu, S., Chen, K.-H., Dasgupta, N. P., Siegel, J. B. & Stefanopoulou, A. G. Evolution of dead lithium growth in lithium metal batteries: experimentally validated model of the apparent capacity loss. *J. Electrochem. Soc.* **166**, A3456–A3463 (2019).

16. 16.

Jin, C. et al. Rejuvenating dead lithium supply in lithium metal anodes by iodine redox. *Nat. Energy* **6**, 378–387 (2021).

17. 17.

Yu, Z. et al. Molecular design for electrolyte solvents enabling energy-dense and long-cycling lithium metal batteries. *Nat. Energy* **5**, 526–533 (2020).

18. 18.

Zeng, Z. et al. Non-flammable electrolytes with high salt-to-solvent ratios for Li-ion and Li-metal batteries. *Nat. Energy* **3**, 674–681 (2018).

19. 19.

Gao, Y. et al. Polymer–inorganic solid–electrolyte interphase for stable lithium metal batteries under lean electrolyte conditions. *Nat. Mater.* **18**, 384–389 (2019).

20. 20.

Weng, Y.-T. et al. An ultrathin ionomer interphase for high-efficiency lithium anode in carbonate-based electrolyte. *Nat. Commun.* **10**, 5824 (2019).

21. 21.

Lin, D. et al. Conformal lithium fluoride protection layer on three-dimensional lithium by nonhazardous gaseous reagent freon. *Nano Lett.* **17**, 3731–3737 (2017).

22. 22.

Lin, D. et al. Layered reduced graphene oxide with nanoscale interlayer gaps as a stable host for lithium metal anodes. *Nat.*

Nanotechnol. **11**, 626–632 (2016).

23. 23.

Lin, D. et al. Three-dimensional stable lithium metal anode with nanoscale lithium islands embedded in ionically conductive solid matrix. *Proc. Natl Acad. Sci. USA* **114**, 4613–4618 (2017).

24. 24.

Chen, H. et al. Tortuosity effects in lithium-metal host anodes. *Joule* **4**, 938–952 (2020).

25. 25.

Mao, C., Ruther, R. E., Li, J., Du, Z. & Belharouak, I. Identifying the limiting electrode in lithium ion batteries for extreme fast charging. *Electrochem. Commun.* **97**, 37–41 (2018).

26. 26.

Belov, D. & Yang, M.-H. Investigation of the kinetic mechanism in overcharge process for Li-ion battery. *Solid State Ionics* **179**, 1816–1821 (2008).

27. 27.

Doyle, M., Fuller, T. F. & Newman, J. Modeling of galvanostatic charge and discharge of the lithium/polymer/insertion cell. *J. Electrochem. Soc.* **140**, 1526–1533 (1993).

28. 28.

Fuller, T. F., Doyle, M. & Newman, J. Simulation and optimization of the dual lithium ion insertion cell. *J. Electrochem. Soc.* **141**, 1–10 (1994).

29. 29.

Zheng, J. et al. Highly stable operation of lithium metal batteries enabled by the formation of a transient high-concentration electrolyte layer. *Adv. Energy Mater.* **6**, 1502151 (2016).

Acknowledgements

Y.C. acknowledges support from the Assistant Secretary for Energy Efficiency and Renewable Energy, Office of Vehicle Technologies of the US Department of Energy, under the Battery Materials Research (BMR), Battery 500 Consortium, and the eXtreme Fast Charge Cell Evaluation of Li-ion batteries (XCEL) programmes.

Author information

Affiliations

1. Department of Materials Science and Engineering, Stanford University, Stanford, CA, USA

Fang Liu, Rong Xu, Ankun Yang, Jinwei Xu, Yangying Zhu, Yusheng Ye, Zewen Zhang, Xin Xiao, Wenxiao Huang, Hansen Wang, Hao Chen & Yi Cui

2. Department of Electrical Engineering, Stanford University, Stanford, CA, USA

Yecun Wu

3. Department of Chemistry, Stanford University, Stanford, CA, USA

David Thomas Boyle

4. Department of Chemical Engineering, Stanford University, Stanford, CA, USA

Zhiao Yu

5. Stanford Institute for Materials and Energy Sciences, SLAC National Accelerator Laboratory, Menlo Park, CA, USA

Yi Cui

Contributions

F.L. and Y.C. conceived the idea and designed the experiments. F.L. conducted the electrochemical characterizations. R.X. conducted numerical simulations. Y.W. and A.Y. helped with the optical-cell fabrication. J.X. and Y.Z. designed and customized the multi-electrode cells. Z.Y. and H.C. prepared the electrolyte. Z.Z. performed SEM characterizations. D.T.B., Y.Y., W.H., H.W. and X.X. helped with the cell fabrication and electrochemical characterizations. F.L. and Y.C. co-wrote the paper. All authors discussed the results and commented on the manuscript.

Corresponding author

Correspondence to [Yi Cui](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Peng Bai and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

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 Configuration of optical cells with i-Li island between electrodes.

a, Experimental setup. **b**, Optical image of the deposited lithium on the Cu island. The scale bar is 100 μm .

Extended Data Fig. 2 Compositional evolution on i-Li island during charge.

a, b, X-ray photoelectron spectroscopy analysis of the ends of i-Li island near Li anode (**a**) and NMC cathode (**b**) after 3 h charging.

[Source data](#)

Extended Data Fig. 3 Morphological evolution of ‘dead Li’ during charge.

Optical images showing the evolution process of micrometre-sized, electrochemically generated ‘dead Li’ filaments under a charging current of 20 μA . The scale bar is 20 μm .

Extended Data Fig. 4 Potential distributions in the electrolyte during battery operations.

a, b, The potential distribution in the electrolyte during charge (**a**) and discharge (**b**). The direction of Li^+ flux is shown in black arrows. The initial and end states are at $t = 100$ s and 2 h, respectively.

Extended Data Fig. 5 The dynamic response and spatial progression of i-Li during battery operations.

a, b, The absolute overpotentials at two ends of i-Li with different width (**a**) and length (**b**) during charge. **c**, The maximum current densities at the ends

of i-Li filaments and Li electrode under various current densities during discharge. **d**, The migration distance of i-Li with different orientations at the end of 1C discharge ($1\text{C} = 3 \text{ mA cm}^{-2}$). The orientation of i-Li is defined as the angle between i-Li and the electric field. 0° and 90° represent the directions along and perpendicular to the electric field, respectively.

[Source data](#)

Extended Data Fig. 6 Progression of i-Li in coin cells during discharge.

a, Voltage–capacity profiles of NMC–G cells with different cell configurations. NMC–polyimide (PI)–G cells with/without Cu present identical voltage profiles, indicating that the addition of Cu does not influence the electrochemical performance of NMC–G cell. Red arrow marks the voltage fluctuation observed in NMC–PI–G cell with Cu/Li due to i-Li’s progression. Meanwhile, the NMC–polypropylene–G cell with i-Li (Cu/Li) exhibits a smooth voltage profile, suggesting that i-Li could not penetrate through the nano-sized pores of commercial polypropylene (PP) separators. **b**, **c**, SEM images of Cu/Li (bottom) and polyimide membrane (top) after 5-min (**b**) and 1-h charge (**c**), showing that i-Li penetrates through the polyimide membrane. Arrows point out the Li dendrites on the edge of polyimide membrane. The scale bars are $10 \mu\text{m}$.

[Source data](#)

Extended Data Fig. 7 The recovery of ‘dead Li;’ during discharge.

a, A representative voltage profile of Cu–Li cell during i-Li formation cycle. **b**, Capacity of ‘dead Li’ on average. Each coloured sphere represents the CE of one Cu–Li cell, the box shows the average CE of ten cells, and the error bar illustrates 1 standard deviation. **c**, The recovery percentage of ‘dead Li’ under different stripping currents.

[Source data](#)

Extended Data Fig. 8 Coulombic efficiency measurements of Cu–Li half-cells with/without activation in the presence of ‘dead Li’.

a–d, The stripping conditions are 0.5 mA cm^{-2} with an activation step of **a**, none, **b**, 3 mA cm^{-2} for 1 min, **c**, 3 mA cm^{-2} for 2 mins, and **d**, 6 mA cm^{-2} for 1 min. For the average CE measurement, a standard protocol is followed: (1) 5 mAh cm^{-2} of Li is deposited onto Cu under 0.5 mA cm^{-2} as a Li reservoir, (2) repeatedly deposit 1 mAh cm^{-2} of Li (0.5 mA cm^{-2}) and strip under different conditions for ten cycles, (3) strip all active Li (0.5 mA cm^{-2}) until 1 V. Accumulated ‘dead Li’ are formed by 5 cycles of Li deposition (3 mA cm^{-2} , 1 h) and stripping (0.5 mA cm^{-2} until 1 V).

[Source data](#)

Extended Data Fig. 9 Electrochemical performance of NMC–Li cells with/without activation.

a, Coulombic efficiency and **b**, cycle life of NMC–Li cells (150%-excessed Li) with/without activation. **c**, Specific capacity and **d**, cycle life of NMC–Li cells (300%-excessed Li) under a fast-charging condition. Each coloured sphere represents the cycle life of one NMC–Li cell, the box shows the average cycle life of three cells, and the error bar illustrates 1 standard deviation. Cycle life is defined as the cycle number when the cell capacity falls below 80% of its initial capacity. TF stands for temperature fluctuation.

[Source data](#)

Extended Data Table 1 Parameters used in the numerical modelling

Source data

[Source Data Fig. 2](#)

[**Source Data Fig. 3**](#)

[**Source Data Fig. 4**](#)

[**Source Data Extended Data Fig. 2**](#)

[**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**](#)

Rights and permissions

[**Reprints and Permissions**](#)

About this article

Cite this article

Liu, F., Xu, R., Wu, Y. *et al.* Dynamic spatial progression of isolated lithium during battery operations. *Nature* **600**, 659–663 (2021).
<https://doi.org/10.1038/s41586-021-04168-w>

- Received: 04 November 2020
- Accepted: 20 October 2021
- Published: 22 December 2021
- Issue Date: 23 December 2021

- DOI: <https://doi.org/10.1038/s41586-021-04168-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

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04168-w>

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

- Article
- [Published: 22 December 2021](#)

High-entropy polymer produces a giant electrocaloric effect at low fields

- [Xiaoshi Qian](#) ORCID: orcid.org/0000-0003-3487-7567^{1,2 na1},
- [Donglin Han](#) ORCID: orcid.org/0000-0003-1897-3727^{1 na1},
- [Lirong Zheng](#) ORCID: orcid.org/0000-0001-6803-5048^{3 na1},
- [Jie Chen](#) ORCID: orcid.org/0000-0002-6134-6246^{4 na1},
- [Madhusudan Tyagi](#)^{5,6},
- [Qiang Li](#)¹,
- [Feihong Du](#)¹,
- [Shanyu Zheng](#)¹,
- [Xingyi Huang](#) ORCID: orcid.org/0000-0002-8919-6884⁴,
- [Shihai Zhang](#)⁷,
- [Junye Shi](#)¹,
- [Houbing Huang](#) ⁸,
- [Xiaoming Shi](#)⁸,
- [Jiangping Chen](#)¹,
- [Hancheng Qin](#) ORCID: orcid.org/0000-0003-0146-8206⁹,
- [Jerzy Bernholc](#) ORCID: orcid.org/0000-0002-9981-8851⁹,
- [Xin Chen](#)¹⁰,
- [Long-Qing Chen](#) ORCID: orcid.org/0000-0003-3359-3781¹⁰,
- [Liang Hong](#) ORCID: orcid.org/0000-0003-0107-336X³ &
- [Q. M. Zhang](#)¹¹

[Nature](#) volume 600, pages 664–669 (2021)

- 4682 Accesses
- 20 Altmetric
- [Metrics details](#)

Subjects

- [Polymers](#)
- [Structure of solids and liquids](#)
- [Thermodynamics](#)
- [Thermoelectric devices and materials](#)

Abstract

More than a decade of research on the electrocaloric (EC) effect has resulted in EC materials and EC multilayer chips that satisfy a minimum EC temperature change of 5 K required for caloric heat pumps^{1,2,3}. However, these EC temperature changes are generated through the application of high electric fields^{4,5,6,7,8} (close to their dielectric breakdown strengths), which result in rapid degradation and fatigue of EC performance. Here we report a class of EC polymer that exhibits an EC entropy change of $37.5 \text{ J kg}^{-1} \text{ K}^{-1}$ and a temperature change of 7.5 K under 50 MV m^{-1} , a 275% enhancement over the state-of-the-art EC polymers under the same field strength. We show that converting a small number of the chlorofluoroethylene groups in poly(vinylidene fluoride-trifluoroethylene-chlorofluoroethylene) terpolymer into covalent double bonds markedly increases the number of the polar entities and enhances the polar–nonpolar interfacial areas of the polymer. The polar phases in the polymer adopt a loosely correlated, high-entropy state with a low energy barrier for electric-field-induced switching. The polymer maintains performance for more than one million cycles at the low fields necessary for practical EC cooling applications, suggesting that this strategy may yield materials suitable for use in caloric heat pumps.

This is a preview of subscription content

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Fig. 1: Entropy-enhanced polymers exhibit giant ECEs under low fields.

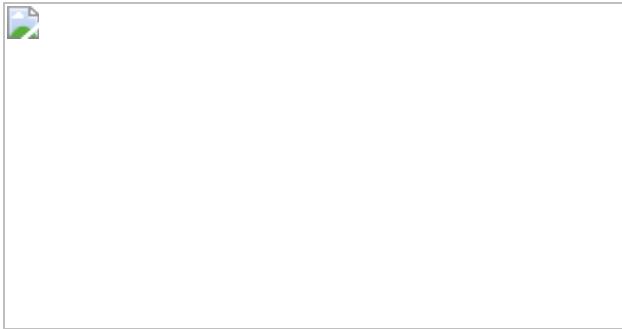


Fig. 2: Structural properties of the modified EC polymers and their base terpolymer.

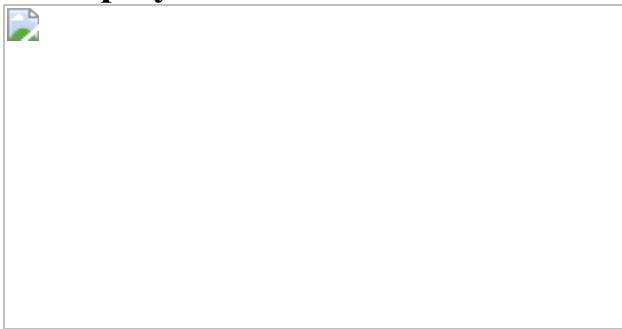


Fig. 3: Dielectric properties of the modified EC polymer and its base terpolymer.

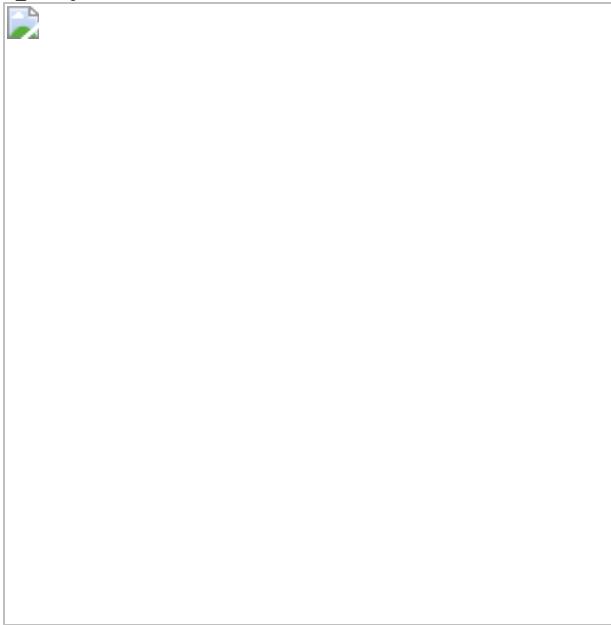
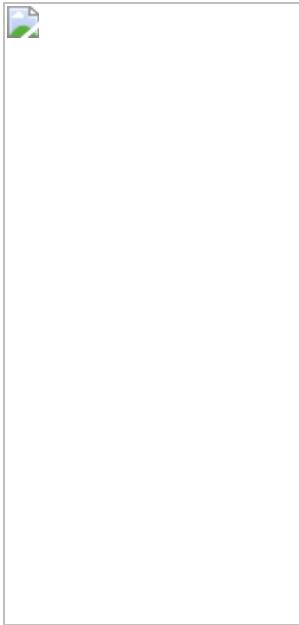


Fig. 4: High-entropy EC polymer and its impact on device performances.



Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding authors on reasonable request.

Code availability

Access to the phase-field model and DFT model codes are available on request from H.H. (hbhuang@bit.edu.cn) and J.B. (bernholc@ncsu.edu), respectively.

References

1. 1.

Qian, S. et al. Not-in-kind cooling technologies: a quantitative comparison of refrigerants and system performance. *Int. J. Refrig.* **62**, 177–192 (2016).

2. 2.

Shi, J. et al. Electrocaloric cooling materials and devices for zero-global-warming-potential, high-efficiency refrigeration. *Joule* **3**, 1200–1225 (2019).

3. 3.

Nair, B. et al. Large electrocaloric effects in oxide multilayer capacitors over a wide temperature range. *Nature* **575**, 468–472 (2019).

4. 4.

Gu, H. et al. A chip scale electrocaloric effect based cooling device. *Appl. Phys. Lett.* **102**, 122904 (2013).

5. 5.

Ma, R. et al. Highly efficient electrocaloric cooling with electrostatic actuation. *Science* **357**, 1130–1134 (2017).

6. 6.

Torelló, A. et al. Giant temperature span in electrocaloric regenerator. *Science* **370**, 125–129 (2020).

7. 7.

Wang, Y. et al. A high-performance solid-state electrocaloric cooling system. *Science* **370**, 129–133 (2020).

8. 8.

Meng, Y. et al. A cascade electrocaloric cooling device for large temperature lift. *Nat. Energy* **5**, 996–1002 (2020).

9. 9.

Hawken, P. *Drawdown: the Most Comprehensive Plan Ever Proposed to Reverse Global Warming* 1st edn (Penguin, 2017).

10. 10.

Kitanovski, A. et al. in *Magnetocaloric Energy Conversion: From Theory to Applications* (eds Kitanovski, A. et al.) 97–166 (Springer, 2015).

11. 11.

Tušek, J. et al. A regenerative elastocaloric heat pump. *Nat. Energy* **1**, 16134 (2016).

12. 12.

Li, B. et al. Colossal barocaloric effects in plastic crystals. *Nature* **567**, 506–510 (2019).

13. 13.

Wang, R. et al. Torsional refrigeration by twisted, coiled, and supercoiled fibers. *Science* **366**, 216–221 (2019).

14. 14.

Zhang, T., Qian, X.-S., Gu, H., Hou, Y. & Zhang, Q. M. An electrocaloric refrigerator with direct solid to solid regeneration. *Appl. Phys. Lett.* **110**, 243503 (2017).

15. 15.

Defay, E. et al. Enhanced electrocaloric efficiency via energy recovery. *Nat. Commun.* **9**, 1827 (2018).

16. 16.

Chen, Y. et al. An all-scale hierarchical architecture induces colossal room-temperature electrocaloric effect at ultralow electric field in polymer nanocomposites. *Adv. Mater.* **32**, 1907927 (2020).

17. 17.

Neese, B. et al. Large electrocaloric effect in ferroelectric polymers near room temperature. *Science* **321**, 821–823 (2008).

18. 18.

Li, X. et al. Tunable temperature dependence of electrocaloric effect in ferroelectric relaxor poly(vinylidene fluoride-trifluoroethylene-chlorofluoroethylene terpolymer. *Appl. Phys. Lett.* **99**, 052907 (2011).

19. 19.

Zhang, G. et al. Ferroelectric polymer nanocomposites for room-temperature electrocaloric refrigeration. *Adv. Mater.* **27**, 1450–1454 (2015).

20. 20.

Klein, R. J., Xia, F., Zhang, Q. M. & Bauer, F. Influence of composition on relaxor ferroelectric and electromechanical properties of poly(vinylidene fluoride-trifluoroethylene- chlorofluoroethylene). *J. Appl. Phys.* **97**, 094105 (2005).

21. 21.

Zhu, G. et al. Self-similar mesocrystals form via interface-driven nucleation and assembly. *Nature* **590**, 416–422 (2021).

22. 22.

Bai, G., Gao, D., Liu, Z., Zhou, X. & Wang, J. Probing the critical nucleus size for ice formation with graphene oxide nanosheets. *Nature* **576**, 437–441 (2019).

23. 23.

Meiners, T., Frolov, T., Rudd, R. E., Dehm, G. & Liebscher, C. H. Observations of grain-boundary phase transformations in an elemental metal. *Nature* **579**, 375–378 (2020).

24. 24.

Borgia, A. et al. Extreme disorder in an ultrahigh-affinity protein complex. *Nature* **555**, 61–66 (2018).

25. 25.

Jaffe, B., Cook, W. R. & Jaffe, H. L. *Piezoelectric Ceramics* (Academic, 1971).

26. 26.

Katsouras, I. et al. The negative piezoelectric effect of the ferroelectric polymer poly(vinylidene fluoride). *Nat. Mater.* **15**, 78–84 (2016).

27. 27.

Tan, S., Li, J., Gao, G., Li, H. & Zhang, Z. Synthesis of fluoropolymer containing tunable unsaturation by a controlled dehydrochlorination of P(VDF-co-CTFE) and its curing for high performance rubber applications. *J. Mater. Chem.* **22**, 18496–18504 (2012).

28. 28.

Le Goupil, F. et al. Enhanced electrocaloric response of vinylidene fluoride-based polymers via one-step molecular engineering. *Adv. Funct. Mater.* **31**, 2007043 (2021).

29. 29.

Kessler, H. Detection of hindered rotation and inversion by NMR spectroscopy. *Angew. Chem. Int. Ed. Engl.* **9**, 219–235 (1970).

30. 30.

Hong, L. et al. Elastic and conformational softness of a globular protein. *Phys. Rev. Lett.* **110**, 028104 (2013).

31. 31.

Pirc, R., Kutnjak, Z., Blinc, R. & Zhang, Q. M. Upper bounds on the electrocaloric effect in polar solids. *Appl. Phys. Lett.* **98**, 021909 (2011).

32. 32.

Yang, L. et al. Novel polymer ferroelectric behavior via crystal isomorphism and the nanoconfinement effect. *Polymer* **54**, 1709–1728 (2013).

33. 33.

Yang, L. et al. Graphene enabled percolative nanocomposites with large electrocaloric efficient under low electric fields over a broad temperature range. *Nano Energy* **22**, 461–467 (2016).

34. 34.

Meyer, A., Dimeo, R. M., Gehring, P. M. & Neumann, D. A. The high-flux backscattering spectrometer at the NIST Center for Neutron Research. *Rev. Sci. Instrum.* **74**, 2759–2777 (2003).

35. 35.

Klug, H. P. & Alexander, L. E. *X-ray Diffraction Procedures for Polycrystalline and Amorphous Materials* 2nd edn (Wiley, 1974).

36. 36.

Tosaka, M. et al. Orientation and crystallization of natural rubber network as revealed by WAXD using synchrotron radiation. *Macromolecules* **37**, 3299–3309 (2004).

37. 37.

Ranjan, V., Nardelli, M. B. & Bernholc, J. Electric field induced phase transitions in polymers: a novel mechanism for high speed energy storage. *Phys. Rev. Lett.* **108**, 087802 (2012).

38. 38.

Chen, X. Z. et al. A nanocomposite approach to tailor electrocaloric effect in ferroelectric polymer. *Polymer* **54**, 5299–5302 (2013).

39. 39.

Lu, Y.-C. et al. Enhanced electrocaloric effect for refrigeration in lead-free polymer composite films with an optimal filler loading. *Appl. Phys. Lett.* **114**, 233901 (2019).

Acknowledgements

This work was supported by the National Key R&D Program of China (2020YFA0711500 and 2020YFA0711503), the National Natural Science Foundation of China (52076127, 51776119, 11974239, 31630002, 51877132, 52003153 and 51972028), the Natural Science Foundation of Shanghai (20ZR1471700) and the China Postdoctoral Science Foundation-funded project 2019M661479. X.Q. acknowledges the support by the Prospective Research Program at Shanghai Jiao Tong University (19X160010008), the Oceanic Interdisciplinary Program of Shanghai Jiao Tong University (project number SL2020MS009), the Student Innovation Center and the Instrumental Analysis Center at Shanghai Jiao Tong University. X.C. and Q.M.Z. acknowledge the support of a seed grant from the Pennsylvania State University MRI-IEE, USA. L.-Q.C. acknowledges the support by the Donald W. Hamer Foundation through the Hamer professorship at Penn State. We thank L. Jiang and B. Zhu for assistance with measurements of energy-dispersive X-ray spectroscopy and attenuated total reflection Fourier-transform infrared spectroscopy. Access to the HFBS was provided by the Center for High Resolution Neutron Scattering, a partnership between the National Institute of Standards and Technology and the National Science Foundation under Agreement No. DMR-1508249. The DFT calculations were performed at the Oak Ridge Leadership Computing Facility, supported by DOE contract DE-AC05-00OR22725. We thank N. Li from the BL19U2 beamline, X. Miu from the BL16B1 beamline and Y. Yang at the BL14B beamline of Shanghai Synchrotron Radiation Facility for the help with synchrotron X-ray measurements. L.H.

thanks the Innovation Program of the Shanghai Municipal Education Commission. H.Q. and J.B. thank B. Zhang and W. Lu for extensive discussions. X.Q. thanks S. Lin for support with simulation software, and X. Yao, Q. Wang and G. Meng for inspiring discussions. Certain commercial material suppliers are identified in this paper to foster understanding. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

Author information

Author notes

1. These authors contributed equally: Xiaoshi Qian, Donglin Han, Lirong Zheng, Jie Chen

Affiliations

1. Institute of Refrigeration and Cryogenics, Interdisciplinary Research Centre for Metamaterials and Intelligent Systems, State Key Laboratory of Mechanical System and Vibration, and MOE Key Laboratory for Power Machinery and Engineering, School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, China

Xiaoshi Qian, Donglin Han, Qiang Li, Feihong Du, Shanyu Zheng, Junye Shi & Jiangping Chen

2. Shanghai Jiao Tong University Jiangsu ZhongGuanCun Research Institute, Liyang, Jiangsu, China

Xiaoshi Qian

3. School of Physics and Astronomy, Institute of Natural Sciences, Shanghai National Center for Applied Mathematics (SJTU Center) and MOE-LSC, Shanghai Jiao Tong University, Shanghai, China

Lirong Zheng & Liang Hong

4. Frontiers Science Center for Transformative Molecules, Department of Polymer Science and Engineering, Shanghai Key Laboratory of Electrical Insulation and Thermal Ageing, State Key Laboratory of Metal Matrix Composites, Shanghai Jiao Tong University, Shanghai, China

Jie Chen & Xingyi Huang

5. NIST Center for Neutron Research, National Institute of Standards and Technology (NIST), Gaithersburg, MD, USA

Madhusudan Tyagi

6. Department of Materials Science and Engineering, University of Maryland, College Park, MD, USA

Madhusudan Tyagi

7. PolyK Technologies, State College, PA, USA

Shihai Zhang

8. School of Materials Science and Engineering and Advanced Research Institute of Multidisciplinary Science, Beijing Institute of Technology, Beijing, China

Houbing Huang & Xiaoming Shi

9. Department of Physics, North Carolina State University, Raleigh, NC, USA

Hancheng Qin & Jerzy Bernholc

10. Department of Materials Science and Engineering, Pennsylvania State University, University Park, PA, USA

Xin Chen & Long-Qing Chen

11. School of Electrical Engineering and Computer Science and Department of Mechanical Engineering, The Pennsylvania State University, University Park, PA, USA

Q. M. Zhang

Contributions

X.Q. conceived the concept, designed the experiment and wrote the manuscript. X.Q., D.H., Jie Chen, Q.L., F.D., X.H., S. Zheng and S. Zhang carried out the material synthesis and characterization. L.Z., D.H., Q.L., L.H. and X.Q. conducted the synchrotron X-ray measurements. L.Z. and M.T. conducted the neutron scattering tests. H.H. and X.S. carried out the phase-field simulation with guidance from L.-Q.C. H.Q. and J.B. carried out the DFT calculation. X.C. conducted the molecular mechanics simulation. X.Q. and D.H. designed the model for cooling devices. J.S. and Jiangping Chen supervised the device modelling. L.H. supervised synchrotron X-ray and neutron scattering tests. X.Q. and Q.M.Z. supervised the project. All authors analysed and interpreted the data.

Corresponding authors

Correspondence to [Xiaoshi Qian](#), [Xingyi Huang](#), [Houbing Huang](#) or [Liang Hong](#).

Ethics declarations

Competing interests

X.Q. and S. Zheng are inventors on a provisional patent application related to the described work. The other authors declare no competing interests.

Peer review information

Nature thanks Brahim Dkhil 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 Weibull plots of EC polymers with various -CH = CF- bond contents.

The breakdown strength was characterized at Shanghai Juter High Voltage Electrical & Equipment Co. Ltd. under a DC voltage at a rate of 200 V s^{-1} . The electrode diameter was 1 mm and the film thickness was $6 \mu\text{m}$. After introducing 0.4 mol% of the double bonds, the Weibull breakdown strength increased from 350 to 395 MV m^{-1} . As the -CH = CF- bond content increased to 0.6%, 0.8%, and 2.0%, the Weibull breakdown strength decreased to 390, 375, and 370 MV m^{-1} . As a result, the maximum cycling field can be extended to 80 MV m^{-1} for TD-0.6%. 50 MV m^{-1} is about 14% of the E_b of the polymer.

Extended Data Fig. 2 Reduction of ECE in EC polymer MLC due to inactive materials.

a, Simulated temperature changes of the MLC with and without the epoxy ($0.5 \mu\text{m}$ thick) and un-electrode margin (with $1 \text{ cm} \times 2 \text{ cm}$ total area and un-electroded margin = 0.5 mm on each side) as inactive materials. b, Schematics of the numerical model composing of 9 layers of EC polymer and 10 layer of epoxy. The electrode thickness for the polymer is below $0.05 \mu\text{m}$ (negligible), and we considered the epoxy layer that mechanically stabilize the interfaces between each layers of MLCs. Hence, the total passive part in the MLC will be less than 15%, causing about 15% reduction of ΔT (and ΔS).

Extended Data Fig. 3 Field-dependent EC-induced entropy changes of the base terpolymer and the modified terpolymer

with different mole percentages of CFE reduction.

EC-induced entropy changes of the base terpolymer and the modified terpolymers with different mole percentages of CFE reduction (*i.e.*, DB concentration ranges from 0, 0.4%, 0.6% and 2.0%). ($n \geq 3$, points are centred on the mean and the bars indicate the s.d., each sample of TD-0.6% in the EC heat flux measurement was verified by independent measurements by an infrared camera, see Supplementary Sections 1.6 and 1.7).

Extended Data Fig. 4 Heat flow signals of EC polymers during the heating and cooling process from the differential scanning calorimetry.

a, base terpolymer, b, TD-0.4%, c, TD-0.6% and d, TD-2.0%.

Extended Data Fig. 5 Rotation energy of carbon-carbon bond in PVDF single chains before (black) and after (red) a CFE group was replaced by a covalent double bond.

a, Rotation energy of the C-C bond adjacent to the Cl atom (on the left of the DB after the CFE reduction), and b, Rotation energy of the C-C bond on the other side of the CFE (DB) group. The simulation was carried out by MM2 force field simulation through ChemOffice (version 19.0). The chain was defined as 5VDF + 1CFE + 5VDF and was relaxed in TGTG' conformation under zero kelvin. A covalent double bond was in-situ introduced by cancel a HCl of the CFE.

Extended Data Fig. 6 Comparison of the temperature-dependent permittivity between the extrinsic blends and intrinsically modified terpolymer.

The blends were composed of terpolymer P(VDF-TrFE-CFE) and copolymer P(VDF-TrFE) 65/35 mol%. The ferroelectric copolymer was intentionally added to generate a larger ECE by inducing internal fields and

providing more polar entities. The composition of blends is noted as TC followed by the percentage of the terpolymer in the blends, *e.g.*, TC90 refers to a polymer blend with 90 mol% terpolymer in the matrix. As the composition of ferroelectric copolymer increases, the dielectric material exhibits higher transition temperatures and stronger polar correlations as shown by the slope of blue arrows, compared with those of the TD series.

Extended Data Fig. 7 Relation between the maximum polarization and the electrocaloric effect of polar materials.

a, Schematic of microscopic ferroelectric-domain switching induced polarization change which does not generate electrocaloric effect. ECE is directly related to the electric field induced polarization change. In ferroelectrics and dielectric in general, the efficiency of inducing electrocaloric entropy change from polarization change can vary over a great range. As an example, macroscopic polarization domain switching, as illustrated here, although generates large polarization change, does not contribute to electrocaloric effect. b, Summarized P-E loops for the base terpolymer and TD series under 120 MVm^{-1} and 10 Hz. c, Ratio of the β coefficient of modified polymeric EC materials to their respective base polymers^{16,28,38,39}, the inset presents values of the β coefficient of the enhanced EC polymeric materials reported in the literature and this work. As shown by Pirc *et al*³¹, $\Delta S = P^2 \ln(\Omega)/(\epsilon_0 \Theta)$, where Ω is the number of polar entities (number of polar-states) which can be accessed by dipoles, ϵ_0 is vacuum permittivity, and Θ is an effective Curie constant, which is directly related to the polar correction in a ferroelectric material. Our results reveal that a large ECE can be achieved (at low fields) by engineering a high-entropy dielectric material which possesses a large β coefficient, and hence a large ECE may not require a material to have a large polarization.

Extended Data Fig. 8 Fatigue behaviour of the base terpolymer P(VDF-TrFE-CFE).

a, In-situ heat flux directly recorded during cycling under 80 MVm^{-1} . b, Initially measured EC entropy change compared to that after 61200 cycles under different electric fields.

Extended Data Fig. 9 Comparison of the EC fatigue performances when the base terpolymer and the TD-0.6% were generating the same amount of ECE.

a, Cyclic performance of TD-0.6% before and after 10^6 cycles, at 50 MVm^{-1} . The integrated peak areas, which represent the total heating and cooling signals, were similar to each other. b, The base terpolymer film experienced a series of breakdown after 50 cycles under 150 MVm^{-1} .

Extended Data Fig. 10 Comparison of hysteresis losses between the base terpolymer and the TD-0.6% high entropy polymer when they were generating the same amount of ECE.

a, Integrated hysteresis loss in percentage. The inset presents PE loops of the base terpolymer under 100 MVm^{-1} and TD-0.6% under 50 MVm^{-1} , at which the two polymers induced similar ECE. b, Hysteresis heat loss for the base terpolymer and TD-0.6% when they induce the same amount of entropy change of 40 and $70 \text{ Jkg}^{-1}\text{K}^{-1}$, respectively.

Supplementary information

Supplementary Information

This file contains four sections and additional references, and includes Supplementary Figs. 1–28 and Tables 1–7. Section 1. Characterization of EC polymers; Section 2. Phase-field simulation; Section 3. DFT calculation of energy barriers in phase transitions; Section 4. Simulation of EC heat pumps using EC polymers.

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Qian, X., Han, D., Zheng, L. *et al.* High-entropy polymer produces a giant electrocaloric effect at low fields. *Nature* **600**, 664–669 (2021).
<https://doi.org/10.1038/s41586-021-04189-5>

- Received: 16 May 2021
- Accepted: 26 October 2021
- Published: 22 December 2021
- Issue Date: 23 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04189-5>

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/s41586-021-04189-5>

- Article
- Open Access
- [Published: 22 December 2021](#)

Rapid microbial methanogenesis during CO₂ storage in hydrocarbon reservoirs

- [R. L. Tyne](#) ORCID: [orcid.org/0000-0003-1316-3119](#)¹,
- [P. H. Barry](#) ORCID: [orcid.org/0000-0002-6960-1555](#)^{1,2},
- [M. Lawson](#) ORCID: [orcid.org/0000-0003-2743-3773](#)³ nAff9,
- [D. J. Byrne](#)⁴,
- [O. Warr](#) ORCID: [orcid.org/0000-0001-8240-7979](#)⁵,
- [H. Xie](#) ORCID: [orcid.org/0000-0001-5656-2035](#)⁶,
- [D. J. Hillegonds](#)¹,
- [M. Formolo](#)⁷,
- [Z. M. Summers](#)⁸,
- [B. Skinner](#)⁷,
- [J. M. Eiler](#) ORCID: [orcid.org/0000-0001-5768-7593](#)⁶ &
- [C. J. Ballentine](#) ORCID: [orcid.org/0000-0001-9382-070X](#)¹

[Nature](#) volume 600, pages 670–674 (2021)

- 10k Accesses
- 180 Altmetric
- [Metrics details](#)

Subjects

- [Environmental impact](#)
- [Hydrology](#)

Abstract

Carbon capture and storage (CCS) is a key technology to mitigate the environmental impact of carbon dioxide (CO_2) emissions. An understanding of the potential trapping and storage mechanisms is required to provide confidence in safe and secure CO_2 geological sequestration^{1,2}. Depleted hydrocarbon reservoirs have substantial CO_2 storage potential^{1,3}, and numerous hydrocarbon reservoirs have undergone CO_2 injection as a means of enhanced oil recovery (CO_2 -EOR), providing an opportunity to evaluate the (bio)geochemical behaviour of injected carbon. Here we present noble gas, stable isotope, clumped isotope and gene-sequencing analyses from a CO_2 -EOR project in the Olla Field (Louisiana, USA). We show that microbial methanogenesis converted as much as 13–19% of the injected CO_2 to methane (CH_4) and up to an additional 74% of CO_2 was dissolved in the groundwater. We calculate an in situ microbial methanogenesis rate from within a natural system of 73–109 millimoles of CH_4 per cubic metre (standard temperature and pressure) per year for the Olla Field. Similar geochemical trends in both injected and natural CO_2 fields suggest that microbial methanogenesis may be an important subsurface sink of CO_2 globally. For CO_2 sequestration sites within the environmental window for microbial methanogenesis, conversion to CH_4 should be considered in site selection.

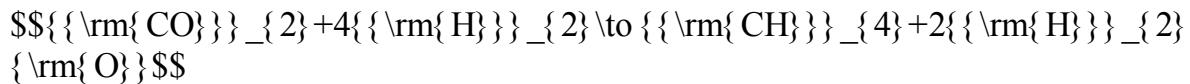
[Download PDF](#)

Main

A possible method for reducing current greenhouse gas emission rates is carbon capture and storage (CCS). Long-term carbon dioxide (CO_2) trapping mechanisms include structural or stratigraphic trapping in stable geological configurations, dissolution into pore fluids (solubility trapping)⁴, carbonate mineralization (precipitation and mineral trapping)⁵ or adsorption (for example, onto coal)⁶. Typically, during CO_2 enhanced oil recovery (CO_2 -EOR), a proportion of the injected CO_2 remains within the reservoir post-injection³, providing an analogue for investigating and quantifying processes within CCS sites, over decadal timescales.

Here we investigate the behaviour of CO_2 within the Olla Oil Field, Louisiana, USA (Fig. 1), which was CO_2 -EOR flooded in the 1980s, and compare its (bio)geochemical composition with the adjacent Nebo-Hemphill Oil Field, which was never subjected to CO_2 -EOR (Methods). Approximately $9 \times 10^7 \text{ m}^3$ (standard temperature and pressure (STP)) of injected CO_2 has been retained in the Olla reservoirs post-injection⁷. The injected CO_2 was sourced from the Black Lake Oil Field, which is located adjacent to the Sabine Island Complex, a basement high and likely conduit for mantle-derived

fluids⁸. Although the Olla Oil Field has been geochemically characterized by several previous studies^{9,10,11} (Methods), this study integrates noble gas and clumped isotope data with stable isotope and microbiological data to investigate the fate of the injected CO₂. Previous studies have suggested that microbial hydrocarbon degradation and methanogenesis may occur within both the Olla and Nebo-Hemphill oil fields^{9,10,11}. Microbial methanogenesis takes acetate, methylated compounds, or hydrogen and CO₂ (hydrogenotrophic methanogenesis, equation (1)) as its starting point; the latter mechanism is most relevant to this study. Hydrogenotrophic methanogenesis results in ¹³C enrichment of residual CO₂ and the formation of methane (CH₄) initially ¹³C depleted compared with the source CO₂, which increases to the value of the consumed CO₂.



(1)

Fig. 1: Map of study area showing locations of the Olla and Nebo-Hemphill oil fields as well as the Black Lake Oil Field, from which the injected CO₂ was sourced.

 figure1

Only the Olla Oil Field contains injected CO₂. The inset on the right shows an expanded view with individual sample locations (nearby urban areas are denoted by stars) and a stratigraphic column showing the relevant lithologic units⁹.

Given their proximity and comparable geological histories (including hydrocarbon production from similar reservoirs; Methods), we assume that the Nebo-Hemphill and

Olla oil fields initially had comparable geochemical compositions, before CO₂-EOR at Olla. For example, the Olla and Nebo-Hemphill oil fields have broadly overlapping pH, temperatures, salinities and ion chemistry (Extended Data Table 1). Despite CO₂ injection ceasing in 1986, CO₂ concentrations within the Olla Oil Field (18 ± 12 mol% (STP) across all reservoirs) are greater than breakthrough concentrations limits (10%)⁷ and significantly greater than those measured at Nebo-Hemphill (1.02 ± 0.69 mol% (STP)). We observe a higher $\delta^{13}\text{C}_{\text{VPDB}}$ of CO₂ at Olla ($13.5 \pm 3.4\text{\textperthousand}$) compared with both Nebo-Hemphill ($4.8 \pm 4.9\text{\textperthousand}$) and the injected CO₂ ($0.85 \pm 0.92\text{\textperthousand}$)¹⁰ (Extended Data Table 2), where $\delta^{13}\text{C} = [({^{13}\text{C}}/{^{12}\text{C}})_{\text{sample}} / ({^{13}\text{C}}/{^{12}\text{C}})_{\text{standard}}] - 1$ and VPDB is Vienna PeeDee Belemnite. Injection alone cannot account for the higher $\delta^{13}\text{C}$ of CO₂ values at Olla; therefore, the integration of complementary datasets is required to evaluate additional in situ processes.

Owing to their inert nature, noble gases provide a powerful tool for tracing and quantifying physiochemical processes associated with CO₂ injection^{4,8,12,13,14,15,16,17,18,19}. The average air-corrected helium isotope (${}^3\text{He}/{}^4\text{He}$) ratio relative to the atmospheric ratio (R/R_A , where $R = {}^3\text{He}/{}^4\text{He}_{\text{sample}}$ and $R_A = {}^3\text{He}/{}^4\text{He}_{\text{air}}$) was determined for the Olla and Nebo-Hemphill oil fields to be $1.76 \pm 0.31 R_A$ and $0.46 \pm 0.12 R_A$, respectively (Extended Data Fig. 1, Extended Data Table 2). The elevated He isotope values in the Olla Oil Field require an enhanced mantle-derived noble gases contribution, which is supported by additional noble gas data (Extended Data Fig. 1, Methods), associated with injected CO₂, which were sourced from the Black Lake Oil Field.

In addition to characterizing fluid origin, a combined noble gas and CO₂ isotope approach provides insight into processes associated with CO₂ trapping (for example, ref. 4) or methanogenesis following injection. ${}^3\text{He}$ is inert and insoluble, with no significant sources within the crust. Thus, variations in CO₂/ ${}^3\text{He}$ post-injection are directly attributable to the addition or removal of CO₂ within the system^{4,20}. Although two samples at Olla (samples O5 and O6) have CO₂/ ${}^3\text{He}$ within the mantle range ($2 \pm 1 \times 10^9$), consistent with the injection of mantle-derived fluids, the majority of the ratios are lower and indicate CO₂ removal (Fig. 2a, Extended Data Table 2). The difference between the highest CO₂/ ${}^3\text{He}$ (O5, considered the most pristine sample) and the remaining samples provides a conservative estimate of post-injection CO₂ trapping/consumption of between 39% and 89%.

Fig. 2: The $\delta^{13}\text{C}$ of CO₂ in the Olla (CO₂ injected field) samples.

 **figure2**

a, The $\delta^{13}\text{C}$ of CO_2 as a function of the $\text{CO}_2/{}^3\text{He}$ ratio. The dashed lines show endmember methanogenesis and dissolution (pH 7) fractionation trajectories. The tick marks represent the total amount of CO_2 trapping within the system, relative to sample O5. The shaded region represents trapping by the combination of both microbial methanogenesis and dissolution. The upper and lower methanogenesis:dissolution ratios (M:D) are 0.33 and 0.19, respectively, showing that dissolution accounts for approximately three times more CO_2 removal (M:D = 0.26) than microbial methanogenesis. The lines labelled ‘consumption’ show the portion of original injected CO_2 that has been removed by net microbial methanogenesis. **b**, The $\delta^{13}\text{C}$ of CO_2 as a function of the $\delta^{13}\text{C}$ of CH_4 . The shaded region represents the CO_2 isotopic composition of the injectate from the Black Lake Oil Field into the Olla system. The Olla data are consistent with thermal re-equilibration with both reservoir temperatures (solid lines) and microbial methanogenesis (dotted lines). The 1σ level of uncertainty is within the symbol size.

A decrease in the $\text{CO}_2/{}^3\text{He}$ alongside increasing ${}^4\text{He}$ and ${}^{20}\text{Ne}$ concentrations is observed within the Olla Oil Field (Extended Data Fig. 2, Methods). In the subsurface, ${}^4\text{He}$ is produced via radiogenic decay of uranium and thorium and accumulates in formation water¹², which also contains atmosphere-derived ${}^{20}\text{Ne}$. An anticorrelation between $\text{CO}_2/{}^3\text{He}$ and ${}^4\text{He}$ and ${}^{20}\text{Ne}$ is observed in natural CCS analogues globally⁴, and indicates that the extent of water contact controls the magnitude of CO_2 trapping via dissolution or precipitation of CO_2 (refs. ^{4,15}). Within the observed pH range (7.0 ± 0.5) at Olla, dissolution will dominate over precipitation as a CO_2 trapping

mechanism^{4,10}. However, as described above, the $\delta^{13}\text{C}$ of CO_2 in the Olla gases is more enriched than can be explained by these processes alone (Fig. 2a).

Independent of the source of the injected CO_2 , such positive $\delta^{13}\text{C}$ values suggest significant modification from a starting composition that is unlikely to be heavier than 2‰ (refs. 21,22). Notably, sorption to coal, hydrocarbon biodegradation to CO_2 , and subsequent consumption and mixing processes could result in an increase in the $\delta^{13}\text{C}$ of CO_2 (Methods). However, even from the most pristine sample (O5), it is unlikely that any of these processes could generate a fluid with the elevated concentrations of the isotopically enriched CO_2 observed at Olla.

Further constraints on the origins and processes affecting these gases are provided by the molecular-average and clumped isotope compositions of the associated CH_4 . We observe a relatively high $\delta^{13}\text{C}$ of CO_2 ($13.6 \pm 3.2\text{\textperthousand}$) and CH_4 ($-47.7 \pm 1.7\text{\textperthousand}$) at Olla compared with those observed at Nebo-Hemphill ($3.5 \pm 3.6\text{\textperthousand}$ and $-56.4 \pm 3.6\text{\textperthousand}$ for CO_2 and CH_4 , respectively) (Fig. 2b, Extended Data Fig. 3a, Extended Data Table 2). The higher $\delta^{13}\text{C}$ of CH_4 at Olla could be interpreted as evidence that the Olla Oil Field contains a higher proportion of thermogenic gas, or thermogenic gas of higher thermal maturity, whereas the Nebo-Hemphill Oil Field contains more microbial gas or thermogenic gas of lower thermal maturity. Using our multi-isotope approach, we present a model that is consistent with the observed higher $\delta^{13}\text{C}$ values of CH_4 and an increase in the $\delta^{13}\text{C}$ of residual CO_2 . In this model, Olla is assumed to have had a pre- CO_2 -EOR composition resembling Nebo-Hemphill, but following CO_2 -EOR, microbial activity converted significant amounts of injected CO_2 to CH_4 (Extended Data Fig. 4). We arrive at this conclusion on the basis of the following observations.

First, we report a difference between the $\delta^{13}\text{C}$ of CH_4 and the $\delta^{13}\text{C}$ of CO_2 of 53.7–64.4‰ at the Olla Oil Field and of 50.8–68.1‰ at the Nebo-Hemphill Oil Field (Fig. 2b). The observed carbon isotope fractionations between co-existing CO_2 and CH_4 are consistent with thermodynamic isotopic equilibrium at temperatures between 44.6 °C and 76.3 °C at Olla and between 29.8 °C and 85.3 °C at Nebo-Hemphill²³, which overlap the present-day reservoir temperatures within these fields (29.2–50.7 °C and 29.7–57.1 °C, respectively), which suggests that the systems are approaching isotopic equilibrium under current reservoir conditions (Methods). The isotopic approach to equilibrium under reservoir conditions appears to be a result of microbial cycling of carbon (that is, methanogenesis and anaerobic oxidation of methane (AOM); evidence for AOM is apparent in the clumped isotopologues, see below). Equilibrium with reservoir conditions between the $\delta^{13}\text{C}$ of CO_2 and the $\delta^{13}\text{C}$ of CH_4 has previously been observed under similar geological conditions²⁴.

Second, we observe a range in the two measurable clumped isotopologues of CH₄, Δ¹³CH₃D and Δ¹²CH₂D₂, at Olla of 3.45–5.62‰ and 9.13–12.4‰ (Fig. 3, Extended Data Table 2), respectively, which correspond to apparent temperature ranges of \(\{29\}_{-12}^{+14}\) °C to \(\{128\}_{-16}^{+17}\) °C and \(\{87\}_{-11}^{+13}\) °C to \(\{132\}_{-23}^{+30}\) °C, respectively, and are approaching equilibrium with current reservoir conditions. Clumped isotope compositions provide an independent constraint on the origin of CH₄ in petroleum systems^{25,26,27,28,29,30}. Thermogenic CH₄ appears to dominantly form under internal isotopic equilibrium^{26,30} whereas microbial CH₄ is highly variable in its clumped isotope compositions, with both non-equilibrium^{28,29} and apparent equilibrium signatures^{24,26,28,29,31,32,33,34}. For Olla, we expect thermogenic-CH₄-generation temperatures to have exceeded 163 ± 18 °C, from independent maturity constraints provided by biomarkers³⁵. Two samples from Olla (O4 and O6) appear to be in internal isotopic equilibrium (about 125 °C and about 95 °C) but have lower apparent temperatures than expected for pure thermogenic CH₄. These temperatures are consistent with a component of microbial CH₄ formed at or close to equilibrium with current reservoir conditions. The remaining samples have Δ¹³CH₃D apparent temperatures of \(\{28.7\}_{-12.3}^{+13.6}\) °C to \(\{74.0\}_{-15.0}^{+16.8}\) °C, within error of the present-day reservoir temperature for these fluids (Methods), consistent with in situ microbial methanogenesis and AOM. Furthermore, these samples exhibit deficits in Δ¹²CH₂D₂, consistent with fluids that are dominated by microbial CH₄^{28,29,34} in a system approaching equilibrium via active methanogenesis and AOM. In contrast, the Δ¹³CH₃D-based temperatures of \(\{77.9\}_{-12.3}^{+13.5}\) °C to \(\{166\}_{-29}^{+34}\) °C at Nebo-Hemphill suggest that these fluids are dominated by thermogenic CH₄, although, similarly to Olla, the Δ¹²CH₂D₂ clumped isotopologues also suggest a microbial CH₄ contribution.

Fig. 3: Δ¹²CH₂D₂ versus Δ¹³CH₃D of the measured Olla and Nebo-Hemphill samples.

 figure3

The clumped isotopologue space illustrates whether measured CH₄ is at internal thermodynamic equilibrium (black line) or not. The thermodynamic equilibrium curve is calculated following ref. ⁴¹. The shaded cross represents thermal equilibration to current reservoir temperatures in each isotopologue. The arrows represent the theoretical trends for methanogenesis (dark grey) and AOM (light grey) ²⁹. The 1 σ level of uncertainty is shown on the measured samples.

Third, we identify evidence for the biodegradation of hydrocarbons in the molecular geochemistry of oils produced from both the Olla and the Nebo-Hemphill oil fields (Methods). In the case of Olla, we also report an elevated $\delta^{13}\text{C}$ of propane (Extended Data Fig. 3b). The presence of oils co-existing with propane in these reservoirs

preclude the possibility of high-temperature cracking as the cause of the positive $\delta^{13}\text{C}$ values and suggests that the propane has been subjected to biodegradation³⁶, consistent with low-temperature microbial activity within the field.

Fourth, small subunit (SSU) ribosomal RNA gene sequencing of microbial communities identifies the presence of hydrogenotrophic methanogens, methanol and methylotrophic methanogens, and anaerobic methanotrophs (ANME) assigned sequences (Extended Data Fig. 5). These were sampled from the formation water in two wells within each field alongside previous microbiology analysis of these fields³⁷. Co-occurrence of both methanogenic and methanotrophic Archaea shows the potential for conversion of CO_2 to CH_4 and vice versa, consistent with isotopic signatures at intramolecular and intermolecular equilibrium. However, equilibration is probably rate limited by AOM, as the abundance of methanogens was 100 times higher than ANME, and AOM probably proceeds at a lower rate per cell.

Combined independent geochemical and microbiological data support our assertion that oil field microbes convert injected CO_2 to CH_4 at Olla. This study identifies the microbial processes associated with the injection of CO_2 into the deep subsurface operating on decadal timescales and the scale of the processing.

To quantify the impact of these microbial processes on the Olla Oil Field, we have constructed a model that considers both dissolution and microbial methanogenesis (Fig. 2a, Methods). From our most pristine sample (O5, highest CO_2/He), we find that to match the isotopic composition of CO_2 with its corresponding CO_2/He , dissolution must account for 3.1 ± 1.1 times more CO_2 removal than hydrogenotrophic methanogenesis. From this model, net microbial methanogenesis is estimated to have consumed a minimum of between 13% and 19% of the post-injection CO_2 , depending on sample location, whereas dissolution is responsible for removal of as much as 74% more CO_2 . This estimate is independently confirmed when CO_2 concentrations are compared with the $\delta^{13}\text{C}$ of CO_2 values (Extended Data Fig. 6). On the basis of the elevated $\delta^{13}\text{C}$ of CO_2 , we recognize that our most pristine sample has probably also undergone a significant degree of modification, consistent with our estimate being conservative, and that CO_2 concentrations post-injection were likely to be higher and variable across the field.

Our combined noble gas and stable isotope approach allows estimates of in situ net microbial methanogenesis rates from within a natural system to be calculated. By extrapolating our results over the 29 years between the cessation of injection (1986) and sampling (2015), and assuming 13–19% microbial consumption of CO_2 since injection (from the hybrid methanogenesis–dissolution model, Fig. 2) with the

remaining injected CO₂ volume, we calculate that a minimum of 1.15×10^7 – 1.72×10^7 m³ (STP) of microbial CH₄ has been produced at a minimum rate of 73–109 mmol CH₄ m⁻³ (STP) yr⁻¹. Previous estimates for CO₂ reduction following methanogenic oil degradation by hydrogenotrophic methanogens in lab microcosm incubations at similar temperatures are significantly lower (about 0.01–0.15 mmol CH₄ m⁻³ (STP) yr⁻¹)³⁸ by comparison.

This identification and quantification of CO₂-EOR-related microbial methanogenesis at Olla has significant implications, and warrants reconsideration of reservoir processes at similar sites worldwide. The fate of injected CO₂ in other oil fields (for example, Belridge Diatomite Formation, Lost Hills Oil Field, California, USA^{17,39}) has never been systematically explored using this integrated approach. Similarly to Olla, the Lost Hills Field has an elevated δ¹³C of dissolved inorganic carbon (DIC) (20.16–23.61‰)⁴⁰ compared with the injected δ¹³C of CO₂ (-30.1‰)³⁹. The fractionation between δ¹³C of DIC and CH₄ suggests that the Lost Hills system is approaching equilibrium with current reservoir conditions and the inverse correlation between the δ¹³C of the post-injection residual CO₂ and concentration is consistent with microbial methanogenesis. Furthermore, data from Nebo-Hemphill, as well as naturally CO₂-rich hydrocarbon systems in the Pannonian Basin, Hungary, are also consistent with significant microbial methanogenesis based on the CO₂/³He and the δ¹³C of CO₂ (Extended Data Fig. 7, Methods). These examples suggest that hydrogenotrophic methanogenesis probably occurs in both natural and injection-perturbed systems when reservoirs have favourable physiochemical conditions.

We conclude that although methanogenesis is not the dominant CO₂ sink, it can represent a substantial process within both natural and perturbed CO₂ fields and may be significantly more prevalent than previously considered. Even with the most conservative estimates of CO₂ conversion by methanogenesis in the Olla Oil Field, we find that, so far, as much as 13% to 19% of the emplaced CO₂ in sampled sections of the field has been consumed by methanogens. This is less than the amount of CO₂ that has been dissolved into water, but similarly occurs at significant rates on engineering timescales. CH₄ is less soluble and more mobile than CO₂ and therefore there is an enhanced risk of gas loss associated with CH₄ production due to microbial methanogenesis. Depleted hydrocarbon reservoirs have the second-largest CO₂ storage potential^{1,3} after deep saline aquifers. If physicochemical and environmental conditions within these reservoirs (for example, low temperature, low salinity and high substrate availability) are conducive to microbial methanogenesis, we suggest that these microbial processes should be considered as criteria for future CCS site selection

to ensure low-risk long-term storage. In addition, integrated studies like this may prove essential to effectively monitor CO₂ storage and the biogeochemical processes that result as a consequence of it.

Methods

Modelling

Here we provide extra details of the δ¹³C-CO₂ fractionation modelling of methanogenesis and dissolution. In our models (Fig. 2a, Extended Data Fig. 6), we use the sample with the highest measured CO₂/³He (most pristine sample, O5) as a reference point to calculate the change in CO₂/³He and δ¹³C-CO₂ with CO₂ processing (that is, methanogenesis and dissolution). We assume there is no ³He loss from the gas phase by dissolution or methanogenesis and therefore changes in the fraction of the CO₂ cause changes in the CO₂/³He.

Dissolution is modelled as open-system Rayleigh fractionation⁴.

$$\delta_{\text{CO}_2i} = \delta_{\text{CO}_2i}^{13} - (\alpha - 1) f \times (\delta_{\text{CO}_2i}^{13} - \delta_{\text{CO}_2i}^{13}) \quad (2)$$

where f is the fraction of the original CO₂ remaining in the reservoir, $\delta_{\text{CO}_2i}^{13}$ and $\delta_{\text{CO}_2i}^{13}$ are the carbon isotopic composition of the CO₂ for post-dissolution and in the initial CO₂ respectively, and α is the fractionation factor. The fractionation factor for dissolution ($\alpha_d = 1.0038 \pm 0.0012$) was calculated based on the average measured pH¹⁰. Precipitation curves (Fig. 2a) can be modelled in the same way using fractionation factors ($\alpha_p = 1.0086 \pm 0.0012$) calculated for the current reservoir temperature²¹.

From consideration of the carbon isotopes, which appear to be approaching equilibrium under current reservoir conditions, net methanogenesis has been modelled as a closed-system equilibrium process (equation (2), Fig. 2a). The fractionation resulting from this cycling, α_m , is 0.9363 ± 0.003 , which is consistent with thermal equilibration at current reservoir temperatures²³.

$$\delta_{\text{CO}_2i} = \delta_{\text{CO}_2i}^{13} - (\alpha_m - 1) (1-f) \times (\delta_{\text{CO}_2i}^{13} - \delta_{\text{CO}_2i}^{13}) \quad (3)$$

Within the Olla Oil Field, it is likely that there is a combination of both microbial cycling and dissolution. As a result, the net effect of microbial methanogenesis and dissolution can be described using equation (3).

$$\delta^{13}\text{C}_{\text{rmCO}_2} = \delta^{13}\text{C}_{\text{rmC}} - \frac{\alpha_A}{2} \times f \times (\alpha_A - 1) \times 1,000$$

(4)

where f is the fraction of CO_2 remaining after trapping/consumption and α_A is the apparent fractionation factor for the system. α_A depends on the degree of dissolution compared with methanogenesis and is described as⁴²:

$$\alpha_A = F_d \alpha_d + (1 - F_d) \alpha_m$$

(5)

where F_d is the fraction of CO_2 of the total CO_2 trapped/consumed due to dissolution.

Sample collection and analysis

Samples for bulk gas composition, stable isotope, noble gas and clumped isotope analysis were collected directly from the wellhead using standard techniques following methods from refs. [8,17,18,24](#). Noble gas isotope determination was conducted in the Noble Laboratory at the University of Oxford, where the analysis of hydrocarbon gases is well established^{[8,17,18,43,44,45](#)}. For bulk composition and C and H stable isotope analysis, gas cylinders were shipped to Isotech in Champaign, Illinois, USA. The standard procedures for these are described in ref. [46](#). CH_4 clumped isotopologues were determined at the California Institute of Technology following the procedures described in ref. [46](#).

Microbiological analysis was conducted at an internal ExxonMobil Research and Engineering Company Facility. The produced water samples were collected and filtered immediately from wells O1, O4, NHH3 and NHH6. Water was passed through 0.22- μm SVGPL10 RC Sterivex GP filters (EMD Millipore) until filters clogged in the following amounts: O1, 1,350 ml; O4, 1,350 ml; NHH3, 1,750 ml; NHH4, 300 ml. Filters were stored on ice immediately and stored at -80°C until DNA extraction. DNA was extracted from the Sterivex filters using the DNeasy PowerWater Sterivex Kit (Qiagen). The SSU rRNA gene was amplified with V4V5 specific primers (<https://vamps2.mbl.edu/>) targeting Bacteria. The SSU rRNA gene was amplified with specific primers targeting Bacteria: forward primers (967F) CTAACCGANGAACCTYACC, CNACCGCGAAGAACCTTANC, CAACGCGMARAACCTTACC and ATACGCGARGAACCTTACC; reverse primer

(1064R) CGACRRCCATGCANCACCT. A second set of amplifications was performed using specific primers targeting Archaea: (Arch2A519F) CAGCMGCCGCGGTAA and (Arch1017R) GGCCATGCACCWCCTCTC. Illumina MiSeq 2x300bp paired-end sequencing was performed by Mr. DNA. Microbiome bioinformatics were performed with QIIME 2 2017.4⁴⁷. Raw sequence data were demultiplexed and quality filtered using the q2-demux plugin followed by denoising with DADA2⁴⁸. Amplicon sequence variants were aligned with mafft and used to construct a phylogeny with fasttree2^{49,50}. Taxonomy was assigned to amplicon sequence variants using the q2-feature-classifier⁵¹ classify-sklearn Naïve Bayes taxonomy classifier against the Greengenes 13_8 99% operational taxonomic unit (OTU) reference sequences⁵².

Geological and production history

The Olla Oil Field is located in La Salle Parish, Louisiana, in the northern Gulf of Mexico and was CO₂ injected between 1983 and 1986 for enhanced oil recovery (CO₂-EOR). The Nebo-Hemphill Oil Field lies 20 km to the southeast of the Olla Oil Field and provides a control field for this study, having had no CO₂ injected (Fig. 1). Both fields are producing from the Palaeocene–Eocene Wilcox Group, which in northern Louisiana is composed of 600–1,500 m of shallow marine clastic sediments to fluvial deltaic deposits and up to 6-m-thick coal beds^{53,54}. The group is bound either side by the Palaeocene Midway Group and the Eocene Clairborne Group confining units⁵⁴. Typically, the Wilcox Group is subdivided into the lower (sandstone with coal beds), middle (sandstone, lignite and clay) and upper (deltaic deposits) intervals, with the ‘Big Shale’ informal unit separating the middle and upper groups^{53,55}. Late Cretaceous intrusive igneous bodies in northern Louisiana, west-central Mississippi and southeastern Arkansas are thought to have affected the regional heat flow, which may have caused an increase in the thermal maturity of the Wilcox Group during the early Palaeocene^{56,57}.

Gas/oil ratios are variable within the field, and do not seem to be associated with proximity to the known gas cap and higher gas/oil ratios may instead reflect gas coming out of solution as bottom hole pressure decreases.

CO₂ is present in low concentrations (1.02 ± 0.69 mol%) within the Nebo-Hemphill Oil Field, which suggests that there is some natural background CO₂ present in both fields. Between March 1983 and April 1986, the 2,800-ft sandstone (within the middle Wilcox Group) in the Olla Oil Field underwent a pilot project for CO₂ flooding. The CO₂ injected into the field was immiscible and injected in the gas phase (that is, not supercritical). It is assumed that a negligible amount of CH₄ was injected alongside the

CO_2 . Previous studies on the Olla Oil Field have assumed that all the injected CO_2 was retained within the 2,800-ft sandstone^{9,10}. However, the initial CO_2 injection report⁷ states that “mole percentages in excess of 10% more than likely indicate production of injected CO_2 ”. Only 1 sample outside of the 2,800-ft sandstone has CO_2 concentrations less than 10% (that is, all other CO_2 concentrations are greater than 10%). Notably, concentrations greater than 10% in other producing zones have previously been reported¹⁰. These signatures are consistent with injected CO_2 migration across the different reservoirs and remaining within the formations upon sampling. If such high concentrations of CO_2 were present in this field before injection, CO_2 would not have been extracted and transported from the Black Lake Oil Field and would have instead have been extracted from the producing fluids of CO_2 rich fluids at Olla. In addition, we see a clear mantle-derived noble gas signature (see below), which is most evident in the He and neon (Ne) isotopes (Extended Data Fig. 1), in the Olla Oil Field, which shows a clear incursion of injection CO_2 into the proximal lithologies.

A total of $2.2 \times 10^8 \text{ m}^3$ of CO_2 was injected into the Olla Oil Field and was sourced from the nearby Black Lake Oil Field (produced from the Lower Cretaceous carbonates of the Petit Formation) in Natchitoches Parish, Louisiana (Fig. 1). Approximately one-third of the injected CO_2 remained within the reservoir after the cessation of the EOR project⁷. It is important to note that given the age of this CO_2 -EOR project, we were unable to obtain constraints such as the ratio of injected CO_2 to reservoir pore volume that are sometimes cited to quantify the scale of the EOR project and the distribution it achieves within the subsurface. Further, we anticipate that the injected CO_2 is not uniformly distributed within the pore volume of the Olla Oil Field due to different phases of EOR and gas cycling.

In 1984, water flooding also started within the field. In contrast, Nebo-Hemphill has never undergone any type of EOR.

Molecular geochemistry

Whole-oil gas chromatography (WOGC) was performed to provide a hydrocarbon fingerprint of each oil, and open-column liquid chromatography was utilized to separate and quantify the saturate, aromatic, resins and asphaltene fractions of each oil. In general, we see similar fluid compositions between the two fields. The WOGC shows evidence of marine-sourced oils with a smooth distribution of *n*-alkanes, pristane/phytane ratios of approximately 2 and low wax content. The notable differences are probably due to varying levels of biodegradation illustrated by the loss of *n*-alkanes in some samples and American Petroleum Institute (API) gravity ranging

from 21 to 42. The similarities in the geochemical signatures of these oils provide a basis to use the Nebo-Hemphill Oil Field as a control site for any biogeochemical differences that may arise as a result of the injection of CO₂ at Olla.

Source facies

Gas chromatography mass spectroscopy (GC/MS) was utilized to analyse the saturate and aromatic fractions to provide the biomarker distributions of each oil. The interpretation of a dominantly marine source is based on the following evidence: (1) a smooth distribution of tricyclic terpanes with C₂₃ being the maximum along with extended hopanes (C₃₁–C₃₅) decreasing in order is a classic marine source signature; (2) the high presence of C₃₀ hopane compared with C₂₉ hopane is indicative of a marine clastic source; (3) a higher abundance of C₂₈ bisnorhopane compared with oleanane suggest a higher marine input compared with the above source⁵⁸; (4) a greater abundance of C₂₆ tricyclic terpane as compared to C₂₅ tricyclic terpane provides evidence of some terrigenous input. Consistent with previous interpretations, we consider this source to be a part of the Wilcox Group oils, more specifically Wilcox Group I as originally described by refs. ^{59,60}.

Thermal maturity

We adopted standard saturate and aromatic biomarker ratios to develop apparent thermal maturity estimates for each of the oil samples analysed. We discuss these in the context of apparent thermal maturities because we recognize these estimates are based on the bulk weighted average composition of the fluid. The interpretation of the GC/MS analyses of the 12 oils yielded a thermal maturity range of R_o equivalent of 0.9–1.25%, which is classified as main-to-late-stage oil. The interpretation of this thermal maturity range is based on the following: (1) the presence of higher concentrations of high-molecular-weight compared with low-molecular-weight triaromatic steriods, described as % ΔTAS (20–57%); (2) the isomerization of C₂₉ $\alpha\alpha\alpha$ -20S relative to C₂₉ $\alpha\alpha\alpha$ -20R steranes (0.45–0.48)⁶¹; (3) the isomerization of 1,2,5-trimethylnaphthalene to 1,3,7-trimethylnaphthalene represented as ratio TMNr (0.48–0.67) as defined by ref. ⁶²; (4) the ratio of C₂₃ tricyclic terpane to C₃₀ hopane, which is calculated as T23/(T23 + H30) (0.10–0.37), where the increased presence of C₂₃ tricyclic represents higher maturity.

Alteration

The use of WOGC and GC/MS analyses help provide a picture of the alteration history of these oils. The shallow Nebo-Hemphill NHH3 fluid is the most biodegraded end

member of the samples tested, based on the apparent lack of *n*-alkanes, an API gravity of 21 and the presence of demethylated hopanes. This sample would be classified as severely biodegraded based on the classification of ref. ⁶³. The remaining Nebo-Hemphill samples have APIs of 32–41 and are classified as unbiodegraded to very slight biodegradation⁶³. The Olla Oil Field fluids show evidence of reduced light *n*-alkanes, display similar *n*-alkane-to-isoprenoid ratios and have APIs of 33 to 37, suggesting similar levels of biodegradation between the fluids. Therefore, the Olla fluids would be classified as having very slight biodegradation⁶³.

Summary of previous findings

The formation water geochemistry, molecular oil composition and produced gas composition (including $\delta^{13}\text{C}$) in the Olla and Nebo-Hemphill oil fields have been previously characterized^{9,10,11}. Microbial methanogenesis was first hypothesised for the Olla Oil Field based on its enriched $\delta^{13}\text{C}$ of CO_2 (ref. ⁹). However, owing to the differences in the isotopic composition between the 2,800-ft sandstone and the other producing formations with the injected material, they concluded that the injected CO_2 was not present in the producing zones¹⁰ and consequently any microbial methanogenesis was independent of injection and is probably associated with microbial oxidation of hydrocarbons. The enhanced methanogenesis within the Olla Oil Field compared with other fields produced from the Wilcox Group was thought to be due to the ideal geochemical conditions (that is, suitable range of pH, temperature and salinity, and lack of alternative electron acceptors)¹⁰. They conclude that the injected CO_2 was most likely dissolved and trapped in a gas phase¹⁰. Further investigation of microbial methanogenesis found that there was little correlation between crude oil biodegradation and methanogenesis; however, the extent of oil biodegradation was correlated with the temperature, salinity and depth of samples¹¹. In addition, there has also been an investigation of the microbial community composition within the Olla reservoir³⁷, which found methanogens present within the system, and no significant difference in microbial communities between the 2,800-ft sandstone and other producing zones³⁷. With the addition of the inert noble gas isotopes and clumped isotopes, we are able to re-investigate the fate of CO_2 within the Olla reservoirs.

Formation water chemistry in the Olla Oil Field

The pH and alkalinity of the formation waters within the Olla Oil Field are between 6.5 and 7.5 and between 16.3 meq kg^{-1} and 57.7 meq kg^{-1} , respectively¹⁰. The most dominant anion was Cl^- ($1,122\text{--}1,621 \text{ mmol l}^{-1}$) and the most dominant cation measured was Na^{2+} ($1,108\text{--}1,630 \text{ mmol l}^{-1}$)

(Extended Data Table 1), which are typical of basinal brines⁶⁴ and in agreement with previous measurements¹⁰. Other cations present include: Mg²⁺ (119–329 ppm), Ca²⁺ (80–500 ppm), K⁺ (223–524 ppm), Ba²⁺ (95–165 ppm), Si⁴⁺ (10.1–12.5 ppm) and Sr²⁺ (77–215 ppm). Measured iron concentrations are less than 0.1 ppm. With the exception of sample O1 where SO₄²⁻ concentrations are 62 ppm, SO₄²⁻ concentrations are less than 2 ppm. Both the iron and SO₄²⁻ concentrations are in agreement with those previously reported¹⁰.

Identification of injected material

CO₂ concentrations in the injected Olla Oil Field ($18 \pm 12\text{mol}\%$) are significantly greater than the Nebo-Hemphill Oil Field ($1.02 \pm 0.69\text{ mol}\%$). CH₄ concentrations at Olla are $77 \pm 13\text{ mol}\%$ (STP) compared with $92.6 \pm 8.2\text{ mol}\%$ (STP) at Nebo-Hemphill. However, the $\delta^{13}\text{C}$ of CO₂ of the injected material ($0.85 \pm 0.91\text{\textperthousand}$) is isotopically distinct from the $\delta^{13}\text{C}$ of CO₂ now within Olla ($13.6 \pm 3.2\text{\textperthousand}$) (Fig. 2b). The injected CO₂ was sourced from the Black Lake Oil Field, which is located close to the Sabine Island Complex. The Sabine Island Complex forms part of the Sabine Uplift basement high and has been suggested to be a conduit for mantle-derived fluids⁸. However, the measured $\delta^{13}\text{C}$ of CO₂ values at Black Lake ($0.85 \pm 0.91\text{\textperthousand}$)¹⁰ are elevated relative to mantle CO₂ ($-2\text{\textperthousand}$ to $-5\text{\textperthousand}$; refs. 21,22), probably reflecting some degree of biological modification when the reservoir was shallower and at temperatures favourable to microbial activity. An alternative interpretation for the elevated $\delta^{13}\text{C}$ of CO₂ at Black Lake is that it is a mixture of mantle-derived and crustal-derived CO₂; however, this would increase the CO₂/³He above that expected in the mantle, which is not observed in the Olla samples.

As the stable isotopes of carbon are frequently modified by chemical or biological processes, the inert nature of noble gases means that they are ideal tracers for injected CO₂ in such scenarios. As injectate samples were not collected during CO₂-EOR, we do not have access to the noble gas compositions of the material injected into the Olla Oil Field. Nevertheless,

it is reasonable that the noble gases record pre-injection conditions at the nearby Nebo-Hemphill Oil Field, and that these are subsequently modified at the Olla Oil Field through mixing of the injected material during the CO₂ injection.

The ⁴He/²⁰Ne ratio is between 3,210 and 20,900 in both fields, which is significantly greater than the atmospheric value of 0.32 and thus there is no significant atmospheric contribution to these samples. The air-corrected He isotopic ratio (³He/⁴He), reported relative to the atmospheric ratio (R_C/R_A where air is $1R_A$), in the Olla Oil Field is $1.84 \pm 0.10R_A$ and $0.46 \pm 0.08R_A$ at Nebo-Hemphill (Extended Data Table 2). These are both higher than a typical hydrocarbon system where the majority of He is derived from crustal radiogenic production (³He/⁴He = $0.02R_A$; ref. 12), and higher than other fields within the Gulf of Mexico region^{8,18,43}. The elevated ³He/⁴He values are the result of a resolvable mantle contribution. Using a simple two endmember mixing model, with subcontinental lithospheric mantle (³He/⁴He = 6.1; ref. 65) and crustal endmembers, we resolve an average 7.8% mantle contribution of mantle He in Nebo-Hemphill Oil Field. It is likely that the mantle fluid contribution is coming from the Little Creek collapse structure, which is situated between the two fields⁶⁶. At Olla, the average mantle contribution to ³He is 30%, which suggests that there is an additional source of mantle fluids to this field, which is reasonably explained by the injectate sourced from the Black Lake Oil Field. Ne isotope ratios are elevated at Olla (²⁰Ne/²²Ne 10.3 ± 0.2 and ²¹Ne/²²Ne 0.0399 ± 0.002) compared with Nebo-Hemphill (9.85 ± 0.09 and 0.0335 ± 0.001 , respectively) and air (9.8 and 0.029) (Extended Data Fig. 1, Extended Data Table 3)^{67,68,69}, again suggesting that the injected material contains mantle-derived fluids.

The ⁴⁰Ar/³⁶Ar values at Nebo-Hemphill are significantly greater than the atmospheric value (298.56; ref. 70), with measured ratios of 370 ± 3 to 604 ± 5 . The Olla Oil Field has a higher ⁴⁰Ar/³⁶Ar ($1,476 \pm 13$ to $2,766 \pm 24$) (Extended Data Table 3). Elevated ⁴⁰Ar/³⁶Ar ratios can be a result of both an excess of radiogenic ⁴⁰Ar and mantle-derived fluids^{12,71}. The difference in the ⁴⁰Ar/³⁶Ar ratios could be due to variable amounts of radiogenic

production or fundamental differences in the extent of contact with formation waters between the two fields. However, as all the wells sampled are within the same geological formation, the two fields are in relatively close proximity and they have indistinguishable ${}^4\text{He}$ concentrations, it seems unlikely that this would be the cause. Thus, the difference is probably a result of the injected CO_2 , which probably had mantle-derived ${}^{40}\text{Ar}/{}^{36}\text{Ar}$ excesses, concordant with the observed mantle-derived He and Ne isotopes. Finally, whereas xenon (Xe) isotope ratios at Nebo-Hemphill are indistinguishable from air, the Olla Oil Field showed excesses in some Xe isotope ratios relative to air. For example, the ${}^{129}\text{Xe}/{}^{130}\text{Xe}$ ratio is 6.59 ± 0.01 compared with an atmospheric value of 6.49 (Extended Data Table 3); this excess in ${}^{129}\text{Xe}$ relative to ${}^{130}\text{Xe}$ is consistent with a contribution from mantle fluids with an end member ${}^{129}\text{Xe}/{}^{130}\text{Xe}$ value of 8.2 (ref. 71). The Olla Oil Field therefore shows a clear mantle component that is not present within the Nebo-Hemphill control field.

The detection of injected fluids within the Olla Oil Field using the conservative noble gas tracers differs from the findings of the previous studies that used the $\delta {}^{13}\text{C}$ of CO_2 to conclude that injected CO_2 is no longer or never was present in the majority of the sampled wells (for example, ref. 10). The subsurface behaviour of the injected CO_2 within the system can therefore be re-examined using a combination of noble gases and stable isotopes.

Dissolution within the Olla Oil Field

CO_2 is highly soluble in water under reservoir conditions. The most likely trapping mechanisms for free-phase subsurface CO_2 are dissolution or precipitation as carbonate minerals mediated by dissolution^{4,21}. A decrease in the $\text{CO}_2/{}^3\text{He}$ correlates with increasing ${}^4\text{He}$ and ${}^{20}\text{Ne}$ concentrations within the Olla Oil Field (Extended Data Fig. 2). ${}^{20}\text{Ne}$ (and often ${}^4\text{He}$)⁴ are derived from the formation water, suggesting that CO_2 trapping in the Olla Oil Field is also proportional to the amount of formation water that has been contacted and degassed, consistent with dissolution within the field⁴. Dissolution (solubility trapping) has previously been found to be a major

sink for CO₂ (ref. 4) and imparts a slight pH-dependent fractionation towards a light isotopic composition²¹. However, this fractionation effect is marginal compared with the amount of CO₂ that can be trapped and the fractionation that can be imparted by methanogenesis–AOM cycling (Fig. 2a).

Given the presence of discontinuous coal beds within the Wilcox Group^{57,72}, it is possible that some of the injected CO₂ may have been sorbed onto these coal beds as CO₂ has a much greater affinity to coal than CH₄ (ref. 9). The resulting isotopic fractionation imparted by this process remains poorly constrained, with some studies suggesting a greater than 5‰ shift^{73,74} whereas others suggest negligible effects^{75,76,77} on the δ¹³C of CO₂. Coal sorption would cause fractionation in the opposition direction of the CH₄ isotopic signature. Given the heavier δ¹³C of CO₂ and δ¹³C of CH₄ isotopic signatures in the injected Olla Oil Field, it is unlikely that CO₂ sorption to coal is a major CO₂ sink in the field, in agreement with previous findings¹⁰.

Subsurface behaviour of carbon at Olla

The Olla samples are enriched in both the δ¹³C of CO₂ and the δ¹³C of CH₄ compared with Nebo-Hemphill and the injected CO₂ values. The offset between the δ¹³C of CH₄ and the δ¹³C of CO₂ in the Olla Oil Field are close to those expected under thermally driven isotope exchange equilibrium for the observed temperatures (Fig. 2b). Carbon isotopes of CO₂ and CH₄ approaching equilibrium in hydrocarbon systems with current reservoir conditions have previously been observed²⁴. However, the thermal equilibrium of CO₂ and CH₄ is kinetically limited at the current temperatures (29.2–50.7 °C) in the Olla reservoirs and would not occur on the timescale of injection, without a process to enhance the reaction rate⁷⁸. A combination of both microbial methanogenesis and AOM active within the field would allow for the system to be approaching equilibrium under the current thermal conditions within the reservoir. Further observations

detailed below provide evidence for this conclusion and allow for the net microbial methanogenesis to be quantified.

The clumped isotopologues of CH₄ ($\Delta^{13}\text{CH}_3\text{D}$ and $\Delta^{12}\text{CH}_2\text{D}_2$) can provide independent constraints on its origin within petroleum systems^{25,26,27,28,29}. Thermogenic CH₄ is expected to have a composition consistent with the source-rock maturation temperature, which is typically at or above 157 °C (ref. ²⁵). At Olla, we expect CH₄ generation to have occurred at 163 ± 18 °C (based on biomarker maturity estimates following ref. ³⁵). Microbial CH₄ is variable in its clumped isotopic compositions with both non-equilibrium^{28,29} and apparent equilibrium signatures reported in samples obtained from the deep subsurface^{26,29}. Microbial methanogenesis forms CH₄ characterized by low $\Delta^{13}\text{CH}_3\text{D}$ and extremely low $\Delta^{12}\text{CH}_2\text{D}_2$ (refs. ^{28,29}), whereas AOM tends to drive the isotopologue composition of CH₄ towards equilibrium, with equilibrium expected to be reached at a faster rate in $\Delta^{13}\text{CH}_3\text{D}$ compared with in $\Delta^{12}\text{CH}_2\text{D}_2$ (refs. ^{29,30,34,79,80,81}). In the Olla Oil Field, we see $\Delta^{13}\text{CH}_3\text{D}$ of 3.45–5.62‰ (apparent temperature of $\{-12\}^{+14}$ °C to $\{-16\}^{+17}$ °C) and $\Delta^{12}\text{CH}_2\text{D}_2$ of 9.13–12.4‰ (apparent temperature of $\{-11\}^{+13}$ °C to $\{-23\}^{+30}$ °C) (Fig. ³, Extended Data Table ²); these temperatures preclude a thermogenic origin for the CH₄ and are instead consistent with a microbial contribution. Samples O4 and O6 are in internal clumped isotopologue equilibrium at temperatures lower than expected for the thermogenic CH₄ and appear to be approaching equilibrium with current reservoir conditions. The remaining samples from Olla have already reached apparent equilibrium with the current reservoir conditions in $\Delta^{13}\text{CH}_3\text{D}$ but have yet to reach apparent equilibrium in $\Delta^{12}\text{CH}_2\text{D}_2$ probably due to kinetic effects. The clumped isotopologues at Olla are approaching equilibrium at reservoir temperatures, which would be accounted for by active AOM within the Olla Oil Field. This supports the conclusion that there is active microbial methanogenesis and AOM driving the carbon isotopes to approach equilibrium under current thermal conditions in the Olla reservoirs.

There is microbial degradation of the higher alkane gases (Extended Data Fig. 3) within the Olla Oil Field. Propane is preferentially consumed during microbial degradation and thus in biodegraded gases, propane will see the most pronounced $\delta^{13}\text{C}$ enrichment³⁶. This $\delta^{13}\text{C}$ enrichment is clear within the Olla gases (Extended Data Table 3) and supports active microbial hydrocarbon degradation within the field.

The presence of methanogenic activity and the idea of a methanogenic ‘hotspot’ in the Olla reservoir has been previously suggested^{9,10}. The environmental conditions for which methanogenesis can occur are relatively well defined. Within nutrient-poor environments such as oil reservoirs or coal beds, methanogens can survive at temperatures up to 80 °C, with the largest growth rates and thus greatest amount of methanogenic activity between 40 °C and 50 °C (refs. ^{36,82,83,84}). Methanogenic activity is greatest at near neutral pH and is inhibited at pH <4 and >9 (ref. ⁸⁴).

Formation water salinity can also inhibit methanogenesis above Cl^- concentrations of 1,500 mM, with greatest growth between 200 nM and 600 mM (refs. ^{82,83,85}). The presence of ferric iron and sulfate at concentrations greater than 1 mM can also limit methanogenesis (as a result of the dominance of other bacteria)^{86,87,88,89}. Notably, the Olla Oil Field exhibits ideal environmental conditions for methanogenesis, in terms of pH (6.5–7.5)¹⁰, temperature (29.2–50.7 °C), salinity (1,094–1,622 mM), sulfate (<0.6 mM l⁻¹)¹⁰ and iron concentrations (\leq 0.1 mM l⁻¹) (Methods).

In addition, geochemical and isotopic parameters associated with microbial methanogenesis in organic-rich settings are also well established^{10,11,83,88,90}. Traditionally, carbon isotopes of CH_4 have been used to determine the presence of biogenic CH_4 (typically less than $-60\text{\textperthousand}$)^{91,92,93} (Extended Data Fig. 3a). However, methanogenesis can result in heavy biogenic CH_4 as the residual CO_2 pool gets increasingly elevated, consistent with observations at Olla (Fig. 2b). In addition, where methanogenesis has occurred, formation water alkalinity is greater than 10 meq kg⁻¹, the acetate concentrations are lower than 1 mM, the Ca/Mg ratios are less than 1.5 and the $\delta^{13}\text{C}$ -DIC is greater than +20‰ (refs. ^{87,88,90}). Within the Olla Oil Field, the alkalinity ranges between 21.4 meq kg⁻¹ and

57.7 meq kg^{-1} , the average Ca/Mg ratio is 0.2 and the $\delta^{13}\text{C}$ -DIC is 14.2– $28.7\text{\textperthousand}$, again indicating that methanogenesis has occurred within the field.

SSU rRNA gene-sequencing analysis of the microbial communities within the Olla Oil Field (samples O1 and O4) indicate that within the archaeal sequences, hydrogenotrophic methanogens, methanol and methylotrophic methanogens, and ANME assigned sequences are present (Extended Data Fig. 5). This microbial array would facilitate both microbial methanogenesis and AOM within the field.

During the cycling of carbon between methanogenesis and AOM, we would not expect a change in the isotopic composition of hydrogen. It is possible that as a result of microbial activity or the reaction kinetics that extracellular hydrogen is added to the system. This hydrogen could be sourced from the water and free hydrogen from the degradation of the higher hydrocarbon gases or from the oil. As a result, we are unable to determine whether the hydrogen isotopes within the system are also approaching equilibrium under current reservoir temperatures.

It is possible that the CO_2 and CH_4 isotopes in the Olla system are reflecting purely the kinetic process of methanogenesis. The difference in the $\delta^{13}\text{C}$ of CO_2 and the $\delta^{13}\text{C}$ of CH_4 at Olla ($53.7\text{--}64.4\text{\textperthousand}$) is consistent with that expected from microbial methanogenesis (about $58\text{\textperthousand}$; ref. 94). In such a scenario, all the CH_4 at Olla would be microbial in origin, thus requiring any thermogenic CH_4 component originally in the system to be overwhelmed. The high concentration and mixed thermogenic–microbial composition of CH_4 within the Nebo-Hemphill Oil Field (Fig. 3, Methods) is inconsistent with such a scenario. In addition, the clumped isotopologues at Olla show a clear need for AOM within the field and 16S rRNA gene sequencing shows active methanotrophs, which is not compatible with a scenario of purely microbial methanogenesis dominating the system. Therefore, future studies investigating the processes affecting CO_2 in potential CCS storage targets will benefit from a complete geological, geochemical and microbiological characterization of the field study site.

Subsurface behaviour of carbon in the Nebo-Hemphill Oil Field

The carbon isotopic composition of CO₂ and CH₄ at Nebo-Hemphill is depleted compared with the Olla Oil Field (Fig. 2b). Despite this, the difference in the δ¹³C of CO₂ and the δ¹³C of CH₄ at Nebo-Hemphill overlaps with that of Olla (50.8–68.1‰ and 53.7–64.4‰, respectively). The degree of isotopic fractionation at Nebo-Hemphill is consistent with thermal equilibrium at 29.8–85.3 °C, in agreement with present-day reservoir temperatures (29.7–57.1 °C). In addition, the δ¹³C of CO₂ at Nebo-Hemphill is enriched compared with that typically expected for crustal or mantle-sourced CO₂ (less than 2‰)^{21,22}, consistent with a small amount of microbial methanogenesis within the field. We propose that there is also microbial methanogenesis (in agreement with previous studies¹⁰) and AOM within the Nebo-Hemphill Oil Field; however, unlike at Olla, this has not been enhanced by any injection to the field.

Microbiological analysis of NHH3 and NHH6 show that hydrogenotrophic methanogens, methanol and methylotrophic methanogens, and ANME assigned sequences are present within Nebo-Hemphill, in agreement with both active methanogenesis and AOM within the field (Extended Data Fig. 5). Clumped isotopologues of CH₄ within the Nebo-Hemphill Oil Field give apparent temperatures for Δ¹³CH₃D of \({78} _{-12}\)^{+14} °C and apparent temperatures for Δ¹²CH₂D₂ that are greater than expected for thermogenic CH₄ in the field (163 ± 18 °C, Fig. 3). They suggest that the CH₄ is likely to be dominantly thermogenic with some microbial contributions. Samples NHH1, NHH2 and NHH8 have apparent temperatures below that expected for thermogenic CH₄ and appear to be approaching low-temperature equilibrium in Δ¹³CH₃D, consistent with active methanogenesis and AOM cycling within the field.

Modelling of the Pannonian Basin

The unequivocal identification and quantification of CO₂-related microbial methanogenesis at Olla has globally significant implications, and warrants the reconsideration of reservoir processes at other sites worldwide. Re-examination of previously published datasets allows the elucidation of the significance of methanogenesis in different reservoir environments, especially in systems that have or had naturally high CO₂ concentrations, such as the Pannonian Basin. A significant proportion of the CO₂ found in natural gas fields within the Pannonian Basin is of mantle origin²⁰. The initial, CO₂/³He in the Szegholm North, Szegholm South, Ebes and Hajduszoboszlo gas fields was estimated to be $7.9 \pm 5.4 \times 10^9$, consistent with the European subcontinental lithospheric mantle range ($0.6\text{--}40 \times 10^9$) and the mid-ocean ridge basalt mantle range ($2 \pm 1 \times 10^9$)^{95,96,97} (Extended Data Fig. 7a). Since emplacement, significant CO₂ trapping in many fields was identified from lower observed CO₂/³He ratios (for example, the Szegholm North and South, Ebes and Hajduszoboszlo gas fields) and speculation that some of the CO₂ may have been converted into CH₄ (ref. 20), although no viable mechanism was identified.

Using an average geothermal gradient of 45 °C km⁻¹ the temperature within these fields is predicted to be between 40 °C and 105 °C (with temperatures being up to 20 °C cooler at the time of CO₂ emplacement)⁹⁸, and thus within the range of microbially driven methanogenesis. In addition, the presence of thermogenic hydrocarbons would also ensure H₂ availability, similar to the Olla and Nebo-Hemphill oil fields. From our work in Olla, we surmise that the Pannonian Basin may be analogous to Olla and that hydrogenotrophic microbial methanogenesis active within the field may provide the missing mechanism for CO₂ removal, which was previously unidentified, and point to this mechanism as an important subsurface process operating on systems naturally enriched in CO₂ as well as anthropogenically CO₂ injected oil and gas fields.

In the absence of supporting clumped isotopologue data, if we assume that the system is methanogenesis driven we can nevertheless construct similar methanogenesis–dissolution models using published data from the

Szegholm South Gas Field (where multiple CO₂ isotope data are available). By comparing the observed CO₂/³He ratios against the CO₂ concentration, a clear correlation emerges (Extended Data Fig. 7a). Attributing this to the effect of dissolution (a pre-requisite for microbial methanogenesis), it is possible to extrapolate back to an initial CO₂/³He end member value of 13.3×10^9 , consistent with previous estimates.

Our model assumes mantle-derived CO₂ in the field with a typical isotopic composition of $\delta^{13}\text{C}$ of CO₂ = $-5\text{\textperthousand}$ (Extended Data Fig. 7b). Taking this as a starting point, 16.4 ± 0.8 times more dissolution than methanogenesis is required in the Szegholm South Gas Field to produce the measured $\delta^{13}\text{C}$ of CO₂. As such, proportionally less methanogenesis, relative to the Olla Oil Field, would be expected here as the gas fields are not being constantly perturbed via water injection and therefore nutrient availability is probably much more limited. The presence of biogenic CH₄ (through CO₂ conversion) can also be clearly identified from the measured C₁/C_N in the Szegholm South Gas Field (0.946 ± 0.004), which is greater than what would be expected in a purely thermogenic system (0.909). Using the excess in the C₁/C_N ratios, we determine that $41 \pm 5\%$ of the CH₄ currently present in the Szegholm South Gas Field part of the basin is biogenic. Similar C₁/C_N excesses are seen in the other gas fields in the Pannonian Basin⁹⁹. In addition, other fields such as Szegholm North^{20,99}, Ebes^{20,99}, Hajduszoboszlo^{20,99}, Kismarja^{20,99}, Répcelak¹⁰⁰ and Mihály¹⁰⁰ (Pannonian Basin), Subei¹⁰¹ (China) and JM Brown Bassett⁹⁵ (Permian Basin) also show geochemical signatures consistent with the occurrence of methanogenesis. Thus, we conclude that although methanogenesis is not the dominant CO₂ loss mechanism, it may play a substantial role that has not previously been considered in many natural systems.

Data availability

The geochemical data that support the findings of this study are available in the NERC EDS National Geoscience Data Centre at <https://doi.org/10.5285/a4070f5d-2064-4caf-a82c-79a786d6af9e>. The

microbial SRA and biosample data can be found at <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA744568>. Source data are provided with this paper.

References

1. 1.

Leung, D. Y. C., Caramanna, G. & Maroto-Valer, M. M. An overview of current status of carbon dioxide capture and storage technologies. *Renew. Sustain. Energy Rev.* **39**, 426–443 (2014).

2. 2.

Bui, M. et al. Carbon capture and storage (CCS): the way forward. *Energy Environ. Sci.* **11**, 1062–1176 (2018).

3. 3.

Stewart, R. J. & Haszeldine, R. S. *Carbon Accounting for Carbon Dioxide Enhanced Oil Recovery* (Scottish Carbon Capture & Storage, 2014); <https://www.sccs.org.uk/images/expertise/misc/SCCS-CO2-EOR-JIP-Carbon-Balance.pdf>

4. 4.

Gilfillan, S. M. V. et al. Solubility trapping in formation water as dominant CO₂ sink in natural gas fields. *Nature* **458**, 614–618 (2009).

5. 5.

Matter, J. M. et al. Rapid carbon mineralization for permanent disposal of anthropogenic carbon dioxide emissions. *Science* **352**, 1312–1314 (2016)

6. 6.

Busch, A., Gensterblum, Y., Krooss, B. M. & Littke, R. Methane and carbon dioxide adsorption–diffusion experiments on coal: upscaling and modeling. *Int. J. Coal Geol.* **60**, 151–168 (2004).

7. 7.

Moffet, C. I. Post-economic Evaluation 2800 foot Sand Pilot Project Olla Field, LaSalle Parish Louisiana (1990).

8. 8.

Byrne, D. J., Barry, P. H., Lawson, M. & Ballentine, C. J. The use of noble gas isotopes to constrain subsurface fluid flow and hydrocarbon migration in the East Texas Basin. *Geochim. Cosmochim. Acta* **268**, 186–208 (2020).

9. 9.

McIntosh, J. C., Warwick, P. D., Martini, A. M. & Osborn, S. G. Coupled hydrology and biogeochemistry of Paleocene–Eocene coal beds, northern Gulf of Mexico. *Geol. Soc. Am. Bull.* **122**, 1248–1264 (2010).

10. 10.

Shelton, J. L., McIntosh, J. C., Warwick, P. D. & Yi, A. L. Z. Fate of injected CO₂ in the Wilcox Group, Louisiana, Gulf Coast Basin: chemical and isotopic tracers of microbial–brine–rock–CO₂ interactions. *Appl. Geochem.* **51**, 155–169 (2014).

11. 11.

Shelton, J. L., McIntosh, J. C., Warwick, P. D. & McCray, J. E. Impact of formation water geochemistry and crude oil biodegradation on microbial methanogenesis. *Org. Geochem.* **98**, 105–117 (2016).

12. 12.

Ballentine, C. J. & Burnard, P. G. Production, release and transport of noble gases in the continental crust. *Rev. Mineral. Geochem.* **47**, 481–538 (2002).

13. 13.

LaForce, T., Ennis-King, J., Boreham, C. & Paterson, L. Residual CO₂ saturation estimate using noble gas tracers in a single-well field test: the CO2CRC Otway project. *Int. J. Greenhouse Gas Control* **26**, 9–21 (2014).

14. 14.

Györe, D., Stuart, F. M., Gilfillan, S. M. V. & Waldron, S. Tracing injected CO₂ in the Cranfield enhanced oil recovery field (MS, USA) using He, Ne and Ar isotopes. *Int. J. Greenhouse Gas Control* **42**, 554–561 (2015).

15. 15.

Sathaye, K. J., Larson, T. E., Hesse, M. A. Noble gas fractionation during subsurface gas migration. *Earth Planet Sci. Lett.* **450**, 1–9 (2016).

16. 16.

Shelton, J. L. et al. Determining CO₂ storage potential during miscible CO₂ enhanced oil recovery: noble gas and stable isotope tracers. *Int. J. Greenhouse Gas Control* **51**, 239–253 (2016).

17. 17.

Barry, P. H. et al. Tracing enhanced oil recovery signatures in casing gases from the Lost Hills Oil Field using noble gases. *Earth Planet Sci. Lett.* **496**, 57–67 (2018).

18. 18.

Barry, P. H., Lawson, M., Meurer, W. P., Cheng, A. & Ballentine, C. J. Noble gases in deepwater oils of the U.S. Gulf of Mexico. *Geochim. Geophys. Geosyst.* **19**, 4218–4235 (2018).

19. 19.

Gilfillan, S. M. V. et al. The noble gas geochemistry of natural CO₂ gas reservoirs from the Colorado Plateau and Rocky Mountain provinces, USA. *Geochim. Cosmochim. Acta* **72**, 1174–1198 (2008).

20. 20.

Sherwood Lollar, B., Ballentine, C. J. & Onions, R. K. The fate of mantle-derived carbon in a continental sedimentary basin: integration of C-He relationships and stable isotope signatures. *Geochim. Cosmochim. Acta* **61**, 2295–2307 (1997).

21. 21.

Clark, I. D. & Fritz, P. *Environmental Isotopes in Hydrogeology* (CRC Press/Lewis Publishers, 1997).

22. 22.

Wycherley, H., Fleet, A. & Shaw, H. Some observations on the origins of large volumes of carbon dioxide accumulations in sedimentary basins. *Mar. Pet. Geol.* **16**, 489–494 (1999).

23. 23.

Bottinga, Y. Calculated fractionation factors for carbon and hydrogen isotope exchange in the system calcite–carbon dioxide–graphite–methane–hydrogen–water vapor. *Geochim. Cosmochim. Acta* **33**, 49–64 (1969).

24. 24.

Thiagarajan, N. et al. Isotopic evidence for quasi-equilibrium chemistry in thermally mature natural gases. *Proc. Natl Acad. Sci. USA* **117**, 3989–3995 (2020).

25. 25.

Stolper, D. A. et al. Formation temperatures of thermogenic and biogenic methane. *Science* **344**, 1500–1503 (2014).

26. 26.

Douglas, P. M. J. et al. Methane clumped isotopes: progress and potential for a new isotopic tracer. *Org. Geochem.* **113**, 262–282 (2017).

27. 27.

Stolper, D. A. et al. The utility of methane clumped isotopes to constrain the origins of methane in natural gas accumulations. *Geol. Soc. Lond. Spec. Publ.* **468**, 23–52 (2018).

28. 28.

Giunta, T. et al. Methane sources and sinks in continental sedimentary systems: new insights from paired clumped isotopologues $\Delta^1_3\text{CH}_3\text{D}$ and $\Delta^1_2\text{CH}_2\text{D}_2$. *Geochim. Cosmochim. Acta* **245**, 327–351 (2019).

29. 29.

Young, E. D. in *Deep Carbon* (eds Orcutt, B. N. et al.) 388–414 (Cambridge Univ. Press, 2019);
<https://doi.org/10.1017/9781108677950.013>

30. 30.

Stolper, D. A. et al. Distinguishing and understanding thermogenic and biogenic sources of methane using multiply substituted isotopologues. *Geochim. Cosmochim. Acta* **161**, 219–247 (2015).

31. 31.

Wang, D. T. et al. Nonequilibrium clumped isotope signals in microbial methane. *Science* **348**, 428–431 (2015).

32. 32.

Gruen, D. S. et al. Experimental investigation on the controls of clumped isotopologue and hydrogen isotope ratios in microbial methane. *Geochim. Cosmochim. Acta* **237**, 339–356 (2018).

33. 33.

Cao, X., Bao, H. & Peng, Y. A kinetic model for isotopologue signatures of methane generated by biotic and abiotic CO₂ methanation. *Geochim. Cosmochim. Acta* **249**, 59–75 (2019).

34. 34.

Ono, S. et al. Clumped isotopologue fractionation by microbial cultures performing the anaerobic oxidation of methane. *Geochim. Cosmochim. Acta* **293**, 70–85 (2021)

35. 35.

Sweeney, J. J. & Burnham, A. K. Evaluation of a simple model of vitrinite reflectance based on chemical kinetics. *AAPG Bull.* **74**, 1559–1570 (1990).

36. 36.

Head, I. M., Jones, D. M. & Larter, S. R. Biological activity in the deep subsurface and the origin of heavy oil. *Nature* **426**, 344–352 (2003).

37. 37.

Shelton, J. L. et al. Microbial community composition of a hydrocarbon reservoir 40 years after a CO₂ enhanced oil recovery flood. *FEMS Microbiol. Ecol.* **94**, fiy153 (2018).

38. 38.

Gray, N. D. et al. Biogenic methane production in formation waters from a large gas field in the North Sea. *Extremophiles* **13**, 511–519 (2009).

39. 39.

Perri, P. R. *Advanced Reservoir Characterization in the Antelope Shale to Establish the Viability of CO₂ Enhanced Oil Recovery in California's Monterey Siliceous Shales* Annual Report (Chevron USA Production Company, 2003).

40. 40.

Gannon, R. S. et al. *Produced Water Chemistry Data for the Lost Hills and Fruitvale Study Areas, Southern San Joaquin Valley, California, 2016* (US Geological Survey, 2020);
<https://doi.org/10.5066/F7X929H9>

41. 41.

Young, E. D. et al. The relative abundances of resolved Δ¹₂CH₂D₂ and Δ¹₃CH₃D and mechanisms controlling isotopic bond ordering in abiotic and biotic methane gases. *Geochim. Cosmochim. Acta* **203**, 235–264 (2017).

42. 42.

Van Breukelen, B. M. Extending the Rayleigh equation to allow competing isotope fractionating pathways to improve quantification of biodegradation. *Environ. Sci. Technol.* **41**, 4004–4010 (2007).

43. 43.

Byrne, D. J., Barry, P. H., Lawson, M. & Ballentine, C. J. Determining gas expulsion vs retention during hydrocarbon generation in the Eagle Ford Shale using noble gases. *Geochim. Cosmochim. Acta* **241**, 240–254 (2018).

44. 44.

Barry, P. H. et al. Noble gases solubility models of hydrocarbon charge mechanism in the Sleipner Vest gas field. *Geochim. Cosmochim. Acta* **194**, 291–309 (2016).

45. 45.

Tyne, R. L. et al. A novel method for the extraction, purification, and characterization of noble gases in produced fluids. *Geochem. Geophys. Geosyst.* **20**, 5588–5597 (2019).

46. 46.

Thiagarajan, N. et al. Identifying thermogenic and microbial methane in deep water Gulf of Mexico reservoirs. *Geochim. Cosmochim. Acta* **275**, 188–208 (2020).

47. 47.

Bolyen, E. et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).

48. 48.

Callahan, B. J. et al. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. Methods* **13**, 581–583 (2016).

49. 49.

Katoh, K., Misawa, K., Kuma, K. & Miyata, T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059–3066 (2002).

50. 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. 51.

Bokulich, N. A. et al. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* **6**, 90 (2018).

52. 52.

McDonald, D. et al. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J.* **6**, 610–618 (2012).

53. 53.

Tye, R. S., Moslow, T. F., Kimbrell, W. C. & Wheeler, C. W. Lithostratigraphy and production characteristics of the Wilcox Group (Paleocene–Eocene) in central Louisiana. *AAPG Bull.* **75**, 1675–1713 (1991).

54. 54.

Johnston, J. E. III, Heinrich, P. V., Lovelace, J. K., McCulloh, R. P. & Zimmerman, R. K. *Stratigraphic Charts of Louisiana* (2000).

55. 55.

Echols, D. J. & Malkin, D. S. Wilcox (Eocene) stratigraphy, a key to production. *AAPG Bull.* **32**, 11–33 (1948).

56. 56.

Heydari, E., Byerly, G. R. & Henry, D. J. Contact metamorphism and over maturation of organic matter associated with an igneous intrusion

in the Smackover Formation, Northeastern Louisiana. *Gulf Coast Assoc. Geol. Soc. Trans.* **47**, 201–213 (1997).

57. 57.

Warwick, P. D., Breland, F. C. & Hackley, P. C. Biogenic origin of coalbed gas in the northern Gulf of Mexico Coastal Plain, U.S.A. *Int. J. Coal Geology* **76**, 119–137 (2008).

58. 58.

Wenger, L. M., Sassen, R. & Shurmacher, D. S. in *Gulf Coast Oils and Gases: Their Characteristics, Origin, Distribution, and Exploration and Production Significance*(eds Schumacher, D. & Perkins, B F.) Ch. 6 (GCSSEPM, 1990).

59. 59.

Sassen, R. Lower Tertiary and Upper Cretaceous source rocks in Louisiana and Mississippi: implications to Gulf of Mexico crude oil. *AAPG Bull.* **74**, 857–878 (1990).

60. 60.

Sassen, R., Tye, R. S., Chinn, E. W. & Lemoine, R. C. Origin of crude oil in the Wilcox trend of Louisiana and Mississippi: evidence of long-range migration. *Gulf Coast Assoc. Geol. Soc. Trans.* **38**, 27–34 (1988).

61. 61.

Seifert, W. K. & Moldowan, J. M. in *Methods in Geochemistry and Geophysics* Vol. 24 (ed. Johns, R. B.) 261–290 (Elsevier, 1986).

62. 62.

van Aarssen, B. G. K., Bastow, T. P., Alexander, R. & Kagi, R. I. Distributions of methylated naphthalenes in crude oils: indicators of

maturity, biodegradation and mixing. *Org. Geochem.* **30**, 1213–1227 (1999).

63. 63.

Wenger, L. M., Davis, C. L. & Isaksen, G. H. Multiple controls on petroleum biodegradation and impact on oil quality. *SPE Reserv. Eval. Eng.* **5**, 375–383 (2002).

64. 64.

Kharaka, Y. K. & Hanor, J. S. Deep fluids in the continents: I. Sedimentary basins. *Treatise Geochem.* **5**, 499–540 (2003).

65. 65.

Day, J. M. D. et al. The helium flux from the continents and ubiquity of low-³He/⁴He recycled crust and lithosphere. *Geochim. Cosmochim. Acta* **153**, 116–133 (2015).

66. 66.

McCulloh, R. P. Little Creek structure, northwestern La Salle Parish, Louisiana. In *Geological Society of America South Central Section 50th Annual Meeting* Paper no. 15-3 (Geological Society of America, 2016); <https://doi.org/10.1130/abs/2016SC-271780>

67. 67.

Ballentine, C. J., Marty, B., Lollar, B. S. & Cassidy, M. Neon isotopes constrain convection and volatile origin in the Earth's mantle. *Nature* **433**, 33–38 (2005).

68. 68.

Kennedy, B. M., Hiyagon, H. & Reynolds, J. H. Crustal neon: a striking uniformity. *Earth Planet. Sci. Lett.* **98**, 277–286 (1990).

69. 69.

Clarke, W. B., Jenkins, W. J. & Top, Z. Determination of tritium by mass spectrometric measurement of ${}^3\text{He}$. *Int. J. Appl. Radiat. Isot.* **27**, 515–522 (1976).

70. 70.

Lee, J.-Y. et al. A redetermination of the isotopic abundances of atmospheric Ar. *Geochim. Cosmochim. Acta* **70**, 4507–4512 (2006).

71. 71.

Graham, D. W. Noble gas isotope geochemistry of mid-ocean ridge and ocean island basalts: characterization of mantle source reservoirs. *Rev. Mineral. Geochem.* **47**, 247–317 (2002).

72. 72.

Warwick, P. D., Breland, F. C. J., Ratchford, M. E., Ingram, S. L. S. & Pashin, J. C. Coal gas resource potential of Cretaceous and Paleogene coals of the eastern Gulf of Mexico Coastal Plain. In *Southeast Section Meeting* (Geological Society of America, 2004).

73. 73.

Prinzhofe, A. A. & Huc, A. Y. Genetic and post-genetic molecular and isotopic fractionations in natural gases. *Chem. Geol.* **126**, 281–290 (1995).

74. 74.

Prinzhofe, A. & Pernaton, É. Isotopically light methane in natural gas: bacterial imprint or diffusive fractionation? *Chem. Geol.* **142**, 193–200 (1997).

75. 75.

Fuex, A. N. Experimental evidence against an appreciable isotopic fractionation of methane during migration. *Phys. Chem. Earth* **12**, 725–732 (1980).

76. 76.

Zhang, T., Wang, X. & Chen, J. Chemical composition of gases as a geochemical tracer of natural gas migration. *Acta Sedimentol. Sinica* **17**, 627–632 (1999).

77. 77.

Xia, X. & Tang, Y. Isotope fractionation of methane during natural gas flow with coupled diffusion and adsorption/desorption. *Geochim. Cosmochim. Acta* **77**, 489–503 (2012).

78. 78.

Giggenbach, W. F. Carbon-13 exchange between CO₂ and CH₄ under geothermal conditions. *Geochim. Cosmochim. Acta* **46**, 159–165 (1982).

79. 79.

50. Ash, J. L. et al. Exchange catalysis during anaerobic methanotrophy revealed by $\Delta^{13}\text{CH}_3\text{D}$ and $\Delta^{12}\text{CH}_2\text{D}_2$ in methane. *Geochem. Persp. Lett.* **10**, 26–30 (2019).

80. 80.

Holler, T. et al. Carbon and sulfur back flux during anaerobic microbial oxidation of methane and coupled sulfate reduction. *Proc. Natl Acad. Sci. USA* **108**, E1484–E1490 (2011).

81. 81.

Yoshinaga, M. Y. et al. Carbon isotope equilibration during sulphate-limited anaerobic oxidation of methane. *Nat. Geosci* **7**, 190–194 (2014).

82. 82.

Doerfert, S. N., Reichlen, M., Iyer, P., Wang, M. & Ferry, J. G. *Methanolobus zinderi* sp. nov., a methylotrophic methanogen isolated from a deep subsurface coal seam. *Int. J. Syst. Evol. Microbiol.* **59**, 1064–1069 (2009).

83. 83.

Larter, S. R. et al. Biodegradation, gas destruction and methane generation in deep subsurface petroleum reservoirs: an overview. *Geol. Soc. Lond. Pet. Geol. Conf. Ser.* **6**, 633–639 (2005).

84. 84.

Zinder, S. H. in *Methanogenesis: Ecology, Physiology, Biochemistry & Genetics* (ed Ferry, J. G.) 128–206 (Springer, 1993);
https://doi.org/10.1007/978-1-4615-2391-8_4

85. 85.

Oren, A. Thermodynamic limits to microbial life at high salt concentrations. *Environ. Microbiol.* **13**, 1908–1923 (2011).

86. 86.

Capone, D. G. & Kiene, R. P. Comparison of microbial dynamics in marine and freshwater sediments: contrasts in anaerobic carbon catabolism. *Limnol. Oceanogr.* **33**, 725–749 (1988).

87. 87.

Lovley, D. R. & Klug, M. J. Sulfate reducers can outcompete methanogens at freshwater sulfate concentrations. *Appl. Environ. Microbiol.* **45**, 187–192 (1983).

88. 88.

Rice, D. D. & Claypool, G. E. Generation, accumulation, and resource potential of biogenic gas. *AAPG Bull.* **65**, 5–25 (1981).

89. 89.

van Bodegom, P. M., Scholten, J. C. M. & Stams, A. J. M. Direct inhibition of methanogenesis by ferric iron. *FEMS Microbiol. Ecol.* **49**, 261–268 (2004).

90. 90.

Mcintosh, J. C. & Martini, A. M. Hydrogeochemical indicators for microbial methane in fractured organic-rich shales: case studies of the Antrim, New Albany, and Ohio shales. *J. Geol.* **13**, 162–174 (2008).

91. 91.

Etiope, G. & Lollar, B. S. Abiotic methane on Earth. *Rev. Geophys.* **51**, 276–299 (2013).

92. 92.

Chung, H. M., Gormly, J. R. & Squires, R. M. Origin of gaseous hydrocarbons in subsurface environments: theoretical considerations of carbon isotope distribution. *Chem. Geol.* **71**, 97–104 (1988).

93. 93.

Whiticar, M. J., Faber, E. & Schoell, M. Biogenic methane formation in marine and freshwater environments: CO₂ reduction vs. acetate fermentation—isotope evidence. *Geochim. Cosmochim. Acta* **50**, 693–709 (1986).

94. 94.

Conrad, R. Quantification of methanogenic pathways using stable carbon isotopic signatures: a review and a proposal. *Org. Geochem.* **36**, 739–752 (2005).

95. 95.

Ballentine, C. J., Schoell, M., Coleman, D. & Cain, B. A. 300-Myr-old magmatic CO₂ in natural gas reservoirs of the west Texas Permian Basin. *Nature* **409**, 327–331 (2001).

96. 96.

Holland, G. & Gilfillan, S. in *The Noble Gases as Geochemical Tracers* (ed. Burnard, P.) 177–223 (Springer, 2013);
https://doi.org/10.1007/978-3-642-28836-4_8

97. 97.

Kämpf, H., Geissler, W. H. & Bäuer, K. in *Mantle Plumes: A Multidisciplinary Approach* (eds Ritter J.R.R. & Christensen U.R.) 127–158 (Springer, 2007).

98. 98.

Royden, L., Horváth, F., Nagymarosy, A. & Stegenga, L. Evolution of the Pannonian Basin system: 2. Subsidence and thermal history. *Tectonics* **2**, 91–137 (1983).

99. 99.

Sherwood Lollar, B., O’Nions, R. K. & Ballentine, C. J. Helium and neon isotope systematics in carbon dioxide-rich and hydrocarbon-rich gas reservoirs. *Geochim. Cosmochim. Acta* **58**, 5279–5290 (1994).

100. 100.

Palcsu, L. et al. In-reservoir mixing of mantle-derived CO₂ and metasedimentary CH₄–N₂ fluids—noble gas and stable isotope study of two multistacked fields (Pannonian Basin System, W-Hungary). *Mar. Pet. Geol.* **54**, 216–227 (2014).

101. 101.

Sheng, X., Nakai, S., Wakita, H., Yongchang, X. & Wang, X. Carbon isotopes of hydrocarbons and carbon dioxide in natural gases in China. *J. Asian Earth Sci.* **15**, 89–101 (1997).

Acknowledgements

R.L.T. was supported by a Natural Environment Research Council studentship (grant reference NE/L002612/1). C.J.B. and P.H.B. acknowledge A. Regberg and B. Meurer for their support of the project and help with sample collection. C.J.B. was part supported by an Earth4D CIFAR fellowship. P.H.B. was supported by NSF awards 1923915 and 2015789. O.W. was supported by Natural Sciences and Engineering Research Council of Canada Discovery and Accelerator grants awarded to the Sherwood Lollar research group and acknowledges B. Sherwood Lollar's support for the project. Z.M.S. acknowledges J. Biddle and G. Christman for their help in generating the microbial data.

Author information

Author notes

1. M. Lawson

Present address: Aker BP, Stavanger, Norway

Affiliations

1. Department of Earth Sciences, University of Oxford, Oxford, UK

R. L. Tyne, P. H. Barry, D. J. Hillegonds & C. J. Ballentine

2. Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, MA, USA

P. H. Barry

3. ExxonMobil Upstream Business Development, Spring, TX, USA

M. Lawson

4. CRPG-CNRS, Université de Lorraine, Nancy, France

D. J. Byrne

5. Department of Earth Sciences, University of Toronto, Toronto, Ontario, Canada

O. Warr

6. Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA, USA

H. Xie & J. M. Eiler

7. ExxonMobil Upstream Integrated Solutions, Spring, TX, USA

M. Formolo & B. Skinner

8. ExxonMobil Research and Engineering Co., Virginia, NJ, USA

Z. M. Summers

Contributions

The project was conceived by R.L.T., P.H.B., M.L. and C.J.B. R.L.T. performed the noble gas isotopic analysis, managed the project and prepared the first draft of the manuscript. P.H.B. collected the samples. Z.M.S., H.X. and B.S. conducted the microbial, clumped isotope and molecular geochemical analysis, respectively. Clumped isotope modelling was developed by H.X. and J.M.E. Integrated noble gas and stable isotope modelling was developed by R.L.T., P.H.B., M.L., C.J.B., O.W. and D.J.B. All authors contributed to the final manuscript.

Corresponding authors

Correspondence to [R. L. Tyne](#), [P. H. Barry](#), [M. Lawson](#) or [C. J. Ballentine](#).

Ethics declarations

Competing interests

M.F. and B.S. are employed by/an employee of ExxonMobil Upstream Integrated Solutions Company. Z.M.S. is employed by/an employee of ExxonMobil Research and Engineering Company. When the manuscript was first submitted, M.L. was employed by/an employee of ExxonMobil Upstream Integrated Solutions Company. M.L. is now employed by/an employee of Aker BP. The views expressed are those of the author(s) and not necessarily those of ExxonMobil Upstream Integrated Solutions Company, ExxonMobil Research and Engineering Company, or Aker BP.

Additional information

Peer review information *Nature* thanks Werner Aeschbach, Jenna Shelton and Marc Strous for their contribution to the peer review of this work. Peer reviewer reports are available.

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 Relationship between noble gas isotopic ratios in the Olla (blue circles) and Nebo-Hemphill (orange triangles) oil fields.

Both fields show a resolvable mantle component that is significantly greater in the Olla Oil Field. One standard deviation error bars are within symbol size. **a**, He isotopic ratios versus ${}^4\text{He}/{}^{20}\text{Ne}$ ratio. Air-corrected He isotopes are reported relative to air ($R_A = 1.38 \times 10^{-6}$ (ref. [69](#))). All samples are in excess of typical radiogenic production ($0.02R_A$), requiring a mantle contribution. High ${}^4\text{He}/{}^{20}\text{Ne}$ indicates negligible air contamination in the samples. **b**, Three neon isotope plot. The crustal production line^{[68](#)} reflects the observed trend in typical crustal fluids. The mantle-air line represents mixing between air and a MORB-like endmember^{[67](#)}. Data are consistent with mixing between air and a mantle-rich component. In both He and Ne isotopes, the mantle signal is stronger in the Olla Oil Field.

[Source data](#)

Extended Data Fig. 2 CO₂/³He variation vs.

a, b, ${}^4\text{He}$ (**a**) and ${}^{20}\text{Ne}$ (**b**) abundances in the Olla (blue circles) and Nebo-Hemphill (orange triangles) oil fields. 1σ level of uncertainty is shown on the measured samples. A negative correlation between ${}^4\text{He}$, which accumulates in formation waters as a function of time and ${}^{20}\text{Ne}$ (which is groundwater derived) and CO₂/³He, has been interpreted in other CO₂-rich natural gas fields^{[4](#)} to support the importance of the formation water in controlling CO₂ removal.

[Source data](#)

Extended Data Fig. 3 Hydrocarbon isotopic composition in the Olla (blue circles) and Nebo-Hemphill (orange triangles) oil fields.

a, Comparison of $\delta^{13}\text{C}$ -CH₄ and δD -CH₄, with typical CH₄ provenances after Whiticar⁹³. One standard deviation error bars are within symbol size. In isolation, the Olla samples (blue circles, this study; unfilled circles, from ref. ¹⁰) are consistent with a mainly thermogenic origin, whereas Nebo-Hemphill (orange triangles, this study; unfilled triangles, from ref. ¹⁰) appears to be a microbial-thermogenic mix. Clumped isotopologues nevertheless provide clear evidence that the Olla CH₄ is microbial in origin (Fig. 3). **b**, Plot of $\delta^{13}\text{C}$ of C_n versus 1/C_n variation in natural gases⁹². One standard deviation error bars are within symbol size. Both the Olla (blue circles) and Nebo-Hemphill (orange triangles) fields show an enrichment in their isotopic signature especially in propane; however, this is most pronounced within the Olla Oil Field.

Source data

Extended Data Fig. 4 Schematic of the processes occurring within the Olla Oil Field and resulting changes in the isotopic composition within the field.

The first panel shows CO₂ injection into the 2,800-ft sandstone (sst) and subsequent lateral and vertical migration (yellow arrows) to all the producing formations within the field. The middle panel demonstrate the different processes (microbial methanogenesis, AOM and dissolution), which are occurring in the field and their effects on the isotopic composition. Yellow circles represent CO₂ dissolution and green circles are the net CH₄ production from microbial activity. Hydrogen for microbial methanogenesis is sourced from the hydrocarbons and water. The final panel shows the current state of the reservoir.

Extended Data Fig. 5

Microbial community analysis of most abundant taxa from the SSU rRNA gene-sequence region. Bar charts represent the most resolved taxa from amplicon sequence variants generated in QIIME 2^{49,50,51} for archaeal primers (**a**) and bacterial primers (**b**). Hydrogenotrophic methanogenesis-capable taxa are represented with bold red font.

[Source data](#)

[Extended Data Fig. 6 Carbon isotopic composition of the Olla samples as a function of CO₂ concentration for data from this study \(dark blue circles\) and those collected without ³He analysis¹⁰ \(unfilled circles\).](#)

One standard deviation error bars are within symbol size. Dashed lines show endmember methanogenesis and dissolution fractionation (at pH 7) trajectories. Tick marks represent the total amount of CO₂ trapping within the system, relative to sample O5. The shaded region represents trapping by the combination of both microbial methanogenesis and dissolution. The same combination of dissolution and microbial methanogenesis has been modelled (grey region, cf. Fig. 2). Lines labelled ‘consumption’ at 5%, 13% and 19 % show the proportion of original CO₂ that has been removed by net microbial methanogenesis. One sample from ref. ¹⁰ has a CO₂ concentration greater than our ‘most pristine’ CO₂/³He sample (O5). This illustrates that by using sample O5 as our least altered composition we calculate a conservative (minimum) amount of CO₂ consumption, and that initial CO₂ concentrations were probably greater.

[Source data](#)

[Extended Data Fig. 7 Carbon composition of the Pannonian Basin samples.](#)

a, Szegholm South^{20,99} (red circles), Szegholm North^{20,99} (purple triangles), Ebcs^{20,99} (blue diamonds) and Hajduszoboszlo^{20,99} (grey squares) gas field

samples as a function of their CO₂ concentrations and CO₂/³He. Dashed lines represent the projection of CO₂ trapping in the fields. **b**, The Szegholm South field had multiple CO₂ isotope data and this field has been used to investigate the processes controlling this pattern using the carbon isotopic composition as a function of CO₂/³He. Dashed lines show endmember methanogenesis and dissolution fractionation (at pH 7) trajectories. Tick marks represent the total amount of CO₂ trapping within the system. The shaded region represents trapping by the combination of both microbial methanogenesis and dissolution. The upper and lower methanogenesis:dissolution ratios (M:D) are 0.050 and 0.060, respectively, showing dissolution accounts for approximately 16 times (M:D = 0.055) more CO₂ removal than microbial methanogenesis. Lines labelled ‘consumption’ show the proportion of original CO₂ that has been removed by net microbial methanogenesis.

Extended Data Table 1 Geochemistry and isotopic composition of the formation waters

Extended Data Table 2 Sample location, reservoir temperature, major gas species, carbon isotopes and clumped isotope data

Extended Data Table 3 Noble gas concentrations and ratios for the sampled wells (with 1 σ errors), as well as measured carbon isotopic compositions of alkanes

Supplementary information

[Peer Review File](#)

Source data

[Source Data Extended Data Fig. 1](#)

[Source Data Extended Data Fig. 2](#)

[Source Data Extended Data Fig. 3](#)

Source Data Extended Data Fig. 5

Source Data Extended Data Fig. 6

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

Tyne, R.L., Barry, P.H., Lawson, M. *et al.* Rapid microbial methanogenesis during CO₂ storage in hydrocarbon reservoirs. *Nature* **600**, 670–674 (2021). <https://doi.org/10.1038/s41586-021-04153-3>

- Received: 06 March 2021
- Accepted: 14 October 2021
- Published: 22 December 2021
- Issue Date: 23 December 2021

- DOI: <https://doi.org/10.1038/s41586-021-04153-3>

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

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04153-3>

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

- Article
- [Published: 09 December 2021](#)

The power of genetic diversity in genome-wide association studies of lipids

- [Sarah E. Graham](#) ORCID: orcid.org/0000-0003-1271-2489¹,
- [Shoa L. Clarke](#) ORCID: orcid.org/0000-0002-6592-1172^{2,3 na1},
- [Kuan-Han H. Wu](#) ORCID: orcid.org/0000-0003-4286-4299^{4 na1},
- [Stavroula Kanoni](#) ORCID: orcid.org/0000-0002-1691-9615^{5 na1},
- [Greg J. M. Zajac](#) ORCID: orcid.org/0000-0001-6411-9666^{6 na1},
- [Shweta Ramdas](#)^{7 na1},
- [Ida Surakka](#)¹,
- [Ioanna Ntalla](#)⁸,
- [Sailaja Vedantam](#)^{9,10},
- [Thomas W. Winkler](#) ORCID: orcid.org/0000-0003-0292-5421¹¹,
- [Adam E. Locke](#)¹²,
- [Eirini Marouli](#) ORCID: orcid.org/0000-0001-6179-1609⁵,
- [Mi Yeong Hwang](#) ORCID: orcid.org/0000-0002-8208-7925¹³,
- [Sohee Han](#) ORCID: orcid.org/0000-0003-2181-2131¹³,
- [Akira Narita](#) ORCID: orcid.org/0000-0002-9076-0710¹⁴,
- [Ananyo Choudhury](#) ORCID: orcid.org/0000-0001-8225-9531¹⁵,
- [Amy R. Bentley](#) ORCID: orcid.org/0000-0002-0827-9101¹⁶,
- [Kenneth Ekoru](#)¹⁶,
- [Anurag Verma](#)⁷,
- [Bhavi Trivedi](#)¹⁷,
- [Hilary C. Martin](#) ORCID: orcid.org/0000-0002-4454-9084¹⁸,
- [Karen A. Hunt](#)¹⁷,
- [Qin Hui](#)^{19,20},
- [Derek Klarin](#)^{21,22,23},
- [Xiang Zhu](#) ORCID: orcid.org/0000-0003-1134-6413^{2,24,25,26},
- [Gudmar Thorleifsson](#)²⁷,

- [Anna Helgadottir](#)²⁷,
- [Daniel F. Gudbjartsson](#) ORCID: orcid.org/0000-0002-5222-9857^{27,28},
- [Hilma Holm](#) ORCID: orcid.org/0000-0002-9517-6636²⁷,
- [Isleifur Olafsson](#)²⁹,
- [Masato Akiyama](#)^{30,31},
- [Saori Sakaue](#)^{30,32,33},
- [Chikashi Terao](#) ORCID: orcid.org/0000-0002-6452-4095³⁴,
- [Masahiro Kanai](#) ORCID: orcid.org/0000-0001-5165-4408^{23,30,35},
- [Wei Zhou](#) ORCID: orcid.org/0000-0001-7719-0859^{4,36,37},
- [Ben M. Brumpton](#) ORCID: orcid.org/0000-0002-3058-1059^{38,39,40},
- [Humaira Rasheed](#)^{38,39},
- [Sanni E. Ruotsalainen](#) ORCID: orcid.org/0000-0002-1631-4069⁴¹,
- [Aki S. Havulinna](#)^{41,42},
- [Yogasudha Veturi](#)⁴³,
- [QiPing Feng](#) ORCID: orcid.org/0000-0002-6213-793X⁴⁴,
- [Elisabeth A. Rosenthal](#)⁴⁵,
- [Todd Lingren](#)⁴⁶,
- [Jennifer Allen Pacheco](#) ORCID: orcid.org/0000-0001-8021-5818⁴⁷,
- [Sarah A. Pendergrass](#)⁴⁸,
- [Jeffrey Haessler](#)⁴⁹,
- [Franco Julianini](#)⁵⁰,
- [Yuki Bradford](#)⁴³,
- [Jason E. Miller](#)⁴³,
- [Archie Campbell](#) ORCID: orcid.org/0000-0003-0198-5078^{51,52},
- [Kuang Lin](#)⁵³,
- [Iona Y. Millwood](#)^{53,54},
- [George Hindy](#)⁵⁵,
- [Asif Rasheed](#)⁵⁶,
- [Jessica D. Faul](#)⁵⁷,
- [Wei Zhao](#) ORCID: orcid.org/0000-0001-7388-0692⁵⁸,
- [David R. Weir](#) ORCID: orcid.org/0000-0002-1661-2402⁵⁷,
- [Constance Turman](#)⁵⁹,
- [Hongyan Huang](#)⁵⁹,
- [Mariaelisa Graff](#) ORCID: orcid.org/0000-0001-6380-1735⁶⁰,

- [Anubha Mahajan⁶¹](#),
- [Michael R. Brown⁶²](#),
- [Weihua Zhang](#) ORCID: orcid.org/0000-0001-8279-3611^{63,64,65},
- [Ketian Yu](#) ORCID: orcid.org/0000-0001-9994-3399⁶⁶,
- [Ellen M. Schmidt](#) ORCID: orcid.org/0000-0003-1198-7536⁶⁶,
- [Anita Pandit⁶⁶](#),
- [Stefan Gustafsson](#) ORCID: orcid.org/0000-0001-5894-0351⁶⁷,
- [Xianyong Yin](#) ORCID: orcid.org/0000-0001-6454-2384⁶⁶,
- [Jian'an Luan](#) ORCID: orcid.org/0000-0003-3137-6337⁶⁸,
- [Jing-Hua Zhao⁶⁹](#),
- [Fumihiko Matsuda⁷⁰](#),
- [Hye-Mi Jang¹³](#),
- [Kyungheon Yoon¹³](#),
- [Carolina Medina-Gomez](#) ORCID: orcid.org/0000-0001-7999-5538^{71,72},
- [Achilleas Pitsillides⁷³](#),
- [Jouke Jan Hottenga](#) ORCID: orcid.org/0000-0002-5668-2368^{74,75},
- [Gonneke Willemsen^{74,76}](#),
- [Andrew R. Wood](#) ORCID: orcid.org/0000-0003-1726-948X⁷⁷,
- [Yingji Ji⁷⁷](#),
- [Zishan Gao](#) ORCID: orcid.org/0000-0003-1299-0286^{78,79,80},
- [Simon Haworth](#) ORCID: orcid.org/0000-0001-7793-7326^{81,82},
- [Ruth E. Mitchell](#) ORCID: orcid.org/0000-0002-3506-160X^{81,83},
- [Jin Fang Chai](#) ORCID: orcid.org/0000-0003-3770-1137⁸⁴,
- [Mette Aadahl⁸⁵](#),
- [Jie Yao⁸⁶](#),
- [Ani Manichaikul⁸⁷](#),
- [Helen R. Warren](#) ORCID: orcid.org/0000-0003-0511-1386^{88,89},
- [Julia Ramirez](#) ORCID: orcid.org/0000-0003-4130-5866⁸⁸,
- [Jette Bork-Jensen⁹⁰](#),
- [Line L. Kårhus⁸⁵](#),
- [Anuj Goel](#) ORCID: orcid.org/0000-0003-2307-4021^{61,91},
- [Maria Sabater-Lleal](#) ORCID: orcid.org/0000-0002-0128-379X^{92,93},
- [Raymond Noordam⁹⁴](#),

- [Carlo Sidore](#) ORCID: [orcid.org/0000-0001-7504-7477⁹⁵](http://orcid.org/0000-0001-7504-7477),
- [Edoardo Fiorillo](#)⁹⁶,
- [Aaron F. McDaid](#)^{97,98},
- [Pedro Marques-Vidal](#) ORCID: [orcid.org/0000-0002-4548-8500⁹⁹](http://orcid.org/0000-0002-4548-8500),
- [Matthias Wielscher](#)¹⁰⁰,
- [Stella Trompet](#)^{101,102},
- [Naveed Sattar](#) ORCID: [orcid.org/0000-0002-1604-2593¹⁰³](http://orcid.org/0000-0002-1604-2593),
- [Line T. Møllehave](#)⁸⁵,
- [Betina H. Thuesen](#)⁸⁵,
- [Matthias Munz](#) ORCID: [orcid.org/0000-0002-4728-3357^{104,105,106}](http://orcid.org/0000-0002-4728-3357),
- [Lingyao Zeng](#)^{107,108},
- [Jianfeng Huang](#)¹⁰⁹,
- [Bin Yang](#)¹⁰⁹,
- [Alaitz Poveda](#)¹¹⁰,
- [Azra Kurbasic](#)¹¹⁰,
- [Claudia Lamina](#)^{111,112},
- [Lukas Forer](#) ORCID: [orcid.org/0000-0003-2139-7329^{111,112}](http://orcid.org/0000-0003-2139-7329),
- [Markus Scholz](#)^{113,114},
- [Tessel E. Galesloot](#)¹¹⁵,
- [Jonathan P. Bradfield](#)¹¹⁶,
- [E. Warwick Daw](#)¹¹⁷,
- [Joseph M. Zmuda](#)¹¹⁸,
- [Jonathan S. Mitchell](#)¹¹⁹,
- [Christian Fuchsberger](#)¹¹⁹,
- [Henry Christensen](#)¹²⁰,
- [Jennifer A. Brody](#)¹²¹,
- [Mary F. Feitosa](#) ORCID: [orcid.org/0000-0002-0933-2410¹¹⁷](http://orcid.org/0000-0002-0933-2410),
- [Mary K. Wojczynski](#)¹¹⁷,
- [Michael Preuss](#) ORCID: [orcid.org/0000-0001-5266-8465¹²²](http://orcid.org/0000-0001-5266-8465),
- [Massimo Mangino](#) ORCID: [orcid.org/0000-0002-2167-7470^{123,124}](http://orcid.org/0000-0002-2167-7470),
- [Paraskevi Christofidou](#)¹²³,
- [Niek Verweij](#) ORCID: [orcid.org/0000-0002-4303-7685¹²⁵](http://orcid.org/0000-0002-4303-7685),
- [Jan W. Benjamins](#)¹²⁵,
- [Jorgen Engmann](#)^{126,127},

- [Rachel L. Kember](#) ORCID: [orcid.org/0000-0001-8820-2659¹²⁸](http://orcid.org/0000-0001-8820-2659),
- [Roderick C. Sliker](#) ORCID: [orcid.org/0000-0003-0961-9152^{129,130}](http://orcid.org/0000-0003-0961-9152),
- [Ken Sin Lo¹³¹](#),
- [Nuno R. Zilhao](#) ORCID: [orcid.org/0000-0002-4648-4288¹³²](http://orcid.org/0000-0002-4648-4288),
- [Phuong Le¹³³](#),
- [Marcus E. Kleber^{134,135}](#),
- [Graciela E. Delgado¹³⁴](#),
- [Shaofeng Huo³⁴¹](#),
- [Daisuke D. Ikeda](#) ORCID: [orcid.org/0000-0001-8403-0327¹³⁶](http://orcid.org/0000-0001-8403-0327),
- [Hiroyuki Iha](#) ORCID: [orcid.org/0000-0003-4261-2067¹³⁶](http://orcid.org/0000-0003-4261-2067),
- [Jian Yang^{137,138}](#),
- [Jun Liu](#) ORCID: [orcid.org/0000-0001-5288-3042¹³⁹](http://orcid.org/0000-0001-5288-3042),
- [Hampton L. Leonard^{140,141}](#),
- [Jonathan Marten¹⁴²](#),
- [Börge Schmidt¹⁴³](#),
- [Marina Arendt^{143,144}](#),
- [Laura J. Smyth](#) ORCID: [orcid.org/0000-0001-6419-4826¹⁴⁵](http://orcid.org/0000-0001-6419-4826),
- [Marisa Cañadas-Garre¹⁴⁵](#),
- [Chaolong Wang^{146,147}](#),
- [Masahiro Nakatuchi](#) ORCID: [orcid.org/0000-0002-1838-4837¹⁴⁸](http://orcid.org/0000-0002-1838-4837),
- [Andrew Wong](#) ORCID: [orcid.org/0000-0003-2079-4779¹⁴⁹](http://orcid.org/0000-0003-2079-4779),
- [Nina Hutil-Kähönen^{150,151}](#),
- [Xueling Sim](#) ORCID: [orcid.org/0000-0002-1233-7642⁸⁴](http://orcid.org/0000-0002-1233-7642),
- [Rui Xia¹⁵²](#),
- [Alicia Huerta-Chagoya¹⁵³](#),
- [Juan Carlos Fernandez-Lopez](#) ORCID: [orcid.org/0000-0003-3680-4193¹⁵⁴](http://orcid.org/0000-0003-3680-4193),
- [Valeriya Lyssenko^{155,156}](#),
- [Meraj Ahmed¹⁵⁷](#),
- [Anne U. Jackson](#) ORCID: [orcid.org/0000-0002-9672-2547⁶](http://orcid.org/0000-0002-9672-2547),
- [Marguerite R. Irvin¹⁵⁸](#),
- [Christopher Oldmeadow¹⁵⁹](#),
- [Han-Na Kim^{160,161}](#),
- [Seungho Ryu^{162,163}](#),

- [Paul R. H. J. Timmers](#) ORCID: orcid.org/0000-0002-5197-1267^{142,164},
- [Liubov Arbeeava](#) ORCID: orcid.org/0000-0001-5004-419X¹⁶⁵,
- [Rajkumar Dorajoo](#) ORCID: orcid.org/0000-0001-6608-2051¹⁴⁷,
- [Leslie A. Lange](#)¹⁶⁶,
- [Xiaoran Chai](#)^{167,168},
- [Gauri Prasad](#)^{169,170},
- [Laura Lorés-Motta](#)¹⁷¹,
- [Marc Pauper](#)¹⁷¹,
- [Jirong Long](#)¹⁷²,
- [Xiaohui Li](#)⁸⁶,
- [Elizabeth Theusch](#) ORCID: orcid.org/0000-0002-5678-759X¹⁷³,
- [Fumihiko Takeuchi](#)¹⁷⁴,
- [Cassandra N. Spracklen](#)^{175,176},
- [Anu Loukola](#)⁴¹,
- [Sailalitha Bollepalli](#)⁴¹,
- [Sophie C. Warner](#)^{177,178},
- [Ya Xing Wang](#)^{179,180},
- [Wen B. Wei](#)¹⁸⁰,
- [Teresa Nutile](#)¹⁸¹,
- [Daniela Ruggiero](#)^{181,182},
- [Yun Ju Sung](#)¹⁸³,
- [Yi-Jen Hung](#)¹⁸⁴,
- [Shufeng Chen](#)¹⁰⁹,
- [Fangchao Liu](#)¹⁰⁹,
- [Jingyun Yang](#) ORCID: orcid.org/0000-0002-3495-3710^{185,186},
- [Katherine A. Kentistou](#)¹⁶⁴,
- [Mathias Gorski](#) ORCID: orcid.org/0000-0002-9103-5860^{11,187},
- [Marco Brumat](#)¹⁸⁸,
- [Karina Meidtner](#)^{189,190},
- [Lawrence F. Bielak](#) ORCID: orcid.org/0000-0002-3443-8030¹⁹¹,
- [Jennifer A. Smith](#) ORCID: orcid.org/0000-0002-3575-5468^{57,191},
- [Prashantha Hebbar](#) ORCID: orcid.org/0000-0001-6560-7063¹⁹²,
- [Aliki-Eleni Farmaki](#)^{193,194},

- [Edith Hofer](#)^{195,196},
- [Maoxuan Lin](#) [ORCID: orcid.org/0000-0001-5191-1857](#)¹⁹⁷,
- [Chao Xue](#)¹,
- [Jifeng Zhang](#)¹,
- [Maria Pina Concas](#) [ORCID: orcid.org/0000-0003-3598-2537](#)¹⁹⁸,
- [Simona Vaccargiu](#)¹⁹⁹,
- [Peter J. van der Most](#) [ORCID: orcid.org/0000-0001-8450-3518](#)²⁰⁰,
- [Niina Pitkänen](#)^{201,202},
- [Brian E. Cade](#) [ORCID: orcid.org/0000-0003-1424-0673](#)^{203,204},
- [Jiwon Lee](#)²⁰³,
- [Sander W. van der Laan](#) [ORCID: orcid.org/0000-0001-6888-1404](#)²⁰⁵,
- [Kumaraswamy Naidu Chitrala](#) [ORCID: orcid.org/0000-0003-0663-9529](#)²⁰⁶,
- [Stefan Weiss](#) [ORCID: orcid.org/0000-0002-3553-4315](#)²⁰⁷,
- [Martina E. Zimmermann](#) [ORCID: orcid.org/0000-0002-6916-4404](#)¹¹,
- [Jong Young Lee](#) [ORCID: orcid.org/0000-0002-0092-9958](#)²⁰⁸,
- [Hyeok Sun Choi](#)²⁰⁹,
- [Maria Nethander](#)^{210,211},
- [Sandra Freitag-Wolf](#)²¹²,
- [Lorraine Southam](#)^{213,214},
- [Nigel W. Rayner](#)^{18,61,213,215},
- [Carol A. Wang](#) [ORCID: orcid.org/0000-0002-4301-3974](#)²¹⁶,
- [Shih-Yi Lin](#)^{217,218,219},
- [Jun-Sing Wang](#)^{220,221},
- [Christian Couture](#)²²²,
- [Leo-Pekka Lyytikäinen](#)^{223,224},
- [Kjell Nikus](#) [ORCID: orcid.org/0000-0002-9345-9851](#)^{225,226},
- [Gabriel Cuellar-Partida](#) [ORCID: orcid.org/0000-0001-7648-4097](#)²²⁷,
- [Henrik Vestergaard](#) [ORCID: orcid.org/0000-0003-3090-269X](#)^{90,228},
- [Bertha Hidalgo](#)²²⁹,
- [Olga Giannakopoulou](#)⁵,
- [Qiuyin Cai](#)¹⁷²,
- [Morgan O. Obura](#)¹²⁹,
- [Jessica van Setten](#) [ORCID: orcid.org/0000-0002-4934-7510](#)²³⁰,

- [Xiaoyin Li²³¹](#),
- [Karen Schwander¹¹⁷](#),
- [Natalie Terzikhan⁷²](#),
- [Jae Hun Shin²⁰⁹](#),
- [Rebecca D. Jackson²³²](#),
- [Alexander P. Reiner ORCID: orcid.org/0000-0002-1427-4470²³³](#),
- [Lisa Warsinger Martin ORCID: orcid.org/0000-0003-4352-0914²³⁴](#),
- [Zhengming Chen^{53,54}](#),
- [Liming Li²³⁵](#),
- [Heather M. Highland ORCID: orcid.org/0000-0002-3583-8239⁶⁰](#),
- [Kristin L. Young ORCID: orcid.org/0000-0003-0070-6145⁶⁰](#),
- [Takahisa Kawaguchi⁷⁰](#),
- [Joachim Thiery^{114,236}](#),
- [Joshua C. Bis ORCID: orcid.org/0000-0002-3409-1110¹²¹](#),
- [Girish N. Nadkarni¹²²](#),
- [Lenore J. Launer ORCID: orcid.org/0000-0002-3238-7612²³⁷](#),
- [Huaxing Li³⁴¹](#),
- [Mike A. Nalls^{140,141}](#),
- [Olli T. Raitakari^{201,202,238}](#),
- [Sahoko Ichihara²³⁹](#),
- [Sarah H. Wild ORCID: orcid.org/0000-0001-7824-2569²⁴⁰](#),
- [Christopher P. Nelson^{177,178}](#),
- [Harry Campbell¹⁶⁴](#),
- [Susanne Jäger ORCID: orcid.org/0000-0001-6619-0861^{189,190}](#),
- [Toru Nabika²⁴¹](#),
- [Fahd Al-Mulla¹⁹²](#),
- [Harri Niinikoski^{242,243}](#),
- [Peter S. Braund ORCID: orcid.org/0000-0001-8540-5709^{177,178}](#),
- [Ivana Kolcic²⁴⁴](#),
- [Peter Kovacs ORCID: orcid.org/0000-0002-0290-5423²⁴⁵](#),
- [Tota Giardoglou²⁴⁶](#),
- [Tomohiro Katsuya^{247,248}](#),
- [Konain Fatima Bhatti⁵](#),
- [Dominique de Kleijn²⁴⁹](#),

- Gert J. de Borst²⁴⁹,
- Eung Kweon Kim²⁵⁰,
- Hieab H. H. Adams^{72,251},

- [M. Arfan Ikram](#)⁷²,
- [Xiaofeng Zhu](#) ORCID: orcid.org/0000-0003-0037-411X²³¹,
- [Folkert W. Asselbergs](#) ORCID: orcid.org/0000-0002-1692-8669²³⁰,
- [Adriaan O. Kraaijeveld](#)²³⁰,
- [Joline W. J. Beulens](#)^{129,252},
- [Xiao-Ou Shu](#)¹⁷²,
- [Loukianos S. Rallidis](#)²⁵³,
- [Oluf Pedersen](#)⁹⁰,
- [Torben Hansen](#)⁹⁰,
- [Paul Mitchell](#)²⁵⁴,
- [Alex W. Hewitt](#) ORCID: orcid.org/0000-0002-5123-5999^{255,256},
- [Mika Kähönen](#)^{257,258},
- [Louis Pérusse](#) ORCID: orcid.org/0000-0001-6440-9698^{222,259},
- [Claude Bouchard](#)²⁶⁰,
- [Anke Tönjes](#)²⁴⁵,
- [Yii-Der Ida Chen](#)⁸⁶,
- [Craig E. Pennell](#) ORCID: orcid.org/0000-0002-0937-6165²¹⁶,
- [Trevor A. Mori](#) ORCID: orcid.org/0000-0002-5264-9229²⁶¹,
- [Wolfgang Lieb](#) ORCID: orcid.org/0000-0003-2544-4460²⁶²,
- [Andre Franke](#)²⁶³,
- [Claes Ohlsson](#)^{210,264},
- [Dan Mellström](#)^{210,265},
- [Yoon Shin Cho](#)²⁰⁹,
- [Hyejin Lee](#)²⁶⁶,
- [Jian-Min Yuan](#)^{267,268},
- [Woon-Puay Koh](#)^{269,270},
- [Sang Youl Rhee](#) ORCID: orcid.org/0000-0003-0119-5818²⁷¹,
- [Jeong-Taek Woo](#)²⁷¹,
- [Iris M. Heid](#)¹¹,
- [Klaus J. Stark](#)¹¹,
- [Henry Völzke](#)²⁷²,
- [Georg Homuth](#)²⁰⁷,
- [Michele K. Evans](#)²⁷³,

- [Alan B. Zonderman](#) ORCID: orcid.org/0000-0002-6523-4778²⁷³,
- [Ozren Polasek](#)²⁴⁴,
- [Gerard Pasterkamp](#)²⁰⁵,
- [Imo E. Hoefer](#)²⁰⁵,
- [Susan Redline](#)^{203,204},
- [Katja Pahkala](#)^{201,202,274},
- [Albertine J. Oldehinkel](#) ORCID: orcid.org/0000-0003-3925-3913²⁷⁵,
- [Harold Snieder](#) ORCID: orcid.org/0000-0003-1949-2298²⁰⁰,
- [Ginevra Biino](#)²⁷⁶,
- [Reinhold Schmidt](#)¹⁹⁵,
- [Helena Schmidt](#)²⁷⁷,
- [Y. Eugene Chen](#)¹,
- [Stefania Bandinelli](#)²⁷⁸,
- [George Dedoussis](#)¹⁹³,
- [Thangavel Alphonse Thanaraj](#) ORCID: orcid.org/0000-0002-4574-5343¹⁹²,
- [Sharon L. R. Kardia](#)¹⁹¹,
- [Norihiro Kato](#) ORCID: orcid.org/0000-0002-2433-8415¹⁷⁴,
- [Matthias B. Schulze](#)^{189,190,279},
- [Giorgia Girotto](#)^{188,280},
- [Bettina Jung](#)¹⁸⁷,
- [Carsten A. Böger](#)^{187,281,282},
- [Peter K. Joshi](#)¹⁶⁴,
- [David A. Bennett](#)^{185,186},
- [Philip L. De Jager](#) ORCID: orcid.org/0000-0002-8057-2505^{283,284},
- [Xiangfeng Lu](#)¹⁰⁹,
- [Vasiliki Mamakou](#)^{285,286},
- [Morris Brown](#)^{89,287},
- [Mark J. Caulfield](#) ORCID: orcid.org/0000-0001-9295-3594^{88,89},
- [Patricia B. Munroe](#)^{88,89},
- [Xiuqing Guo](#)⁸⁶,
- [Marina Ciullo](#)^{181,182},
- [Jost B. Jonas](#) ORCID: orcid.org/0000-0003-2972-5227^{288,289,290},
- [Nilesh J. Samani](#)^{177,178},

- [Jaakko Kaprio](#) ORCID: orcid.org/0000-0002-3716-2455⁴¹,
- [Päivi Pajukanta](#)²⁹³,
- [Linda S. Adair](#)^{299,300},
- [Sonny Augustin Bechayda](#)^{301,302},
- [H. Janaka de Silva](#)³⁰³,
- [Ananda R. Wickremasinghe](#)³⁰⁴,
- [Ronald M. Krauss](#)³⁰⁵,
- [Jer-Yuarn Wu](#)³⁰⁶,
- [Wei Zheng](#)¹⁷²,
- [Anneke I. den Hollander](#)¹⁷¹,
- [Dwaipayan Bharadwaj](#)^{170,307},
- [Adolfo Correa](#)³⁰⁸,
- [James G. Wilson](#) ORCID: orcid.org/0000-0001-5751-9178³⁰⁹,
- [Lars Lind](#)³¹⁰,
- [Chew-Kiat Heng](#) ORCID: orcid.org/0000-0002-7309-9473^{311,312},
- [Amanda E. Nelson](#)^{165,313},
- [Yvonne M. Golightly](#)^{165,314,315,316},
- [James F. Wilson](#)^{142,164},
- [Brenda Penninx](#)^{317,318},
- [Hyung-Lae Kim](#) ORCID: orcid.org/0000-0002-3260-8884³¹⁹,
- [John Attia](#)^{159,320},
- [Rodney J. Scott](#) ORCID: orcid.org/0000-0001-7724-3404^{159,320},
- [D. C. Rao](#)³²¹,
- [Donna K. Arnett](#)³²²,
- [Mark Walker](#)³²³,
- [Heikki A. Koistinen](#) ORCID: orcid.org/0000-0001-7870-070X^{324,325,326},
- [Giriraj R. Chandak](#) ORCID: orcid.org/0000-0002-3095-9453^{157,327},
- [Chittaranjan S. Yajnik](#)³²⁸,
- [Josep M. Mercader](#) ORCID: orcid.org/0000-0001-8494-3660^{292,329,330},
- [Teresa Tusié-Luna](#)^{294,295,296},
- [Carlos A. Aguilar-Salinas](#)^{297,298},
- [Clicerio Gonzalez Villalpando](#)³³¹,

- [Lorena Orozco](#) ORCID: orcid.org/0000-0002-5801-9180³³²,
- [Myriam Fornage](#)^{152,333},
- [E. Shyong Tai](#)^{84,334},
- [Rob M. van Dam](#)^{84,334},
- [Terho Lehtimäki](#)^{223,224},
- [Nish Chaturvedi](#)¹⁴⁹,
- [Mitsuhiko Yokota](#)³³⁵,
- [Jianjun Liu](#) ORCID: orcid.org/0000-0002-3255-3019¹⁴⁷,
- [Dermot F. Reilly](#)³³⁶,
- [Amy Jayne McKnight](#) ORCID: orcid.org/0000-0002-7482-709X¹⁴⁵,
- [Frank Kee](#)¹⁴⁵,
- [Karl-Heinz Jöckel](#)¹⁴³,
- [Mark I. McCarthy](#)^{48,61,337},
- [Colin N. A. Palmer](#)³³⁸,
- [Veronique Vitart](#)¹⁴²,
- [Caroline Hayward](#) ORCID: orcid.org/0000-0002-9405-9550¹⁴²,
- [Eleanor Simonsick](#)³³⁹,
- [Cornelia M. van Duijn](#)¹³⁹,
- [Fan Lu](#)³⁴⁰,
- [Jia Qu](#)³⁴⁰,
- [Haretsugu Hishigaki](#)¹³⁶,
- [Xu Lin](#)³⁴¹,
- [Winfried März](#)^{134,342,343},
- [Esteban J. Parra](#) ORCID: orcid.org/0000-0002-2057-8577¹³³,
- [Miguel Cruz](#)³⁴⁴,
- [Vilmundur Gudnason](#)^{132,345},
- [Jean-Claude Tardif](#)^{131,346},
- [Guillaume Lettre](#)^{131,347},
- [Leen M. 't Hart](#)^{129,130,348},
- [Petra J. M. Elders](#)³⁴⁹,
- [Scott M. Damrauer](#)^{352,353},
- [Meena Kumari](#)³⁵⁴,
- [Mika Kivimaki](#) ORCID: orcid.org/0000-0002-4699-5627¹²⁷,
- [Pim van der Harst](#) ORCID: orcid.org/0000-0002-2713-686X¹²⁵,

- [Tim D. Spector](#) ORCID: orcid.org/0000-0002-9795-0365¹²³,
- [Ruth J. F. Loos](#) ORCID: orcid.org/0000-0002-8532-5087^{122,355},
- [Michael A. Province](#) ORCID: orcid.org/0000-0001-6102-774X¹¹⁷,
- [Bruce M. Psaty](#) ORCID: orcid.org/0000-0002-7278-2190^{356,357},
- [Ivan Brandslund](#)^{120,358},
- [Peter P. Pramstaller](#)¹¹⁹,
- [Kaare Christensen](#) ORCID: orcid.org/0000-0002-5429-5292³⁵⁹,
- [Samuli Ripatti](#) ORCID: orcid.org/0000-0002-0504-1202^{41,360,361},
- [Elisabeth Widén](#)⁴¹,
- [Hakon Hakonarson](#) ORCID: orcid.org/0000-0003-2814-7461^{362,363},
- [Struan F. A. Grant](#)^{350,363,364},
- [Lambertus A. L. M. Kiemeney](#)¹¹⁵,
- [Jacqueline de Graaf](#)¹¹⁵,
- [Markus Loeffler](#)^{113,114},
- [Florian Kronenberg](#) ORCID: orcid.org/0000-0003-2229-1120^{112,365},
- [Dongfeng Gu](#)^{109,366},
- [Jeanette Erdmann](#) ORCID: orcid.org/0000-0002-4486-6231³⁶⁷,
- [Heribert Schunkert](#) ORCID: orcid.org/0000-0001-6428-3001^{368,369},
- [Paul W. Franks](#) ORCID: orcid.org/0000-0002-0520-7604¹¹⁰,
- [Allan Linneberg](#) ORCID: orcid.org/0000-0002-0994-0184^{85,370},
- [J. Wouter Jukema](#) ORCID: orcid.org/0000-0002-3246-8359^{101,371},
- [Amit V. Khera](#)^{372,373,374,375},
- [Minna Männikkö](#)³⁷⁶,
- [Marjo-Riitta Jarvelin](#)^{100,377,378},
- [Zoltan Katalik](#)^{98,379},
- [Francesco Cucca](#)^{380,381},
- [Dennis O. Mook-Kanamori](#)^{382,383},
- [Ko Willems van Dijk](#)^{384,385,386},
- [Hugh Watkins](#)^{61,387},
- [David P. Strachan](#)³⁸⁸,
- [Niels Grarup](#) ORCID: orcid.org/0000-0001-5526-1070⁹⁰,
- [Peter Sever](#)³⁸⁹,
- [Neil Poulter](#)³⁹⁰,
- [Jerome I. Rotter](#)⁸⁶,

- [Thomas M. Dantoft](#)⁸⁵,
- [Fredrik Karpe](#) ORCID: orcid.org/0000-0002-2751-1770^{391,392},
- [Matt J. Neville](#)^{391,392},
- [Nicholas J. Timpson](#)^{81,83},
- [Ching-Yu Cheng](#)^{167,393},
- [Tien-Yin Wong](#)^{167,393},
- [Chiea Chuen Khor](#)¹⁴⁷,
- [Charumathi Sabanayagam](#) ORCID: orcid.org/0000-0002-4042-4719^{167,393},
- [Annette Peters](#) ORCID: orcid.org/0000-0001-6645-0985^{80,190,394},
- [Christian Gieger](#)^{79,80,190},
- [Andrew T. Hattersley](#) ORCID: orcid.org/0000-0001-5620-473X³⁹⁵,
- [Nancy L. Pedersen](#)³⁹⁶,
- [Patrik K. E. Magnusson](#)³⁹⁶,
- [Dorret I. Boomsma](#)^{74,397,398},
- [Eco J. C. de Geus](#) ORCID: orcid.org/0000-0001-6022-2666^{74,318},
- [L. Adrienne Cupples](#)^{73,399},
- [Joyce B. J. van Meurs](#)^{71,72},
- [Mohsen Ghanbari](#)^{72,400},
- [Penny Gordon-Larsen](#) ORCID: orcid.org/0000-0001-5322-4188^{299,300},
- [Wei Huang](#)⁴⁰¹,
- [Young Jin Kim](#) ORCID: orcid.org/0000-0002-4132-4437¹³,
- [Yasuharu Tabara](#)⁷⁰,
- [Nicholas J. Wareham](#)⁶⁸,
- [Claudia Langenberg](#)⁶⁸,
- [Eleftheria Zeggini](#) ORCID: orcid.org/0000-0003-4238-659X^{213,214,402},
- [Johanna Kuusisto](#)⁴⁰³,
- [Markku Laakso](#) ORCID: orcid.org/0000-0002-3394-7749⁴⁰³,
- [Erik Ingelsson](#)^{3,67,404,405},
- [Goncalo Abecasis](#)^{66,406},
- [John C. Chambers](#)^{63,64,407,408},
- [Jaspal S. Kooner](#)^{64,65,409,410},

- [Paul S. de Vries](#)⁶²,
- [Alanna C. Morrison](#) ORCID: [orcid.org/0000-0001-6381-4296](#)⁶²,
- [Kari E. North](#)⁶⁰,
- [Martha Daviglus](#)⁴¹¹,
- [Peter Kraft](#)^{59,412},
- [Nicholas G. Martin](#) ORCID: [orcid.org/0000-0003-4069-8020](#)⁴¹³,
- [John B. Whitfield](#) ORCID: [orcid.org/0000-0002-1103-0876](#)⁴¹³,
- [Shahid Abbas](#)^{56,414},
- [Danish Saleheen](#)^{56,415,416},
- [Robin G. Walters](#) ORCID: [orcid.org/0000-0002-9179-0321](#)^{53,54,417},
- [Michael V. Holmes](#)^{53,54,418},
- [Corri Black](#)⁴¹⁹,
- [Blair H. Smith](#) ORCID: [orcid.org/0000-0002-5362-9430](#)⁴²⁰,
- [Anne E. Justice](#)⁴²¹,
- [Aris Baras](#)⁴⁰⁶,
- [Julie E. Buring](#)^{50,292},
- [Paul M. Ridker](#)^{50,292},
- [Daniel I. Chasman](#)^{50,291},
- [Charles Kooperberg](#)⁴⁹,
- [Wei-Qi Wei](#)⁴²²,
- [Gail P. Jarvik](#)⁴²³,
- [Bahram Namjou](#)⁴²⁴,
- [M. Geoffrey Hayes](#)^{425,426,427},
- [Marylyn D. Ritchie](#)⁴³,
- [Pekka Jousilahti](#)⁴²,
- [Veikko Salomaa](#) ORCID: [orcid.org/0000-0001-7563-5324](#)⁴²,
- [Kristian Hveem](#)^{38,428,429},
- [Bjørn Olav Åsvold](#) ORCID: [orcid.org/0000-0003-3837-2101](#)^{38,428,430},
- [Michiaki Kubo](#)⁴³¹,
- [Yoichiro Kamatani](#)^{30,432},
- [Yukinori Okada](#) ORCID: [orcid.org/0000-0002-0311-8472](#)^{30,32,433,434},
- [Yoshinori Murakami](#)⁴³⁵,
- [Unnur Thorsteinsdottir](#)^{27,345},
- [Kari Stefansson](#)^{27,345},

- Yuk-Lam Ho⁴³⁶,
- Julie A. Lynch^{437,438},
- Daniel J. Rader^{350,351},
- Philip S. Tsao^{2,3,439},
- Kyong-Mi Chang^{351,440},
- Kelly Cho^{436,441},
- Christopher J. O'Donnell^{436,441},
- John M. Gaziano^{436,441},
- Peter Wilson^{442,443},
- Charles N. Rotimi ORCID: orcid.org/0000-0001-5759-053X¹⁶,
- Scott Hazelhurst^{15,444},
- Michèle Ramsay ORCID: orcid.org/0000-0002-4156-4801^{15,445},
- Richard C. Trembath⁴⁴⁶,
- David A. van Heel ORCID: orcid.org/0000-0002-0637-2265¹⁷,
- Gen Tamiya¹⁴,
- Masayuki Yamamoto ORCID: orcid.org/0000-0002-9073-9436¹⁴,
- Bong-Jo Kim¹³,
- Karen L. Mohlke ORCID: orcid.org/0000-0001-6721-153X¹⁷⁵,
- Timothy M. Frayling⁷⁷,
- Joel N. Hirschhorn ORCID: orcid.org/0000-0002-1650-7901^{9,10,447},
- Sekar Kathiresan^{373,375,448},
- VA Million Veteran Program,
- Global Lipids Genetics Consortium*,
- Michael Boehnke ORCID: orcid.org/0000-0002-6442-7754⁶,
- Pradeep Natarajan^{37,449,450,451},
- Gina M. Peloso ORCID: orcid.org/0000-0002-5355-8636^{73 na2},
- Christopher D. Brown^{7 na2},
- Andrew P. Morris ORCID: orcid.org/0000-0002-6805-6014^{452 na2},
- Themistocles L. Assimes ORCID: orcid.org/0000-0003-2349-0009^{2,3,439 na2},
- Panos Deloukas ORCID: orcid.org/0000-0001-9251-070X^{5,89,453 na2},
- Yan V. Sun ORCID: orcid.org/0000-0002-2838-1824^{19,20 na2} &
- Cristen J. Willer ORCID: orcid.org/0000-0001-5645-4966^{1,4,453,454 na2}

Nature volume **600**, pages 675–679 (2021)

- 11k Accesses

- 244 Altmetric
- [Metrics details](#)

Subjects

- [Computational models](#)
- [Genome-wide association studies](#)
- [Preventive medicine](#)
- [Risk factors](#)

Abstract

Increased blood lipid levels are heritable risk factors of cardiovascular disease with varied prevalence worldwide owing to different dietary patterns and medication use¹. Despite advances in prevention and treatment, in particular through reducing low-density lipoprotein cholesterol levels², heart disease remains the leading cause of death worldwide³. Genome-wide association studies (GWAS) of blood lipid levels have led to important biological and clinical insights, as well as new drug targets, for cardiovascular disease. However, most previous GWAS^{4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23} have been conducted in European ancestry populations and may have missed genetic variants that contribute to lipid-level variation in other ancestry groups. These include differences in allele frequencies, effect sizes and linkage-disequilibrium patterns²⁴. Here we conduct a multi-ancestry, genome-wide genetic discovery meta-analysis of lipid levels in approximately 1.65 million individuals, including 350,000 of non-European ancestries. We quantify the gain in studying non-European ancestries and provide evidence to support the expansion of recruitment of additional ancestries, even with relatively small sample sizes. We find that increasing diversity rather than studying additional individuals of European ancestry results in substantial improvements in fine-mapping functional variants and portability of polygenic prediction (evaluated in approximately 295,000 individuals from 7 ancestry groupings). Modest gains in the number of discovered loci and ancestry-specific variants were also achieved. As GWAS expand emphasis beyond the identification of genes and fundamental biology towards the use of genetic variants for preventive and precision medicine²⁵, we anticipate that increased diversity of participants will lead to more accurate and equitable²⁶ application of polygenic scores in clinical practice.

This is a preview of subscription content

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Fig. 1: Comparison of identified loci across ancestry groups.



Fig. 2: Inclusion of multiple ancestries drives improved fine-mapping.

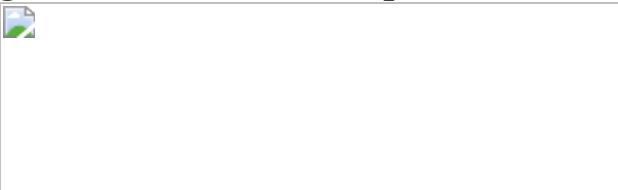
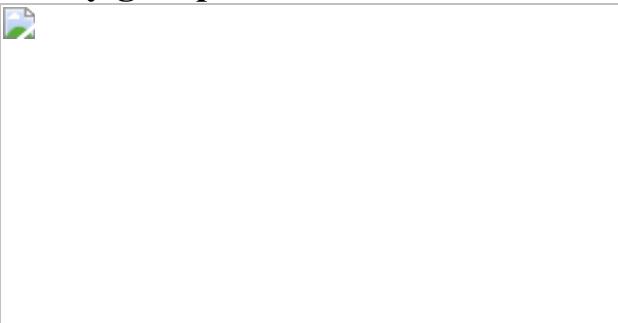


Fig. 3: Multi-ancestry LDL-C PRS show similar performance across ancestry groups.



Data availability

The GWAS meta-analysis results (including both ancestry-specific and multi-ancestry analyses) and risk score weights are available at <http://csg.sph.umich.edu/willer/public/glgc-lipids2021>. The optimized multi-ancestry and single-ancestry PRS weights are deposited in the PGS

Catalogue (<https://www.pgscatalog.org/>) accession numbers [PGS000886](#)–[PGS000897](#) (all intervening numbers).

Code availability

The code EasyQC is available at www.genepi-regensburg.de/easyqc, and Raremetal is available at <https://github.com/SailajaVeda/raremetal>.

References

1. 1.

Taddei, C. et al. Repositioning of the global epicentre of non-optimal cholesterol. *Nature* **582**, 73–77 (2020).

2. 2.

Ference, B. A. et al. Low-density lipoproteins cause atherosclerotic cardiovascular disease. 1. Evidence from genetic, epidemiologic, and clinical studies. A consensus statement from the European Atherosclerosis Society Consensus Panel. *Eur. Heart J.* **38**, 2459–2472 (2017).

3. 3.

Roth, G. A. et al. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet* **392**, 1736–1788 (2018).

4. 4.

Teslovich, T. M. et al. Biological, clinical and population relevance of 95 loci for blood lipids. *Nature* **466**, 707–713 (2010).

5. 5.

Willer, C. J. et al. Discovery and refinement of loci associated with lipid levels. *Nat. Genet.* **45**, 1274–1283 (2013).

6. 6.

Liu, D. J. et al. Exome-wide association study of plasma lipids in >300,000 individuals. *Nat. Genet.* **49**, 1758–1766 (2017).

7. 7.

Lu, X. et al. Exome chip meta-analysis identifies novel loci and East Asian-specific coding variants that contribute to lipid levels and coronary artery disease. *Nat. Genet.* **49**, 1722–1730 (2017).

8. 8.

Kathiresan, S. et al. A genome-wide association study for blood lipid phenotypes in the Framingham Heart Study. *BMC Med. Genet.* **8**, S17 (2007).

9. 9.

Kathiresan, S. et al. Polymorphisms associated with cholesterol and risk of cardiovascular events. *N. Engl. J. Med.* **358**, 1240–1249 (2008).

10. 10.

Peloso, G. M. et al. Association of low-frequency and rare coding-sequence variants with blood lipids and coronary heart disease in 56,000 whites and blacks. *Am. J. Hum. Genet.* **94**, 223–232 (2014).

11. 11.

Hoffmann, T. J. et al. A large electronic-health-record-based genome-wide study of serum lipids. *Nat. Genet.* **50**, 401–413 (2018).

12. 12.

Surakka, I. et al. The impact of low-frequency and rare variants on lipid levels. *Nat. Genet.* **47**, 589–597 (2015).

13. 13.

Klarin, D. et al. Genetics of blood lipids among ~300,000 multi-ethnic participants of the Million Veteran Program. *Nat. Genet.* **50**, 1514–1523 (2018).

14. 14.

Holmen, O. L. et al. Systematic evaluation of coding variation identifies a candidate causal variant in *TM6SF2* influencing total cholesterol and myocardial infarction risk. *Nat. Genet.* **46**, 345–351 (2014).

15. 15.

Asselbergs, F. W. et al. Large-scale gene-centric meta-analysis across 32 studies identifies multiple lipid loci. *Am. J. Hum. Genet.* **91**, 823–838 (2012).

16. 16.

Albrechtsen, A. et al. Exome sequencing-driven discovery of coding polymorphisms associated with common metabolic phenotypes. *Diabetologia* **56**, 298–310 (2013).

17. 17.

Saxena, R. et al. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* **316**, 1331–1336 (2007).

18. 18.

Iotchkova, V. et al. Discovery and refinement of genetic loci associated with cardiometabolic risk using dense imputation maps. *Nat. Genet.* **48**, 1303–1312 (2016).

19. 19.

Tachmazidou, I. et al. A rare functional cardioprotective *APOC3* variant has risen in frequency in distinct population isolates. *Nat. Commun.* **4**, 2872 (2013).

20. 20.

Tang, C. S. et al. Exome-wide association analysis reveals novel coding sequence variants associated with lipid traits in Chinese. *Nat. Commun.* **6**, 10206 (2015).

21. 21.

van Leeuwen, E. M. et al. Genome of the Netherlands population-specific imputations identify an *ABCA6* variant associated with cholesterol levels. *Nat. Commun.* **6**, 6065 (2015).

22. 22.

Spracklen, C. N. et al. Association analyses of East Asian individuals and trans-ancestry analyses with European individuals reveal new loci associated with cholesterol and triglyceride levels. *Hum. Mol. Genet.* **26**, 1770–1784 (2017).

23. 23.

Kanai, M. et al. Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. *Nat. Genet.* **50**, 390–400 (2018).

24. 24.

Sirugo, G., Williams, S. M. & Tishkoff, S. A. The missing diversity in human genetic studies. *Cell* **177**, 26–31 (2019).

25. 25.

Khera, A. V. et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat. Genet.* **50**, 1219–1224 (2018).

26. 26.

Duncan, L. et al. Analysis of polygenic risk score usage and performance in diverse human populations. *Nat. Commun.* **10**, 3328 (2019).

27. 27.

Bunielo, A. et al. The NHGRI–EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res.* **47**, D1005–D1012 (2019).

28. 28.

Tishkoff, S. A. et al. The genetic structure and history of Africans and African Americans. *Science* **324**, 1035–1044 (2009).

29. 29.

Mägi, R. et al. Trans-ethnic meta-regression of genome-wide association studies accounting for ancestry increases power for discovery and improves fine-mapping resolution. *Hum. Mol. Genet.* **26**, 3639–3650 (2017).

30. 30.

Lee, S. H., Yang, J., Goddard, M. E., Visscher, P. M. & Wray, N. R. Estimation of pleiotropy between complex diseases using single-nucleotide polymorphism-derived genomic relationships and restricted maximum likelihood. *Bioinformatics* **28**, 2540–2542 (2012).

31. 31.

Brown, B. C., Ye, C. J., Price, A. L. & Zaitlen, N. Transethnic genetic-correlation estimates from summary statistics. *Am. J. Hum. Genet.* **99**,

76–88 (2016).

32. 32.

Guo, J. et al. Quantifying genetic heterogeneity between continental populations for human height and body mass index. *Sci. Rep.* **11**, 5240 (2021).

33. 33.

Ge, T., Chen, C.-Y., Ni, Y., Feng, Y.-C. A. & Smoller, J. W. Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat. Commun.* **10**, 1776 (2019).

34. 34.

Majara, L. et al. Low generalizability of polygenic scores in African populations due to genetic and environmental diversity. Preprint at *bioRxiv* <https://doi.org/10.1101/2021.01.12.426453> (2021).

35. 35.

Lehmann, B. C. L., Mackintosh, M., McVean, G. & Holmes, C. C. High trait variability in optimal polygenic prediction strategy within multiple-ancestry cohorts. Preprint at *bioRxiv* <https://doi.org/10.1101/2021.01.15.426781> (2021).

36. 36.

Shi, H. et al. Population-specific causal disease effect sizes in functionally important regions impacted by selection. *Nat. Commun.* **12**, 1098 (2021).

37. 37.

Martin, A. R. et al. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat. Genet.* **51**, 584–591 (2019).

38. 38.

Cavazos, T. B. & Witte, J. S. Inclusion of variants discovered from diverse populations improves polygenic risk score transferability. *HGG Adv.* **2**, 100017 (2021).

39. 39.

Wojcik, G. L. et al. Genetic analyses of diverse populations improves discovery for complex traits. *Nature* **570**, 514–518 (2019).

40. 40.

Bentley, A. R. et al. Multi-ancestry genome-wide gene–smoking interaction study of 387,272 individuals identifies new loci associated with serum lipids. *Nat. Genet.* **51**, 636–648 (2019).

41. 41.

Taliun, D. et al. Sequencing of 53,831 diverse genomes from the NHLBI TOPMed Program. *Nature* **590**, 290–299 (2021).

42. 42.

Kowalski, M. H. et al. Use of >100,000 NHLBI Trans-Omics for Precision Medicine (TOPMed) Consortium whole genome sequences improves imputation quality and detection of rare variant associations in admixed African and Hispanic/Latino populations. *PLoS Genet.* **15**, e1008500 (2019).

43. 43.

Das, S. et al. Next-generation genotype imputation service and methods. *Nat. Genet.* **48**, 1284–1287 (2016).

44. 44.

Baigent, C. et al. Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90 056 participants in 14 randomised trials of statins. *Lancet* **366**, 1267–1278 (2005).

45. 45.

Loh, P.-R. et al. Efficient Bayesian mixed-model analysis increases association power in large cohorts. *Nat. Genet.* **47**, 284–290 (2015).

46. 46.

Zhou, W. et al. Efficiently controlling for case-control imbalance and sample relatedness in large-scale genetic association studies. *Nat. Genet.* **50**, 1335–1341 (2018).

47. 47.

Winkler, T. W. et al. Quality control and conduct of genome-wide association meta-analyses. *Nat. Protoc.* **9**, 1192–1212 (2014).

48. 48.

Feng, S., Liu, D., Zhan, X., Wing, M. K. & Abecasis, G. R. RAREMETAL: fast and powerful meta-analysis for rare variants. *Bioinformatics* **30**, 2828–2829 (2014).

49. 49.

Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).

50. 50.

Loh, P.-R., Palamara, P. F. & Price, A. L. Fast and accurate long-range phasing in a UK Biobank cohort. *Nat. Genet.* **48**, 811–816 (2016).

51. 51.

Liu, X. et al. WGSAn: an annotation pipeline for human genome sequencing studies. *J. Med. Genet.* **53**, 111–112 (2016).

52. 52.

Cingolani, P. et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* **6**, 80–92 (2012).

53. 53.

Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).

54. 54.

Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

55. 55.

Liu, D. J. et al. Meta-analysis of gene-level tests for rare variant association. *Nat. Genet.* **46**, 200–204 (2014).

56. 56.

Maller, J. B. et al. Bayesian refinement of association signals for 14 loci in 3 common diseases. *Nat. Genet.* **44**, 1294–1301 (2012).

57. 57.

Kass, R. E. & Raftery, A. E. Bayes factors. *J. Am. Stat. Assoc.* **90**, 773–795 (1995).

58. 58.

Machiela, M. J. & Chanock, S. J. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics* **31**, 3555–3557 (2015).

59. 59.

McLaren, W. et al. The Ensembl Variant Effect Predictor. *Genome Biol.* **17**, 122 (2016).

60. 60.

Sherry, S. T. et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* **29**, 308–311 (2001).

61. 61.

Giambartolomei, C. et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet.* **10**, e1004383 (2014).

62. 62.

GTEx Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330 (2020).

63. 63.

Purcell, S. et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* **81**, 559–575 (2007).

64. 64.

Berisa, T. & Pickrell, J. K. Approximately independent linkage disequilibrium blocks in human populations. *Bioinformatics* **32**, 283–285 (2016).

65. 65.

Finer, S. et al. Cohort Profile: East London Genes &Health (ELGH), a community-based population genomics and health study in British Bangladeshi and British Pakistani people. *Int. J. Epidemiol.* **49**, 20–21i (2019).

66. 66.

Moon, S. et al. The Korea Biobank Array: design and identification of coding variants associated with blood biochemical traits. *Sci. Rep.* **9**, 1382 (2019).

67. 67.

Alexander, D. H., Novembre, J. & Lange, K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* **19**, 1655–1664 (2009).

Acknowledgements

Funding for the Global Lipids Genetics Consortium was provided by the NIH (R01-HL127564). This research was conducted using the UK Biobank Resource under application number 24460. Computing support and file management for central meta-analysis by S. Caron is acknowledged. This research is based on data from the MVP, Office of Research and Development, Veterans Health Administration, and was supported by awards 2I01BX003362-03A1 and 1I01BX004821-01A1. This publication does not represent the views of the Department of Veteran Affairs or the United States Government. Study-specific acknowledgements are provided in the Supplementary Information.

Author information

Author notes

1. These authors contributed equally: Shoa L. Clarke, Kuan-Han H. Wu, Stavroula Kanoni, Greg J. M. Zajac, Shweta Ramdas
2. These authors jointly supervised this work: Gina M. Peloso, Christopher D. Brown, Andrew P. Morris, Themistocles L. Assimes, Panos Deloukas, Yan V. Sun, Cristen J. Willer

Affiliations

1. Department of Internal Medicine, Division of Cardiology, University of Michigan, Ann Arbor, MI, USA

Sarah E. Graham, Ida Surakka, Chao Xue, Jifeng Zhang, Y. Eugene Chen & Cristen J. Willer

2. VA Palo Alto Health Care System, Palo Alto, CA, USA

Shoa L. Clarke, Xiang Zhu, Philip S. Tsao & Themistocles L. Assimes

3. Department of Medicine, Division of Cardiovascular Medicine, Stanford University School of Medicine, Stanford, CA, USA

Shoa L. Clarke, Erik Ingelsson, Philip S. Tsao & Themistocles L. Assimes

4. Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA

Kuan-Han H. Wu, Wei Zhou & Cristen J. Willer

5. William Harvey Research Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK

Stavroula Kanoni, Eirini Marouli, Olga Giannakopoulou, Konain Fatima Bhatti & Panos Deloukas

6. Department of Biostatistics and Center for Statistics Genetics, University of Michigan, Ann Arbor, MI, USA

Greg J. M. Zajac, Anne U. Jackson & Michael Boehnke

7. Department of Genetics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

Shweta Ramdas, Anurag Verma & Christopher D. Brown

8. Clinical Pharmacology, William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK

Ioanna Ntalla

9. Endocrinology, Boston Children's Hospital, Boston, MA, USA

Sailaja Vedantam & Joel N. Hirschhorn

10. Medical and Population Genetics, Broad Institute, Cambridge, MA, USA

Sailaja Vedantam & Joel N. Hirschhorn

11. Department of Genetic Epidemiology, University of Regensburg, Regensburg, Germany

Thomas W. Winkler, Mathias Gorski, Martina E. Zimmermann, Iris M. Heid & Klaus J. Stark

12. McDonnell Genome Institute and Department of Medicine, Washington University, St Louis, MO, USA

Adam E. Locke

13. Department of Precision Medicine, Division of Genome Science, National Institute of Health, Cheongju-si, South Korea

Mi Yeong Hwang, Sohee Han, Hye-Mi Jang, Kyungheon Yoon, Young Jin Kim & Bong-Jo Kim

14. Tohoku Medical Megabank Organization, Tohoku University, Sendai, Japan

Akira Narita, Gen Tamiya & Masayuki Yamamoto

15. Sydney Brenner Institute for Molecular Bioscience, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa
Ananyo Choudhury, Scott Hazelhurst & Michèle Ramsay
16. Center for Research on Genomics and Global Health, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA
Amy R. Bentley, Kenneth Ekoru & Charles N. Rotimi
17. Blizard Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK
Bhavi Trivedi, Karen A. Hunt & David A. van Heel
18. Wellcome Sanger Institute, Hinxton, UK
Hilary C. Martin & Nigel W. Rayner
19. Department of Epidemiology, Emory University Rollins School of Public Health, Atlanta, GA, USA
Qin Hui & Yan V. Sun
20. Atlanta VA Health Care System, Decatur, GA, USA
Qin Hui & Yan V. Sun
21. Malcolm Randall VA Medical Center, Gainesville, FL, USA
Derek Klarin
22. Division of Vascular Surgery and Endovascular Therapy, University of Florida College of Medicine, Gainesville, FL, USA
Derek Klarin

23. Program in Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA

Derek Klarin & Masahiro Kanai

24. Department of Statistics, The Pennsylvania State University, University Park, PA, USA

Xiang Zhu

25. Huck Institutes of the Life Sciences, The Pennsylvania State University, University Park, PA, USA

Xiang Zhu

26. Department of Statistics, Stanford University, Stanford, CA, USA

Xiang Zhu

27. deCODE genetics/Amgen, Reykjavik, Iceland

Gudmar Thorleifsson, Anna Helgadottir, Daniel F. Gudbjartsson, Hilma Holm, Unnur Thorsteinsdottir & Kari Stefansson

28. School of Engineering and Natural Sciences, University of Iceland, Reykjavik, Iceland

Daniel F. Gudbjartsson

29. Department of Clinical Biochemistry, Landspitali–National University Hospital of Iceland, Reykjavik, Iceland

Isleifur Olafsson

30. Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan

Masato Akiyama, Saori Sakaue, Masahiro Kanai, Yoichiro Kamatani & Yukinori Okada

31. Department of Ophthalmology, Graduate School of Medical Sciences,
Kyushu University, Fukuoka, Japan

Masato Akiyama

32. Department of Statistical Genetics, Osaka University Graduate School
of Medicine, Osaka, Japan

Saori Sakaue & Yukinori Okada

33. Department of Allergy and Rheumatology, Graduate School of
Medicine, The University of Tokyo, Tokyo, Japan

Saori Sakaue

34. Laboratory for Statistical and Translational Genetics, RIKEN Center
for Integrative Medical Sciences, Yokohama, Japan

Chikashi Terao

35. Department of Biomedical Informatics, Harvard Medical School,
Boston, MA, USA

Masahiro Kanai

36. Analytic and Translational Genetics Unit, Massachusetts General
Hospital, Boston, MA, USA

Wei Zhou

37. Program in Medical and Population Genetics, Broad Institute of
Harvard and MIT, Cambridge, MA, USA

Wei Zhou & Pradeep Natarajan

38. K.G. Jebsen Center for Genetic Epidemiology, Department of Public
Health and Nursing, NTNU, Norwegian University of Science and
Technology, Trondheim, Norway

Ben M. Brumpton, Humaira Rasheed, Kristian Hveem & Bjørn Olav Åsvold

39. MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK

Ben M. Brumpton & Humaira Rasheed

40. Department of Thoracic Medicine, St Olavs Hospital, Trondheim University Hospital, Trondheim, Norway

Ben M. Brumpton

41. Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland

Sanni E. Ruotsalainen, Aki S. Havulinna, Anu Loukola, Sailalitha Bollepalli, Jaakko Kaprio, Samuli Ripatti & Elisabeth Widén

42. Department of Public Health and Welfare, Finnish Institute for Health and Welfare, Helsinki, Finland

Aki S. Havulinna, Pekka Jousilahti & Veikko Salomaa

43. Department of Genetics, Institute for Biomedical Informatics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

Yogasudha Veturi, Yuki Bradford, Jason E. Miller & Marylyn D. Ritchie

44. Division of Clinical Pharmacology, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, USA

QiPing Feng

45. Department of Medicine (Medical Genetics), University of Washington, Seattle, WA, USA

Elisabeth A. Rosenthal

46. Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, Cincinnati, OH, USA

Todd Lingren

47. Center for Genetic Medicine, Northwestern University, Evanston, IL, USA

Jennifer Allen Pacheco

48. Genentech, South San Francisco, CA, USA

Sarah A. Pendergrass & Mark I. McCarthy

49. Fred Hutchinson Cancer Research Center, Division of Public Health Sciences, Seattle, WA, USA

Jeffrey Haessler & Charles Kooperberg

50. Division of Preventive Medicine, Brigham and Women's Hospital, Boston, MA, USA

Franco Giulianini, Julie E. Buring, Paul M. Ridker & Daniel I. Chasman

51. Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Western General Hospital, Edinburgh, UK

Archie Campbell

52. Usher Institute for Population Health Sciences and Informatics, The University of Edinburgh, Edinburgh, UK

Archie Campbell

53. Clinical Trial Service Unit and Epidemiological Studies Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK

Kuang Lin, Iona Y. Millwood, Zhengming Chen, Robin G. Walters & Michael V. Holmes

54. Medical Research Council Population Health Research Unit, Nuffield Department of Population Health, University of Oxford, Oxford, UK

Iona Y. Millwood, Zhengming Chen, Robin G. Walters & Michael V. Holmes

55. Department of Population Medicine, Qatar University College of Medicine, QU Health, Doha, Qatar

George Hindy

56. Center for Non-Communicable Diseases, Karachi, Pakistan

Asif Rasheed, Shahid Abbas & Danish Saleheen

57. Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI, USA

Jessica D. Faul, David R. Weir & Jennifer A. Smith

58. Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI, USA

Wei Zhao

59. Program in Genetic Epidemiology and Statistical Genetics, Department of Epidemiology, Harvard T. H. Chan School of Public Health, Boston, MA, USA

Constance Turman, Hongyan Huang & Peter Kraft

60. Department of Epidemiology, University of North Carolina, Chapel Hill, NC, USA

Mariaelisa Graff, Heather M. Highland, Kristin L. Young & Kari E. North

61. Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK

Anubha Mahajan, Anuj Goel, Nigel W. Rayner, Mark I. McCarthy & Hugh Watkins

62. Human Genetics Center, Department of Epidemiology, Human Genetics, and Environmental Sciences, School of Public Health, The University of Texas Health Science Center at Houston, Houston, TX, USA

Michael R. Brown, Paul S. de Vries & Alanna C. Morrison

63. Department of Epidemiology and Biostatistics, Imperial College London, London, UK

Weihua Zhang & John C. Chambers

64. Department of Cardiology, Ealing Hospital, London North West University Healthcare NHS Trust, Southall, UK

Weihua Zhang, John C. Chambers & Jaspal S. Kooner

65. Imperial College Healthcare NHS Trust, London, UK

Weihua Zhang & Jaspal S. Kooner

66. Department of Biostatistics, Center for Statistical Genetics, University of Michigan, Ann Arbor, MI, USA

Ketian Yu, Ellen M. Schmidt, Anita Pandit, Xianyong Yin & Goncalo Abecasis

67. Department of Medical Sciences, Molecular Epidemiology and Science for Life Laboratory, Uppsala University, Uppsala, Sweden

Stefan Gustafsson & Erik Ingelsson

68. MRC Epidemiology Unit, University of Cambridge School of Clinical Medicine, Cambridge, UK

Jian'an Luan, Nicholas J. Wareham & Claudia Langenberg

69. Cardiovascular Epidemiology Unit, Department of Public Health and Primary Care, University of Cambridge Strangeways Research Laboratory, Cambridge, UK

Jing-Hua Zhao

70. Center for Genomic Medicine, Kyoto University Graduate School of Medicine, Kyoto, Japan

Fumihiko Matsuda, Takahisa Kawaguchi & Yasuharu Tabara

71. Department of Internal Medicine, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

Carolina Medina-Gomez & Joyce B. J. van Meurs

72. Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, The Netherlands

Carolina Medina-Gomez, Natalie Terzikhan, Hieab H. H. Adams, M. Arfan Ikram, Joyce B. J. van Meurs & Mohsen Ghanbari

73. Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA

Achilleas Pitsillides, L. Adrienne Cupples & Gina M. Peloso

74. Department of Biological Psychology, Behavioral and Movement Sciences, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

Jouke Jan Hottenga, Gonke Willemsen, Dorret I. Boomsma & Eco J. C. de Geus

75. Amsterdam Public Health, VU Medical Center Amsterdam,
Amsterdam, The Netherlands

Jouke Jan Hottenga

76. Amsterdam Public Health Research Institute, VU Medical Center
Amsterdam, Amsterdam, The Netherlands

Gonneke Willemsen

77. Genetics of Complex Traits, University of Exeter Medical School,
University of Exeter, Exeter, UK

Andrew R. Wood, Yingji Ji & Timothy M. Frayling

78. Department of Clinical Acupuncture and Moxibustion, Nanjing
University of Chinese Medicine, Nanjing, China

Zishan Gao

79. Research Unit of Molecular Epidemiology, Helmholtz Zentrum
München, German Research Center for Environmental Health,
Neuherberg, Germany

Zishan Gao & Christian Gieger

80. Institute of Epidemiology, Helmholtz Zentrum München, German
Research Center for Environmental Health, Neuherberg, Germany

Zishan Gao, Annette Peters & Christian Gieger

81. MRC Integrative Epidemiology Unit at the University of Bristol,
Bristol, UK

Simon Haworth, Ruth E. Mitchell & Nicholas J. Timpson

82. Bristol Dental School, University of Bristol, Bristol, UK

Simon Haworth

83. Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

Ruth E. Mitchell & Nicholas J. Timpson

84. Saw Swee Hock School of Public Health, National University of Singapore and National University Health System, Singapore, Singapore

Jin Fang Chai, Xueling Sim, E. Shyong Tai & Rob M. van Dam

85. Center for Clinical Research and Prevention, Bispebjerg and Frederiksberg Hospital, Copenhagen, Denmark

Mette Aadahl, Line L. Kårhus, Line T. Møllehave, Betina H. Thuesen, Allan Linneberg & Thomas M. Dantoft

86. The Institute for Translational Genomics and Population Sciences, Department of Pediatrics, Lundquist Institute for Biomedical Innovations (Formerly LABioMed) at Harbor-UCLA Medical Center, Torrance, CA, USA

Jie Yao, Xiaohui Li, Yii-Der Ida Chen, Xiuqing Guo & Jerome I. Rotter

87. Center for Public Health Genomics, University of Virginia, Charlottesville, VA, USA

Ani Manichaikul

88. William Harvey Research Institute, Barts and The London School of Medicine and Dentistry, John Vane Science Centre, Queen Mary University of London, London, UK

Helen R. Warren, Julia Ramirez, Mark J. Caulfield & Patricia B. Munroe

89. NIHR Barts Cardiovascular Biomedical Research Centre, Barts and The London School of Medicine and Dentistry, Queen Mary

University of London, London, UK

Helen R. Warren, Morris Brown, Mark J. Caulfield, Patricia B. Munroe & Panos Deloukas

90. Novo Nordisk Foundation Center for Basic Metabolic Research,
Faculty of Health and Medical Sciences, University of Copenhagen,
Copenhagen, Denmark

Jette Bork-Jensen, Henrik Vestergaard, Oluf Pedersen, Torben Hansen & Niels Grarup

91. Division of Cardiovascular Medicine, Radcliffe Department of Medicine, John Radcliffe Hospital, University of Oxford, Oxford, UK

Anuj Goel

92. Group of Genomics of Complex Diseases, Research Institute of Hospital de la Santa Creu i Sant Pau (IIB Sant Pau), Barcelona, Spain

Maria Sabater-Lleal

93. Cardiovascular Medicine Unit, Department of Medicine, Karolinska Institutet, Center for Molecular Medicine, Karolinska University Hospital, Stockholm, Sweden

Maria Sabater-Lleal

94. Department of Internal Medicine, Section Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands

Raymond Noordam

95. Institute for Genetic and Biomedical Research, Italian National Council of Research (IRGB CNR), Cagliari,, Italy

Carlo Sidore

96. Institute for Genetic and Biomedical Research, Italian National Council of Research (IRGB CNR), Lanusei, Italy

Edoardo Fiorillo

97. University Center for Primary Care and Public Health, University of Lausanne, Lausanne, Switzerland

Aaron F. McDaid

98. Swiss Institute of Bioinformatics, Lausanne, Switzerland

Aaron F. McDaid & Zoltan Katalik

99. Department of Medicine, Internal Medicine, Lausanne University Hospital and University of Lausanne, Lausanne, Switzerland

Pedro Marques-Vidal

100. Department of Epidemiology and Biostatistics, MRC–PHE Centre for Environment and Health, School of Public Health, Imperial College London, London, UK

Matthias Wielscher & Marjo-Riitta Jarvelin

101. Department of Cardiology, Leiden University Medical Center, Leiden, The Netherlands

Stella Trompet & J. Wouter Jukema

102. Department of Internal Medicine, Section of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands

Stella Trompet

103. BHF Glasgow Cardiovascular Research Centre, Faculty of Medicine, Glasgow, UK

Naveed Sattar

104. Institute for Cardiogenetics, University of Lübeck, DZHK (German Research Centre for Cardiovascular Research), Partner site Hamburg/Lübeck/Kiel, University Heart Center Lübeck, Lübeck, Germany

Matthias Munz

105. Charité–University Medicine Berlin, Member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany

Matthias Munz

106. Department of Periodontology and Synoptic Dentistry, Berlin Institute of Health, Institute for Dental and Craniofacial Sciences, Berlin, Germany

Matthias Munz

107. Deutsches Herzzentrum München, Klinik für Herz- und Kreislauferkrankungen, Technische Universität München, Munich, Germany

Lingyao Zeng

108. Deutsches Zentrum für Herz-Kreislauf-Forschung (DZHK), Partner site Munich Heart Alliance, Munich, Germany

Lingyao Zeng

109. Key Laboratory of Cardiovascular Epidemiology and Department of Epidemiology, State Key Laboratory of Cardiovascular Disease, Fuwai Hospital, National Center for Cardiovascular Diseases, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China

Jianfeng Huang, Bin Yang, Shufeng Chen, Fangchao Liu, Xiangfeng Lu & Dongfeng Gu

110. Lund University Diabetes Centre, Malmo, Sweden

Alaitz Poveda, Azra Kurbasic & Paul W. Franks

111. Department of Genetics and Pharmacology, Institute of Genetic Epidemiology, Medical University of Innsbruck, Innsbruck, Austria

Claudia Lamina & Lukas Forer

112. German Chronic Kidney Disease Study, Innsbruck, Austria

Claudia Lamina, Lukas Forer & Florian Kronenberg

113. Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, Leipzig, Germany

Markus Scholz & Markus Loeffler

114. LIFE Research Centre for Civilization Diseases, University of Leipzig, Leipzig, Germany

Markus Scholz, Joachim Thiery & Markus Loeffler

115. Radboud University Medical Center, Radboud Institute for Health Sciences, Nijmegen, The Netherlands

Tessel E. Galesloot, Lambertus A. L. M. Kiemeney & Jacqueline de Graaf

116. Quantinuum Research, Wayne, PA, USA

Jonathan P. Bradfield

117. Division of Statistical Genomics, Department of Genetics, Washington University School of Medicine, St Louis, MO, USA

E. Warwick Daw, Mary F. Feitosa, Mary K. Wojczynski, Karen Schwander & Michael A. Province

118. Department of Epidemiology, University of Pittsburgh, Pittsburgh, PA, USA

Joseph M. Zmuda

119. Institute for Biomedicine, Eurac Research, Affiliated Institute of the University of Lübeck, Bolzano, Italy

Jonathan S. Mitchell, Christian Fuchsberger & Peter P. Pramstaller

120. Department of Clinical Biochemistry, Lillebaelt Hospital, Vejle, Denmark

Henry Christensen & Ivan Brandslund

121. Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA, USA

Jennifer A. Brody & Joshua C. Bis

122. The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Michael Preuss, Girish N. Nadkarni & Ruth J. F. Loos

123. Department of Twin Research and Genetic Epidemiology, King's College London, London, UK

Massimo Mangino, Paraskevi Christofidou & Tim D. Spector

124. NIHR Biomedical Research Centre at Guy's and St Thomas' Foundation Trust, London, UK

Massimo Mangino

125. Department of Cardiology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

Niek Verweij, Jan W. Benjamins & Pim van der Harst

126. Institute of Cardiovascular Sciences, University College London, London, UK

Jorgen Engmann

127. Department of Epidemiology and Public Health, University College London, London, UK

Jorgen Engmann & Mika Kivimaki

128. Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, USA

Rachel L. Kember

129. Amsterdam UMC, Department of Epidemiology and Biostatistics, Amsterdam Public Health Research Institute, Amsterdam, The Netherlands

Roderick C. Sliker, Morgan O. Obura, Joline W. J. Beulens & Leen M. 't Hart

130. Department of Cell and Chemical Biology, Leiden University Medical Center, Leiden, The Netherlands

Roderick C. Sliker & Leen M. 't Hart

131. Montreal Heart Institute, Montreal, Quebec, Canada

Ken Sin Lo, Jean-Claude Tardif & Guillaume Lettre

132. Icelandic Heart Association, Reykjavik, Iceland

Nuno R. Zilhao & Vilmundur Gudnason

133. Department of Anthropology, University of Toronto at Mississauga, Mississauga, Ontario, Canada

Phuong Le & Esteban J. Parra

134. Vth Department of Medicine, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

Marcus E. Kleber, Graciela E. Delgado & Winfried März

135. Synlab MVZ Humangenetik Mannheim, Mannheim, Germany

Marcus E. Kleber

136. Biomedical Technology Research Center, Tokushima Research Institute, Otsuka Pharmaceutical, Tokushima, Japan

Daisuke D. Ikeda, Hiroyuki Iha & Haretsugu Hishigaki

137. Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland, Australia

Jian Yang

138. Institute for Advanced Research, Wenzhou Medical University, Wenzhou, China

Jian Yang

139. Nuffield Department of Population Health, University of Oxford, Oxford, UK

Jun Liu & Cornelia M. van Duijn

140. Laboratory of Neurogenetics, National Institute on Aging, NIH, Bethesda, MD, USA

Hampton L. Leonard & Mike A. Nalls

141. Data Tectica International, Glen Echo, MD, USA

Hampton L. Leonard & Mike A. Nalls

142. MRC Human Genetics Unit, Institute of Genetics and Cancer, University of Edinburgh, Western General Hospital, Edinburgh, UK

Jonathan Marten, Paul R. H. J. Timmers, James F. Wilson, Veronique Vitart & Caroline Hayward

143. Institute for Medical Informatics, Biometrie and Epidemiology,
University of Duisburg-Essen, Essen, Germany

Börge Schmidt, Marina Arendt & Karl-Heinz Jöckel

144. Department of Computer Science, University of Applied Sciences and Arts Dortmund, Dortmund, Germany

Marina Arendt

145. Centre for Public Health, Queen's University of Belfast, Belfast, UK

Laura J. Smyth, Marisa Cañadas-Garre, Amy Jayne
McKnight & Frank Kee

146. Department of Epidemiology and Biostatistics, School of Public Health, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

Chaolong Wang

147. Genome Institute of Singapore, Agency for Science, Technology and Research, Singapore, Singapore

Chaolong Wang, Rajkumar Dorajoo, Jianjun Liu & Chiea Chuen Khor

148. Public Health Informatics Unit, Department of Integrated Health Sciences, Nagoya University Graduate School of Medicine, Nagoya, Japan

Masahiro Nakatomi

149. MRC Unit for Lifelong Health and Ageing at UCL, London, UK

Andrew Wong & Nish Chaturvedi

150. Department of Pediatrics, Tampere University Hospital, Tampere, Finland

Nina Hutri-Kähönen

151. Department of Pediatrics, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland

Nina Hutri-Kähönen

152. Brown Foundation Institute of Molecular Medicine, McGovern Medical School, University of Texas Health Science Center at Houston, Houston, TX, USA

Rui Xia & Myriam Fornage

153. Departamento de Medicina Genómica y Toxicología Ambiental, Instituto de Investigaciones Biomédicas, UNAM, Ciudad de Mexico, Mexico, Mexico

Alicia Huerta-Chagoya

154. Departamento de Genómica Computacional, Instituto Nacional de Medicina Genómica, Ciudad de Mexico, Mexico, Mexico

Juan Carlos Fernandez-Lopez

155. Center for Diabetes Research, University of Bergen, Bergen, Norway

Valeriya Lyssenko

156. Lund University Diabetes Center, Lund University, Malmo, Sweden

Valeriya Lyssenko

157. Genomic Research on Complex Diseases (GRC Group), CSIR–Centre for Cellular and Molecular Biology, Hyderabad, India

Meraj Ahmed & Giriraj R. Chandak

158. Epidemiology, School of Public Health, University of Alabama at Birmingham, Birmingham, AL, USA

Marguerite R. Irvin

159. Hunter Medical Research Institute, Newcastle, New South Wales, Australia

Christopher Oldmeadow, John Attia & Rodney J. Scott

160. Medical Research Institute, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul, Korea

Han-Na Kim

161. Department of Clinical Research Design and Evaluation, SAIHST, Sungkyunkwan University, Seoul, Korea

Han-Na Kim

162. Center for Cohort Studies, Total Healthcare Center, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul, Korea

Seungho Ryu

163. Department of Occupational and Environmental Medicine, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul, Korea

Seungho Ryu

164. Centre for Global Health Research, Usher Institute, University of Edinburgh, Edinburgh, UK

Paul R. H. J. Timmers, Katherine A. Kentistou, Harry Campbell, Peter K. Joshi & James F. Wilson

165. Thurston Arthritis Research Center, University of North Carolina, Chapel Hill, NC, USA

Liubov Arbeeva, Amanda E. Nelson & Yvonne M. Golightly

166. Division of Biomedical Informatics and Personalized Medicine,
Department of Medicine, Anschutz Medical Campus, University of
Colorado, Denver, Aurora, CO, USA

Leslie A. Lange

167. Ocular Epidemiology, Singapore Eye Research Institute, Singapore
National Eye Centre, Singapore, Singapore

Xiaoran Chai, Ching-Yu Cheng, Tien-Yin Wong & Charumathi
Sabanayagam

168. Department of Ophthalmology, National University of Singapore and
National University Health System, Singapore, Singapore

Xiaoran Chai

169. Genomics and Molecular Medicine Unit, CSIR–Institute of Genomics
and Integrative Biology, New Delhi, India

Gauri Prasad

170. Academy of Scientific and Innovative Research, CSIR–Institute of
Genomics and Integrative Biology Campus, New Delhi, India

Gauri Prasad & Dwaipayan Bharadwaj

171. Departments of Ophthalmology and Human Genetics, Radboud
University Nijmegen Medical Center, Nijmegen, The Netherlands

Laura Lorés-Motta, Marc Pauper & Anneke I. den Hollander

172. Vanderbilt Epidemiology Center, Division of Epidemiology, Vanderbilt
University Medical Center, Nashville, TN, USA

Jirong Long, Qiuyin Cai, Xiao-Ou Shu & Wei Zheng

173. Department of Pediatrics, University of California, San Francisco,
Oakland, CA,, USA

Elizabeth Theusch

174. National Center for Global Health and Medicine, Tokyo, Japan

Fumihiko Takeuchi & Norihiro Kato

175. Department of Genetics, University of North Carolina, Chapel Hill,
NC, USA

Cassandra N. Spracklen & Karen L. Mohlke

176. Department of Biostatistics and Epidemiology, University of
Massachusetts-Amherst, Amherst, MA,, USA

Cassandra N. Spracklen

177. Department of Cardiovascular Sciences, University of Leicester,
Leicester, UK

Sophie C. Warner, Christopher P. Nelson, Peter S. Braund & Nilesh J.
Samani

178. NIHR Leicester Biomedical Research Centre, Glenfield Hospital,
Leicester, UK

Sophie C. Warner, Christopher P. Nelson, Peter S. Braund & Nilesh J.
Samani

179. Beijing Institute of Ophthalmology, Beijing Key Laboratory of
Ophthalmology and Visual Sciences, Beijing, China

Ya Xing Wang

180. Beijing Tongren Eye Center, Beijing Tongren Hospital, Capital
Medical University, Beijing, China

Ya Xing Wang & Wen B. Wei

181. Institute of Genetics and Biophysics “Adriano Buzzati-Traverso”—CNR, Naples, Italy

Teresa Nutile, Daniela Ruggiero & Marina Ciullo

182. IRCCS Neuromed, Pozzilli, Isernia, Italy

Daniela Ruggiero & Marina Ciullo

183. Department of Psychiatry, Washington University, St Louis, MO, USA

Yun Ju Sung

184. Division of Endocrinology and Metabolism, Tri-Service General Hospital Songshan Branch, Taipei, Taiwan

Yi-Jen Hung

185. Rush Alzheimer’s Disease Center, Rush University Medical Center, Chicago, IL, USA

Jingyun Yang & David A. Bennett

186. Department of Neurological Sciences, Rush University Medical Center, Chicago, IL, USA

Jingyun Yang & David A. Bennett

187. Department of Nephrology, University Hospital Regensburg, Regensburg, Germany

Mathias Gorski, Bettina Jung & Carsten A. Böger

188. Department of Medicine, Surgery and Health Sciences, University of Trieste, Trieste, Italy

Marco Brumat & Giorgia Girotto

189. Department of Molecular Epidemiology, German Institute of Human Nutrition Potsdam-Rehbruecke, Nuthetal, Germany

Karina Meidtner, Susanne Jäger & Matthias B. Schulze

190. German Center for Diabetes Research (DZD), Neuherberg, Germany

Karina Meidtner, Susanne Jäger, Matthias B. Schulze, Annette Peters & Christian Gieger

191. Department of Epidemiology, School of Public Health, University of Michigan, Ann Arbor, MI, USA

Lawrence F. Bielak, Jennifer A. Smith & Sharon L. R. Kardia

192. Department of Genetics and Bioinformatics, Dasman Diabetes Institute, Kuwait, Kuwait

Prashantha Hebbar, Fahd Al-Mulla & Thangavel Alphonse Thanaraj

193. Department of Nutrition and Dietetics, School of Health Science and Education, Harokopio University of Athens, Athens, Greece

Aliki-Eleni Farmaki & George Dedoussis

194. Department of Population Science and Experimental Medicine, University College London, London, UK

Aliki-Eleni Farmaki

195. Department of Neurology, Clinical Division of Neurogeriatrics, Medical University of Graz, Graz, Austria

Edith Hofer & Reinhold Schmidt

196. Institute for Medical Informatics, Statistics and Documentation, Medical University of Graz, Graz, Austria

Edith Hofer

197. Department of Bioinformatics and Genomics, University of North Carolina at Charlotte, Charlotte, NC, USA

Maoxuan Lin

198. Institute for Maternal and Child Health, IRCCS Burlo Garofolo, Trieste, Italy

Maria Pina Concas

199. Institute of Genetic and Biomedical Research, National Research Council of Italy, UOS of Sassari, Sassari, Italy

Simona Vaccargiu

200. Department of Epidemiology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

Peter J. van der Most & Harold Snieder

201. Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland

Niina Pitkänen, Olli T. Raitakari & Katja Pahkala

202. Centre for Population Health Research, University of Turku and Turku University Hospital, Turku, Finland

Niina Pitkänen, Olli T. Raitakari & Katja Pahkala

203. Sleep Medicine and Circadian Disorders, Brigham and Women's Hospital, Boston, MA, USA

Brian E. Cade, Jiwon Lee & Susan Redline

204. Division of Sleep Medicine, Harvard Medical School, Boston, MA, USA

Brian E. Cade & Susan Redline

205. Central Diagnostics Laboratory, Division Laboratories, Pharmacy, and Biomedical Genetics, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

Sander W. van der Laan, Gerard Pasterkamp & Imo E. Hoefer

206. Laboratory of Epidemiology and Population Science, National Institute on Aging Intramural Research Program, NIH Biomedical Research Center, NIA, Baltimore, MD, USA

Kumaraswamy Naidu Chitrala

207. Department of Functional Genomics, Interfaculty Institute for Genetics and Functional Genomics, University of Greifswald and University Medicine Greifswald, Greifswald, Germany

Stefan Weiss & Georg Homuth

208. Oneomics, Bucheon, Korea

Jong Young Lee

209. Department of Biomedical Science, Hallym University, Chuncheon, Korea

Hyeok Sun Choi, Jae Hun Shin & Yoon Shin Cho

210. Centre for Bone and Arthritis Research, Department of Internal Medicine and Clinical Nutrition, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

Maria Nethander, Claes Ohlsson & Dan Mellström

211. Bioinformatics Core Facility, Sahlgrenska Academy, University of Gothenburg, Gothenburg, Sweden

Maria Nethander

212. Institute of Medical Informatics and Statistics, Kiel University, Kiel, Germany

Sandra Freitag-Wolf

213. Institute of Translational Genomics, Helmholtz Zentrum München—German Research Center for Environmental Health, Neuherberg, Germany

Lorraine Southam, Nigel W. Rayner & Eleftheria Zeggini

214. Wellcome Trust Sanger Institute, Hinxton, UK

Lorraine Southam & Eleftheria Zeggini

215. Oxford Centre for Diabetes Endocrinology and Metabolism, Oxford, UK

Nigel W. Rayner

216. School of Medicine and Public Health, College of Health, Medicine and Wellbeing, University of Newcastle, Newcastle, New South Wales, Australia

Carol A. Wang & Craig E. Pennell

217. Center for Geriatrics and Gerontology, Division of Endocrinology and Metabolism, Department of Internal Medicine, Taichung Veterans General Hospital, Taichung, Taiwan

Shih-Yi Lin

218. School of Medicine, National Yang-Ming University, Taipei, Taiwan

Shih-Yi Lin

219. School of Medicine, National Defense Medical Center, Taipei, Taiwan

Shih-Yi Lin

220. Division of Endocrinology and Metabolism, Department of Internal Medicine, Taichung Veterans General Hospital, Taichung, Taiwan

Jun-Sing Wang

221. Department of Medicine, School of Medicine, National Yang-Ming University, Taipei, Taiwan

Jun-Sing Wang

222. Department of Kinesiology, Université Laval, Quebec, Quebec, Canada

Christian Couture & Louis Pérusse

223. Department of Clinical Chemistry, Fimlab Laboratories, Tampere, Finland

Leo-Pekka Lyytikäinen & Terho Lehtimäki

224. Department of Clinical Chemistry, Finnish Cardiovascular Research Center–Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland

Leo-Pekka Lyytikäinen & Terho Lehtimäki

225. Department of Cardiology, Heart Center, Tampere University Hospital, Tampere, Finland

Kjell Nikus

226. Department of Cardiology, Finnish Cardiovascular Research Center–Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland

Kjell Nikus

227. University of Queensland Diamantina Institute, Translational Research Institute, Brisbane, Queensland, Australia

Gabriel Cuellar-Partida

1. Department of Medicine, Bornholms Hospital, Ronne, Denmark
Henrik Vestergaard
2. School of Public Health, University of Alabama at Birmingham,
Birmingham, AL, USA
Bertha Hidalgo
3. Cardiology, Division Heart and Lungs, University Medical Center
Utrecht, Utrecht University, Utrecht, The Netherlands
Jessica van Setten, Folkert W. Asselbergs & Adriaan O. Kraaijeveld
4. Department of Population and Quantitative Health Sciences, Case
Western Reserve University, Cleveland, OH, USA
Xiaoyin Li & Xiaofeng Zhu
5. Division of Endocrinology, Ohio State University, Columbus, OH,
USA
Rebecca D. Jackson
6. Department of Epidemiology, University of Washington, Seattle, WA,
USA
Alexander P. Reiner
7. School of Medicine and Health Sciences, George Washington
University, Washington, DC, USA
Lisa Warsinger Martin
8. Department of Epidemiology, School of Public Health, Peking
University Health Science Center, Beijing, China
Liming Li

9. Institute for Laboratory Medicine, University Hospital Leipzig,
Leipzig, Germany

Joachim Thiery

10. Laboratory of Epidemiology and Population Sciences, National
Institute on Aging, NIH, Baltimore, MD, USA

Lenore J. Launer

11. Department of Clinical Physiology and Nuclear Medicine, Turku
University Hospital, Turku, Finland

Olli T. Raitakari

12. Department of Environmental and Preventive Medicine, Jichi Medical
University School of Medicine, Shimotsuke, Japan

Sahoko Ichihara

13. Centre for Population Health Sciences, Usher Institute, University of
Edinburgh, Edinburgh, UK

Sarah H. Wild

14. Department of Functional Pathology, Shimane University School of
Medicine, Izumo, Japan

Toru Nabika

15. Department of Pediatrics and Adolescent Medicine, Turku University
Hospital and University of Turku, Turku, Finland

Harri Niinikoski

16. Department of Physiology, University of Turku, Turku, Finland

Harri Niinikoski

17. Faculty of Medicine, University of Split, Split, Croatia

Ivana Kolcic & Ozren Polasek

18. Medical Department III–Endocrinology, Nephrology, Rheumatology,
University of Leipzig Medical Center, Leipzig, Germany

Peter Kovacs & Anke Tönjes

19. Department of Nutrition–Dietetics, Harokopio University, Eleftheriou Venizelou, Athens, Greece

Tota Giardoglou

20. Department of Clinical Gene Therapy, Osaka University Graduate School of Medicine, Suita, Japan

Tomohiro Katsuya

21. Department of Geriatric and General Medicine, Osaka University Graduate School of Medicine, Suita, Japan

Tomohiro Katsuya

22. Department of Vascular Surgery, Division of Surgical Specialties,
University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands

Dominique de Kleijn & Gert J. de Borst

23. Corneal Dystrophy Research Institute, Department of Ophthalmology,
Yonsei University College of Medicine, Seoul, Korea

Eung Kweon Kim

24. Department of Radiology and Nuclear Medicine, Erasmus MC,
University Medical Center Rotterdam, Rotterdam, The Netherlands

Hieab H. H. Adams

25. Julius Centre for Health Sciences and Primary Care, University Medical Centre Utrecht, Utrecht, The Netherlands
Joline W. J. Beulens
26. Second Department of Cardiology, Medical School, National and Kapodistrian University of Athens, Attikon University Hospital, Athens, Greece
Loukianos S. Rallidis
27. Center for Vision Research, Department of Ophthalmology and The Westmead Institute, University of Sydney, Sydney, New South Wales, Australia
Paul Mitchell
28. Menzies Institute for Medical Research, School of Medicine, University of Tasmania, Hobart, Tasmania, Australia
Alex W. Hewitt
29. Centre for Eye Research Australia, University of Melbourne, Melbourne, Victoria, Australia
Alex W. Hewitt
30. Department of Clinical Physiology, Tampere University Hospital, Tampere, Finland
Mika Kähönen
31. Department of Clinical Physiology, Finnish Cardiovascular Research Center–Tampere, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland
Mika Kähönen

32. Institute of Nutrition and Functional Foods (INAF), Université Laval,
Quebec, Quebec, Canada

Louis Pérusse

33. Pennington Biomedical Research Center, Baton Rouge, LA, USA

Claude Bouchard

34. Medical School, Faculty of Health and Medical Sciences, University
of Western Australia, Perth, Western Australia, Australia

Trevor A. Mori

35. Institute of Epidemiology, Kiel University, Kiel, Germany

Wolfgang Lieb

36. Institute of Clinical Molecular Biology, Kiel University, Kiel,
Germany

Andre Franke

37. Department of Drug Treatment, Sahlgrenska University Hospital,
Gothenburg, Sweden

Claes Ohlsson

38. Geriatric Medicine, Institute of Medicine, Sahlgrenska Academy,
University of Gothenburg, Gothenburg, Sweden

Dan Mellström

39. Department of Internal Medicine, Ewha Womans University School of
Medicine, Seoul, Korea

Hyejin Lee

40. Division of Cancer Control and Population Sciences, UPMC Hillman Cancer Center, University of Pittsburgh, Pittsburgh, PA, USA

Jian-Min Yuan

41. Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA

Jian-Min Yuan

42. Healthy Longevity Translational Research Programme, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

Woon-Puay Koh

43. Singapore Institute for Clinical Sciences, Agency for Science Technology and Research (A*STAR), Singapore, Singapore

Woon-Puay Koh

44. Department of Endocrinology and Metabolism, Kyung Hee University School of Medicine, Seoul, Korea

Sang Youl Rhee & Jeong-Taek Woo

45. Institute for Community Medicine, University Medicine Greifswald, Greifswald, Germany

Henry Völzke

46. Laboratory of Epidemiology and Population Science, National Institute on Aging Intramural Research Program, NIH Biomedical Research Center, Baltimore, MD, USA

Michele K. Evans & Alan B. Zonderman

47. Paavo Nurmi Centre, Sports and Exercise Medicine Unit, Department of Physical Activity and Health, University of Turku, Turku, Finland

Katja Pahkala

48. Interdisciplinary Center Psychopathology and Emotion Regulation (ICPE), University of Groningen, University Medical Center Groningen, Groningen, The Netherlands

Albertine J. Oldehinkel

49. Institute of Molecular Genetics, National Research Council of Italy, Pavia, Italy

Ginevra Biino

50. Gottfried Schatz Research Center for Cell Signaling, Metabolism and Aging, Medical University of Graz, Graz, Austria

Helena Schmidt

51. Local Health Unit Toscana Centro, Firenze, Italy

Stefania Bandinelli

52. Institute of Nutritional Science, University of Potsdam, Nuthetal, Germany

Matthias B. Schulze

53. Institute for Maternal and Child Health IRCCS “Burlo Garofolo”, Trieste, Italy

Giorgia Girotto

54. Department of Nephrology, Diabetology, Rheumatology, Traunstein Hospital, Traunstein, Germany

Carsten A. Böger

55. KfH Kidney Center Traunstein, Traunstein, Germany

Carsten A. Böger

56. Center for Translational and Systems Neuroimmunology, Department of Neurology, Columbia University Medical Center, New York, NY, USA

Philip L. De Jager

57. Program in Medical and Population Genetics, Broad Institute, Cambridge, MA, USA

Philip L. De Jager

58. Medical School, National and Kapodistrian University Athens, Athens, Greece

Vasiliki Mamakou

59. Dromokaiteio Psychiatric Hospital, Athens, Greece

Vasiliki Mamakou

60. Clinical Pharmacology, William Harvey Research Institute, Queen Mary University of London, London, UK

Morris Brown

61. Department of Ophthalmology, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

Jost B. Jonas

62. Beijing Institute of Ophthalmology, Beijing Key Laboratory of Ophthalmology and Visual Sciences, Beijing Tongren Eye Center, Beijing Tongren Hospital, Capital Medical University, Beijing, China

Jost B. Jonas

63. Institute of Molecular and Clinical Ophthalmology Basel, Basel, Switzerland

Jost B. Jonas

64. Harvard Medical School, Boston, MA, USA

Daniel I. Chasman

65. Harvard Medical School, Boston, MA, USA

Josep M. Mercader, Julie E. Buring & Paul M. Ridker

66. Department of Human Genetics, David Geffen School of Medicine at UCLA, University of California, Los Angeles, Los Angeles, CA, USA

Päivi Pajukanta

67. Unidad de Biología Molecular y Medicina Genómica, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico, Mexico

Teresa Tusié-Luna

68. Instituto de Investigaciones Biomédicas, UNAM, Ciudad de México, Mexico, Mexico

Teresa Tusié-Luna

69. Unidad de Biología Molecular y Medicina Genómica, Instituto de Investigaciones Biomédicas UNAM/ Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico, Mexico

Teresa Tusié-Luna

70. Departamento de Endocrinología y Metabolismo, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico, Mexico

Carlos A. Aguilar-Salinas

71. Dirección de Nutrición and Unidad de Estudios de Enfermedades Metabólicas, Instituto Nacional de Ciencias Médicas y Nutrición Salvador Zubirán, Mexico, Mexico

Carlos A. Aguilar-Salinas

72. Department of Nutrition, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC, USA

Linda S. Adair & Penny Gordon-Larsen

73. Carolina Population Center, University of North Carolina, Chapel Hill, NC, USA

Linda S. Adair & Penny Gordon-Larsen

74. USC–Office of Population Studies Foundation, University of San Carlos, Cebu City, Philippines

Sonny Augustin Bechayda

75. Department of Anthropology, Sociology, and History, University of San Carlos, Cebu City, Philippines

Sonny Augustin Bechayda

76. Department of Medicine, Faculty of Medicine, University of Kelaniya, Ragama, Sri Lanka

H. Janaka de Silva

77. Department of Public Health, Faculty of Medicine, University of Kelaniya, Ragama, Sri Lanka

Ananda R. Wickremasinghe

78. Departments of Pediatrics and Medicine, University of California, San Francisco, San Francisco, CA, USA

Ronald M. Krauss

79. Institute of Biomedical Sciences, Academia Sinica, Taipei, Taiwan

Jer-Yuarn Wu

80. Systems Genomics Laboratory, School of Biotechnology, Jawaharlal Nehru University, New Delhi, India

Dwaipayan Bharadwaj

81. Department of Medicine, University of Mississippi Medical Center, Jackson, MS, USA

Adolfo Correa

82. Department of Physiology and Biophysics, University of Mississippi Medical Center, Jackson, MS, USA

James G. Wilson

83. Department of Medical Sciences, Uppsala University, Uppsala, Sweden

Lars Lind

84. Department of Paediatrics, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore

Chew-Kiat Heng

85. Khoo Teck Puat–National University Children’s Medical Institute, National University Health System, Singapore, Singapore

Chew-Kiat Heng

86. Department of Medicine, University of North Carolina, Chapel Hill, NC, USA

Amanda E. Nelson

87. Department of Epidemiology, Gillings School of Global Public Health,
University of North Carolina, Chapel Hill, NC, USA

Yvonne M. Golightly

88. Injury Prevention Research Center, University of North Carolina,
Chapel Hill, NC, USA

Yvonne M. Golightly

89. Division of Physical Therapy, University of North Carolina, Chapel
Hill, NC, USA

Yvonne M. Golightly

90. Department of Psychiatry, Amsterdam UMC, Vrije Universiteit
Amsterdam, Amsterdam, The Netherlands

Brenda Penninx

91. Amsterdam Public Health Research Institute, VU Medical Center
Amsterdam, Amsterdam, The Netherlands

Brenda Penninx & Eco J. C. de Geus

92. Department of Biochemistry, College of Medicine, Ewha Womans
University, Seoul, Korea

Hyung-Lae Kim

93. Faculty of Health and Medicine, University of Newcastle, Newcastle,
New South Wales, Australia

John Attia & Rodney J. Scott

94. Division of Biostatistics, Washington University School of Medicine,
St Louis, MO, USA

D. C. Rao

95. University of Kentucky, College of Public Health, Lexington, KY, USA

Donna K. Arnett

96. Institute of Cellular Medicine (Diabetes), The Medical School, Newcastle University, Newcastle upon Tyne, UK

Mark Walker

97. Department of Population Health, Finnish Institute for Health and Welfare, Helsinki, Finland

Heikki A. Koistinen

98. University of Helsinki and Department of Medicine, Helsinki University Hospital, Helsinki, Finland

Heikki A. Koistinen

99. Minerva Foundation Institute for Medical Research, Helsinki, Finland

Heikki A. Koistinen

100. Academy of Scientific and Innovative Research (AcSIR), New Delhi, India

Giriraj R. Chandak

101. Diabetology Research Centre, KEM Hospital and Research Centre, Pune, India

Chittaranjan S. Yajnik

102. Programs in Metabolism and Medical and Population Genetics, Broad Institute of MIT and Harvard, Cambridge, MA, USA

Josep M. Mercader

103. Diabetes Unit and Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA

Josep M. Mercader

104. Instituto Nacional de Salud Publica y Centro de Estudios en Diabetes, Cuernavaca, Mexico

Clicerio Gonzalez Villalpando

105. Instituto Nacional de Medicina Genómica, Mexico, Mexico

Lorena Orozco

106. Human Genetics Center, School of Public Health, University of Texas Health Science Center at Houston, Houston, TX, USA

Myriam Fornage

107. Yong Loo Lin School of Medicine, National University of Singapore and National University Health System, Singapore, Singapore

E. Shyong Tai & Rob M. van Dam

108. Kurume University School of Medicine, Kurume, Japan

Mitsuhiko Yokota

109. Genetics, Merck Sharp & Dohme, Kenilworth, NJ, USA

Dermot F. Reilly

110. Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford, UK

Mark I. McCarthy

111. Population Health and Genomics, University of Dundee, Ninwells Hospital and Medical School, Dundee, UK

Colin N. A. Palmer

112. Intramural Research Program, National Institute on Aging, Baltimore, MD, USA

Eleanor Simonsick

113. The Eye Hospital, School of Ophthalmology and Optometry, Wenzhou Medical University, Wenzhou, China

Fan Lu & Jia Qu

114. Shanghai Institute of Nutrition and Health, University of Chinese Academy of Sciences, Chinese Academy of Sciences, Shanghai, China

Shaofeng Huo, Huaixing Li & Xu Lin

115. Synlab Academy, Synlab, Mannheim, Germany

Winfried März

116. Clinical Institute of Medical and Chemical Laboratory Diagnostics, Medical University of Graz, Graz, Austria

Winfried März

117. Unidad de Investigacion Medica en Bioquimica, Hospital de Especialidades, Centro Medico Nacional Siglo XXI, Instituto Mexicano del Seguro Social, Mexico, Mexico

Miguel Cruz

118. Faculty of Medicine, University of Iceland, Reykjavik, Iceland

Vilmundur Gudnason, Unnur Thorsteinsdottir & Kari Stefansson

119. Department of Medicine, Faculty of Medecine, Université de Montréal, Quebec, Quebec, Canada

Jean-Claude Tardif

120. Department of Medicine, Faculty of Medicine, Université de Montréal, Montreal, Quebec, Canada

Guillaume Lettre

121. Department of Biomedical Data Sciences, Section Molecular Epidemiology, Leiden University Medical Center, Leiden, The Netherlands

Leen M. 't Hart

122. Amsterdam UMC, Department of General Practice and Elderly Care, Amsterdam Public Health Research Institute, Amsterdam, The Netherlands

Petra J. M. Elders

123. Department of Genetics, University of Pennsylvania, Philadelphia, PA, USA

Struan F. A. Grant & Daniel J. Rader

124. Department of Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA

Daniel J. Rader & Kyong-Mi Chang

125. Department of Surgery, University of Pennsylvania, Philadelphia, PA, USA

Scott M. Damrauer

126. Corporal Michael Crescenz VA Medical Center, Philadelphia, PA, USA

Scott M. Damrauer

127. Institute of Social and Economic Research, University of Essex, Essex, UK

Meena Kumari

128. The Mindich Child Health and Development Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA

Ruth J. F. Loos

129. Cardiovascular Health Research Unit, Departments of Medicine, Epidemiology and Health Services, University of Washington, Seattle, WA, USA

Bruce M. Psaty

130. Kaiser Permanent Washington Health Research Institute, Seattle, WA, USA

Bruce M. Psaty

131. Institute of Regional Health Research, University of Southern Denmark, Odense, Denmark

Ivan Brandlund

132. Danish Aging Research Center, University of Southern Denmark, Odense, Denmark

Kaare Christensen

133. Public Health, Faculty of Medicine, University of Helsinki, Helsinki, Finland

Samuli Ripatti

134. Broad Institute of MIT and Harvard, Cambridge, MA, USA

Samuli Ripatti

135. Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, PA, USA

Hakon Hakonarson

136. Department of Pediatrics, Perelman School of Medicine, The University of Pennsylvania, Philadelphia, PA, USA

Hakon Hakonarson & Struan F. A. Grant

137. Division of Human Genetics, Children's Hospital of Philadelphia, Philadelphia, PA, USA

Struan F. A. Grant

138. Institute of Genetic Epidemiology, Department of Genetics and Pharmacology, Medical University of Innsbruck, Innsbruck, Austria

Florian Kronenberg

139. School of Medicine, Southern University of Science and Technology, Shenzhen, China

Dongfeng Gu

140. Institute for Cardiogenetics, University of Lübeck, DZHK (German Research Centre for Cardiovascular Research), Partner site Hamburg/Lübeck/Kiel, and University Heart Center Lübeck, Lübeck, Germany

Jeanette Erdmann

141. Deutsches Herzzentrum München, Klinik für Herz- und Kreislauferkrankungen, Technische Universität München, Munich, Germany

Heribert Schunkert

142. Deutsches Zentrum für Herz-Kreislauf-Forschung (DZHK) e.V.,
Partner site Munich Heart Alliance, Munich, Germany

Heribert Schunkert

143. Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

Allan Linneberg

144. Netherlands Heart Institute, Utrecht, The Netherlands

J. Wouter Jukema

145. Division of Cardiology, Department of Medicine, Massachusetts General Hospital, Boston, MA, USA

Amit V. Khera

146. Program of Medical and Population Genetics, Broad Institute, Cambridge, MA, USA

Amit V. Khera & Sekar Kathiresan

147. Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA

Amit V. Khera

148. Department of Medicine, Harvard Medical School, Boston, MA, USA

Amit V. Khera & Sekar Kathiresan

149. Northern Finland Birth Cohorts, Infrastructure for Population Studies, Faculty of Medicine, University of Oulu, Oulu, Finland

Minna Männikkö

150. Center for Life Course Health Research, Faculty of Medicine,
University of Oulu, Oulu, Finland

Marjo-Riitta Jarvelin

151. Biocenter of Oulu, University of Oulu, Oulu, Finland

Marjo-Riitta Jarvelin

152. University Center for Primary Care and Public Health, Lausanne,
Switzerland

Zoltan Katalik

153. Institute for Genetic and Biomedical Research, Italian National
Council of Research (IRGB CNR), Cagliari, Italy

Francesco Cucca

154. University of Sassari, Sassari, Italy

Francesco Cucca

155. Department of Clinical Epidemiology, Leiden University Medical
Center, Leiden, The Netherlands

Dennis O. Mook-Kanamori

156. Department of Public Health and Primary Care, Leiden University
Medical Center, Leiden, The Netherlands

Dennis O. Mook-Kanamori

157. Department of Internal Medicine, Division of Endocrinology, Leiden
University Medical Center, Leiden, the Netherlands

Ko Willems van Dijk

158. Einthoven Laboratory for Experimental Vascular Medicine, Leiden University Medical Center, Leiden, the Netherlands

Ko Willems van Dijk

159. Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands

Ko Willems van Dijk

160. Division of Cardiovascular Medicine, Radcliffe Department of Medicine, John Radcliffe Hospital, University of Oxford, Oxford, UK

Hugh Watkins

161. Population Health Research Institute, St George's University of London, London, UK

David P. Strachan

162. National Heart and Lung Institute, Imperial College London, London, UK

Peter Sever

163. School of Public Health, Imperial College London, London, UK

Neil Poulter

164. OCDEM, University of Oxford, Churchill Hospital, Oxford, UK

Fredrik Karpe & Matt J. Neville

165. NIHR Oxford Biomedical Research Centre, Churchill Hospital, Oxford, UK

Fredrik Karpe & Matt J. Neville

166. Ophthalmology and Visual Sciences Academic Clinical Program (Eye ACP), Duke–NUS Medical School, Singapore, Singapore

Ching-Yu Cheng, Tien-Yin Wong & Charumathi Sabanayagam

167. DZHK (German Centre for Cardiovascular Research), Munich Heart Alliance Partner Site, Munich, Germany

Annette Peters

168. University of Exeter Medical School, University of Exeter, Exeter, UK

Andrew T. Hattersley

169. Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

Nancy L. Pedersen & Patrik K. E. Magnusson

170. Amsterdam Public Health, VU Medical Center Amsterdam, Amsterdam, The Netherlands

Dorret I. Boomsma

171. Amsterdam Reproduction and Development Research Institute, VU Medical Center Amsterdam, Amsterdam, the Netherlands

Dorret I. Boomsma

172. Framingham Heart Study, National Heart, Lung, and Blood Institute, US National Institutes of Health, Bethesda, MD, USA

L. Adrienne Cupples

173. Department of Genetics, School of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran

Mohsen Ghanbari

174. Department of Genetics, Shanghai-MOST Key Laboratory of Health and Disease Genomics, Chinese National Human Genome Center at Shanghai, Shanghai, China

Wei Huang

175. TUM School of Medicine, Technical University of Munich (TUM) and Klinikum Rechts der Isar, Munich, Germany

Eleftheria Zeggini

176. Institute of Clinical Medicine, Internal Medicine, University of Eastern Finland and Kuopio University Hospital, Kuopio, Finland

Johanna Kuusisto & Markku Laakso

177. Stanford Cardiovascular Institute, Stanford University, Stanford, CA, USA

Erik Ingelsson

178. Stanford Diabetes Research Center, Stanford University, Stanford, CA, USA

Erik Ingelsson

179. Regeneron Pharmaceuticals, Tarrytown, NY, USA

Goncalo Abecasis & Aris Baras

180. Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

John C. Chambers

181. Imperial College Healthcare NHS Trust, Imperial College London, London, UK

John C. Chambers

182. MRC–PHE Centre for Environment and Health, Imperial College London, London, UK

Jaspal S. Kooner

183. National Heart and Lung Institute, Imperial College London, London, UK

Jaspal S. Kooner

184. Institute for Minority Health Research, University of Illinois College of Medicine, Chicago, IL, USA

Martha Daviglus

185. Department of Biostatistics, Harvard T. H. Chan School of Public Health, Boston, MA, USA

Peter Kraft

186. QIMR Berghofer Medical Research Institute, Brisbane, Queensland, Australia

Nicholas G. Martin & John B. Whitfield

187. Faisalabad Institute of Cardiology, Faisalabad, Pakistan

Shahid Abbas

188. Department of Medicine, Columbia University Irving Medical Center, New York, NY, USA

Danish Saleheen

189. Department of Cardiology, Columbia University Irving Medical Center, New York, NY, USA

Danish Saleheen

190. Big Data Institute, University of Oxford, Oxford, UK
Robin G. Walters
191. National Institute for Health Research Oxford Biomedical Research Centre, Oxford University Hospitals, Oxford, UK
Michael V. Holmes
192. Aberdeen Centre for Health Data Science, School of Medicine, Medical Science and Nutrition, University of Aberdeen, Aberdeen, UK
Corri Black
193. Division of Population Health and Genomics, Ninewells Hospital and Medical School, University of Dundee, Dundee, UK
Blair H. Smith
194. Biomedical and Translational Informatics, Geisinger Health, Danville, PA, USA
Anne E. Justice
195. Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN, USA
Wei-Qi Wei
196. Departments of Medicine (Medical Genetics) and Genome Sciences, University of Washington Medical Center, Seattle, WA, USA
Gail P. Jarvik
197. Center for Autoimmune Genomics and Etiology, Cincinnati Children's Hospital Medical Center (CCHMC), Cincinnati, OH, USA
Bahram Namjou

198. Division of Endocrinology, Metabolism, and Molecular Medicine,
Department of Medicine, Northwestern University, Feinberg School of
Medicine, Chicago, IL, USA

M. Geoffrey Hayes

199. Department of Anthropology, Northwestern University, Evanston, IL,
USA

M. Geoffrey Hayes

200. Center for Genetic Medicine, Northwestern University, Feinberg
School of Medicine, Chicago, IL, USA

M. Geoffrey Hayes

201. HUNT Research Centre, Department of Public Health and Nursing,
NTNU, Norwegian University of Science and Technology, Levanger,
Norway

Kristian Hveem & Bjørn Olav Åsvold

202. Department of Medicine, Levanger Hospital, Nord-Trøndelag Hospital
Trust, Levanger, Norway

Kristian Hveem

203. Department of Endocrinology, St Olavs Hospital, Trondheim
University Hospital, Trondheim, Norway

Bjørn Olav Åsvold

204. RIKEN Center for Integrative Medical Sciences, Yokohama, Japan

Michiaki Kubo

205. Laboratory of Complex Trait Genomics, Department of Computational
Biology and Medical Sciences, Graduate School of Frontier Sciences,
The University of Tokyo, Tokyo, Japan

Yoichiro Kamatani

206. Laboratory of Statistical Immunology, WPI Immunology Frontier Research Center, Osaka University, Osaka, Japan

Yukinori Okada

207. Integrated Frontier Research for Medical Science Division, Institute for Open and Transdisciplinary Research Initiatives, Osaka University, Osaka, Japan

Yukinori Okada

208. Division of Molecular Pathology, Institute of Medical Science, The University of Tokyo, Tokyo, Japan

Yoshinori Murakami

209. VA Boston Healthcare System, Boston, MA, USA

Yuk-Lam Ho, Kelly Cho, Christopher J. O'Donnell & John M. Gaziano

210. VA Informatics and Computing Infrastructure, VA Salt Lake City Health Care System, Salt Lake City, UT, USA

Julie A. Lynch

211. University of Massachusetts, Boston, MA, USA

Julie A. Lynch

212. Cardiovascular Institute, Stanford University School of Medicine, Stanford, CA, USA

Philip S. Tsao & Themistocles L. Assimes

213. Corporal Michael J. Crescenz VA Medical Center, Philadelphia, PA, USA

Kyong-Mi Chang

214. Department of Medicine, Brigham Women's Hospital, Boston, MA, USA

Kelly Cho, Christopher J. O'Donnell & John M. Gaziano

215. Atlanta VA Medical Center, Atlanta, GA, USA

Peter Wilson

216. Division of Cardiology, Emory University School of Medicine, Atlanta, GA, USA

Peter Wilson

217. School of Electrical and Information Engineering, University of the Witwatersrand, Johannesburg, South Africa

Scott Hazelhurst

218. Division of Human Genetics, National Health Laboratory Service and School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

Michèle Ramsay

219. School of Basic and Medical Biosciences, Faculty of Life Sciences and Medicine, King's College London, London, UK

Richard C. Trembath

220. Departments of Pediatrics and Genetics, Harvard Medical School, Boston, MA, USA

Joel N. Hirschhorn

221. Center for Genomic Medicine, Department of Medicine, Massachusetts General Hospital, Boston, MA, USA

Sekar Kathiresan

222. Cardiology Division, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

Pradeep Natarajan

223. Department of Medicine, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

Pradeep Natarajan

224. Cardiovascular Research Center and Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA

Pradeep Natarajan

225. Centre for Genetics and Genomics Versus Arthritis, Centre for Musculoskeletal Research, Division of Musculoskeletal and Dermatological Sciences, The University of Manchester, Manchester, UK

Andrew P. Morris

226. Princess Al-Jawhara Al-Brahim Centre of Excellence in Research of Hereditary Disorders (PACER-HD), King Abdulaziz University, Jeddah, Saudi Arabia

Panos Deloukas & Cristen J. Willer

227. Department of Human Genetics, University of Michigan, Ann Arbor, MI, USA

Cristen J. Willer

Consortia

VA Million Veteran Program

- Shoa L. Clarke
- , Qin Hui
- , Derek Klarin
- , Xiang Zhu
- , Scott M. Damrauer
- , Yuk-Lam Ho
- , Julie A. Lynch
- , Daniel J. Rader
- , Philip S. Tsao
- , Kyong-Mi Chang
- , Kelly Cho
- , Christopher J. O'Donnell
- , John M. Gaziano
- , Peter Wilson
- , Themistocles L. Assimes
- , Yan V. Sun
- & Ioanna Ntalla

Global Lipids Genetics Consortium*

- Sarah E. Graham
- , Shoa L. Clarke
- , Kuan-Han H. Wu
- , Stavroula Kanoni
- , Greg J. M. Zajac
- , Shweta Ramdas
- , Ida Surakka
- , Ioanna Ntalla
- , Sailaja Vedantam
- , Thomas W. Winkler
- , Adam E. Locke
- , Eirini Marouli
- , Mi Yeong Hwang
- , Sohee Han

- , Akira Narita
- , Ananyo Choudhury
- , Amy R. Bentley
- , Kenneth Ekoru
- , Anurag Verma
- , Bhavi Trivedi
- , Hilary C. Martin
- , Karen A. Hunt
- , Qin Hui
- , Derek Klarin
- , Xiang Zhu
- , Gudmar Thorleifsson
- , Anna Helgadottir
- , Daniel F. Gudbjartsson
- , Hilma Holm
- , Isleifur Olafsson
- , Masato Akiyama
- , Saori Sakaue
- , Chikashi Terao
- , Masahiro Kanai
- , Wei Zhou
- , Ben M. Brumpton
- , Humaira Rasheed
- , Sanni E. Ruotsalainen
- , Aki S. Havulinna
- , Yogasudha Veturi
- , QiPing Feng
- , Elisabeth A. Rosenthal
- , Todd Lingren
- , Jennifer Allen Pacheco
- , Sarah A. Pendergrass
- , Jeffrey Haessler
- , Franco Giulianini
- , Yuki Bradford
- , Jason E. Miller
- , Archie Campbell
- , Kuang Lin

- , Iona Y. Millwood
- , George Hindy
- , Asif Rasheed
- , Jessica D. Faul
- , Wei Zhao
- , David R. Weir
- , Constance Turman
- , Hongyan Huang
- , Mariaelisa Graff
- , Anubha Mahajan
- , Michael R. Brown
- , Weihua Zhang
- , Ketian Yu
- , Ellen M. Schmidt
- , Anita Pandit
- , Stefan Gustafsson
- , Xianyong Yin
- , Jian'an Luan
- , Jing-Hua Zhao
- , Fumihiko Matsuda
- , Hye-Mi Jang
- , Kyungheon Yoon
- , Carolina Medina-Gomez
- , Achilleas Pitsillides
- , Jouke Jan Hottenga
- , Gonneke Willemse
- , Andrew R. Wood
- , Yingji Ji
- , Zishan Gao
- , Simon Haworth
- , Ruth E. Mitchell
- , Jin Fang Chai
- , Mette Aadahl
- , Jie Yao
- , Ani Manichaikul
- , Helen R. Warren
- , Julia Ramirez

- , Jette Bork-Jensen
- , Line L. Kårhus
- , Anuj Goel
- , Maria Sabater-Lleal
- , Raymond Noordam
- , Carlo Sidore
- , Edoardo Fiorillo
- , Aaron F. McDaid
- , Pedro Marques-Vidal
- , Matthias Wielscher
- , Stella Trompet
- , Naveed Sattar
- , Line T. Møllehave
- , Betina H. Thuesen
- , Matthias Munz
- , Lingyao Zeng
- , Jianfeng Huang
- , Bin Yang
- , Alaitz Poveda
- , Azra Kurbasic
- , Claudia Lamina
- , Lukas Forer
- , Markus Scholz
- , Tessel E. Galesloot
- , Jonathan P. Bradfield
- , E. Warwick Daw
- , Joseph M. Zmuda
- , Jonathan S. Mitchell
- , Christian Fuchsberger
- , Henry Christensen
- , Jennifer A. Brody
- , Mary F. Feitosa
- , Mary K. Wojczynski
- , Michael Preuss
- , Massimo Mangino
- , Paraskevi Christofidou
- , Niek Verweij

- , Jan W. Benjamins
- , Jorgen Engmann
- , Rachel L. Kember
- , Roderick C. Slieker
- , Ken Sin Lo
- , Nuno R. Zilhao
- , Phuong Le
- , Marcus E. Kleber
- , Graciela E. Delgado
- , Shaofeng Huo
- , Daisuke D. Ikeda
- , Hiroyuki Iha
- , Jian Yang
- , Jun Liu
- , Hampton L. Leonard
- , Jonathan Marten
- , Börge Schmidt
- , Marina Arendt
- , Laura J. Smyth
- , Marisa Cañadas-Garre
- , Chaolong Wang
- , Masahiro Nakatuchi
- , Andrew Wong
- , Nina Hutil-Kähönen
- , Xueling Sim
- , Rui Xia
- , Alicia Huerta-Chagoya
- , Juan Carlos Fernandez-Lopez
- , Valeriya Lyssenko
- , Meraj Ahmed
- , Anne U. Jackson
- , Marguerite R. Irvin
- , Christopher Oldmeadow
- , Han-Na Kim
- , Seungho Ryu
- , Paul R. H. J. Timmers
- , Liubov Arbeeva

- , Rajkumar Dorajoo
- , Leslie A. Lange
- , Xiaoran Chai
- , Gauri Prasad
- , Laura Lorés-Motta
- , Marc Pauper
- , Jirong Long
- , Xiaohui Li
- , Elizabeth Theusch
- , Fumihiko Takeuchi
- , Cassandra N. Spracklen
- , Anu Loukola
- , Sailalitha Bollepalli
- , Sophie C. Warner
- , Ya Xing Wang
- , Wen B. Wei
- , Teresa Nutile
- , Daniela Ruggiero
- , Yun Ju Sung
- , Yi-Jen Hung
- , Shufeng Chen
- , Fangchao Liu
- , Jingyun Yang
- , Katherine A. Kentistou
- , Mathias Gorski
- , Marco Brumat
- , Karina Meidtner
- , Lawrence F. Bielak
- , Jennifer A. Smith
- , Prashantha Hebbar
- , Aliki-Eleni Farmaki
- , Edith Hofer
- , Maoxuan Lin
- , Chao Xue
- , Jifeng Zhang
- , Maria Pina Concas
- , Simona Vaccargiu

- , Peter J. van der Most
- , Niina Pitkänen
- , Brian E. Cade
- , Jiwon Lee
- , Sander W. van der Laan
- , Kumaraswamy Naidu Chitrala
- , Stefan Weiss
- , Martina E. Zimmermann
- , Jong Young Lee
- , Hyeok Sun Choi
- , Maria Nethander
- , Sandra Freitag-Wolf
- , Lorraine Southam
- , Nigel W. Rayner
- , Carol A. Wang
- , Shih-Yi Lin
- , Jun-Sing Wang
- , Christian Couture
- , Leo-Pekka Lyytikäinen
- , Kjell Nikus
- , Gabriel Cuellar-Partida
- , Henrik Vestergaard
- , Bertha Hildalgo
- , Olga Giannakopoulou
- , Qiuyin Cai
- , Morgan O. Obura
- , Jessica van Setten
- , Xiaoyin Li
- , Karen Schwander
- , Natalie Terzikhan
- , Jae Hun Shin
- , Rebecca D. Jackson
- , Alexander P. Reiner
- , Lisa Warsinger Martin
- , Zhengming Chen
- , Liming Li
- , Heather M. Highland

- , Kristin L. Young
- , Takahisa Kawaguchi
- , Joachim Thiery
- , Joshua C. Bis
- , Girish N. Nadkarni
- , Lenore J. Launer
- , Huaixing Li
- , Mike A. Nalls
- , Olli T. Raitakari
- , Sahoko Ichihara
- , Sarah H. Wild
- , Christopher P. Nelson
- , Harry Campbell
- , Susanne Jäger
- , Toru Nabika
- , Fahd Al-Mulla
- , Harri Niinikoski
- , Peter S. Braund
- , Ivana Kolcic
- , Peter Kovacs
- , Tota Giardoglou
- , Tomohiro Katsuya
- , Konain Fatima Bhatti
- , Dominique de Kleijn
- , Gert J. de Borst
- , Eung Kweon Kim
- , Hieab H. H. Adams
- , M. Arfan Ikram
- , Xiaofeng Zhu
- , Folkert W. Asselbergs
- , Adriaan O. Kraaijeveld
- , Joline W. J. Beulens
- , Xiao-Ou Shu
- , Loukianos S. Rallidis
- , Oluf Pedersen
- , Torben Hansen
- , Paul Mitchell

- , Alex W. Hewitt
- , Mika Kähönen
- , Louis Pérusse
- , Claude Bouchard
- , Anke Tönjes
- , Yii-Der Ida Chen
- , Craig E. Pennell
- , Trevor A. Mori
- , Wolfgang Lieb
- , Andre Franke
- , Claes Ohlsson
- , Dan Mellström
- , Yoon Shin Cho
- , Hyejin Lee
- , Jian-Min Yuan
- , Woon-Puay Koh
- , Sang Youl Rhee
- , Jeong-Taek Woo
- , Iris M. Heid
- , Klaus J. Stark
- , Henry Völzke
- , Georg Homuth
- , Michele K. Evans
- , Alan B. Zonderman
- , Ozren Polasek
- , Gerard Pasterkamp
- , Imo E. Hoefer
- , Susan Redline
- , Katja Pahkala
- , Albertine J. Oldehinkel
- , Harold Snieder
- , Ginevra Biino
- , Reinhold Schmidt
- , Helena Schmidt
- , Y. Eugene Chen
- , Stefania Bandinelli
- , George Dedoussis

- , Thangavel Alphonse Thanaraj
- , Sharon L. R. Kardia
- , Norihiro Kato
- , Matthias B. Schulze
- , Giorgia Girotto
- , Bettina Jung
- , Carsten A. Böger
- , Peter K. Joshi
- , David A. Bennett
- , Philip L. De Jager
- , Xiangfeng Lu
- , Vasiliki Mamakou
- , Morris Brown
- , Mark J. Caulfield
- , Patricia B. Munroe
- , Xiuqing Guo
- , Marina Ciullo
- , Jost B. Jonas
- , Nilesh J. Samani
- , Jaakkko Kaprio
- , Päivi Pajukanta
- , Linda S. Adair
- , Sonny Augustin Bechayda
- , H. Janaka de Silva
- , Ananda R. Wickremasinghe
- , Ronald M. Krauss
- , Jer-Yuarn Wu
- , Wei Zheng
- , Anneke I. den Hollander
- , Dwaipayan Bharadwaj
- , Adolfo Correa
- , James G. Wilson
- , Lars Lind
- , Chew-Kiat Heng
- , Amanda E. Nelson
- , Yvonne M. Golightly
- , James F. Wilson

- , Brenda Penninx
- , Hyung-Lae Kim
- , John Attia
- , Rodney J. Scott
- , D. C. Rao
- , Donna K. Arnett
- , Mark Walker
- , Heikki A. Koistinen
- , Giriraj R. Chandak
- , Chittaranjan S. Yajnik
- , Josep M. Mercader
- , Teresa Tusié-Luna
- , Carlos A. Aguilar-Salinas
- , Clicerio Gonzalez Villalpando
- , Lorena Orozco
- , Myriam Fornage
- , E. Shyong Tai
- , Rob M. van Dam
- , Terho Lehtimäki
- , Nish Chaturvedi
- , Mitsuhiro Yokota
- , Jianjun Liu
- , Dermot F. Reilly
- , Amy Jayne McKnight
- , Frank Kee
- , Karl-Heinz Jöckel
- , Mark I. McCarthy
- , Colin N. A. Palmer
- , Veronique Vitart
- , Caroline Hayward
- , Eleanor Simonsick
- , Cornelia M. van Duijn
- , Fan Lu
- , Jia Qu
- , Haretsugu Hishigaki
- , Xu Lin
- , Winfried März

- , Esteban J. Parra
- , Miguel Cruz
- , Vilmundur Gudnason
- , Jean-Claude Tardif
- , Guillaume Lettre
- , Leen M. 't Hart
- , Petra J. M. Elders
- , Scott M. Damrauer
- , Meena Kumari
- , Mika Kivimaki
- , Pim van der Harst
- , Tim D. Spector
- , Ruth J. F. Loos
- , Michael A. Province
- , Bruce M. Psaty
- , Ivan Brandslund
- , Peter P. Pramstaller
- , Kaare Christensen
- , Samuli Ripatti
- , Elisabeth Widén
- , Hakon Hakonarson
- , Struan F. A. Grant
- , Lambertus A. L. M. Kiemeney
- , Jacqueline de Graaf
- , Markus Loeffler
- , Florian Kronenberg
- , Dongfeng Gu
- , Jeanette Erdmann
- , Heribert Schunkert
- , Paul W. Franks
- , Allan Linneberg
- , J. Wouter Jukema
- , Amit V. Khera
- , Minna Männikkö
- , Marjo-Riitta Jarvelin
- , Zoltan Kutalik
- , Francesco Cucca

- , Dennis O. Mook-Kanamori
- , Ko Willems van Dijk
- , Hugh Watkins
- , David P. Strachan
- , Niels Grarup
- , Peter Sever
- , Neil Poulter
- , Jerome I. Rotter
- , Thomas M. Dantoft
- , Fredrik Karpe
- , Matt J. Neville
- , Nicholas J. Timpson
- , Ching-Yu Cheng
- , Tien-Yin Wong
- , Chiea Chuen Khor
- , Charumathi Sabanayagam
- , Annette Peters
- , Christian Gieger
- , Andrew T. Hattersley
- , Nancy L. Pedersen
- , Patrik K. E. Magnusson
- , Dorret I. Boomsma
- , Eco J. C. de Geus
- , L. Adrienne Cupples
- , Joyce B. J. van Meurs
- , Mohsen Ghanbari
- , Penny Gordon-Larsen
- , Wei Huang
- , Young Jin Kim
- , Yasuharu Tabara
- , Nicholas J. Wareham
- , Claudia Langenberg
- , Eleftheria Zeggini
- , Johanna Kuusisto
- , Markku Laakso
- , Erik Ingelsson
- , Goncalo Abecasis

- , John C. Chambers
- , Jaspal S. Kooner
- , Paul S. de Vries
- , Alanna C. Morrison
- , Kari E. North
- , Martha Daviglus
- , Peter Kraft
- , Nicholas G. Martin
- , John B. Whitfield
- , Shahid Abbas
- , Danish Saleheen
- , Robin G. Walters
- , Michael V. Holmes
- , Corri Black
- , Blair H. Smith
- , Anne E. Justice
- , Aris Baras
- , Julie E. Buring
- , Paul M. Ridker
- , Daniel I. Chasman
- , Charles Kooperberg
- , Wei-Qi Wei
- , Gail P. Jarvik
- , Bahram Namjou
- , M. Geoffrey Hayes
- , Marylyn D. Ritchie
- , Pekka Jousilahti
- , Veikko Salomaa
- , Kristian Hveem
- , Bjørn Olav Åsvold
- , Michiaki Kubo
- , Yoichiro Kamatani
- , Yukinori Okada
- , Yoshinori Murakami
- , Unnur Thorsteinsdottir
- , Kari Stefansson
- , Yuk-Lam Ho

- , Julie A. Lynch
- , Daniel J. Rader
- , Philip S. Tsao
- , Kyong-Mi Chang
- , Kelly Cho
- , Christopher J. O'Donnell
- , John M. Gaziano
- , Peter Wilson
- , Charles N. Rotimi
- , Scott Hazelhurst
- , Michèle Ramsay
- , Richard C. Trembath
- , David A. van Heel
- , Gen Tamiya
- , Masayuki Yamamoto
- , Bong-Jo Kim
- , Karen L. Mohlke
- , Timothy M. Frayling
- , Joel N. Hirschhorn
- , Sekar Kathiresan
- , Michael Boehnke
- , Pradeep Natarajan
- , Gina M. Peloso
- , Christopher D. Brown
- , Andrew P. Morris
- , Themistocles L. Assimes
- , Panos Deloukas
- , Yan V. Sun
- & Cristen J. Willer

Contributions

S.L.C., K.-H.H.W., S. Kanoni, G.J.M.Z. and S. Ramdas contributed equally to this work as co-second authors. All authors reviewed the manuscript. Consortium management: G.M.P., P.N., T.L.A., M.B., S.Kathireshan and C.J.W. Study design, interpretation of results and drafting of the manuscript: S.E.G., S.L.C., K.-H.H.W., S. Kanoni, G.J.M.Z., S. Ramdas, I.S., I.N., E.M., K.L.M., T.M.F., J.N.H., S. Kathiresan, M. Boehnke, P.N., G.M.P., C.D.B., A.P.M., Y.V.S., P.D., T.L.A. and C.J.W. Primary meta-analysis and QC: S.E.G., S. Vedantam, T.W.W. and A.E.L. PRS analysis and development: S.E.G., S.L.C., K.-H.H.W., S. Kanoni, M.Y.H., S.H., A.N., A. Choudhury, A.R.B., K.E., A.V., B.T., H.C.M., K.A.H., C.N.R., S.H., M.R., R.C.T., D.A.v.H., G.T., M.Y. and B.-J.K. Individual study genetic analysis: S.E.G., S. Kanoni, S. Vedantam, A.E.L., K.L.M., G.M.P., P.D., C.J.W., Q.H., D.K., X. Zhu, G.T., A. Helgadottir, D.F.G., H. Holm, I.O., M. Akiyama, S.S., C. Terao, M. Kanai, W. Zhou, B.M.B., H.R., S.E.R., A.S.H., Y.V., Q.F., E.A.R., T. Lingren, J.A.P., S.A.P., J. Haessler, F.G., Y.B., J.E.M., A. Campbell, K. Lin, I.Y.M., G. Hindy, A.R., J.D.F., W. Zhao, D.R.W., C. Turman, H. Huang, M. Graff, A. Mahajan, M.R.B., W. Zhang, K. Yu, E.M.S., A. Pandit, S.G., X.Y., J. Luan, J.-H. Zhao, F.M., H.-M.J., K. Yoon, C.M.-G., A. Pitsillides, J.J.H., G.W., A.R. Wood, Y.J., Z.G., S. Haworth, R.E.M., J.F.C., M. Aadahl, J.Yao, A. Manichaikul, H.R.W., J.R., J.B.-J., L.L.K., A.G., M.S.-L., R.N., C. Sidore, E.F., A.F.M., P.M.-V., M. Wielscher, S.T., N.S., L.T.M., B.H.T., M. Munz, L.Z., J. Huang, B.Y., A. Poveda, A.K., C. Lamina, L.F., M. Scholz, T.E.G., J.P.B., E.W.D., J.M.Z., J.S.M., C.F., H. Christensen, J.A.B., M.F.F., M.K.W., M. Preuss, M. Mangino, P.C., N.V., J.W. Benjamins, J. Engmann, R.L.K., R.C.S., K.S.L., N.R.Z., P.L., M.E.K., G.E.D., S. Huo, D.D.I., H.I., J. Yang, Jun Liu, H.L.L., J.M., B.S., M. Arendt, L.J.S., M.C.-G., C.W., M. Nakatuchi, A.W., N.H.-K., X.Sim, R.X., A.H.-C., J.C.F.-L., V.L., M. Ahmed, A.U.J., M.R.I., C. Oldmeadow, H.-N.K., S. Ryu, P.R.H.J.T., L.A., R.D., L.A.L., X.C., G. Prasad, L.L.-M., M. Pauper, J. Long, X. Li, E. Theusch, F.T., C.N.S., A. Loukola, S. Bollepalli, S.C.W., Y.X.W., W.B.W., T. Nutile, D. Ruggiero, Y.J.S., Y.-J. Hung, S.C., F. Liu, Jingyun Yang, K.A.K., M. Gorski, M. Brumat, K.M., L.F.B., J.A.S., P.H., A.-E.F., E.H., M. Lin, C.X., J. Zhang, M.P.C., S. Vaccargiu, P.J.v.d.M., N. Pitkänen, B.E.C., J. Lee,

S.W.v.d.L., K.N.C., S.W., M.E.Z., J.Y.L., H.S.C., M. Nethander, S.F.-W., L.S., N.W.R., C.A.W., S.-Y.L., J.-S.W., C. Couture, L.-P.L., K.N., G.C.-P., H. Vestergaard, B.H., O.G., Q.C., M.O.O., J.v.S., Xiaoyin Li, K. Schwander, N.T., J.H.S., R.D.J., A.P.R., L.W.M., Z.C., L.Li, H.M.H., K.L.Y., T. Kawaguchi, J. Thiery, J.C.B., G.N.N., L.J.L., H.Li, M.A.N., O.T.R., S.I., S.H.W., C.P.N., H. Campbell, S.J., T. Nabika, F.A.-M., H.N., P.S.B., I.K., P. Kovacs, T.G., T. Katsuya, K.F.B., D.d.K., G.J.d.B., E.K.K., H.H.H.A., M.A.I., Xiaofeng Zhu, F.W.A., A.O.K., J.W.J.B., X.-O.S., L.S.R., O. Pedersen, T.H., P. Mitchell, A.W.H., M. Kkähönen, L.P., C. Bouchard, A.T., Y.-D.I.C., C.E.P., T.A.M., W.L., A. Franke, C. Ohlsson, D.M., Y.S.C., H. Lee, J.-M.Y., W.-P.K., S.Y.R., J.-T.W., I.M.H., K.J.S., H. Völzke, G. Homuth, M.K.E., A.B.Z., O. Polasek, G. Pasterkamp, I.E.H., S. Redline, K.P., A.J.O., H. Snieder, G.B., R.S., H. Schmidt, Y.E.C., S. Bandinelli, G. Dedoussis, T.A.T., S.L.R.K., N.K., M.B.S., G.G., B.J., C.A.B., P.K.J., D.A.B., P.L.D.J., X. Lu, V.M., M. Brown, M.J.C., P.B.M., X.G., M. Ciullo, J.B.J., N.J.S., J. Kaprio, P.P., L.S.A., S.A.B., H.J.d.S., A.R.W., R.M.K., J.-Y.W., W. Zheng, A.I.d.H., D.B., A. Correa, J.G.W., L. Lind, C.-K.H., A.E.N., Y.M.G., J.F.W., B.P., H.-L.K., J.A., R.J.S., D.C.R., D.K.A., M. Walker, H.A.K., G.R.C., C.S.Y., J.M.M., T.T.-L., C.A.-S., C.G.V., L.O., M.F., E.S.T., R.M.v.D., T. Lehtimäki, N.C., M.Y., Jianjun Liu, D.F.R., A.J.M., F. Kee, K.-H.J., M.I.M., C.N.A.P., V.V., C. Hayward, E.S., C.M.v.D., F. Lu, J.Q., H. Hishigaki, X. Lin, W.M., E.J.P., M. Cruz, V.G., J.-C.T., G.L., L.M.t.H., P.J.M.E., S.M.D., M. Kumari, M. Kivimaki, P.v.d.H., T.D.S., R.J.F.L., M.A.P., B.M.P., I.B., P.P.P., K. Christensen, S. Ripatti, E.W., H. Hakonarson, S.F.A.G., L.A.L.M.K., J.d.G., M. Loeffler, F. Kronenberg, D.G., J. Erdmann, H. Schunkert, P.W.F., A. Linneberg, J.W.J., A.V.K., M. Männikkö, M.-R.J., Z.K., F.C., D.O.M.-K., K.W.v.D., H.W., D.P.S., N.G., P.S., N. Poulter, J.I.R., T.M.D., F. Karpe, M.J.N., N.J.T., C.-Y.C., T.-Y.W., C.C.K., C. Sabanayagam, A. Peters, C.G., A.T.H., N.L.P., P.K.E.M., D.I.B., E.J.C.d.G., L.A.C., J.B.J.v.M., M. Ghanbari, P.G.-L., W.H., Y.J.K., Y.T., N.J.W., C. Langenberg, E.Z., J. Kuusisto, M. Laakso, E.I., G.A., J.C.C., J.S.K., P.S.d.V., A.C.M., K.E.N., M.D., P. Kraft, N.G.M., J.B.W., S.A., D.S., R.G.W., M.V.H., C.Black, B.H.S., A.E.J., A.B., J.E.B., P.M.R., D.I.C., C. Kooperberg, W.-Q.W., G.P.J., B.N., M.G.H., M.D.R., P.J., V.S., K.H., B.O.A., M. Kubo, Y. Kamatani, Y.O., Y.M., U.T., K. Stefansson, Y.-L.H., J.A.L., D. Rader, P.S.T., K.-M.C., K. Cho, C.J.O., J.M.G., and P.W.

Corresponding authors

Correspondence to [Themistocles L. Assimes](#) or [Cristen J. Willer](#).

Ethics declarations

Competing interests

G.J.M.Z. is an employee of Incyte Corporation. G.C.-P. is currently an employee of 23andMe. M.J.C. is the Chief Scientist for Genomics England, a UK Government company. B.M.P. serves on the steering committee of the Yale Open Data Access Project funded by Johnson & Johnson. G.T., A. Helgadottir, D.F.G., H. Holm., U.T. and K. Stefansson are employees of deCODE/Amgen. V.S. has received honoraria for consultations from Novo Nordisk and Sanofi and has an ongoing research collaboration with Bayer. M. McCarthy has served on advisory panels for Pfizer, NovoNordisk and Zoe Global, has received honoraria from Merck, Pfizer, Novo Nordisk and Eli Lilly, and research funding from Abbvie, Astra Zeneca, Boehringer Ingelheim, Eli Lilly, Janssen, Merck, NovoNordisk, Pfizer, Roche, Sanofi Aventis, Servier and Takeda. M. McCarthy and A. Mahajan are employees of Genentech and are holders of Roche stock. M. Scholz receives funding from Pfizer for a project unrelated to this work. M.E.K. is employed by Synlab. W.M. has received grants from Siemens Healthineers, grants and personal fees from Aegerion Pharmaceuticals, grants and personal fees from Amgen, grants from AstraZeneca, grants and personal fees from Sanofi, grants and personal fees from Alexion Pharmaceuticals, grants and personal fees from BASF, grants and personal fees from Abbott Diagnostics, grants and personal fees from Numares, grants and personal fees from Berlin-Chemie, grants and personal fees from Akzea Therapeutics, grants from Bayer Vital, grants from bestbion dx, grants from Boehringer Ingelheim, grants from Immundiagnostik, grants from Merck Chemicals, grants from MSD Sharp and Dohme, grants from Novartis Pharma, grants from Olink Proteomics, and is employed by Synlab Holding Deutschland, all outside the submitted work. A.V.K. has served as a consultant to Sanofi, Medicines Company, Maze Pharmaceuticals, Navitor Pharmaceuticals, Verve Therapeutics, Amgen and Color Genomics;

received speaking fees from Illumina, the Novartis Institute for Biomedical Research; received sponsored research agreements from the Novartis Institute for Biomedical Research and IBM Research; and reports a patent related to a genetic risk predictor (20190017119). S.K. is an employee of Verve Therapeutics and holds equity in Verve Therapeutics, Maze Therapeutics, Catabasis and San Therapeutics. He is a member of the scientific advisory boards for Regeneron Genetics Center and Corvidia Therapeutics; he has served as a consultant for Acceleron, Eli Lilly, Novartis, Merck, Novo Nordisk, Novo Ventures, Ionis, Alnylam, Aegerion, Haug Partners, Noble Insights, Leerink Partners, Bayer Healthcare, Illumina, Color Genomics, MedGenome, Quest and Medscape; and reports patents related to a method of identifying and treating a person having a predisposition to or afflicted with cardiometabolic disease (20180010185) and a genetics risk predictor (20190017119). D.K. accepts consulting fees from Regeneron Pharmaceuticals. D.O.M.-K. is a part-time clinical research consultant for Metabolon. D.S. has received support from the British Heart Foundation, Pfizer, Regeneron, Genentech and Eli Lilly pharmaceuticals. The spouse of C.J.W. is employed by Regeneron.

Additional information

Peer review information *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

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 sizes of identified index variants from multi-ancestry meta-analysis.

Index variants associated with a) HDL cholesterol, b) LDL cholesterol, c) triglycerides, d) nonHDL cholesterol and e) total cholesterol include both common variants of small to moderate effect and low frequency variants of moderate to large effect.

Extended Data Fig. 2 Comparison of the number of index variants by sample size.

a) Comparison of the number of index variants reaching genome-wide significance ($p < 5 \times 10^{-8}$) from meta-analysis of LDL-C in each ancestry group. A meta-analysis of five random subsets of European cohorts selected to reach sample sizes of approximately 100,000, 200,000, 400,000, 600,000, or 800,000 individuals is also shown. b) Comparison of chi-squared values from meta-analysis of LDL-C for each possible combination of ancestry groups (without genomic-control correction) for variants with minor allele frequency (MAF) $\geq 5\%$. The colored lines indicate a linear regression model of all meta-analyses for a specific ancestry (eg. all analyses including European individuals). c) Comparison of chi-squared values from meta-analysis of LDL-C for variants with MAF $\leq 5\%$. d) Comparison of chi-squared valued for variants with MAF $\geq 5\%$ for LDL-C without genomic-control correction in a meta-analysis of all European cohorts as well as five subsets selected to reach sample sizes of approximately 100,000, 200,000, 400,000, 600,000, or 800,000 individuals.

Extended Data Fig. 3 Effect sizes by ancestry for unique index variants from ancestry-specific meta-analysis.

Comparison of effect sizes and standard errors for the 389 unique variants reaching genome-wide significance (p -value $< 5 \times 10^{-8}$ as given by

RAREMETAL) in two ancestry groups. Variants with discordant directions of effect between ancestries are labeled by chromosome and position (build 37). Association results for all index variants are given in [Supplementary Table 3](#). The red line depicts an equivalent European ancestry and non-European ancestry effect size while the black line depicts a linear regression model. $R^2 = 0.93$.

Extended Data Fig. 4 Comparison of credible set size.

The number of variants in the 99% credible sets for each association signal are compared between a) admixed African ancestry and multi-ancestry analysis and b) European ancestry and multi-ancestry analysis.

Extended Data Fig. 5 Overview of LDL-C polygenic score generation and validation.

Polygenic scores were calculated separately in each ancestry group (or in all ancestries) using either pruning and thresholding or PRS-CS. The polygenic scores were then taken forward for testing in ancestry-matched participants followed by validation in independent data sets.

Extended Data Fig. 6 Optimal polygenic score threshold by ancestry group for either PRS-CS or pruning and thresholding based LDL-C polygenic scores.

Adjusted R^2 estimated upon testing in UK Biobank ancestry-matched participants (who were not included in GWAS summary statistics). a) Admixed African, East Asian and South Asian ancestry polygenic scores. b) European and multi-ancestry polygenic scores. c) European ancestry (GLGC 2010) and multi-ancestry polygenic scores. d) All polygenic scores across all thresholds used for score construction. e) Comparison of adjusted R^2 across ancestry groups relative to the maximum for covariates alone, polygenic scores from PRS-CS or polygenic scores from pruning and thresholding.

Extended Data Fig. 7 Improvement in PRS performance in African Americans when starting with ancestry-mismatched European ancestry scores by updating weights, updating variant lists, or updating both variants and weights to be ancestry-matched.

By comparison to the gold-standard performance of the multi-ancestry-derived PRS in African Americans (adjusted R² = 0.12), a European ancestry derived score capture only 47% of the variance explained by the multi-ancestry PRS. When LD and association information from the target population is used to optimize the list of variants for inclusion in the PRS, but with ancestry-mismatched weights from European ancestry GWAS, the variance explained reaches 71% of the gold standard. If the PRS variant list selected in European ancestry individuals were genotyped in the target population, and PRS weights were updated using a GWAS from the target population, the variance explained reached 87% of the gold standard. Finally, deriving both the marker list and weights from the target population (single-ancestry GWAS of admixed African individuals) explained 94% of the variance relative to the gold-standard trans-ancestry PRS.

Extended Data Fig. 8 Comparison of PRS performance by admixture quartile.

We divided the testing cohorts into quartiles by proportion of African ancestry and estimated the performance of the PRS separately within each quartile in a) the Michigan Genomics Initiative ($N= 1,341$), and b) the Million Veteran Program ($N= 18,251$). Error bars represent 95% confidence intervals.

Supplementary information

Supplementary Information

This file contains acknowledgements for each cohort, VA Million Veteran Program and Global Lipids Genetics Consortium authors, Supplementary

Tables 2, 4, 8, 13 and 21–23, Supplementary Figs. 1–10, the Supplementary Notes and Supplementary Methods.

Reporting Summary

Peer Review File

Supplementary Tables

This file contains Supplementary Tables 1, 3, 5–7, 9–12 and 14–20.

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Graham, S.E., Clarke, S.L., Wu, K.H.H. *et al.* The power of genetic diversity in genome-wide association studies of lipids. *Nature* **600**, 675–679 (2021). <https://doi.org/10.1038/s41586-021-04064-3>

- Received: 15 September 2020
- Accepted: 27 September 2021
- Published: 09 December 2021
- Issue Date: 23 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04064-3>

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

Major cholesterol study reveals benefits of examining diverse populations

Research Briefing 09 Dec 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04064-3>

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

- Article
- [Published: 17 November 2021](#)

Cortical responses to touch reflect subcortical integration of LTMR signals

- [Alan J. Emanuel](#) ORCID: orcid.org/0000-0002-6479-4263^{1,2},
- [Brendan P. Lehnert](#) ORCID: orcid.org/0000-0001-7719-5924^{1,2},
- [Stefano Panzeri](#) ORCID: orcid.org/0000-0003-1700-8909^{3,4},
- [Christopher D. Harvey](#) ORCID: orcid.org/0000-0001-9850-2268¹ &
- [David D. Ginty](#) ORCID: orcid.org/0000-0001-9723-8530^{1,2}

Nature volume 600, pages 680–685 (2021)

- 3213 Accesses
- 29 Altmetric
- [Metrics details](#)

Subjects

- [Cortex](#)
- [Neural circuits](#)
- [Sensory processing](#)
- [Thalamus](#)
- [Touch receptors](#)

Abstract

Current models to explain how signals emanating from cutaneous mechanoreceptors generate representations of touch are based on comparisons of the tactile responses of mechanoreceptor subtypes and neurons in somatosensory cortex^{1,2,3,4,5,6,7,8}. Here we used mouse genetic manipulations to investigate the contributions of peripheral mechanoreceptor subtypes to cortical responses to touch. Cortical neurons exhibited remarkably homogeneous and transient responses to skin indentation that resembled rapidly adapting (RA) low-threshold mechanoreceptor (LTMR) responses. Concurrent disruption of signals from both A β RA-LTMRs and A β slowly adapting (SA)-LTMRs eliminated cortical responses to light indentation forces. However, disruption of either LTMR subtype alone caused opposite shifts in cortical sensitivity but otherwise largely unaltered tactile responses, indicating that both subtypes contribute to normal cortical responses. Selective optogenetic activation of single action potentials in A β RA-LTMRs or A β SA-LTMRs drove low-latency responses in most mechanically sensitive cortical neurons. Similarly, most somatosensory thalamic neurons were also driven by activation of A β RA-LTMRs or A β SA-LTMRs. These findings support a model in which signals from physiologically distinct mechanoreceptor subtypes are extensively integrated and transformed within the subcortical somatosensory system to generate cortical representations of touch.

This is a preview of subscription content

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Fig. 1: Hindpaw S1 responses to step indentations.

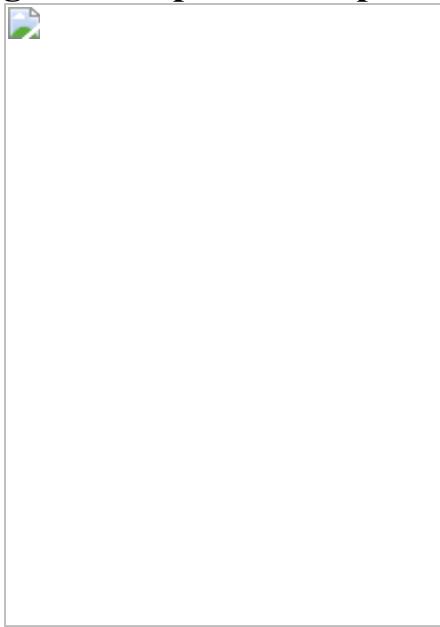


Fig. 2: S1 in mice lacking Meissner corpuscles and/or Merkel cells exhibits shifted sensitivity.

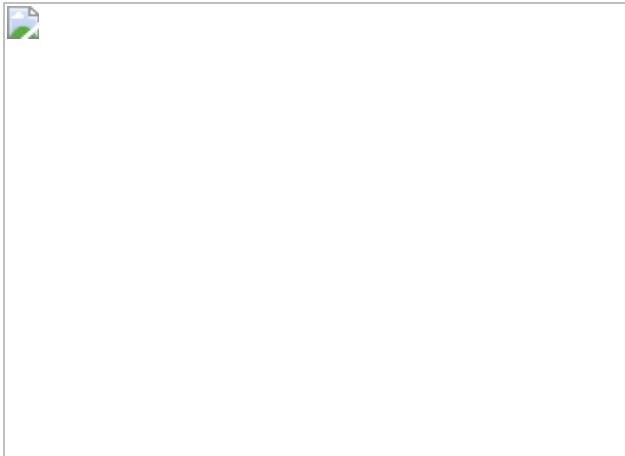


Fig. 3: Selective activation of A β LTMR subtypes drives the majority of mechanically sensitive S1 neurons.

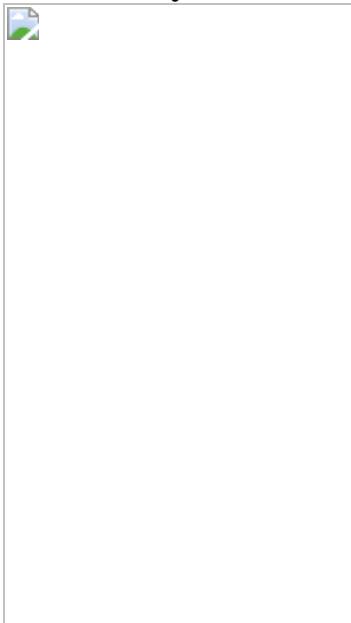
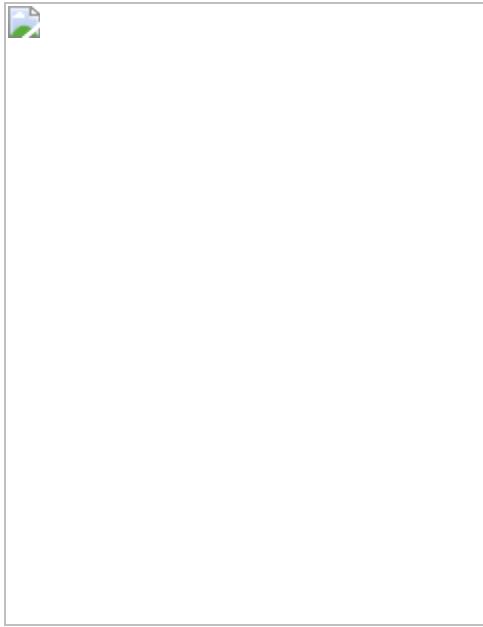


Fig. 4: Most VPL neurons receive convergent input from A β RA-LTMRs and A β SA-LTMRs.



Data availability

Data are available upon request to the corresponding authors. [Source data](#) are provided with this paper.

Code Availability

Code is available upon request to the corresponding authors. Analysis scripts are available at <https://github.com/ajemanuel/analyzeMEA>. [Source data](#) are provided with this paper.

References

1. 1.

Dykes, R. W., Rasmusson, D. D. & Hoeltzell, P. B. Organization of primary somatosensory cortex in the cat. *J. Neurophysiol.* **43**, 1527–1546 (1980).

2. 2.

Mountcastle, V. B. Modality and topographic properties of single neurons of cat's somatic sensory cortex. *J. Neurophysiol.* **20**, 408–434 (1957).

3. 3.

Paul, R. L., Merzenich, M. & Goodman, H. Representation of slowly and rapidly adapting cutaneous mechanoreceptors of the hand in Brodmann's areas 3 and 1 of *Macaca mulatta*. *Brain Res.* **36**, 229–249 (1972).

4. 4.

Phillips, J. R., Johnson, K. O. & Hsiao, S. S. Spatial pattern representation and transformation in monkey somatosensory cortex. *Proc. Natl Acad. Sci. USA* **85**, 1317–1321 (1988).

5. 5.

Sur, M., Wall, J. T. & Kaas, J. H. Modular segregation of functional cell classes within the postcentral somatosensory cortex of monkeys. *Science* **212**, 1059–1061 (1981).

6. 6.

Sur, M., Wall, J. T. & Kaas, J. H. Modular distribution of neurons with slowly adapting and rapidly adapting responses in area 3b of somatosensory cortex in monkeys. *J. Neurophysiol.* **51**, 724–744 (1984).

7. 7.

Kaas, J. H., Nelson, R. J., Sur, M., Dykes, R. W. & Merzenich, M. M. The somatotopic organization of the ventroposterior thalamus of the squirrel monkey, *Saimiri sciureus*. *J. Comp. Neurol.* **226**, 111–140 (1984).

8. 8.

Pei, Y. C., Denchev, P. V., Hsiao, S. S., Craig, J. C. & Bensmaia, S. J. Convergence of submodality-specific input onto neurons in primary somatosensory cortex. *J. Neurophysiol.* **102**, 1843–1853 (2009).

9. 9.

Handler, A. & Ginty, D. D. The mechanosensory neurons of touch and their mechanisms of activation. *Nat. Rev. Neurosci.* **22**, 521–537 (2021).

10. 10.

Johnson, K. O. The roles and functions of cutaneous mechanoreceptors. *Curr. Opin. Neurobiol.* **11**, 455–461 (2001).

11. 11.

Bai, L. et al. Genetic identification of an expansive mechanoreceptor sensitive to skin stroking. *Cell* **163**, 1783–1795 (2015).

12. 12.

Moehring, F. et al. Keratinocytes mediate innocuous and noxious touch via ATP-P2X4 signaling. *eLife* **7**, e31684 (2018).

13. 13.

Walcher, J. et al. Specialized mechanoreceptor systems in rodent glabrous skin. *J. Physiol.* **596**, 4995–5016 (2018).

14. 14.

Neubarth, N. L. et al. Meissner corpuscles and their spatially intermingled afferents underlie gentle touch perception. *Science* **368**, eabb2751 (2020).

15. 15.

Lynn, B. & Shakhanbeh, J. Properties of A δ high threshold mechanoreceptors in the rat hairy and glabrous skin and their response to heat. *Neurosci. Lett.* **85**, 71–76 (1988).

16. 16.

Chapin, J. K. Laminar differences in sizes, shapes, and response profiles of cutaneous receptive fields in the rat SI cortex. *Exp. Brain Res.* **62**, 549–559 (1986).

17. 17.

Enander, J. M. D. & Jörntell, H. Somatosensory cortical neurons decode tactile input patterns and location from both dominant and non-dominant digits. *Cell Rep.* **26**, 3551–3560.e3554 (2019).

18. 18.

Maricich, S. M. et al. Merkel cells are essential for light-touch responses. *Science* **324**, 1580–1582 (2009).

19. 19.

Hooks, B. M., Lin, J. Y., Guo, C. & Svoboda, K. Dual-channel circuit mapping reveals sensorimotor convergence in the primary motor cortex. *J. Neurosci.* **35**, 4418–4426 (2015).

20. 20.

Luo, W., Enomoto, H., Rice, F. L., Milbrandt, J. & Ginty, D. D. Molecular identification of rapidly adapting mechanoreceptors and their developmental dependence on ret signaling. *Neuron* **64**, 841–856 (2009).

21. 21.

Kuehn, E. D., Meltzer, S., Abraira, V. E., Ho, C. Y. & Ginty, D. D. Tiling and somatotopic alignment of mammalian low-threshold mechanoreceptors. *Proc. Natl Acad. Sci. USA* **116**, 9168–9177 (2019).

22. 22.

Gabernet, L., Jadhav, S. P., Feldman, D. E., Carandini, M. & Scanziani, M. Somatosensory integration controlled by dynamic thalamocortical feed-forward inhibition. *Neuron* **48**, 315–327 (2005).

23. 23.

Bruno, R. M. & Simons, D. J. Feedforward mechanisms of excitatory and inhibitory cortical receptive fields. *J. Neurosci.* **22**, 10966–10975 (2002).

24. 24.

Abraira, V. E. et al. The cellular and synaptic architecture of the mechanosensory dorsal horn. *Cell* **168**, 295–310.e219 (2017).

25. 25.

Li, L. et al. The functional organization of cutaneous low-threshold mechanosensory neurons. *Cell* **147**, 1615–1627 (2011).

26. 26.

Prsa, M., Morandell, K., Cuenu, G. & Huber, D. Feature-selective encoding of substrate vibrations in the forelimb somatosensory cortex. *Nature* **567**, 384–388 (2019).

27. 27.

Johnson, K. O. & Hsiao, S. S. Neural mechanisms of tactal form and texture perception. *Annu. Rev. Neurosci.* **15**, 227–250 (1992).

28. 28.

Saal, H. P. & Bensmaia, S. J. Touch is a team effort: interplay of submodalities in cutaneous sensibility. *Trends Neurosci.* **37**, 689–697 (2014).

29. 29.

Sathian, K. Tactile sensing of surface features. *Trends Neurosci.* **12**, 513–519 (1989).

30. 30.

Johansson, R. S. & Flanagan, J. R. Coding and use of tactile signals from the fingertips in object manipulation tasks. *Nat. Rev. Neurosci.* **10**, 345–359 (2009).

31. 31.

Choi, S. et al. Parallel ascending spinal pathways for affective touch and pain. *Nature* **587**, 258–263 (2020).

32. 32.

Madisen, L. et al. A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat. Neurosci.* **13**, 133–140 (2010).

33. 33.

da Silva, S. et al. Proper formation of whisker barrelettes requires periphery-derived Smad4-dependent TGF β signaling. *Proc. Natl Acad. Sci. USA* **108**, 3395–3400 (2011).

34. 34.

Liu, Y. et al. Sexually dimorphic BDNF signaling directs sensory innervation of the mammary gland. *Science* **338**, 1357–1360 (2012).

35. 35.

Shroyer, N. F. et al. Intestine-specific ablation of *Mouse atonal homolog 1* (*Math1*) reveals a role in cellular homeostasis. *Gastroenterology* **132**, 2478–2488 (2007).

36. 36.

Ramirez, A. et al. A keratin K5Cre transgenic line appropriate for tissue-specific or generalized Cre-mediated recombination. *Genesis* **39**, 52–57 (2004).

37. 37.

Rutlin, M. et al. The cellular and molecular basis of direction selectivity of A δ -LTMRs. *Cell* **159**, 1640–1651 (2014).

38. 38.

Wang, V. Y., Rose, M. F. & Zoghbi, H. Y. Math1 expression redefines the rhombic lip derivatives and reveals novel lineages within the brainstem and cerebellum. *Neuron* **48**, 31–43 (2005).

39. 39.

Jun, J. J. et al. Real-time spike sorting platform for high-density extracellular probes with ground-truth validation and drift correction. Preprint at <https://doi.org/10.1101/101030> (2017).

40. 40.

Rodriguez, A. & Laio, A. Clustering by fast search and find of density peaks. *Science* **344**, 1492–1496 (2014).

41. 41.

Niell, C. M. & Stryker, M. P. Highly selective receptive fields in mouse visual cortex. *J. Neurosci.* **28**, 7520–7536 (2008).

42. 42.

Bartho, P. et al. Characterization of neocortical principal cells and interneurons by network interactions and extracellular features. *J. Neurophysiol.* **92**, 600–608 (2004).

43. 43.

Shannon, C. E. A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423, 623–656 (1948).

44. 44.

Magri, C., Whittingstall, K., Singh, V., Logothetis, N. K. & Panzeri, S. A toolbox for the fast information analysis of multiple-site LFP, EEG and spike train recordings. *BMC Neurosci.* **10**, 81 (2009).

45. 45.

Panzeri, S., Senatore, R., Montemurro, M. A. & Petersen, R. S. Correcting for the sampling bias problem in spike train information measures. *J. Neurophysiol.* **98**, 1064–1072 (2007).

46. 46.

Panzeri, S. & Treves, A. Analytical estimates of limited sampling biases in different information measures. *Network* **7**, 87–107 (1996).

47. 47.

Lin, J. Y., Knutsen, P. M., Muller, A., Kleinfeld, D. & Tsien, R. Y. ReaChR: a red-shifted variant of channelrhodopsin enables deep transcranial optogenetic excitation. *Nat. Neurosci.* **16**, 1499–1508 (2013).

48. 48.

Keller, D., Ero, C. & Markram, H. Cell densities in the mouse brain: a systematic review. *Front. Neuroanat.* **12**, 83 (2018).

49. 49.

Cases, O. et al. Lack of barrels in the somatosensory cortex of monoamine oxidase A-deficient mice: role of a serotonin excess during the critical period. *Neuron* **16**, 297–307 (1996).

Acknowledgements

We thank J. Hua, T. Monteiro, and S. Shea for assistance with mouse husbandry and histology, and J. Assad, S. Choi, M. Do, G. Fishell, S. Meltzer, L. Orefice, G. Rankin, C. Santiago and M. Springel for comments on the manuscript. This work was supported by NIH grants F32NS105324 (A.J.E.), K99NS119739 (A.J.E.), R01NS089521 (C.D.H.), DP1MH125776 (C.D.H.), R01NS108410 (C.D.H. and S.P.), NS097344 (D.D.G.), the Fondation Bertarelli (S.P. and D.D.G.) and a Harvard Medical School Dean's Initiative Award (C.D.H. and D.D.G.). D.D.G. is an investigator of the Howard Hughes Medical Institute.

Author information

Affiliations

1. Department of Neurobiology, Harvard Medical School, Boston, MA, USA

Alan J. Emanuel, Brendan P. Lehnert, Christopher D. Harvey & David D. Ginty

2. Howard Hughes Medical Institute, Harvard Medical School, Boston, MA, USA

Alan J. Emanuel, Brendan P. Lehnert & David D. Ginty

3. Department of Excellence for Neural Information Processing, Center for Molecular Neurobiology (ZMNH), University Medical Center Hamburg-Eppendorf (UKE), Hamburg, Germany

Stefano Panzeri

4. Neural Computation Laboratory, Istituto Italiano di Tecnologia, Genova, Italy

Stefano Panzeri

Contributions

A.J.E., C.D.H. and D.D.G. conceptualized experiments. A.J.E. performed experiments. A.J.E. analysed data with assistance from S.P. B.P.L. (Waabishkigidaagamigizi in ojibwemowin) developed methods for activation of skin sensory neurons. A.J.E., C.D.H. and D.D.G. wrote the manuscript and all authors contributed to review and editing.

Corresponding authors

Correspondence to [Christopher D. Harvey](#) or [David D. Ginty](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Cheryl Stucky and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

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 A \$\beta\$ LTMR Responses to Force-Controlled Step Indentations.](#)

a, Raster plot showing cutaneous A β RA-LTMR and A β SA-LTMR responses to a series of step indentations ranging from 1 to 75 mN applied to the most responsive skin region for each neuron. Markers are colored

according to how the neurons were labeled (Blue, unlabeled; Orange, Ret⁺; Green, TrkB⁺; Red, TrkC⁺). A subset of these recordings (unlabeled neurons that were recorded in littermate controls for *TrkB*^{cKO} mice, Ret⁺ neurons, and TrkB⁺ neurons) were previously published¹⁴. **b**, Example RFs of an Aβ RA-LTMR (top) and an Aβ SA-LTMR (bottom) to 10-mN step indentations superimposed on a schematic of the hindpaw. Dashed lines outline pedal pads. Unfilled markers represent stimulus locations that did not evoke a response. Color represents the total number of action potentials evoked during the step indentation. **c**, RF sizes for Aβ RA-LTMRs ($n = 17$) and Aβ SA-LTMRs ($n = 14$) that were responsive to 10-mN step indentations. Markers are colored according to how the neurons were labeled. Mean \pm s.e.m. areas of 2.3 ± 0.5 and 0.9 ± 0.2 mm² for Aβ RA-LTMRs and Aβ SA-LTMRs, respectively and median \pm i.q.r. of 1.3 ± 3.4 and 0.8 ± 0.9 for Aβ RA-LTMRs and Aβ SA-LTMRs, respectively (Two-sided Mann-Whitney $U = 81.0$, $p = 0.07$). **d**, Force threshold for step indentation response for Aβ RA-LTMRs ($n = 25$) and Aβ SA-LTMRs ($n = 20$). Markers are colored according to how the neurons were labeled. Mean \pm s.e.m. thresholds of 9.0 ± 2.1 and 9.2 ± 2.1 mN for Aβ RA-LTMRs and Aβ SA-LTMRs, respectively and median \pm i.q.r. of 5.0 ± 8.0 and 5.0 ± 11.5 for Aβ RA-LTMRs and Aβ SA-LTMRs, respectively (Two-sided Mann-Whitney $U = 253.5$, $p = 0.45$). **e**, Individual (gray) and mean (black) waveforms recorded from a Pacinian corpuscle-innervating Aβ LTMR labeled with a *Ret*^{CreER}; *PVFlpO* intersectional strategy (3 mg tamoxifen administered at embryonic day 11.5). These neurons were not labeled with dye-conjugated CTB, which was injected into the pedal and digit pads 48 h prior to recording. **f**, A 100-Hz sine ramp stimulus was applied to multiple locations across the glabrous hindpaw to assess the responsive region for the Pacinian corpuscle-innervating Aβ LTMR. Top left: Response of neuron to most sensitive region. Top right: sine stimulus response threshold for each probed location. This Aβ LTMR likely innervated a Pacinian corpuscle in the ankle joint. Bottom left: Response to step indentations at most sensitive location. Bottom right: Response to step indentations at all locations overlaid. In some locations action potentials are generated as the probe initially comes into contact with the skin but are never generated in response to the step indentations (which were low-pass filtered at 33 Hz). **g**, Frequency-response relationships for sine stimuli delivered in a ramp (top) or in a 1-s step (bottom). All Pacinian corpuscle-innervating Aβ LTMRs

were most sensitive to high frequency stimulation. Ankle and digit terminal locations were inferred based on the regions of the paw that responded to a handheld vibrating metal probe. Ankle neurons ($n = 3$) responded when the probe was applied to most regions of the paw, including digits and pedal pads. Digit neurons ($n = 3$) only responded when the probe was applied to a single digit.

Extended Data Fig. 2 Depth Calibration and Validation for S1 Recordings.

a, Current source density (CSD) plots of an exemplar hindpaw wild-type S1 recording. Sources (red) and sinks (blue) are apparent soon after the onset of the step indentation. The depth of an early, prolonged sink (marked by an asterisk) was used to rigidly adjust the depth of the probe so that this sink was at the center of layer IV. **b**, Optotagging protocol (top) and corresponding action potential timing of an example optotagged unit (bottom) from an *Scnn1a-tg3-Cre;R26^{LSL-ChR2}* mouse. NBQX (5 mM) was applied to the surface of the brain to block excitatory synaptic transmission starting on trial 16. **c**, Probability distributions of the latency to the first spike after LED pulses for two optotagged units. Shaded region represents 95% confidence interval of shuffled distribution. **d**, Probability distributions of the latency to the first spike after LED pulses for two non-optotagged units. Shaded region represents 95% confidence interval of shuffled distribution. **e**, Mechanical responses to 75-mN step indentations for each unit in a recording from a *Scnn1a-tg3-Cre;R26^{LSL-ChR2}* mouse before (left) and after (right) application of NBQX. Optotagged units (blue) had similar mechanical response profiles to non-optotagged units (gray). **f**, Depth distribution (after CSD calibration) of all units (gray, $n = 866$) recorded from wild-type hindpaw S1 compared with the depth of all optotagged units (blue, $n = 24$ from 5 recordings from 3 mice). The majority of optotagged units were within layer IV (416.5 – 535.5 μm deep). **g**, Typical location of an electrode array in hindpaw S1 superimposed upon post hoc histology of a mouse expressing ChR2-EYFP in layer IV neurons (*Scnn1a-tg3-Cre;R26^{LSL-ChR2-EYFP}*). The probe was coated in DiI prior to recording. Scale bar, 500 μm .

Extended Data Fig. 3 Movement Subtraction.

a, Frame of video (taken at 10 Hz) of paw during stimulation. Dashed box outlines region of interest (ROI) used for movement analysis. The ROI was binarized using Otsu thresholding and the difference from frame to frame was calculated. **b**, The first derivative of the frame to frame difference for an example recording. When this derivative exceeded three standard deviations from 0.25 s before the step to 0.25 s after the step, the entire step was excluded from subsequent analyses. **c**, Firing rate histograms in response to 75-mN step indentations (from 0 to 0.5 s) without (left) and with (middle) subtraction from the recording in b. The difference is shown on the right.

Extended Data Fig. 4 Receptive Fields and Spatial Information of Units in Hindpaw S1 and Forepaw S1.

a, Left: stimulus locations used to probe receptive fields with 10-mN step indentations for hindpaw S1. The grid size is 5×5 mm. Right: Mean (\pm s.e.m.) spatial information for all hindpaw S1 units during the course of the 10-mN step indentation. **b**, Distribution of spatial information across all hindpaw S1 units at the onset (ON; 20-70 ms after step onset), sustained (SUS; 250-500 ms after step onset), and offset (OFF; 20-70 ms after step offset) portions of the 10-mN step indentation. **c**, A correlation (Pearson $r = 0.79$, $p = 1.4 \times 10^{-179}$) was apparent between the amount of spatial information at the onset and offset of the step indentation, but the amount of spatial information at the onset was reliably greater than that at the offset. Colors represent units from different recordings. Gray dashed line is the unity line. **d**, RF sizes of FS hindpaw S1 units ($n = 201$ units) were larger than those of RS hindpaw S1 units ($n = 649$ units). Two-sided Mann-Whitney $U = 56,703$, $p = 1.6 \times 10^{-10}$. Box plot element definitions: center line, median; box limits, upper and lower quartiles; whiskers, $1.5 \times$ interquartile range; points, outliers. **e**, Example RFs for units with varying degrees of spatial information at the onset. The mean spatial information over the 20-70 ms after the onset of the step indentation is displayed above the heatmap for each unit. The magnitude of the response (rather than the size of the RF) appeared to account for most of the differences in spatial

information. **f**, Top: Mean (\pm s.e.m.) spatial information for all units in a recording from an *Scnn1a-tg3-Cre; R26^{LSL-ChR2}* mouse. The spatial information for two optotagged layer IV units (cyan and magenta) are overlaid. Bottom: Example RFs for these two optotagged units are qualitatively similar to those displayed in e. **g**, As in a, for forepaw S1. Stimuli were applied to 16 locations in a 4×4 mm grid. Thus, the spatial information calculated from forepaw S1 recordings is not directly comparable to that calculated from hindpaw S1 recordings. **h**, As in b, for forepaw S1. **i**, As in c, for forepaw S1. (Pearson $r = 0.71$, $p = 1.2 \times 10^{-99}$). **j**, As in d, for forepaw S1 ($n = 435$ RS and FS units, respectively; Two-sided Mann-Whitney $U = 31,348$). **k**, As in e, for forepaw S1. **l**, Receptive field sizes for hindpaw ($n = 850$ units from 12 recordings in 8 mice) and forepaw S1 ($n = 677$ units from 9 recordings in 4 mice) units. Two-sided Mann-Whitney $U = 266932$. Box plot element definitions as in d.

Extended Data Fig. 5 Hindpaw S1 FS Responses and Sensitivity Measurements using Fits to Saturating Exponential.

a, The hindpaw was tethered over a 7.6-mm diameter circular aperture through which step indentations were applied to the glabrous skin of the forepaw. **b**, Distribution of trough-to-peak times of action potential waveforms for hindpaw S1 units. The dashed red line demarcates the threshold (0.55 ms) used for classifying RS from FS units. **c**, Baseline firing rate for hindpaw S1 RS ($n = 658$ units from 12 recordings in 7 mice) and FS units ($n = 181$ units) from layers II/III, IV and V. Box plot element definitions: center line, median; box limits, upper and lower quartiles; whiskers, $1.5 \times$ interquartile range; points, outliers. **d**, Top: Heatmaps of the Z-scored firing rate for 181 FS units. Bottom: Grand mean firing rate (\pm s.e.m) for hindpaw S1 FS units from each layer. Shaded region indicates timing of step indentation. **e**, Left: Peak-normalized, baseline-subtracted firing rate at the onset of the step indentation for each force for an example hindpaw S1 RS unit. The step indentation begins at 0 s. Right: Fit of the intensity-response relation for this unit to a saturating exponential ($R=1\backslash mbox{--}\{e\}^{\{-I/\{I\}_{\{0\}}\}}$). **f**, As in e for another example hindpaw S1 RS unit. **g**, Sensitivity as measured by the mean (\pm 95% confidence interval) I_0 fit parameter for A β RA-LTMRs and A β SA-LTMRs (left) and

well-fit ($\sum |\text{residuals}| < 1.2$) RS and FS units in each layer of hindpaw (middle) and forepaw (right) S1. DRG n= 24 and 19 neurons for A β RA-LTMRs and A β SA-LTMRs, respectively. Hindpaw S1 n = 75 and 74 RS and FS units for layer II/III, respectively; n = 74 and 20 RS and FS units for layer IV, respectively; n = 224 and 62 RS and FS units for layer V, respectively. Forepaw S1 n = 24 and 35 RS and FS units for layer II/III, respectively; n = 27 and 13 RS and FS units for layer IV, respectively; n = 149 and 78 for layer V, respectively. No significant differences apparent within areas (DRG: Two-sided Mann-Whitney U = 183, p = 0.14; hindpaw S1: Kruskal-Wallis H = 7.04, p = 0.22; forepaw S1: Kruskal-Wallis H = 10.67, p = 0.06) but I_0 differs between all DRG neurons and hindpaw S1 units (Two-sided Mann-Whitney U = 7594, p = 0.0004) and between all hindpaw S1 units and all forepaw S1 units (Two-sided Mann-Whitney U = 52,285, p = 9.6×10^{-18}).

Extended Data Fig. 6 Forepaw S1 Responses to Step Indentation.

a, The forepaw was tethered over a 6.4-mm diameter circular aperture through which step indentations were applied to the glabrous skin of the forepaw. **b**, Distribution of trough-to-peak times of action potential waveforms for forepaw S1 units. The dashed red line demarcates the threshold (0.55 ms) used for classifying RS from FS units. **c**, Baseline firing rate for forepaw RS (n = 576 units from 12 recordings in 6 mice) and FS units (n = 258 units) from layers II/III, IV and V. Box plot element definitions: center line, median; box limits, upper and lower quartiles; whiskers, 1.5 \times interquartile range; points, outliers. **d**, Grand mean (\pm s.e.m.) baseline-subtracted firing rate for all forepaw S1 units (cyan; n = 834) and all hindpaw S1 units (gray; n = 866) in response to step indentations. The shaded regions represent the timing of step indentations and the numbers at the stop signify the intensity of the indentation (in mN). While a sustained response is generated within forepaw S1 to high forces, this response is dwarfed by the transients at the onset and offset of the step indentations. **e**, Top: Heatmaps of the Z-scored firing rate for forepaw S1 RS (left) and FS (right) units. Bottom: Grand mean firing rate (\pm s.e.m) for RS (left) and FS (right) units from each layer. Shaded region indicates timing of step indentation. **f**, Cumulative distributions of baseline-subtracted firing rate for

all forepaw S1 units at each step intensity for the onset (ON; 10-50 ms after step onset), offset (OFF; 10-50 ms after step offset), sustained (SUS; 250-500 ms after step onset) periods. **g**, Sensitivity as measured by the I_0 parameter for RS and FS units in each layer of forepaw S1 fit well ($\sum|\text{residuals}| < 1.2$) with a saturating exponential. Number of units indicated on each bar. **h**, Density histograms of the β coefficients for the A β RA-LTMR and A β SA-LTMR profiles that best fit hindpaw S1 units at forces designated above each plot. Heatmap colors represent number of units per bin. Only units with significant R^2 values, as determined by permutation of the LTMR response profiles, were included. **i**, As in h, for forepaw S1 units.

Extended Data Fig. 7 DKO Histology and S1 Receptive Fields and Response Durations in Knockout Mice.

a, Example pedal pad glabrous skin section in a littermate control (AET10) immunostained for NFH (magenta) to identify axons, Tromal1 (yellow) to identify Merkel cells (indicated by arrowheads), and S100 (cyan) to identify Meissner corpuscles (indicated by arrows). Scale bar: 50 μm . Similar pattern observed in four littermate controls. **b**, Pedal pad glabrous skin section in a DKO (AEV5) immunostained for NFH (magenta), Tromal1 (yellow), and S100 (cyan). No Merkel cells or Meissner corpuscles were apparent in this section. Scale bar: 50 μm . Similar pattern observed in four DKOs. **c**, Quantification of the density of Merkel cells within pedal pads for littermate controls (gray markers) and DKOs (purple markers). **d**, Quantification of the density of Meissner corpuscles within pedal pads for littermate controls (gray markers) and DKOs (purple markers). **e**, Durations of responses (violin plot shows kernel destiny estimate of underlying distribution) at the onset of step indentations (calculated by multiplying the number of consecutive bins with a Z score > 2 by the bin size [20 ms]) for hindpaw S1 units sensitive to each force in control $TrkB^{fl/fl}$ and $TrkB^{cKO}$ mice (top), in control $Atoh1^{fl/fl}$ and $Atoh1^{cKO}$ mice (middle), and in littermate controls and DKO mice (bottom). Plots shown only for forces to which at least 20 units responded at the indentation onset. * $p < 0.05$, Two-sided Mann-Whitney U test. **f**, As in e, for responses at the offset of step indentations. **g**, Schematic of RF measurements. 10-mN indentations were delivered to 36 locations in a 5×5 mm grid for wild-type animals, $TrkB^{fl/fl}$

controls, *TrkB*^{cKO} mice, *Atoh1*^{f/f} controls, and *Atoh1*^{cKO} mice. 40-mN indentations were delivered at each location for DKOs and their littermate controls. **h**, Mean (\pm s.e.m.) spatial information (left; blue: *TrkB*^{cKO}, gray: *TrkB*^{f/f}) and mean (\pm s.e.m.) RF areas (right) for S1 units from wild-type (n = 649 units), *TrkB*^{f/f} (230 units), and *TrkB*^{cKO} (182 units) animals. *p < 0.05, ***p < 0.0001, Two-sided Mann-Whitney U test, Bonferroni corrections for multiple comparisons applied. **i**, As in h, for RF measurements in *Atoh1*^{f/f} (gray; n = 48 units) and *Atoh1*^{cKO} (red; 128 units) animals. **j**, As in h, for RF measurements (made at 40 mN) in DKOs (purple; n = 369 units) and their littermate controls (gray; 217 units). Two-sided Mann-Whitney U = 5,412. Box plot element definitions (h–j): center line, median; box limits, upper and lower quartiles; whiskers, 1.5 \times interquartile range; points, outliers.

Extended Data Fig. 8 Optical Responses in A β LTMRs and Controls for Optogenetic Gain-of-Function Experiments.

a, A β LTMR subtypes selectively labeled in a *TrkB*^{CreER}; *R26*^{LSL-ReaChR-mCitrine} (*TrkB*::*ReaChR*) mouse (left), a *Ret*^{CreER}; *Advillin*^{FlpO}; *R26*^{LSL-FSF-ReaChR-mCitrine} (*Ret*::*ReaChR*) mouse (middle), and a *TrkC*^{CreER}; *R26*^{LSL-ReaChR-mCitrine} (*TrkC*::*ReaChR*) mouse (right). Arrows indicate mCitrine $^{+}$ fibers within S100 $^{+}$ Meissner corpuscles (left and middle) or abutting Tromal1 $^{+}$ Merkel cells (right). Similar patterns observed in all 7 *TrkB*::*ReaChR*, 8 *Ret*::*ReaChR*, and 12 *TrkC*::*ReaChR* mice. Scale bars: 40 μ m. **b**, Top: 33 μ J light pulses were directed to the skin at each location indicated by a marker. A ReaChR-expressing TrkB $^{+}$ A β RA-LTMR responded with, in most cases, a single action potential when the pulses were directed onto the mechanical RF of the neuron. Bottom: Histogram showing the distribution of latencies to the first spike for all locations in which an action potential was evoked by optical stimulation. Scale bar: 1 mm. **c**, As in a, for a ReaChR-expressing Ret $^{+}$ A β LTMR. Scale bar: 1 mm. **d**, As in a, for a ReaChR-expressing TrkC $^{+}$ A β SA-LTMR. Scale bar: 1 mm. **e**, As in a, for a ReaChR-expressing TrkC $^{+}$ proprioceptor. This proprioceptor responded to movement of a digit. Light did not evoke action potentials, even during ongoing activity. Similar results obtained in 4

additional proprioceptors from 3 mice. Scale bar: 1 mm. **f**, Hindpaw S1 recordings from mice ($n = 3$) in which proprioceptors expressed ReaChR, driven intersectionally using the Cux^{CreER} and PV^{FlpO} driver lines. No responses to optical stimulation were observed in S1 despite responsivity to mechanical stimulation. Dashed lines demarcate cortical layers. **g**, Native mCitrine fluorescence in Clark's column and the dorsal column of the cervical spinal cord of a $Cux2^{CreER}; PV^{FlpO}; R26^{LSL-FSF-ReaChR-mCitrine}$ animal. Similar pattern observed in 2 additional mice. Scale bar: 500 μm . **h**, No optical responses were observed in hindpaw S1 of an $R26^{LSL-FSF-ReaChR}$ animal lacking Cre recombinase. Dashed lines demarcate cortical layers.

Extended Data Fig. 9 Response Profiles of S1 Units Sensitive to Selective Optogenetic Stimulation and Receptive Fields and Intensity-Response Relationships in VPL.

a, Grand mean (\pm s.e.m.) firing rate responses to step indentations for forepaw S1 (left) and hindpaw S1 (right) wild-type (cyan) units compared to that of units in each driver line that were responsive to optical stimulation. The response profiles are similar across all driver lines and wild-type units. Shaded regions indicate the timing of step indentations. **b**, Top: Schematic of probe position within VPL. Bottom: *Post-hoc* histology showing the location of the electrode tract (DiI, red) in relation to thalamic structures. Hoechst 33258 nuclear stain shown in blue. Scale bar: 500 μm . Similar histology observed in 3 mice. VPL: ventroposterolateral nucleus of the thalamus, VPM: ventroposteromedial nucleus of the thalamus, dLGN: dorsal lateral geniculate nucleus, vLGN: ventral lateral geniculate nucleus. **c**, Mean (\pm s.e.m.) spatial information of mechanically sensitive VPL units ($n=56$) in relation to 500-ms, 10-mN step indentations applied to 25 locations in a 4×4 mm grid. **d**, Left: Cumulative distribution of mean spatial information of VPL units at the onset (ON; 20-70 ms after step onset), sustained (SUS; 250-500 ms after step onset), and offset (OFF; 20-70 ms after step offset) portions of the 10-mN step indentation. Right: The amount of information at the onset and offset is correlated (Pearson $r = 0.70$, $p = 1.3 \times 10^{-9}$). **e**, Spatial information in VPL units ($n = 56$) and forepaw S1 units ($n = 306$) differs at the onset (left) but is indistinguishable at the offset (right) of the step indentation. Two-sided Mann-Whitney $U =$

6,520 and 8,071 for onset and offset comparisons, respectively. **f**, RFs of example VPL units with varying amounts of mean spatial information at the onset of the step indentation, noted above each heat map. **g**, RF area for S1 and VPL units ($n = 850$, 644, and 56 units for hindpaw S1, forepaw S1, and VPL, respectively). Kruskal-Wallis $H = 3.12$. Box plot element definitions: center line, median; box limits, upper and lower quartiles; whiskers, $1.5 \times$ interquartile range; points, outliers. **h**, Distribution of action potential waveform trough to peak times for mechanically sensitive VPL units. **i**, Sensitivity of VPL units compared to forepaw S1 units at the onset of the step indentation as assessed by I_0 fits to saturating exponentials (left; $n = 80$ and 342 units for VPL and S1, respectively; $U = 9,026$) and maximum response firing rate (right; $n = 174$ and 599 units for VPL and S1, respectively; $U = 51,650$). Mann-Whitney U test. **j**, Sensitivity of VPL units at the offset of the step indentation as assessed by I_0 values (left; $n = 64$ and 316 units for VPL and S1, respectively; $U = 9,502$) and maximum response firing rate (right; $n = 174$ and 599 units for VPL and S1, respectively; $U = 40,986$). Mann-Whitney U test.

Extended Data Fig. 10 Optogenetic Activation of A β RA-LTMRs and A β SA-LTMRs Modulate Firing Rates in VPL Units with Heterogeneous Response Profiles to Mechanical Indentations.

a, Mean (\pm s.e.m.) firing rate of two example VPL units that respond to 75-mN step indentations by transiently increasing their firing rate at the onset and offset of the step indentation, much like typical S1 units. The unit on the left can be driven by selective optical activation of A β RA-LTMRs and the unit on the right can be driven by selective optical activation of A β SA-LTMRs. **b**, Mean (\pm s.e.m.) of two VPL units with prominent sustained responses to 75-mN step indentations. The unit on the left can be driven by optical activation of A β RA-LTMRs and the unit on the right can be driven by optical activation of A β SA-LTMRs. **c**, Mean (\pm s.e.m.) firing rate of two VPL units that respond to the 75-mN step indentation with decreases in their firing rates. The firing rate of both units can be modulated by optical activation of A β SA-LTMRs but the unit on the left increases its firing rate while the unit on the right decreases its firing rate. **d**, Time of peak |firing

rate] relative to the laser pulse for optically sensitive units in VPL and each layer of forepaw S1 *TrkB::ReaChR* (left; $U = 9,038$, $p = 2.6 \times 10^{-6}$ for all VPL units compared to all forepaw S1 units, Mann-Whitney U test; $n = 119, 34, 21, 152$, and 10 units in VPL, layers II/III, IV, V, and VI, respectively) and *TrkC::ReaChR* (right; $U = 4,828$, $p = 2.8 \times 10^{-9}$ for all VPL units compared to all forepaw S1 units, Mann-Whitney U test; $n = 84, 34, 39, 120$, and 11 units in VPL, layers II/III, IV, V, and VI, respectively) mice. Box plot element definitions: center line, median; box limits, upper and lower quartiles; whiskers, $1.5 \times$ interquartile range; points, outliers. **e**, Each panel shows mechanical (75 mN step indentation from 0 to 0.5 s) and optical (0.3 ms pulse at 0 ms) responses of one of four clusters determined by K-means clustering of the first three principal components of the Z-scored response to 75-mN step indentations. Individual (thin lines) and mean (thick lines) Z-scored responses from units from *TrkC::ReaChR* (red) and *TrkB::ReaChR* (black) mice. Both the mechanical and optical responses generated by stimulation of mice from either genotype are similar within clusters. **f**, The majority of units from each cluster respond to selective optical activation of A β SA-LTMRs (*TrkC::ReaChR*, red) or A β RA-LTMRs (*TrkB::ReaChR*, black). Total number of units indicated on each bar.

Supplementary information

Reporting Sumary

Source data

Source Data Fig. 1

Source Data Fig. 2

Source Data Fig. 3

Source Data Fig. 4

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Emanuel, A.J., Lehnert, B.P., Panzeri, S. *et al.* Cortical responses to touch reflect subcortical integration of LTMR signals. *Nature* **600**, 680–685 (2021). <https://doi.org/10.1038/s41586-021-04094-x>

- Received: 23 October 2020
- Accepted: 04 October 2021
- Published: 17 November 2021
- Issue Date: 23 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04094-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

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

- Article
- [Published: 24 November 2021](#)

Mechanical actions of dendritic-spine enlargement on presynaptic exocytosis

- [Hasan Ucar](#) ORCID: [orcid.org/0000-0001-9037-533X^{1,2}](https://orcid.org/0000-0001-9037-533X),
- [Satoshi Watanabe^{1,3}](#),
- [Jun Noguchi^{1,3}](#),
- [Yuichi Morimoto](#) ORCID: [orcid.org/0000-0002-1750-5297^{1,2}](https://orcid.org/0000-0002-1750-5297),
- [Yusuke Iino^{1,2}](#),
- [Sho Yagishita^{1,2}](#),
- [Noriko Takahashi](#) ORCID: [orcid.org/0000-0003-1025-1690^{1,4}](https://orcid.org/0000-0003-1025-1690) &
- [Haruo Kasai](#) ORCID: [orcid.org/0000-0003-2327-9027^{1,2}](https://orcid.org/0000-0003-2327-9027)

[Nature](#) volume 600, pages 686–689 (2021)

- 11k Accesses
- 1 Citations
- 269 Altmetric
- [Metrics details](#)

Subjects

- [Exocytosis](#)
- [Long-term potentiation](#)
- [Nanoscale biophysics](#)

- [Spine structure](#)
- [Synaptic vesicle exocytosis](#)

Abstract

Synaptic transmission involves cell-to-cell communication at the synaptic junction between two neurons, and chemical and electrical forms of this process have been extensively studied. In the brain, excitatory glutamatergic synapses are often made on dendritic spines that enlarge during learning^{1,2,3,4,5}. As dendritic spines and the presynaptic terminals are tightly connected with the synaptic cleft⁶, the enlargement may have mechanical effects on presynaptic functions⁷. Here we show that fine and transient pushing of the presynaptic boutons with a glass pipette markedly promotes both the evoked release of glutamate and the assembly of SNARE (soluble *N*-ethylmaleimide-sensitive factor attachment protein receptor) proteins^{8,9,10,11,12}—as measured by Förster resonance transfer (FRET) and fluorescence lifetime imaging—in rat slice culture preparations¹³. Both of these effects persisted for more than 20 minutes. The increased presynaptic FRET was independent of cytosolic calcium (Ca^{2+}), but dependent on the assembly of SNARE proteins and actin polymerization in the boutons. Notably, a low hypertonic solution of sucrose (20 mM) had facilitatory effects on both the FRET and the evoked release without inducing spontaneous release, in striking contrast with a high hypertonic sucrose solution (300 mM), which induced exocytosis by itself¹⁴. Finally, spine enlargement induced by two-photon glutamate uncaging enhanced the evoked release and the FRET only when the spines pushed the boutons by their elongation. Thus, we have identified a mechanosensory and transduction mechanism¹⁵ in the presynaptic boutons, in which the evoked release of glutamate is enhanced for more than 20 min.

This is a preview of subscription content

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Fig. 1: Enhancement of glutamate release by fine pushing of a presynaptic terminal.

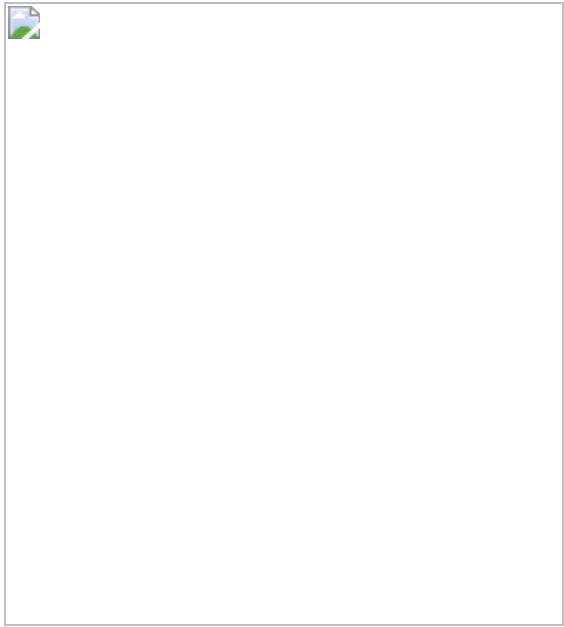


Fig. 2: Effects of fine pushing on SNARE assembly.

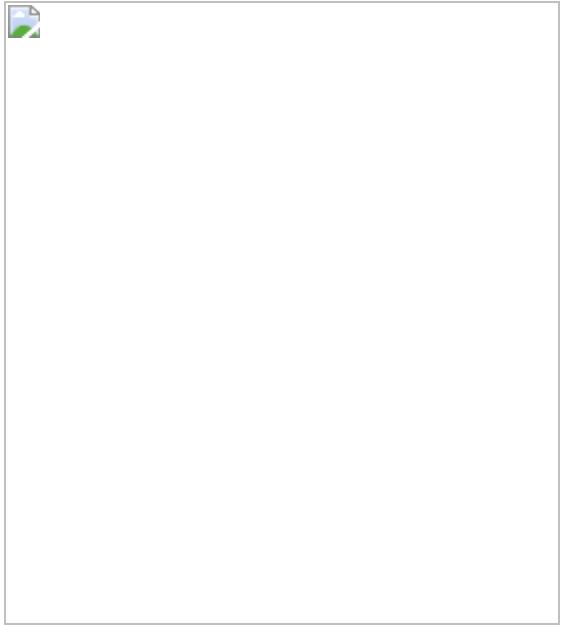


Fig. 3: Enhanced glutamate release and SNARE assembly induced by a moderate increase in the osmotic pressure with 20 mM sucrose.

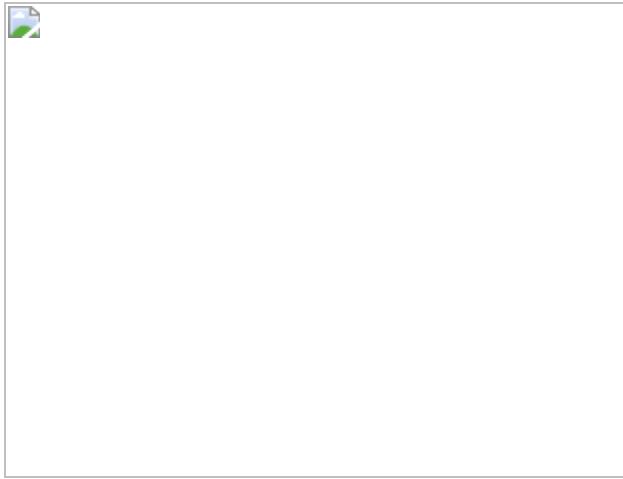
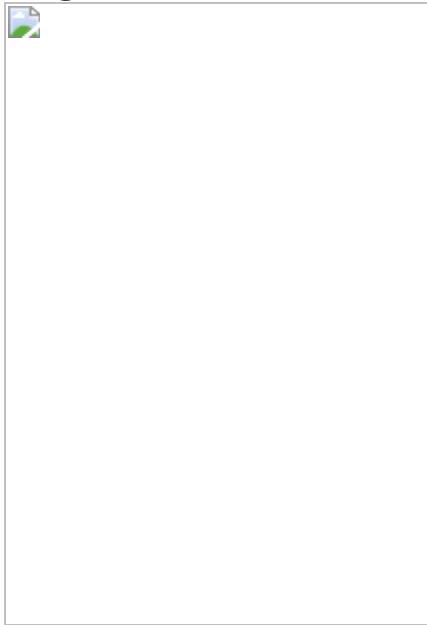


Fig. 4: Increases in the presynaptic SNARE assembly caused by spine enlargement.



Data availability

All data are available from the corresponding author upon reasonable request. [Source data](#) are provided with this paper.

Code availability

The MATLAB scripts for the FRET/FLIM analysis together with other codes used in this study are available at <https://github.com/sw->

[swatanabe/ucar-et-al-2021](#).

References

1. 1.

Matsuzaki, M., Honkura, N., Ellis-Davies, G. C. & Kasai, H. Structural basis of long-term potentiation in single dendritic spines. *Nature* **429**, 761–766 (2004).

2. 2.

Harvey, C. D., Yasuda, R., Zhong, H. & Svoboda, K. The spread of Ras activity triggered by activation of a single dendritic spine. *Science* **321**, 136–140 (2008).

3. 3.

Herring, B. E. & Nicoll, R. A. Long-term potentiation: from CaMKII to AMPA receptor trafficking. *Annu. Rev. Physiol.* **78**, 351–365 (2016).

4. 4.

Noguchi, J. et al. Bidirectional in vivo structural dendritic spine plasticity revealed by two-photon glutamate uncaging in the mouse neocortex. *Sci. Rep.* **9**, 13922 (2019).

5. 5.

Kasai, H., Ziv, N. E., Okazaki, H., Yagishita, S. & Toyoizumi, T. Spine dynamics in the brain, mental disorders and artificial neural networks. *Nat. Rev. Neurosci.* **22**, 407–422 (2021).

6. 6.

Missler, M., Sudhof, T. C. & Biederer, T. Synaptic cell adhesion. *Cold Spring Harb. Perspect. Biol.* **4**, a005694 (2012).

7. 7.

Kasai, H., Fukuda, M., Watanabe, S., Hayashi-Takagi, A. & Noguchi, J. Structural dynamics of dendritic spines in memory and cognition. *Trends Neurosci.* **33**, 121–129 (2010).

8. 8.

Sudhof, T. C. & Rothman, J. E. Membrane fusion: grappling with SNARE and SM proteins. *Science* **323**, 474–477 (2009).

9. 9.

Jahn, R. & Scheller, R. H. SNAREs—engines for membrane fusion. *Nat. Rev. Mol. Cell Biol.* **7**, 631–643 (2006).

10. 10.

Augustine, G. J. How does calcium trigger neurotransmitter release? *Curr. Opin. Neurobiol.* **11**, 320–326 (2001).

11. 11.

Tucker, W. C., Weber, T. & Chapman, E. R. Reconstitution of Ca^{2+} -regulated membrane fusion by synaptotagmin and SNAREs. *Science* **304**, 435–438 (2004).

12. 12.

Kasai, H., Takahashi, N. & Tokumaru, H. Distinct initial SNARE configurations underlying the diversity of exocytosis. *Physiol. Rev.* **92**, 1915–1964 (2012).

13. 13.

Takahashi, N. et al. Two-photon fluorescence lifetime imaging of primed SNARE complexes in presynaptic terminals and β cells. *Nat. Commun.* **6**, 8531 (2015).

14. 14.

Fatt, P. & Katz, B. Spontaneous subthreshold activity at motor nerve endings. *J. Physiol.* **117**, 109–128 (1952).

15. 15.

Hannezo, E. & Heisenberg, C. P. Mechanochemical feedback loops in development and disease. *Cell* **178**, 12–25 (2019).

16. 16.

Imig, C. et al. The morphological and molecular nature of synaptic vesicle priming at presynaptic active zones. *Neuron* **84**, 416–431 (2014).

17. 17.

Rosenmund, C. & Stevens, C. F. Definition of the readily releasable pool of vesicles at hippocampal synapses. *Neuron* **16**, 1197–1207 (1996).

18. 18.

Marvin, J. S. et al. An optimized fluorescent probe for visualizing glutamate neurotransmission. *Nat. Methods* **10**, 162–170 (2013).

19. 19.

Klapoetke, N. C. et al. Independent optical excitation of distinct neural populations. *Nat. Methods* **11**, 338–346 (2014).

20. 20.

Murakoshi, H., Shibata, A. C., Nakahata, Y. & Nabekura, J. A dark green fluorescent protein as an acceptor for measurement of Förster resonance energy transfer. *Sci. Rep.* **5**, 15334 (2015).

21. 21.

Delmas, P. & Coste, B. Mechano-gated ion channels in sensory systems. *Cell* **155**, 278–284 (2013).

22. 22.

Tapley, P., Lamballe, F. & Barbacid, M. K252a is a selective inhibitor of the tyrosine protein kinase activity of the trk family of oncogenes and neurotrophin receptors. *Oncogene* **7**, 371–381 (1992).

23. 23.

Doussau, F. & Augustine, G. J. The actin cytoskeleton and neurotransmitter release: an overview. *Biochimie* **82**, 353–363 (2000).

24. 24.

Cingolani, L. A. & Goda, Y. Actin in action: the interplay between the actin cytoskeleton and synaptic efficacy. *Nat. Rev. Neurosci.* **9**, 344–356 (2008).

25. 25.

Tanaka, J. et al. Protein synthesis and neurotrophin-dependent structural plasticity of single dendritic spines. *Science* **319**, 1683–1687 (2008).

26. 26.

Iino, M. Tension responses of chemically skinned fibre bundles of the guinea-pig taenia caeci under varied ionic environments. *J. Physiol.* **320**, 449–467 (1981).

27. 27.

Harris, K. M. & Stevens, J. K. Dendritic spines of CA 1 pyramidal cells in the rat hippocampus: serial electron microscopy with reference to their biophysical characteristics. *J. Neurosci.* **9**, 2982–2997 (1989).

28. 28.

de Vivo, L. et al. Ultrastructural evidence for synaptic scaling across the wake/sleep cycle. *Science* **355**, 507–510 (2017).

29. 29.

Watanabe, S. et al. Ultrafast endocytosis at *Caenorhabditis elegans* neuromuscular junctions. *Elife* **2**, e00723 (2013).

30. 30.

Honkura, N., Matsuzaki, M., Noguchi, J., Ellis-Davies, G. C. & Kasai, H. The subspine organization of actin fibers regulates the structure and plasticity of dendritic spines. *Neuron* **57**, 719–729 (2008).

31. 31.

Chang, J. Y. et al. CaMKII autophosphorylation is necessary for optimal integration of Ca^{2+} signals during LTP induction, but not maintenance. *Neuron* **94**, 800–808 (2017).

32. 32.

Kishimoto, T. et al. Vacuolar sequential exocytosis of large dense-core vesicles in adrenal medulla. *EMBO J.* **25**, 673–682 (2006).

33. 33.

Ninomiya, Y., Kishimoto, T., Miyashita, Y. & Kasai, H. Ca^{2+} -dependent exocytotic pathways in Chinese hamster ovary fibroblasts revealed by a caged- Ca^{2+} compound. *J. Biol. Chem.* **271**, 17751–17754 (1996).

34. 34.

Iino, Y. et al. Dopamine D2 receptors in discrimination learning and spine enlargement. *Nature* **579**, 555–560 (2020).

35. 35.

Shibata, A. C. E. et al. Photoactivatable CaMKII induces synaptic plasticity in single synapses. *Nat. Commun.* **12**, 751 (2021).

36. 36.

Dana, H. et al. High-performance calcium sensors for imaging activity in neuronal populations and microcompartments. *Nat. Methods* **16**, 649–657 (2019).

37. 37.

Grieger, J. C., Choi, V. W. & Samulski, R. J. Production and characterization of adeno-associated viral vectors. *Nat. Protoc.* **1**, 1412–1428 (2006).

38. 38.

Leung, R. W., Yeh, S. C. & Fang, Q. Effects of incomplete decay in fluorescence lifetime estimation. *Biomed. Opt. Express* **2**, 2517–2531 (2011).

39. 39.

Drobizhev, M., Makarov, N. S., Tillo, S. E., Hughes, T. E. & Rebane, A. Two-photon absorption properties of fluorescent proteins. *Nat. Methods* **8**, 393–399 (2011).

40. 40.

Kasai, H. et al. A new quantitative (TEPIQ) analysis for diameters of exocytic vesicles and its application to mouse pancreatic islets. *J. Physiol* **568**, 891–903 (2005).

41. 41.

Hayama, T. et al. GABA promotes the competitive selection of dendritic spines by controlling local Ca^{2+} signaling. *Nat. Neurosci.* **16**, 1409–1416 (2013).

Acknowledgements

We thank H. Murakoshi for providing ShadowG and paCaMKII; M. Iino for discussion; and S. Fujii, C. Fujinami, Y. Sakata, R. Yonemoto, H. Ohno and A. Kurabayashi for their technical assistance. This work was supported by CREST (JPMJCR21E2 and JPMJCR1652 to H.K.) from JST, KAKENHI (JP20H05685 to H.K., JP21K15203 to H.U., JP21K07535 to S.W., JP18K06497 to J.N., JP19J10613 and JP21K20682 to Y.M., JP20K15893 to Y.I., JP21H05176 and JP21H02594 to S.Y., and JP17K08530 and JP20K07280 to N.T.) from JSPS, SRPBS (JP20dm0107120 to H.K.) and Brain/MINDS (JP21dm0207069 to S.Y.) from AMED, the World Premier International Research Center Initiative (WPI) from MEXT, Takeda Science Foundation (to S.Y.) and a grant-in-aid of The Fugaku Trust for Medical Research (to H.K.).

Author information

Affiliations

1. Laboratory of Structural Physiology, Center for Disease Biology and Integrative Medicine, Faculty of Medicine, The University of Tokyo, Tokyo, Japan

Hasan Ucar, Satoshi Watanabe, Jun Noguchi, Yuichi Morimoto, Yusuke Iino, Sho Yagishita, Noriko Takahashi & Haruo Kasai

2. International Research Center for Neurointelligence (WPI-IRCN), UTIAS, The University of Tokyo, Tokyo, Japan

Hasan Ucar, Yuichi Morimoto, Yusuke Iino, Sho Yagishita & Haruo Kasai

3. Department of Ultrastructural Research, National Institute of Neuroscience, National Center of Neurology and Psychiatry, Tokyo, Japan

Satoshi Watanabe & Jun Noguchi

4. Department of Physiology, Kitasato University School of Medicine,
Sagamihara, Japan

Noriko Takahashi

Contributions

H.U. and H.K. designed the experiments. H.U., J.N., S.W. and Y.I. performed the slice experiments. S.W. and S.Y. assisted in the programming of MATLAB and LabView. Y.M. and N.T. helped with FRET data analysis. H.U. and H.K. wrote the manuscript, and all authors edited the manuscript.

Corresponding author

Correspondence to [Haruo Kasai](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Edwin Chapman, Ryohei Yasuda and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

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 Suppression of glutamate release by the strong pushing of a presynaptic terminal.

a, Strong pushing of a presynaptic bouton labelled with iGluSnFR (green) by a glass pipette (magenta). White dashed-lines indicate the outline of the boutons. Orange lines represent the axial line of the axon. Pushing was applied over the axial line. Scale bar 0.5 μm . **b**, Fluorescence profiles along the yellow line (inset) before (black), during (magenta) and after (cyan) the strong pushing of the terminal shown in **a**. **c, d**, The Pr measurements before and after the strong pushing of a bouton (**c**) and their averaged traces (**d**). Orange bars (LED) indicate optogenetic stimulation (5 ms). **e**, The mean time courses of Pr (left) and the peak of the averaged traces (right). $n = 12$ boutons (8 slices, 8 rats). For the baseline fluctuations (Baseline), 3 consecutive values (5 min apart) from all boutons were obtained. For Pr, two-sided Kruskal–Wallis test ($H = 22.2$, $**P = 1.85 \times 10^{-4}$) followed by Dunn’s multiple comparison test (vs the baseline). 6.5 min, $**P = 2.62 \times 10^{-3}$; 11.5 min, $**P = 4.2 \times 10^{-4}$; 20 min, $**P = 1.9 \times 10^{-3}$. For the averaged trace, ($H = 13.9$, $**P = 7.65 \times 10^{-3}$) followed by post-hoc Dunn’s multiple comparison test (vs the baseline). 6.5 min, $**P = 3.13 \times 10^{-3}$; 11.5 min, $*P = 0.015$; 20 min, $*P = 0.0194$. **f**, Time courses of ΔFWHM caused by the fine pushing (black) and strong pushing (orange). For fine pushing, the trace is the replicate of Fig. 1g. For strong pushing, $n = 13$ boutons (11 slices, 10 rats). Magenta bars in **e** and **f** indicate pushing. For strong pushing two-sided Kruskal–Wallis test ($H = 28.6$, $**P = 6.22 \times 10^{-7}$) followed by Dunn’s multiple comparison test (vs the baseline). Push (10 s), $**P = 4.3 \times 10^{-5}$; push (110 s), $**P = 4.77 \times 10^{-6}$. **g**, ΔFWHMs obtained from the same data as **f**. Two-tailed Mann–Whitney U test, $U = 232$, $**P = 7.28 \times 10^{-3}$. Error bars indicate s.e.m.

[Source data](#)

Extended Data Fig. 2 Pharmacological blockage of increases in FRET induced by pushing.

a, EPSCs recorded from the whole-cell clamped CA1 neurons. Optogenetic stimulations (5 ms, orange bars) were given to CA3 neurons, expressing

FRET probes (Control) or FRET probes and the light chain of BoNT^E (BoNT^E). **b, f**, FRET intensities (A1%, see [Methods](#)) of the boutons that were finely pushed in the presence of BoNT^E expression (**b**), or that were treated with K252a (2 μM, **f**). Scale bar indicates 0.5 μm. **c**, Time courses of ΔFWHM during fine pushing from HF side (black), LF side (open circle), in the absence of calcium (EGTA 100 μM, BAPTA-AM 30 μM, blue), in the presence of K252a (orange) or LatA (cyan) from the same data as Fig. [2e](#). The magenta bar indicates the duration of pushing. **d**, ΔFWHM of data in **c** during pushing. For the baseline fluctuations (Baseline), 2 consecutive FWHM values (20 s apart) from all boutons were plotted from the same data as Fig. [2c](#). Two-sided Kruskal–Wallis test ($H = 127$, ** $P = 6.41 \times 10^{-15}$) followed by Dunn's multiple comparison test (vs Baseline). HF side, $n = 17$, ** $P = 2.32 \times 10^{-10}$; LF side, $n = 10$, ** $P = 2.3 \times 10^{-8}$; BoNT^E, $n = 12$, ** $P = 4.7 \times 10^{-10}$; Ca²⁺-free, $n = 13$, ** $P = 2.8 \times 10^{-10}$; K252a, $n = 8$, ** $P = 7.07 \times 10^{-7}$; LatA, $n = 10$, ** $P = 1.02 \times 10^{-10}$. Kruskal–Wallis test ($H = 0.492$, $P = 0.992$) followed by Dunn's multiple comparison test (vs HF side). LF side, $P = 0.943$; BoNT^E, $P = 0.758$; Ca²⁺-free, $P = 0.819$; K252a, $P = 0.628$; LatA, $P = 0.563$. **e**, FRET images of a pushed bouton and in a neighbouring bouton. Scale bar 1 μm. Error bars indicate s.e.m. **g, h**, Replots of the same data shown in **e** and **f** using the average lifetime, instead of A1%

[Source data](#).

[Extended Data Fig. 3 The absence of the effect of pushing on the resting calcium concentrations and evoked calcium transients.](#)

a, GCaMP7s fluorescence images of a bouton (green) before and during pushing with a glass pipette (magenta). Scale bar 0.5 μm. **b**, The mean time course of GCaMP7s or 6s intensities before and during pushing (magenta bar) among 9 boutons. **c**, Evoked calcium transients measured with GCaMP7s fluorescence from a bouton before, during and after pushing. Orange arrowheads and dashed lines indicate the timing of optogenetic stimulation with a duration of 5 ms. **d**, The peak values of evoked calcium responses (Exp#1–5) and their average (black). $n = 5$ boutons (5 slices, 3

rats). Two-sided Kruskal–Wallis test ($H = 3.46$, $P = 0.177$) followed by Dunn’s multiple comparison test (vs Before). Pushing, $P = 0.752$; after, $P = 0.095$. Error bars indicate s.e.m

[Source data](#).

Extended Data Fig. 4 Alterations of mEPSC frequencies and evoked EPSC amplitudes by varying concentrations of sucrose.

a, mEPSCs recorded from a whole-cell clamped CA1 neuron. By adding 20, 40, 60 or 150 mM sucrose to ACSF (320 mOsm), we made a series of sucrose solutions with the osmolarities of 340 mM, 360 mM, 380 mM and 470 mM. They were sequentially applied, each followed by 10 min wash periods with ACSF. **b**, The mean frequencies of mEPSC in the sucrose solutions normalized by the one in ACSF. $n = 5$ cells (5 slices, 3 rats). Two-sided Kruskal–Wallis test ($H = 41.5$, $P = 2.11 \times 10^{-8}$) followed by Dunn’s multiple comparison test (vs the baseline). 20 mM, $P = 0.656$; 40 mM, $**P = 1.81 \times 10^{-3}$; 60 mM, $**P = 7.15 \times 10^{-4}$; 150 mM, $**P = 3.50 \times 10^{-8}$. **c**, Shrinkage of a bouton (left) by 20 mM sucrose quantified by the fluorescent profile (right). The profiles were aligned at the right side. $n = 9$ boutons (6 slices, 3 rats). Two-sided Kruskal–Wallis test ($H = 17.3$, $**P = 1.76 \times 10^{-4}$) followed by Dunn’s multiple comparison test (vs the baseline). 5 min, $**P = 5.09 \times 10^{-5}$. **d**, Evoked EPSCs recorded from CA1 neurons by optogenetic stimulation of the CA3 region (orange bar, 5 ms) in ACSF, isotonic sucrose (10 mM NaCl was replaced with 20 mM sucrose to yield an osmolarity the same as ACSF, 320 mOsm) and hypertonic sucrose (20 mM). We used the EPSCs from samples whose PPRs, with an interstimulus interval of 70 ms, were larger than 1.0 (see [Methods](#)). **e**, The mean time course for evoked EPSC amplitudes. Brown and magenta bars indicate stimulation with isotonic sucrose ($n = 6$) and hypertonic ($n = 7$) sucrose, respectively. **f**, Mean EPSC amplitudes during sucrose stimulations. For the baseline fluctuations (Baseline), 2 consecutive EPSC values (5 min apart) from all boutons were plotted. $n = 6$ cells (6 slices, 6 rats) for isotonic sucrose, $n = 7$ cells (7 slices, 7 rats) for hypertonic sucrose. Two-sided Kruskal–Wallis test ($H = 16.1$, $**P = 3.28 \times 10^{-4}$) followed by Dunn’s multiple comparison test

(vs the baseline). Isotonic, $P = 0.542$, Hypertonic, $**P = 6.5 \times 10^{-5}$. Error bars indicate s.e.m

[Source data](#).

Extended Data Fig. 5 Effects of the three types of pressure on cell structures.

Fine pushing and the osmotic pressure distort the cellular structures, whereas the hydrostatic pressure does not.

Extended Data Fig. 6 Reduction of the paired-pulse ratio by the low-sucrose solution.

a, b, iGluSnFR fluorescence intensity traces (**a**) in response to optogenetic (orange bars) paired-pulse stimulation and their averaged traces (**b**) before, during and after application of the ACSF solution containing 20 mM sucrose. The inter-stimulus interval was 70 ms. **c, d**, The mean time courses for PPR (**a**) and Δ PPR (**d**) for 5 boutons, in which PPR was calculated as the ratio of the peak values of the second and first pulses from the averaged traces (**b**). Magenta bars indicate sucrose application. Green horizontal bars at the x-axis indicate the time windows for the Pr analysis shown in **a**. $n = 5$ boutons (5 slices, 3 rats). Two-tailed Mann–Whitney U test, ACSF–sucrose (**c**), $U = 6$, $*P = 4.78 \times 10^{-2}$; ACSF–sucrose (**d**), $U = 2$, $**P = 7.99 \times 10^{-3}$. **e**, Δ PPR during ACSF and sucrose solution. For the baseline fluctuations (Baseline), 2 consecutive PPR values (5 min apart) from all boutons were plotted. $n = 4$ boutons. Two-tailed Mann–Whitney U test, $U = 2$, $**P = 7.99 \times 10^{-3}$. Error bars indicate s.e.m

[Source data](#).

Extended Data Fig. 7 cis-SNARE formation by a high-sucrose solution but not by a low-sucrose solution.

a, The design of the FRET probe for measuring the *cis*-SNARE assembly. Venus (Vn) and mTurquoise (mTq1) were fused with the C-terminals of the

SNAREs, syntaxin-1A and VAMP2. **b**, mEPSCs induced by the 300 mM sucrose solution (620 mOsm, HSS) recorded from the whole-cell clamped CA1 neuron. **c**, FRET images of *cis*-SNARE probe during application of LSS (20 mM sucrose) and after application of HSS (300 mM sucrose). Imaging was not possible during the application of 300 mM sucrose solution, because it caused an increase in the refractive index. Scale bar 0.5 μ m. **d**, The mean time course of Δ FRET, representing *cis*-SNARE formation (**a**). Magenta and purple bars indicate 20 mM and 300 mM sucrose solutions. **b1–b4** indicate time points when the images in **c** were obtained. $n = 11$ boutons (4 slices, 3 rats) for both 20 mM sucrose and 300 mM sucrose. **e**, Δ FRET of the *cis*-SNARE probe by the 20 mM and 300 mM sucrose solutions from the same dataset analysed in **d**. For the baseline fluctuations (Baseline), 3 consecutive FRET values (5 min apart) from all boutons were plotted. Two-sided Kruskal–Wallis test ($H = 48.27$, $**P = 8.27 \times 10^{-10}$) followed by Dunn’s multiple comparison test (vs Baseline). 20 mM sucrose, $P = 0.394$; 300 mM sucrose, $**P = 1.33 \times 10^{-7}$. Error bars indicate s.e.m

[Source data.](#)

[Extended Data Fig. 8 Identification of synaptic contacts.](#)

a–c, A bouton-spine pair which showed a precise horizontal alignment and a substantial overlap. The bouton and spine were labelled with iGluSnFR and Alexa594, respectively. The X-axis was redefined according to the bouton-spine axis, and fluorescent profiles along the X-axis of XY- and XZ projections were plotted ([Methods](#)). The fluorescent profiles were normalized at their peaks. The crossing of the two profiles defines Fz and Fx, which were 0.93 and 0.9 in this example, and met the criteria of $Fz > 0.8$ and $Fx > 0.6$. **d–f**, An example of bouton-spine pairs which did not fulfil the criteria ($Fz = 0.43$ and $Fx = 0.6$). **g**, The bouton and spines were also labelled with mTq2-syntaxin-1A and Alexa594, which fulfilled the criteria with $Fz = 0.95$ and $Fx = 0.7$ in this example. Horizontal scale bars indicate 1 μ m, vertical scale bars indicate 2 μ m. **h**, Detection of functional connectivity in a bouton-spine pair by optogenetic presynaptic stimulation with CsChrimsonR and Ca^{2+} responses (Cal520) in postsynaptic spines. **i**,

The Ca^{2+} responses were induced in 3 out of 10 stimulations in the pair shown in **h**. LED pulses of 5 ms were delivered every 15 s.

Extended Data Fig. 9 ΔFWHM and ΔFRET of the boutons during spine enlargement.

a, The values of FWHM of the bouton at 6 min after the onset of STDP from the data shown in Fig. 4e. $n = 14$ pushing spines (14 slices, 11 rats) and $n = 9$ non-pushing spines (9 slices, 8 rats). Two-sided Wilcoxon signed-rank test vs 0 μm , $W = 4$, $**P = 2.58 \times 10^{-3}$ and $W = 41$, $*P = 3.3 \times 10^{-2}$). **b**, Correlation between ΔFWHM of the bouton and ΔLength of the spine at 6 min after STDP induction from the data shown in **a**. Red and black circles indicate pushing and non-pushing spines, respectively. Spearman's correlation with two tails, $\rho = 0.406$, $**P = 1.38 \times 10^{-3}$. **c**, Correlation between ΔFRET and ΔFWHM of the boutons at 6 min after STDP induction from the data shown in **a**. Spearman's correlation with two tails, $\rho = 0.557$, $**P = 6.8 \times 10^{-5}$. **d**, Correlation between ΔFRET of the bouton and ΔV of the spine at 6 min after STDP induction from the data shown in **a**. Spearman's correlation with two tails, $\rho = 0.002$, $P = 0.848$. **e, f**, Absence of spine enlargement (**e**) and FRET changes (**f**) when uncaging was applied without postsynaptic spikes (green, $n = 7$ boutons, 7 slices, 5 rats). The control STDP traces are the replicates from the red traces in Fig. 4d, g. Two-sided Kruskal–Wallis test ($H = 4.39$, $P = 0.356$) for (**e**) and ($H = 0.48$, $P = 0.975$) for (**f**). Error bars indicate s.e.m

[Source data](#).

Extended Data Fig. 10 Enhancement of glutamate release by spine enlargement.

a, d, Fluorescence images (single XY- and XZ-planes) of the bouton-spine pairs, and dendritic spines before and after stimulation with STDP protocol for a pushing (**a**) and a non-pushing (**d**) spine. Horizontal scale bars indicate 1 μm , vertical scale bars indicate 2 μm . In all bouton-spine pairs shown in this figure, the criteria ($F_z > 0.8$ and $F_x > 0.6$) were met ([Methods](#)). **b–f**, The Pr measurements (**b, e**) and the averaged iGluSnFR traces (**c, f**) before

and after induction of STDP. Orange bars indicate optogenetic stimulation (5 ms). **g–j**, The time courses of spine enlargement (**g**), spine length (**h**), Pr (**i**) and the peak of the averaged iGluSnFR traces (**j**) in pushing (red) and non-pushing (black) spines for the same set of data shown in **k–n**. **k–n**, The increases in spine volume (ΔV , **k**), spine length (ΔL , **l**), Pr (ΔPr , **m**) and average release ($\Delta Release$, **n**) in the synapses with pushing and without pushing. $n = 6$ (6 slices, 6 rats) and 5 (5 slices, 5 rats) for pushing and non-pushing spines at 10 min after STDP induction. Two-sided Mann–Whitney *U* test ($U = 11$, $P = 0.537$ [**k**]). Two-sided Wilcoxon signed-rank test ($W = 21$, $*P = 0.036$ [**l**]; $W = 21$, $*P = 0.036$ [**m**]; $W = 21$, $*P = 0.036$ [**n**]) for pushing synapses and ($W = 3$, $P = 0.281$ [**l**]; $W = 1$, $P = 0.104$ [**m**]; $W = 1$, $P = 0.106$ [**n**]) for non-pushing synapses. **o**, **p**, Correlation for ΔPr (**o**) and $\Delta Release$ (**p**) vs ΔV from the same dataset as **k–n**. Spearman’s correlation, rho = 0.257, $P = 0.116$ (**o**) and rho = 0.132, $P = 0.273$ (**p**). **q**, **r**, Correlation for ΔPr (**q**) and $\Delta Release$ (**r**) vs ΔL with (red) and without (black) spine elongation from the same dataset as **k–n**. Spearman’s correlation, rho = 0.469, $*P = 0.025$ (**q**) and rho = 0.746, $**P = 1.25 \times 10^{-3}$ (**r**). Error bars indicate s.e.m.

[Source data](#).

Supplementary information

[Reporting Summary](#)

[Peer Review File](#)

[Supplementary Video 1](#)

Fine pushing of a bouton (iGluSnFR). An example of fine pushing of a presynaptic bouton, which expresses iGluSnFR (green), with a glass pipette (magenta). Note the change in the bouton shape and its immediate recovery after pipette retraction, as analysed in Fig. 1g. The distortion of the bouton (Fig. 1f) was very small (0.1 μm) but significant.

[Supplementary Video 2](#)

Fine pushing of a bouton (iSLIM). An example of the fine pushing of the presynaptic bouton, which expresses mTurquoise2-syntaxin-1A, as shown in Fig. 2b. The distortion of the bouton was very small but significant, as analysed in Extended Data Fig. 2c, d. The pipette was retracted 1 min after pushing.

Supplementary Video 3

Stimulation and pushing of a bouton (GCaMP7s). GCaMP7s imaging of a presynaptic bouton in the CA1 region. The bouton was optogenetically stimulated at the CA3 region. Three out of ten stimulations are displayed, and indicated as stim1, stim2 and stim3. The white dot indicates the LED stimulation. Despite responding to the optogenetic stimulation (Extended Data Fig. 3c), the bouton showed no increase in fluorescence intensity by pushing (Extended Data Fig. 3b).

Source data

Source Data Fig. 1

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. 6

[**Source Data Extended Data Fig. 7**](#)

[**Source Data Extended Data Fig. 9**](#)

[**Source Data Extended Data Fig. 10**](#)

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Ucar, H., Watanabe, S., Noguchi, J. *et al.* Mechanical actions of dendritic-spine enlargement on presynaptic exocytosis. *Nature* **600**, 686–689 (2021). <https://doi.org/10.1038/s41586-021-04125-7>

- Received: 01 August 2020
- Accepted: 12 October 2021
- Published: 24 November 2021
- Issue Date: 23 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04125-7>

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

Further reading

- **Push the bouton**

- Sian Lewis

Nature Reviews Neuroscience (2021)

Forceful synapses reveal mechanical interactions in the brain

Research Briefing 24 Nov 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04125-7>

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

- Article
- [Published: 08 December 2021](#)

Collective durotaxis along a self-generated stiffness gradient *in vivo*

- [Adam Shellard ORCID: orcid.org/0000-0002-7609-9049¹](#) &
- [Roberto Mayor ORCID: orcid.org/0000-0001-9053-9613¹](#)

Nature volume **600**, pages 690–694 (2021)

- 6975 Accesses
- 155 Altmetric
- [Metrics details](#)

Subjects

- [Cell biology](#)
- [Developmental biology](#)

Abstract

Collective cell migration underlies morphogenesis, wound healing and cancer invasion^{1,2}. Most directed migration *in vivo* has been attributed to chemotaxis, whereby cells follow a chemical gradient^{3,4,5}. Cells can also follow a stiffness gradient *in vitro*, a process called durotaxis^{3,4,6,7,8}, but evidence for durotaxis *in vivo* is lacking⁶. Here we show that in *Xenopus laevis* the neural crest—an embryonic cell population—self-generates a stiffness gradient in the adjacent placodal tissue, and follows this gradient by durotaxis. The gradient moves with the neural crest, which is continually

pursuing a retreating region of high substrate stiffness. Mechanistically, the neural crest induces the gradient due to N-cadherin interactions with the placodes and senses the gradient through cell–matrix adhesions, resulting in polarized Rac activity and actomyosin contractility, which coordinates durotaxis. Durotaxis synergizes with chemotaxis, cooperatively polarizing actomyosin machinery of the cell group to prompt efficient directional collective cell migration *in vivo*. These results show that durotaxis and dynamic stiffness gradients exist *in vivo*, and gradients of chemical and mechanical signals cooperate to achieve efficient directional cell migration.

This is a preview of subscription content

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Fig. 1: A dynamic self-generated stiffness gradient in vivo.

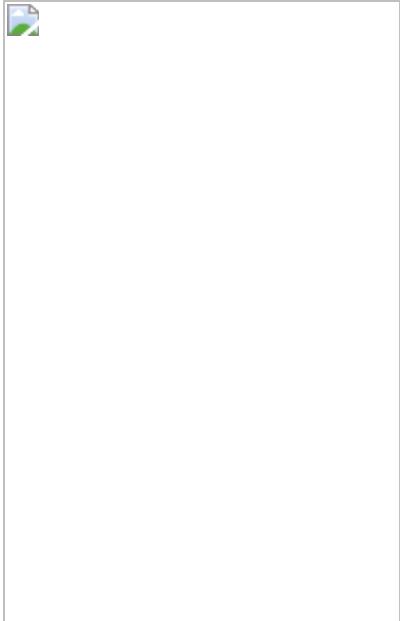


Fig. 2: Neural crest durotaxis in vivo.

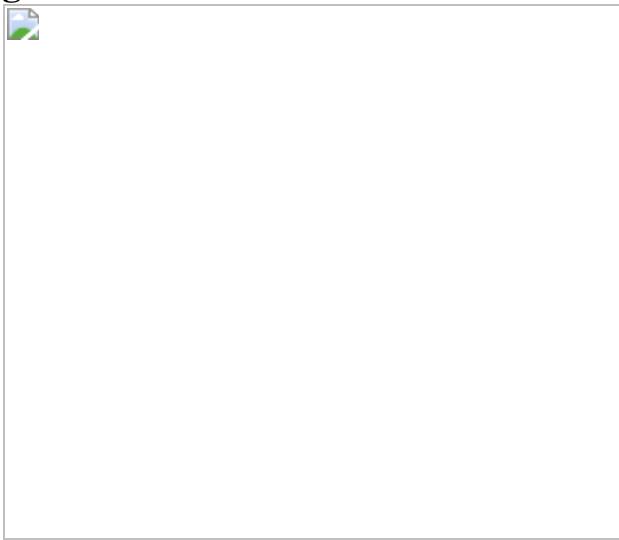


Fig. 3: Polarized actomyosin contraction and Rac during durotaxis.

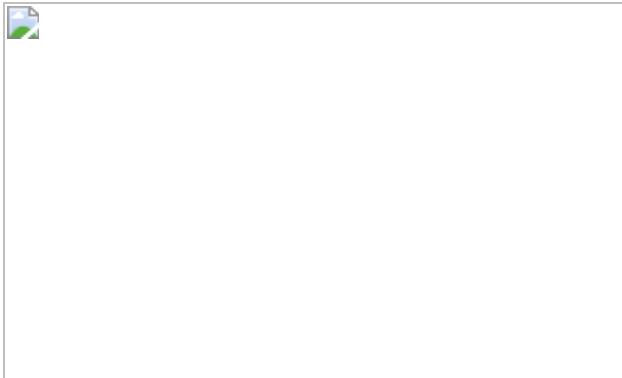
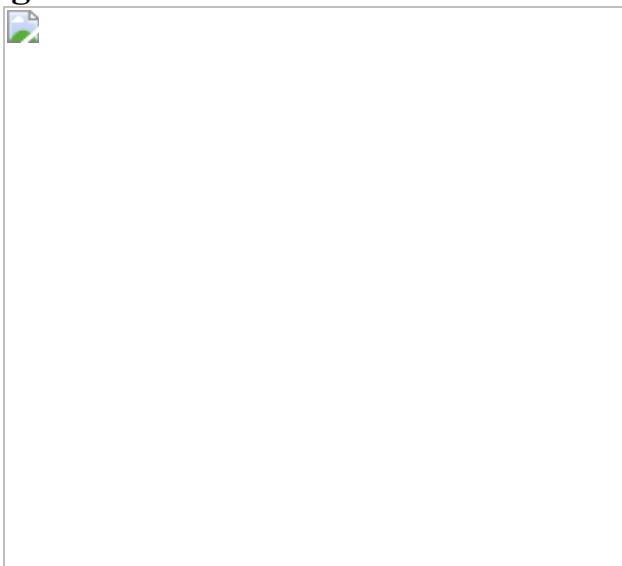


Fig. 4: Durotaxis and chemotaxis cooperatively coordinate neural crest migration.



Data availability

The data supporting the findings of this study are available within the Article and its Supplementary Information. Source data are provided with the paper.

References

1. 1.

Yamada, K. M. & Sixt, M. Mechanisms of 3D cell migration. *Nat. Rev. Mol. Cell Biol.* **20**, 738–752 (2019).

2. 2.

Friedl, P. & Gilmour, D. Collective cell migration in morphogenesis, regeneration and cancer. *Nat. Rev. Mol. Cell Biol.* **10**, 445–457 (2009).

3. 3.

Shellard, A. & Mayor, R. All roads lead to directional cell migration. *Trends Cell Biol.* **30**, 852–868 (2020).

4. 4.

SenGupta, S., Parent, C. A. & Bear, J. E. The principles of directed cell migration. *Nat. Rev. Mol. Cell Biol.* **22**, 529–547 (2021).

5. 5.

Insall, R. H. Understanding eukaryotic chemotaxis: a pseudopod-centred view. *Nat. Rev. Mol. Cell Biol.* **11**, 453–458 (2010).

6. 6.

Shellard, A. & Mayor, R. Durotaxis: the hard path from in vitro to in vivo. *Dev. Cell* **56**, 227–239 (2021).

7. 7.

Lo, C. M., Wang, H. B., Dembo, M. & Wang, Y. L. Cell movement is guided by the rigidity of the substrate. *Biophys. J.* **79**, 144–152 (2000).

8. 8.

Sunyer, R. et al. Collective cell durotaxis emerges from long-range intercellular force transmission. *Science* **353**, 1157–1161 (2016).

9. 9.

Charras, G. & Sahai, E. Physical influences of the extracellular environment on cell migration. *Nat. Rev. Mol. Cell Biol.* **15**, 813–824

(2014).

10. 10.

Guimaraes, C. F., Gasperini, L., Marques, A. P. & Reis, R. L. The stiffness of living tissues and its implications for tissue engineering. *Nat. Rev. Mater.* **5**, 351–370 (2020).

11. 11.

Barriga, E. H., Franze, K., Charras, G. & Mayor, R. Tissue stiffening coordinates morphogenesis by triggering collective cell migration in vivo. *Nature* **554**, 523–527 (2018).

12. 12.

Theveneau, E. et al. Chase-and-run between adjacent cell populations promotes directional collective migration. *Nat. Cell Biol.* **15**, 763–772 (2013).

13. 13.

Szabo, A. & Mayor, R. Mechanisms of neural crest migration. *Ann. Rev. Genet.* **52**, 43–63 (2018).

14. 14.

Shellard, A. & Mayor, R. Chemotaxis during neural crest migration. *Semin. Cell Dev. Biol.* **55**, 111–118 (2016).

15. 15.

Alfandari, D., Cousin, H., Gaultier, A., Hoffstrom, B. G. & DeSimone, D. W. Integrin $\alpha 5\beta 1$ supports the migration of *Xenopus* cranial neural crest on fibronectin. *Dev. Biol.* **260**, 449–464 (2003).

16. 16.

Wang, S. J. et al. Tiam1 interaction with the PAR complex promotes talin-mediated Rac1 activation during polarized cell migration. *J. Cell Biol.* **199**, 331–345 (2012).

17. 17.

Schiller, H. B. et al. β 1- and α v-class integrins cooperate to regulate myosin II during rigidity sensing of fibronectin-based microenvironments. *Nat. Cell Biol.* **15**, 625–636 (2013).

18. 18.

Shellard, A., Szabo, A., Trepat, X. & Mayor, R. Supracellular contraction at the rear of neural crest cell groups drives collective chemotaxis. *Science* **362**, 339–343 (2018).

19. 19.

Guilluy, C., Garcia-Mata, R. & Burridge, K. Rho protein crosstalk: another social network? *Trends Cell Biol.* **21**, 718–726 (2011).

20. 20.

Theveneau, E. et al. Collective chemotaxis requires contact-dependent cell polarity. *Dev. Cell* **19**, 39–53 (2010).

21. 21.

Koser, D. E. et al. Mechanosensing is critical for axon growth in the developing brain. *Nat. Neurosci.* **19**, 1592–1598 (2016).

22. 22.

Zhu, M. et al. Spatial mapping of tissue properties *in vivo* reveals a 3D stiffness gradient in the mouse limb bud. *Proc. Natl Acad. Sci. USA* **117**, 4781–4791 (2020).

23. 23.

Tse, J. R. & Engler, A. J. Stiffness gradients mimicking in vivo tissue variation regulate mesenchymal stem cell fate. *Plos ONE* **6**, e15978 (2011).

24. 24.

Tweedy, L. et al. Seeing around corners: cells solve mazes and respond at a distance using attractant breakdown. *Science* **369**, eaay9792 (2020).

25. 25.

Dona, E. et al. Directional tissue migration through a self-generated chemokine gradient. *Nature* **503**, 285–289 (2013).

26. 26.

Thompson, A. J. et al. Rapid changes in tissue mechanics regulate cell behaviour in the developing embryonic brain. *eLife* **8**, e39356 (2019).

27. 27.

Plotnikov, S. V., Pasapera, A. M., Sabass, B. & Waterman, C. M. Force fluctuations within focal adhesions mediate ECM-rigidity sensing to guide directed cell migration. *Cell* **151**, 1513–1527 (2012).

28. 28.

Nieuwkoop, P. & Faber, J. *Normal Table of Xenopus laevis (Daudin): A Systematical and Chronological Survey of the Development from the Fertilized Egg Till the End of Metamorphosis* 2nd edn (1967).

29. 29.

Shi, J. L., Severson, C., Yang, J. X., Wedlich, D. & Klymkowsky, M. W. Snail2 controls mesodermal BMP/Wnt induction of neural crest. *Development* **138**, 3135–3145 (2011).

30. 30.

Tse, J. R. & Engler, A. J. in *Current Protocols in Cell Biology* Vol. 47 (Wiley Interscience, 2010).

31. 31.

Yue, L., Wang, S., Wulf, V. & Willner, I. Stiffness-switchable DNA-based constitutional dynamic network hydrogels for self-healing and matrix-guided controlled chemical processes. *Nat. Commun.* **10**, 4774 (2019).

32. 32.

Giobbe, G. G. et al. Extracellular matrix hydrogel derived from decellularized tissues enables endodermal organoid culture. *Nat. Commun.* **10**, 5658 (2019).

33. 33.

Karoutas, A. et al. The NSL complex maintains nuclear architecture stability via lamin A/C acetylation. *Nat. Cell Biol.* **21**, 1248–1260 (2019).

34. 34.

Chen, C. Y. et al. Suppression of detyrosinated microtubules improves cardiomyocyte function in human heart failure. *Nat. Med.* **24**, 1225–1233 (2018).

35. 35.

Ding, Y., Xu, G. K. & Wang, G. F. On the determination of elastic moduli of cells by AFM based indentation. *Sci. Rep.* **7**, 45575 (2017).

36. 36.

Gouignard, N., Rouvière, C. & Theveneau, E. in *The Epithelial-to Mesenchymal Transition* Vol. 2179 (eds Campbell, K. & Theveneau, E.) 257–274 (2021).

37. 37.

Barriga, E. H., Shellard, A. & Mayor, R. in *Neural Crest Cells* Vol. 1976 (eds Schwarz, Q. & Wiszniak, S.) 135–152 (Humana, 2019).

38. 38.

Dubaissi, E. et al. A secretory cell type develops alongside multiciliated cells, ionocytes and goblet cells, and provides a protective, anti-infective function in the frog embryonic mucociliary epidermis. *Development* **141**, 1514–1525 (2014).

39. 39.

Zhang, C., Basta, T., Jensen, E. D. & Klymkowsky, M. W. The beta-catenin/VegT-regulated early zygotic gene *Xnr5* is a direct target of SOX3 regulation. *Development* **130**, 5609–5624 (2003).

40. 40.

Schlosser, G. & Ahrens, K. Molecular anatomy of placode development in *Xenopus laevis*. *Dev. Biol.* **271**, 439–466 (2004).

41. 41.

Schlosser, G. et al. Eya1 and Six1 promote neurogenesis in the cranial placodes in a SoxB1-dependent fashion. *Dev. Biol.* **320**, 199–214 (2008).

42. 42.

Cao, X. et al. A phosphorylation switch controls the spatiotemporal activation of Rho GTPases in directional cell migration. *Nat. Commun.* **6**, 7721 (2015).

43. 43.

Kanoldt, V. et al. Metavinculin modulates force transduction in cell adhesion sites. *Nat. Commun.* **11**, 6403 (2020).

44. 44.

Karimi, K. et al. Xenbase: a genomic, epigenomic and transcriptomic model organism database. *Nucleic Acids Res.* **46**, D861–D868 (2018).

Acknowledgements

We thank G. Charras for the PDMS and discussions; J. Hartmann for assistance with calculating synergy statistics; E. Theveneau for gifting the Eya1 and Foxi1c probes; M. Klymkowsky for gifting the anti-Sox3 antibody; E. Barriga for preliminary experiments; and G. Charras, B. Stramer and J. Hartmann for comments on the manuscript. Work in R.M.’s laboratory is supported by grants from the Medical Research Council (MR/S007792/1), Biotechnology and Biological Sciences Research Council (M008517, BB/T013044) and Wellcome Trust (102489/Z/13/Z).

Author information

Affiliations

1. Department of Cell and Developmental Biology, University College London, London, UK

Adam Shellard & Roberto Mayor

Contributions

A.S. and R.M. conceived the project and designed the experiments. A.S. performed and analysed all of the experiments. R.M. contributed to experimental repeats and analysis of the data in Figs. [1c](#) and [2b, e](#). A.S. and R.M. wrote and edited the manuscript.

Corresponding author

Correspondence to [Roberto Mayor](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work.

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 Neural crest use Placodes as substrate for migration.](#)

a, Schematic of lateral view of an embryo showing the position of neural crest (pink) and placodes (blue) before neural crest migration. The black line represents the position of a cross section shown in c. **b**, Double *in situ* hybridization against the placodal marker *Six1* and the neural crest marker *Twist* at pre-migratory stage²³. Scale bar is 500 µm (b). **c, e**, Diagram illustrations of the neural crest environment *in vivo*, engaged in “chase and run”¹² interaction with placodes. The diagram is illustrative of a cross-section through the embryo, has illustrated by the black line in a. Neural crest (pink) chemotax toward Sdf1-secreting placode (Sdf1, purple; placode, blue). The two cell types interact causing placode to run away. **d, f**, Cryosection images showing fluorescent *in situ* hybridization for *Twist* (neural crest, magenta), and immunostaining for *Sox3* (placodes, cyan) and fibronectin (grey in single channel; yellow in merge). The fibronectin and merge panels are zooms of the white dashed boxes. Note that the neural crest migrates toward the placodes and that a fibronectin layer interfaces the two tissues. Scale bar is 50 µm (d, f). **g, i**, Fibronectin surround neural crest

stream. **g**, Schematic of fibronectin (green) surrounding the neural crest at the interface with mesoderm and placodes. **h**, Double ISH (Twist) and immunostaining (fibronectin) at the interfaces shown in the squares in g. Scale bar is 20 μm (h). **i**, Quantification of fibronectin levels at the interface of neural crest with Placodes and Mesoderm. Note the higher levels of fibronectin at the interface of neural crest and placodes than with mesoderm. Thick bars (i) represent mean; error bars (i) represent s.d.; unpaired two-tailed *t*-test (i); *** $P\leq 0.0001$; $n = 29$ embryos each (i). Statistics and reproducibility are in the source data and [Methods](#)

[Source Data](#).

[**Extended Data Fig. 2 Dynamic stiffness gradient in placode cells.**](#)

a–b, Diagram illustrating removal of the epidermis to expose the placodes for atomic force microscopy (a). Cryosection images (b) showing that the placodes are the most superficial tissue after epidermis removal. The right merge panel is a zoom of the white dashed box. Scale bar is 50 μm (b). **c**, *in situ* hybridization against the placodal marker *Eya1* at migratory stages on a representative embryo on the control side and side where the epidermis was dissected. Note that the placodal tissue is unaffected by epidermal dissection. Scale bar is 200 μm (c). **d–g**, Apparent elasticity measurements *in vivo* at the start of neural crest migration (c, d). Note that b and c represent the same data set. Stiffness measurements of the representative heat map in Fig. [1d](#), a stage 22 embryo (e) and where the epidermis was not removed (f). **h–j**, Placode deletion. **h**, Schematics of the different treatments. **i**, *in situ* hybridization against the placodal marker *Eya1* after each treatment. Stiffness values after each treatment are shown in Fig [1d](#). Scale bar is 200 μm (i). **j**, *in situ* hybridization against the neural crest marker *Twist* in conditions stated. Note that neural crest migration depends on cranial placodes. Scale bar is 200 μm (j). **k**, Quantification of the distance of stiffness gradients measured by nanoindentation and cranial placode tissue *in vivo*. **l–n**, Diagram illustration of the neural crest and placodes *in vivo* (l). The area indicated represent the placode whose stiffness was measured after epidermis removal (not shown). Quantification of apparent elasticity measurements of the cranial placodes (m). Note that

the gradient emerges at the onset of neural crest migration, and that the gradient persists as the neural crest migrate over time. A corresponding heat map representing averaged stiffness for embryos at different stages along the dorsoventral axis is shown in Fig 1e. Comparison of mesodermal and placodal stiffness over the dorsoventral axis (n). The dashed line indicates the rear of the placode. Thick bars (d, g, k, m, n) and circles (m) represent mean; error bars represent s.d. (d, g, k, m, n); Tukey's test (d), Dunn's test (g), two-tailed Mann-Whitney *U* test (k), Wilcoxon match-pairs signed rank test (m, st. 21), paired two-tailed *t*-test (m, st. 22, st. 23, st. 24); ns, $P>0.05$, *** $P\leq0.0001$; $n = 11$ (d, e), 20 (g), 17 (k), 10 (m) embryos, 10 (f) linear lines. Diagrams in h are adapted from *Normal table of Xenopus laevis (Daudin): a systematical and chronological survey of the development from the fertilized egg till the end of metamorphosis*. Copyright © 1994, Nieuwkoop and Faber. Reproduced by permission of Taylor and Francis books US. Statistics and reproducibility are in the source data and [Methods](#)

[Source Data](#).

[**Extended Data Fig. 3 The neural crest self-generates the stiffness gradient through N-Cadherin.**](#)

a, b, Schematics indicating the different treatments (a) and embryos stained by in situ hybridization for the neural crest marker, *Slug* (b). Heat maps of this experiments are shown in Fig. 1f. Scale bar is 200 μm (b). **c**, Embryos stained by in situ hybridization for the neural crest marker, *Slug*. Black boxes represent the region in which nanoindentation was performed; mean stiffness heat map of each condition is shown in Fig. 1h. Scale bar is 200 μm (c). **d**, Diagram illustrating “chase and run”¹². Neural crest (pink) chemotax toward Sdf1-secreting placode (Sdf1, purple; placode, blue). The two cell types interact through N-Cadherin (green), causing placode to run away (migration indicated by black arrow). **e**, Diagram illustrating placode cultured on fibronectin (grey) or fibronectin with N-Cadherin. Placodal stiffness measured in these conditions is shown in Fig. 1i. **f**, Schematics indicating the different treatments. Results are shown in Fig. 1l–n. **g, h**, exogenous stiffness gradient formation. Stiffness heat map from a representative embryo in which an exogenous local pressure was applied ventral to the neural crest as depicted in g (bottom of the heat map, g), and

quantification along the axis (h). Thick lines (h) represent mean; error bars (h) represent s.d.; Dunn's test (h); * $P\leq 0.05$, ** $P\leq 0.01$; $n = 4$ (h) linear lines. Diagrams in **a, f** are adapted from *Normal table of Xenopus laevis (Daudin)*. Copyright © 1994, Nieuwkoop and Faber. Reproduced by permission of Taylor and Francis books US. Statistics and reproducibility are in the source data and [Methods](#)

[Source Data](#).

Extended Data Fig. 4 Neural crest durotaxis in vivo and on steep gradients ex vivo.

a, b, Graft of fluorescently labelled neural crest into control embryos (cyan) or ablated embryos (magenta; **a**, centred cell tracks; **b**, track angles). Grafted embryos are shown in Fig. [2d, e](#). **c**, Apparent elasticity measurements of steep stiffness (blue) and uniform (grey) gradient gels. **d–g**, Neural crest explants with labelled nuclei (magenta) and membrane (cyan) (**d, f**) and time-coded projected tracks (**e, g**) on gels of uniform stiffness (**d, e**) and graded stiffness (**f, g**). Scale bar is 50 μm (**d–g**). **h–k**, Cell tracks (**h, j**) and angles of movement (**i, k**) from clusters on gels of uniform (**h, i**) or graded (**j, k**) stiffness; **l–n**, Formula for tactic index (**l**), quantification of tactic index (**m**) and speed (**n**). Thick bars (**c, m, n**) represent mean; error bars (**c, m, n**) represent s.d.; two-tailed Mann-Whitney U test (**m, n**); *** $P\leq 0.0001$; $n = 6$ (**c**) gels, 360 (**i, k**) cells from 18 clusters, 17 (**m, n**) explants. Statistics and reproducibility are in the source data and [Methods](#)

[Source Data](#).

Extended Data Fig. 5 Rear actomyosin contraction is essential for collective durotaxis.

a, Neural crest in vivo expressing fluorescently tagged myosin light chain II (MLC) and LifeAct. Note that a supracellular actomyosin cable exists at the edge of the neural crest cell group in vivo. Arrowheads indicate the actomyosin cable. Scale bar is 25 μm (**a**). **b**, Kymograph of the edge of the neural crest cluster in vivo. Green represents MLC, which is absent from

cell-cell contacts, and red represents LifeAct. Note the in vivo contraction of the actomyosin cable. Cell-cell contact contraction is indicated with the black arrowheads. **c**, Neural crest ex vivo expressing fluorescently tagged myosin (MLC, myosin light chain), LifeAct and membrane marker. Corresponding low magnifications are shown in Fig. [4a](#). Scale bar is 25 μm (c). **d**, Two time points from the edge of a neural crest cluster on a stiffness gradient. Yellow arrowheads mark cell-cell contacts. Note the contraction of the actomyosin cable and reduction in length between cell contacts. Scale bar is 10 μm (d). **e**, Heat map of an example actomyosin contraction at the edge of clusters in control, durotaxis and chemotaxis. Time point zero represents the start of actomyosin contraction. **f**, Quantification of actomyosin contraction by cable length. Note that the amplitude of contraction is the same in all conditions. **g, h**, An illustrative diagram showing rear contraction and front movement of a cluster (circle outline) at three time points (g), and a histogram representing the time at which front movement occurs relative to rear contraction (h). The dashed line indicates the rear contraction time point reference, $t = 0$. **i, j**, Tactic index (i) and speed (j) of clusters on shallow stiffness gradient gels exposed to the myosin II inhibitor, blebbistatin. **k-m**, Pictures of membrane and merge of LifeAct and membrane for the example ablation shown in Fig. [3c-g](#) (k, top, middle) or for cytoplasmic ablation (k, bottom). Yellow arrowheads indicate location of ablation. Scale bar is 10 μm (k). **l**, Migration of neural crest clusters on physiological stiffness gradients. The dashed line (start of ablations) separates before and during laser ablation of the cytoplasm. **m**, Tactic index of clusters before and during actomyosin cable ablation in the front or rear portion of migrating cell groups, or in the cytoplasm. Thick bars (f, i, j, l, m) represent mean; error bars (f, i, j, l, m) represent s.d.; unpaired two-tailed *t*-test (i), two-tailed Mann-Whitney *U* test (j), Dunn's test (m); ns, $P > 0.05$, * $P \leq 0.05$, *** $P \leq 0.0001$; $n = 20$ (f), 15 (h), 30 (i, j), 6 (l, m) clusters. Statistics and reproducibility are in the source data and [Methods](#)

[Source Data](#)

[Extended Data Fig. 6 Rac during durotaxis.](#)

a–b, Cells at the front and rear of neural crest explants on a stiffness gradient, with Rac-GTP, Phalloidin and DAPI labelling (a), and quantification of Rac-GTP along the axis from the edge of the cluster inwards (b). Note that Rac-GTP preferentially accumulates at the cell edge irrespective of its position within the cluster, consistent with previous observations of Rac polarity. Scale bar is 10 μm (a). **c, d**, Immunostaining of vinculin with Phalloidin in explants on uniform stiffness (c, top; m) or a physiological stiffness gradient (c, bottom) and polarity of the number of vinculin spots quantified (d). Scale bar is 5 μm (c). **e, f**, Immunostaining of Rac-GTP in control or Integrin- $\beta 1$ knockdown (e) and quantification of its polarity (f). Scale bar is 50 μm (e). Thick bars (b, d, f) represent mean; error bars (b, d, f) represent s.d.; two-tailed Mann-Whitney U test (d, f), *** $P \leq 0.0001$; $n = 20$ (b, d, f) clusters. Statistics and reproducibility are in the source data and [Methods](#)

[Source Data](#).

[Extended Data Fig. 7 Durotaxis and chemotaxis cooperatively coordinate neural crest migration in vivo.](#)

a, Schematics indicating the different treatments. Results are shown in Fig. 4a, b. **b**, Stiffness measurements in control and ablated embryos. Quantification of in vivo migration in each condition is shown in Fig 5b. **c**, An example heat map of stiffness from local pressure treatment (c) as depicted in Fig. 4c and quantification of the exogenous stiffness gradient (d). Results are shown in Fig. 4c. **e–l**, Ectopic migration analysis. **e**, Schematic illustrating how the ectopic migration index (emi) was calculated. For each neural crest stream a vector was drawn from their origin to the final position of migration. The control and experimental side of the same embryos were analysed. For the control side the vector always lays in the migratory pathway, while for the experimental side some vectors point to ectopic locations. **f**, The difference between these two vectors $(\mathop{\rm a}\limits^{} - \mathop{\rm b}\limits^{})$ generates the ectopic migration, which normalized by the control vector corresponds to the emi, which is shown as an scalar value in Fig. 4d. **g, h**, emi vectors for the experiment described in Fig. 4c. **i–l**, In situ hybridization for *Twist* of

control and experimental side of embryos treated with exogenous local pressure (i, j) or an SDF1 bead (k, l) and the associated vectors. Scale bar is 250 μm (i, k). Thick bars (b, d) represent mean; error bars (b, d) represent s.d.; Tukey's test (b, control; d), Dunn's test (b, ablation); ns, $P>0.05$, * $P\leq0.05$, ** $P\leq0.001$, *** $P\leq0.0001$; $n = 11$ (b, control), 10 (b, ablation), 12 (d), 9 (g, h, j, l) embryos. Diagrams in **a** are adapted from *Normal table of Xenopus laevis (Daudin)*. Copyright © 1994, Nieuwkoop and Faber. Reproduced by permission of Taylor and Francis books US. Statistics and reproducibility are in the source data and [Methods](#)

[Source Data](#).

Extended Data Fig. 8 Fabrication of physiological stiffness gradient gels.

a, Apparent elasticity measurements of shallow (physiological) stiffness gradient gels and uniform stiffness gels. **b**, Gradient of stiffness from in vivo embryo measurements and ex vivo polyacrylamide gels. **c–d**, Immunostaining of fibronectin and fluorescent microspheres in gel. Images represent soft and stiff sides of the same gel exhibiting a stiffness gradient, in either top view (c) or side view (d). Scale bar is 100 μm (c, d). **e–f**, Quantification of fibronectin thickness (e) and mean fluorescence (f). **g**, Tactic index of clusters seeded on different regions of physiological gradient gels. **h–j**, Immunostaining of Rac-GTP of clusters on a stiffness gradient or uniformly high stiffness (h), quantification of total Rac-GTP (i) and Rac-GTP polarity (j) on different portions of the gradient gel. Scale bar is 20 μm (h). Thick bars (a, b, e-g, i, j) represent mean; error bars (a, b, e-g, i, j) represent s.d.; unpaired two-tailed *t*-test (b, e), two-tailed Mann-Whitney *U* test (f), Dunn's test (g, i, j); $n = 8$ (a), 30 (e), 20 (f), 28 (g, 0.6, 1.4), 25 (g, 1), 33 (g, 1.8) gels, 15 (b) clusters and embryos, 29 (i, j) clusters. Statistics and reproducibility are in the source data and [Methods](#)

[Source Data](#).

Extended Data Fig. 9 Synergistic effects of chemotactic and durotactic gradients and comparison of collective versus single cell durotaxis.

a–i, Time-coded projected tracks of example clusters (a, c, e, g), angles of the tracks (b, d, f, h), quantification of directional migration (i). Scale bar is 100 μm (a, c, e, g). **j–m**, Synergy analysis, real combination of chemotaxis and durotaxis is compared with the inferred combination. Clusters exposed to durotactic and chemotactic gradients simultaneously and the inferred addition of track angles (j), speed (k), tactic index (l), and migration distance (m) of clusters exposed to durotactic and chemotactic gradients simultaneously based on the data with either gradient alone. **n**, A time-coded projected track of an individually migrating neural crest cell on a physiological stiffness gradient. Scale bar is 40 μm (n). **o**, A circular histogram showing the angles of the tracks by single cells plated on physiological (shallow) shallow stiffness gradients. **p**, Quantification of tactic index by clusters and single cells on physiological (shallow) stiffness gradients. Box plots (k-m) show the median, box edges represent the 25th and 75th percentiles, and whiskers show spread of data; thick bars (p) represent mean; error bars (p) represent s.d.; resampling test (j-m), two-tailed Mann-Whitney U test (p); * $P \leq 0.05$, ** $P \leq 0.001$, *** $P \leq 0.0001$; $n = 42$ (a, real combination), 600 (a, inferred combination; b, inferred combination), 94 (b, real combination), 16 (c, real combination), 570 (c, inferred combination), 43 (p, cell cluster) clusters, 600 (o) angles, 44 (p, single cell) cells. Statistics and reproducibility are in the source data and [Methods](#)

[Source Data](#).

[Extended Data Fig. 10 Chemotaxis and Durotaxis synergically control actomyosin contraction polarity.](#)

a, Heat maps derived from the change in actomyosin cable length. Actomyosin contraction pulses are cyan/purple rectangles. Note that front contractions are inhibited when clusters are exposed to chemical and mechanical gradients. **b**, Quantification of the frequency of actomyosin contractions at the rear and front of cell clusters in control (purple), durotaxis (lilac), chemotaxis (green) and both (blue). Thick bars represent mean (b); error bars represent s.d. (b); Dunn's test (b); ns, $P > 0.05$, * $P \leq 0.05$, ** $P \leq 0.001$, *** $P \leq 0.0001$; $n = 32$ (b, control rear), 35 (b, control front), 41 (b, durotaxis rear), 26 (b, durotaxis front), 29 (b, chemotaxis rear), 28 (b,

chemotaxis front), 25 (b, both rear), 26 (b, both front) clusters. Statistics and reproducibility are in the source data and [Methods](#)

[Source Data](#).

Supplementary information

[Reporting Summary](#)

[Supplementary Video 1](#)

Stiffness gradient ablation inhibits neural crest migration in vivo. Fluorescently labelled neural crest grafts in a control host embryo (left) and host embryo that is mechanically ablated to abrogate the stiffness gradient (right).

[Supplementary Video 2](#)

Stiffness gradient ablation inhibits directional neural crest migration in vivo. Fluorescently labelled neural crest grafts in a control host embryo (left) and host embryo that is mechanically ablated to abrogate the stiffness gradient (right).

[Supplementary Video 3](#)

Neural crest durotaxis ex vivo. Neural crest clusters expressing fluorescent nuclei (magenta) and membrane (cyan) markers explanted on a gel of uniform (left) and steep gradient (right) stiffness.

Source data

[Source Data Fig. 1](#)

[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. 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

Shellard, A., Mayor, R. Collective durotaxis along a self-generated stiffness gradient in vivo. *Nature* **600**, 690–694 (2021).

<https://doi.org/10.1038/s41586-021-04210-x>

- Received: 09 November 2020
- Accepted: 02 November 2021
- Published: 08 December 2021
- Issue Date: 23 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04210-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/s41586-021-04210-x>

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

- Article
- [Published: 08 December 2021](#)

Unrepresentative big surveys significantly overestimated US vaccine uptake

- [Valerie C. Bradley](#)¹ na1,
- [Shiro Kuriwaki](#) [ORCID: orcid.org/0000-0002-5687-2647](#)² na1,
- [Michael Isakov](#) [ORCID: orcid.org/0000-0001-6468-5908](#)³,
- [Dino Sejdinovic](#)¹,
- [Xiao-Li Meng](#)⁴ &
- [Seth Flaxman](#) ⁵

[Nature](#) volume 600, pages 695–700 (2021)

- 43k Accesses
- 1 Citations
- 1130 Altmetric
- [Metrics details](#)

Subjects

- [Infectious diseases](#)
- [Research data](#)
- [Statistics](#)

Abstract

Surveys are a crucial tool for understanding public opinion and behaviour, and their accuracy depends on maintaining statistical representativeness of their target populations by minimizing biases from all sources. Increasing data size shrinks confidence intervals but magnifies the effect of survey bias: an instance of the Big

Data Paradox¹. Here we demonstrate this paradox in estimates of first-dose COVID-19 vaccine uptake in US adults from 9 January to 19 May 2021 from two large surveys: Delphi–Facebook^{2,3} (about 250,000 responses per week) and Census Household Pulse⁴ (about 75,000 every two weeks). In May 2021, Delphi–Facebook overestimated uptake by 17 percentage points (14–20 percentage points with 5% benchmark imprecision) and Census Household Pulse by 14 (11–17 percentage points with 5% benchmark imprecision), compared to a retroactively updated benchmark the Centers for Disease Control and Prevention published on 26 May 2021. Moreover, their large sample sizes led to minuscule margins of error on the incorrect estimates. By contrast, an Axios–Ipsos online panel⁵ with about 1,000 responses per week following survey research best practices⁶ provided reliable estimates and uncertainty quantification. We decompose observed error using a recent analytic framework¹ to explain the inaccuracy in the three surveys. We then analyse the implications for vaccine hesitancy and willingness. We show how a survey of 250,000 respondents can produce an estimate of the population mean that is no more accurate than an estimate from a simple random sample of size 10. Our central message is that data quality matters more than data quantity, and that compensating the former with the latter is a mathematically provable losing proposition.

[Download PDF](#)

Main

Governments, businesses and researchers rely on survey data to inform the provision of government services⁷, steer business strategy and guide the response to the COVID-19 pandemic^{8,9}. With the ever-increasing volume and accessibility of online surveys and organically collected data, the line between traditional survey research and Big Data is becoming increasingly blurred¹⁰. Large datasets enable the analysis of fine-grained subgroups, which are in high demand for designing targeted policy interventions¹¹. However, counter to common intuition¹², larger sample sizes alone do not ensure lower error. Instead, small biases are compounded as sample size increases¹.

We see initial evidence of this in the discrepancies in estimates of first-dose COVID-19 vaccine uptake, willingness and hesitancy from three online surveys in the US. Two of them—Delphi–Facebook’s COVID-19 symptom tracker^{2,3} (around 250,000 responses per week and with over 4.5 million responses from January to May 2021) and the Census Bureau’s Household Pulse survey⁴ (around 75,000 responses per survey wave and with over 600,000 responses from January to May 2021)—have large enough sample sizes to render standard uncertainty intervals negligible; however, they report significantly different estimates of vaccination behaviour with nearly identically

worded questions (Table 1). For example, Delphi–Facebook’s state-level estimates for willingness to receive a vaccine from the end of March 2021 are 8.5 percentage points lower on average than those from the Census Household Pulse (Extended Data Fig. 1a), with differences as large as 16 percentage points.

Table 1 Comparison of survey designs

The US Centers for Disease Control and Prevention (CDC) compiles and reports vaccine uptake statistics from state and local offices¹³. These figures serve as a rare external benchmark, permitting us to compare survey estimates of vaccine uptake to those from the CDC. The CDC has noted the discrepancies between their own reported vaccine uptake and that of the Census Household Pulse^{14,15}, and we find even larger discrepancies between the CDC and Delphi–Facebook data (Fig. 1a). By contrast, the Axios–Ipsos Coronavirus Tracker⁵ (around 1,000 responses per wave, and over 10,000 responses from January to May 2021) tracks the CDC benchmark well. None of these surveys use the CDC benchmark to adjust or assess their estimates of vaccine uptake, thus by examining patterns in these discrepancies, we can infer each survey’s accuracy and statistical representativeness, a nuanced concept that is critical for the reliability of survey findings^{16,17,18,19}.

Fig 1: Errors in estimates of vaccine uptake.

 figure1

a, Estimates of vaccine uptake for US adults in 2021 compared to CDC benchmark data, plotted by the end date of each survey wave. Points indicate each study’s weighted estimate of first-dose vaccine uptake, and intervals are 95% confidence intervals using reported standard errors and design effects. Delphi–Facebook has $n = 4,525,633$ across 19 waves, Census Household Pulse has $n = 606,615$ across 8 waves and Axios–Ipsos has $n = 11,421$ across 11 waves. Delphi–Facebook’s confidence

intervals are too small to be visible. **b**, Total error $\{(\overline{Y})_n - (\overline{Y})_N\}$. **c**, Data defect correlation $\{\hat{\rho}_{Y,R}\}$. **d**, Data scarcity $\{\sqrt{(N-n)/n}\}$. **e**, Inherent problem difficulty σ_Y . Shaded bands represent scenarios of $\pm 5\%$ (darker) and $\pm 10\%$ (lighter) imprecision in the CDC benchmark relative to reported values (points). **b–e** comprise the decomposition in equation (1).

The Big Data Paradox in vaccine uptake

We focus on the Delphi–Facebook and Census Household Pulse surveys because their large sample sizes (each greater than 10,000 respondents²⁰) present an opportunity to examine the Big Data Paradox¹ in surveys. The Census Household Pulse is an experimental product designed to rapidly measure pandemic-related behaviour. Delphi–Facebook has stated that the intent of their survey is to make comparisons over space, time and subgroups, and that point estimates should be interpreted with caution³. However, despite these intentions, Delphi–Facebook has reported point estimates of vaccine uptake in its own publications^{11,21}.

Delphi–Facebook and Census Household Pulse surveys persistently overestimate vaccine uptake relative to the CDC’s benchmark (Fig. 1a) even taking into account Benchmark Imprecision (Fig. 1b) as explained in ‘Decomposing Error in COVID Surveys’. Despite being the smallest survey by an order of magnitude, the estimates of Axios–Ipsos track well with the CDC rates (Fig. 1a), and their 95% confidence intervals contain the benchmark estimate from the CDC in 10 out of 11 surveys (an empirical coverage probability of 91%).

One might hope that estimates of changes in first-dose vaccine uptake are correct, even if each snapshot is biased. However, errors have increased over time, from just a few percentage points in January 2021 to Axios-Ipsos’ 4.2 percentage points [1–7 percentage points with 5% benchmark imprecision (BI)], Census Household Pulse’s 14 percentage points [5% BI: 11–17] and Delphi-Facebook’s 17 percentage points [5% BI: 14–20] by mid-May 2021 (Fig. 1b). For context, for a state that is near the herd immunity threshold (70–80% based on recent estimates²²), a discrepancy of 10 percentage points in vaccination rates could be the difference between containment and uncontrolled exponential growth in new SARS-CoV-2 infections.

Conventional statistical formulas for uncertainty further mislead when applied to biased big surveys because as sample size increases, bias (rather than variance) dominates estimator error. Figure 1a shows 95% confidence intervals for vaccine uptake based on the reported sampling standard errors and weighting design effects of each survey²³. Axios–Ipsos has the widest confidence intervals, but also the smallest design effects (1.1–1.2), suggesting that its accuracy is driven more by minimizing

bias in data collection rather than post-survey adjustment. The 95% confidence intervals of Census Household Pulse are widened by large design effects (4.4–4.8) but they are still too narrow to include the true rate of vaccine uptake in almost all survey waves. The confidence intervals for Delphi–Facebook are extremely small, driven by large sample size and moderate design effects (1.4–1.5), and give us a negligible chance of being close to the truth.

One benefit of such large surveys might be to compare estimates of spatial and demographic subgroups^{24,25,26}. However, relative to the CDC’s contemporaneously reported state-level estimates, which did not include retroactive corrections, Delphi–Facebook and Census Household Pulse overestimated CDC state-level vaccine uptake by 16 and 9 percentage points, respectively (Extended Data Fig. 1g,h) in March 2021, and by equal or larger amounts by May 2021 (Extended Data Fig. 2g,h). Relative estimates were no better than absolute estimates in March of 2021: there is little agreement in a survey’s estimated state-level rankings with the CDC (a Kendall rank correlation of 0.31 for Delphi–Facebook in Extended Data Fig. 1j and 0.26 for Census Household Pulse in Extended Data Fig. 1j) but they improved in May of 2021 (correlations of 0.78 and 0.74, respectively, in Extended Data Fig. 2i,j). Among 18–64-year-olds, both Delphi–Facebook and Census Household Pulse overestimate uptake, with errors increasing over time (Extended Data Fig. 6).

These examples illustrate a mathematical fact. That is, when biased samples are large, they are doubly misleading: they produce confidence intervals with incorrect centres and substantially underestimated widths. This is the Big Data Paradox: “the bigger the data, the surer we fool ourselves”¹ when we fail to account for bias in data collection.

A framework for quantifying data quality

Although it is well-understood that traditional confidence intervals capture only survey sampling errors²⁷ (and not total error), the traditional survey framework lacks analytic tools for quantifying nonsampling errors separately from sampling errors. A previously formulated statistical framework¹ permits us to exactly decompose the total error of a survey estimate into three components:

$$\begin{array}{c} \text{\rm Total} \\ \text{\rm error} \\ = \\ \text{\rm Data} \\ \times \\ \text{\rm quality} \\ \times \\ \text{\rm defect} \\ \times \\ \text{\rm Data} \\ \times \\ \text{\rm scarcity} \\ \times \\ \text{\rm Inherent} \\ \times \\ \text{\rm problem} \\ \times \\ \text{\rm difficulty} \end{array} \quad (1)$$

This framework has been applied to COVID-19 case counts²⁸ and election forecasting²⁹. Its full application requires ground-truth benchmarks or their estimates

from independent sources¹.

Specifically, the ‘total error’ is the difference between the observed sample mean $\langle \overline{Y} \rangle_n$ as an estimator of the ground truth, the population mean $\langle \overline{Y} \rangle_N$. The ‘data quality defect’ is measured using $\hat{\rho}_{Y,R}$, called the ‘data defect correlation’ (ddc)¹, which quantifies total bias (from any source), measured by the correlation between the event that an individual’s response is recorded and its value, Y . The effect of data quantity is captured by ‘data scarcity’, which is a function of the sample size n and the population size N , measured as $\sqrt{(N-n)/n}$ and hence what matters for error is the relative sample size—that is, how close n is to N —rather than the absolute sample size n . The third factor is the ‘inherent problem difficulty’, which measures the population heterogeneity (via the standard deviation σ_Y of Y), because the more heterogeneous a population is, the harder it is to estimate its average well. Mathematically, equation (1) is given by $\langle \overline{Y} \rangle_n - \langle \overline{Y} \rangle_N = \hat{\rho}_{Y,R} \sqrt{(N-n)/n} \sigma_Y$. This expression was inspired by the Hartley–Ross inequality for biases in ratio estimators³⁰. More details on the decomposition are provided in ‘Calculation and interpretation of ddc’ in the Methods, in which we also present a generalization for weighted estimators.

Decomposing error in COVID surveys

Although the ddc is not directly observed, COVID-19 surveys present a rare case in which it can be deduced because all of the other terms in equation (1) are known (see ‘Calculation and interpretation of ddc’ in the Methods for an in-depth explanation). We apply this framework to the aggregate error shown in Fig. 1b, and the resulting components of error from the right-hand side of equation (1) are shown in Fig. 1c–e.

We use the CDC’s report of the cumulative count of first doses administered to US adults as the benchmark^{8,13}, $\langle \overline{Y} \rangle_N$. This benchmark time series may be affected by administrative delays and slippage in how the CDC centralizes information from states^{31,32,33,34}. The CDC continuously updates their entire time series retroactively for such delays as they are reported. But to account for potentially unreported delays, we present our results with Benchmark Imprecision (BI) in case the CDC’s numbers from our study period, 9 January to 26 May 2021, as reported on 26 May by the CDC suffer from $\pm 5\%$ and $\pm 10\%$ imprecision. These scenarios were chosen on the basis of analysis of the magnitude by which the CDC’s initial estimate for vaccine uptake by a particular day increases as the CDC receives delayed reports of vaccinations that occurred on that day (Extended Data Fig. 3, [Supplementary Information A.2](#)). That said, these scenarios may not capture latent systemic issues that affect CDC vaccination reporting.

The total error of each survey's estimate of vaccine uptake (Fig. 1b) increases over time for all studies, most markedly for Delphi–Facebook. The data quality defect, measured by the ddc, also increases over time for Census Household Pulse and for Delphi–Facebook (Fig. 1c). The ddc for Axios–Ipsos is much smaller and steady over time, consistent with what one would expect from a representative sample. The data scarcity, $\sqrt{(N-n)/n}$, for each survey is roughly constant across time (Fig. 1d). Inherent problem difficulty is a population quantity common to all three surveys that peaks when the benchmark vaccination rate approaches 50% in April 2021 (Fig. 1e). Therefore, the decomposition suggests that the increasing error in estimates of vaccine uptake in Delphi–Facebook and Census Household Pulse is primarily driven by increasing ddc, which captures the overall effect of the bias in coverage, selection and response.

Equation (1) also yields a formula for the bias-adjusted effective sample size n_{eff} , which is the size of a simple random sample that we would expect to exhibit the same level of mean squared error (MSE) as what was actually observed in a given study with a given ddc. Unlike the classical effective sample size²³, this quantity captures the effect of bias as well as that of an increase in variance from weighting and sampling. For details of this calculation, see ‘Error decomposition with survey weights’ in the Methods.

For estimating the US vaccination rate, Delphi–Facebook has a bias-adjusted effective sample size of less than 10 in April 2021, a 99.99% reduction from the raw average weekly sample size of 250,000 (Fig. 2). The Census Household Pulse is also affected by over 99% reductions in effective sample size by May 2021. A simple random sample would have controlled estimation errors by controlling ddc. However, once this control is lost, small increases in ddc beyond what is expected in simple random samples can result in marked reductions of effective sample sizes for large populations¹.

Fig 2: Bias-adjusted effective sample size.

 **figure2**

An estimate's bias-adjusted effective sample size (different from the classic Kish effective sample size) is the size of a simple random sample that would have the same MSE as the observed estimate. Effective sample sizes are shown here on the \log_{10} scale. The original sample size was $n = 4,525,633$ across 19 waves for Delphi–Facebook, $n = 606,615$ across 8 waves for Census Household Pulse and $n = 11,421$ across 11 waves for Axios–Ipsos. Shaded bands represent scenarios of $\pm 5\%$ benchmark imprecision in the CDC benchmark.

Comparing study designs

Understanding why bias occurs in some surveys but not others requires an understanding of the sampling strategy, modes, questionnaire and weighting scheme of each survey. Table 1 compares the design of each survey (more details in ‘Additional survey methodology’ in the Methods, Extended Data Table 1).

All three surveys are conducted online and target the US adult population, but vary in the methods that they use to recruit respondents³⁵. The Delphi–Facebook survey recruits respondents from active Facebook users (the Facebook active user base, or FAUB) using daily unequal-probability stratified random sampling². The Census Bureau uses a systematic random sample to select households from the subset of the master address file (MAF) of the Census for which they have obtained either cell

phone or email contact information (approximately 81% of all households in the MAF)⁴.

In comparison, Axios–Ipsos relies on inverse response propensity sampling from Ipsos’ online KnowledgePanel. Ipsos recruits panellists using an address-based probabilistic sample from USPS’s delivery sequence file (DSF)⁵. The DSF is similar to the MAF of the Census. Unlike the Census Household Pulse, potential respondents are not limited to the subset for whom email and phone contact information is available. Furthermore, Ipsos provides internet access and tablets to recruited panellists who lack home internet access. In 2021, this ‘offline’ group typically comprises 1% of the final survey (Extended Data Table 1).

All three surveys weight on age and gender; that is, assign larger weights to respondents of underrepresented age by gender subgroups and smaller weights to those of overrepresented subgroups^{2,4,5} (Table 1). Axios–Ipsos and Census Household Pulse also weight on education and race and/or ethnicity (hereafter, race/ethnicity). Axios–Ipsos additionally weights to the composition of political partisanship measured by “recent ABC News/Washington Post telephone polls”⁵ in 6 of the 11 waves we study. Education—a known correlate of propensity to respond to surveys³⁶ and social media use³⁷ are notably absent from Delphi–Facebook’s weighting scheme, as is race/ethnicity. As noted before, none of the surveys use the CDC benchmark to adjust or assess estimates of vaccine uptake.

Explanations for error

Table 2 illustrates some consequences of these design choices. Axios–Ipsos samples mimic the actual breakdown of education attainment among US adults even before weighting, whereas those of Census Household Pulse and Delphi–Facebook do not. After weighting, Axios–Ipsos and Census Household Pulse match the population benchmark, by design. Delphi–Facebook does not explicitly weight on education, and hence the education bias persists in their weighted estimates: those without a college degree are underrepresented by nearly 20 percentage points. The case is similar for race/ethnicity. Delphi–Facebook’s weighting scheme does not adjust for race/ethnicity, and hence their weighted sample still overrepresents white adults by 8 percentage points, and underrepresents the proportions of Black and Asian individuals by around 50% of their size in the population (Table 2).

Table 2 Composition of survey respondents by educational attainment and race/ethnicity

The overrepresentation of white adults and people with college degrees explains part of the error of Delphi–Facebook. The racial groups that Delphi–Facebook

underrepresents tend to be more willing and less vaccinated in the samples (Table 2). In other words, reweighting the Delphi–Facebook survey to upweight racial minorities will bring willingness estimates closer to Household Pulse and the vaccination rate closer to CDC. The three surveys also report that people without a four-year college degree are less likely to have been vaccinated compared to those with a degree (Table 2, Supplementary Table 1). If we assume that vaccination behaviours do not differ systematically between non-respondents and respondents within each demographic category, underrepresentation of less-vaccinated groups would contribute to the bias found here. However, this alone cannot explain the discrepancies in all the outcomes. Census Household Pulse weights on both race and education⁴ and still overestimates vaccine uptake by over ten points in late May of 2021 (Fig. 1b).

Delphi–Facebook and Census Household Pulse may be unrepresentative with respect to political partisanship, which has been found to be correlated with vaccine behaviour³⁸ and with survey response³⁹, and thus may contribute to observed bias. However, neither Delphi–Facebook nor Census Household Pulse collects partisanship of respondents. US Census agencies cannot ask about political preference, and no unequivocal population benchmark for partisanship in the general adult population exists.

Rurality may also contribute to the errors, because it correlates with vaccine status⁸ and home internet access⁴⁰. Neither Census Household Pulse nor Delphi–Facebook weights on sub-state geography, which may mean that adults in more rural areas who are less likely to be vaccinated are also underrepresented in the two surveys, leading to overestimation of vaccine uptake.

Axios–Ipsos weights to metropolitan status and also recruits a fraction of its panellists from an ‘offline’ population of individuals without internet access. We find that dropping these offline respondents ($n = 21$, or 1% of the sample) in their 22 March 2021 wave increases Axios–Ipsos’ overall estimate of the vaccination rate by 0.5 percentage points, thereby increasing the total error (Extended Data Table 2). However, this offline population is too small to explain the entirety of the difference in accuracy between Axios–Ipsos and either Census Household Pulse (6 percentage points) or Delphi–Facebook (14 percentage points), in this time period.

Careful recruitment of panellists is at least as important as weighting. Weighting on observed covariates alone cannot explain or correct the discrepancies we observe. For example, reweighting Axios–Ipsos 22 March 2021 wave using only Delphi–Facebook’s weighting variables (age group and gender) increased the error in their vaccination estimates by 1 percentage point, but this estimate with Axios–Ipsos data is still more accurate than that from Delphi–Facebook during the same period (Extended Data Table 2). The Axios–Ipsos estimate with Delphi–Facebook weighting

overestimated vaccination by 2 percentage points, whereas Delphi–Facebook overestimated it by 11 percentage points.

The key implication is that there is no silver bullet: every small part of panel recruitment, sampling and weighting matters for controlling the data quality measured as the correlation between an outcome and response—what we call the ddc. In multi-stage sampling, which includes for example the selection of potential respondents followed by non-response, bias in even a single step can substantially affect the final result ('Population size in multi-stage sampling' in the [Methods](#), Extended Data Table 3). A total quality control approach, inspired by the total survey error framework⁴¹, is a better strategy than trying to prioritize some components over others to improve data quality. This emphasis is a reaffirmation of the best practice for survey research as advocated by the American Association for Public Opinion Research:⁶ "The quality of a survey is best judged not by its size, scope, or prominence, but by how much attention is given to [preventing, measuring and] dealing with the many important problems that can arise."⁴²

Addressing common misperceptions

The three surveys discussed in this article demonstrate a seemingly paradoxical phenomenon—the two larger surveys that we studied are more statistically confident, but also more biased, than the smaller, more traditional Axios–Ipsos poll. These findings are paradoxical only when we fall into the trap of the intuition that estimation errors necessarily decrease in larger datasets¹².

A limitation of our vaccine uptake analysis is that we only examine ddc with respect to an outcome for which a benchmark is available: first-dose vaccine uptake. One might hope that surveys biased on vaccine uptake are not biased on other outcomes, for which there may not be benchmarks to reveal their biases. However, the absence of evidence of bias for the remaining outcomes is not evidence of its absence. In fact, mathematically, when a survey is found to be biased with respect to one variable, it implies that the entire survey fails to be statistically representative. The theory of survey sampling relies on statistical representativeness for all variables achieved through probabilistic sampling⁴³. Indeed, Neyman's original introduction of probabilistic sampling showed the limits of purposive sampling, which attempted to achieve overall representativeness by enforcing it only on a set of variables^{18,44}.

In other words, when a survey loses its overall statistical representativeness (for example, through bias in coverage or non-response), which is difficult to repair (for example, by weighting or modelling on observable characteristics) and almost impossible to verify⁴⁵, researchers who wish to use the survey for scientific studies must supply other reasons to justify the reliability of their survey estimates, such as

evidence about the independence between the variable of interest and the factors that are responsible for the unrepresentativeness. Furthermore, scientific journals that publish studies based on surveys that may be unrepresentative¹⁷—especially those with large sizes such as Delphi–Facebook (biased with respect to vaccination status (Fig. 1), race and education (Table 2))—need to ask for reasonable effort from the authors to address the unrepresentativeness.

Some may argue that bias is a necessary trade-off for having data that are sufficiently large for conducting highly granular analysis, such as county-level estimation of vaccine hesitancy²⁶. Although high-resolution inference is important, we warn that this is a double-edged argument. A highly biased estimate with a misleadingly small confidence interval can do more damage than having no estimate at all. We further note that bias is not limited to population point estimates, but also affects estimates of changes over time (contrary to published guidance³). Both Delphi–Facebook and Census Household Pulse significantly overestimate the slope of vaccine uptake relative to that of the CDC benchmark (Fig. 1b).

The accuracy of our analysis does rely on the accuracy of the CDC’s estimates of COVID vaccine uptake. However, if the selection bias in the CDC’s benchmark is significant enough to alter our results, then that itself would be another example of the Big Data Paradox.

Discussion

This is not the first time that the Big Data Paradox has appeared: Google Trends predicted more than twice the number of influenza-like illnesses than the CDC in February 2013⁴⁶. This analysis demonstrates that the Big Data Paradox applies not only to organically collected Big Data, like Google Trends, but also to surveys. Delphi–Facebook is “the largest public health survey ever conducted in the United States”⁴⁷. The Census Household Pulse is conducted in collaboration between the US Census Bureau and eleven statistical government partners, all with enormous resources and survey expertise. Both studies take steps to mitigate selection bias, but substantially overestimate vaccine uptake. As we have shown, the effect of bias is magnified as relative sample size increases.

By contrast, Axios–Ipsos records only about 1,000 responses per wave, but makes additional efforts to prevent selection bias. Small surveys can be just as wrong as large surveys in expectation—of the three other small-to-medium online surveys additionally analysed, two also miss the CDC vaccination benchmark (Extended Data Fig. 5). The overall lesson is that investing in data quality (particularly during collection, but also in analysis) minimizes error more efficiently than does increasing data quantity. Of course, a sample size of 1,000 may be too small (that is, leading to

unhelpfully large uncertainty intervals) for the kind of 50-state analyses made possible by big surveys. However, small-area methods that borrow information across subgroups⁴⁸ can perform better with higher-quality—albeit few—data, and whether that approach would outperform the large, biased surveys is an open question.

There are approaches to correct for these biases in both probability and nonprobability samples alike. For COVID-19 surveys in particular, since June 2021, the AP–NORC multimode panel has weighted their COVID-19 related surveys to the CDC benchmark, so that the weighted ddc for vaccine uptake is zero by design⁴⁹. More generally, there is an extensive literature on approaches for making inferences from data collected from nonprobability samples^{50,51,52}. Other promising approaches include integrating surveys of varying quality^{53,54}, and leveraging the estimated ddc in one outcome to correct bias in others under several scenarios ([Supplementary Information D](#)).

Although more needs to be done to fully examine the nuances of large surveys, organically collected administrative datasets and social media data, we hope that this comparative study of ddc highlights the concerning implications of the Big Data Paradox—how large sample sizes magnify the effect of seemingly small defects in data collection, which leads to overconfidence in incorrect inferences.

Methods

Calculation and interpretation of ddc

The mathematical expression for equation (1) is given here for completeness:

$$\text{Equation (2)}: \text{The first factor } \hat{\rho}_{Y,R} = \frac{\overline{Y}_n - \overline{Y}_N}{\sqrt{\frac{N-n}{n}} \sigma_Y}$$

The first factor $\hat{\rho}_{Y,R}$ is called the data defect correlation (ddc)¹. It is a measure of data quality represented by the correlation between the recording indicator R ($R = 1$ if an answer is recorded and $R = 0$ otherwise) and its value, Y . Given a benchmark, the ddc $\hat{\rho}_{Y,R}$ can be calculated by substituting known quantities into equation (2). In the case of a single survey wave of a COVID-19 survey, n is the sample size of the survey wave, N is the population size of US adults from US Census estimates⁵⁵, \overline{Y}_n is the survey estimate of vaccine uptake and \overline{Y}_N is the estimate of vaccine uptake for the corresponding period taken from the CDC's report of the cumulative count of first doses administered to US adults^{8,13}. We calculate σ_Y

$\}_{\{Y\}}=\sqrt{\overline{Y}(1-\overline{Y})}$ because Y is binary (but equation (2) is not restricted to binary Y).

We calculate $\hat{\rho}_{Y,R}$ by using total error $\sqrt{\overline{Y}(1-\overline{Y})}$, which captures not only selection bias but also any measurement bias (for example, from question wording). However, with this calculation method, $\hat{\rho}_{Y,R}$ lacks the direct interpretation as a correlation between Y and R , and instead becomes a more general index of data quality directly related to classical design effects (see ‘Bias-adjusted effective sample size’).

It is important to point out that the increase in ddc does not necessarily imply that the response mechanisms for Delphi–Facebook and Census Household Pulse have changed over time. The correlation between a changing outcome and a steady response mechanism could change over time, hence changing the value of ddc. For example, as more individuals become vaccinated, and vaccination status is driven by individual behaviour rather than eligibility, the correlation between vaccination status and propensity to respond could increase even if the propensity to respond for a given individual is constant. This would lead to large values of ddc over time, reflecting the increased impact of the same response mechanism.

Error decomposition with survey weights

The data quality framework given by equations (1) and (2) is a special case of a more general framework for assessing the actual error of a weighted estimator $\hat{\rho}_{Y,R} = \frac{\sum_i w_i R_i Y_i}{\sum_i w_i}$, where w_i is the survey weight assigned to individual i . It is shown in Meng¹ that

$$\text{Equation 3: } \hat{\rho}_{Y,R} = \frac{\sum_i w_i R_i Y_i}{\sqrt{\sum_i w_i} \sqrt{\sum_i w_i (R_i - \bar{R})^2}},$$

where $\text{Corr}(Y, R) = \frac{\sum_i w_i R_i Y_i}{\sqrt{\sum_i w_i} \sqrt{\sum_i w_i (R_i - \bar{R})^2}}$ is the finite population correlation between Y_i and R_i (over $i = 1, \dots, N$). The ‘hat’ on ρ reminds us that this correlation depends on the specific realization of R_i . The term n_w is the classical ‘effective sample size’ due to weighting²³, that is, $n_w = \frac{n}{1 + \frac{\sum_i w_i (R_i - \bar{R})^2}{\sum_i w_i}}$, where CV_w is the coefficient of variation of the weights for all individuals in the observed sample, that is, the standard deviation of weights normalized by their mean. It is common for surveys to rescale their weights to

have mean 1, in which case $\left(\{\rm CV\}\right)_w^2$ is simply the sample variance of W .

When all weights are the same, equation (3) reduces to equation (2). In other words, the ddc term $\hat{\rho}_{Y,R} \left(\{\rm w\}\right)$ now also takes into account the effect of the weights as a means to combat the selection bias represented by the recording indicator R . Intuitively, if $\hat{\rho}_{Y,R} = \text{Corr}(Y, R)$ is high (in magnitude), then some Y_i 's have a higher chance of entering our dataset than others, thus leading to a sample average that is a biased estimator for the population average. Incorporating appropriate weights can reduce $\hat{\rho}_{Y,R}$ to $\hat{\rho}_{Y,R} \left(\{\rm w\}\right)$, with the aim of reducing the effect of the selection bias. However, this reduction alone may not be sufficient to improve the accuracy of \bar{Y}_w because the use of weight necessarily reduces the sampling fraction $f = n_w/N$ to $f \left(\{\rm w\}\right) = n_w/N$ as well, as $n_w < n$. Equation (3) precisely describes this trade-off, providing a formula to assess when the reduction of ddc is significant to outweigh the reduction of the effective sample size.

Measuring the correlation between Y and R is not a new idea in survey statistics (though note that ddc is the population correlation between Y and R , not the sample correlation), nor is the observation that as sample size increases, error is dominated by bias instead of variance^{56,57}. The new insight is that ddc is a general metric to index the lack of representativeness of the data we observe, regardless of whether or not the sample is obtained through a probabilistic scheme, or weighted to mimic a probabilistic sample. As discussed in ‘Addressing common misperceptions’ in the main text, any single ddc deviating from what is expected under representative sampling (for example, probabilistic sampling) is sufficient to establish that the sample is not representative (but the converse is not true). Furthermore, the ddc framework refutes the common belief that increasing sample size necessarily improves statistical estimation^{1,58}.

Bias-adjusted effective sample size

By matching the mean-squared error of \bar{Y}_w with the variance of the sample average from simple random sampling, Meng¹ derives the following formula for calculating a bias-adjusted effective sample size, or n_{eff} :

$$\begin{array}{l} \$\$ \begin{array}{r} n_{\text{eff}} = \frac{n \left(\{\rm w\}\right)}{N - n \left(\{\rm w\}\right)} \times \frac{1}{E[\hat{\rho}_Y]_R^2} \\ \end{array} \$\$ \end{array}$$

Given an estimator \overline{Y}_w with expected total MSE T due to data defect, sampling variability and weighting, this quantity n_{eff} represents the size of a simple random sample such that its mean \bar{Y}_N , as an estimator for the same population mean \overline{Y}_N , would have the identical MSE T . The term $E[\hat{\rho}_Y^2]$ represents the amount of selection bias (squared) expected on average from a particular recording mechanism R and a chosen weighting scheme.

For each survey wave, we use $\hat{\rho}_Y^2$ to approximate $E[\hat{\rho}_Y^2]$. This estimation is unbiased by design, as we use an estimator to estimate its expectation. Therefore, the only source of error is the sampling variation, which is typically negligible for large surveys such as Delphi–Facebook and the Census Household Pulse. This estimation error may have more impact for smaller surveys such as the Axios–Ipsos survey, an issue that we will investigate in subsequent work.

We compute $\hat{\rho}_Y^2$ by using the benchmark \overline{Y}_N , namely, by solving equation (3) for $\hat{\rho}_Y^2$,

$$\hat{\rho}_Y^2 = \frac{\text{var}(\text{where}_Z)}{\text{var}(\text{where}_Z)} = \frac{\text{var}(\overline{Y}_N - \overline{Y}_w)}{\text{var}(\overline{Y}_N)} = \frac{\text{var}(\frac{1-f}{n_w} \text{rm}(w))}{\text{var}(\text{rm}(w))} = \sigma_Y^2$$

We introduce this notation Z_w because it is the quantity that determines the well-known survey efficiency measure, the so-called ‘design effect’, which is the variance of Z_w for a probabilistic sampling design²³ (when we assume the weights are fixed). For the more general setting in which \overline{Y}_w may be biased, we replace the variance by MSE, and hence the bias-adjusted design effect $D_e = E[\text{var}(Z_w)]$, which is the MSE relative to the benchmark measured in the unit of the variance of an average from a simple random sample of size n_w . Hence $D_I \equiv E[\hat{\rho}_Y^2]$, which was termed as ‘data defect index’¹, is simply the bias-adjusted design effect per unit, because $D_I = D_e / n_w$.

Furthermore, because Z_w is the standardized actual error, it captures any kind of error inherited in \overline{Y}_w . This observation is important because when Y is subject to measurement errors, $\frac{\text{var}(Z_w)}{\text{var}(N)}$ no longer has the simple interpretation as a correlation. But because we estimate D_I by $\frac{\text{var}(Z_w)}{n_w}$ directly, our effective sample size calculation is still valid even when equation (3) does not hold.

Asymptotic behaviour of ddc

As shown in Meng¹, for any probabilistic sample without selection biases, the ddc is on the order of $\frac{1}{\sqrt{N}}$. Hence the magnitude of $(\hat{\rho}_{Y,R})$ (or $(\hat{\rho}_{Y,R})$) is small enough to cancel out the effect of $\sqrt{N-n}$ (or $\sqrt{N-n}$) in the data scarcity term on the actual error, as seen in equation (2) (or equation (3)). However, when a sample is unrepresentative; for example, when those with $Y=1$ are more likely to enter the dataset than those with $Y=0$, then $(\hat{\rho}_{Y,R})$ can far exceed $\frac{1}{\sqrt{N}}$ in magnitude. In this case, error will increase with \sqrt{N} for a fixed ddc and growing population size N (equation (2)). This result may be counterintuitive in the traditional survey statistics framework, which often considers how error changes as sample size n grows. The ddc framework considers a more general set-up, taking into account individual response behaviour, including its effect on sample size itself.

As an example of how response behaviour can shape both total error and the number of respondents n , suppose individual response behaviour is captured by a logistic regression model

$$\text{logit}(\Pr(R=1|Y)) = \alpha + \beta Y \quad (4)$$

This is a model for a response propensity score. Its value is determined by α , which drives the overall sampling fraction ($f=\frac{n}{N}$), and by β , which controls how strongly Y influences whether a participant will respond or not.

In this logit response model, when $(\beta \neq 0)$, $(\hat{\rho}_{Y,R})$ is determined by individual behaviour, not by population size N . In [Supplementary Information B.1](#), we prove that ddc cannot vanish as N grows, nor can the observed sample size n ever approach 0 or N for a given set of (finite and plausible) values of $\{\alpha, \beta\}$, because there will always be a non-trivial percentage of non-respondents. For example, an f of 0.01 can be obtained under this model for either $\alpha = -0.46$, $\beta = 0$ (no influence of individual behaviour on response propensity), or for $\alpha = -3.9$, $\beta = -4.84$. However, despite the same f , the implied ddc and consequently the MSE will differ. For example, the MSE for the former (no correlation with Y) is 0.0004, whereas the MSE for the latter (a -4.84 coefficient on Y) is 0.242, over 600 times larger.

See [Supplementary Information B.2](#) for the connection between ddc and a well-studied non-response model from econometrics, the Heckman selection model⁵⁹.

Population size in multi-stage sampling

We have shown that the asymptotic behaviour of error depends on whether the data collection process is driven by individual response behaviour or by survey design. The reality is often a mix of both. Consequently, the relevant ‘population size’ N depends on when and where the representativeness of the sample is destroyed; that is, when the individual response behaviours come into play. Real-world surveys that are as complex as the three surveys we analyse here have multiple stages of sample selection.

Extended Data Table 3 takes as an example the sampling stages of the Census Household Pulse, which has the most extensive set of documentation among the three surveys we analyse. As we have summarized (Table 1, Extended Data Table 1), the Census Household Pulse (1) first defines the sampling frame as the reachable subset of the MAF, (2) takes a random sample of that population to prompt (send a survey questionnaire) and (3) waits for individuals to respond to that survey. Each of these stages reduces the desired data size, and the corresponding population size is the intended sample size from the prior stage (in notation, $N_s = n_{s-1}$, for $s = 2, 3$). For example, in stage 3, the population size N_3 is the size of the intended sample size n_2 from the second stage (random sample of the outreach list), because only the sampled individuals have a chance to respond.

Although all stages contribute to the overall ddc, the stage that dominates is the first stage at which the representativeness of our sample is destroyed—the size of which will be labelled as the dominating population size (dps)—when the relevant population size decreases markedly at each step. However, we must bear in mind that dps refers to the worst-case scenario, when biases accumulate, instead of (accidentally) cancelling each other out.

For example, if the 20% of the MAFs excluded from the Census Household Pulse sampling frame (because they had no cell phone or email contact information) is not representative of the US adult population, then the dps is N_1 , or 255 million adults contained in 144 million households. Then the increase in bias for given ddc is driven by the rate of $\sqrt{N_1}$ where $N_1 = 2.55 \times 10^8$ and is large indeed ($\sqrt{2.5 \times 10^8} \approx 15,000$). By contrast, if the sampling frame is representative of the target population and the outreach list is representative of the frame (and hence representative of the US adult population) but there is non-response bias, then dps is $N_3 = 10^6$ and the impact of ddc is amplified by the square root of that number ($\sqrt{10^6} = 1,000$). By contrast, Axios–Ipsos reports a response rate of about 50%, and obtains a sample of $n = 1,000$, so the dps could be as small as $N_3 = 2,000$ (with $\sqrt{2,000} \approx 45$).

This decomposition is why our comparison of the surveys is consistent with the ‘Law of Large Populations’¹ (estimation error increases with \sqrt{N}), even though all three surveys ultimately target the same US adult population. Given our existing knowledge about online–offline populations⁴⁰ and our analysis of Axios–Ipsos’ small ‘offline’ population, Census Household Pulse may suffer from unrepresentativeness at Stage 1 of Extended Data Table 3, where $N=255$ million, and Delphi–Facebook may suffer from unrepresentativeness at the initial stage of starting from the Facebook user base. By contrast, the main source of unrepresentativeness for Axios–Ipsos may be at a later stage at which the relevant population size is orders of magnitude smaller.

CDC estimates of vaccination rates

Our analysis of the nationwide vaccination rate covers the period between 9 January 2021 and 19 May 2021. We used CDC’s vaccination statistics published on their data tracker as of 26 May 2021. This dataset is a time series of counts of 1st dose vaccinations for every day in our time period, reported for all ages and disaggregated by age group.

This CDC time series obtained on 26 May 2021 included retroactive updates to dates covering our entire study period, as does each daily update provided by the CDC daily update. For example, the CDC benchmark we use for March 2021 is not only the vaccination counts originally reported in March but also includes the delayed reporting for March that the CDC became aware of by 26 May 2021. Analyzing several snapshots before 26 May 2021, we find that these retroactive updates 40 days out could change the initial estimate by about 5% (Extended Data Fig. 3), hence informing our sensitivity analysis of $\pm 5\%$ and 10% benchmark imprecision.

To match the sampling frame of the surveys we analyze, US adults 18 years and older, we must restrict the CDC vaccination counts to those administered to those adults. However, because of the different way states and jurisdiction report their vaccination statistics, the CDC did not possess age-coded counts for some jurisdictions, such as Texas, at the time of our study. The number of vaccinations with missing age data reached about 10 percent of the total US vaccinations at its peak at the time of our study. We therefore assume that the day by day fraction of adults among individuals for whom age is reported as missing is equal to the fraction of adults among individuals with age reported. Because minors became eligible for vaccinations only towards the end of our study period, the fraction of adults in data reporting age never falls below 97%.

Additional survey methodology

The Census Household Pulse and Delphi–Facebook surveys are the first of their kind for each organization, whereas Ipsos has maintained their online panel for 12 years.

Question wording

All three surveys ask whether respondents have received a COVID-19 vaccine (Extended Data Table 1). Delphi–Facebook and Census Household Pulse ask similar questions (“Have you had/received a COVID-19 vaccination/vaccine?”). Axios–Ipsos asks “Do you personally know anyone who has already received the COVID-19 vaccine?”, and respondents are given response options including “Yes, I have received the vaccine.” The Axios–Ipsos question wording might pressure respondents to conform to their communities’ modal behaviour and thus misreport their true vaccination status, or may induce acquiescence bias from the multiple ‘yes’ options presented⁶⁰. This pressure may exist both in high- and low-vaccination communities, so its net effect on Axios–Ipsos’ results is unclear. Nonetheless, Axios–Ipsos’ question wording does differ from that of the other two surveys, and may contribute to the observed differences in estimates of vaccine uptake across surveys.

Population of interest

All three surveys target the US adult population, but with different sampling and weighting schemes. Census Household Pulse sets the denominator of their percentages as the household civilian, non-institutionalized population in the United States of 18 years of age or older, excluding Puerto Rico or the island areas. Axios–Ipsos designs samples to be representative of the US general adult population of 18 or older. For Delphi–Facebook, the US target population reported in weekly contingency tables is the US adult population, excluding Puerto Rico and other US territories. For the CDC Benchmark, we define the denominator as the US 18+ population, excluding Puerto Rico and other US territories. To estimate the size of the total US population, we use the US Census Bureau Annual Estimates of the Resident Population for the United States and Puerto Rico, 2019⁵⁵. This is also what the CDC uses as the denominator in calculating rates and percentages of the US population⁶⁰.

Axios–Ipsos and Delphi–Facebook generate target distributions of the US adult population using the Current Population Survey (CPS), March Supplement, from 2019 and 2018, respectively. Census Household Pulse uses a combination of 2018 1-year American Community Survey (ACS) estimates and the Census Bureau’s Population Estimates Program (PEP) from July 2020. Both the CPS and ACS are well-established large surveys by the Census and the choice between them is largely inconsequential.

Axios–Ipsos data

The Axios–Ipsos Coronavirus tracker is an ongoing, bi-weekly tracker intended to measure attitudes towards COVID-19 of adults in the US. The tracker has been running since 13 March 2020 and has released results from 45 waves as of 28 May 2021. Each wave generally runs over a period of 4 days. The Axios–Ipsos data used in this analysis were scraped from the topline PDF reports released on the Ipsos website⁵. The PDF reports also contain Ipsos’ design effects, which we have confirmed are calculated as 1 plus the variance of the (scaled) weights.

Census Household Pulse data

The Census Household Pulse is an experimental product of the US Census Bureau in collaboration with eleven other federal statistical agencies. We use the point estimates presented in Data Tables, as well as the standard errors calculated by the Census Bureau using replicate weights. The design effects are not reported, however we can calculate it as $\sqrt{1 + \text{CV}_w^2}$, where CV_w is the coefficient of variation of the individual-level weights included in the microdata²³.

Delphi–Facebook COVID symptom survey

The Delphi–Facebook COVID symptom survey is an ongoing survey collaboration between Facebook, the Delphi Group at Carnegie Mellon University (CMU), and the University of Maryland². The survey is intended to track COVID-like symptoms over time in the US and in over 200 countries. We use only the US data in this analysis. The study recruits respondents using daily stratified random samples recruiting a cross-section of Facebook active users. New respondents are obtained each day, and aggregates are reported publicly on weekly and monthly frequencies. The Delphi–Facebook data used here were downloaded directly from CMU’s repository for weekly contingency tables with point estimates and standard errors.

Ethical compliance

According to HRA decision tools (<http://www.hra-decasiontools.org.uk/research/>), our study is considered Research, and according to the NHS REC review tool (<http://www.hra-decasiontools.org.uk/ethics/>), we do not need NHS Research Ethics Committee (REC) review, as we used only (1) publicly available, (2) anonymized and (3) aggregated data outside of clinical settings.

Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

Data availability

Raw data have been deposited in the Harvard Dataverse, at <https://doi.org/10.7910/DVN/GKBUUK>. Data were collected from publicly available repositories of survey data by downloading the data directly or using APIs.

Code availability

Code to replicate the findings is available in the repository <https://github.com/vcbradley/ddc-vaccine-US>. The main decomposition of the ddc is available on the package ‘ddi’ from the Comprehensive R Archive Network (CRAN).

References

1. 1.

Meng, X.-L. Statistical paradises and paradoxes in big data (I): Law of large populations, big data paradox, and the 2016 US presidential election. *Ann. Appl. Stat.* **12**, 685–726 (2018).

2. 2.

Barkay, N. et al. Weights and methodology brief for the COVID-19 Symptom Survey by University of Maryland and Carnegie Mellon University, in partnership with Facebook. Preprint at <https://arxiv.org/abs/2009.14675> (2020).

3. 3.

Kreuter, F. et al. Partnering with Facebook on a university-based rapid turn-around global survey. *Surv. Res. Methods* **14**, 159–163 (2020).

4. 4.

Fields, J. F. et al. *Design and Operation of the 2020 Household Pulse Survey* (U.S. Census Bureau, 2020).

5. 5.

Jackson, C., Newall, M. & Yi, J. Axios Ipsos Coronavirus Index (2021).

6. 6.

American Association for Public Opinion Research (AAPOR). Best practices for survey research. <https://www.aapor.org/Standards-Ethics/Best-Practices.aspx> (2021).

7. 7.

Hastak, M., Mazis, M. B. & Morris, L. A. The role of consumer surveys in public policy decision making. *J. Public Policy Mark.* **20**, 170–185 (2001).

8. 8.

B. P. Murthy, et al. Disparities in COVID-19 vaccination coverage between urban and rural counties: United States, December 14, 2020–April 10, 2021. *MMWR Morb. Mortal. Wkly Rep.* **70**, 759–764 (2021).

9. 9.

Arrieta, A., Gakidou, E., Larson, H., Mullany, E. & Troeger, C. Through understanding and empathy, we can convince women to get the COVID-19 vaccine. *Think Global Health* <https://www.thinkglobalhealth.org/article/through-understanding-and-empathy-we-can-convince-women-get-covid-19-vaccine> (2021).

10. 10.

Japec, L. et al. *AAPOR Report on Big Data* (American Association of Public Opinion Researchers, 2015).

11. 11.

Reinhart, A., Kim, E., Garcia, A. & LaRocca, S. Using the COVID-19 Symptom Survey to track vaccination uptake and sentiment in the United States. *CMU Delphi Group* <https://delphi.cmu.edu/blog/2021/01/28/using-the-covid-19-symptom-survey-to-track-vaccination-uptake-and-sentiment-in-the-united-states> (2021).

12. 12.

Mayer-Schönberger, V. & Cukier, K. *Big Data: A Revolution That Will Transform How We Live, Work, and Think* (Houghton Mifflin Harcourt, 2013).

13. 13.

CDC. Trends in number of COVID-19 vaccinations (2021).

14. 14.

Nguyen, K. H. et al. Comparison of COVID-19 vaccination coverage estimates from the Household Pulse Survey, Omnibus Panel Surveys, and COVID-19 vaccine administration data, United States, March 2021. *CDC AdultVaxView* <https://www.cdc.gov/vaccines/imz-managers/coverage/adultvaxview/pubs-resources/covid19-coverage-estimates-comparison.html> (2021).

15. 15.

Santibanez, T. A. et al. Sociodemographic factors associated with receipt of COVID-19 vaccination and intent to definitely get vaccinated, adults aged 18 years or above—Household Pulse Survey, United States, April 28–May 10, 2021. *CDC AdultVaxView* <https://www.cdc.gov/vaccines/imz-managers/coverage/adultvaxview/pubs-resources/sociodemographic-factors-covid19-vaccination.html> (2021).

16. 16.

Kruskal, W. & Mosteller, F. Representative sampling, I: Non-scientific literature. *Int. Stat. Rev.* **47**, 13–24 (1979).

17. 17.

Kruskal, W. & Mosteller, F. Representative sampling, II: Scientific literature, excluding statistics. *Int. Stat. Rev.* **47**, 111–127 (1979).

18. 18.

Kruskal, W. & Mosteller, F. Representative sampling, III: The current statistical literature. *Int. Stat. Rev.* **47**, 245–265 (1979).

19. 19.

Kruskal, W. & Mosteller, F. Representative sampling, IV: The history of the concept in statistics, 1895–1939. *Int. Stat. Rev.* **48**, 169–195 (1980).

20. 20.

American Association for Public Opinion Research (AAPOR). Margin of sampling error/credibility interval. <https://www.aapor.org/Education-Resources/Election-Polling-Resources/Margin-of-Sampling-Error-Credibility-Interval.aspx> (2021).

21. 21.

The Delphi Group at Carnegie Mellon University in partnership with Facebook. *Topline Report on COVID-19 Vaccination in the United States* (2021).

22. 22.

Haas, E. J. et al. Impact and effectiveness of mRNA BNT162b2 vaccine against SARS-CoV-2 infections and COVID-19 cases, hospitalisations, and deaths following a nationwide vaccination campaign in Israel: an observational study using national surveillance data. *Lancet* **397**, 1819–1829 (2021).

23. 23.

Kish, L. *Survey Sampling* (Wiley, 1965).

24. 24.

Institute for Health Metrics and Evaluation (IHME). COVID-19 vaccine hesitancy. <https://vaccine-hesitancy.healthdata.org/> (2021).

25. 25.

King, W. C., Rubinstein, M., Reinhart, A. & Mejia, R. J. Time trends and factors related to COVID-19 vaccine hesitancy from january-may 2021 among US adults: findings from a large-scale national survey. Preprint at <https://doi.org/10.1101/2021.07.20.21260795> (2021).

26. 26.

CDC. Estimates of vaccine hesitancy for COVID-19 (2021).

27. 27.

Groves, R. M. et al. *Survey Methodology* Vol. 561 (Wiley, 2011).

28. 28.

Dempsey, W. The hypothesis of testing: paradoxes arising out of reported coronavirus case-counts. Preprint at <https://arxiv.org/abs/2005.10425> (2020).

29. 29.

Isakov, M. & Kuriwaki, S. Towards principled unskewing: viewing 2020 election polls through a corrective lens from 2016. *Harvard Data Science Review* **2** <https://doi.org/10.1162/99608f92.86a46f38> (2020).

30. 30.

Hartley, H. O. & Ross, A. Unbiased ratio estimators. *Nature* **174**, 270–271 (1954).

31. 31.

Tiu, A., Susswein, Z., Merritt, A. & Bansal, S. Characterizing the spatiotemporal heterogeneity of the COVID-19 vaccination landscape. Preprint at <https://doi.org/10.1101/2021.10.04.21263345> (2021).

32. 32.

Groen, J. Sources of error in survey and administrative data: the importance of reporting procedures. *J. Off. Stat.* **28**, 173–198 (2012).

33. 33.

Tu, X. M., Meng, X.-L. & Pagano, M. The AIDS epidemic: estimating survival after AIDS diagnosis from surveillance data. *J. Am. Stat. Assoc.* **88**, 26–36 (1993).

34. 34.

Barnes, O. & Burn-Murdoch, J. COVID response hampered by population data glitches. *Financial Times* (11 October 2021).

35. 35.

Kennedy, C. et al. Evaluating online nonprobability surveys. *Pew Research Center* <https://www.pewresearch.org/methods/2016/05/02/evaluating-online-nonprobability-surveys/> (2016).

36. 36.

Kennedy, C. et al. An evaluation of the 2016 election polls in the United States. *Public Opin. Q.* **82**, 1–33 (2018).

37. 37.

Auxier, B. & Anderson, M. Social media use in 2021. *Pew Research Center* <https://www.pewresearch.org/internet/2021/04/07/social-media-use-in-2021/> (2021).

38. 38.

Gadarian, S. K., Goodman, S. W. & Pepinsky, T. B. Partisanship, health behavior, and policy attitudes in the early stages of the COVID-19 pandemic. *PLoS ONE* **16**, e0249596 (2021).

39. 39.

Mercer, A., Lau, A. & Kennedy, C. For weighting online opt-in samples, what matters most? *Pew Research Center* <https://www.pewresearch.org/methods/2018/01/26/for-weighting-online-opt-in-samples-what-matters-most/> (2018).

40. 40.

Ryan, C. *Computer and Internet Use in the United States: 2016*. American Community Survey Report No. ACS-39 (U.S. Census Bureau, 2017).

41. 41.

Biemer, P. P. & Lyberg, L. E. *Introduction to Survey Quality* (Wiley, 2003).

42. 42.

Scheuren, F. *What is a Survey?* (American Statistical Association, 2004).

43. 43.

Sukhatme, P. V. *Sampling Theory of Surveys with Applications* (1954).

44. 44.

Neyman, J. On the two different aspects of the representative method: the method of stratified sampling and the method of purposive selection. *J. R. Stat. Soc.* **97**, 558–625 (1934).

45. 45.

Groves, R. M. Nonresponse rates and nonresponse bias in household surveys. *Public Opin. Q.* **70**, 646–675 (2006).

46. 46.

Lazer, D., Kennedy, R., King, G. & Vespignani, A. The parable of Google Flu: traps in big data analysis. *Science* **343**, 1203–1205 (2014).

47. 47.

Salomon, J. A. et al. The US COVID-19 Trends and Impact Survey, 2020-2021: continuous real-time measurement of COVID-19 symptoms, risks, protective behaviors, testing and vaccination. Preprint at <https://doi.org/10.1101/2021.07.24.21261076> (2021).

48. 48.

Park, D. K., Gelman, A. & Bafumi, J. Bayesian multilevel estimation with poststratification: State-level estimates from national polls. *Polit. Anal.* **12**, 375–385 (2004).

49. 49.

Associated Press–NORC Center for Public Affairs Research. The June 2021 AP–NORC center poll (2021).

50. 50.

Wang, W., Rothschild, D., Goel, S. & Gelman, A. Forecasting elections with non-representative polls. *Int. J. Forecast.* **31**, 980–991 (2015).

51. 51.

Elliott, M. R. & Valliant, R. Inference for nonprobability samples. *Stat. Sci.* **32**, 249–264 (2017).

52. 52.

Little, R. J., West, B. T., Boonstra, P. S. & Hu, J. Measures of the degree of departure from ignorable sample selection. *J. Surv. Stat. Methodol.* **8**, 932–964 (2020).

53. 53.

Wiśniowski, A., Sakshaug, J. W., Perez Ruiz, D. A. & Blom, A. G. Integrating probability and nonprobability samples for survey inference. *J. Surv. Stat. Methodol.* **8**, 120–147 (2020).

54. 54.

Yang, S., Kim, J. K. & Song, R. Doubly robust inference when combining probability and non-probability samples with high dimensional data. *J. R. Stat. Soc. B* **82**, 445–465 (2020).

55. 55.

U.S. Census Bureau. Methodology for the United States population estimates: Vintage 2019.

56. 56.

Bethlehem, J. in *Survey Nonresponse* (eds Groves, R. M. et al.) 275–288 (Wiley, 2002).

57. 57.

Meng, X.-L. in *Past, Present, and Future of Statistical Science* (eds Lin, X. et al.) 537–562 (CRC Press, 2014).

58. 58.

Meng, X.-L. & Xie, X. I got more data, my model is more refined, but my estimator is getting worse! Am I just dumb? *Econom. Rev.* **33**, 218–250 (2014).

59. 59.

Heckman, J. J. Sample selection bias as a specification error. *Econometrica* 153–161 (1979).

60. 60.

CDC. Reporting COVID-19 vaccination demographic data.
<https://www.cdc.gov/coronavirus/2019-ncov/vaccines/distributing/demographics-vaccination-data.html> (2021).

Acknowledgements

We thank F. Kreuter, A. Reinhart and the Delphi Group at CMU; Facebook's Demography and Survey Science group; F. Barlas, C. Jackson, C. Morris, M. Newall and the Public Affairs team at Ipsos; and J. Fields and J. Hunter Childs at the US Census Bureau for productive conversations about their surveys. We further thank the

Delphi Group at CMU for their help in computing weekly design effects for their survey; the Ipsos team for providing data on their ‘offline’ respondents; the CDC for responding to our questions; S. Paddock, other participants at the JPSM 2021 lecture (delivered by X.-L.M.) and S. Finch for providing comments; A. Edwards-Levy for inspiring our interest in this topic; and R. Born for suggesting more intuitive terms used in equation (1). V.C.B. is funded by the University of Oxford’s Clarendon Fund and the EPSRC and MRC through the OxWaSP CDT programme (EP/L016710/1). X.-L.M. acknowledges partial financial support by NSF. S.F. acknowledges the support of the EPSRC (EP/V002910/1).

Author information

Author notes

1. These authors contributed equally: Valerie C. Bradley, Shiro Kuriwaki

Affiliations

1. Department of Statistics, University of Oxford, Oxford, UK

Valerie C. Bradley & Dino Sejdinovic

2. Department of Political Science, Stanford University, Stanford, CA, USA

Shiro Kuriwaki

3. Harvard College, Harvard University, Cambridge, MA, USA

Michael Isakov

4. Department of Statistics, Harvard University, Cambridge, MA, USA

Xiao-Li Meng

5. Department of Computer Science, University of Oxford, Oxford, UK

Seth Flaxman

Contributions

V.C.B. and S.F. conceived and formulated the research questions. V.C.B. and S.K. contributed equally to data analysis, writing and visualization. X.-L.M. conceived and

formulated the methodology. All authors contributed to methodology, writing, visualization, editing and data analysis. S.F. supervised the work.

Corresponding author

Correspondence to [Seth Flaxman](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Michael Elliot, Alex Reinhart and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

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 Comparisons of state-level vaccine uptake, hesitancy and willingness across surveys and the CDC for March 2021.](#)

Comparison of Delphi-Facebook and Census Household Pulse's state-level point estimates (**a–c**) and rankings (**d–f**) for vaccine hesitancy, willingness and uptake. Dotted black lines show agreement and red points show the average of 50 states. During our study period, the CDC published daily reports of the cumulative number of vaccinations by state that had occurred up to a certain date. Due to reporting delays, these may be an underestimate, but retroactively updated data was not available to us. **g–j** compare state-level point estimates and rankings for the same survey waves to CDC benchmark estimates from 31 March 2021. The Delphi–Facebook data are from the week ending 27 March 2021 and the Census Household Pulse is the wave ending 29 March 2021. See Extended Data Fig. [3](#) for details on the degree of retroactive updates we could expect, and [Supplementary Information A.2](#) for details.

Extended Data Fig. 2 Comparisons of state-level vaccine uptake, hesitancy and willingness across surveys and the CDC for May 2021.

Comparison of Delphi-Facebook and Census Household Pulse's state-level point estimates (**a–c**) and rankings (**d–f**) for vaccine hesitancy, willingness and uptake. Dotted black lines show agreement and red points show the average of 50 states. During our study period, the CDC published daily reports of the cumulative number of vaccinations by state that had occurred up to a certain date. Due to reporting delays, these may be an underestimate, but retroactively updated data was not available to us. **g–j** compare state-level point estimates and rankings for the same survey waves to CDC benchmark estimates from 15 May 2021. The Delphi–Facebook data are from the wave week ending 8 May 2021 and the Census Household Pulse is the wave ending 10 May 2021. See Extended Data Fig. [3](#) for details on the degree of retroactive updates we could expect, and [Supplementary Information A.2](#) for details.

Extended Data Fig. 3 Retroactive adjustment of CDC vaccine uptake figures for 3–12 April 2021, over the 45 days from 12 April.

Increase is shown as a percentage of the vaccine uptake reported on 12 April. Most of the retroactive increases in reported estimates appear to occur in the first 10 days after an estimate is first reported. By about 40 days after the initial estimates for a particular day are reported, the upward adjustment plateaus at around 5–6% of the initial estimate. We use this analysis to guide the choice of 5% and 10% threshold for the possible imprecision in the CDC benchmark when computing Benchmark Imprecision (BI) intervals.

Extended Data Fig. 4 Revised estimates of hesitancy and willingness after accounting for survey errors for vaccination uptake.

The grey point shows the reported value at the last point of the time series. Each line shows a different scenario for what might be driving the error in uptake estimate, derived using hypothetical ddc values for willingness and hesitancy based on the observed ddc value for uptake. Access scenario: willingness suffers from at least as much, if not more, bias than uptake. Hesitancy scenario: hesitancy suffers from at least as much, if not more, bias than uptake. Uptake scenario: the error is split roughly equally between hesitancy and willingness. See [Supplementary Information D](#) for more details.

Extended Data Fig. 5 Vaccination rates compared with CDC benchmark for four online polls.

Ribbons indicate traditional 95% confidence intervals, which are twice the standard error reported by the poll. Grey line is the CDC benchmark. Data for Progress asks “As of today, have you been vaccinated for Covid-19?”; Morning Consult asks “Have you gotten the vaccine, or not?”; Harris Poll asks “Which of the following best describes your mindset when it comes to getting the COVID-19 vaccine when it becomes available to you?”. YouGov surveys are not analysed because they explicitly examined how their surveys tracked CDC vaccine uptake. See [Supplementary Information C.3](#) for the sampling methodology of each survey and discussion of differences.

[Extended Data Fig. 6 Survey error by age group \(18–64-year-olds, and those aged 65 and over\).](#)

a, Estimates of vaccine uptake from Delphi–Facebook (blue) and Census Household Pulse (green) for 18–64-year-olds (left) and those aged 65 or older (right). Bounds on the CDC’s estimate of vaccine uptake for those groups are shown in grey. The CDC receives vaccination-by-age data only from some jurisdictions. We do know, however, the total number of vaccinations in the US. Therefore, we calculate the bounds by allocating all the vaccine doses for which age is unknown to either 18–64 or 65+. **b**, Unweighted ddc for each Delphi–Facebook and Census Household Pulse calculated for the 18–64 group using the bounds on the CDC’s estimates of uptake. ddc for 65+ is not shown due to large uncertainty in the bounded CDC estimates of uptake.

Extended Data Table 1 Methodologies of Axios–Ipsos, Census Household Pulse and Delphi–Facebook studies

Extended Data Table 2 Contribution of offline recruitment and weighting schemes to discrepancies between surveys

Extended Data Table 3 Example of multi-stage population selection

Supplementary information

[Supplementary Information](#)

This file contains a Supplementary Discussion and Supplementary Table 1.

[Reporting Summary](#)

[Peer Review File](#)

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Bradley, V.C., Kuriwaki, S., Isakov, M. *et al.* Unrepresentative big surveys significantly overestimated US vaccine uptake. *Nature* **600**, 695–700 (2021). <https://doi.org/10.1038/s41586-021-04198-4>

- Received: 18 June 2021
- Accepted: 29 October 2021
- Published: 08 December 2021
- Issue Date: 23 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04198-4>

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

Further reading

- [What surveys really say](#)

- Frauke Kreuter

Nature (2021)

[What surveys really say](#)

- Frauke Kreuter

News & Views 08 Dec 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04198-4>

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

- Article
- [Published: 21 October 2021](#)

Immunogenicity and efficacy of heterologous ChAdOx1–BNT162b2 vaccination

- [Bruno Pozzetto](#)^{1,2 na1},
- [Vincent Legros](#) [ORCID: orcid.org/0000-0002-6118-7564](#)^{1,3 na1},
- [Sophia Djebali](#) [ORCID: orcid.org/0000-0002-0567-5387](#)^{1 na1},
- [Véronique Barateau](#)^{1 na1},
- [Nicolas Guibert](#) [ORCID: orcid.org/0000-0002-1192-8301](#)⁴,
- [Marine Villard](#) [ORCID: orcid.org/0000-0002-3905-8993](#)¹,
- [Loïc Peyrot](#)¹,
- [Omran Allatif](#) [ORCID: orcid.org/0000-0002-2539-972X](#)¹,
- [Jean-Baptiste Fassier](#)⁴,
- [Amélie Massardier-Pilonchéry](#)⁴,
- [Karen Brengel-Pesce](#)⁵,
- [Melyssa Yaugel-Novoa](#)¹,
- [Solène Denolly](#) [ORCID: orcid.org/0000-0003-1739-0079](#)¹,
- [Bertrand Boson](#) [ORCID: orcid.org/0000-0003-1920-5172](#)¹,
- [Thomas Bourlet](#)¹,
- [Antonin Bal](#)^{1,6},
- [Martine Valette](#)⁶,
- [Thibault Andrieu](#)⁷,
- [Bruno Lina](#)^{1,6},
- [Covid-Ser study group](#),
- [François-Loïc Cosset](#) [ORCID: orcid.org/0000-0001-8842-3726](#)^{1 na2},
- [Stéphane Paul](#) [ORCID: orcid.org/0000-0002-8830-4273](#)^{1 na2},
- [Thierry Defrance](#) [ORCID: orcid.org/0000-0002-6943-358X](#)^{1 na2},
- [Jacqueline Marvel](#) [ORCID: orcid.org/0000-0001-6241-459X](#)^{1 na2},
- [Thierry Walzer](#) [ORCID: orcid.org/0000-0002-0857-8179](#)^{1 na2} &
- [Sophie Trouillet-Assant](#) ^{1,5 na2}

Nature volume 600, pages 701–706 (2021)

- 35k Accesses
- 7 Citations
- 1005 Altmetric
- [Metrics details](#)

Subjects

- [RNA vaccines](#)
- [SARS-CoV-2](#)
- [Viral infection](#)

Abstract

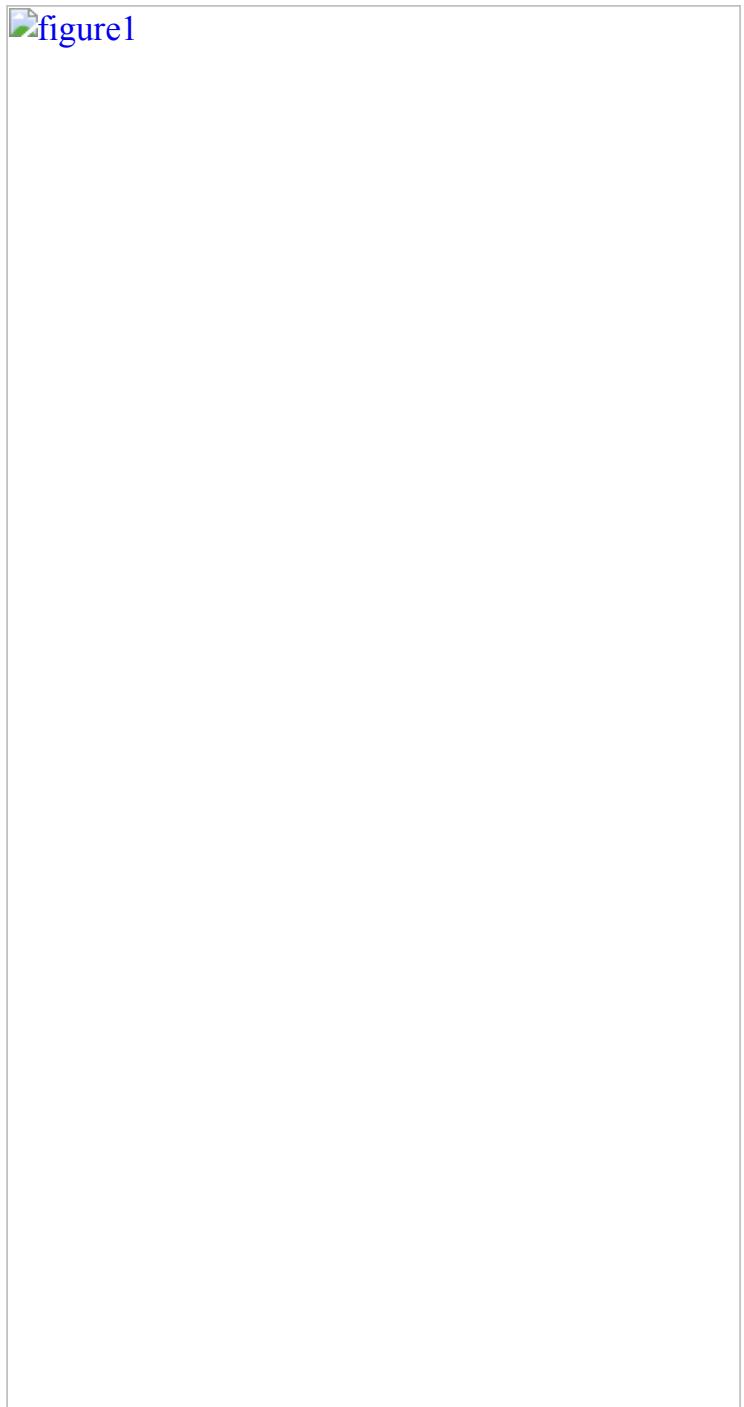
Following severe adverse reactions to the AstraZeneca ChAdOx1-S-nCoV-19 vaccine^{1,2}, European health authorities recommended that patients under the age of 55 years who received one dose of ChAdOx1-S-nCoV-19 receive a second dose of the Pfizer BNT162b2 vaccine as a booster. However, the effectiveness and the immunogenicity of this vaccination regimen have not been formally tested. Here we show that the heterologous ChAdOx1-S-nCoV-19 and BNT162b2 combination confers better protection against severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection than the homologous BNT162b2 and BNT162b2 combination in a real-world observational study of healthcare workers ($n = 13,121$). To understand the underlying mechanism, we conducted a longitudinal survey of the anti-spike immunity conferred by each vaccine combination. Both combinations induced strong anti-spike antibody responses, but sera from heterologous vaccinated individuals displayed a stronger neutralizing activity regardless of the SARS-CoV-2 variant. This enhanced neutralizing potential correlated with increased frequencies of switched and activated memory B cells that recognize the SARS-CoV-2 receptor binding domain. The ChAdOx1-S-nCoV-19 vaccine induced a weaker IgG response but a stronger T cell response than the BNT162b2 vaccine after the priming dose, which could explain the complementarity of both vaccines when used in combination. The heterologous vaccination regimen could therefore be particularly suitable for immunocompromised individuals.

[Download PDF](#)

Main

The ChAdOx1-S-nCoV-19 and BNT162b2 vaccines (hereafter referred to as ChAd and BNT, respectively) have been shown to confer strong protection against SARS-CoV-2 infection and to induce robust immune responses directed against the spike protein of SARS-CoV-2 when administered in a homologous setting^{3,4,5}. However, the effectiveness of a ChAd–BNT combination as a prime–boost regimen has not been formally tested in vaccine trials. To compare the risk of SARS-CoV-2 infection following heterologous ChAd–BNT or homologous BNT–BNT vaccination schedules, we extracted data from the occupational medicine database of the University Hospital of Lyon (Hospices Civils de Lyon), France. Specifically, we analysed the number of individuals who received each vaccination regimen (which started in January 2021 in both groups) and the number of SARS-CoV-2 infections (documented by a positive PCR with reverse transcription (RT–PCR) result) that occurred at least 2 weeks after the booster dose in each group (Table 1). Ten infections were identified out of 2,512 individuals (0.40%) in the heterologous vaccination group compared with 81 infections out of 10,609 individuals (0.76%) in the homologous vaccination group (Fig. 1). As the age of the individual and the vaccination regimen are known covariates, we used a multiple logistic regression model to model the probability of being infected. Our predictors were age (set as a continuous variable, considered as a possible confounder) and the vaccination regimen. Results showed that age is not associated with infection status ($P = 0.4514$). In particular, patients aged 60 years or above (715 out of 13,121) were not more infected (3 out of 715; 0.42%) than patients aged 60 years or below (88 out of 12,406 (0.71%); two-sided Fisher's exact test, $P = 0.4890$). However, the vaccination regimen was significantly associated with the probability of being infected after vaccination ($P = 0.0384$). Individuals vaccinated with BNT–BNT were twice as likely to be infected than those vaccinated with ChAd–BNT (relative risk of 2.03).

Table 1 Infection rate in vaccinated HCWs at the Lyon University Hospital
Fig. 1: Incidence of SARS-CoV-2 infection after different vaccination regimens.



 **figure1**

Histograms show the infection rate (as documented by a positive SARS-CoV-2 RT-PCR result) among groups of HCWs who were vaccinated with the homologous BNT-BNT combination ($n = 81$ out of 10,609) within the recommended 4-week timeframe between the two doses or with the BNT boost after receiving the first ChAd dose ($n = 10$ out of 2,512) approximately 12 weeks before, as recorded by the service of occupational medicine, Hospices Civils de Lyon. Data show the infection rates that occurred 14 days after receiving booster up to the end of recording (15 August 2021).

Statistical significance was calculated using a logistic regression model adjusted for age. Demographic data and other statistics are available in Table 1.

[Source data](#).

Antibodies and virus neutralization

To understand the immunological basis of this difference, we assembled a cohort of healthcare workers (HCWs) without comorbidity (Covid-Ser-Vac). Half of this HCW cohort received the homologous BNT–BNT combination within the recommended 4-week interval between the two doses, whereas the second half received the BNT booster after receiving the first ChAd dose approximately 12 weeks before. For both groups, the mean age was 41 years and 34 years, respectively, and about 70% of participants were female (Table 2). Three blood samples were drawn: before vaccination, before the second dose and 4 weeks after the second dose. Immunological analyses were performed longitudinally (Extended Data Fig. 1), and a linear regression model was used to compare immune parameters between groups to adjust for age differences.

Table 2 Clinical characteristics of patients in the Covid-Ser-Vac study

As previously shown^{6,7}, the BNT vaccine induced a stronger level of SARS-CoV-2 receptor binding domain (RBD) or spike S1-specific IgG than the ChAd vaccine after the first dose. However, these IgG titres were no longer statistically different between groups after the second dose (Fig. 2a,b). Moreover, the positivity rate was 100% for both heterologous and homologous vaccinations. Serum IgA levels measured after the booster vaccination tended to be greater in the homologous than in the heterologous setting (median (interquartile range (IQR)) of 46.7 ng m^{-1} equivalent (Eq) (36.30–78.70 ng m^{-1} ng Eq) versus 37.4 ng ml^{-1} Eq (25.40–59.80 ng ml^{-1} Eq), respectively; equivalence was to the standard used ([Methods](#))) (Fig. 2c).

Fig. 2: Heterologous vaccination induces higher neutralizing antibody titres.

 **figure2**

a–c, Sera from ChAd–BNT ($n = 29$) or BNT–BNT ($n = 31$) individuals were assayed for S1-specific IgG (a), RBD-specific IgG (b) or S1-specific IgA (c) levels using commercial or custom-made ELISA tests at different times during the vaccination process as indicated (a, b) or 4 weeks after full vaccination (c). In a–c, concentrations are expressed in binding antibody units per ml (BAU ml⁻¹) or ng ml⁻¹ Eq of immunoglobulin as indicated. Each serum sample was evaluated as a single measurement (a, b) or in triplicate (c). Dotted lines in a and b indicate positive detection according to the manufacturer. **d–f**, Sera from ChAd–BNT ($n = 29$) or BNT–BNT ($n = 31$) vaccinated individuals were assayed in triplicate for their capacity to neutralize the entry of virus-like particles pseudotyped with the Wuhan strain SARS-CoV-2 envelope (d) or in duplicate for their capacity to neutralize infection of Vero E6 cells by different SARS-CoV-2 strains (e, f), as indicated. Data show the per cent of neutralization relative to a positive control (d) or the 50% plaque reduction neutralization test (PRNT₅₀) (e), and are expressed as dot plots, with one dot corresponding to one individual. The limit of detection is shown as a dotted line in

e, FC, fold change in the mean of the indicated groups. In all panels, box-and-whiskers plots are shown (see [Methods](#) for details), and the median is represented by the magenta middle line. A linear regression model was used to compare values between groups, and this model was corrected for age. Exact *P* values are shown for the indicated comparisons when significant or nearly significant. **f**, Comparison of serum neutralizing activity against the reference lineage 19A and against the variants of concern for each group. *P* values are shown and calculated using the linear regression model described in **e**.

[Source data](#).

We then measured the ability of serum samples to neutralize SARS-CoV-2 spike-pseudotyped viral particles⁸. None of the sera displayed a detectable neutralizing activity before vaccination or after the first dose. However, most of the sera displayed a neutralizing activity after the second dose, and those from the ChAd–BNT vaccinated individuals displayed a higher neutralizing efficacy than those from the BNT–BNT vaccinated individuals (median (IQR) of 62% (34–93%) versus 99% (89–100%) of neutralization, respectively; *P*<0.001) (Fig. [2d](#)).

To validate these results, we tested the ability of the sera to prevent Vero E6 cell infection by SARS-CoV-2 isolates belonging to various clades, including 19A (B38 lineage), Alpha (B.1.1.7 lineage), Beta (B.1.351 lineage), Gamma (P1) and Delta (B.1.617.2 lineage)⁹. Each SARS-CoV-2 isolate was sequenced and confirmed to contain the characteristic mutations of its viral clade. The ChAd–BNT vaccinated individuals exhibited 2.3-fold to 3.6-fold higher serum neutralizing antibody titres against the different variants than BNT–BNT vaccinated individuals (Fig. [2e](#)). Moreover, in the latter group, the neutralizing responses observed against the Beta, Gamma and Delta variants were significantly reduced compared with that against the 19A clade. By contrast, the neutralizing activity detected in sera from ChAd–BNT vaccinees was similarly high against all strains except the Beta clade (Fig. [2e,f](#)). These data demonstrate that the neutralizing potential of antibodies generated by heterologous vaccination is less affected by spike mutations appearing in variants of concern than that of antibodies induced by homologous vaccination.

Vaccine-induced memory cells

We then studied the memory B cell (mBC) compartments using fluorescently labelled tetrameric RBDs to track RBD-specific mBCs and to analyse their phenotype (the gating analysis is presented in Extended Data Fig. [2a](#)). As shown in Fig. [3a](#) and Extended Data Fig. [2b](#), which correspond to concatenated flow cytometry plots, the frequency of RBD-binding mBCs was comparable for both vaccination groups after the prime dose but was on average twofold higher for the ChAd–BNT vaccination

schedule after the booster dose. Moreover, switched ($\text{IgD}^- \text{CD27}^+$) mBCs accounted for about 67% and 47% of the RBD-specific mBC pool in the ChAd–BNT group and the BNT–BNT group, respectively (Fig. 3b,c, Extended Data Fig. 2c). The expression pattern of IgM and IgG isotypes was next used to compare the immunoglobulin isotype distribution among the switched mBCs. The results showed that the homologous vaccination schedule promoted more IgG-switched mBCs (62% versus 48%) than the heterologous vaccination schedule (Fig. 3d, Extended Data Fig. 2d). Finally, to address the activation status of mBCs, we analysed their expression of CD21 and CD11c. It is generally considered that the concurrent loss of CD21 and the upregulation of CD11c are phenotypic features associated with B cell activation independent of the developmental stage¹⁰. Resting mBCs were therefore defined as $\text{CD21}^+ \text{CD11c}^-$ cells, whereas $\text{CD21}^+ \text{CD11c}^+$, $\text{CD21}^- \text{CD11c}^+$ and $\text{CD21}^- \text{CD11c}^-$ cells were globally considered as activated mBCs. The frequencies of activated RBD-specific mBCs were significantly enhanced in the ChAd–BNT group compared with the BNT–BNT group (Fig. 3e, Extended Data Fig. 2e). In particular, the proportions of the $\text{CD21}^- \text{CD11c}^+$ subset were almost three times higher in the ChAd–BNT group. Altogether, our results indicate that compared with the homologous vaccination schedule, the ChAd–BNT combination induces significantly more RBD-specific B cells, increases the frequencies of post-switch mBCs and induces a more active mBC generation process.

Fig. 3: Analysis of RBD-specific memory B cells and spike-specific memory T cells in vaccinated individuals.

 **figure3**

a–e, Flow cytometry analyses of spike-specific mBCs. **a**, Frequencies of SARS-CoV-2 RBD-binding mBCs among the pool of polyclonal mBCs before vaccination, after prime and after boost injections. **b, c**, Proportions of unswitched (**b**) and switched (**c**) RBD-binding mBCs in per cent of RBD-binding mBCs. **d**, Pattern of immunoglobulin isotypes expressed by switched mBCs. **e**, Proportions of resting ($CD21^+CD11c^-$) and activated ($CD21^-CD11c^+$, $CD21^-CD11c^-$ and $CD21^+CD11c^+$) mBCs among the entire pool of RBD-binding mBCs. **f–h**, T cell responses. **f**, IFN γ release after whole-blood

stimulation with RBD peptides. IFN γ was measured by ELISA in the supernatant. **g**, **h**, Flow cytometry analyses of the percentage of cells positive for intracellular IFN γ expression among CD4 (**g**) or CD8 (**h**) T cells after PBMC stimulation with commercial spike peptides. In all panels, box-and-whiskers plots (see [Methods](#) for details) of $n = 29$ and $n = 31$ participants in heterologous and homologous vaccination groups, respectively, are shown. The median value is shown in each group as a purple line. A high cellular mortality rate at the thawing step led to the exclusion of a few samples in each panel; each individual dot corresponds to one participant, and each measurement was performed once. In all panels, a linear regression model was used to compare values between groups, and this model was corrected for age. Exact P values are shown for the indicated comparisons when significant or nearly significant.

[Source data.](#)

We then longitudinally monitored the T cell response of vaccinees against RBD peptides using a whole-blood interferon- γ (IFN γ) release assay. The T cell response was higher after ChAd immunization than after BNT priming and was similarly increased by the BNT booster dose in both groups (median of 0.33 UI ml $^{-1}$ versus 0.43 UI ml $^{-1}$) (Fig. [3f](#)). To confirm and consolidate these results, we measured the spike-specific CD4 and CD8 T cell response at the single-cell level by flow cytometry after stimulating peripheral blood mononuclear cells (PBMCs) from vaccinees with a pool of commercial peptides spanning the entire spike protein and then staining for intracellular IFN γ (the gating strategy is presented in Extended Data Fig. [3](#)). Both vaccination regimens were able to induce a progressive increase in the frequency of S-reactive CD4 and CD8 T cells from the pre-vaccine to the post-booster phase (Fig. [3g](#), [h](#)). However, the heterologous combination resulted in a stronger CD4 T cell response, both after priming and boosting, and in an increasing trend in the CD8 T cell response after the booster dose.

Finally, to understand interrelations between immune parameters, we performed a matrix analysis (Extended Data Fig. [4](#)). This confirmed that the neutralizing antibody activity correlated with the titres of spike-specific IgG, regardless of the variant analysed, and not correlated with IgA levels. Moreover, the neutralization activity correlated with the percentage of switched RBD $^+$ mBCs and with the percentage of activated RBD $^+$ mBCs. This provides further support that better neutralizing antibodies are produced as a result of sustained B cell activation.

Discussion

Heterologous prime–boost vaccinations have been reported to be more immunogenic than homologous ones in experimental settings^{[11](#)}. This has also been proven for some human vaccines^{[12,13](#)}. Moreover, studies using mice have demonstrated the strong

immunogenicity of the ChAd–BNT combination¹⁴. Here we report that heterologous ChAd–BNT vaccination confers better protection against infection—which is associated with more switched mBCs and higher virus neutralizing antibody titres—irrespective of the variant analysed. This finding is of particular importance considering the global increase in the SARS-CoV-2 Delta variant¹⁵. Other studies reported better neutralizing potential of sera from ChAd–BNT vaccinated individuals compared with BNT–BNT vaccinated individuals, but these studies only used surrogate^{16,17} or pseudovirus¹⁸ neutralization assays. Moreover, the ability of such assays to predict neutralizing activity against authentic clinical virus isolates is still debated¹⁹.

Heterologous vaccination induced antibodies with a stronger neutralization potential than homologous vaccination, but the spike-specific IgG antibody titres were comparable. This result suggests that it is the quality of the antibodies that underlies the neutralization potential of the antibody response induced by heterologous vaccination. Assuming that the secondary antibody response is primarily derived from the mBC clones generated by primo-immunization, we can postulate that ChAd and BNT formulations elicit different mBC compartments. mBCs generated by the ChAd formulation could, for example, bear antigen receptors displaying a larger breadth of epitope recognition or a better fit for the SARS-CoV-2 spike protein. This could be linked to the different conformations of the spike protein, as that of the BNT vaccine bears a mutation stabilizing the protein in its pre-fusion conformation²⁰. Our data showed that mBCs still exhibit features of activation 4 weeks after the booster dose in the heterologous regimen, which suggests that the post-boost process of mBC differentiation is more active in this schedule. This could result from a more efficient or long-lasting germinal centre reaction that facilitates an extended cycle of somatic hypermutations and possibly output of B cell clones with higher affinity antigen receptors. An enlarged pool of CD4⁺ T helper cells could contribute to an amplified or prolonged germinal centre reaction. In keeping with this notion, we report here a higher frequency of spike-specific CD4 T cells after priming with the ChAd vaccine, which has been observed in other studies^{6,7}. Moreover, stronger T cell responses may contribute to better protection independently of B cell responses²¹.

The better neutralization potential of sera from ChAd–BNT vaccinated individuals could also be linked to their relatively lower level of S1-specific IgA levels. Indeed, plasma IgA monomers specific to SARS-CoV-2 proteins are twofold less potent than IgG equivalents²². Furthermore, IgA levels persist better than IgG after infection, which coincides with a reduction in the serum neutralizing potential²³. Following a possible competition between serum IgA and IgG, the decrease in IgA levels in heterologous vaccination conditions may lead to better neutralization effects. This phenomenon has already been observed in the context of the RV144 vaccination trial against HIV-1, in which the level of envelope-specific IgA correlated with the risk of

infection²⁴. In this context, IgA was also found to inhibit antibody-dependent cell cytotoxicity by competing with IgG²⁵.

Our data do not allow the discarding of the possibility that the later timing of administration of the booster dose contributed to the high efficacy of the heterologous regimen. Still, it is noteworthy that the size of the RBD⁺ mBC pool after the prime dose was comparable for both groups of vaccinees, despite the disparity in the analysis time point after the prime vaccination (4 versus 12 weeks). This observation indicates that the extended interval between prime and boost in the ChAd–BNT cohort has not been translated into a greater size of the RBD-specific mBC compartment. It has been recently demonstrated for both the ChAd vaccine²⁶ and the BNT vaccine²⁷ that long injection time intervals (12 weeks or more) provide higher binding and neutralizing antibody titres than shorter intervals (less than 6 weeks). For a comparable long vaccination interval (more than 9 weeks), the heterologous ChAd–BNT vaccination schedule induced better humoral immunity (titres and neutralization) than the homologous ChAd–ChAd combination²⁸. To our knowledge, there is only one study in which ChAd–BNT and BNT–BNT vaccination schedules were compared with the same prime–boost interval (4 weeks)⁶. That study⁶ concluded that for short intervals, the humoral responses induced by both regimens were comparable. However, a longer interval (that is, 12 weeks) between prime and boost could be needed to allow the synergy between heterologous vaccines.

Together, we present a real-world observational study of HCWs showing that the heterologous ChAd–BNT vaccination regimen confers stronger protective immunity than the homologous BNT–BNT prime–boost schedule. As hidden confounding factors might be present in our study (for example, different levels of exposure to the virus), confirmatory studies and a longer follow-up of vaccinated participants are warranted. However, our data suggest that the heterologous combination could be particularly suitable for immunocompromised individuals.

Methods

Ethical statement and cohort description

The use and analysis of data from the occupational health medical file were authorized after a regulatory declaration to the National Commission for Information Technology and Civil Liberties according to the reference methodology (declaration MR004 number 20-121 of 30 April 2020). The declaration of SARS-CoV-2 infection is compulsory for all staff to obtain daily allowances without loss of salary during the imposed quarantine.

For the Covid-Ser-Vac study, clinical data were recorded by a trained clinical research associate using Clinsight software (v.Csonline 7.5.720.1). Blood samples were processed and stored at the Centre de Ressource Biologique Neurobiotec, 69500 Bron. Sixty naive HCWs for COVID-19 and vaccinated with BNT and/or ChAd vaccines were included in a prospective longitudinal cohort study conducted at the Hospices Civils de Lyon. Blood sampling was performed before vaccination, before the second dose of vaccine and 4 weeks after the end of vaccination schedule. The absence of previous SARS-CoV-2 infection was confirmed using the Wantai SARS-CoV-2 Ab total assay in the pre-vaccine sample. Demographic characteristics and delays between doses are depicted in Table 2 and Extended Data Fig. 1. Written informed consent was obtained from all participants. Ethics approval was obtained from the national review board for biomedical research in April 2020 (Comité de Protection des Personnes Sud Méditerranée I, Marseille, France; ID RCB 2020-A00932-37), and the study was registered at ClinicalTrials.gov ([NCT04341142](#)).

Measurement of IgG titres

Sera were immediately stored at -80°C after blood sampling. S1-specific IgG and RBD-specific IgG were measured using Siemens Atellica IM SARS-CoV-2 IgG (sCOVG) and bioMérieux Vidas SARS-CoV-2 IgG diagnosis kits, respectively, according to the manufacturers' recommendations. For standardization of these assays to the first World Health Organization international standard, the concentrations were transformed into binding antibody units per ml (BAU ml $^{-1}$) using the conversion factors provided by the manufacturers.

Measurement of spike-specific IgA

ELISAs to evaluate IgA binding to the SARS-CoV-2 spike protein were performed as previously described²⁹. High-binding 96-half-well plates (2310M, Nunc) were coated with 100 μl per well of a spike protein solution (1 $\mu\text{g ml}^{-1}$; 40591-V08H spike S1-RBD, Sino Biologicals) in PBS overnight at 4°C . Plates were washed with washing buffer containing 1× PBS with 0.05% Tween 20 (Sigma-Aldrich) and incubated with 170 μl of blocking buffer per well containing 1× PBS with 3% fat milk powder and 0.05% Tween 20 (Sigma-Aldrich) for 1 h at room temperature. Immediately after blocking, recombinant anti-RBD IgA (B Cell Design, IB3C4 PV) or serum samples diluted in PBS were added and incubated for 1 h at 37°C . Plasma samples were assayed at a 1:100 starting dilution and 7 additional 3-fold serial dilutions.

Recombinant human anti-RBD IgA was used to perform a calibration curve starting at 1.5 $\mu\text{g ml}^{-1}$. Plates were washed and then incubated with anti-human IgA (A0295, Sigma-Aldrich) secondary antibody conjugated to horseradish peroxidase in blocking buffer at 1:10,000. Plates were developed by the addition of the horseradish peroxidase substrate 3,3',5,5'-tetramethylbenzidine (TMB; 34021, Thermo Fisher

Scientific) for 10 min, and the developing reaction was stopped by adding 50 µl of 1 M HCl. Optical density units were measured at 450 nm in a microplate reader (Tecan). For serum samples, a positive control (serum pool from critically ill patients with COVID-19, diluted 200-fold in PBS) and a negative control (pool of historical serum samples) were added in duplicate to each run. After deduction of the background, a relative content in IgA equivalent (ng ml⁻¹ Eq) was calculated using the calibration curve. The limit of detection of the assay was 0.1 ng ml⁻¹ Eq.

Live-virus neutralization experiments

A plaque reduction neutralization test (PRNT) was used for the detection and titration of neutralizing antibodies. A tenfold dilution of each serum specimen in culture medium was first heated for 30 min at 56 °C to avoid complement-linked reduction of viral activity. Serial twofold dilutions (tested in duplicate) of the serum specimens in culture medium were mixed in equal volume with the live SARS-CoV-2 virus. After gentle shaking and a contact of 30 min at room temperature in plastic microplates, 150 µl of the mix was transferred into 96-well microplates covered with Vero E6 cells (American Type Culture Collection (ATCC), CRL-1586, not authenticated but regularly tested for mycoplasma contamination). The plates were incubated at 37 °C in a 5% CO₂ atmosphere. Infection efficiency was evaluated by microscopy 5 days later when the cytopathic effect of the virus control reached 100–500 TCID₅₀ (median culture infectious dose) per 150 µl. Neutralization was recorded if more than 50% of the cells present in the well were preserved. The neutralizing titre was expressed as the inverse of the higher serum dilution that exhibited neutralizing activity; a threshold of 20 was used (PRNT₅₀ titre ≥ 20). All experiments were performed in a biosafety level 3 laboratory. The different viral strains that were used were sequenced and deposited at GISAID (<https://www.gisaid.org/>) (accession numbers EPI_ISL_1707038 19A (B.38 lineage); EPI_ISL_1707039 Alpha (B.1.1.7 lineage); EPI_ISL_768828 Beta (B.1.351 lineage); EPI_ISL_1359892 Gamma (P.1 lineage); and EPI_ISL_1904989 Delta (B.1.617.2 lineage)).

Monitoring of T cell responses using whole-blood IFNγ release assay

Fresh blood collected in heparinized tubes was stimulated for 22 h at 37 °C under 5% CO₂ with SARS-CoV-2 peptide pools (derived from the prototype Wuhan strain [NC_045512.2](#)) targeting RBD (46 peptides, [C] = 8 µg ml⁻¹) (bioMérieux) diluted in IGRA solution (bioMérieux). The peptides (15-mer) encompassed the entire protein sequence and overlapped by five residues. The IGRA solution was used as a negative control and a mitogen as a positive control. The concentration of IFNγ in the supernatant was measured using a VIDAS automated platform (VIDAS IFNγ RUO, bioMérieux). The measuring range was 0.08–8 IU ml⁻¹ and WB IGRA positivity

thresholds were defined at 0.08 IU ml^{-1} . The IFN γ response was defined as detectable when the IFN γ concentration of the test was above threshold and the negative control was below threshold or when the IFN γ concentration of the test minus IFN γ concentration of the negative control was above threshold.

Monitoring of T cell responses by flow cytometry

Overnight-rested PBMCs were stimulated with SARS-CoV-2 overlapping peptide pools against SARS-CoV-2 spike protein (PepTivator, Miltenyi Biotec) at a final concentration of $1 \mu\text{g ml}^{-1}$ for 1 h in the presence of $1 \mu\text{g ml}^{-1}$ monoclonal antibodies CD28 and CD49d, and then for an additional 5 h with GolgiPlug and GolgiStop (BD Biosciences). Dead cells were labelled using Fixable Viability eFluor 780 dye (Thermo Fisher Scientific). Surface markers were stained using BV786-conjugated anti-CD3 (BD Biosciences, 565491; diluted 1:100), BUV486-conjugated anti-CD4 (BD Biosciences, 612937; 1:50), PE-Cy7-conjugated anti-CD8 (BioLegend, 301012; 1:100), APC-Cy7-conjugated anti-CD14 (BioLegend, 301820; 1:100), APC-Cy7-conjugated anti-CD56 (BioLegend, 362512; 1:100) and APC-Cy7-conjugated anti-CD19 (BioLegend, 302218; 1:100). Cells were then washed, fixed with Cytofix/Cytoperm (BD Biosciences) and stained with PE-conjugated anti-IFN γ (BioLegend). Negative controls without peptide stimulation were run for each sample. All results were acquired on a BD LSRFortessa (BD Biosciences) flow cytometer using the BD FACSDIVA v.8.01 software and analysed using FlowJo v.10.6.1 software.

SARS-CoV-2 pseudoparticle preparation and neutralization

SARS-CoV-2 spike-pseudotyped murine leukaemia virus retrovirus particles were produced as described for SARS-CoV³⁰. In brief, HEK293T cells (ATCC, CRL-1573, not authenticated but regularly tested for mycoplasma contamination) were transfected with constructs expressing murine leukaemia virus Gag-Pol, the green fluorescent protein (GFP) reporter and the SARS-CoV-2 spike protein (a gift from D. Lavillette (CAS Key Laboratory of Molecular Virology & Immunology, Institut Pasteur of Shanghai Chinese Academy of Sciences, Pasteurien College, Soochow University, Jiangsu, China)). Control particles pseudotyped with the unrelated RD114 virus surface glycoprotein (from a cat endogenous virus) were generated as previously described³¹. For neutralization assays, a sample of approximately 1×10^3 pseudoparticles was incubated with a 100-fold dilution of sera or control antibodies for 1 h at 37°C , spinoculated for 2 h at $2,500\text{g}$ before infection of Vero E6 cells. After 72 h of infection, the percentage of GFP-positive cells was determined by flow cytometry. As a control, the same procedure was performed using RD114 pseudoparticles. Anti-spike SARS-CoV-2 RBD (Sino Biological, 40150-V08B2) and anti-gp70 RD114

(ViroMed Biosafety Labs) antibodies were used with a 100-fold dilution as positive and negative control, respectively.

Generation of fluorescent SARS-CoV-2 RBD tetramers

Biotinylated recombinant RBD domain of SARS-CoV-2 RBD was purchased from Miltenyi Biotech (130-127-457) and tetramerized with either streptavidin-PE (Becton Dickinson (BD)) or with streptavidin-APC (BioLegend, 105243).

Flow cytometry analysis of SARS-CoV-2 RBD-specific B cells

Cryopreserved PBMCs were centrifuged and suspended in PEB buffer (PBS with 0.5% BSA and 2 mM EDTA) and incubated with Fc receptor block (Miltenyi, 130-059-901) for 15 min at 4 °C (dilution 1:10). Next, cells were washed in PEB and stained for 30 min in brilliant stain buffer at 4 °C in the dark using the following antibodies together with both the PE- and APC-conjugated recombinant RBD tetramers: anti-CD3-APC Fire 810 (BioLegend, 344858; diluted 1:100); anti-CD11c-BV785 (BioLegend, 301644; 1:50); anti-CD19-PE Vio770 (Miltenyi, 130-113-170; 1:100); anti-CD20-BV421 (BD, 562873; 1:100); anti-CD21-BUV496 (BD, 750614; 1:50); anti-CD27-PercP-Vio700 (Miltenyi, 130-113-632; 1:100); anti-CD38-Viobright FITC (Miltenyi, 130-113-433; 1:50); anti-IgM-PE-CF594 (BD, 562539; 1:50); anti-IgD-BV605 (BioLegend, 348232; 1:50); and anti-IgG-BV480 (BD, 746341; 1:50). Cells were washed in PEB and resuspended in a PEB dilution (1:500) of the fixable viability dye eFluor 780 (eBiosciences, 65-0865-18). They were next washed and fixed with 4% paraformaldehyde for 20 min at 4 °C in the dark before a final wash and resuspension for analysis. Cells were then acquired on a Cytek Aurora spectral flow cytometer equipped with five lasers operating at 355 nm, 405 nm, 488 nm, 561 nm and 640 nm using the SpectroFlo V.2.2.0 (Cytek) software. Data were analysed using FlowJo 10.6.1 software (BD).

Statistical analyses

Statistical modelling for epidemiological data

To model the probability of being infected knowing the subject age and the used vaccination regimen, we used the generalized linear model (glm) function in R to perform a multiple logistic regression model with the logit link function. The outcome variable was the infection status, whereby infected = 1 and not infected = 0. Our predictors were the age, as a continuous variable (considered as a possible confounder), and the vaccination regimen, whereby Chad–BNT = 0 and BNT–BNT = 1. On the basis of the model coefficients for the vaccination regimen ($B_0 = -5.306$ and

$B_1 = 0.712$), we also assessed the relative risk (RR) as follows: $RR = (1 + \exp(-B_0)) / (1 + \exp(-B_0 - B_1))$.

Statistical analysis for immunological and virological analyses

ChAd–BNT and BNT–BNT populations showed a significant difference in the average ages. To test for the consequence of the vaccination regimen on the different immune parameters, we fitted a multiple linear regression model, which allows for the correction of age (as a possible confounding factor) by including age and vaccination regimen simultaneously as predictors in the right explanatory side of the lm function in R (one model per immune parameter). Results of statistical significance displayed in Figs. 2 and 3 are from these adjusted models, for which age showed no effect. For box and whiskers plots, the upper and lower bounds of the box are the 75th and the 25th percentile (the third Q3 and first Q1 quartile), respectively. The $IQR = Q3 - Q1$; an observation was considered an outlier if it was above $Q3 + 1.5 \times IQR$ or below $Q1 - 1.5 \times IQR$. Upper and lower whiskers represent the maximum and the minimum values, respectively, without taking into account the outliers.

Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

Data availability

Viral sequences are available at GISAID (accession numbers EPI_ISL_1707038 19A (B.38 lineage); EPI_ISL_1707039 Alpha (B.1.1.7 lineage); EPI_ISL_768828 Beta (B.1.351 lineage); EPI_ISL_1359892 Gamma (P.1 lineage); and EPI_ISL_1904989 Delta (B.1.617.2 lineage). [Source data](#) are provided with this paper.

References

1. 1.

Greinacher, A. et al. Thrombotic thrombocytopenia after ChAdOx1 nCov-19 vaccination. *N. Engl. J. Med.* **384**, 2092–2101 (2021).

2. 2.

Scully, M. et al. Pathologic antibodies to platelet factor 4 after ChAdOx1 nCoV-19 vaccination. *N. Engl. J. Med.* **384**, 2202–2211 (2021).

3. 3.

Polack, F. P. et al. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. *N. Engl. J. Med.* **383**, 2603–2615 (2020).

4. 4.

Ramasamy, M. N. et al. Safety and immunogenicity of ChAdOx1 nCoV-19 vaccine administered in a prime–boost regimen in young and old adults (COV002): a single-blind, randomised, controlled, phase 2/3 trial. *Lancet* **396**, 1979–1993 (2021).

5. 5.

Voysey, M. et al. Safety and efficacy of the ChAdOx1 nCoV-19 vaccine (AZD1222) against SARS-CoV-2: an interim analysis of four randomised controlled trials in Brazil, South Africa, and the UK. *Lancet* **397**, 99–111 (2021).

6. 6.

Liu, X. et al. Safety and immunogenicity of heterologous versus homologous prime–boost schedules with an adenoviral vectored and mRNA COVID-19 vaccine (Com-COV): a single-blind, randomised, non-inferiority trial. *Lancet* **398**, 856–869 (2021).

7. 7.

Hillus, D. et al. Safety, reactogenicity, and immunogenicity of homologous and heterologous prime–boost immunisation with ChAdOx1 nCoV-19 and BNT162b2: a prospective cohort study. *Lancet Respir. Med.* **9**, 1255–1265 (2021).

8. 8.

Legros, V. et al. A longitudinal study of SARS-CoV-2-infected patients reveals a high correlation between neutralizing antibodies and COVID-19 severity. *Cell. Mol. Immunol.* **18**, 318–327 (2021).

9. 9.

Gonzalez, C. et al. Live virus neutralisation testing in convalescent patients and subjects vaccinated against 19A, 20B, 20I/501Y.V1 and 20H/501Y.V2 isolates of SARS-CoV-2. *Emerg. Microbes Infect.* **10**, 1499–1502 (2021).

10. 10.

Sanz, I. et al. Challenges and opportunities for consistent classification of human B cell and plasma cell populations. *Front. Immunol.* **10**, 2458 (2019).

11. 11.

Lu, S. Heterologous prime–boost vaccination. *Curr. Opin. Immunol.* **21**, 346–351 (2009).

12. 12.

Jou, J., Harrington, K. J., Zocca, M.-B., Ehrnrooth, E. & Cohen, E. E. W. The changing landscape of therapeutic cancer vaccines—novel platforms and neoantigen identification. *Clin. Cancer Res.* **27**, 689–703 (2021).

13. 13.

Levine, M. Z. et al. Heterologous prime–boost with A(H5N1) pandemic influenza vaccines induces broader cross-clade antibody responses than homologous prime-boost. *NPJ Vaccines* **4**, 22 (2019).

14. 14.

Spencer, A. J. et al. Heterologous vaccination regimens with self-amplifying RNA and adenoviral COVID vaccines induce robust immune responses in mice. *Nat. Commun.* **12**, 2893 (2021).

15. 15.

Alizon, S. et al. Rapid spread of the SARS-CoV-2 Delta variant in some French regions, June 2021. *Euro Surveill.* **26**, 2100573 (2021).

16. 16.

Schmidt, T. et al. Immunogenicity and reactogenicity of heterologous ChAdOx1 nCoV-19/mRNA vaccination. *Nat. Med.* **27**, 1530–1535 (2021).

17. 17.

Tenbusch, M. et al. Heterologous prime–boost vaccination with ChAdOx1 nCoV-19 and BNT162b2. *Lancet Infect. Dis.* **21**, 1212–1213 (2021).

18. 18.

Groß, R. et al. Heterologous ChAdOx1 nCoV-19 and BNT162b2 prime–boost vaccination elicits potent neutralizing antibody responses and T cell reactivity. Preprint at <https://doi.org/10.1101/2021.05.30.21257971> (2021).

19. 19.

von Rhein, C. et al. Comparison of potency assays to assess SARS-CoV-2 neutralizing antibody capacity in COVID-19 convalescent plasma. *J. Virol. Methods* **288**, 114031 (2021).

20. 20.

Juraszek, J. et al. Stabilizing the closed SARS-CoV-2 spike trimer. *Nat. Commun.* **12**, 244 (2021).

21. 21.

Sette, A. & Crotty, S. Adaptive immunity to SARS-CoV-2 and COVID-19. *Cell* **184**, 861–880 (2021).

22. 22.

Wang, Z. et al. Enhanced SARS-CoV-2 neutralization by dimeric IgA. *Sci. Transl. Med.* **13**, eabf1555 (2021).

23. 23.

Wang, Z. et al. Naturally enhanced neutralizing breadth against SARS-CoV-2 one year after infection. *Nature* **595**, 426–431 (2021).

24. 24.

Haynes, B. F. et al. Immune-correlates analysis of an HIV-1 vaccine efficacy trial. *N. Engl. J. Med.* **366**, 1275–1286 (2012).

25. 25.

Tomaras, G. D. et al. Vaccine-induced plasma IgA specific for the C1 region of the HIV-1 envelope blocks binding and effector function of IgG. *Proc. Natl Acad. Sci. USA* **110**, 9019–9024 (2013).

26. 26.

Voysey, M. et al. Single-dose administration and the influence of the timing of the booster dose on immunogenicity and efficacy of ChAdOx1 nCoV-19 (AZD1222) vaccine: a pooled analysis of four randomised trials. *Lancet* **397**, 881–891 (2021).

27. 27.

Payne, R. P. et al. Immunogenicity of standard and extended dosing intervals of BNT162b2 mRNA vaccine. *Cell* **184**, 5699–5714 (2021).

28. 28.

Barros-Martins, J. et al. Immune responses against SARS-CoV-2 variants after heterologous and homologous ChAdOx1 nCoV-19/BNT162b2 vaccination. *Nat. Med.* **27**, 1525–1529 (2021).

29. 29.

Amanat, F. et al. A serological assay to detect SARS-CoV-2 seroconversion in humans. *Nat. Med.* **26**, 1033–1036 (2020).

30. 30.

Broer, R., Boson, B., Spaan, W., Cosset, F.-L. & Corver, J. Important role for the transmembrane domain of severe acute respiratory syndrome coronavirus spike protein during entry. *J. Virol.* **80**, 1302–1310 (2006).

31. 31.

Lavillette, D. et al. Human serum facilitates hepatitis C virus infection, and neutralizing responses inversely correlate with viral replication kinetics at the acute phase of hepatitis C virus infection. *J. Virol.* **79**, 6023–6034 (2005).

Acknowledgements

This study was supported by IDEX Lyon (HPI-2019), operated by the French National Research Agency (ANR-11-IDEX-0007), INSERM-Transfert and by institutional grants from Fondation des Hospices Civils de Lyon, INSERM, CNRS, UCBL1 and ENS de Lyon. We thank the following staff members and individuals: all the staff members of the Occupational Health and Medicine Department of the Hospices Civils de Lyon (who contributed to the sample collection) and all the clinical research associates for their excellent work; the staff members of SFR Biosciences (UAR3444/CNRS, US8/INSERM, ENS de Lyon, UCBL) AniRA-cytometry, E. Devêvre, S. Dussurgey and V. Malassigné; K. Brahima and all the members of the

clinical research and innovation department for their reactivity (DRCI, Hospices Civils de Lyon); and all the HCWs for their participation in this clinical study. Human biological samples and associated data were obtained from NeuroBioTec (CRB HCL, Lyon France, Biobank BB-0033-00046).

Author information

Author notes

1. These authors contributed equally: Bruno Pozzetto, Vincent Legros, Sophia Djebali, Véronique Barateau
2. These authors jointly supervised this work: François-Loïc Cosset, Stéphane Paul, Thierry Defrance, Jacqueline Marvel, Thierry Walzer, Sophie Trouillet-Assant

Affiliations

1. CIRI (Centre International de Recherche en Infectiologie), Université de Lyon, Université Claude Bernard Lyon 1, INSERM U1111, CNRS, UMR5308, ENS Lyon, Université Jean Monnet de Saint-Etienne, Lyon, France

Bruno Pozzetto, Vincent Legros, Sophia Djebali, Véronique Barateau, Marine Villard, Loïc Peyrot, Omran Allatif, Melyssa Yaugel-Novoa, Solène Denolly, Bertrand Boson, Thomas Bourlet, Antonin Bal, Bruno Lina, Vanessa Escuret, Florence Morfin, Margaux Prieux, Valérie Dubois, Laurence Josset, François-Loïc Cosset, Stéphane Paul, Thierry Defrance, Jacqueline Marvel, Thierry Walzer & Sophie Trouillet-Assant

2. Immunology laboratory, CIC1408, CHU Saint-Etienne, Saint-Etienne, France

Bruno Pozzetto

3. Campus Vétérinaire de Lyon, VetAgro Sup, Université de Lyon, Marcy-l'Etoile, France

Vincent Legros

4. Occupational Health and Medicine Department, Hospices Civils de Lyon, Université Claude Bernard Lyon1, Ifsttar, UMRESTTE, UMR T_9405, Université de Lyon, Lyon, France

Nicolas Guibert, Jean-Baptiste Fassier, Amélie Massardier-Pilonchéry, Constance d'Aubarede & Virginie Pitiot

5. Joint Research Unit Civils Hospices of Lyon–bioMérieux, Hospices Civils de Lyon, Lyon Sud Hospital, Pierre-Bénite, France

Karen Brengel-Pesce, Kahina Saker, Christelle Compagnon, Bouchra Mokdad, Soizic Daniel & Sophie Trouillet-Assant

6. Virology Laboratory, Institute of Infectious Agents, Laboratory Associated with the National Reference Centre for Respiratory Viruses, Hospices Civils de Lyon, Lyon, France

Antonin Bal, Martine Valette, Bruno Lina, Vanessa Escuret, Florence Morfin, Mary-Anne Trabaud & Laurence Josset

7. Centre de Recherche en Cancérologie de Lyon, INSERM U1052, CNRS UMR5286, Université de Lyon, Université Claude Bernard Lyon 1, Centre Léon Bérard, Lyon, France

Thibault Andrieu

Consortia

Covid-Ser study group

- Kahina Saker
- , Christelle Compagnon
- , Bouchra Mokdad
- , Constance d'Aubarede
- , Virginie Pitiot
- , Vanessa Escuret
- , Florence Morfin
- , Mary-Anne Trabaud
- , Margaux Prieux
- , Valérie Dubois
- , Laurence Josset
- & Soizic Daniel

Contributions

B.P., V.L., S. Djebali, V.B., M. Villard, L.P., M.Y.-N., S. Denolly, B.B., T.B., M. Valette, A.B. and K.B.-P. performed experiments and analysed the data. T.A. helped with the analysis of flow cytometry data and O.A. performed statistical analyses. N.G., J.-B.F. and A.M.-P. analysed the clinical infection data. The Covid-Ser study members

collected the samples and performed experiments. B.L. supervised some of the experimental tasks. F.-L.C., S.P., T.D., J.M., T.W. and S.T.-A. designed the study, analysed the data and wrote the manuscript.

Corresponding authors

Correspondence to [François-Loïc Cosset](#), [Stéphane Paul](#), [Thierry Defrance](#), [Jacqueline Marvel](#), [Thierry Walzer](#) or [Sophie Trouillet-Assant](#).

Ethics declarations

Competing interests

K.B.-P. is a bioMérieux employee.

Additional information

Peer review information *Nature* thanks Yunda Huang, Pei-Yong Shi and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

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 Study design.](#)

Sixty HCWs naïve for COVID-19 and vaccinated with the Pfizer ChAd/BNT or the BNT/BNT combination were included in a prospective longitudinal cohort study. Blood sampling was performed as described, before vaccination, before the second dose of vaccine and 4 weeks after the end of vaccination schedule. Different immunological analyses were performed on the blood samples, including serological investigations (Spike-specific IgA and IgG), serum neutralization assays (both plaque reduction neutralization test (PRNT) and pseudoneutralization), analysis of Spike-specific T cells and RBD-specific memory B cells, at various time points indicated in the text. Parts of the figure were drawn by using pictures from Servier Medical Art (<http://smart.servier.com/>), licensed under a Creative Commons Attribution 3.0 Unported License (<https://creativecommons.org/licenses/by/3.0/>).

Extended Data Fig. 2 Analysis of RBD-specific memory B cells.

(A) Gating strategy. From left to right, the first three pseudo-color plots show the successive gates applied for single cells and viable cells. B cells were gated as CD19⁺/CD3⁻ cells (4th plot). Within the B cell gate, memory B cells (mBCs) were defined as non-naïve B cells on the same biparameter plot (5th plot). (B-E) Concatenated phenotypic profiles of RBD-specific mBCs for all individuals in the ChAd/BNT and BNT/BNT groups at three time points (before vaccination, before the 2nd dose, after the second dose). For each group, the phenotypic profiles correspond to the concatenation of the FACS analysis data collected from 30 individual vaccinees. (B) Visualization of RBD-binding B cells in the mBC gate. (C) Visualization of the proportion of switched (IgD⁻) and unswitched (IgD⁺) B cells among the RBD-specific mBC pool. (D) Visualization of the pattern of Ig isotype expression by RBD-specific mBCs. (E) Visualization of the distribution of RBD-binding B cells within the resting (CD21⁺CD11c⁻) and “activated” (CD21⁻CD11c⁺/CD21⁻CD11c⁻/ CD21⁺CD11c⁺) mBC compartments. The mean frequencies of RBD-specific mBCs falling into each quadrant are indicated in red.

Extended Data Fig. 3 Analysis of SARS-CoV-2-specific T cells.

PBMCs were stimulated as detailed in the methods and stained for CD3, CD4, CD8, CD14, CD56 and CD19 expression and with a dead cell marker. Cells were gated from left to right and top to bottom. The Viab/dump channel includes antibodies against CD14, CD19, CD56 and a viability marker. The frequency of CD4 or CD8 T cells positive for intracellular IFNg was measured as shown in the central and right bottom panels.

Extended Data Fig. 4 Analysis of inter-relationships between immune parameters.

A matrix file with all immunological measurements for all vaccinated individuals showing the Spearman correlations and associated p-values. The size and the color of the circles correspond to the correlation, as indicated, and the stars indicate the significance of the correlation.

Supplementary information

Reporting Summary

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

Pozzetto, B., Legros, V., Djebali, S. *et al.* Immunogenicity and efficacy of heterologous ChAdOx1–BNT162b2 vaccination. *Nature* **600**, 701–706 (2021). <https://doi.org/10.1038/s41586-021-04120-y>

- Received: 19 July 2021
- Accepted: 10 October 2021
- Published: 21 October 2021
- Issue Date: 23 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04120-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

- [SARS-CoV-2 specific cellular response following COVID-19 vaccination in patients with chronic lymphocytic leukemia](#)

- Sibylle C. Mellinghoff
- Sandra Robrecht
- Petra Langerbeins

Leukemia (2021)

- [Mix-and-match COVID vaccines ace the effectiveness test](#)

- Ewen Callaway

Nature (2021)

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04120-y>.

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

- Article
- [Published: 01 December 2021](#)

Antigen-presenting innate lymphoid cells orchestrate neuroinflammation

- [John B. Grigg^{1,2,3},](#)
- [Arathi Shanmugavadi^{4 na1},](#)
- [Tommy Regen^{4 na1},](#)
- [Christopher N. Parkhurst^{1,2,3},](#)
- [Anees Ahmed^{1,2,3},](#)
- [Ann M. Joseph^{1,2,3},](#)
- [Michael Mazzucco⁵,](#)
- [Konrad Gronke^{6,7},](#)
- [Andreas Diefenbach ORCID: orcid.org/0000-0002-9176-9530^{6,7},](#)
- [Gerard Eberl ORCID: orcid.org/0000-0002-1119-5638⁸,](#)
- [Timothy Vartanian⁵,](#)
- [Ari Waisman ORCID: orcid.org/0000-0003-4304-8234^{4 na2} &](#)
- [Gregory F. Sonnenberg ORCID: orcid.org/0000-0003-0948-5908^{1,2,3 na2}](#)

[Nature](#) volume 600, pages 707–712 (2021)

- 9722 Accesses
- 205 Altmetric
- [Metrics details](#)

Subjects

- [Autoimmunity](#)
- [Innate lymphoid cells](#)
- [Neuroimmunology](#)

Abstract

Pro-inflammatory T cells in the central nervous system (CNS) are causally associated with multiple demyelinating and neurodegenerative diseases^{1,2,3,4,5,6}, but the pathways that control these responses remain unclear. Here we define a population of inflammatory group 3 innate lymphoid cells (ILC3s) that infiltrate the CNS in a mouse model of multiple sclerosis. These ILC3s are derived from the circulation, localize in proximity to infiltrating T cells in the CNS, function as antigen-presenting cells that restimulate myelin-specific T cells, and are increased in individuals with multiple sclerosis. Notably, antigen presentation by inflammatory ILC3s is required to promote T cell responses in the CNS and the development of multiple-sclerosis-like disease in mouse models. By contrast, conventional and tissue-resident ILC3s in the periphery do not appear to contribute to disease induction, but instead limit autoimmune T cell responses and prevent multiple-sclerosis-like disease when experimentally targeted to present myelin antigen. Collectively, our data define a population of inflammatory ILC3s that is essential for directly promoting T-cell-dependent neuroinflammation in the CNS and reveal the potential of harnessing peripheral tissue-resident ILC3s for the prevention of autoimmune disease.

This is a preview of subscription content

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

Rent or Buy

All prices are NET prices.

Additional access options:

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

Fig. 1: A unique subset of ILC3s infiltrates the CNS during neuroinflammation.

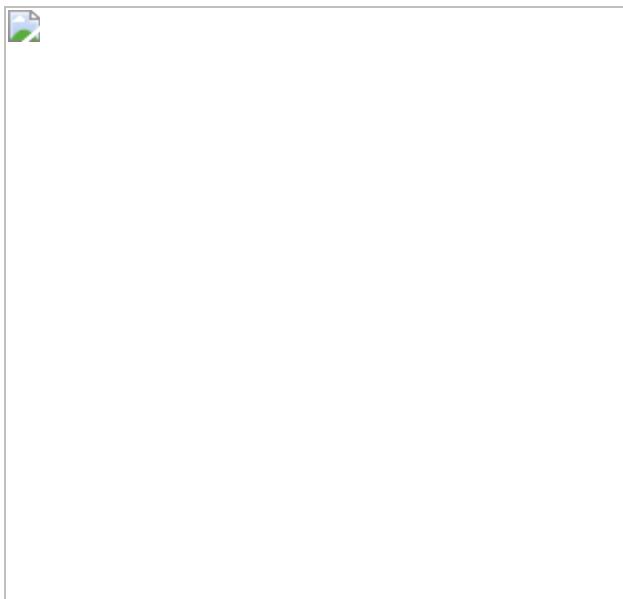


Fig. 2: CNS-associated ILC3s restimulate encephalitogenic T cells in the CNS.

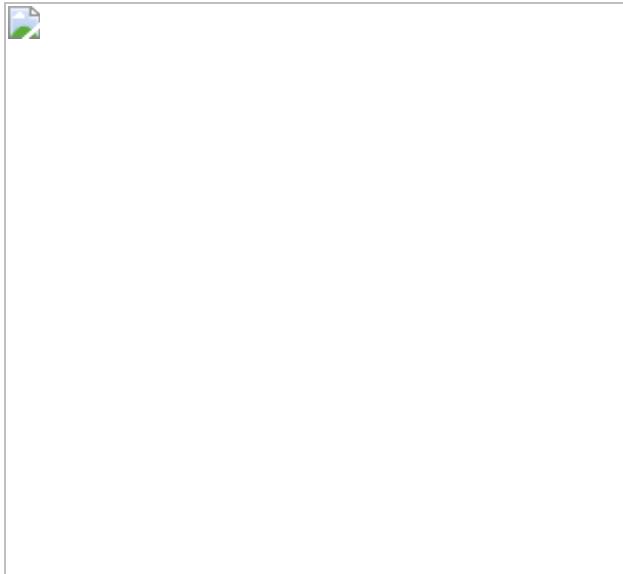


Fig. 3: MHCII⁺ ILC3s in the CNS promote autoimmune T cells and demyelinating disease.

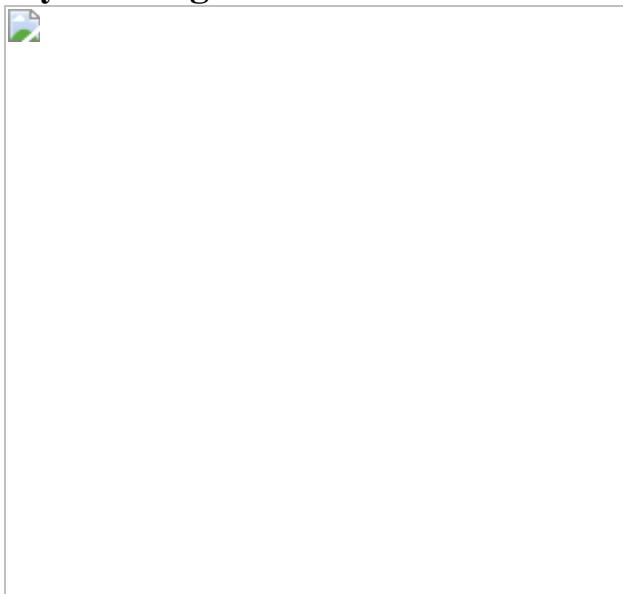
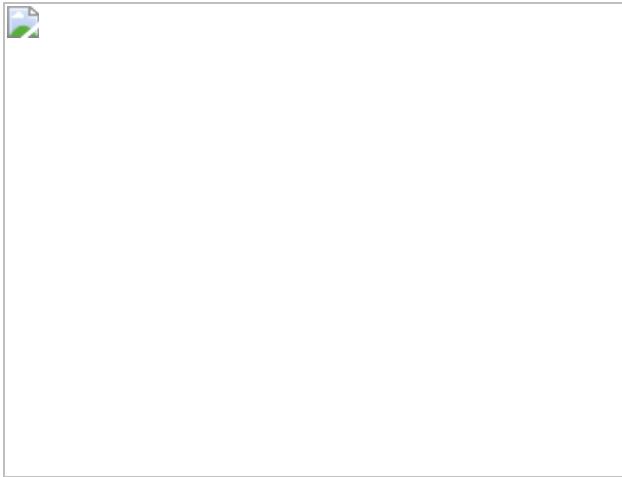


Fig. 4: Peripheral tissue-resident ILC3s can be harnessed to prevent neuroinflammation.



Data availability

ILC3 RNA-sequencing data are available at the Gene Expression Omnibus under accession number [GSE165580](#). Faecal 16S data are deposited at the NCBI Sequence Read Archive under BioProject accession number [PRJNA761714](#). [Source data](#) are provided with this paper.

References

1. 1.
Ota, K. et al. T-cell recognition of an immunodominant myelin basic protein epitope in multiple sclerosis. *Nature* **346**, 183–187 (1990).
2. 2.
Dendrou, C. A., Fugger, L. & Friese, M. A. Immunopathology of multiple sclerosis. *Nat. Rev. Immunol.* **15**, 545–558 (2015).
3. 3.
Togo, T. et al. Occurrence of T cells in the brain of Alzheimer's disease and other neurological diseases. *J. Neuroimmunol.* **124**, 83–92 (2002).
4. 4.

Monsonego, A. et al. Increased T cell reactivity to amyloid β protein in older humans and patients with Alzheimer disease. *J. Clin. Invest.* **112**, 415–422 (2003).

5. 5.

Sulzer, D. et al. T cells from patients with Parkinson’s disease recognize α -synuclein peptides. *Nature* **546**, 656–661 (2017).

6. 6.

Lindestam Arlehamn, C. S. et al. alpha-Synuclein-specific T cell reactivity is associated with preclinical and early Parkinson’s disease. *Nat. Commun.* **11**, 1875 (2020).

7. 7.

Lincoln, M. R. et al. A predominant role for the HLA class II region in the association of the MHC region with multiple sclerosis. *Nat. Genet.* **37**, 1108–1112 (2005).

8. 8.

Hamza, T. H. et al. Common genetic variation in the HLA region is associated with late-onset sporadic Parkinson’s disease. *Nat. Genet.* **42**, 781–785 (2010).

9. 9.

Jansen, I. E. et al. Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer’s disease risk. *Nat. Genet.* **51**, 404–413 (2019).

10. 10.

Nalls, M. A. et al. Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson’s disease. *Nat. Genet.* **46**, 989–993 (2014).

11. 11.

Fallis, R. J., Raine, C. S. & McFarlin, D. E. Chronic relapsing experimental allergic encephalomyelitis in SJL mice following the adoptive transfer of an epitope-specific T cell line. *J. Neuroimmunol.* **22**, 93–105 (1989).

12. 12.

Brochard, V. et al. Infiltration of CD4⁺ lymphocytes into the brain contributes to neurodegeneration in a mouse model of Parkinson disease. *J. Clin. Invest.* **119**, 182–192 (2009).

13. 13.

Browne, T. C. et al. IFN- γ production by amyloid β -specific Th1 cells promotes microglial activation and increases plaque burden in a mouse model of Alzheimer's disease. *J. Immunol.* **190**, 2241–2251 (2013).

14. 14.

Lodygin, D. et al. β -Synuclein-reactive T cells induce autoimmune CNS grey matter degeneration. *Nature* **566**, 503–508 (2019).

15. 15.

Dulken, B. W. et al. Single-cell analysis reveals T cell infiltration in old neurogenic niches. *Nature* **571**, 205–210 (2019).

16. 16.

Vivier, E. et al. Innate lymphoid cells: 10 years on. *Cell* **174**, 1054–1066 (2018).

17. 17.

Sonnenberg, G. F. & Hepworth, M. R. Functional interactions between innate lymphoid cells and adaptive immunity. *Nat. Rev. Immunol.* **19**, 599–613 (2019).

18. 18.

Mair, F. & Becher, B. Thy1⁺ Sca1⁺ innate lymphoid cells infiltrate the CNS during autoimmune inflammation, but do not contribute to disease development. *Eur. J. Immunol.* **44**, 37–45 (2014).

19. 19.

Hatfield, J. K. & Brown, M. A. Group 3 innate lymphoid cells accumulate and exhibit disease-induced activation in the meninges in EAE. *Cell. Immunol.* **297**, 69–79 (2015).

20. 20.

Kwong, B. et al. T-bet-dependent NKp46⁺ innate lymphoid cells regulate the onset of T_H17-induced neuroinflammation. *Nat. Immunol.* **18**, 1117–1127 (2017).

21. 21.

Yamano, T. et al. Aire-expressing ILC3-like cells in the lymph node display potent APC features. *J. Exp. Med.* **216**, 1027–1037 (2019).

22. 22.

Gasteiger, G., Fan, X., Dikiy, S., Lee, S. Y. & Rudensky, A. Y. Tissue residency of innate lymphoid cells in lymphoid and nonlymphoid organs. *Science* **350**, 981–985 (2015).

23. 23.

Takeshita, Y. & Ransohoff, R. M. Inflammatory cell trafficking across the blood–brain barrier: chemokine regulation and in vitro models. *Immunol. Rev.* **248**, 228–239 (2012).

24. 24.

Perry, J. S. et al. Inhibition of LT_i cell development by CD25 blockade is associated with decreased intrathecal inflammation in multiple sclerosis. *Sci. Transl. Med.* **4**, 145ra106 (2012).

25. 25.

Lin, Y. C. et al. Daclizumab reverses intrathecal immune cell abnormalities in multiple sclerosis. *Ann. Clin. Transl. Neurol.* **2**, 445–455 (2015).

26. 26.

Degn, M. et al. Increased prevalence of lymphoid tissue inducer cells in the cerebrospinal fluid of patients with early multiple sclerosis. *Mult. Scler.* **22**, 1013–1020 (2016).

27. 27.

Serafini, B. et al. ROR γ t expression and lymphoid neogenesis in the brain of patients with secondary progressive multiple sclerosis. *J. Neuropathol. Exp. Neurol.* **75**, 877–888 (2016).

28. 28.

Hepworth, M. R. et al. Innate lymphoid cells regulate CD4 $^{+}$ T-cell responses to intestinal commensal bacteria. *Nature* **498**, 113–117 (2013).

29. 29.

Hepworth, M. R. et al. Immune tolerance. Group 3 innate lymphoid cells mediate intestinal selection of commensal bacteria-specific CD4 $^{+}$ T cells. *Science* **348**, 1031–1035 (2015).

30. 30.

von Burg, N. et al. Activated group 3 innate lymphoid cells promote T-cell-mediated immune responses. *Proc. Natl Acad. Sci. USA* **111**,

12835–12840 (2014).

31. 31.

Ting, J. P. & Trowsdale, J. Genetic control of MHC class II expression. *Cell* **109**, S21–S33, (2002).

32. 32.

Schroder, K., Hertzog, P. J., Ravasi, T. & Hume, D. A. Interferon- γ : an overview of signals, mechanisms and functions. *J. Leukoc. Biol.* **75**, 163–189 (2004).

33. 33.

Bryant, P. W., Lennon-Dumenil, A. M., Fiebiger, E., Lagaudriere-Gesbert, C. & Ploegh, H. L. Proteolysis and antigen presentation by MHC class II molecules. *Adv. Immunol.* **80**, 71–114 (2002).

34. 34.

Zhang, Q. & Vignali, D. A. Co-stimulatory and co-inhibitory pathways in autoimmunity. *Immunity* **44**, 1034–1051 (2016).

35. 35.

Lee, J. Y. et al. Serum amyloid A proteins induce pathogenic Th17 cells and promote inflammatory disease. *Cell* **180**, 79–91 (2020).

36. 36.

Koda, T. et al. Sema4A is implicated in the acceleration of Th17 cell-mediated neuroinflammation in the effector phase. *J. Neuroinflammation* **17**, 82 (2020).

37. 37.

Hur, E. M. et al. Osteopontin-induced relapse and progression of autoimmune brain disease through enhanced survival of activated T

cells. *Nat. Immunol.* **8**, 74–83 (2007).

38. 38.

Giles, D. A., Duncker, P. C., Wilkinson, N. M., Washnock-Schmid, J. M. & Segal, B. M. CNS-resident classical DCs play a critical role in CNS autoimmune disease. *J. Clin. Invest.* **128**, 5322–5334 (2018).

39. 39.

Mundt, S. et al. Conventional DCs sample and present myelin antigens in the healthy CNS and allow parenchymal T cell entry to initiate neuroinflammation. *Sci. Immunol.* **4**, eaau8380 (2019).

40. 40.

Korn, T. & Kallies, A. T cell responses in the central nervous system. *Nat. Rev. Immunol.* **17**, 179–194 (2017).

41. 41.

Waismann, A. & Johann, L. Antigen-presenting cell diversity for T cell reactivation in central nervous system autoimmunity. *J. Mol. Med.* **96**, 1279–1292 (2018).

42. 42.

Frommer, F. et al. Tolerance without clonal expansion: self-antigen-expressing B cells program self-reactive T cells for future deletion. *J. Immunol.* **181**, 5748–5759 (2008).

43. 43.

Buonocore, S. et al. Innate lymphoid cells drive interleukin-23-dependent innate intestinal pathology. *Nature* **464**, 1371–1375 (2010).

44. 44.

Huang, Y. et al. IL-25-responsive, lineage-negative KLRG1^{hi} cells are multipotential ‘inflammatory’ type 2 innate lymphoid cells. *Nat. Immunol.* **16**, 161–169 (2015).

45. 45.

Huang, Y. et al. S1P-dependent interorgan trafficking of group 2 innate lymphoid cells supports host defense. *Science* **359**, 114–119 (2018).

46. 46.

Montaldo, E. et al. Human ROR γ t⁺CD34⁺ cells are lineage-specified progenitors of group 3 ROR γ t⁺ innate lymphoid cells. *Immunity* **41**, 988–1000 (2014).

47. 47.

Lim, A. I. et al. Systemic human ILC precursors provide a substrate for tissue ILC differentiation. *Cell* **168**, 1086–1100 (2017).

48. 48.

Scoville, S. D. et al. A progenitor cell expressing transcription factor ROR γ t generates all human innate lymphoid cell subsets. *Immunity* **44**, 1140–1150 (2016).

49. 49.

Jordao, M. J. C. et al. Single-cell profiling identifies myeloid cell subsets with distinct fates during neuroinflammation. *Science* **363**, eaat7554 (2019).

50. 50.

Hashimoto, K., Joshi, S. K. & Koni, P. A. A conditional null allele of the major histocompatibility IA-beta chain gene. *Genesis* **32**, 152–153 (2002).

51. 51.

Bettelli, E. et al. Myelin oligodendrocyte glycoprotein-specific T cell receptor transgenic mice develop spontaneous autoimmune optic neuritis. *J. Exp. Med.* **197**, 1073–1081 (2003).

52. 52.

Srinivas, S. et al. Cre reporter strains produced by targeted insertion of *EYFP* and *ECFP* into the *ROSA26* locus. *BMC Dev. Biol.* **1**, 4 (2001).

53. 53.

Lee, P. P. et al. A critical role for Dnmt1 and DNA methylation in T cell development, function, and survival. *Immunity* **15**, 763–774 (2001).

54. 54.

Dobes, J. et al. A novel conditional *Aire* allele enables cell-specific ablation of the immune tolerance regulator *Aire*. *Eur. J. Immunol.* **48**, 546–548 (2018).

55. 55.

Hirota, K. et al. Fate mapping of IL-17-producing T cells in inflammatory responses. *Nat. Immunol.* **12**, 255–263 (2011).

56. 56.

Ahlfors, H. et al. IL-22 fate reporter reveals origin and control of IL-22 production in homeostasis and infection. *J. Immunol.* **193**, 4602–4613 (2014).

57. 57.

Lochner, M. et al. In vivo equilibrium of proinflammatory IL-17⁺ and regulatory IL-10⁺ Foxp3⁺ ROR γ t⁺ T cells. *J. Exp. Med.* **205**, 1381–1393 (2008).

58. 58.

Croxford, A. L., Kurschus, F. C. & Waisman, A. Cutting edge: an IL-17F-Cre^{EYFP} reporter mouse allows fate mapping of Th17 cells. *J. Immunol.* **182**, 1237–1241 (2009).

59. 59.

Polman, C. H. et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Ann. Neurol.* **69**, 292–302 (2011).

60. 60.

Miller, S. D., Karpus, W. J. & Davidson, T. S. Experimental autoimmune encephalomyelitis in the mouse. *Curr. Protoc. Immunol.* **88**, 15.1.1–15.1.20 (2010).

61. 61.

Lee, Y. et al. Induction and molecular signature of pathogenic TH17 cells. *Nat. Immunol.* **13**, 991–999 (2012).

62. 62.

Kamran, P. et al. Parabiosis in mice: a detailed protocol. *J. Vis. Exp.* **80**, e50556 (2013).

63. 63.

Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).

64. 64.

Edgar, R. C. SINTAX: a simple non-Bayesian taxonomy classifier for 16S and ITS sequences. Preprint at <https://doi.org/10.1101/074161> (2016).

65. 65.

Cole, J. R. et al. Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* **42**, D633–D642 (2014).

66. 66.

McMurdie, P. J. & Holmes, S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8**, e61217 (2013).

67. 67.

Louveau, A., Filiano, A. J. & Kipnis, J. Meningeal whole mount preparation and characterization of neural cells by flow cytometry. *Curr. Protoc. Immunol.* **121**, e50 (2018).

Acknowledgements

We thank members of the Sonnenberg laboratory for discussions and critical reading of the manuscript; the Microbiome Core and Epigenomics Core of Weill Cornell Medicine and G. G. Putzel for technical assistance; C. Gai of Weill Cornell’s Center of Comparative Medicine and Pathology (CCMP) for performing the parabiosis surgeries; and I. Ivanov of Columbia University for sharing mouse lines. The Sonnenberg laboratory is supported by the National Institutes of Health (R01AI143842, R01AI123368, R01AI145989, R01AI162936, R21CA249284 and U01AI095608), the National Institute of Allergy and Infectious Diseases (NIAID) Mucosal Immunology Studies Team (MIST), the Crohn’s and Colitis Foundation, the Searle Scholars Program, the American Asthma Foundation Scholar Award, Pilot Project Funding from the Center for Advanced Digestive Care (CADC), an Investigators in the Pathogenesis of Infectious Disease Award from the Burroughs Wellcome Fund, a Wade F.B. Thompson–Cancer Research Institute (CRI) CLIP Investigator grant, the Meyer Cancer Center Collaborative Research Initiative, the Dalton Family Foundation, L. and G. Greenberg, and the Jill Roberts Institute for Research in Inflammatory Bowel Disease. G.F.S. is a CRI Lloyd J. Old STAR. J.B.G is supported by the NIAID of the National Institutes of Health under award number F31AI138389-01A1; A.M.J is supported by T32DK116970; and the

Waisman laboratory is supported by the Deutsche Forschungsgemeinschaft (DFG) grants AW1600/10-1, AW1600/11-1 and AW1600/14-1, as well as by the National Multiple Sclerosis Society (NMSS) grant RG 1707-28780.

Author information

Author notes

1. These authors contributed equally: Arthi Shanmugavadi, Tommy Regen
2. These authors jointly supervised this work: Ari Waisman, Gregory F. Sonnenberg

Affiliations

1. Jill Roberts Institute for Research in Inflammatory Bowel Disease, Weill Cornell Medicine, Cornell University, New York, NY, USA
John B. Grigg, Christopher N. Parkhurst, Anees Ahmed, Ann M. Joseph & Gregory F. Sonnenberg
2. Joan and Sanford I. Weill Department of Medicine, Weill Cornell Medicine, Cornell University, New York, NY, USA
John B. Grigg, Christopher N. Parkhurst, Anees Ahmed, Ann M. Joseph & Gregory F. Sonnenberg
3. Department of Microbiology and Immunology, Weill Cornell Medicine, Cornell University, New York, NY, USA
John B. Grigg, Christopher N. Parkhurst, Anees Ahmed, Ann M. Joseph & Gregory F. Sonnenberg
4. Institute for Molecular Medicine, University Medical Center of the Johannes Gutenberg University Mainz, Mainz, Germany
Arthi Shanmugavadi, Tommy Regen & Ari Waisman

5. Brain and Mind Research Institute, Weill Cornell Medicine, Cornell University, New York, NY, USA

Michael Mazzucco & Timothy Vartanian

6. Laboratory of Innate Immunity, Department of Microbiology, Infectious Diseases and Immunology, Charité—Universitätsmedizin Berlin, Berlin, Germany

Konrad Gronke & Andreas Diefenbach

7. Mucosal and Developmental Immunology, Deutsches Rheuma-Forschungszentrum (DRFZ), Berlin, Germany

Konrad Gronke & Andreas Diefenbach

8. Microenvironment and Immunity Unit, Institut Pasteur, Paris, France

Gerard Eberl

Contributions

J.B.G. and G.F.S. conceived the project. J.B.G. performed most of the experiments and analysed the data. A.S., T.R., A.A. and A.M.J. also helped to process tissues, generated experimental datasets and analysed data from EAE in IiMOG mice (A.S. and T.R.) or Rorc^{cre} mice (A.A. and A.M.J.). C.N.P. processed tissues and generated the immunofluorescence data. M.M. coordinated the selection and T.V. provided and supervised the selection and analysis of samples from patients with RRMS, as well as providing scientific discussion. G.E., K.G. and A.D. provided essential mouse models, scientific advice and expertise. A.W. jointly supervised the research, provided essential mouse strains and contributed to the formation of the project. J.B.G. and G.F.S. wrote the manuscript, with input from all of the authors.

Corresponding author

Correspondence to [Gregory F. Sonnenberg](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work.

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 Characterization of ILC3s in the CNS.

a, Representative time course and clinical disease categorization of active EAE ($n = 4$, 5 mice/timepoint). **b-c**, Quantification of ILC3 frequency and absolute counts within indicated tissues at steady state (Naive) ($n = 6$ mice) versus EAE onset (d11, $n = 7$), acute (d15, $n = 8$) or chronic (d20, $n = 9$) phase (**b**) and reverse flow cytometry gating strategy defining all GFP⁺ cells in the CNS (d15) ($n = 4$ mice/group) (**c**) during EAE in Rorc-eGFP mice. **d-e**, Naive C57BL/6 mice ($n = 4$ mice/group) were immunized with either PBS (Naive), CFA/PTx alone, or with CFA/PTx/MOGp. At day 15 post-immunization, ILC3s in the CNS, inLN and cLN were enumerated by flow cytometry: **d**, Representative flow cytometry gating strategy for ILC3s in CNS (Lin1 = CD3, CD5, CD8, Lin2 = CD11b, CD11c, B220), **e**, Quantitation of frequencies and absolute counts of ILC3s. **f-g**, Representative flow cytometry on YFP⁺ ILCs (**f**) and quantification of differential expression of ILC heterogeneity in YFP^{+/−} ILCs (**g**) in the CNS of Rorc-cre^{eYFP} mice during active EAE ($n = 6$ mice/group) (Lin 1 = B220, CD11b, CD11c, Lin 2 = CD3ε, CD5, CD8, Ly6C). Data in **a-f** are representative of two independent experiments with similar results and data

in **b** are pooled from two independent experiments. Results are shown as mean \pm s.d. Statistics are calculated by one-way (**b**, **e**) analysis of variance (ANOVA) with Sidak's multiple comparisons test.

[Source data](#)

Extended Data Fig. 2 Phenotypic analyses of ILC3s in the CNS and peripheral blood.

a-c, Heat maps showing absolute log normalized counts (one plus log 2) (**a**, **b**) or relative expression Z-scores (**c**) from RNA sequencing of indicated lineage-specifying genes in sorted ILC3s from the cLN, CNS, or small intestine lamina propria (SI-LP) of Rorc-eGFP mice during peak of active EAE ($n = 4$ mice). **d**, Quantification of cytokine production by ILC3s ($CD45^+$, $CD3\epsilon^-$, $CD5^-$, $CD8\alpha^-$, $TCR\gamma\delta^-$, $NK1.1^-$, $CD11b^-$, $CD11c^-$, $B220^-$, $CD127^+$, $CD90.2^+$, $KLRG1^-$, $ROR\gamma t^+$) in indicated tissues during EAE (day 15) C57BL/6 mice ($n = 4$ mice). **e**, Representative flow cytometry gating strategy to detect ILC3s in human PBMCs (Lineage = $CD19$, $CD94$, $CD14$, $CD123$, $FcR1a$, $CD11c$) and frequency quantification of ILC3s in healthy control (HC) or RRMS samples ($n = 18$ sample-pairs – Supplementary Table 1). Data in **d** are representative of two independent experiments with similar results. Data in **e** are pooled and representative of two independent flow cytometry experiments on cryopreserved PBMC sample sets with similar results. Results are shown as mean \pm s.d. Statistics are calculated by one-way analysis of variance (ANOVA) with Sidak's multiple comparisons test (**d**).

[Source data](#)

Extended Data Fig. 3 Expression of MHCII and co-stimulatory molecules on ILC3s during neuroinflammation.

a, d, Representative histograms and quantification of MHCII (**a**), and CD80 or CD86 (**d**) expression on ILC3s in indicated tissues of Rorc-eGFP mice at steady state (Naive) ($n = 3$ mice) versus EAE onset (d11, $n = 4$), acute (d15, $n = 4$) or chronic (d20, $n = 5$) phase. **b**, Naive C57BL/6 mice were immunized with either PBS (Naive), CFA/PTx or with CFA/PTx/MOGp (n

= 4 mice/group). At day 15 EAE, frequency and counts of MHCII⁺ ILC3s in the CNS and cLN were enumerated by flow cytometry. **c, e** Representative staining and quantification of HLA-DR (**c**) and co-stimulatory molecules (**e**) on human blood ILC3s (n = 18 samples/group). Data in **a, b, d** are representative of two independent experiments with similar results. Results are shown as mean ± s.d. Statistics are calculated by one-way (**b**) analysis of variance (ANOVA) with Sidak's multiple comparisons test.

[Source data](#)

[Extended Data Fig. 4 ILC3s in the periphery do not express co-stimulatory molecules.](#)

a-b, Naive C57BL/6 mice were immunized with either PBS (Naive), CFA/PTx or with CFA/PTx/MOGp (EAE) (n = 4 mice/group). At day 15 post immunization, expression of co-stimulatory molecules CD80, CD86 (**a**) and OX40L, CD40, and CD30L (**b**) by MHCII⁺ ILC3s in the mLN, inLN, cLN and CNS were enumerated by flow cytometry. Data are representative of three independent experiments with similar results. Results are shown as mean ± s.d.

[Source data](#)

[Extended Data Fig. 5 ILC3s are found in the cerebral spinal fluid of patients with RRMS and express HLA-DR and CD86.](#)

a, Validation of anti-ROR γ t antibody staining on single cell suspensions from human donor tonsils compared to FMO (Fluorescence Minus One) control indicating staining without anti-ROR γ t antibody. **b-e**, Cerebral spinal fluid (CSF) was obtained from 7 patients with RRMS and one control individual (other neurological disease, OND). Individuals were further stratified by the presence or absence of contrast enhancing lesions (CEL) as well as CEL number (Supplementary Table 2). CSF was processed immediately by centrifuging for 10 min at 400 x g and staining for ILC3s as indicated (**b**) (Lineage = CD19, CD94, CD14, CD123, FcR1a, CD11c). Gated ROR γ t⁺ populations in the indicated tissues exhibited dim

staining for CD45, which is a defining feature of ILC3s, relative to CD4 T cells (**c**). Indicated frequencies of ILC3s in the CSF were quantified (**d**). PBMCs were used as controls during each collection and used for comparison of HLA-DR or CD86 expression on ILC3s in the CSF (**e**). Results are shown as mean (**d**) and RRMS flow cytometry is representative flow cytometry from CSF of patient #2 (Supplementary Table 2). Human tonsil data is representative of 3 individual tonsil samples.

[Source data](#)

[Extended Data Fig. 6 Interrogation of ILC3 and T cell interactions in the CNS.](#)

a, Representative immunofluorescence staining of fixed dura meninges, brain cerebellum and spinal cord (edge denoted by dashed white line) from Rorc-eGFP mice during day 11 or day 18 of active EAE showing enrichment of GFP⁺ cells in focal lesions of the parenchyma (representative of n = 3 mice/timepoint). **b-c**, Frequency of cytokine producing 2D2 T cells after 72 h co-culture with *ex vivo* sorted cDC or alone (2D2 T cells alone) from the CNS (**b**) or the mLN (**c**) (pooled from 5 mice, d18 EAE) in the presence of MOGp +/- α -MHCII blocking antibody. **d**, To determine the ability to process full-length antigen 2D2 T cells and ILC3s were sorted from the CNS (pooled from n = 5 mice, d19 EAE) and co-cultured as indicated for 72 h prior to determination of cell counts, staining of IFN γ or CD25 for flow cytometry. Resulting 2D2 cell counts were measured and normalized as a fold-change in comparison to myelin-specific 2D2 T cells cultured alone (no APC) and treated with MOGp₁₋₁₂₅ (dashed line) (left graph). Data in **d** are pooled from two independent experiments and data are representative of two (**b-c**) independent experiments with similar results. Data were necessarily pooled in noted experiments due to limited cell numbers. Results are shown as mean \pm s.d. Statistics are calculated by one-way analysis of variance (ANOVA) with Sidak's multiple comparisons test. Data points indicate technical well replicates and dashed lines (**b-d**) indicate baseline cytokine production by 2D2 T cells alone from indicated tissues.

[Source data](#)

Extended Data Fig. 7 Selective targeting of T cells and ILC3s through RORyt.

a, Expression of MHCII was quantified by flow cytometry in indicated antigen-presenting cells from the cLN or CNS of indicated mice at steady state ($n = 4$ mice/group). **b-c** Representative flow cytometry gating strategy for indicated immune cell populations in the CNS (**b**) and expression of YFP in indicated cell populations and tissues (**c**) of *Rorc-cre^{eYFP}* mice at day 15 of active EAE. **d**, Reverse flow cytometry gating strategy defining all YFP⁺ fate-mapped cells in the CNS of *Rorc-cre^{eYFP}* mice (day 15 EAE) (Lin 1 = B220, CD11b, CD11c, Lin 2 = CD3ε, CD5, CD8, Ly6C). Results are shown as mean ± s.d. Statistics are calculated by one-way analysis of variance (ANOVA) with Sidak's multiple comparisons test (**a**). Data in **a-d** are representative of two independent experiments with similar results

[Source data](#).

Extended Data Fig. 8 Effect of ILC3-specific MHCII and generalized intestinal inflammation on neuroinflammation.

a-d, Naive 2D2 T cells ($\text{Thy}1.1^+$) were transferred into recipient mice, which were immunized 24 h later to induce active EAE. At day 14 EAE, frequencies and counts (**a**), activation/proliferation (**b**), and polarization (**c**) of donor $\text{Thy}1.1^+$ 2D2 T cells (**c**, upper panel) or endogenous T cells (**c**, lower panel) were analysed by intracellular flow cytometry in the cLN or inLN ($n = 9$ mice/group (**a, b**), $n = 5$ mice/group (**c**)). **d**, Frequencies of IFNγ-producing endogenous T cells in the CNS were quantified by intracellular flow cytometry cytokine staining ($n = 8$ mice/group). **e-f**, Passive EAE was induced in recipient mice and day 15 frequencies of $\text{Thy}1.1^+$ 2D2 T cells in the cLN and CNS were determined by flow cytometry ($n = 4$ mice/group) (**e**). Representative H&E of fixed transverse spinal cord sections of mice shown at 4X magnification (top, scale bar = 300 μM) with 20X inset magnification (bottom, scale bar = 75 μm) of cellular infiltration near ventromedial fissure at day 15 passive EAE (**f**). **g-h**, Average clinical scores in indicated mice ($\text{MHCII}^{\Delta\text{Tcell}} = \text{CD4-cre}^+ \times \text{H2-Ab1}^{\text{fl/fl}}$) after induction of active EAE ($n = 4$ mice/group (**g**), $n = 5$

mice/group (**h**)). **i**, C57BL/6 mice were treated with either 3% D.S.S. in the drinking water for 7 days or were orally gavaged with *C. rodentium* (n = 5 mice/group). Active EAE was induced 14 days later and clinical scores were taken on mice or controls. Data are representative of two (**g-i**) or three (**a-f**) independent experiments with similar results. Data are pooled from two independent experiments in **a**, **b**, **d**. Results are shown as mean ± s.d. (**a-e**) or s.e.m (**g-i**) Statistics are calculated by two-way analysis of variance (ANOVA) (**c**, **d**) with Sidak's multiple comparisons test or unpaired, two-tailed *t*-test (Mann–Whitney U-test) (**a**, **b**, **e**).

[Source data](#)

[Extended Data Fig. 9 Targeting peripheral ILC3s and their role in neuroinflammation.](#)

a-d, Littermate H2-Ab1^{fl/fl} and MHCII^{ΔILC3} mice were treated with Vancomycin (Vanco) in the drinking water for 7 days prior to induction of active EAE and Vanco was continued throughout EAE. **a**, PCoA (Weighted UniFrac) of 16S rRNA gene sequencing of faecal samples at indicated treatment timepoints (n = 5 mice/group, rep. of N = 2). At the end of EAE plus vancomycin treatment, **b**, colon lengths, spleen mass, **c**, and total frequencies of endogenous neutrophils, CD4 T cells, and transferred Thy1.1+ 2D2 T cells in the indicated gastro-intestinal tissues. LI-LP = large intestine lamina propria. **d**, Daily (n = 5 mice/group) and cumulative clinical scores of mice (n = 9 mice/group, pooled from N = 2). **e-f**, Splenic antigen-presenting cells (APCs: DCs or B cells) were sort-purified from steady state (**e**) or CFA-treated mice (**f**, d10) and co-cultured with 2D2 T cells previously stained with violet cell trace violet. Groups included 2D2 T cells alone (no APC), 2D2 T cells plus APC, or 2D2 T cells plus APC and MOG peptide. After 72 h, proliferation of 2D2 T cells was determined by flow cytometry to measure cell trace violet dilution from cell division. **g**, Average clinical scores in indicated mice after induction of active EAE (n = 18-21 mice/group left panel pooled from three independent experiments with similar results, n = 13-4 mice/group middle panel pooled from three independent experiments with similar results, right panel n = 3 mice/group representative of two independent experiments with similar results). Results are shown as mean ± s.e.m (**d** left panel, **g**) or mean ± s.d. (**a-c**, **d** right

panel). Statistics are calculated by two-tailed t-test (Mann–Whitney U-test) (**b**, **d**) or two-way analysis of variance (ANOVA) (**c**) with Sidak’s multiple comparisons test. Results in **a** are representative of two independent experiments ($n = 4\text{--}5$ mice/group) and data in **b-d** are pooled from $N = 2$.

Source data

Extended Data Fig. 10 Antigen presenting group 3 innate lymphoid cells orchestrate neuroinflammation.

a, ILC3 family heterogeneity with inflammatory ILC3s (iILC3s) depicted in yellow. **b**, During autoimmune neuroinflammation, iILC3s enter the CNS from the circulation and are essential to promote pro-inflammatory T cell responses and demyelinating disease through antigen presentation. **c**, Tissue-resident and peripheral ILC3s retain tolerogenic potential, and when targeted to express myelin peptide can eliminate self-specific T cells and prevent demyelinating disease.

Supplementary information

Reporting Summary

Supplementary Tables

Supplementary Table 1 contains the sex and age range of patients with RRMS who provided PBMC samples. Supplementary Table 2 provides information on the patients with RRMS.

Source data

Source Data Fig. 1

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. 6**](#)

[**Source Data Extended Data Fig. 7**](#)

[**Source Data Extended Data Fig. 8**](#)

[**Source Data Extended Data Fig. 9**](#)

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Grigg, J.B., Shanmugavadiu, A., Regen, T. *et al.* Antigen-presenting innate lymphoid cells orchestrate neuroinflammation. *Nature* **600**, 707–712 (2021). <https://doi.org/10.1038/s41586-021-04136-4>

- Received: 11 October 2019
- Accepted: 14 October 2021

- Published: 01 December 2021
- Issue Date: 23 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04136-4>

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

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04136-4>

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

- Article
- [Published: 08 December 2021](#)

Gut microbiota modulates weight gain in mice after discontinued smoke exposure

- [Leviel Fluhr^{1 na1}](#),
- [Uria Mor ORCID: orcid.org/0000-0002-2918-6406^{1 na1}](#),
- [Aleksandra A. Kolodziejczyk¹](#),
- [Mally Dori-Bachash¹](#),
- [Avner Leshem ORCID: orcid.org/0000-0003-4082-2036^{1,2}](#),
- [Shlomik Itav¹](#),
- [Yotam Cohen¹](#),
- [Jotham Suez ORCID: orcid.org/0000-0003-3836-972X¹](#),
- [Niv Zmora^{1,3,4}](#),
- [Claudia Moresi¹](#),
- [Shahar Molina¹](#),
- [Niv Ayalon¹](#),
- [Rafael Valdés-Mas¹](#),
- [Shanni Hornstein¹](#),
- [Hodaya Karbi¹](#),
- [Denise Kviatcovsky ORCID: orcid.org/0000-0002-9697-0380¹](#),
- [Adi Livne¹](#),
- [Aurelie Bukimer¹](#),
- [Shimrit Eliyahu-Miller¹](#),
- [Alona Metz¹](#),
- [Alexander Brandis⁵](#),
- [Tevie Mehlman⁵](#),
- [Yael Kuperman ORCID: orcid.org/0000-0003-3830-1856⁶](#),

- [Michael Tsoory⁶](#),
- [Noa Stettner⁶](#),
- [Alon Harmelin⁶](#),
- [Hagit Shapiro](#) ^{1 na2} &
- [Eran Elinav](#) [ORCID: orcid.org/0000-0002-5775-2110](#)^{1,7 na2}

[*Nature*](#) volume **600**, pages 713–719 (2021)

- 9751 Accesses
- 2 Citations
- 261 Altmetric
- [Metrics details](#)

Subjects

- [Metabolic disorders](#)
- [Microbiome](#)

Abstract

Cigarette smoking constitutes a leading global cause of morbidity and preventable death¹, and most active smokers report a desire or recent attempt to quit². Smoking-cessation-induced weight gain (SCWG; 4.5 kg reported to be gained on average per 6–12 months, $>10\text{ kg year}^{-1}$ in 13% of those who stopped smoking³) constitutes a major obstacle to smoking abstinence⁴, even under stable^{5,6} or restricted⁷ caloric intake. Here we use a mouse model to demonstrate that smoking and cessation induce a dysbiotic state that is driven by an intestinal influx of cigarette-smoke-related metabolites. Microbiome depletion induced by treatment with antibiotics prevents SCWG. Conversely, fecal microbiome transplantation from mice previously exposed to cigarette smoke into germ-free mice naive to smoke exposure induces excessive weight gain across diets and mouse strains. Metabolically, microbiome-induced SCWG involves a concerted host and

microbiome shunting of dietary choline to dimethylglycine driving increased gut energy harvest, coupled with the depletion of a cross-regulated weight-lowering metabolite, *N*-acetyl glycine, and possibly by the effects of other differentially abundant cigarette-smoke-related metabolites. Dimethylglycine and *N*-acetyl glycine may also modulate weight and associated adipose-tissue immunity under non-smoking conditions.

Preliminary observations in a small cross-sectional human cohort support these findings, which calls for larger human trials to establish the relevance of this mechanism in active smokers. Collectively, we uncover a microbiome-dependent orchestration of SCWG that may be exploitable to improve smoking-cessation success and to correct metabolic perturbations even in non-smoking settings.

This is a preview of subscription content

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

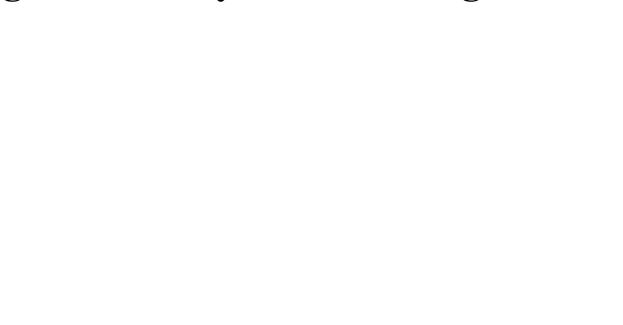
Fig. 1: Microbiome depletion modulates SCWG.



Fig. 2: Cigarette-smoke-induced dysbiosis drives SCWG.



Fig. 3: DMG dynamics during smoke exposure and cessation.



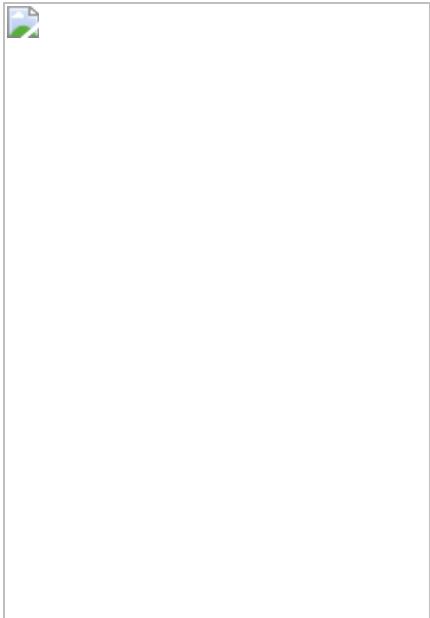
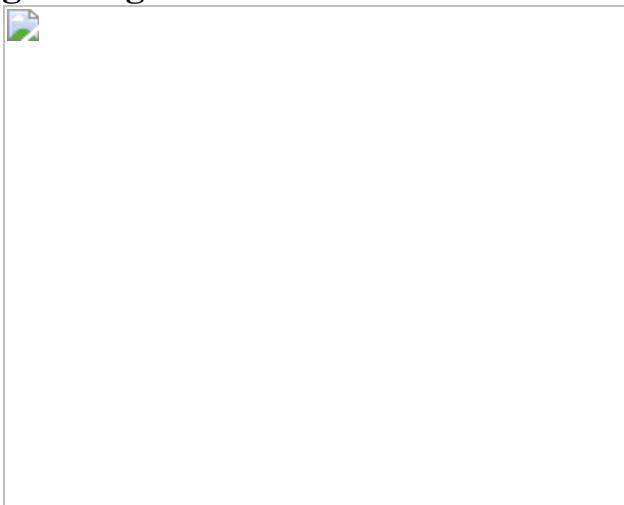


Fig. 4: Cigarette-smoke-related metabolites affect SCWG.



Data availability

Raw metagenomic sequences and resolved MAGs are available at the European Nucleotide Archive under project accession number [PRJEB40095](#). Single-cell RNA raw sequences are deposited in ArrayExpress with accession [E-MTAB-10869](#). Metabolomic data are provided in Supplementary Table 4 and in the publicly accessible platform of Mendeley at <https://doi.org/10.17632/539zh45tw2.1>. [Source data](#) are provided with this paper.

References

1. 1.

Centers for Disease Control and Prevention (CDC). Vital signs: current cigarette smoking among adults aged ≥ 18 years with mental illness—United States, 2009–2011. *MMWR Morb. Mortal Wkly Rep.* **62**, 81–87 (2013).

2. 2.

Centers for Disease Control and Prevention (CDC). Quitting smoking among adults—United States, 2001–2010. *MMWR Morb. Mortal Wkly Rep.* **60**, 1513–1519 (2011).

3. 3.

Harris, K. K., Zopey, M. & Friedman, T. C. Metabolic effects of smoking cessation. *Nat. Rev. Endocrinol.* **12**, 299–308 (2016).

4. 4.

Aubin, H. J., Farley, A., Lycett, D., Lahmek, P. & Aveyard, P. Weight gain in smokers after quitting cigarettes: meta-analysis. *BMJ* **345**, e4439 (2012).

5. 5.

Biedermann, L. et al. Smoking cessation induces profound changes in the composition of the intestinal microbiota in humans. *PLoS ONE* **8**, e59260 (2013).

6. 6.

Rodin, J. Weight change following smoking cessation: the role of food intake and exercise. *Addict. Behav.* **12**, 303–317 (1987).

7. 7.

Hankey, C. & Leslie, W. Obesity: is weight gain after smoking cessation an important concern? *Nat. Rev. Endocrinol.* **8**, 630–632 (2012).

8. 8.

Biedermann, L. et al. Smoking cessation alters intestinal microbiota: insights from quantitative investigations on human fecal samples using FISH. *Inflamm. Bowel Dis.* **20**, 1496–1501 (2014).

9. 9.

Lee, S. H. et al. Association between cigarette smoking status and composition of gut microbiota: population-based cross-sectional study. *J. Clin. Med.* **7**, 282 (2018).

10. 10.

Lim, M. Y. et al. Analysis of the association between host genetics, smoking, and sputum microbiota in healthy humans. *Sci. Rep.* **6**, 23745 (2016).

11. 11.

Russell, M. A., Wilson, C., Patel, U. A., Feyerabend, C. & Cole, P. V. Plasma nicotine levels after smoking cigarettes with high, medium, and low nicotine yields. *BMJ* **2**, 414–416 (1975).

12. 12.

Benowitz, N. L. & Jacob, P. 3rd Daily intake of nicotine during cigarette smoking. *Clin. Pharmacol. Ther.* **35**, 499–504 (1984).

13. 13.

Bush, T., Lovejoy, J. C., Deprey, M. & Carpenter, K. M. The effect of tobacco cessation on weight gain, obesity, and diabetes risk. *Obesity* **24**, 1834–1841 (2016).

14. 14.

Graff-Iversen, S., Hewitt, S., Forsen, L., Grotvedt, L. & Ariansen, I. Associations of tobacco smoking with body mass distribution; a population-based study of 65,875 men and women in midlife. *BMC Public Health* **19**, 1439 (2019).

15. 15.

Crooks, P. A., Bardo, M. T. & Dwoskin, L. P. Nicotinic receptor antagonists as treatments for nicotine abuse. *Adv. Pharmacol.* **69**, 513–551 (2014).

16. 16.

Turnbaugh, P. J. et al. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444**, 1027–1031 (2006).

17. 17.

Langille, M. G. et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat. Biotechnol.* **31**, 814–821 (2013).

18. 18.

Teng, Y. W., Mehedint, M. G., Garrow, T. A. & Zeisel, S. H. Deletion of betaine-homocysteine *S*-methyltransferase in mice perturbs choline and 1-carbon metabolism, resulting in fatty liver and hepatocellular carcinomas. *J. Biol. Chem.* **286**, 36258–36267 (2011).

19. 19.

Hill, D. A. et al. Distinct macrophage populations direct inflammatory versus physiological changes in adipose tissue. *Proc. Natl Acad. Sci. USA* **115**, E5096–E5105 (2018).

20. 20.

Jaitin, D. A. et al. Lipid-associated macrophages control metabolic homeostasis in a Trem2-dependent manner. *Cell* **178**, 686–698.e14 (2019).

21. 21.

Coats, B. R. et al. Metabolically activated adipose tissue macrophages perform detrimental and beneficial functions during diet-induced obesity. *Cell Rep.* **20**, 3149–3161 (2017).

22. 22.

Crewe, C., An, Y. A. & Scherer, P. E. The ominous triad of adipose tissue dysfunction: inflammation, fibrosis, and impaired angiogenesis. *J. Clin. Invest.* **127**, 74–82 (2017).

23. 23.

Flaherty, S. E. 3rd et al. A lipase-independent pathway of lipid release and immune modulation by adipocytes. *Science* **363**, 989–993 (2019).

24. 24.

Koeth, R. A. et al. Intestinal microbiota metabolism of l-carnitine, a nutrient in red meat, promotes atherosclerosis. *Nat. Med.* **19**, 576–585 (2013).

25. 25.

Witschi, H., Espiritu, I., Dance, S. T. & Miller, M. S. A mouse lung tumor model of tobacco smoke carcinogenesis. *Toxicol. Sci.* **68**, 322–330 (2002).

26. 26.

Bolyen, E. et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).

27. 27.

Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

28. 28.

Wood, D. E., Lu, J. & Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome Biol.* **20**, 257 (2019).

29. 29.

Buchfink, B., Xie, C. & Huson, D. H. Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* **12**, 59–60 (2015).

30. 30.

Delmont, T. O. et al. Nitrogen-fixing populations of Planctomycetes and Proteobacteria are abundant in surface ocean metagenomes. *Nat. Microbiol.* **3**, 804–813 (2018).

31. 31.

Li, D., Liu, C. M., Luo, R., Sadakane, K. & Lam, T. W. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* **31**, 1674–1676 (2015).

32. 32.

Eddy, S. R. Accelerated Profile HMM Searches. *PLoS Comput. Biol.* **7**, e1002195 (2011).

33. 33.

Kim, D., Song, L., Breitwieser, F. P. & Salzberg, S. L. Centrifuge: rapid and sensitive classification of metagenomic sequences. *Genome Res.* **26**, 1721–1729 (2016).

34. 34.

Alneberg, J. et al. Binning metagenomic contigs by coverage and composition. *Nat. Methods* **11**, 1144–1146 (2014).

35. 35.

Jain, C., Rodriguez, R. L., Phillippy, A. M., Konstantinidis, K. T. & Aluru, S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* **9**, 5114 (2018).

36. 36.

Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888-1902.e21 (2019).

37. 37.

Tabula Muris, C. et al. Single-cell transcriptomics of 20 mouse organs creates a *Tabula Muris*. *Nature* **562**, 367–372 (2018).

38. 38.

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).

39. 39.

Raudvere, U. et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* **47**, W191–W198 (2019).

Acknowledgements

We thank the members of the Elinav laboratory, Weizmann Institute of Science, and members of the DKFZ Cancer–Microbiome Division for insightful discussions; R. Alon, A. Bar-Shai and M. Hatzav for assistance

with operation of the smoking machine; M. Eren for fruitful discussions and guidance on making the MAG reference catalogue; and C. Bar-Natan for care of the GF animals and animal work. A.A.K. is a recipient of EMBO Long Term Fellowship number 2016-1088 and the European Union's Horizon 2020 research and innovation programme under Marie Skłodowska-Curie grant agreement number 747114. M.T. is the Incumbent of the Carolito Stiftung Research Fellow Chair in Neurodegenerative Diseases. Y.K. is the incumbent of the Sarah and Rolando Uziel Research Associate Chair. H.S. is the incumbent of the Vera Rosenberg Schwartz Research Fellow Chair. E.E. is supported by the Leona M. and Harry B. Helmsley Charitable Trust, the Adelis Foundation, the Pearl Welinsky Merlo Scientific Progress Research Fund, the Park Avenue Charitable Fund, The Hanna and Dr. Ludwik Wallach Cancer Research Fund, the Daniel Morris Trust, The Wolfson Family Charitable Trust & The Wolfson Foundation, the Ben B. and Joyce E. Eisenberg Foundation, the White Rose International Foundation, the Estate of Malka Moskowitz, the Estate of Myron H. Ackerman, the Estate of Bernard Bishin for the WIS-Clalit Program, the Else Kroener Fresenius Foundation, the Jeanne and Joseph Nissim Center for Life Sciences Research, Aliza Moussaieff, Miel de Botton, the Vainboim Family, Alex Davidoff, the V. R. Schwartz Research Fellow Chair, the Swiss Society Institute for Cancer Prevention Research at the Weizmann Institute of Science, Rehovot, Israel and by grants funded by the European Research Council, the Israel Science Foundation, the Israel Ministry of Science and Technology, the Israel Ministry of Health, the Helmholtz Foundation, the Garvan Institute, the European Crohn's and Colitis Organization, the Deutsch-Israelische Projektkooperation, the IDSA Foundation, the Sagol Institute for Longevity Research Program, the Charlie Teo Foundation, and the Wellcome Trust. E.E. is the incumbent of the Sir Marc and Lady Tania Feldmann Professorial Chair, is a senior fellow of the Canadian Institute of Advanced Research (CIFAR), and is an international scholar of The Bill & Melinda Gates Foundation and Howard Hughes Medical Institute (HHMI).

Author information

Author notes

1. These authors contributed equally: Leviel Fluhr, Uria Mor
2. These authors jointly supervised this work: Hagit Shapiro, Eran Elinav

Affiliations

1. Immunology Department, Weizmann Institute of Science, Rehovot, Israel

Leviel Fluhr, Uria Mor, Aleksandra A. Kolodziejczyk, Mally Dori-Bachash, Avner Leshem, Shlomik Itav, Yotam Cohen, Jotham Suez, Niv Zmora, Claudia Moresi, Shahar Molina, Niv Ayalon, Rafael Valdés-Mas, Shanni Hornstein, Hodaya Karbi, Denise Kviatcovsky, Adi Livne, Aurelie Bukimer, Shimrit Eliyahu-Miller, Alona Metz, Hagit Shapiro & Eran Elinav

2. Department of Surgery, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel

Avner Leshem

3. Research Center for Digestive Tract and Liver Diseases, Tel Aviv Sourasky Medical Center, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel

Niv Zmora

4. Internal Medicine Department, Tel Aviv Sourasky Medical Center, Tel Aviv, Israel

Niv Zmora

5. Department of Biological Services, Weizmann Institute of Science, Rehovot, Israel

Alexander Brandis & Tevie Mehlman

6. Department of Veterinary Resources, Weizmann Institute of Science,
Rehovot, Israel

Yael Kuperman, Michael Tsoory, Noa Stettner & Alon Harmelin

7. Division of Cancer-Microbiome Research, DKFZ, Heidelberg,
Germany

Eran Elinav

Contributions

L.F. and U.M. contributed equally to the study. L.F. designed, performed, interpreted the experiments and wrote the manuscript. U.M. designed and headed all computational and analytical aspects related to this study.

A.A.K., performed and analysed the single-cell transcriptomics analysis. M.D.-B., A. Leshem, S.I., S.H., H.K., D.K., A. Livne, C.M., S.M., N.A. and H.S. helped with experiments. Y.C., R.V.-M. and N.Z. assisted with data analysis. A. Brandis and T.M. performed the metabolomic experiments. Y.K. assisted with the metabolic cage experiments. M.T. assisted with the smoking-chamber experiments. N.S. and A.H. supervised all the GF experiments. J.S., A. Bukimer, S.E.-M. and A.M. designed and conducted the human trial, and assisted in the analysis of its results. H.S. and E.E. conceived the study, jointly supervised the participants, interpreted the experiments and wrote the manuscript.

Corresponding authors

Correspondence to [Hagit Shapiro](#) or [Eran Elinav](#).

Ethics declarations

Competing interests

E.E. is the scientific founder of Daytwo & BiomX, and a consultant to HELLO INSIDE in topics unrelated to the subject of this work. L.F., U.M., A.A.K., M.D.-B., A. Leshem, S.I., Y.C., J.S., N.Z., C.M., S.M., N.A., R.V.-

M., S.H., H.K., D.K., A. Livne, A. Brandis, S.E.-M., A.M., A. Bukimer, T.M., Y.K., M.T., N.S., A.H. and H.S. do not have any financial or non-financial competing interests. Weizmann Institute of Science is in the process of applying for a patent application covering weight modulatory effects of discovered microbes and metabolites that lists L.F., U.M. and E.E. as inventors.

Additional information

Peer review information *Nature* thanks the anonymous reviewer(s) for their contribution to the peer review of this work.

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 smoking cessation mouse model.

(a) Experimental scheme, HFD-fed mice. Mice not exposed to cigarette smoke (Non-SMK), mice exposed to smoke and cessation (SMK) and mice continuously exposed to smoke (continuous SMK). SMK group was exposed to cigarette smoke for 3 weeks, 5 days/week. Continuously smoke exposed group was exposed to cigarette-smoke throughout the duration of the experiment. **(b)** Weight change of HFD-fed mice during smoke exposure and following discontinuation of smoke exposure (cessation). Non-SMK, SMK and continuous SMK ($n = 10$), mean \pm s.e.m.; day 63: one-way ANOVA and Tukey correction. Inset: iAUC of weight change at smoke exposure or cessation, one-way ANOVA and Tukey correction, mean \pm s.e.m. **(c–e)** Ammonia **(c)**, cholesterol **(d)** and HDL **(e)** levels during smoke exposure and cessation: Non-SMK (at smoke exposure $n = 9$ for ammonia and $n = 10$ for d-e; at cessation $n = 8$ for all groups), SMK and continuous SMK (at smoke exposure $n = 10$, at cessation $n = 8$), mean values. Smoke exposure phase: two-sided Mann-Whitney *U*-test **(c, e)** or unpaired two-sided *t*-test **(d)**; cessation: Kruskal-Wallis and Dunn's test. In the smoke exposure phase, SMK and continuous SMK were considered as

one group. **(f–g)** Aspartate aminotransferase (AST, **f**) and alanine aminotransferase (ALT, **g**) levels during smoke exposure and cessation: Non-SMK, SMK, continuous SMK ($n = 10$ per group during smoke exposure; $n = 8$ per group during cessation); two-sided Mann-Whitney U -test for smoke exposure, Kruskal-Wallis and Dunn's test for cessation, mean values. **(h–j)** non-fasting glucose (**h**), triglycerides (**i**) and albumin (ALB, **j**) levels during smoke exposure and cessation: Non-SMK, SMK and continuous SMK ($n = 10$ per group during smoke exposure, $n = 8$ per group during cessation). Glucose: unpaired two-sided t -test for smoke exposure and one-way ANOVA with Tukey correction for cessation; triglycerides and albumin: unpaired two-sided t -test for smoke exposure and Kruskal-Wallis and Dunn's test for cessation. Plasma samples were collected 21 days during smoke exposure and on day 63 during cessation, mean values. Grey backgrounds represent cessation period. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$, exact P values presented in Supplementary Table 1.

Extended Data Fig. 2 Impact of microbiome depletion on SCWG.

(a) Lean (top) and fat (bottom) mass during smoke exposure (day 19) and cessation (day 35). Non-SMK ($n = 30$ at smoke exposure, $n = 6$ at cessation), SMK ($n = 38$ at smoke exposure, $n = 5$ at cessation), Non-SMK+abx (at smoke exposure $n = 20$, at cessation $n = 5$), SMK+abx (at smoke exposure $n = 15$, at cessation $n = 5$), HFD-fed mice, three-way during smoke exposure and cessation, mean values. **(b)** Weight change during smoke exposure and cessation. HFD-fed mice, Non-SMK and Non-SMK+abx ($n = 10$), SMK, SMK+abx and SMK+abx at cessation ($n = 9$), mean \pm s.e.m. For **(b–d)**: last day: one-way ANOVA and Sidak correction. Inset: iAUC, one-way ANOVA and Sidak correction, mean \pm s.e.m. **(c)** Weight change during smoke exposure and cessation in mice exposed to low-nicotine cigarette smoke and fed with HFD. Non-SMK, SMK and Non-SMK+abx ($n = 10$), SMK+abx ($n = 9$), mean \pm s.e.m. **(d)** Weight change during smoke exposure and cessation. HFD-fed mice were implanted with osmotic pumps containing vehicle (PBS) or 0.5 mg/kg/day varenicline (var). Non-SMK and Non-SMK+var ($n = 5$), SMK and SMK+var ($n = 10$), mean \pm s.e.m. Grey background in graphs refers to cessation period.

* $P<0.05$; ** $P<0.01$; *** $P<0.001$; **** $P<0.0001$, exact P values presented in Supplementary Table 1.

Extended Data Fig. 3 Metabolic impacts of antibiotics treatment on SCWG.

(a–c) Metabolic cage analysis during smoke exposure cessation over a period of 260h, $n=4$ for all mice groups, HFD-fed mice. Locomotion (**a**, ‘cnts’ refers to light beam break count), food intake (**b**, total kcal), respiratory exchange ratio (**c**, RER); inset: AUC (baseline = 0, baseline of RER = 0.7). One-way ANOVA and Sidak correction, mean \pm s.e.m. **(d)** Energy expenditure analysis during cessation over a period of 260h. Non-SMK and Non-SMK+abx ($n=8$), SMK ($n=5$), SMK+abx ($n=6$). Inset: AUC (baseline= minimum score, 14.1), One-way ANOVA and Sidak correction, mean \pm s.e.m. **(e)** Pearson correlation between fecal calories and weight gain rate, smoke exposure (circles) and cessation (triangles). SMK (smoke exposure period $n=8$, cessation period $n=9$, red), SMK+abx (smoke exposure period $n=9$, cessation period $n=9$, yellow). Lines represent linear regression slopes and intercepts. Grey background in graphs refers to the dark cycle. * $P<0.05$; ** $P<0.01$; exact P values presented in Supplementary Table 1.

Extended Data Fig. 4 Microbiome changes during smoke exposure and cessation.

(a–b) 16S rDNA-based principal coordinate analysis (PCoA) for **(a)** Bray-Curtis dissimilarity or **(b)** Jaccard similarity for HFD-fed Non-SMK and SMK mice at day 21 (smoke exposure). Non-SMK ($n=10$), SMK ($n=8$) mice. Inset: PERMANOVA. **(c)** KEGG modules fully covered by KEGG orthologs that significantly differed upon comparisons between day 0 and day 21. HFD-fed mice, Non-SMK ($n=10$ for all times), SMK ($n=10$ for day 0, $n=8$ for day 21); asterisks denote significant differences between day 21 versus day 0 for each group ($Q<0.1$); two-sided Mann-Whitney U -test with BKY correction. **(d)** Alpha diversity quantified as Shannon index in fecal samples at day 0 (baseline), 21 (smoke exposure) and 35 (cessation). HFD-fed mice, Non-SMK and SMK $n=10$ mice per group;

PERMANOVA, mean values. **(e–f)** PCA of species-level relative abundance at day 0 (**e**, baseline) and day 35 (**f**, cessation), HFD-fed mice $n = 10$ per group; inset: PERMANOVA. **(g)** Alpha diversity quantified as Shannon index in fecal samples at day 0 (baseline), 21 (smoke exposure) and 35 (cessation). HFD-fed mice, SMK ($n = 10$ for all times), SMK+abx ($n = 10$ at days 0 and 21, $n = 9$ at day 35); PERMANOVA, mean values. **(h)** Heatmap representing differentially abundant bacteria (genus level) in HFD-fed Non-SMK versus SMK mice at day 21 (smoke exposure). $n = 10$ mice per group; DESeq2; coloured bar represents z score. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$, exact P values presented in Supplementary Table 1.

Extended Data Fig. 5 Nicotine impacts on the microbiome and SCWG.

(a) Nicotine and cotinine levels in peripheral plasma, portal vein plasma and stool in naive SPF and GF mice. SPF ($n = 6$), GF ($n = 5$ for plasma and portal vein, $n = 6$ for stool); two-sided Mann-Whitney U test, mean values. **(b)** Plasma nicotine and cotinine levels during smoke exposure. Non-SMK ($n = 8$), SMK and SMK+abx ($n = 10$), Non-SMK+abx ($n = 9$), HFD-fed mice, Kruskal-Wallis and Dunn's test, mean values. **(c)** Fecal nicotine levels in HFD-fed mice exposed for 3 weeks to cigarettes smoke or to nicotine by intra-peritoneal (i.p.) injections; $n = 10$ mice per group; Kruskal-Wallis and Dunn's test, mean values. **(d)** PCoA depicting Jaccard similarity in mice exposed to cigarette smoke or to i.p. nicotine; Naive ($n = 9$), SMK ($n = 10$), nicotine-administered (i.p., $n = 8$) mice; pairwise PERMANOVA. **(e)** Weight change in mice exposed to cigarette smoke or to i.p. nicotine; Naïve and nicotine ($n = 9$), SMK ($n = 10$); day 21: one-way ANOVA and Tukey correction, mean values. **(f–g)** 16S rDNA-based PCoA depicting **(f)** Bray-Curtis dissimilarity and **(g)** Jaccard distance; Naive and SMK ($n = 10$), nicotine-administered via drinking water ($n = 9$), pairwise PERMANOVA. **(h)** Weight change in mice exposed for 3 weeks to cigarette smoke or to nicotine via drinking water. Naive and SMK ($n = 10$) nicotine ($n = 9$); day 21: one-way ANOVA and Tukey correction, mean values. **(i)** Plasma levels of nicotine and cotinine in HFD-fed mice administered with PBS or nicotine by osmotic pumps for 3 weeks, $n = 10$ mice per group; two-sided Mann-Whitney U -test, mean values. **(j)** Weight change in HFD-fed mice

administered with PBS or nicotine by osmotic pumps for 4 weeks. $n = 10$ mice per group; day 28: unpaired two-sided *t*-test, mean values. **P*<0.05; ***P*<0.01; ****P*<0.001; *****P*<0.0001, *P* values presented in Supplementary Table 1.

Extended Data Fig. 6 Fecal microbiome transplantation (FMT) from mice exposed to smoke.

(a) Experimental scheme of FMT experiments performed using fecal donations from mice exposed to smoke. Fecal samples from Non-SMK and SMK mice at day 7 (early smoke exposure) or day 18 (during smoke exposure) were transferred into GF mice of the designated strains, fed with the designated diets. **(b-d)** Weight change of GF mice following FMT. Fecal donations were collected at day 7 from HFD-fed C57BL/6 mice. For **b, c, d, f**: day 28, unpaired two-sided *t*-test, mean ± s.e.m.; upper inset: shotgun metagenomic sequencing showing PCA of species composition (for b Non-SMK $n = 5$, SMK $n = 7$ mice, for c, d, f $n = 6$ mice per group), PERMANOVA. **(b)** Recipients: HFD-fed SW mice. Non-SMK ($n = 11$), and SMK ($n = 13$) mice; day 28: unpaired two-sided *t*-test, mean ± s.e.m. **(c)** Recipients: HFD-fed C57BL/6 mice. Non-SMK ($n = 7$) and SMK ($n = 6$) mice, mean ± s.e.m. **(d)** Recipients: NC-fed C57BL/6 mice, $n = 6$ mice per group, mean ± s.e.m. **(e)** Experimental scheme of FMT experiments performed during cessation. Fecal samples from Non-SMK and SMK mice at day 24 (cessation) were transferred into GF mice of the designated strains, fed the designated diets. **(f)** Weight change of GF mice receiving FMT from C57BL/6 NC-fed mice at day 4 of cessation (total day 24). Recipients: NC-fed C57BL/6 mice, $n = 6$ mice per group, mean ± s.e.m. **(g)** Heatmap presenting log2 fold change of bacteria that were significantly changed ($Q < 0.05$) at least in two conditions (presented in Extended Data Fig. 6e); DESeq2; ns.: non-significant; coloured bar represents log fold change. **P*<0.05; ***P*<0.01, *****P*<0.0001, exact *P* values presented in Supplementary Table 1.

Extended Data Fig. 7 Functional features of transplanted microbiomes from mice exposed to smoke and at cessation.

(a) Heatmaps describing the functional potential of metagenomic features (EC annotated) in shotgun metagenomics, corresponding to the conditions of Extended Data Fig. 6e. Features were grouped according to BRITE hierarchy; colored bar represents *z* score. **(b–c)** Weight change of GF C57BL/6 mice following FMT. Fecal donations were collected at day 24 of smoke exposure and cessation from abx-treated or untreated HFD-fed C57BL/6 mice (**b**) or NC (**c**); Day 28, unpaired two-sided *t*-test, mean ± s.e.m. **(b)** Recipients: HFD-fed C57BL/6 mice, SMK (*n* = 7) and SMK+abx (*n* = 6), mean ± s.e.m. **(c)** Recipients: NC-fed C57BL/6 mice, *n* = 6 per group, mean ± s.e.m. **(d–e)** Fecal calories of GF recipients after FMT from early smoke exposure (**d**, Non-SMK *n* = 5, SMK *n* = 6) or cessation (**e**, *n* = 5 mice per group); unpaired two-sided *t*-test, mean values. **P*<0.05, ***P*<0.01, *P* values presented in Supplementary Table 1.

Extended Data Fig. 8 Candidate metabolites associated with SCWG.

(a) Heatmaps depicting plasma metabolites of HFD-fed mice during smoke exposure (day 15) and cessation (day 30). Non-SMK, SMK+abx (*n* = 7 smoke exposure, *n* = 6 cessation), SMK (*n* = 6 smoke exposure, *n* = 5 cessation), Non-SMK+abx (*n* = 8 smoke exposure, *n* = 6 cessation); heatmap intensities represent normalized metabolite levels (*z*-score). **(b)** Heatmaps depicting fecal metabolites during smoke exposure (day 21) and cessation (day 35). Non-SMK and Non-SMK+abx (*n* = 9 for smoke exposure and *n* = 6 for cessation), SMK (*n* = 8 for smoke exposure and *n* = 5 for cessation) and SMK+abx (*n* = 8 for smoke exposure and *n* = 6 for cessation) HFD-fed mice; heatmap intensities represent normalized metabolite levels (*z*-score). **(c)** LMM, utilizing both smoke and antibiotics alterations throughout time and quantifying their interactive impacts on plasma metabolite levels at smoke exposure. Venn diagram (left) represents significant (*p*<0.05) metabolites impacted by smoke exposure, antibiotics and their interactions. Post-hoc: Tukey correction for multiple comparisons. Venn diagram (right): metabolites fulfilling both mixed linear model and statistical hypothesis testing criteria. **(d)** Predicted gene copy numbers (PICRUSt2) of fecal choline sulfatase: Non-SMK (*n* = 10), SMK (*n* = 8) HFD-fed mice; unpaired two-sided *t*-test; mean values. **(e–f)** Fecal metabolites, untargeted mass spectrometry, HFD-fed mice, Non-SMK (*n* =

9), SMK ($n = 8$ at smoke exposure), unpaired two-sided t -test, mean \pm s.e.m. values of (e) choline and (f) betaine; SmD- scaled imputed data, log scale values (methods). (g) qPCR quantification of liver betaine-homocysteine S-methyltransferase (BHMT) transcripts during smoke exposure. HFD-fed mice, Non-SMK ($n = 4$), SMK ($n = 6$); two-sided Mann-Whitney U -test, mean \pm s.e.m. (h) Plasma DMG levels, assessed by targeted mass spectrometry. Non-SMK, Non-SMK+abx and SMK+abx ($n = 10$), SMK ($n = 9$), mice fed with HFD; Kruskal-Wallis and Dunn's test, mean values. * $P < 0.05$, ** $P < 0.01$, exact P values presented in Supplementary Table 1.

Extended Data Fig. 9 Metabolite impacts on SCWG.

(a–b) Weight change during smoke exposure and cessation upon addition of N-formylanthranilic acid (a, N-FAN acid) or trigonelline (b, TRI). a: $n = 10$ HFD-fed mice per group, mean \pm s.e.m; day 42: unpaired 2-sided t test; iAUC: unpaired 2-sided t test, mean \pm s.e.m. b: Non-SMK+abx+PBS, Non-SMK+abx+TRI and SMK+abx+TRI ($n = 10$), SMK+abx+PBS ($n = 9$), HFD-fed mice, mean \pm s.e.m; day 35: one-way ANOVA Sidak correction; inset: iAUC; one-way ANOVA Sidak correction, mean \pm s.e.m. (c) iAUC of weight change at smoke exposure or cessation of Fig 4a. NS+abx ($n = 9$), SMK+abx+PBS ($n = 10$ smoke exposure, $n = 9$ cessation), all other groups $n = 10$; one-way ANOVA Sidak correction; box plots represent max-min values. Rx- treatment. (d) Plasma levels of metabolites of the DMG biosynthesis pathway assessed by targeted mass spectrometry in mice consuming choline-deficient diet (CDD), $n = 10$; two-sided Mann-Whitney U -test, mean values. (e) Hepatic BHMT qPCR during smoke exposure in mice consuming CDD, $n = 10$; unpaired two-sided t -test, mean values. (f) Targeted mass spectrometry of plasma DMG. PBS ($n = 9$), DMG ($n = 10$) HFD-fed mice; two-sided Mann-Whitney U -test, mean values. (g–h) Weight change during DMG supplementation on day 14 (g, $n = 10$) or day 49 (h, PBS: $n = 19$, DMG: $n = 18$, pool of 2 independent repeats), mean \pm s.e.m., HFD-fed mice; last time point: unpaired two-sided t -test. Inset: iAUC, unpaired two-sided t -test, mean \pm s.e.m. (i) Plasma levels of HFD-fed mice treated with PBS ($n = 8$ for all parameters except $n = 7$ for NH₃) or DMG ($n = 10$ for all parameters except $n = 9$ for ALT); unpaired two-sided t -test, except two-sided Mann Whitney U test for NH₃ and HDL;

mean values. **(j–m)** Metabolic cage analysis for 172h in DMG versus PBS supplemented mice. Locomotion (**j**, $n = 8$), respiratory exchange ratio (**k**, RER, $n = 4$), energy expenditure (**l**, $n = 4$) and food intake (Total kcal, **m**, $n = 8$). Inset: AUC; unpaired two-sided *t*-test, mean \pm s.e.m. Grey background depicts the cessation period (**a–c**) or the dark cycle (**j–i**). * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$; **** $P < 0.0001$, *P* values presented in Supplementary Table 1.

Extended Data Fig. 10 ACG supplementation ameliorates weight gain.

(a) Plasma levels of ACG during smoke exposure (day 15) and cessation (day 30). Non-SMK ($n = 7$ for smoke exposure, $n = 5$ for cessation), SMK ($n = 6$ for smoke exposure, $n = 5$ for cessation), Non-SMK+abx ($n = 8$ for smoke exposure, $n = 6$ for cessation), SMK+abx ($n = 7$ for smoke exposure, $n = 6$ for cessation), HFD-fed mice; two-way ANOVA and BH correction ($Q < 0.1$), mean values. **(b)** Calories measured in rodent diet. $n = 3$ mice per group; one-way ANOVA and Tukey correction, mean values. **(c)** Plasma levels of ACG in HFD (Control, $n = 10$) or HFD+ACG (ACG, $n = 9$) mice, assessed by targeted mass spectrometry. Unpaired two-sided *t*-test, mean values. **(d)** Weight change during ACG supplementation. $n = 5$ HFD-fed mice per group, mean \pm s.e.m.; day 14: unpaired two-sided *t*-test. Inset: iAUC, unpaired two-sided *t*-test, mean \pm s.e.m. **(e)** GTT during ACG administration (day 49) $n = 10$ HFD-fed mice per group. Inset: AUC, Unpaired two-sided *t*-test, mean \pm s.e.m. **(f)** Plasma biochemical levels Control ($n = 19$ except $n = 17$ for TG & urea, $n = 18$ for albumin). ACG ($n = 19$ except for $n = 18$ for NH3) HFD-fed mice. Two-sided Mann-Whitney *U*-test, except for unpaired two-sided *t*-test for glucose, total cholesterol, urea and total protein, mean values. **(g–j)** Metabolic cage analysis for 384h, $n = 4$ HFD-fed mice per group, dashed line indicates the starting day of the diet. **(g)** Food intake (total kcal); **(h)** locomotion ('cnts' refers to light beam break count); **(i)** respiratory exchange ratio (RER); and **(j)** energy expenditure; inset: AUC, unpaired two-sided *t*-test except two-sided Mann-Whitney *U*-test for energy expenditure, mean \pm s.e.m. Grey background depicts the dark cycle. **(k)** Plasma DMG levels in Control ($n = 7$) or ACG ($n = 6$) mice, unpaired two-sided *t*-test, mean values. * $P < 0.05$; ** $P < 0.01$;

*** $P<0.001$; **** $P<0.0001$, exact P values presented in Supplementary Table 1.

Extended Data Fig. 11 Single cell transcriptomic analysis of adipose tissue immune cells of HFD-fed, DMG- and ACG-supplemented mice not exposed to smoke.

(a) UMAP representation of single cell data showing stepwise clustering of adipose immune cell populations. **(b)** Balloonplot showing the mean of normalized and scaled gene expression of key markers used to annotate clusters to cell types. **(c)** Balloonplot showing the mean of normalized and scaled gene expression of key markers of macrophage and monocyte populations. **(d)** Balloonplot showing the mean of normalized and scaled gene expression of key markers of chemokines and cytokines in macrophage and monocyte populations. A larger size of the dot in the balloonplot represent increased expression. See Supplementary Table 6 for single cell transcriptomic data.

Extended Data Fig. 12 Adipose tissue immune cell features of HFD-fed, DMG- and ACG-supplemented mice not exposed to smoke.

(a) Heatmap showing -log10 of BH adjusted p-value GO functional enrichment analysis of genes differentially expressed in macrophage subpopulations upon treatment with ACG and DMG. See Supplementary Table 6 for single cell transcriptomic data.

Extended Data Fig. 13 Potential microbiome and metabolite associations with human who smoke.

(a) Experimental outline of the human cohort. **(b-f)** human fecal microbiome analysis (Non-SMK $n=40$, SMK $n=20$). **(b)** PCA of metagenomically-assembled genomes (MAGs) relative abundances in human stool; inset: PERMANOVA. **(c)** Differentially abundant MAGs; asterisks denote significant differences ($p<0.05$); two-sided Mann-Whitney U -test. **(d)** PCA of KO annotated reads; inset: PERMANOVA. **(e)** KEGG

orthologs of highest effect (feature with negative values enriched in SMK), two-sided Mann-Whitney *U*-test, highest (and lowest), 2-fold log change. **(f)** ROC curves for binary classifier (methods). **(g-i)** Targeted mass spectrometry of metabolites from the DMG biosynthesis pathway: choline (**g**), betaine (**h**) and DMG (**i**); Non-SMK ($n = 62$), SMK ($n = 34$); two-sided Mann-Whitney *U*-test, mean values. * $P < 0.05$; ** $P < 0.01$, exact P values presented in Supplementary Table 1, raw data presented in Supplementary Table 7.

Supplementary information

Reporting Summary

Supplementary information 1

Absolute weight panels.

Supplementary information 2

Individual repeats of pooled main figure panels.

Supplementary Table 1

P values and mixed linear model.

Supplementary Table 2

16S rDNA.

Supplementary Table 3

Shotgun metagenomic.

Supplementary Table 4

Metabolomics.

Supplementary Table 5

Candidate metabolites identified by LMM.

Supplementary Table 6

Single-cell transcriptomics.

Supplementary Table 7

Human metadata and shotgun metagenomics.

Source data

[**Source Data Fig. 1**](#)

[**Source Data Fig. 2**](#)

[**Source Data Fig. 3**](#)

[**Source Data Fig. 4**](#)

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Fluhr, L., Mor, U., Kolodziejczyk, A.A. *et al.* Gut microbiota modulates weight gain in mice after discontinued smoke exposure. *Nature* **600**, 713–719 (2021). <https://doi.org/10.1038/s41586-021-04194-8>

- Received: 25 August 2020

- Accepted: 28 October 2021
- Published: 08 December 2021
- Issue Date: 23 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04194-8>

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

Further reading

- [**Smoking, dysbiosis and weight gain**](#)
 - Ursula Hofer

Nature Reviews Microbiology (2021)
- [**Gut clues to weight gain after quitting smoking**](#)
 - Matthew P. Spindler
 - Jeremiah J. Faith
 - Paul J. Kenny

Nature (2021)

[**Gut clues to weight gain after quitting smoking**](#)

- Matthew P. Spindler
- Jeremiah J. Faith
- Paul J. Kenny

News & Views Forum 08 Dec 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04194-8>

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

- Article
- [Published: 08 December 2021](#)

A hormone complex of FABP4 and nucleoside kinases regulates islet function

- [Kacey J. Prentice](#)¹,
- [Jani Saksi](#)¹,
- [Lauren T. Robertson](#)¹,
- [Grace Y. Lee](#)¹,
- [Karen E. Inouye](#)¹,
- [Kosei Eguchi](#)¹,
- [Alexandra Lee](#)¹,
- [Ozgur Cakici](#)¹,
- [Emily Otterbeck](#)¹,
- [Paulina Cedillo](#)¹,
- [Peter Achenbach](#) [ORCID: orcid.org/0000-0001-6720-2684](#)²,
- [Anette-Gabriele Ziegler](#)²,
- [Ediz S. Calay](#)¹,
- [Feyza Engin](#) [ORCID: orcid.org/0000-0002-7987-5381](#)^{1,3} &
- [Gökhan S. Hotamisligil](#) [ORCID: orcid.org/0000-0003-2906-1897](#)^{1,4}

[Nature](#) volume 600, pages 720–726 (2021)

- 9177 Accesses
- 750 Altmetric
- [Metrics details](#)

Subjects

- [Extracellular signalling molecules](#)
- [Mechanisms of disease](#)
- [Peptide hormones](#)
- [Type 1 diabetes](#)
- [Type 2 diabetes](#)

Abstract

The liberation of energy stores from adipocytes is critical to support survival in times of energy deficit; however, uncontrolled or chronic lipolysis associated with insulin resistance and/or insulin insufficiency disrupts metabolic homeostasis^{1,2}. Coupled to lipolysis is the release of a recently identified hormone, fatty-acid-binding protein 4 (FABP4)³. Although circulating FABP4 levels have been strongly associated with cardiometabolic diseases in both preclinical models and humans^{4,5,6,7}, no mechanism of action has yet been described^{8,9,10}. Here we show that hormonal FABP4 forms a functional hormone complex with adenosine kinase (ADK) and nucleoside diphosphate kinase (NDPK) to regulate extracellular ATP and ADP levels. We identify a substantial effect of this hormone on beta cells and given the central role of beta-cell function in both the control of lipolysis and development of diabetes, postulate that hormonal FABP4 is a key regulator of an adipose–beta-cell endocrine axis. Antibody-mediated targeting of this hormone complex improves metabolic outcomes, enhances beta-cell function and preserves beta-cell integrity to prevent both type 1 and type 2 diabetes. Thus, the FABP4–ADK–NDPK complex, Fabkin, represents a previously unknown hormone and mechanism of action that integrates energy status with the function of metabolic organs, and represents a promising target against metabolic disease.

This is a preview of subscription content

Access options

[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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Fig. 1: Targeting FABP4 increases beta-cell mass and function.

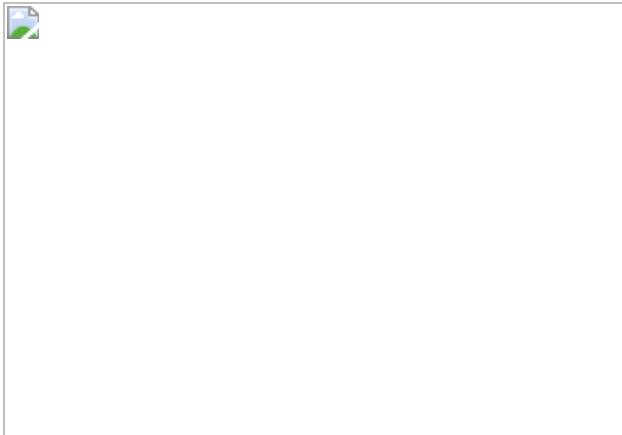


Fig. 2: Circulating FABP4 forms a hormonal complex with NDPK and ADK to regulate extracellular nucleosides.

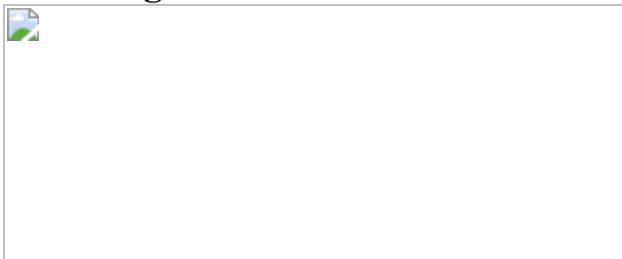


Fig. 3: Fabkin inhibits GSIS and is neutralized by a-Ab.

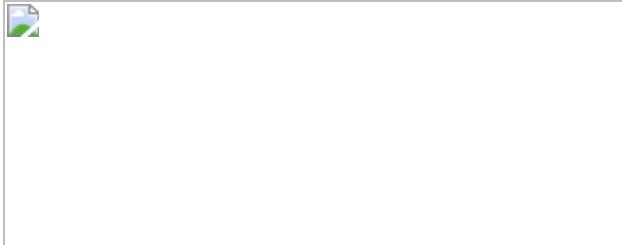
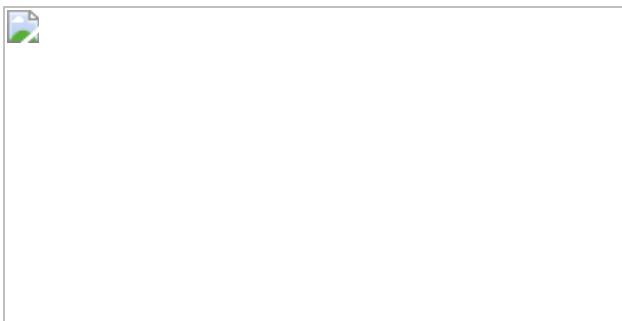


Fig. 4: Fabkin alters beta-cell calcium dynamics and promotes cell death.



Data availability

The datasets generated during the current study are available in the FigShare repository at <https://doi.org/10.6084/m9.figshare.16547853>. Source data are provided with this paper.

References

1. 1.
Jensen, M. D., Caruso, M., Heiling, V. & Miles, J. M. Insulin regulation of lipolysis in nondiabetic and IDDM subjects. *Diabetes* **38**, 1595–1601 (1989).
2. 2.
Kim, J. Y. et al. Increased lipolysis, diminished adipose tissue insulin sensitivity, and impaired beta-cell function relative to adipose tissue insulin sensitivity in obese youth with impaired glucose tolerance. *Diabetes* **66**, 3085–3090 (2017).
3. 3.
Cao, H. et al. Adipocyte lipid chaperone AP2 is a secreted adipokine regulating hepatic glucose production. *Cell Metab.* **17**, 768–778 (2013).
4. 4.
Prentice, K. J., Saksi, J. & Hotamisligil, G. S. Adipokine FABP4 integrates energy stores and counterregulatory metabolic responses. *J. Lipid Res.* **60**, 734–740 (2019).
5. 5.
Saksi, J. et al. Low-expression variant of fatty acid-binding protein 4 favors reduced manifestations of atherosclerotic disease and increased plaque stability. *Circ. Cardiovasc. Genet* **7**, 588–598 (2014).

6. 6.

Tuncman, G. et al. A genetic variant at the fatty acid-binding protein aP2 locus reduces the risk for hypertriglyceridemia, type 2 diabetes, and cardiovascular disease. *Proc. Natl Acad. Sci. USA* **103**, 6970–6975 (2006).

7. 7.

Zhao, W. et al. Identification of new susceptibility loci for type 2 diabetes and shared etiological pathways with coronary heart disease. *Nat. Genet.* **49**, 1450–1457 (2017).

8. 8.

Ertunc, M. E. et al. Secretion of fatty acid binding protein aP2 from adipocytes through a nonclassical pathway in response to adipocyte lipase activity. *J. Lipid Res.* **56**, 423–434 (2015).

9. 9.

Schlottmann, I., Ehrhart-Bornstein, M., Wabitsch, M., Bornstein, S. R. & Lamounier-Zepter, V. Calcium-dependent release of adipocyte fatty acid binding protein from human adipocytes. *Int. J. Obes.* **38**, 1221–1227 (2014).

10. 10.

Villeneuve, J. et al. Unconventional secretion of FABP4 by endosomes and secretory lysosomes. *J. Cell Biol.* **217**, 649–665 (2018).

11. 11.

Burak, M. F. et al. Development of a therapeutic monoclonal antibody that targets secreted fatty acid-binding protein aP2 to treat type 2 diabetes. *Sci. Transl. Med.* **7**, 319ra205 (2015).

12. 12.

Thumer, L. et al. German new onset diabetes in the young incident cohort study: DiMelli study design and first-year results. *Rev. Diabet. Stud.* **7**, 202–208 (2010).

13. 13.

Warncke, K. et al. Does diabetes appear in distinct phenotypes in young people? Results of the diabetes mellitus incidence Cohort Registry (DiMelli). *PLoS ONE* **8**, e74339 (2013).

14. 14.

Ziegler, A. G., Hummel, M., Schenker, M. & Bonifacio, E. Autoantibody appearance and risk for development of childhood diabetes in offspring of parents with type 1 diabetes: the 2-year analysis of the German BABYDIAB Study. *Diabetes* **48**, 460–468 (1999).

15. 15.

Xiao, Y. et al. Fatty acid binding protein 4 promotes autoimmune diabetes by recruitment and activation of pancreatic islet macrophages. *JCI Insight* **6**, e141814 (2021).

16. 16.

Kralisch, S. et al. Circulating adipocyte fatty acid-binding protein induces insulin resistance in mice in vivo. *Obesity* **23**, 1007–1013 (2015).

17. 17.

Wu, L. E. et al. Identification of fatty acid binding protein 4 as an adipokine that regulates insulin secretion during obesity. *Mol. Metab.* **3**, 465–473 (2014).

18. 18.

Annes, J. P. et al. Adenosine kinase inhibition selectively promotes rodent and porcine islet beta-cell replication. *Proc. Natl Acad. Sci. USA* **109**, 3915–3920 (2012).

19. 19.

Navarro, G. et al. Genetic disruption of adenosine kinase in mouse pancreatic beta-cells protects against high-fat diet-induced glucose intolerance. *Diabetes* **66**, 1928–1938 (2017).

20. 20.

Furuhashi, M. et al. Treatment of diabetes and atherosclerosis by inhibiting fatty-acid-binding protein aP2. *Nature* **447**, 959–965 (2007).

21. 21.

Khan, S. et al. Autocrine activation of P2Y1 receptors couples Ca^{2+} influx to Ca^{2+} release in human pancreatic beta-cells. *Diabetologia* **57**, 2535–2545 (2014).

22. 22.

Leon, C. et al. The P2Y1 receptor is an ADP receptor antagonized by ATP and expressed in platelets and megakaryoblastic cells. *FEBS Lett.* **403**, 26–30 (1997).

23. 23.

Detimary, P., Jonas, J. C., & Henquin, J. C. Stable and diffusible pools of nucleotides in pancreatic islet cells. *Endocrinology* **137**, 4671–4676 (1996).

24. 24.

Bauer, C. et al. ATP mediates a negative autocrine signal on stimulus-secretion coupling in mouse pancreatic beta-cells. *Endocrine* **63**, 270–283 (2019).

25. 25.

Baurand, A. et al. Inhibition of platelet function by administration of MRS2179, a P2Y1 receptor antagonist. *Eur. J. Pharmacol.* **412**, 213–221 (2001).

26. 26.

Seo, J. B., Jung, S. R., Hille, B. & Koh, D. S. Extracellular ATP protects pancreatic duct epithelial cells from alcohol-induced damage through P2Y1 receptor-cAMP signal pathway. *Cell Biol. Toxicol.* **32**, 229–247 (2016).

27. 27.

Engin, F. et al. Restoration of the unfolded protein response in pancreatic beta-cells protects mice against type 1 diabetes. *Sci. Transl. Med.* **5**, 211ra156 (2013).

28. 28.

Engin, F., Nguyen, T., Yermalovich, A. & Hotamisligil, G. S. Aberrant islet unfolded protein response in type 2 diabetes. *Sci. Rep.* **4**, 4054 (2014).

29. 29.

Burnstock, G., Vaughn, B. & Robson, S. C. Purinergic signalling in the liver in health and disease. *Purinergic Signal.* **10**, 51–70 (2014).

30. 30.

Burnstock, G. Purinergic signaling in the cardiovascular system. *Circ. Res.* **120**, 207–228 (2017).

31. 31.

Liu, Y. et al. Rapid Elevation in CMPF may act as a tipping point in diabetes development. *Cell Rep.* **14**, 2889–2900 (2016).

32. 32.

Batchuluun, B. et al. Elevated medium-chain acylcarnitines are associated with gestational diabetes mellitus and early progression to type 2 diabetes and induce pancreatic beta-cell dysfunction. *Diabetes* **67**, 885–897 (2018).

33. 33.

Prentice, K. J. et al. The furan fatty acid metabolite CMPF is elevated in diabetes and induces beta-cell dysfunction. *Cell Metab.* **19**, 653–666 (2014).

34. 34.

Rezania, A. et al. Reversal of diabetes with insulin-producing cells derived in vitro from human pluripotent stem cells. *Nat. Biotechnol.* **32**, 1121–1133 (2014).

35. 35.

Potel, C. M., Facci, D. & Heck, A. J. R. Mix and match of the tumor metastasis suppressor Nm23 protein isoforms in vitro and in vivo. *FEBS J.* **285**, 2856–2868 (2018).

36. 36.

Han, B. G., Min, K., Lee, B. I. & Lee, S. Refined structure of human NM23-H1 from a hexagonal crystal. *Bull. Korean Chem. Soc.* **31**, 1397–1399 (2010).

37. 37.

Prentice, K. J. et al. CMPF, a metabolite formed upon prescription omega-3-acid ethyl ester supplementation, prevents and reverses steatosis. *EBioMedicine* **27**, 200–213 (2018).

38. 38.

Basford, C. L. et al. The functional and molecular characterisation of human embryonic stem cell-derived insulin-positive cells compared with adult pancreatic beta-cells. *Diabetologia* **55**, 358–371 (2012).

Acknowledgements

We thank members of the Hotamisligil Laboratory and the Sabri Ülker Center, past and present, for their contributions to our understanding of FABP4 and their helpful discussions. We thank staff at the Islet Core and Clinical Islet Laboratory (Alberta Islet Distribution Program, University of Alberta) for providing us with human islets from review-board-approved donors. UCB generated and purified the antibodies used in this study and independently reproduced the interaction data between NDPK, FABP4 and a-Ab. The Hotamisligil laboratory is supported by grants from the National Institutes of Health (NIH DK123458) and the Juvenile Diabetes Research Foundation (JDRF; 2-SRA-2019-660-S-B). A JDRF Postdoctoral Fellowship (3-PDF-2017-400-A-N) funds K.J.P. The Sigrid Juselius Foundation, the Finnish Foundation for Cardiovascular Research and the Otto A. Malm Foundation support J.S. F.E. was supported by grants from the JDRF (JDRF-5-CDA-2014-184-A-N) and the NIH (NIH 5K01DK102488-03).

Author information

Affiliations

1. Sabri Ülker Center for Metabolic Research, Harvard T. H. Chan School of Public Health, Department of Molecular Metabolism, Boston, MA, USA

Kacey J. Prentice, Jani Saksi, Lauren T. Robertson, Grace Y. Lee, Karen E. Inouye, Kosei Eguchi, Alexandra Lee, Ozgur Cakici, Emily Otterbeck, Paulina Cedillo, Ediz S. Calay, Feyza Engin & Gökhan S. Hotamisligil

2. Institute of Diabetes Research, Helmholtz Zentrum Munchen, German Research Center for Environmental Health, Munich-Neuherberg, Germany

Peter Achenbach & Anette-Gabriele Ziegler

3. Departments of Biomolecular Chemistry and Medicine, Division of Endocrinology, Diabetes and Metabolism, University of Wisconsin–Madison School of Medicine and Public Health, Madison, WI, USA

Feyza Engin

4. Broad Institute of Harvard and MIT, Cambridge, MA, USA

Gökhan S. Hotamisligil

Contributions

K.J.P. designed and performed the in vitro and in vivo experiments, analysed the data, prepared the figures and wrote and revised the manuscript. F.E., K.E.I., A.L., P.C. and L.T.R. designed and performed the in vivo experiments, analysed data and revised the manuscript. J.S., G.Y.L., K.E., E.O. and E.S.C. designed and performed the in vitro experiments, analysed data and revised the manuscript. O.C. analysed the crystal structure and generated recombinant proteins used in this study, analysed the data and revised the manuscript. P.A. and A.-G.Z. collected and analysed human serum samples and revised the manuscript. G.S.H. conceived, supervised and supported the project, designed experiments, interpreted results and revised the manuscript.

Corresponding author

Correspondence to [Gökhan S. Hotamisligil](#).

Ethics declarations

Competing interests

The Hotamisligil Lab has generated intellectual property (assigned to Harvard University) related to hormonal FABP4 and its therapeutic targeting and receives funding for this project from Lab1636, LLC, an affiliate of Deerfield Management. G.S.H. is on the Scientific Advisory Board of Crescinta Pharmaceuticals and holds equity. Other authors have no conflicts of interest to declare.

Additional information

Peer review information *Nature* thanks Matthias Hebrok, Andrew Stewart, Thomas Wieland and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

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 FABP4-/- mice have no difference in alpha cell mass, and FABP4 is not expressed in beta cells.

(a) Dithizone-stained islets in vivo from FABP4^{-/-} mice (Representative image from N = 3). (b) Immunohistochemical staining for glucagon of pancreatic sections from 7-week-old WT or FABP4^{-/-} mice, and (c) quantification of percentage of glucagon positive area per total pancreatic area based on IHC (N = 5/group). (d) Immunofluorescent staining for insulin (green), FABP4 (red) and nuclei (DAPI, blue) in primary isolated mouse islets (N = 15 islets). (e) Western blot for FABP4 and B-Tubulin (B-Tubb) from INS1 cells with and without treatment with FABP4 (N = 3). Baseline matching parameters for 10 week old WT female NOD mice prior to the initiation of dosing showing no difference in (f) body weight, (g) blood glucose, (h) plasma insulin, or (i) plasma FABP4 (N = 46/group). (j) 6 h fasting blood glucose, (k) body weight, and (l) plasma insulin levels among NOD mice that remained non-diabetic for the duration of the treatment period (N = 11 (PBS), 23 (a-Ab), 13 (c-Ab)). (m) Immunofluorescent staining for insulin (green) and glucagon (red) in

pancreatic sections from NOD mice treated with a-Ab or c-Ab for 5 wks (N = 4 mice/group). Scale bars are 100um. Data are mean +/- SEM. Two-tailed unpaired t-test (c); One-way ANOVA (f-i); Two-way ANOVA (j-l).

[Source data](#)

Extended Data Fig. 2 FABP4 exerts differential activity *in vivo* and *in vitro*, and only interacts with a-Ab through the light chain.

(a) GSIS from islets from FABP4^{-/-} mice treated with increasing doses of FABP4 (N = 3). (b) Plasma FABP4 levels following acute injection of 10 µg FABP4 (N = 3 (PBS), 6 (FABP4)). (c) Blood glucose ($P = 0.0300$) and (d) plasma insulin levels 20 min post-injection ($P = 0.0160$) with FABP4 or PBS (N = 3 (PBS), 6 (FABP4)).(e) Crystal structure showing FABP4 binding through the light chain of a-Ab. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. Data are presented as mean +/- SEM. Two-way ANOVA (a,b); Welch's t-test (c,d).

[Source data](#)

Extended Data Fig. 3 Determination of the Binding Epitope for a-Ab on NDK-A.

(a) Amino acid sequence for human NDK-A, and indication of the location of each peptide generated for epitope mapping (red). Each peptide is 15 amino acids long with an N-terminal 6xHis Tag for peptide labeling. Each peptide has a 5 amino acid overlap with the preceding peptide sequentially from the N-terminus of the protein.(b) Table summarizing binding affinity as determined by MicroScale Thermophoresis (MST). Low affinity binding was observed between peptide 2 and peptide 3 with a-Ab. High affinity binding was observed between peptide 8 and a-Ab, comparable to full-length NDK-A protein. No binding was detected (NBD) between a-Ab and any other peptide examined. The lack of binding between peptide 7 or peptide 9 and a-Ab indicated that the central 5 amino acids of peptide 8, non-overlapping in sequence with either peptide 7 or 9, is likely the primary epitope for a-Ab binding. (c) Example MST binding

curves for full-length NDPK-A, NDPK-A peptide 9 and NDPK-A peptide 8 with a-Ab. No binding is observed between NDPK-A peptide 9 and a-Ab. (d) Crystal structure of human NDPK-A in hexamer conformation (source PDB: 3L7U) with the potential binding region of peptide 8 highlighted in pink and peptide 3 highlighted in blue. The protein structure indicates the potential binding region of peptide 8 is on the surface of the protein, amenable for a-Ab interaction. The tertiary folding of NDPK-A places peptide 3 and peptide 8 in close proximity, suggesting binding of a-Ab may be primarily occurring through residues in peptide 8, and partially facilitated through interaction with residues in peptides 2 and 3. EC50 calculated by Hill Slope.

Extended Data Fig. 4 a-Ab interacts with Fabkin.

(a–e) Representative mesoscale thermophoresis experiments showing 1:1 protein interactions between complex components and a-Ab ($N = 6$ /interaction). EC50 was calculated using Hill Slope. (f) Representative Western blot and (g) quantification showing relative abundance of complex components interacting with GST-NDPK-A ($N = 4$ replicates; $P < 0.0001$). Kinase activity of recombinant ADK to generate (h) ATP and (i) ADP in the presence of complex components ($N = 3$ /group). Kinase activity of recombinant NDPK to generate (j) ADP and (k) ADP in the presence of complex components ($N = 3$ /group). Activity of recombinant (l) ADK and (m) NDPK in the presence of a-Ab or c-Ab alone ($N = 3$ /group). * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$. Data are presented as mean \pm SEM. Two-way ANOVA (g-m).

[Source data](#)

Extended Data Fig. 5 Complex components alone do not alter GSIS.

(a) GSIS from WT mouse islets treated with each of the proposed complex components alone ($N = 4$ /condition). (b) GSIS from INS1 cells treated with NDPK-ADK with FABP4, lipid binding mutant (LBM) FABP4, or FABP4 pre-treated with inhibitor BMS-309403 ($N = 4$ /group). * $P < 0.05$, *** $P < 0.001$. Data are mean \pm SEM. One-way ANOVA (a,b).

[Source data](#)

Extended Data Fig. 6 Treatment with a-Ab preserves beta cell mass and islet number with minimal alterations to the pancreas immune profile.

Cytosolic calcium flux in INS1 cells from (a,b) the extracellular space (N = 2-3 coverslips/treatment; 88 (Control), 100 (NDPK-ADK-FABP4), 47 (MRS2365) cells) and (c,d) the ER in response to thapsigargin as determined by Fura-2 AM staining in control conditions or pretreatment with FABP4-ADK-NDPK with or without MRS2365 (N = 2-3 coverslips/treatment; 50 cells). (e) Cytosolic calcium flux from the ER in INS1 cells under control conditions, or pretreatment with FABP4-ADK-NDPK with or without the adenylyl cyclase inhibitor NKY80 (N = 2-3 coverslips/treatment; 100 (Control, NDPK-ADK-FABP4), 150 (NKY80) cells). (f) Quantification of Western blots from Fig. 4g (N = 4/condition). (g) Gene expression of ER stress markers BIP following 2hr treatment with NDPK-ADK, FABP4-ADK-NDPK, or FABP4-ADK-NDPK with a-Ab in the presence or absence of Tg (N = 3). (h) Cleaved caspase 3/7 activity in INS1 cells treated with increasing concentrations of Tg with or without FABP4-ADK-NDPK (N = 4). (i) Percentage of CD45+ cells (N = 12), (j) regulatory T-cells (N = 6), (k) cytotoxic T-cells (N = 6), (l) T-Helper cells (N = 5), (m) B-cells (N = 6), (n) dendritic cells (N = 6), and (o) granulocytes (N = 6) in whole pancreas from mice treated with PBS, a-Ab or c-Ab for 14 weeks by one-way ANOVA. *P < 0.05, **P < 0.01, ***P < 0.001. Data are presented as mean +/- SEM. One-way ANOVA (b,d,i-o); Two-tailed unpaired t-test (f); Two-way ANOVA (g,h).

[Source data](#)

Extended Data Table 1 Comparative data on T1D and control groups (BABYDIAB and DIMELLI cohorts)

Extended Data Table 2 Comparative data on T1D and control group (BRI cohort)

Extended Data Table 3 Spearman correlations of circulating FABP4 in the BRI cohort

Extended Data Table 4 Nucleoside abundance in INS1 cell supernatant with high glucose

Extended Data Table 5 Immunophenotyping of peripheral tissues in non-diabetic NOD mice

Extended Data Table 6 List of antibodies used in the study

Supplementary information

Supplementary Figure 1

This file contains the uncropped blots.

Reporting Summary

Source data

Source Data Fig. 1

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. 4

Source Data Extended Data Fig. 5

Source Data Extended Data Fig. 6

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Prentice, K.J., Saksi, J., Robertson, L.T. *et al.* A hormone complex of FABP4 and nucleoside kinases regulates islet function. *Nature* **600**, 720–726 (2021). <https://doi.org/10.1038/s41586-021-04137-3>

- Received: 12 August 2020
- Accepted: 14 October 2021
- Published: 08 December 2021
- Issue Date: 23 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04137-3>

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

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

- Article
- [Published: 15 December 2021](#)

The KEYNOTE-811 trial of dual PD-1 and HER2 blockade in HER2-positive gastric cancer

- [Yelena Y. Janjigian](#) ORCID: [orcid.org/0000-0002-5357-5858¹](https://orcid.org/0000-0002-5357-5858),
- [Akihito Kawazoe²](#),
- [Patricio Yañez³](#),
- [Ning Li⁴](#),
- [Sara Lonardi⁵](#),
- [Oleksii Kolesnik⁶](#),
- [Olga Barajas⁷](#),
- [Yuxian Bai⁸](#),
- [Lin Shen⁹](#),
- [Yong Tang¹⁰](#),
- [Lucjan S. Wyrwicz¹¹](#),
- [Jianming Xu¹²](#),
- [Kohei Shitara²](#),
- [Shukui Qin¹³](#),
- [Eric Van Cutsem¹⁴](#),
- [Josep Tabernero](#) ORCID: [orcid.org/0000-0002-2495-8139¹⁵](https://orcid.org/0000-0002-2495-8139),
- [Lie Li¹⁶](#),
- [Sukrut Shah¹⁶](#),
- [Pooja Bhagia¹⁶](#) &
- [Hyun Cheol Chung¹⁷](#)

[Nature](#) volume 600, pages 727–730 (2021)

- 9155 Accesses
- 2 Citations
- 206 Altmetric
- [Metrics details](#)

Subjects

- [Cancer immunotherapy](#)
- [Chemotherapy](#)
- [Gastric cancer](#)
- [Targeted therapies](#)

Abstract

Human epidermal growth factor receptor 2 (*HER2*, also known as *ERBB2*) amplification or overexpression occurs in approximately 20% of advanced gastric or gastro-oesophageal junction adenocarcinomas^{1,2,3}. More than a decade ago, combination therapy with the anti-*HER2* antibody trastuzumab and chemotherapy became the standard first-line treatment for patients with these types of tumours⁴. Although adding the anti-programmed death 1 (PD-1) antibody pembrolizumab to chemotherapy does not significantly improve efficacy in advanced *HER2*-negative gastric cancer⁵, there are preclinical^{6,7,8,9,10,11,12,13,14,15,16,17,18,19} and clinical^{20,21} rationales for adding pembrolizumab in *HER2*-positive disease. Here we describe results of the protocol-specified first interim analysis of the randomized, double-blind, placebo-controlled phase III KEYNOTE-811 study of pembrolizumab plus trastuzumab and chemotherapy for unresectable or metastatic, *HER2*-positive gastric or gastro-oesophageal junction adenocarcinoma²² (<https://clinicaltrials.gov>, NCT03615326). We show that adding pembrolizumab to trastuzumab and chemotherapy markedly reduces tumour size, induces complete responses in some participants, and significantly improves objective response rate.

This is a preview of subscription content

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Fig. 1: Best percentage change from baseline in the size of target lesions among participants in the efficacy population.



Data availability

Merck Sharp & Dohme Corp., a subsidiary of Merck & Co., Inc., Kenilworth, NJ, USA (MSD) is committed to providing qualified scientific researchers access to anonymized data and clinical study reports from the company's clinical trials for the purpose of conducting legitimate scientific research. MSD is also obligated to protect the rights and privacy of trial participants and, as such, has a procedure in place for evaluating and fulfilling requests for sharing company clinical trial data with qualified external scientific researchers. The MSD data-sharing documentation (http://engagezone.msd.com/ds_documentation.php) outlines the process and requirements for submitting a data request. Feasible requests will be reviewed by a committee of MSD subject matter experts to assess the scientific validity of the request and the qualifications of the requestors. In line with data privacy legislation, submitters of approved requests must enter into a standard data-sharing agreement with MSD before data access is granted. Data will be made available for request after product approval in the US and EU or after product development is discontinued; an exception will be made for this ongoing trial such that data will be made available for request after the protocol-specified primary endpoint analyses have been reported. There are circumstances that may prevent MSD from sharing requested data, including country or region-specific regulations. If the

request is declined, it will be communicated to the investigator. Access to genetic or exploratory biomarker data requires a detailed statistical analysis plan that is collaboratively developed by the requestor and MSD subject matter experts; after approval of the statistical analysis plan and execution of a data-sharing agreement, MSD will either perform the proposed analyses and share the results with the requestor or will construct biomarker covariates and add them to a file with clinical data that is uploaded to a SAS portal so that the requestor can perform the proposed analyses.

References

1. 1.
Cancer Genome Atlas Research Network. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature* **513**, 202–209 (2014).
2. 2.
Van Cutsem, E. et al. HER2 screening data from ToGA: targeting HER2 in gastric and gastroesophageal junction cancer. *Gastric Cancer* **18**, 476–484 (2015).
3. 3.
Cancer Genome Atlas Research Network. Integrated genomic characterization of oesophageal carcinoma. *Nature* **541**, 169–175 (2017).
4. 4.
Bang, Y. J. et al. Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. *Lancet* **376**, 687–697 (2010).
5. 5.

Shitara, K. et al. Efficacy and safety of pembrolizumab or pembrolizumab plus chemotherapy vs chemotherapy alone for patients with first-line, advanced gastric cancer: the KEYNOTE-062 phase 3 randomized clinical trial. *JAMA Oncol.* **6**, 1571–1580 (2020).

6. 6.

Gall, V. A. et al. Trastuzumab increases HER2 uptake and cross-presentation by dendritic cells. *Cancer Res.* **77**, 5374–5383 (2017).

7. 7.

Park, S. et al. The therapeutic effect of anti-HER2/neu antibody depends on both innate and adaptive immunity. *Cancer Cell* **18**, 160–170 (2010).

8. 8.

Taylor, C. et al. Augmented HER-2 specific immunity during treatment with trastuzumab and chemotherapy. *Clin. Cancer Res.* **13**, 5133–5143 (2007).

9. 9.

zum Buschenfelde, C. M., Hermann, C., Schmidt, B., Peschel, C. & Bernhard, H. Antihuman epidermal growth factor receptor 2 (HER2) monoclonal antibody trastuzumab enhances cytolytic activity of class I-restricted HER2-specific T lymphocytes against HER2-overexpressing tumor cells. *Cancer Res.* **62**, 2244–2247 (2002).

10. 10.

Chaganty, B. K. R. et al. Trastuzumab upregulates PD-L1 as a potential mechanism of trastuzumab resistance through engagement of immune effector cells and stimulation of IFN γ secretion. *Cancer Lett.* **430**, 47–56 (2018).

11. 11.

Varadan, V. et al. Immune signatures following single dose trastuzumab predict pathologic response to preoperative trastuzumab and chemotherapy in HER2-positive early breast cancer. *Clin. Cancer Res.* **22**, 3249–3259 (2016).

12. 12.

Triulzi, T. et al. HER2 signaling regulates the tumor immune microenvironment and trastuzumab efficacy. *Oncoimmunology* **8**, e1512942 (2019).

13. 13.

Triulzi, T. et al. Early immune modulation by single-agent trastuzumab as a marker of trastuzumab benefit. *Br. J. Cancer* **119**, 1487–1494 (2018).

14. 14.

Stagg, J. et al. Anti-ErbB-2 mAb therapy requires type I and II interferons and synergizes with anti-PD-1 or anti-CD137 mAb therapy. *Proc. Natl Acad. Sci. USA* **108**, 7142–7147 (2011).

15. 15.

Apetoh, L. et al. Toll-like receptor 4-dependent contribution of the immune system to anticancer chemotherapy and radiotherapy. *Nat. Med.* **13**, 1050–1059 (2007).

16. 16.

Ghiringhelli, F. et al. Activation of the NLRP3 inflammasome in dendritic cells induces IL-1 β -dependent adaptive immunity against tumors. *Nat. Med.* **15**, 1170–1178 (2009).

17. 17.

Tesniere, A. et al. Immunogenic death of colon cancer cells treated with oxaliplatin. *Oncogene* **29**, 482–491 (2010).

18. 18.

Mortenson, E. D., Park, S., Jiang, Z., Wang, S. & Fu, Y. X. Effective anti-neu-initiated antitumor responses require the complex role of CD4⁺ T cells. *Clin. Cancer Res.* **19**, 1476–1486 (2013).

19. 19.

Muller, P. et al. Trastuzumab emtansine (T-DM1) renders HER2⁺ breast cancer highly susceptible to CTLA-4/PD-1 blockade. *Sci. Transl. Med.* **7**, 315ra188 (2015).

20. 20.

Janjigian, Y. Y. et al. First-line pembrolizumab and trastuzumab in HER2-positive oesophageal, gastric, or gastro-oesophageal junction cancer: an open-label, single-arm, phase 2 trial. *Lancet Oncol.* **21**, 821–831 (2020).

21. 21.

Rha, S. Y. et al. Targeting HER2 in combination with anti-PD-1 and chemotherapy confers a significant tumor shrinkage of gastric cancer: A multi-institutional phase Ib/II trial of first-line triplet regimen (pembrolizumab, trastuzumab, chemotherapy) for HER2-positive advanced gastric cancer (AGC). *J. Clin. Oncol.* **38**, abstract 3081 (2020).

22. 22.

Chung, H. C. et al. First-line pembrolizumab/placebo plus trastuzumab and chemotherapy in HER2-positive advanced gastric cancer: KEYNOTE-811. *Future Oncol.* **17**, 491–501 (2021).

23. 23.

Barrueto, L. et al. Resistance to checkpoint inhibition in cancer immunotherapy. *Transl. Oncol.* **13**, 100738 (2020).

24. 24.

Shitara, K. et al. Pembrolizumab versus paclitaxel for previously treated, advanced gastric or gastro-oesophageal junction cancer (KEYNOTE-061): a randomised, open-label, controlled, phase 3 trial. *Lancet* **392**, 123–133 (2018).

25. 25.

Boku, N. et al. Nivolumab plus chemotherapy versus chemotherapy alone in patients with previously untreated advanced or recurrent gastric/gastroesophageal junction (G/GEJ) cancer: ATTRACTION-4 (ONO-4538-37) study. *Ann. Oncol.* **31**, S1192–S1192 (2020).

26. 26.

Janjigian, Y. Y. et al. First-line nivolumab plus chemotherapy versus chemotherapy alone for advanced gastric, gastro-oesophageal junction, and oesophageal adenocarcinoma (CheckMate 649): a randomised, open-label, phase 3 trial. *Lancet* **398**, 27–40 (2021).

27. 27.

Chao, J. et al. Assessment of pembrolizumab therapy for the treatment of microsatellite instability-high gastric or gastroesophageal junction cancer among patients in the KEYNOTE-059, KEYNOTE-061, and KEYNOTE-062 clinical trials. *JAMA Oncol.* **7**, 895–902 (2021).

28. 28.

Janjigian, Y. Y. et al. Genetic predictors of response to systemic therapy in esophagogastric cancer. *Cancer Discov.* **8**, 49–58 (2018).

29. 29.

Eisenhauer, E. A. et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur. J. Cancer* **45**, 228–247 (2009).

30. 30.

Oken, M. M. et al. Toxicity and response criteria of the Eastern Cooperative Oncology Group. *Am. J. Clin. Oncol.* **5**, 649–655 (1982).

31. 31.

Kulangara, K. et al. Clinical utility of the combined positive score for programmed death ligand-1 expression and the approval of pembrolizumab for treatment of gastric cancer. *Arch. Pathol. Lab. Med.* **143**, 330–337 (2019).

32. 32.

Seymour, L. et al. iRECIST: guidelines for response criteria for use in trials testing immunotherapeutics. *Lancet Oncol.* **18**, e143–e152 (2017).

Acknowledgements

This study was funded by Merck Sharp & Dohme Corp., a subsidiary of Merck & Co., Inc., Kenilworth, NJ, USA. Y.Y.J. was additionally supported by Memorial Sloan Kettering Cancer Center National Institutes of Health/National Cancer Institute Cancer Center Support Grant P30 CA008748. We thank the patients and their families and caregivers for participating in the study; the investigators and site personnel; and the following employees of Merck Sharp & Dohme Corp., a subsidiary of Merck & Co., Inc., Kenilworth, NJ, USA: C. S. Shih for study oversight and M. A. Leiby for medical writing and editorial assistance.

Author information

Affiliations

1. Memorial Sloan Kettering Cancer Center and Weill Cornell Medical College, New York, NY, USA

Yelena Y. Janjigian

2. National Cancer Hospital East, Kashiwa, Japan

Akihito Kawazoe & Kohei Shitara

3. Universidad de La Frontera, James Lind Cancer Research Center, Temuco, Chile

Patricio Yañez

4. Henan Cancer Hospital, the Affiliated Cancer Hospital of Zhengzhou University, Zhengzhou, China

Ning Li

5. Istituto Oncologico Veneto IOV-IRCCS, Padova, Italy

Sara Lonardi

6. Medical Center ‘Oncolife’, Zaporizhzhia, Ukraine

Oleksii Kolesnik

7. Arturo López Pérez Foundation, Santiago, Chile

Olga Barajas

8. Harbin Medical University Cancer Hospital, Harbin, China

Yuxian Bai

9. Department of Gastrointestinal Oncology, Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Peking University Cancer Hospital & Institute, Beijing, China

Lin Shen

10. Cancer Hospital Affiliated to Xinjiang Medical University, Xinjiang, China

Yong Tang

11. Maria Sklodowska-Curie National Cancer Research Institute, Warsaw, Poland

Lucjan S. Wyrwicz

12. The Fifth Medical Center, Chinese PLA General Hospital, Beijing, China

Jianming Xu

13. Cancer Center of People's Liberation Army, Nanjing, China

Shukui Qin

14. University Hospitals Gasthuisberg and KU Leuven, Leuven, Belgium

Eric Van Cutsem

15. Vall d'Hebron Hospital Campus and Institute of Oncology (VHIO), IOB-Quiron, UVic-UCC, Barcelona, Spain

Josep Tabernero

16. Merck & Co., Kenilworth, NJ, USA

Lie Li, Sukrut Shah & Pooja Bhagia

17. Yonsei Cancer Center, Yonsei University College of Medicine, Seoul, South Korea

Hyun Cheol Chung

Contributions

Y.Y.J., L.S., K.S., S.Q., E.V.C., J.T., P.B. and H.C.C. conceived, designed and planned the study. Y.Y.J., A.K., P.Y., N.L., S.L., O.K., O.B., Y.B., L.S., Y.T., L.S.W., J.X. and H.C.C. collected data. P.B. provided study oversight. L.L. performed the statistical analysis. Y.Y.J., L.L., S.S. and P.B. prepared the first draft of the manuscript. Y.Y.J., A.K., P.Y., N.L., S.L., O.K., O.B., Y.B., L.S., Y.T., L.S.W., J.X., K.S., S.Q., E.V.C., J.T., L.L., S.S., P.B. and H.C.C. interpreted the results, provided critical review and revision of the drafts, and approved the decision to submit for publication.

Corresponding author

Correspondence to [Yelena Y. Janjigian](#).

Ethics declarations

Competing interests

Y.Y.J. has received research funding to the institution from Merck Sharp & Dohme, the National Cancer Institute, the US Department of Defense, Cycle for Survival, Fred's Team, Rgenix, Bayer, Genentech/Roche, Bristol-Myers Squibb, and Eli Lilly, has served on advisory boards for Rgenix, Merck Serono, Bristol-Myers Squibb, Eli Lilly, Pfizer, Bayer, Imugene, Merck Sharp & Dohme, Daiichi-Sankyo, Zymeworks, SeaGen, Basilea Pharmaceutical and AstraZeneca, and has equity in Rgenix. A.K. has received research funding to the institution from Merck Sharp & Dohme, Ono Pharmaceutical Co., Ltd. and Taiho Pharmaceutical Co., Ltd., and received honoraria for lectures, presentations, speakers' bureaus, or educational events from Ono Pharmaceutical Co., Ltd. and Taiho Pharmaceutical Co., Ltd. P.Y. has received research funding to the institution and travel support from Merck Sharp & Dohme. N.L. has received research funding to the institution from Merck Sharp & Dohme. S.L. has received research funding to the institution from Merck Sharp & Dohme, Amgen, Merck Serono, Bayer, Roche, Lilly, AstraZeneca, and Bristol-Myers Squibb, consulting fees from Amgen, Merck Serono, Lilly, AstraZeneca, Incyte, Daiichi Sankyo, Bristol-Myers Squibb, and Servier, and payment or honoraria for lectures, presentations, speakers' bureaus, or

educational events from Roche, Lilly, Bristol-Myers Squibb, Servier, Merck Serono, Pierre-Fabre, GSK, and Amgen. O.K. has received research funding to the institution from Merck Sharp & Dohme. O.B. has received research funding to the institution from Merck Sharp & Dohme, has received honoraria for serving as a speaker from Eli Lilly, travel support from Bristol-Myers Squibb and Roche, participated as an advisor for Bristol-Myers Squibb and Roche, and has a leadership role in the Sociedad Chilena de Oncología Medica. Y.B. has received research funding to the institution from Merck Sharp & Dohme. L.S. has received research funding to the institution from Merck Sharp & Dohme, Beijing Xiantong Biomedical Technology, Qilu Pharmaceutical, ZaiLab Pharmaceutical (Shanghai), Jacobio Pharmaceuticals, and Beihai Kangcheng (Beijing) Medical Technology, has received consulting fees from Merck Sharp & Dohme, Merck, Mingji Biopharmaceutical, Haichuang Pharmaceutical, Harbour BioMed, and Boehringer Ingelheim, has received payment for speakers bureaus from Hutchison Whampoa, Henrui, ZaiLab, and CSTONE Pharmaceutical, and has participated on an advisory board for Rongchang Pharmaceutical, Zailab, and CSTONE Pharmaceutical. Y.T. has received research funding to the institution from Merck Sharp & Dohme. L.S.W. has received research funding to the institution from Merck Sharp & Dohme. J.X. has received research funding to the institution from Merck Sharp & Dohme. K.S. has received research funding to the institution from Merck Sharp & Dohme, Astellas, Lilly, Ono, Sumitomo Dainippon, Daiichi Sankyo, Taiho, Chugai, Medi Science, and Eisai, has received consulting fees from Astellas, Lilly, Bristol-Myers Squibb, Takeda, Pfizer, Ono, Merck Sharp & Dohme, Taiho, Novartis, AbbVie, GlaxoSmithKline, Daiichi Sankyo, Amgen, and Boehringer Ingelheim, and has received honoraria for lectures from Novartis, AbbVie, and Yalkult. S.Q. has received consulting fees from Merck Sharp & Dohme for serving on an advisory board. E.V.C. has received funding to the institution from Merck Sharp & Dohme, Amgen, Bayer, Boehringer Ingelheim, Bristol-Myers Squibb, Celgene, Ipsen, Lilly, Merck KGaA, Novartis, Roche, and Servier and has received consulting fees from Array, Astellas, AstraZeneca, Bayer, Beigene, Biocartis, Boehringer Ingelheim, Bristol-Myers Squibb, Celgene, Daiichi Sankyo, Halozyme, GSK, Incyte, Ipsen, Merck Sharp & Dohme, Merck KGaA, Novartis, Pierre-Fabre, Roche, Servier, Sirtex, and Taiho. J.T. has received funding to the institution from Merck Sharp & Dohme, has

received consulting fees from Array Biopharma, AstraZeneca, Avvinity, Bayer, Boehringer Ingelheim, Chugai, Daiichi Sankyo, F. Hoffmann-La Roche Ltd., Genentech Inc., HalioDX SAS, Hutchison MediPharma International, Ikena Oncology, IQVIA, Lilly, Menarini, Merck Serono, Merus, Merck Sharp & Dohme, Mirati, Neophore, Novartis, Orion Biotechnology, Peptomyc, Pfizer, Pierre Fabre, Samsung Bioepis, Sanofi, SeaGen, Servier, Taiho, Tessa Therapeutics, and TheraMyc, and has received payment for educational collaboration with Imedex, Medscape Education, MJH Life Sciences, PeerView Institute for Medical Education, and Physicians Education Resource (PER). L.L. receives salary for full-time employment from Merck Sharp & Dohme Corp., a subsidiary of Merck & Co., Inc., Kenilworth, NJ, USA. S.S. receives salary for full-time employment from Merck Sharp & Dohme Corp., a subsidiary of Merck & Co., Inc., Kenilworth, NJ, USA and owns stock and/or has stock options in Merck & Co., Inc., Kenilworth, NJ, USA. P.B. receives salary for full-time employment from Merck Sharp & Dohme Corp., a subsidiary of Merck & Co., Inc., Kenilworth, NJ, USA and owns stock and/or has stock options in Merck & Co., Inc., Kenilworth, NJ, USA. H.C.C. has received funding to the institution from Merck Sharp & Dohme, Lilly, GlaxoSmithKline, Merck-Serono, Bristol-Myers Squibb/Ono, Taiho, Amgen, BeiGene, Incyte, and Zymeworks, has served on advisory boards for Taiho, Celltrion, Merck Sharp & Dohme, Lilly, Bristol-Myers Squibb, Merck-Serono, Gloria, BeiGene, Amgen, and Zymeworks, and has received honoraria for lectures from Merck-Serono and Lilly. Y.Y.J., A.K., P.Y., N.L., S.L., O.K., O.B., Y.B., L.S., Y.T., L.S.W., J.X., K.S., S.Q., E.V.C., J.T., L.L., S.S., P.B., and H.C.C. received medical writing support for this manuscript (no payment) from Merck Sharp & Dohme.

Additional information

Peer review information *Nature* thanks Myriam Chalabi, Florian Lordick, Vicky Wu and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

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 Treatment difference in objective response in subgroups of the efficacy population.

Response was assessed per RECIST, version 1.1, by blinded, independent central review. The estimated treatment difference between the pembrolizumab and placebo groups in the overall population was calculated using the Miettinen and Nurminen method stratified by geographic region (Australia/Europe/Israel/North America [Aus/Eur/Isr/NAm] vs Asia vs rest of world), PD-L1 combined positive score ([CPS]; ≥ 1 vs <1), and chemotherapy choice (5-fluorouracil plus cisplatin [FP] vs capecitabine plus oxaliplatin [CAPOX]); differences in subgroups were calculated using the unstratified Miettinen and Nurminen method. The efficacy population included the first 264 participants randomly allocated to treatment. The treatment regimen included trastuzumab and chemotherapy in both groups. Diamonds represent the estimated treatment difference in objective response, the error bars represent the 95% confidence interval (CI) of the estimated treatment difference, and the region shaded in dark grey represents the 95% CI for the treatment difference in the overall population. ECOG, Eastern Cooperative Oncology Group; GEJ, gastro-oesophageal junction; IHC, immunohistochemistry; ISH, in situ hybridization.

Extended Data Table 1 Follow-up duration calculated as time from randomization to the date of death or the database cutoff date, whichever was earliest

Extended Data Table 2 Disposition of study treatment

Extended Data Table 3 Demographic and disease characteristics at baseline

Extended Data Table 4 Adverse events that occurred in 10% or more of participants in either group of the as-treated population

Extended Data Table 5 Adverse events with a possible immune-mediated cause and infusion reactions in the as-treated population

Extended Data Table 6 Adverse events leading to death in the as-treated population

Supplementary information

Supplementary Information

This file contains a list of investigators, sites, and ethics committees; data monitoring committee; protocol cover letter and study protocol

Reporting Summary

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Janjigian, Y.Y., Kawazoe, A., Yañez, P. *et al.* The KEYNOTE-811 trial of dual PD-1 and HER2 blockade in HER2-positive gastric cancer. *Nature* **600**, 727–730 (2021). <https://doi.org/10.1038/s41586-021-04161-3>

- Received: 25 May 2021
- Accepted: 30 September 2021
- Published: 15 December 2021
- Issue Date: 23 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04161-3>

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

- [**Pembrolizumab for HER2⁺ gastric cancer**](#)

- David Killock

Nature Reviews Clinical Oncology (2021)

- [**Stomach cancer gets a triple punch of therapy**](#)

- Myriam Chalabi

Nature (2021)

[**Stomach cancer gets a triple punch of therapy**](#)

- Myriam Chalabi

News & Views 15 Dec 2021

[**Pembrolizumab for HER2⁺ gastric cancer**](#)

- David Killock

Research Highlight 22 Dec 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04161-3>

- Article
- [Published: 24 November 2021](#)

ecDNA hubs drive cooperative intermolecular oncogene expression

- [King L. Hung¹](#), [na1](#),
- [Kathryn E. Yost¹](#), [na1](#),
- [Liangqi Xie^{2,3,4}](#), [na1](#),
- [Quanming Shi¹](#),
- [Konstantin Helmsauer](#) [ORCID: orcid.org/0000-0003-0773-4114⁵](#),
- [Jens Luebeck](#) [ORCID: orcid.org/0000-0003-4391-979X^{6,7}](#),
- [Robert Schöpflin^{8,9,10}](#),
- [Joshua T. Lange](#) [ORCID: orcid.org/0000-0002-4246-6058^{11,12,13}](#),
- [Rocío Chamorro González](#) [ORCID: orcid.org/0000-0002-5653-7259⁵](#),
- [Natasha E. Weiser^{1,13}](#),
- [Celine Chen](#) [ORCID: orcid.org/0000-0002-1579-024X⁵](#),
- [Maria E. Valieva](#) [ORCID: orcid.org/0000-0002-4419-2839^{8,9}](#),
- [Ivy Tsz-Lo Wong^{12,13}](#),
- [Sihan Wu](#) [ORCID: orcid.org/0000-0001-8329-7492¹⁴](#),
- [Siavash R. Dehkordi⁷](#),
- [Connor V. Duffy¹](#),
- [Katerina Kraft¹](#),
- [Jun Tang^{12,13}](#),
- [Julia A. Belk](#) [ORCID: orcid.org/0000-0003-4724-6158^{13,15}](#),
- [John C. Rose¹](#),
- [M. Ryan Corces](#) [ORCID: orcid.org/0000-0001-7465-7652¹](#),
- [Jeffrey M. Granja](#) [ORCID: orcid.org/0000-0001-8402-3442¹](#),
- [Rui Li¹](#),
- [Utkrisht Rajkumar⁷](#),

- [Jordan Friedlein¹⁶](#),
- [Anindya Bagchi¹⁶](#),
- [Ansuman T. Satpathy](#) ORCID: [orcid.org/0000-0002-5167-537X¹³](#),
- [Robert Tjian^{3,4}](#),
- [Stefan Mundlos^{8,9,17}](#),
- [Vineet Bafna](#) ORCID: [orcid.org/0000-0002-5810-6241⁷](#),
- [Anton G. Henssen](#) ORCID: [orcid.org/0000-0003-1534-778X^{5,18,19,20}](#),
- [Paul S. Mischel](#) ORCID: [orcid.org/0000-0002-4560-2211^{12,13}](#),
- [Zhe Liu](#) ORCID: [orcid.org/0000-0002-3592-3150²](#) &
- [Howard Y. Chang](#) ORCID: [orcid.org/0000-0002-9459-4393^{1,21}](#)

[Nature](#) volume 600, pages 731–736 (2021)

- 14k Accesses
- 2 Citations
- 174 Altmetric
- [Metrics details](#)

Subjects

- [Cancer epigenetics](#)
- [Oncogenes](#)
- [Tumour heterogeneity](#)

Abstract

Extrachromosomal DNA (ecDNA) is prevalent in human cancers and mediates high expression of oncogenes through gene amplification and altered gene regulation¹. Gene induction typically involves *cis*-regulatory elements that contact and activate genes on the same chromosome^{2,3}. Here we show that ecDNA hubs—clusters of around 10–100 ecDNAs within the nucleus—enable intermolecular enhancer–gene interactions to promote oncogene overexpression. ecDNAs that encode multiple distinct oncogenes

form hubs in diverse cancer cell types and primary tumours. Each ecDNA is more likely to transcribe the oncogene when spatially clustered with additional ecDNAs. ecDNA hubs are tethered by the bromodomain and extraterminal domain (BET) protein BRD4 in a *MYC*-amplified colorectal cancer cell line. The BET inhibitor JQ1 disperses ecDNA hubs and preferentially inhibits ecDNA-derived-oncogene transcription. The BRD4-bound *PVT1* promoter is ectopically fused to *MYC* and duplicated in ecDNA, receiving promiscuous enhancer input to drive potent expression of *MYC*. Furthermore, the *PVT1* promoter on an exogenous episome suffices to mediate gene activation in *trans* by ecDNA hubs in a JQ1-sensitive manner. Systematic silencing of ecDNA enhancers by CRISPR interference reveals intermolecular enhancer–gene activation among multiple oncogene loci that are amplified on distinct ecDNAs. Thus, protein-tethered ecDNA hubs enable intermolecular transcriptional regulation and may serve as units of oncogene function and cooperative evolution and as potential targets for cancer therapy.

This is a preview of subscription content

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Fig. 1: ecDNA imaging correlates ecDNA clustering with transcriptional bursting.

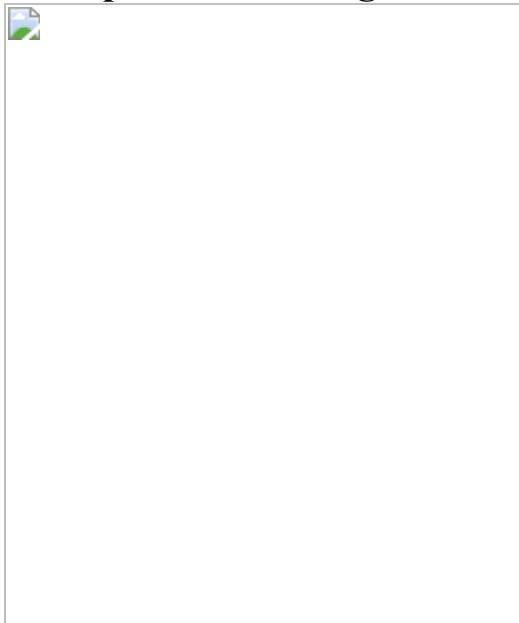


Fig. 2: BET proteins mediate ecDNA hub formation and transcription.

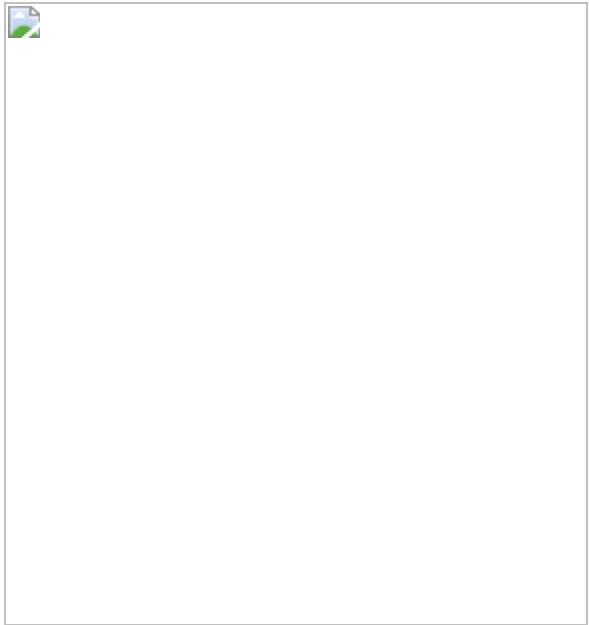


Fig. 3: Intermolecular activation of an episomal luciferase reporter in ecDNA hubs.

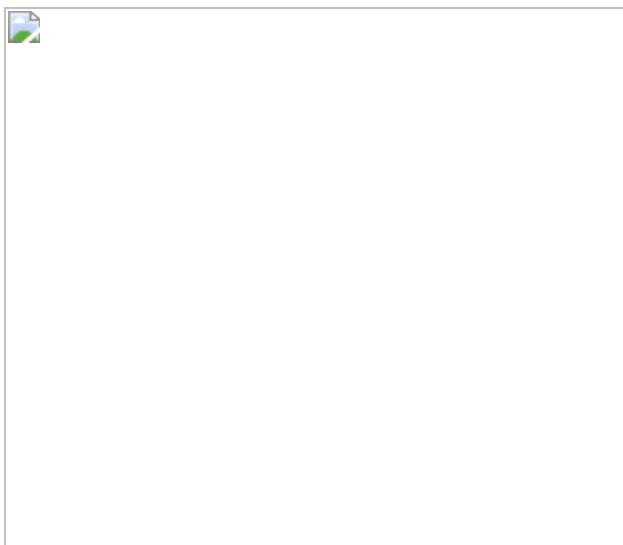
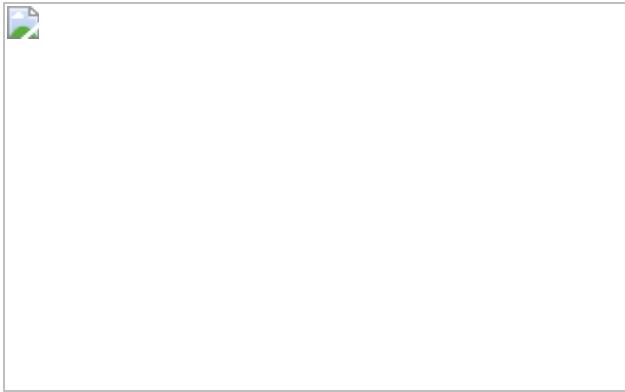


Fig. 4: ecDNA hubs mediate intermolecular enhancer–gene interactions.



Data availability

ChIP-seq, HiChIP, Hi-C, RNA-seq and single-cell multiomics (10x Chromium Single Cell Multiome ATAC + Gene Expression) data generated in this study have been deposited in the GEO and are available under accession number [GSE159986](#). Nanopore sequencing data, WGS data, sgRNA sequencing data and targeted ecDNA sequencing data after CRISPR–Cas9 digestion and PFGE generated in this study have been deposited in the SRA and are available under accession number [PRJNA670737](#). Optical mapping data generated in this study have been deposited in GenBank with BioProject code [PRJNA731303](#). The following publicly available data were also used in this study: TR14 H3K27ac ChIP-seq⁹³ (GEO: [GSE90683](#)); COLO320-DM, COLO320-HSR and PC3 WGS¹ (SRA: [PRJNA506071](#)); SNU16 WGS⁶⁰ (SRA: [PRJNA523380](#)); and HK359 WGS⁶ (SRA: [PRJNA338012](#)). Microscopy image files are available on figshare at <https://doi.org/10.6084/m9.figshare.c.5624713>.

Code availability

Custom code used in this study is available at
<https://github.com/ChangLab/ecDNA-hub-code-2021>.

References

1. 1.

Wu, S. et al. Circular ecDNA promotes accessible chromatin and high oncogene expression. *Nature* **575**, 699–703 (2019).

2. 2.

Gorkin, D. U., Leung, D. & Ren, B. The 3D genome in transcriptional regulation and pluripotency. *Cell Stem Cell* **14**, 762–775 (2014).

3. 3.

Zheng, H. & Xie, W. The role of 3D genome organization in development and cell differentiation. *Nat. Rev. Mol. Cell Biol.* **20**, 535–550 (2019).

4. 4.

Bailey, C., Shoura, M. J., Mischel, P. S. & Swanton, C. Extrachromosomal DNA—relieving heredity constraints, accelerating tumour evolution. *Ann. Oncol.* **31**, 884–893 (2020).

5. 5.

Kim, H. et al. Extrachromosomal DNA is associated with oncogene amplification and poor outcome across multiple cancers. *Nat. Genet.* **52**, 891–897 (2020).

6. 6.

Turner, K. M. et al. Extrachromosomal oncogene amplification drives tumour evolution and genetic heterogeneity. *Nature* **543**, 122–125 (2017).

7. 7.

Verhaak, R. G. W., Bafna, V. & Mischel, P. S. Extrachromosomal oncogene amplification in tumour pathogenesis and evolution. *Nat. Rev. Cancer* **19**, 283–288 (2019).

8. 8.

Cox, D., Yuncken, C. & Spriggs, A. I. Minute chromatin bodies in malignant tumours of childhood. *Lancet* **286**, 55–58 (1965).

9. 9.

van der Bliek, A. M., Lincke, C. R. & Borst, P. Circular DNA of 3T6R50 double minute chromosomes. *Nucleic Acids Res.* **16**, 4841–4851 (1988).

10. 10.

Hamkalo, B. A., Farnham, P. J., Johnston, R. & Schimke, R. T. Ultrastructural features of minute chromosomes in a methotrexate-resistant mouse 3T3 cell line. *Proc. Natl Acad. Sci.* **82**, 1126–1130 (1985).

11. 11.

Maurer, B. J., Lai, E., Hamkalo, B. A., Hood, L. & Attardi, G. Novel submicroscopic extrachromosomal elements containing amplified genes in human cells. *Nature* **327**, 434–437 (1987).

12. 12.

VanDevanter, D. R., Piaskowski, V. D., Casper, J. T., Douglass, E. C. & Von Hoff, D. D. Ability of circular extrachromosomal DNA molecules to carry amplified MYCN protooncogenes in human neuroblastomas in vivo. *J Natl Cancer Inst.* **82**, 1815–1821 (1990).

13. 13.

Nathanson, D. A. et al. Targeted therapy resistance mediated by dynamic regulation of extrachromosomal mutant EGFR DNA. *Science* **343**, 72–76 (2014).

14. 14.

Ståhl, F., Wettergren, Y. & Levan, G. Amplicon structure in multidrug-resistant murine cells: a nonrearranged region of genomic DNA

corresponding to large circular DNA. *Mol. Cell. Biol.* **12**, 1179–1187 (1992).

15. 15.

Vicario, R. et al. Patterns of HER2 gene amplification and response to anti-HER2 therapies. *PLoS ONE* **10**, e0129876 (2015).

16. 16.

Carroll, S. M. et al. Double minute chromosomes can be produced from precursors derived from a chromosomal deletion. *Mol. Cell. Biol.* **8**, 1525–1533 (1988).

17. 17.

Kitajima, K., Haque, M., Nakamura, H., Hirano, T. & Utiyama, H. Loss of irreversibility of granulocytic differentiation induced by dimethyl sulfoxide in HL-60 sublines with a homogeneously staining region. *Biochem. Biophys. Res. Commun.* **288**, 1182–1187 (2001).

18. 18.

Quinn, L. A., Moore, G. E., Morgan, R. T. & Woods, L. K. Cell lines from human colon carcinoma with unusual cell products, double minutes, and homogeneously staining regions. *Cancer Res.* **39**, 4914–4924 (1979).

19. 19.

Storlazzi, C. T. et al. Gene amplification as double minutes or homogeneously staining regions in solid tumors: origin and structure. *Genome Res.* **20**, 1198–1206 (2010).

20. 20.

Wahl, G. M. The importance of circular DNA in mammalian gene amplification. *Cancer Res.* **49**, 1333–1340 (1989).

21. 21.

Kumar, P. et al. ATAC-seq identifies thousands of extrachromosomal circular DNA in cancer and cell lines. *Sci. Adv.* **6**, eaba2489 (2020).

22. 22.

Morton, A. R. et al. Functional enhancers shape extrachromosomal oncogene amplifications. *Cell* **179**, 1330–1341 (2019).

23. 23.

Helmsauer, K. et al. Enhancer hijacking determines extrachromosomal circular *MYCN* amplicon architecture in neuroblastoma. *Nat. Commun.* **11**, 5823 (2020).

24. 24.

Itoh, N. & Shimizu, N. DNA replication-dependent intranuclear relocation of double minute chromatin. *J. Cell Sci.* **111**, 3275–3285 (1998).

25. 25.

Kanda, T., Sullivan, K. F. & Wahl, G. M. Histone-GFP fusion protein enables sensitive analysis of chromosome dynamics in living mammalian cells. *Curr. Biol.* **8**, 377–385 (1998).

26. 26.

Oobatake, Y. & Shimizu, N. Double-strand breakage in the extrachromosomal double minutes triggers their aggregation in the nucleus, micronucleation, and morphological transformation. *Genes Chromosomes Cancer* **59**, 133–143 (2020).

27. 27.

Beliveau, B. J. et al. Versatile design and synthesis platform for visualizing genomes with Oligopaint FISH probes. *Proc. Natl Acad.*

Sci. USA **109**, 21301–21306 (2012).

28. 28.

Koche, R. P. et al. Extrachromosomal circular DNA drives oncogenic genome remodeling in neuroblastoma. *Nat. Genetics* **52**, 29–34 (2019).

29. 29.

Parker, S. C. J. et al. Chromatin stretch enhancer states drive cell-specific gene regulation and harbor human disease risk variants. *Proc. Natl Acad. Sci. USA* **110**, 17921–17926 (2013).

30. 30.

Whyte, W. A. et al. Master transcription factors and mediator establish super-enhancers at key cell identity genes. *Cell* **153**, 307–319 (2013).

31. 31.

Lovén, J. et al. Selective inhibition of tumor oncogenes by disruption of super-enhancers. *Cell* **153**, 320–334 (2013).

32. 32.

Filippakopoulos, P. et al. Selective inhibition of BET bromodomains. *Nature* **468**, 1067–1073 (2010).

33. 33.

Sabari, B. R. et al. Coactivator condensation at super-enhancers links phase separation and gene control. *Science* **361**, eaar3958 (2018).

34. 34.

Ren, C. et al. Spatially constrained tandem bromodomain inhibition bolsters sustained repression of BRD4 transcriptional activity for TNBC cell growth. *Proc. Natl Acad. Sci. USA* **115**, 7949–7954 (2018).

35. 35.

Deshpande, V. et al. Exploring the landscape of focal amplifications in cancer using AmpliconArchitect. *Nat. Commun.* **10**, 392 (2019).

36. 36.

Luebeck, J. et al. AmpliconReconstructor integrates NGS and optical mapping to resolve the complex structures of focal amplifications. *Nat. Commun.* **11**, 4374 (2020).

37. 37.

Schwab, M., Klempnauer, K. H., Alitalo, K., Varmus, H. & Bishop, M. Rearrangement at the 5' end of amplified c-myc in human COLO 320 cells is associated with abnormal transcription. *Mol. Cell. Biol.* **6**, 2752–2755 (1986).

38. 38.

L'Abbate, A. et al. Genomic organization and evolution of double minutes/homogeneously staining regions with MYC amplification in human cancer. *Nucleic Acids Res.* **42**, 9131–9145 (2014).

39. 39.

Hann, S. R., King, M. W., Bentley, D. L., Anderson, C. W. & Eisenman, R. N. A non-AUG translational initiation in c-myc exon 1 generates an N-terminally distinct protein whose synthesis is disrupted in Burkitt's lymphomas. *Cell* **52**, 185–195 (1988).

40. 40.

Carramus, L. et al. The PVT-1 oncogene is a Myc protein target that is overexpressed in transformed cells. *J. Cell. Physiol.* **213**, 511–518 (2007).

41. 41.

Cho, S. W. et al. Promoter of lncRNA gene *PVT1* is a tumor-suppressor DNA boundary element. *Cell* **173**, 1398–1412 (2018).

42. 42.

Tolomeo, D., Agostini, A., Visci, G., Traversa, D. & Storlazzi, C. T. *PVT1*: a long non-coding RNA recurrently involved in neoplasia-associated fusion transcripts. *Gene* **779**, 145497 (2021).

43. 43.

Mumbach, M. R. et al. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat. Methods* **13**, 919 (2016).

44. 44.

Fulco, C. P. et al. Activity-by-contact model of enhancer–promoter regulation from thousands of CRISPR perturbations. *Nat. Genet.* **51**, 1664–1669 (2019).

45. 45.

Park, J. et al. A reciprocal regulatory circuit between CD44 and FGFR2 via c-myc controls gastric cancer cell growth. *Oncotarget* **7**, 28670–28683 (2016).

46. 46.

Furlong, E. E. M. & Levine, M. Developmental enhancers and chromosome topology. *Science* **361**, 1341–1345 (2018).

47. 47.

Zhu, Y. et al. Oncogenic extrachromosomal DNA functions as mobile enhancers to globally amplify chromosomal transcription. *Cancer Cell* **39**, 694–707 (2021).

48. 48.

Xue, K. S., Hooper, K. A., Ollodart, A. R., Dingens, A. S. & Bloom, J. D. Cooperation between distinct viral variants promotes growth of H3N2 influenza in cell culture. *Elife* **5**, e13974 (2016).

49. 49.

Vignuzzi, M., Stone, J. K., Arnold, J. J., Cameron, C. E. & Andino, R. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature* **439**, 344–348 (2006).

50. 50.

Henssen, A. et al. Targeting MYCN-driven transcription by BET-bromodomain inhibition. *Clin. Cancer Res.* **22**, 2470–2481 (2016).

51. 51.

Xie, L. et al. 3D ATAC-PALM: super-resolution imaging of the accessible genome. *Nat. Methods* **17**, 430–436 (2020).

52. 52.

Ambros, P. F. et al. International consensus for neuroblastoma molecular diagnostics: report from the International Neuroblastoma Risk Group (INRG) Biology Committee. *Br. J. Cancer* **100**, 1471–1482 (2009).

53. 53.

Balaban-Malenbaum, G. & Gilbert, F. Double minute chromosomes and the homogeneously staining regions in chromosomes of a human neuroblastoma cell line. *Science* **198**, 739–741 (1977).

54. 54.

Marrano, P., Irwin, M. S. & Thorner, P. S. Heterogeneity of *MYCN* amplification in neuroblastoma at diagnosis, treatment, relapse, and metastasis. *Genes Chromosomes Cancer* **56**, 28–41 (2017).

55. 55.

Villamón, E. et al. Genetic instability and intratumoral heterogeneity in neuroblastoma with *MYCN* amplification plus 11q deletion. *PLoS ONE* **8**, e53740 (2013).

56. 56.

Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).

57. 57.

Rajkumar, U. et al. EcSeg: semantic segmentation of metaphase images containing extrachromosomal DNA. *Iscience* **21**, 428–435 (2019).

58. 58.

Veatch, S. L. et al. Correlation functions quantify super-resolution images and estimate apparent clustering due to over-counting. *PLoS ONE* **7**, e31457 (2012).

59. 59.

Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114 (2014).

60. 60.

Ghandi, M. et al. Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* **569**, 503–508 (2019).

61. 61.

Normanno, D. et al. Probing the target search of DNA-binding proteins in mammalian cells using TetR as model searcher. *Nat. Commun.* **6**, 7357 (2015).

62. 62.

Mirkin, E. V., Chang, F. S. & Kleckner, N. Protein-mediated chromosome pairing of repetitive arrays. *J. Mol. Biol.* **426**, 550–557 (2014).

63. 63.

Grimm, J. B. et al. A general method to optimize and functionalize red-shifted rhodamine dyes. *Nat. Methods* **17**, 815–821 (2020).

64. 64.

Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).

65. 65.

Zhang, Y. et al. Model-based analysis of ChIP-seq (MACS). *Genome Biol.* **9**, R137 (2008).

66. 66.

Ramírez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).

67. 67.

Sedlazeck, F. J. et al. Accurate detection of complex structural variations using single-molecule sequencing. *Nat. Methods* **15**, 461–468 (2018).

68. 68.

Overhauser, J. in *Pulsed-Field Gel Electrophoresis, Methods in Molecular Biology* Vol. 12 (eds. Burmeister, M. & Ulanovsky, L.) 129–134 (Humana Press, 1992).

69. 69.

Picelli, S. et al. Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. *Genome Res.* **24**, 2033–2040 (2014).

70. 70.

Corces, M. R. et al. An improved ATAC-seq protocol reduces background and enables interrogation of frozen tissues. *Nat. Methods* **14**, 959–962 (2017).

71. 71.

Talevich, E., Shain, A. H., Botton, T. & Bastian, B. C. CNVkit: genome-wide copy number detection and visualization from targeted DNA sequencing. *PLoS Comput. Biol.* **12**, e1004873 (2016).

72. 72.

Raeisi Dehkordi, S., Luebeck, J. & Bafna, V. FaNDOM: fast nested distance-based seeding of optical maps. *Patterns* **2**, 100248 (2021).

73. 73.

Haas, B. J. et al. Accuracy assessment of fusion transcript detection via read-mapping and de novo fusion transcript assembly-based methods. *Genome Biol.* **20**, 213 (2019).

74. 74.

Hahne, F. & Ivanek, R. in *Statistical Genomics, Methods in Molecular Biology* Vol. 1418 (eds. Mathé, E. & Davis, S.) 335–351 (Humana Press, 2016).

75. 75.

Butler, A., Hoffman, P., Smibert, P., Papalexis, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).

76. 76.

Granja, J. M. et al. ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nat. Genet.* **53**, 403–411 (2021).

77. 77.

Satpathy, A. T. et al. Massively parallel single-cell chromatin landscapes of human immune cell development and intratumoral T cell exhaustion. *Nat. Biotechnol.* **37**, 925–936 (2019).

78. 78.

Mumbach, M. R. et al. Enhancer connectome in primary human cells identifies target genes of disease-associated DNA elements. *Nat. Genet.* **49**, 1602–1612 (2017).

79. 79.

Mumbach, M. R. et al. HiChIRP reveals RNA-associated chromosome conformation. *Nat. Methods* **16**, 489–492 (2019).

80. 80.

Servant, N. et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol.* **16**, 259 (2015).

81. 81.

Bhattacharyya, S., Chandra, V., Vijayanand, P. & Ay, F. Identification of significant chromatin contacts from HiChIP data by FitHiChIP. *Nat. Commun.* **10**, 4221 (2019).

82. 82.

Rao, S. S. P. et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665–1680 (2014).

83. 83.

Vidal, E. et al. OneD: increasing reproducibility of Hi-C samples with abnormal karyotypes. *Nucleic Acids Res.* **46**, e49 (2018).

84. 84.

Flynn, R. A. et al. Discovery and functional interrogation of SARS-CoV-2 RNA–host protein interactions. *Cell* **184**, 2394–2411 (2021).

85. 85.

Li, W. et al. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol.* **15**, 554 (2014).

86. 86.

Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).

87. 87.

Scheinin, I. et al. DNA copy number analysis of fresh and formalin-fixed specimens by shallow whole-genome sequencing with identification and exclusion of problematic regions in the genome assembly. *Genome Res.* **24**, 2022–2032 (2014).

88. 88.

Hadi, K. et al. Distinct classes of complex structural variation uncovered across thousands of cancer genome graphs. *Cell* **183**, 197–210 (2020).

89. 89.

Blumrich, A. et al. The *FRA2C* common fragile site maps to the borders of *MYCN* amplicons in neuroblastoma and is associated with

gross chromosomal rearrangements in different cancers. *Hum. Mol. Genet.* **20**, 1488–1501 (2011).

90. 90.

Gogolin, S. et al. CDK4 inhibition restores G₁-S arrest in MYCN-amplified neuroblastoma cells in the context of doxorubicin-induced DNA damage. *Cell Cycle* **12**, 1091–1104 (2013).

91. 91.

Durand, N. C. et al. Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst.* **3**, 95–98 (2016).

92. 92.

Knight, P. A. & Ruiz, D. A fast algorithm for matrix balancing. *IMA J. Numer. Anal.* **33**, 1029–1047 (2013).

93. 93.

Boeva, V. et al. Heterogeneity of neuroblastoma cell identity defined by transcriptional circuitries. *Nat. Genet.* **49**, 1408–1413 (2017).

Acknowledgements

We thank members of the Chang, Liu, Mischel, and Bafna laboratories for discussions; R. Zermenio, M. Weglarz and L. Nichols at the Stanford Shared FACS Facility for assistance with cell sorting experiments; X. Ji, D. Wagh and J. Coller at the Stanford Functional Genomics Facility for assistance with high-throughput sequencing; and A. Pang of Bionano Genomics for assistance with optical mapping. H.Y.C. was supported by NIH R35-CA209919 and RM1-HG007735; K.L.H. was supported by a Stanford Graduate Fellowship; and K.E.Y. was supported by the National Science Foundation Graduate Research Fellowship Program (NSF DGE-1656518), a Stanford Graduate Fellowship and a NCI Predoctoral to Postdoctoral Fellow Transition Award (NIH F99CA253729). Cell sorting for this project

was done on instruments in the Stanford Shared FACS Facility. Sequencing was performed by the Stanford Functional Genomics Facility (supported by NIH grants S10OD018220 and 1S10OD021763). Microscopy was performed on instruments in the UCSD Microscopy Core (supported by NINDS NS047101). A.G.H. is supported by the Deutsche Forschungsgemeinschaft (DFG; German Research Foundation) (398299703) and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement no. 949172). Z.L. is a Janelia Group Leader, and H.Y.C. and R.T. are Investigators of the Howard Hughes Medical Institute.

Author information

Author notes

1. These authors contributed equally: King L. Hung, Kathryn E. Yost, Liangqi Xie

Affiliations

1. Center for Personal Dynamic Regulomes, Stanford University School of Medicine, Stanford, CA, USA

King L. Hung, Kathryn E. Yost, Quanming Shi, Natasha E. Weiser, Connor V. Duffy, Katerina Kraft, John C. Rose, M. Ryan Corces, Jeffrey M. Granja, Rui Li & Howard Y. Chang

2. Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA, USA

Liangqi Xie & Zhe Liu

3. Department of Molecular and Cell Biology, Li Ka Shing Center for Biomedical and Health Sciences, CIRM Center of Excellence, University of California, Berkeley, Berkeley, CA, USA

Liangqi Xie & Robert Tjian

4. Howard Hughes Medical Institute, Berkeley, CA, USA

Liangqi Xie & Robert Tjian

5. Department of Pediatric Oncology and Hematology, Charité—Universitätsmedizin Berlin, Berlin, Germany

Konstantin Helmsauer, Rocío Chamorro González, Celine Chen & Anton G. Henssen

6. Bioinformatics and Systems Biology Graduate Program, University of California, San Diego, La Jolla, CA, USA

Jens Luebeck

7. Department of Computer Science and Engineering, University of California, San Diego, La Jolla, CA, USA

Jens Luebeck, Siavash R. Dehkordi, Utkrisht Rajkumar & Vineet Bafna

8. Development and Disease Research Group, Max Planck Institute for Molecular Genetics, Berlin, Germany

Robert Schöpflin, Maria E. Valieva & Stefan Mundlos

9. Institute for Medical and Human Genetics, Charité—Universitätsmedizin Berlin, Berlin, Germany

Robert Schöpflin, Maria E. Valieva & Stefan Mundlos

10. Department of Computational Molecular Biology, Max Planck Institute for Molecular Genetics, Berlin, Germany

Robert Schöpflin

11. Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla, CA, USA

Joshua T. Lange

12. ChEM-H, Stanford University, Stanford, CA, USA

Joshua T. Lange, Ivy Tsz-Lo Wong, Jun Tang & Paul S. Mischel

13. Department of Pathology, Stanford University, Stanford, CA, USA

Joshua T. Lange, Natasha E. Weiser, Ivy Tsz-Lo Wong, Jun Tang, Julia A. Belk, Ansuman T. Satpathy & Paul S. Mischel

14. Children's Medical Center Research Institute, University of Texas Southwestern Medical Center, Dallas, TX, USA

Sihan Wu

15. Department of Computer Science, Stanford University, Stanford, CA, USA

Julia A. Belk

16. Tumor Initiation and Maintenance Program, Sanford Burnham Prebys Medical Discovery Institute, La Jolla, CA, USA

Jordan Friedlein & Anindya Bagchi

17. Berlin-Brandenburg Center for Regenerative Therapies (BCRT), Charité—Universitätsmedizin Berlin, Berlin, Germany

Stefan Mundlos

18. Experimental and Clinical Research Center (ECRC), Max Delbrück Center for Molecular Medicine and Charité—Universitätsmedizin Berlin, Berlin, Germany

Anton G. Henssen

19. German Cancer Consortium (DKTK), partner site Berlin, and German Cancer Research Center DKFZ, Heidelberg, Germany

Anton G. Henssen

20. Berlin Institute of Health, Berlin, Germany

Anton G. Henssen

21. Howard Hughes Medical Institute, Stanford University School of Medicine, Stanford, CA, USA

Howard Y. Chang

Contributions

K.L.H., K.E.Y. and H.Y.C. conceived the project. K.L.H. performed and analysed CRISPRi and in vitro ecDNA digestion and PFGE experiments, and analysed single-cell multiomics, RNA-seq and ATAC-seq experiments. K.E.Y. performed and analysed metaphase DNA FISH imaging, ChIP-seq, HiChIP, WGS, COLO320-DM nanopore sequencing and JQ1 perturbation experiments. L.X. performed and analysed interphase DNA and RNA FISH imaging, TetO-eGFP cell line generation and live-cell imaging, and PVT1p-nLuc imaging experiments. Q.S. performed and analysed all luciferase reporter experiments, with the exception of PVT1p-nLuc RNA and DNA FISH, and helped with CRISPRi experiments. K.H. and R.S. analysed TR14 Hi-C data and amplicon reconstruction. J.L. and S.R.D. analysed COLO320-DM WGS, nanopore sequencing, optical mapping data and amplicon reconstruction. J.T.L., S.W., C.C. and J.T. performed and analysed DNA FISH imaging. R.C.G. generated Hi-C, DNA FISH, WGS and nanopore sequencing data for TR14. N.E.W. performed and analysed small-molecule inhibitor experiments and DNA FISH imaging after MS645 treatment. M.E.V. performed Hi-C experiments and data analysis for TR14. I.T.-L.W. performed metaphase DNA FISH imaging. C.V.D. performed and analysed ChIP-seq experiments. K.K. performed HiChIP experiments. J.A.B. helped with CRISPRi experimental design and cloning of the sgRNA pool. R.L. performed RNA-seq experiments. U.R. analysed metaphase DNA FISH data. J.F. generated COLO320-DM WGS data. M.R.C. and J.M.G. wrote the HiChIP data processing pipeline. M.R.C., J.C.R., A.B., A.T.S., R.T., S.M., V.B., A.G.H., P.S.M., Z.L. and H.Y.C. guided data

analysis and provided feedback on experimental design. K.L.H., K.E.Y. and H.Y.C. wrote the manuscript with input from all authors.

Corresponding author

Correspondence to [Howard Y. Chang](#).

Ethics declarations

Competing interests

H.Y.C. is a co-founder of Accent Therapeutics, Boundless Bio and Cartography Biosciences, and an advisor of 10x Genomics, Arsenal Biosciences and Spring Discovery. P.S.M. is a co-founder of Boundless Bio. He has equity and chairs the scientific advisory board, for which he is compensated. V.B. is a co-founder and advisor of Boundless Bio. A.T.S. is a founder of Immunai and Cartography Biosciences. K.E.Y. is a consultant for Cartography Biosciences.

Additional information

Peer review information *Nature* thanks Charles Lin and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

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 ecDNA FISH strategies and copy number estimation.](#)

a, WGS tracks with DNA FISH probe locations. For COLO320-DM and PC3, a 1.5 Mb *MYC* FISH probe (Fig. [1a, b](#)), a 100 kb *MYC* FISH probe

(Fig. 1d–f), or a 1.5 Mb chromosome 8 FISH probe was used. Commercial probes were used in SNU16 and HK359 cells. **b**, Representative DNA FISH image using chromosomal and 1.5 Mb *MYC* probes in non-ecDNA amplified HCC1569 showing paired signals as expected from the chromosomal loci. **c**, ecDNA clustering of individual COLO320-DM cells by autocorrelation $g(r)$. **d**, Representative FISH images showing ecDNA clustering in primary neuroblastoma tumours (patients 11 and 17). **e**, ecDNA clustering of individual primary tumour cells from all three patients using autocorrelation $g(r)$. **f**, Comparison of *MYC* copy number in COLO320-DM calculated based on WGS ($n=7$ genomic bins overlapping with DNA FISH probes), metaphase FISH ($n=82$ cells) and interphase FISH ($n=47$ cells). P-values determined by two-sided Wilcoxon test. **g**, Representative images of nascent *MYC* RNA FISH showing overlap of nascent RNA (intrinsic) and total RNA (exonic) FISH probes in PC3 cells (independently repeated twice). **h**, Representative images from combined DNA FISH for *MYC* ecDNA (100 kb probe) and chromosomal DNA with nascent *MYC* RNA FISH in COLO320-DM cells (independently repeated four times). **i**, *MYC* transcription probability measured by nascent RNA FISH normalized to DNA copy number by FISH comparing singleton ecDNAs to those found in hubs in COLO320-DM (box centre line, median; box limits, upper and lower quartiles; box whiskers, 1.5x interquartile range). To control for noise in transcriptional probability for small numbers of ecDNAs, we randomly re-sampled RNA FISH data grouped by hub size and calculated transcription probability. The violin plot represents transcriptional probability per ecDNA hub based on the hub size matched sampling. P-value determined by two-sided Wilcoxon test.

Extended Data Fig. 2 Generation of TetR-GFP COLO320-DM cells for ecDNA imaging in live cells.

a, ecDNA imaging based on TetO array knock-in and labelling with TetR-eGFP (left). Representative images of TetR-eGFP signal in TetO-eGFP COLO320-DM cells at indicated timepoints in a time course (right; independently repeated twice). **b**, GFP signal in ecDNA-TetO COLO320-DM cells. TetR-eGFP and monomeric TetR-eGFP(A206K)-labelled ecDNA hubs appear to be smaller in living cells than in DNA FISH studies of fixed cells, probably because the TetO array is not integrated in all ecDNA

molecules and there are potential differences caused by denaturation during DNA FISH and eGFP dimerization. **c**, ecDNA hub diameter in microns (box centre line, median; box limits, upper and lower quartiles; box whiskers, 1.5x interquartile range). Tet-eGFP-labelled hubs are slightly smaller than monomeric TetR-eGFP(A206K)-labelled hubs, potentially due to eGFP dimerization effects (Methods). P-value determined by two-sided Wilcoxon test. **d**, ecDNA hub number per cell. Line represents median. P-value determined by two-sided Wilcoxon test. **e**, TetR-eGFP signal in chr8-chromosomal-TetO (chr8:116,860,000–118,680,000, left) and ecDNA-TetO (TetO-eGFP COLO320-DM, right) COLO320-DM cells. **f**, Fluorescence intensity for chr8-chromosomal-TetO and ecDNA-TetO foci. **g**, **h**, Inferred ecDNA copy number per foci (g; n = number of foci/cell) and per cell (h; n = number of cells) for ecDNA-TetO labelled cells based on summed fluorescence intensity relative to chr8-chromosomal-TetO foci. Line represents median. **i**, Representative images of TetR-GFP signal in parental COLO320-DM without TetO array integration which shows minimal TetR-GFP foci. **j**, Mean fluorescence intensities for ecDNA (TetO-eGFP) and BRD4 (HaloTag) foci across a line drawn across the centre of the largest ecDNA (TetO-eGFP) signal. Data are mean ± SEM for n=5 ecDNA foci. **k**, Representative image of TetR-eGFP signal in COLO320-DM cells without TetO array integration overlaid with BRD4-HaloTag signal. Dashed line indicates nucleus boundary. We noted cytoplasmic TetR-eGFP signal in a subset of COLO320-DM cells without TetO array integration but it did not colocalize with BRD4-HaloTag. **l**, MYC RNA measured by RT-qPCR for parental COLO320-DM and BRD4-HaloTag COLO320-DM cells treated with DMSO or 500 nM JQ1 for 6 h which shows similar levels of MYC transcription and sensitivity to JQ1 inhibition following epitope tagging of BRD4. Data are mean ± SD between 3 biological replicates. P values determined by two-sided Student's t-test.

Extended Data Fig. 3 BET inhibition leads to ecDNA hub dispersal.

a, Representative metaphase FISH images and schematic showing ecDNA in COLO320-DM and chromosomal HSRs in COLO320-HSR (independently repeated twice for COLO320-DM and not repeated for COLO320-HSR). **b**, Ranked BRD4 ChIP-seq signal. Peaks in ecDNA or

HSR amplifications are highlighted and labelled with nearest gene. **c**, ATAC-seq, BRD4 ChIP-seq, H3K27ac ChIP-seq and WGS at amplified MYC locus. **d**, Number of ecDNA locations (including ecDNA hubs with >1 ecDNA and singleton ecDNAs) from interphase FISH imaging for individual COLO320-DM cells after treatment with DMSO or 500 nM JQ1 for 6 h. N = number of cells quantified per condition. P-value determined by two-sided Wilcoxon test. **e**, ecDNA copies in each ecDNA location from interphase FISH imaging in COLO320-DM after treatment with DMSO or 500 nM JQ1 for 6 h (box centre line, median; box limits, upper and lower quartiles; box whiskers, 1.5x interquartile range). N = number of ecDNA locations quantified per condition. P-value determined by two-sided Wilcoxon test. **f**, Representative live images of TetR-eGFP-labelled ecDNA after treatment with DMSO or 500 nM JQ1 at indicated timepoints in a time course (top; independently repeated twice) and ecDNA hub zoom-ins (bottom). **g**, Representative image from combined DNA/RNA FISH in COLO320-DM cells treated with DMSO, 500 nM JQ1, or 1% 1,6-hexanediol for 6 h. **h**, MYC transcription probability measured by dual DNA/RNA FISH after treatment with DMSO, 1% 1,6-hexanediol, or 100 µg/mL alpha-amanitin for 6 h (box centre line, median; box limits, upper and lower quartiles; box whiskers, 1.5x interquartile range; n = number of cells). P-values determined by two-sided Wilcoxon test. **i**, Representative DNA FISH images for MYC ecDNA in interphase COLO320-DM treated with either 1% 1,6-hexanediol or 100 µg/mL alpha-amanitin for 6 h. **j**, ecDNA clustering in interphase cells by autocorrelation $g(r)$ for COLO320-DM treated with DMSO, 1% 1,6-hexanediol, or 100 µg/mL alpha-amanitin for 6 h. Data are mean ± SEM (n = 10 cells quantified per condition). **k**, Averaged BRD4 ChIP-seq signal and heat map over all BRD4 peaks for cells treated with DMSO or 500 nM JQ1 for 6 h. **l**, Cell viability measured by ATP levels (CellTiterGlo) after treatment with different JQ1 concentrations for 48 h normalized to DMSO-treated cells. Data are mean ± SD between 3 biological replicates. P values determined by two-sided Student's t-test. **m**, Cell proliferation after treatment with different JQ1 concentrations over 72 h. Data are mean ± SD between 3 biological replicates. **n**, Cell doubling times after treatment with different JQ1 concentrations over 72 h in hours (top) or after normalization to DMSO-treated cells (bottom). Data are mean ± SD between 3 biological replicates. P values determined by two-sided Student's t-test. **o**, MYC RNA measured

by RT-qPCR after treatment with indicated inhibitors for 6 h (top; each point represents a biological replicate, n=6 for DMSO and JQ1 treatments, n=3 for all other drug treatments). Data are mean \pm SD. P values determined by two-sided Student's t-test. Details of inhibitor panel, protein target, significance of effect on *MYC* transcription, and comparison of effect on ecDNA and HSR transcription (bottom). **p, q**, Representative DNA FISH images (**p**) and clustering by autocorrelation $g(r)$ (**q**) for *MYC* ecDNAs in COLO320-DM treated with DMSO or 500 nM MS645 for 6 h. Data are mean \pm SEM. P-value determined by two-sided Wilcoxon test at radius = 0.

Extended Data Fig. 4 Reconstruction of COLO320-DM ecDNA amplicon structure.

a, Structural variant (SV) view of AmpliconArchitect (AA) reconstruction of the *MYC* amplicon in COLO320-DM cells. **b**, Nanopore sequencing of COLO320-DM cells (left) and distribution of read lengths. **c**, WGS for COLO320-DM with junctions detected by WGS and nanopore sequencing. **d**, Molecule lengths used for optical mapping and statistics. **e**, Reconstructed COLO320-DM ecDNA after integrating WGS, optical mapping, and in-vitro ecDNA digestion. Chromosomes of origin and corresponding coordinates (hg19) are labelled. Three inner circular tracks (light tan, slate and brown in colour, guides A, B and C, respectively) representing expected fragments as a result of Cas9 cleavage using three distinct sgRNAs and their expected sizes. Guide sequences are in Supplementary Table 2 (PFGE_guide_A-C). **f**, In-vitro Cas9 digestion of COLO320-DM ecDNA followed by PFGE (left). Fragment sizes were determined based on *H. wingei* and *S. cerevisiae* ladders. Uncropped gel image is in Supplementary Fig. 1. Middle panel shows short-read sequencing of the *MYC* ecDNA amplicon for all isolated fragments, ordered by fragment size. Right panel shows concordance of expected fragment sizes by optical mapping reconstruction, and observed fragment sizes by in-vitro Cas9 digestion (discordant fragments circled). Each sgRNA digestion was performed in one independent experiment. **g**, Metaphase FISH images showing colocalization of *MYC*, *PCAT1* and *PLUT* as predicted by optical mapping and in-vitro digestion. N = 20 cells and 1,270 ecDNAs quantified for *MYC/PCAT1* DNA FISH and n = 15 cells and 678 ecDNAs for

MYC/PLUT DNA FISH from one experiment. **h**, RNA expression measured by RT–qPCR for indicated transcripts in COLO320-DM cells stably expressing dCas9-KRAB and indicated sgRNAs (n=2 biological replicates). Canonical *MYC* was amplified with primers *MYC_exon1_f*w and *MYC_exon2_r*v; fusion *PVT1-MYC* was amplified with *PVT1_exon1_f*w and *MYC_exon2_r*v; total *MYC* was amplified with *total_MYC_exon2_f*w and *total_MYC_exon2_r*v. All primer sequences are in Supplementary Table 1 and guide sequences are in Supplementary Table 2. **i**, Alignment of junction reads at the *PVT1-MYC* breakpoint.

Extended Data Fig. 5 Single-cell multiomic analysis reveals combinatorial and heterogeneous ecDNA regulatory element activities associated with *MYC* expression.

a, Joint single-cell RNA and ATAC–seq for simultaneously assaying gene expression and chromatin accessibility and identifying regulatory elements associated with *MYC* expression. **b**, Unique ATAC–seq fragments and RNA features for cells passing filter (both log2-transformed). **c**, Correlation between *MYC* accessibility score and normalized RNA expression. **d**, UMAP from the RNA or the ATAC–seq data (left). Log-normalized and scaled *MYC* RNA expression (top right) and *MYC* accessibility scores (bottom right) were visualized on the ATAC–seq UMAP, showing cell-level heterogeneity in *MYC* RNA-seq and ATAC-seq signals in ecDNA-containing COLO320-DM. **e**, Gene expression scores (calculated using Seurat in R) of *MYC*-upregulated genes (Gene Set M6506, Molecular Signatures Database; MSigDB) across all *MYC* RNA quantile bins. Horizontal line marks median. Population variances for all individual cells are shown (top). P-value determined by two-sided F-test. **f**, *MYC* expression levels of top and bottom bins (left). Normalized ATAC–seq coverages are shown (right). **g**, Number of variable elements identified on COLO320-DM ecDNAs compared to chromosomal HSRs in COLO320-HSR (left). 45 variable elements were uniquely observed on ecDNA. All variable elements on ecDNA are shown on the right (y-axis shows -log10(FDR) and dot size represents log2 fold change. Five most significantly variable elements are highlighted and named based on relative position in kb to the *MYC* TSS (negative, 5'; positive, 3'). **h**, Correlation between estimated *MYC* copy numbers and normalized log2-transformed *MYC* expression of all

individual cells showing a high level of copy number variability associated with increased expression, in particular for COLO320-DM. **i**, Estimated *MYC* amplicon copy number of all cell bins separated by *MYC* RNA expression. **j**, Zoom-ins of the ATAC-seq coverage of each of the five most significantly variable elements identified in **g** (marked by dashed boxes). **k**, Similar distributions of TSS enrichment in the high and low cell bins, indicating differences in accessibility at variable elements are not an artifact of differences in data quality. **l**, Mean copy number regressed, log-normalized, scaled ATAC-seq coverage of the differential peaks against mean *MYC* RNA (log-normalized, mean-centred, scaled) for each cell bin in orange. Same number of random non-differential peaks from the same amplicon interval and shown in grey. Error bands show 95% confidence intervals for the linear models. **m**, Cumulative probability of *MYC* amplicon copy number distributions (mean-centred, scaled) of single-cell ATAC-seq data and DNA FISH data, suggesting that copy number estimates from single cell ATAC-seq data reflect heterogeneity in ecDNA copy number as measured by DNA FISH. P-values determined by Kolmogorov-Smirnov test (1,000 bootstrap simulations).

Extended Data Fig. 6 Endogenous enhancer connectome of COLO320-DM *MYC* ecDNA amplicon and effect of promoter sequence, *cis* enhancers, and BET inhibition on episomal reporter activation.

a, Top to bottom: COLO320-DM H3K27ac HiChIP contact map (KR-normalized read counts, 10-kb resolution), reconstructed COLO320-DM amplicon, H3K27ac ChIP-seq signal, BRD4 ChIP-seq signal, WGS coverage, interaction profile of *PVT1* (top, dark pink) and *MYC* (bottom, light pink) promoters at 10-kb resolution with FitHiChIP loops shown below, coloured by adjusted p-value. Active elements identified by scATAC and overlapping H3K27ac HiChIP contacts named by genomic distance to *MYC* start site: -1132E, -1087E, -679E, -655E, -401E, -328E, -85E. **b**, Comparison of HiChIP matrix normalization methods for COLO320-DM H3K27ac HiChIP at 10-kb resolution. HiChIP signal is robust to different normalization methods. **c**, Quantification of NanoLuc luciferase signal for plasmids with *PVT1p*-, *minp*-, or *MYCp*-driven NanoLuc reporter

expression. Luciferase signal was calculated by normalizing NanoLuc readings to Firefly readings. Bar plot shows mean \pm SEM. P values were calculated using a two-sided Student's t -test ($n=3$ biological replicates). **d**, Violin plots showing mean fluorescence intensities and signal sizes of the NanoLuc reporter RNA in PVT1p-reporter and minp-reporter transfected cells. P -values were calculated using a two-sided Wilcoxon test. **e**, Schematic of *PVT1* promoter-driven luciferase reporter plasmid with a *cis*-enhancer. Details of *cis*-enhancer are in Methods. **f**, Bar plot showing luciferase signal driven by PVT1p, MYCp or the constitutive TKp with or without a *cis*-enhancer (mean \pm SEM). All values are normalized to the corresponding promoter-only construct without a *cis*-enhancer. P values were calculated using a two-sided Student's t -test ($n=3$ biological replicates). **g**, Dot plots showing fold change in luciferase signal (Firefly-normalized NanoLuc signal) in JQ1-treated over DMSO-treated COLO320-DM and COLO320-HSR cells after transfection with the PVT1p or the MYCp plasmid with or without a *cis*-enhancer. P values were calculated using a two-sided Student's t -test ($n=3$ biological replicates).

Extended Data Fig. 7 Generation of monoclonal SNU16-dCas9-KRAB with reduced ecDNA fusions.

a, Representative DNA FISH images showing extrachromosomal single-positive *MYC* and *FGFR2* amplifications (top left and top middle) and double-positive *MYC* and *FGFR2* amplifications in metaphase spreads in parental SNU16 cells (top right) with zoom in (top right). $N = 42$ cells and 8,222 ecDNAs. Representative DNA FISH images showing distinct extrachromosomal *MYC* and *FGFR2* amplifications in metaphase spreads in SNU16-dCas9-KRAB cells (bottom). $N = 29$ cells and 3,893 ecDNAs. **b**, Ranked plot showing number of junction reads supporting each breakpoint in AmpliconArchitect. Breakpoints are coloured based on whether they span regions from the same amplicon (*MYC/FGFR2*) or regions from two distinct amplicons. **c**, HiChIP contact matrices at 10-kb resolution with KR normalization for parental SNU16 cell line (left) and SNU16-dCas9-KRAB cell line (right). Contact matrix for parental cells contains regions of increased *cis*-contact frequency between chr8 and chr10 as indicated, as compared to SNU16-dCas9-KRAB cells with highly reduced contact frequency between chr8 and chr10. Regions of increased focal interaction

overlapping low frequency structural rearrangements between chr8 and chr10 described in **b** indicated with boxes.

Extended Data Fig. 8 Perturbations of ecDNA enhancers by CRISPRi reveal functional intermolecular enhancer–gene interactions.

a, CRISPRi experiments perturbing candidate enhancers in SNU16-dCas9-KRAB cells. Single-guide RNAs (sgRNAs) were designed to target candidate enhancers on *FGFR2* and *MYC* ecDNAs based on chromatin accessibility. **b**, Experimental workflow for pooled CRISPRi repression of putative enhancers. Stable SNU16-dCas9-KRAB cells were generated from a single cell clone. Cells were transduced with a lentiviral pool of sgRNAs, selected with antibiotics and oncogene RNA was assessed by flowFISH. Cells were sorted into six bins by fluorescence-activated cell sorting (FACS) based on oncogene expression. sgRNAs were quantified for cells in each bin. **c**, FACS gating strategy. **d**, Log₂ fold changes of sgRNAs for each candidate enhancer element compared to unsorted cells for CRISPRi libraries targeting either *MYC* or *FGFR2* ecDNAs, followed by cell sorting based on expression levels of *MYC* or *FGFR2*. Each dot represents the mean log₂ fold change of 20 sgRNAs targeting a candidate element. Elements negatively correlated with oncogene expression as compared to the negative control sgRNA distributions in the same pools are marked in red. **e**, Bar plot showing significance of CRISPRi repression of candidate enhancer elements as in Fig. 4e (top). Significant in-*trans* and in-*cis* enhancers are coloured as indicated. SNU16-dCas9-KRAB H3K27ac HiChIP 1D signal track and interaction profiles of *FGFR2* and *MYC* promoters at 10-kb resolution with *cis* FitHiChIP loops shown below. Interaction profiles in *cis* shown in purple and in *trans* shown in orange. **f**, Spearman correlations of individual sgRNAs that target *MYC* TSS across fluorescence bins corresponding to *MYC* and *FGFR2* expression. P values using the lower-tailed t-test comparing target sgRNAs with negative control sgRNAs (negcontrols) are shown. Each dot represents an independent sgRNA.

Extended Data Fig. 9 Intermolecular enhancers and *MYC* are located on distinct molecules for the vast majority of ecDNAs.

a, Top: two-colour DNA FISH on metaphase spreads for quantifying the frequency of colocalization of the *MYC* gene and intermolecular enhancers shown in Fig. 4e. Above-random colocalization would indicate fusion events. Bottom: representative DNA FISH images. DNA FISH probes target the following hg19 genomic coordinates: E1, chr10:122,635,712–122,782,544 (RP11-95I16; n = 11 cells); E2, chr10:122,973,293–123,129,601 (RP11-57H2; n = 12 cells); E3/E4/E5, chr10:123,300,005–123,474,433 (RP11-1024G22; n = 10 cells). **b**, Top: numbers of distinct and colocalized FISH signals. To estimate random colocalization, 100 simulated images were generated with matched numbers of signals and mean simulated frequencies were compared with observed colocalization. P values determined by two-sided t-test (Bonferroni-adjusted). Bottom: number of colocalized signals significantly above random chance. Colocalization above simulated random distributions is the sum of colocalized molecules in excess of random means in all FISH images in which total colocalization was above the random mean plus 95% confidence interval (100 simulated images per FISH image). **c**, In vitro Cas9 digestion of *MYC*-containing ecDNA in SNU16-dCas9-KRAB followed by PFGE (one independent experiment). Fragment sizes were determined based on *H. wingei* and *S. cerevisiae* ladders. Uncropped gel image is in Supplementary Fig. 1. *MYC* CDS guide corresponds to guide B in Supplementary Table 2. **d**, Enrichment of enhancer DNA sequences in isolated *MYC* ecDNAs bands from c over background (DNA isolated from a separate PFGE lane in the corresponding size range resulting from undigested genomic DNA) based on normalized reads in 5kb windows. Each dot represents DNA from a distinct gel band. Red indicates fold change above 4. **e**, Sequencing track for a gel-purified *MYC* ecDNA showing enrichment of the *MYC* amplicon and depletion of the *FGFR2* amplicon containing enhancers E1-E5.

Extended Data Fig. 10 Reconstruction of four distinct amplicons in TR14 neuroblastoma cell line and intermolecular amplicon interaction patterns associated with H3K27ac marks.

a, Top to bottom: long read-based reconstruction of four different amplicons; genome graph with long read-based structural variants of >10kb size and >20 supporting reads indicated by red edges; copy number variation and coverage from short-read whole-genome sequencing, positions of the selected genes. **b**, A representative DNA FISH image of *MYCN* ecDNAs in interphase TR14 cells (top) and ecDNA clustering compared to DAPI control in the same cells assessed by autocorrelation $g(r)$ (bottom). Data are mean \pm SEM ($n = 14$ cells). **c**, Custom Hi-C map of reconstructed TR14 amplicons. The *MYCN/CDK4* amplicon and the *MYCN* ecDNA share sequences, which prevented an unambiguous short-read mapping in these regions and appear as white areas. *Trans* interactions appear locally elevated between *MYCN* ecDNA and *ODC1* amplicon (indicated by arrows). *Cis*- and *trans*-contact frequencies are coloured as indicated. **d**, Read support for structural variants identified by long read sequencing linking amplicons. Only one structural variant between distinct amplicons (*MYCN* and *MDM2* amplicons) was identified with 3 supporting reads. **e**, Variant allele frequency for structural variants linking amplicons. **f**, *Trans*-interaction pattern between enhancers on a *MYCN* amplicon fragment (vertical) and an *ODC1* amplicon fragment (horizontal). Short-read WGS coverage (grey), H3K27ac ChIP-seq track showing mean fold change over input in 1kb bins (yellow) and Hi-C contact map showing (KR-normalized counts in 5kb bins). **g**, Top to bottom: three amplicon reconstructions, virtual 4C interaction profile of the enhancer-rich *HPCAL1* locus on the *ODC1* amplicon with loci on other amplicons (red), and H3K27ac ChIP-seq (fold change over input; yellow). **h**, *Trans* interaction between different amplicons (KR-normalized counts in 5kb bins) depending on H3K27ac signal of the interaction loci (left; box centre line, median; box limits, upper and lower quartiles; box whiskers, 1.5x interquartile range). *Trans* interaction (KR-normalized counts in 5kb bins) separated by amplicon pair (right). H3K27ac High vs. Low denotes at least vs. less than 3-fold mean enrichment over input in 5kb bins. $N = 114,636$ H3K27ac Low + Low pairs, $n = 11,990$ H3K27ac High + Low pairs, $n = 296$ H3K27ac High + High pairs.

Supplementary information

Supplementary Information

This file contains Supplementary Tables 1 and 2 and accompanying legends for Supplementary Tables 1–3.

Reporting Summary

Supplementary Figure 1

Raw images of agarose gels. Related to Extended Data Figs. 4f, 9c.

Supplementary Table 3

See Supplementary Information for Supplementary Table 3 legend.

Supplementary Video 1

Live-cell imaging with untreated TetO–eGFP COLO320-DM cells. Snapshots of an untreated cell are shown over the course of 30 minutes. GFP labels TetO-knock-in *MYC* ecDNAs.

Supplementary Video 2

Live-cell imaging with DMSO-treated TetO–eGFP COLO320-DM cells. A control cell treated with DMSO was tracked over the course of 1 hour. GFP labels TetO-knock-in *MYC* ecDNAs.

Supplementary Video 3

Live-cell imaging with TetO-GFP COLO320-DM cells after JQ1 treatment. A cell treated with 500 nM JQ1 was tracked over the course of 1 hour. GFP labels TetO-knock-in *MYC* ecDNAs.

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Hung, K.L., Yost, K.E., Xie, L. *et al.* ecDNA hubs drive cooperative intermolecular oncogene expression. *Nature* **600**, 731–736 (2021). <https://doi.org/10.1038/s41586-021-04116-8>

- Received: 16 November 2020
- Accepted: 08 October 2021
- Published: 24 November 2021
- Issue Date: 23 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04116-8>

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

Further reading

- [**Hubs of transcriptional activity**](#)
 - Sarah Seton-Rogers

Nature Reviews Cancer (2021)

Hubs of transcriptional activity

- Sarah Seton-Rogers

Research Highlight 14 Dec 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04116-8>

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

- Article
- [Published: 08 December 2021](#)

Sex-specific chromatin remodelling safeguards transcription in germ cells

- [Tien-Chi Huang](#) [ORCID: orcid.org/0000-0002-8464-166X^{1,2}](#),
- [Yi-Fang Wang¹](#) [ORCID: orcid.org/0000-0002-1285-2214^{1,2}](#),
- [Eric Vazquez-Ferrer^{1,2}](#) [ORCID: orcid.org/0000-0001-5239-9162^{1,2}](#),
- [Ina Theofel](#) [ORCID: orcid.org/0000-0002-4063-5575^{3,4}](#) &
- [Cristina E. Requena](#) [ORCID: orcid.org/0000-0002-9762-5634^{3,4}](#)
- [Courtney W. Hanna](#) [ORCID: orcid.org/0000-0003-4145-1468^{1,2}](#)
- [Gavin Kelsey](#) [ORCID: orcid.org/0000-0003-4145-1468^{1,2}](#)
- [Petra Hajkova](#) [ORCID: orcid.org/0000-0003-4145-1468^{1,2}](#)

[Nature](#) volume 600, pages 737–742 (2021)

- 8717 Accesses
- 117 Altmetric
- [Metrics details](#)

Subjects

- [DNA methylation](#)
- [Gene silencing](#)
- [Germline development](#)
- [Histone post-translational modifications](#)

Abstract

Stability of the epigenetic landscape underpins maintenance of the cell-type-specific transcriptional profile. As one of the main repressive epigenetic systems, DNA methylation has been shown to be important for long-term gene silencing; its loss leads to ectopic and aberrant transcription in differentiated cells and cancer¹. The developing mouse germ line endures global changes in DNA methylation in the absence of widespread transcriptional activation. Here, using an ultra-low-input native chromatin immunoprecipitation approach, we show that following DNA demethylation the gonadal primordial germ cells undergo remodelling of repressive histone modifications, resulting in a sex-specific signature in mice. We further demonstrate that Polycomb has a central role in transcriptional control in the newly hypomethylated germline genome as the genetic loss of *Ezh2* leads to aberrant transcriptional activation, retrotransposon derepression and dramatic loss of developing female germ cells. This sex-specific effect of *Ezh2* deletion is explained by the distinct landscape of repressive modifications observed in male and female germ cells. Overall, our study provides insight into the dynamic interplay between repressive chromatin modifications in the context of a developmental reprogramming system.

This is a preview of subscription content

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Fig. 1: Sex-specific remodelling of repressive histone modifications following genome-wide DNA demethylation in PGCs.

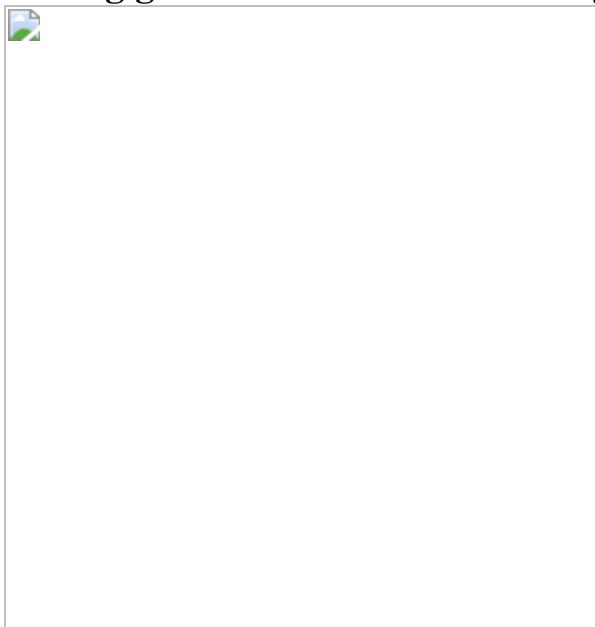


Fig. 2: Base composition determines H3K27me3 enrichment during gonadal DNA demethylation.



Fig. 3: Loss of *Ezh2* leads to widespread transcriptional derepression and loss of germ cells in the female germline.



Fig. 4: EZH2-mediated H3K27me3 regulates retrotransponson repression.



Data availability

ChIP-seq and RNA-seq data have been deposited in Gene Expression Omnibus (GEO) under [GSE141182](#). [Source data](#) are provided with this paper.

References

1. 1.

Robertson, K. D. DNA methylation and human disease. *Nat. Rev. Genet.* **6**, 597–610 (2005).

2. 2.

Guibert, S., Forne, T. & Weber, M. Global profiling of DNA methylation erasure in mouse primordial germ cells. *Genome Res.* **22**, 633–641 (2012).

3. 3.

Hackett, J. A. et al. Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. *Science* **339**, 448–452 (2013).

4. 4.

Seisenberger, S. et al. The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. *Mol. Cell* **48**, 849–862 (2012).

5. 5.

Seki, Y. et al. Cellular dynamics associated with the genome-wide epigenetic reprogramming in migrating primordial germ cells in mice. *Development* **134**, 2627–2638 (2007).

6. 6.

Hajkova, P. et al. Chromatin dynamics during epigenetic reprogramming in the mouse germ line. *Nature* **452**, 877–881 (2008).

7. 7.

Hajkova, P. Epigenetic reprogramming in the germline: towards the ground state of the epigenome. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **366**, 2266–2273 (2011).

8. 8.

Spiller, C., Koopman, P. & Bowles, J. Sex determination in the mammalian germline. *Annu. Rev. Genet.* **51**, 265–285 (2017).

9. 9.

Hill, P. W. S. et al. Epigenetic reprogramming enables the transition from primordial germ cell to gonocyte. *Nature* **555**, 392–396 (2018).

10. 10.

Brind'Amour, J. et al. An ultra-low-input native ChIP-seq protocol for genome-wide profiling of rare cell populations. *Nat. Commun.* **6**, 6033 (2015).

11. 11.

Hanna, C. W. et al. MLL2 conveys transcription-independent H3K4 trimethylation in oocytes. *Nat. Struct. Mol. Biol.* **25**, 73–82 (2018).

12. 12.

Hajkova, P. et al. Epigenetic reprogramming in mouse primordial germ cells. *Mech. Dev.* **117**, 15–23 (2002).

13. 13.

Juan, A. H. et al. Roles of H3K27me2 and H3K27me3 examined during fate specification of embryonic stem cells. *Cell Rep.* **17**, 1369–1382 (2016).

14. 14.

Li, H. et al. Polycomb-like proteins link the PRC2 complex to CpG islands. *Nature* **549**, 287–291 (2017).

15. 15.

Wang, X. et al. Molecular analysis of PRC2 recruitment to DNA in chromatin and its inhibition by RNA. *Nat. Struct. Mol. Biol.* **24**, 1028–1038 (2017).

16. 16.

Mendenhall, E. M. et al. GC-rich sequence elements recruit PRC2 in mammalian ES cells. *PLoS Genet.* **6**, e1001244 (2010).

17. 17.

Ohinata, Y. et al. Blimp1 is a critical determinant of the germ cell lineage in mice. *Nature* **436**, 207–213 (2005).

18. 18.

Li, Z. et al. The Sm protein methyltransferase PRMT5 is not required for primordial germ cell specification in mice. *EMBO J.* **34**, 748–758 (2015).

19. 19.

Soh, Y. Q. et al. A gene regulatory program for meiotic prophase in the fetal ovary. *PLoS Genet.* **11**, e1005531 (2015).

20. 20.

Bowles, J. & Koopman, P. Retinoic acid, meiosis and germ cell fate in mammals. *Development* **134**, 3401–3411 (2007).

21. 21.

De Cecco, M. et al. L1 drives IFN in senescent cells and promotes age-associated inflammation. *Nature* **566**, 73–78 (2019).

22. 22.

Jackson-Grusby, L. et al. Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation. *Nat. Genet.* **27**, 31–

39 (2001).

23. 23.

Liu, S. et al. Setdb1 is required for germline development and silencing of H3K9me3-marked endogenous retroviruses in primordial germ cells. *Genes Dev.* **28**, 2041–2055 (2014).

24. 24.

Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S. & Saitou, M. Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* **146**, 519–532 (2011).

25. 25.

Kawabata, Y. et al. Sex-specific histone modifications in mouse fetal and neonatal germ cells. *Epigenomics* **11**, 543–561 (2019).

26. 26.

Riising, E. M. et al. Gene silencing triggers Polycomb repressive complex 2 recruitment to CpG islands genome wide. *Mol. Cell* **55**, 347–360 (2014).

27. 27.

Yokobayashi, S. et al. PRC1 coordinates timing of sexual differentiation of female primordial germ cells. *Nature* **495**, 236–240 (2013).

28. 28.

Chiacchiera, F. et al. Polycomb complex PRC1 preserves intestinal stem cell identity by sustaining Wnt/β-catenin transcriptional activity. *Cell Stem Cell* **18**, 91–103 (2016).

29. 29.

Cohen, I. et al. PRC1 fine-tunes gene repression and activation to safeguard skin development and stem cell specification. *Cell Stem Cell* **22**, 726–739.e7 (2018).

30. 30.

Malki, S., van der Heijden, G. W., O'Donnell, K. A., Martin, S. L. & Bortvin, A. A role for retrotransposon LINE-1 in fetal oocyte attrition in mice. *Dev. Cell* **29**, 521–533 (2014).

31. 31.

Pezic, D., Manakov, S. A., Sachidanandam, R. & Aravin, A. A. piRNA pathway targets active LINE1 elements to establish the repressive H3K9me3 mark in germ cells. *Genes Dev.* **28**, 1410–1428 (2014).

32. 32.

Inoue, K., Ichiyanagi, K., Fukuda, K., Glinka, M. & Sasaki, H. Switching of dominant retrotransposon silencing strategies from posttranscriptional to transcriptional mechanisms during male germ-cell development in mice. *PLoS Genet.* **13**, e1006926 (2017).

33. 33.

Walter, M., Teissandier, A., Perez-Palacios, R. & Bourc'his, D. An epigenetic switch ensures transposon repression upon dynamic loss of DNA methylation in embryonic stem cells. *Elife* **5**, e11418 (2016).

34. 34.

Frapporti, A. et al. The Polycomb protein Ezl1 mediates H3K9 and H3K27 methylation to repress transposable elements in *Paramecium*. *Nat. Commun.* **10**, 2710 (2019).

35. 35.

Wachter, E. et al. Synthetic CpG islands reveal DNA sequence determinants of chromatin structure. *Elife* **3**, e03397 (2014).

36. 36.

McLaughlin, K. et al. DNA methylation directs Polycomb-dependent 3D genome re-organization in naive pluripotency. *Cell Rep.* **29**, 1974–1985.e6 (2019).

37. 37.

van Mierlo, G. et al. Integrative proteomic profiling reveals PRC2-dependent epigenetic crosstalk maintains ground-state pluripotency. *Cell Stem Cell* **24**, 123–137.e8 (2019).

38. 38.

Ehrlich, M. DNA hypomethylation in cancer cells. *Epigenomics* **1**, 239–259 (2009).

39. 39.

Jones, P. A., Issa, J. P. & Baylin, S. Targeting the cancer epigenome for therapy. *Nat. Rev. Genet.* **17**, 630–641 (2016).

40. 40.

Su, I. H. et al. Ezh2 controls B cell development through histone H3 methylation and IgH rearrangement. *Nat. Immunol.* **4**, 124–131 (2003).

41. 41.

Malki, S., Tharp, M. E. & Bortvin, A. A whole-mount approach for accurate quantitative and spatial assessment of fetal oocyte dynamics in mice. *Biol. Reprod.* **93**, 113 (2015).

42. 42.

Hayashi, K. & Saitou, M. Generation of eggs from mouse embryonic stem cells and induced pluripotent stem cells. *Nat. Protoc.* **8**, 1513–1524 (2013).

43. 43.

Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

44. 44.

Kim, D. et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).

45. 45.

Liao, Y., Smyth, G. K. & Shi, W. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. *Nucleic Acids Res.* **41**, e108 (2013).

46. 46.

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).

47. 47.

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).

48. 48.

Young, M. D., Wakefield, M. J., Smyth, G. K. & Oshlack, A. Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* **11**, R14 (2010).

49. 49.

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).

50. 50.

Risso, D., Ngai, J., Speed, T. P. & Dudoit, S. Normalization of RNA-seq data using factor analysis of control genes or samples. *Nat. Biotechnol.* **32**, 896–902 (2014).

51. 51.

Wingett, S. W. & Andrews, S. FastQ Screen: a tool for multi-genome mapping and quality control. *F1000Res.* **7**, 1338 (2018).

52. 52.

Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

53. 53.

The ENCODE Project Consortium. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).

54. 54.

Amemiya, H. M., Kundaje, A. & Boyle, A. P. The ENCODE blacklist: identification of problematic regions of the genome. *Sci. Rep.* **9**, 9354 (2019).

55. 55.

Ramirez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).

56. 56.

Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).

57. 57.

Lawrence, M. et al. Software for computing and annotating genomic ranges. *PLoS Comput. Biol.* **9**, e1003118 (2013).

58. 58.

Yu, G., Wang, L. G. & He, Q. Y. ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics* **31**, 2382–2383 (2015).

59. 59.

Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).

60. 60.

Krueger, F. & Andrews, S. R. Bismark: a flexible aligner and methylation caller for Bisulfite-seq applications. *Bioinformatics* **27**, 1571–1572 (2011).

61. 61.

Gardiner-Garden, M. & Frommer, M. CpG islands in vertebrate genomes. *J. Mol. Biol.* **196**, 261–282 (1987).

62. 62.

Hon, G. C. et al. 5mC oxidation by Tet2 modulates enhancer activity and timing of transcriptome reprogramming during differentiation. *Mol. Cell* **56**, 286–297 (2014).

63. 63.

Miyauchi, H. et al. Bone morphogenetic protein and retinoic acid synergistically specify female germ-cell fate in mice. *EMBO J.* **36**, 3100–3119 (2017).

Acknowledgements

We thank J. Elliot, T. Adejumo and B. Patel for help with FACS; C. Whilding for help with microscopy and immunofluorescence data analysis; L. Game for help with next-generation sequencing; Z. Agate-Bacon and G. Zimmerman for mouse husbandry; T. Carell (LMU Munich) for providing isotopically labelled deoxynucleoside standards; the members of the Hajkova laboratory for discussions and revisions of the manuscript; and A. Hocher and T. Warnecke for additional computational analysis. Work in the Hajkova laboratory is supported by MRC funding (MC_US_A652_5PY70) and an ERC grant (ERC-CoG-648879—dynamic modifications) to P.H. T.-C.H. was a recipient of an Imperial College London/Taiwan Top University Strategic Alliance PhD Scholarship.

Author information

Author notes

1. These authors contributed equally: Yi-Fang Wang, Eric Vazquez-Ferrer

Affiliations

1. MRC London Institute of Medical Sciences (LMS), London, UK

Tien-Chi Huang, Yi-Fang Wang, Eric Vazquez-Ferrer, Ina Theofel, Cristina E. Requena & Petra Hajkova

2. Institute of Clinical Sciences (ICS), Faculty of Medicine, Imperial College London, London, UK

Tien-Chi Huang, Eric Vazquez-Ferrer, Ina Theofel, Cristina E. Requena & Petra Hajkova

3. Epigenetics Programme, Babraham Institute, Cambridge, UK

Courtney W. Hanna & Gavin Kelsey

4. Centre for Trophoblast Research, University of Cambridge,
Cambridge, UK

Courtney W. Hanna & Gavin Kelsey

Contributions

T.-C.H. and P.H. conceived the study. T.-C.H. performed the experiments and analysed the data. Y.-F.W. analysed the next-generation sequencing data. E.V.-F. carried out the ULI-nChIP with the help of C.W.H. and G.K. I.T. and T.-C.H. carried out the PGCLC experiments. C.E.R. carried out LC–MS/MS. T.-C.H. and P.H. wrote the manuscript.

Corresponding author

Correspondence to [Petra Hajkova](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work.

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 Summary of ULI-nChIP-seq and genomic distribution of H3K27me3 and H3K9me3 enrichment.

a, Experimental scheme of PGC isolation using ΔPE Oct4-GFP mice (GOF-18/ΔPE-GFP)¹². ~1000 PGCs were used for each ULI-nChIP-seq. **b**, Characteristics of the H3K27me3 and H3K9me3 ULI-nChIP-seq datasets. Fraction of paired-end reads based on the mappability. Uniquely aligned (dark green). Uniquely aligned duplicates (light green). Multiple aligned (blue). Unaligned (grey). R: biological replicates. **c, d**, Genome-wide correlation of biological replicates for H3K27me3 and H3K9me3 ULI-nChIP-seq (Pearson correlation coefficient, 10 kb bins)²⁵. **e**, Genomic distribution of H3K27me3 and H3K9me3 peaks. **f**, Distance of H3K27me3 and H3K9me3 peaks relative to transcription start site (TSS). **g**, Violin plot showing the peak intensity distribution in E10.5 and E13.5 male and female PGCs. Peaks identified by MACS2 peak calling pipeline with broad peak setting ([Methods](#)). Box plots plotted using Tukey's method. The upper and lower hinges: the first and the third quartiles. The central line: the median. The upper end of the whisker represents the lowest value among either the third quartile plus $1.5 \times \text{IQR}$ or the maximum value from the data set. The lower end of the whisker represents the largest value among either the first quartile minus $1.5 \times (\text{IQR})$ or the minimum value from the data set. **h**, Bar chart showing the fraction of mapped paired-end reads of H3K27me3 and H3K9me3 associated with TEs in the genome.

Extended Data Fig. 2 Dynamics of DNA methylation and H3K27me3 at promoters during gonadal reprogramming.

a, Whole-genome bisulphite sequencing (WGBS) data from E10.5 and E12.5 female and male PGCs⁹. Density plot depicting DNA methylation levels at all promoters. **b**, Density plot depicting H3K27me3 enrichment at all promoters. **c**, Violin plot of H3K27me3 levels at promoters that lost DNA methylation (DNA methylation > 0.2 at E10.5). **d, f**, Dynamics of DNA methylation and H3K27me3 at all promoters (**d**) and group A (**f**). DNA methylation and H3K27me3 enrichment are shown by colour gradient. Distribution of each dot's value is shown using rug plot along x and y axis. **e**, Box plot shows H3K27me3 enrichment of low CpG density

(CpG < 4.1%) promoters which gained H3K27me3 in the female PGCs following global loss of DNA methylation. P values were calculated using 2 tailed Mann-Whitney U test with Bonferroni correction. ***: P < 0.001. **g**, Comparison of sequence characteristics and H3K27me3 enrichment of the group A from E13.5 male and female germ cells. P values were calculated by 2 tailed Mann-Whitney U test. Box plots were plotted using Tukey's method. The upper and lower hinges: the first and the third quartiles. The central line: the median. The upper end of the whisker represents the lowest value among either the third quartile plus $1.5 \times$ IQR or the maximum value from the data set. The lower end of the whisker represents the largest value among either the first quartile minus $1.5 \times$ (IQR) or the minimum value from the data set.

[Extended Data Fig. 3 Dynamics of DNA methylation, H3K9me3 and H3K27me3 at promoters during gonadal reprogramming.](#)

a, Heat map depicting the H3K27me3 and H3K9me3 enrichment (ULI-nChIP-seq), and DNA methylation rate (WGBS) at promoters. The promoters were grouped based on the pattern of dynamic change between DNA methylation and H3K27me3 (See also Fig. 2). Group A: loss of DNA methylation, gain H3K27me3 at E13.5. Group B: median loss of DNA methylation, high H3K27me3. Group C: low DNA methylation, High H3K27me3. Group D: loss of DNA methylation, low H3K27me3 at E13.5. Promoters that did not meet the criteria were grouped into non-classified. The expression levels of promoter-associated genes from RNA-seq are presented by TPM (Transcripts Per Kilobase Million) or z-score. The total number of promoters in each group are shown on the left. **b**, Box plot showing the quantitative measurement in each category, female and male PGCs, respectively. Box plots were plotted using Tukey's method. The upper and lower hinges: the first and the third quartiles. The central line: the median. The upper end of the whisker represents the lowest value among either the third quartile plus $1.5 \times$ IQR or the maximum value from the data set. The lower end of the whisker represents the largest value among either the first quartile minus $1.5 \times$ (IQR) or the minimum value from the data set. P values were calculated using 2 tailed Mann-Whitney U test. **c**, Venn diagram showing the number of overlapped promoters between male and

female PGCs. **d**, Gene ontology of expressed genes in group A (Supplementary Table 4).

Extended Data Fig. 4 Generation of the germline specific *Ezh2* conditional knockout.

a, Representative immunofluorescence (IF) staining shows EZH2 expression at different embryonic stages. EZH2 is highly expressed in PGCs, compared with surrounding somatic cells. Biological replicates $n = 3$. OCT4: PGC marker. DAPI indicates DNA. Scale bar: 10 μm . **b**, Functional domains of EZH2 protein and targeting strategy of *Ezh2* allele. Open boxes: exons. Black arrowhead: loxP sites. **c**, Breeding scheme to generate germline *Ezh2* knockout. *Ezh2* ^{Δ/Δ} , Tg (*Blimp1-Cre*) refers to CKO in the figures. f: allele flanked by loxP sites (floxed). Δ : Deleted allele generated using Cre-mediated recombination. Tg (*Blimp1-Cre*): transgenic mice express Cre recombinase under the control of *Blimp1* (*Prdm1*) promoter. **d**, Deleted alleles were confirmed by PCR genotyping using the primers shown by black arrows in **b**. #1: *Ezh2*^{f/ Δ} , Tg (*Blimp1-Cre*)^{+/−}. #2: *Ezh2*^{f/+}. This experiment was repeated at least 3 times independently with similar results.

Extended Data Fig. 5 Global H3K9me3, H2A119ub and DNA methylation are not altered in PGCs following the loss of EZH2.

a–c, Representative IF staining for H3K9me3, H2A119ub and TET1 using cryosectioned genital ridges. MVH and OCT4: PGC marker. DAPI indicates DNA. Biological replicates $n = 2$. Each dot represents one cell. **d**, IF staining for 5-methylcytosine (5mC). 5mC is enriched in pericentromeric regions in the nucleus of somatic cells but depleted in both Ctrl and *Ezh2* CKO germ cells. 3 independent experiments were repeated with similar results. DNA was stained by Propidium Iodide (PI). Scale Bar: 10 μm . **e**, Global 5mC and 5hmC levels (mean \pm s.d.) were measured by LC-MS/MS. Each dot represents an individual biological replicate. adj. P values were calculated using one-way ANOVA and Tukey's post-hoc multiple comparison test (2 tailed). **f**, Representative IF images of E13.5 Ctrl and

Ezh2 CKO female and male gonads. The bar chart shows the total number of germ cells per female gonad. Biological replicates $n = 2$. Each dot represents an individual biological replicate. MVH: PGC marker. DAPI indicates DNA. Scale bar: 100 μm .

[Source data](#)

Extended Data Fig. 6 Transcriptome analysis of Ctrl and *Ezh2* CKO PGCs.

a, Catalytic core and accessory subunits of mammalian PRC2. **b**, RNA expression levels (TPM) of PRC2 components in PGCs at different embryonic stages. **c**, RNA expression levels (TPM) of *Ezh1* in the Ctrl and *Ezh2* CKO germ cells. **d**, Sample distance matrix of RNA-seq samples by non-supervised clustering ([Methods](#)). **e**, PCA Plot shows the distance of transcriptomes from different PGC developmental stages. Dash line circle indicates samples of the same developmental stage

[Source data](#).

Extended Data Fig. 7 Sex-specific transcription factor repertoire determines transcriptional activation upon loss of EZH2.

a, Gene ontology (GO) terms associated with E13.5 male differentially expressed (DE) genes (Ctrl vs *Ezh2* CKO). **b**, Integrative Genomics Viewer (IGV) plot shows the H3K27me3 enrichment and RNA-seq read counts of *Stra8*. Mouse genome: mm9. **c**, Heat map depicting gene expression and the chromatin dynamics at promoters of meiotic DE genes. Box plot shows the H3K27me3 enrichment and RNA expression (TPM) of meiotic DE genes in male and female PGCs. Z scores were calculated for male and female separately. Box plots were plotted using Tukey's method. The upper and lower hinges: the first and the third quartiles. The central line: the median. The upper end of the whisker represents the lowest value among either the third quartile plus $1.5 \times \text{IQR}$ or the maximum value from the data set. The lower end of the whisker represents the largest value among either the first quartile minus $1.5 \times \text{IQR}$ or the minimum value from the data set. P

values were calculated using 2 tailed Mann-Whitney U test. ***: P < 0.01
d, Bar chart showing the odds ratio of 104 meiosis prophase genes¹⁹ in each groups of promoters (Fig. 2). P values were calculated by Fisher exact test.
*: P < 0.05. **: P < 0.01. ***: P < 0.001. **e**, Promoters of upregulated genes in female *Ezh2* CKO are significantly enriched for transcription factor motifs that relate to retinoid acid signalling pathway. Motif analysis was performed using Bioconductor package PWMEnrich. **f**, Heat map showing the relative gene expression of identified transcription factors in Ctrl and *Ezh2* CKO PGC samples. **g**, Heat map showing gene expression of 45 Germline Reprogramming Responsive (GRR) genes⁹ in Ctrl and *Ezh2* CKO samples. Differentially expressed genes are shown on the top (adj. P < 0.05) (Two-tailed Wald test using DESeq2; See [Methods](#)).

Extended Data Fig. 8 Loss of Ezh2 does not lead to precocious meiotic prophase.

a, Representative images of the female germ cells immunostained with meiosis specific synaptonemal complex protein SCP3 and germ cell marker MVH in Ctrl and *Ezh2* CKO embryonic ovaries. n = 2 independent animals analysed. **b**, Representative images of leptotene/zygotene and pachytene germ cells from female gonads at E15.5 and E16.5. Meiotic prophase I progression was classified based on SCP1 and SYCP3 staining. Leptotene/zygotene: DNA aggregates/ knobs were identified with numerous thin threads of SYCP3 staining. Pachytene: thick threads stained with strong SCP1 and SCP3 signal. Percentage of staged germ cells in Ctrl and *Ezh2* CKO gonads was shown. Each dot represents a biological replicate. n = 2 independent animals analysed. At least 200 germ cells were calculated for each biological replicate. **c**, **d**, Representative IF images of γH2AX positive cells in cryosectioned gonads. During normal meiosis progression, γH2AX signal shows DNA double strand breaks (DSBs) occurring during homologous recombination. Accumulation of γH2AX signal was identified in E15.5 Ctrl and *Ezh2* CKO germ cells and was greatly reduced at E18.5. A number RAD51 foci can be identified at E16.5 but greatly decreased at E18.5. Filament-like, RAD51-positive structure was identified in about 10% of the *Ezh2* CKO germ cells but not in Ctrl germ cells. MVH positive germ cells are indicated by yellow arrowhead. DAPI indicates DNA. Scale

bar: 10 µm. $n = 2$ independent animals analysed. These experiments were repeated 2 times independently with similar results.

[Source data](#)

Extended Data Fig. 9 EZH2-mediated H3K27me3 regulates TE repression.

a, Volcano plot of TEs in E13.5 male PGCs (*Ezh2* CKO vs Ctrl). Significantly upregulated TEs are labelled in red. Fold change > 1 , FDR < 0.1 . **b**, Multidimensional scaling (MDS) plot showing distance of Ctrl and *Ezh2* CKO PGC RNA-seq samples based on TE expression. Only uniquely mapped reads were considered. **c**, Gene Set Enrichment Analysis (GSEA) of RNA-seq (E13.5♀ *Ezh2* CKO vs Ctrl). Number of genes enriched in each gene set is shown by the circle size. **d**, GSEA plot showing genes upregulated in female *Ezh2* CKO PGCs are enriched in p53 pathway and interferon alpha response. FDR q value < 0.25 was considered significant. NES: normalized enrichment score. **e**, Distance of H3K9me3 *de novo* peaks to transcription start sites (TSS) and transposable elements (TEs). **f**, Representative IGV plot showing H3K9me3 enrichment on IAP Ez elements. **g**, TE subfamilies enriched predominantly for H3K9me3 or H3K27me3. Each row represents one TE subfamily. Multiple mapped and uniquely mapped reads were taken into account. **h, i**, Heat map showing DNA methylation, H3K9me3 and H3K27me3 enrichment at individual copies of IAPLTR2_Mm and L1Md_Gf. Each row represents one uniquely mapped, single TE copy belonging to the respective TE subfamily.

Extended Data Fig. 10 Loss of Ezh2 does not lead to loss of male germ cells.

a, Representative images of γH2AX staining in male gonads. OCT4 or MVH positive germ cells are indicated by the yellow arrowhead. γH2AX positive cells are indicated by the white arrowhead. Less than 1% of MVH positive were γH2AX positive in E18.5 male *Ezh2* CKO gonads. $n = 2$ independent animals analysed. **b**, Gross appearance of Ctrl and *Ezh2* CKO male gonads at E18.5. **c**, Numbers of Ctrl and *Ezh2* CKO male germ cells per section. Data are represented as mean ± s.d. Biological replicate $n = 2$.

d, e, Representative IF images of DAZL and MILI expression in MVH positive male germ cells in Ctrl and *Ezh2* CKO embryos. DAPI indicates DNA. Scale bar: 10 μ m. $n = 2$ independent animals analysed.

[Source data](#)

Extended Data Fig. 11 EZH2 is dispensable for transcriptional regulation in PGC-like cells (PGCLCs).

a, Experimental scheme of PGCLC induction and 4-OHT induced EZH2 depletion in PGCLCs. Equivalent mouse germ line stages are shown above the experimental timeline. *Ezh2*^{f/f}, CreERT2 embryonic stem cells (ESCs) were firstly induced into epiblast-like cells (EpiLCs) ([Methods](#)). Ethanol (Ctrl) or 4-Hydroxytamoxifen (4-OHT) were then added to induce Cre-mediated deletion of *Ezh2* during PGCLC induction. At day 4 of treatment, PGCLCs were sorted by FACS and subjected to RNA-seq. **b**, Representative IF images of whole mount day 4 Ctrl and *Ezh2* KO PGCLC aggregates. Biological replicates = 3. Experiment were repeated independently with similar results. Scale bar: 50 μ m. **c**, Histone modifications (H3K27me3, H3K9me3 and H2AK119ub) in Ctrl and *Ezh2* KO PGCLCs. AP2 γ positive cells: PGCLCs. Biological replicates = 3. Experiment were repeated independently with similar results. DNA was stained by DAPI (blue). Scale bar: 10 μ m. **d**, Expression of germline associated genes and pluripotency associated genes from this study and published datasets^{62,63}. TPM: Transcripts Per Kilobase Million. Each data point represents one biological replicate (mean \pm s.d.). $n = 3$ biological independent experiments. **e**, IGV plot shows RNA-seq read coverage of *Ezh2* transcripts. Reads from deleted exons were depleted in the KO samples. Mouse genome: mm9. **f**, Sample distance matrix of RNA-seq samples by non-supervised clustering. **g, h**, Volcano plot showing differentially expressed (DE) genes (**g**) and TEs (Multiple mapped plus uniquely mapped reads) (**h**) in Ctrl and *Ezh2* KO PGCLCs. Significantly DE genes and TEs are labelled in red (P values calculated by two-tailed Wald test using DESeq2; See [Methods](#)).

Extended Data Fig. 12 Summary of transcriptional regulation by EZH2 in the germ line.

a, Venn diagram depicting overlap of genes significantly upregulated in the *Ezh2* CKO female PGCs and the *Rnf2* CKO PGCs²⁷. **b**, Upper Venn diagram showing the overlap between the EZH2 and SETDB1 regulated TEs in E13.5 mouse germ cells. Lower Venn diagram shows overlap between EZH2 regulated TEs in E13.5 germ cells with EED regulated TEs in ESCs^{23,33}. **c**, Model depicting the relationship between DNA demethylation and changes in histone repressive modifications in gonadal PGCs undergoing epigenetic reprogramming.

Supplementary information

Supplementary information

This file contains uncropped immunoblots in Fig 1g and the gating strategy for isolating OCT4-GFP positive primordial germ cells.

Reporting Summary

Supplementary Table 1

Summary for ULI-nChIP libraries and RNA-seq libraries.

Supplementary Table 2

This table contains a summary of the CG content of promoters and their associated genes.

Supplementary Table 3

This file contains (1) TPM and lists of differentially expressed genes (*Ezh2* CKO versus *Ezh2* Ctrl PGCs) and (2) a summary of expression levels of DNA methyltransferases and relevant chromatin modifiers.

Supplementary Table 4

This table contains gene ontology terms related to Fig. 2 and RNA-seq data.

Supplementary Table 5

This table contains differentially expressed retrotransposons (*Ezh2* CKO vs *Ezh2* Ctrl PGCs).

Supplementary Table 6

This file contains detailed information for antibodies, primer sequences, mouse strains and essential kits.

Source data

Source Data Fig. 1

Source Data Fig. 3

Source Data Fig. 4

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

Huang, TC., Wang, YF., Vazquez-Ferrer, E. *et al.* Sex-specific chromatin remodelling safeguards transcription in germ cells. *Nature* **600**, 737–742 (2021). <https://doi.org/10.1038/s41586-021-04208-5>

- Received: 17 January 2020
- Accepted: 01 November 2021
- Published: 08 December 2021
- Issue Date: 23 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04208-5>

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/s41586-021-04208-5>

- Article
- Open Access
- [Published: 26 October 2021](#)

A conserved mechanism for regulating replisome disassembly in eukaryotes

- [Michael Jenkyn-Bedford](#) [ORCID: orcid.org/0000-0003-2335-8278¹](#) na1,
- [Morgan L. Jones](#) [ORCID: orcid.org/0000-0003-1948-4925¹](#) na1,
- [Yasemin Baris¹](#),
- [Karim P. M. Labib²](#),
- [Giuseppe Cannone¹](#),
- [Joseph T. P. Yeeles](#) [ORCID: orcid.org/0000-0002-2255-2119¹](#) &
- [Tom D. Deegan](#) [ORCID: orcid.org/0000-0001-9639-7636²](#) nAff3

Nature volume 600, pages 743–747 (2021)

- 7012 Accesses
- 81 Altmetric
- [Metrics details](#)

Subjects

- [Electron microscopy](#)
- [Replisome](#)

Abstract

Replisome disassembly is the final step of eukaryotic DNA replication and is triggered by ubiquitylation of the CDC45–MCM–GINS (CMG) replicative helicase^{1,2,3}. Despite being driven by evolutionarily diverse E3 ubiquitin ligases in different eukaryotes (SCF^{Dia2} in budding yeast¹, CUL2^{LRR1} in metazoa^{4,5,6,7}), replisome disassembly is governed by a common regulatory principle, in which ubiquitylation of CMG is suppressed before replication termination, to prevent replication fork collapse. Recent evidence suggests that this suppression is mediated by replication fork DNA^{8,9,10}.

However, it is unknown how SCF^{Dia2} and CUL2^{LRR1} discriminate terminated from elongating replisomes, to selectively ubiquitylate CMG only after termination. Here we used cryo-electron microscopy to solve high-resolution structures of budding yeast and human replisome–E3 ligase assemblies. Our structures show that the leucine-rich repeat domains of Dia2 and LRR1 are structurally distinct, but bind to a common site on CMG, including the MCM3 and MCM5 zinc-finger domains. The LRR–MCM interaction is essential for replisome disassembly and, crucially, is occluded by the excluded DNA strand at replication forks, establishing the structural basis for the suppression of CMG ubiquitylation before termination. Our results elucidate a conserved mechanism for the regulation of replisome disassembly in eukaryotes, and reveal a previously unanticipated role for DNA in preserving replisome integrity.

[Download PDF](#)

Main

The eukaryotic replisome is assembled around the CMG helicase at replication origins during replication initiation. Once assembled, CMG remains stably associated with replication forks until two forks emanating from adjacent origins converge, or a single fork encounters the end of a linear chromosome or a template discontinuity, at which point replication terminates (Fig. 1a). Upon termination, the replisome is disassembled in two steps: first, CMG is ubiquitylated on its Mcm7 subunit by a cullin-RING E3 ubiquitin ligase (SCF^{Dia2} in budding yeast, CUL2^{LRR1} in metazoa); second, ubiquitylated Mcm7 is unfolded by the Cdc48 ATPase (also known as p97 in higher eukaryotes), leading to disassembly of the replisome^{1,2,3,8,9,10}. As there is no known mechanism for origin-independent CMG assembly in S phase, premature disassembly of CMG must be avoided, to prevent replication fork collapse and genome instability¹¹. CMG translocates on the leading-strand template while excluding the lagging-strand template from its central channel¹². It has been suggested that this ‘excluded’ DNA strand, which is lost upon termination (Fig. 1a), inhibits ubiquitylation of CMG at replication forks^{8,9,10}. However, because there are currently no structures of terminated replisomes in complex with SCF^{Dia2} or CUL2^{LRR1}, how ubiquitylation of CMG is regulated to restrict replisome disassembly to termination remains a key unanswered question.

Fig. 1: Cryo-EM structures of terminated replisomes from *Saccharomyces cerevisiae* bound by SCF^{Dia2}.

 **figure1**

a, Schematic of the regulation of replisome disassembly. For clarity, replisomes are depicted as CMG. CMG ubiquitylation and replisome disassembly are inhibited at replication forks by an as yet unknown mechanism, dependent on the excluded DNA strand (in red box). This inhibition is relieved following translocation onto dsDNA (in green box, left and middle) or off DNA (in green box, right). **b**, Slice-through view of cryo-EM density for complexes assembled on dsDNA. The density shown is a composite of focused maps (refer to Extended Data Fig. 2). **c**, DNA engagement within the MCM C-tier motor domains by complexes assembled on dsDNA (coloured) or on a replication fork (grey; PDB: 6SKL¹³). **d**, Cryo-EM density as in **b** (left) and corresponding atomic model (right) for complexes assembled on dsDNA. For the atomic model, only SCF^{Dia2}, DNA and MCM subunits that interact with SCF^{Dia2} are coloured. **e**, Alternative view of the atomic model in **d**. **f**, Cryo-EM density for complexes assembled in the absence of DNA, derived from multibody refinement. **g**, Comparison of the MCM–Dia2^{LRR} interface from complexes assembled on dsDNA (**b–e**), off DNA (**f**) or on a replication fork (PDB: 6SKL¹³). For the regions of MCM at this interface, the root mean square deviation (r.m.s.d.) of the replication fork-bound complex compared with the dsDNA-bound or off-DNA complexes is 1.39 Å and 0.93 Å, respectively.

Terminated yeast replisome structures

To determine the molecular basis for the regulation of CMG ubiquitylation, we aimed to solve the structure of a terminated replisome, by adapting our system for reconstituting budding yeast replisomes for structural analysis¹³. After convergence of two replication forks, CMG translocates onto nascent double-stranded DNA (dsDNA) produced by the converging replisome^{3,9,14} (Fig. 1a). To trap a replisome bound around dsDNA, we used a DNA substrate that lacked a 5' flap and contained a short

stretch methylphosphonate modifications embedded in dsDNA, which slow translocation of CMG¹⁵ (Extended Data Fig. 1a). This DNA substrate was incubated with CMG, the replisome factors Tof1–Csm3, Mrc1 and Ctf4, SCF^{Dia2} (Hrt1–Cdc53–Skp1–Dia2), an E2–ubiquitinconjugate (Cdc34–Ub)¹⁶, and the leading-strand DNA polymerase Pol-ε, in the presence of ATP (Extended Data Fig. 1a). After glycerol gradient sedimentation, complexes containing all replisome and SCF^{Dia2} subunits were isolated (Extended Data Fig. 1b). Cdc34–Ub did not associate with the complex, perhaps reflecting the absence of neddylation on the Cdc53 cullin subunit of SCF^{Dia2} (refs. 17,18).

After gradient fixation, samples were prepared for cryo-electron microscopy (cryo-EM), yielding three-dimensional (3D) reconstructions at average resolutions of 3.2–4.0 Å (Fourier shell correlation (FSC) = 0.143 criterion; Extended Data Fig. 1c–h, Extended Data Table 1). DNA binding was heterogenous across the dataset, with the majority of particles still engaging single-stranded DNA (Extended Data Fig. 2). Nonetheless, we identified a subset of particles, which was subsequently subclassified into two conformations (conformations I and II), that had unambiguously translocated onto dsDNA, representative of bona fide termination intermediates produced after fork convergence (Fig. 1b, Extended Data Figs. 2, 3a–e). While the configuration of the MCM C-tier differed between conformations I and II (Extended Data Fig. 3), in both cases the incoming dsDNA was bent by approximately 46° between the MCM N-tier and C-tier, necessitating distortion of the DNA duplex within the N-tier (Extended Data Fig. 3f). For conformation I, the nucleotide occupancy and interactions with the phosphate backbone of the leading-strand template are similar to replication fork-bound CMG¹³ (Fig. 1c, Extended Data Fig. 3b,g), suggesting a shared mechanism for translocation of CMG over single-stranded DNA and dsDNA¹⁵.

Having identified particles that had translocated onto dsDNA, we were able to build an atomic model of a terminated replisome (Fig. 1d). The overall architecture of CMG, Ctf4, Tof1–Csm3 and the non-catalytic module of Pol-ε (Pol-ε^{non-Cat}) was almost indistinguishable from previous structures^{13,19,20,21} (for details of the structure of Pol-ε, see Extended Data Fig. 4). We observed an additional, large region of density at the N-tier face of CMG beside Mcm3 and Mcm7, which closely approaches Csm3 and the dsDNA ahead of CMG, before extending away from the core of the complex, forming an elongated arm characteristic of the cullin subunit (Cdc53) of SCF^{Dia2} (Fig. 1d,e). The resolution of the cullin arm is relatively poor (precluding model building for Cdc53–Hrt1), due to a large degree of flexibility in this region, as highlighted by comparison of 3D classes (Extended Data Fig. 5a). We predict that this flexibility is important for conjugating the long K48-linked polyubiquitin chains required for Cdc48-dependent replisome disassembly⁸. Regardless, the orientation of SCF^{Dia2} can be unambiguously defined, placing the Cdc53 C terminus and Hrt1 ~45–70 Å from the

primary ubiquitylation site on Mcm7 (Lys29)^{8,22} (Extended Data Fig. 5a, b), consistent with previous structures of un-neddylated cullin-RING E3 ligases²³.

Density corresponding to the E3 ligase substrate-recognition module (Skp1–Dia2) is adjacent to the N-tier face of CMG (Fig. 1d, e). The N-terminal tetratricopeptide repeat domain of Dia2, which binds Ctf4 and Mrc1 (refs. 8,24,25), was not visible in our structure. However, clearsecondary structure and side chain density enabled us to build a de novo atomic model for the remainder of Dia2, encompassing the F-box (residues 211–247), 15 tandem leucine-rich repeats (LRRs) (248–716) and a C-terminal tail (717–732), which folds back onto the concave surface of the horseshoe-shaped LRRs (Extended Data Fig. 5c–k). The C-terminal end of the LRR domain forms an extensive interface with the N-tier of the Mcm3, Mcm5 and Mcm7 subunits of CMG (Fig. 1d, Extended Data Fig. 5l–n; see text below for a detailed description), demonstrating that Dia2 binds directly to CMG bound around dsDNA, equivalent to the situation after convergence of two replication forks.

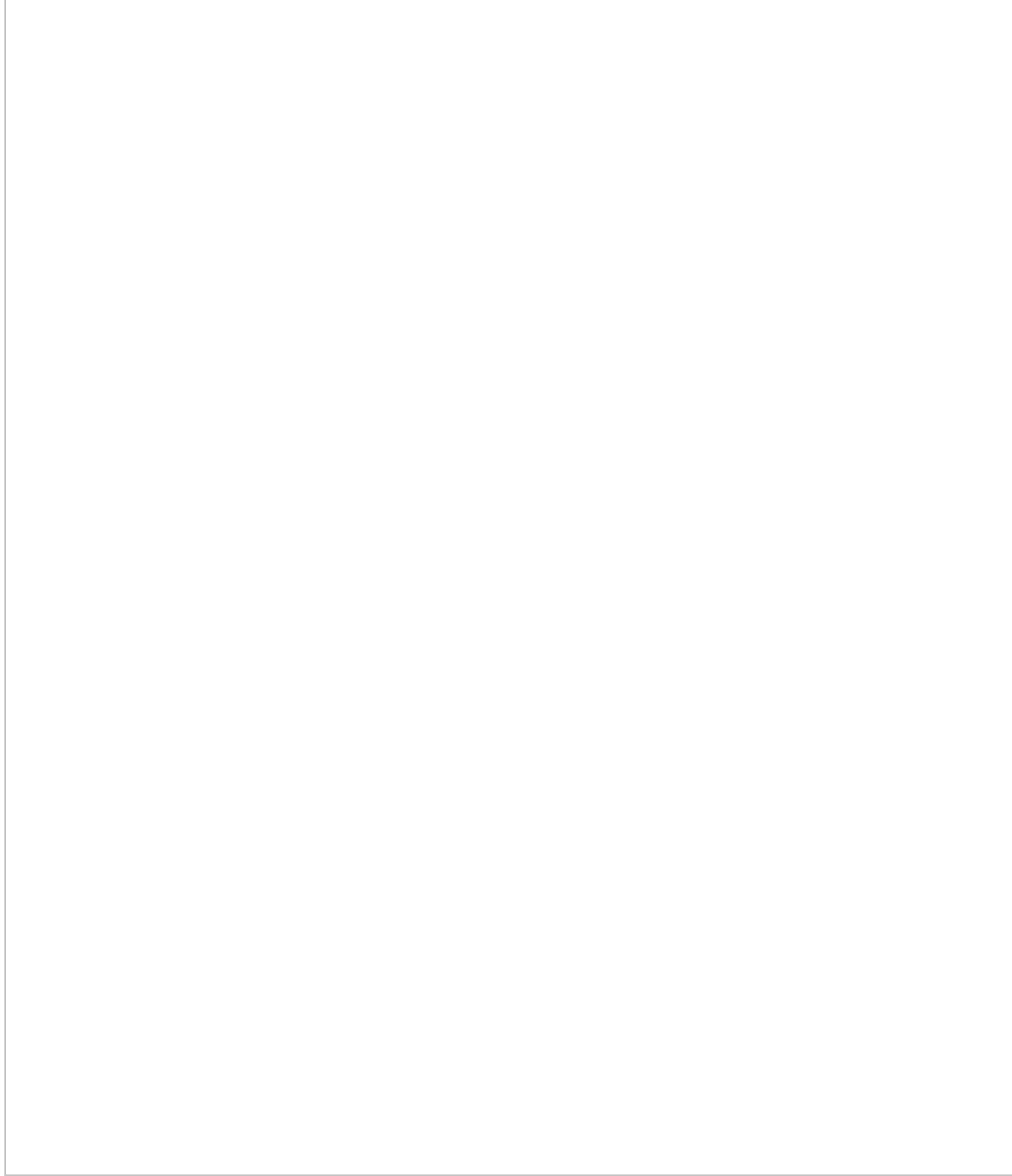
When DNA replication terminates at the end of linear chromosomes, CMG is thought to dissociate from DNA, at which point the loss of the excluded strand triggers CMG ubiquitylation^{8,9} (Fig. 1a). To establish how SCF^{Dia2} engages the replisome following termination at chromosome ends, we repeated cryo-EM sample preparation as described above, except in the absence of DNA. This yielded a 3D reconstruction of an ‘off DNA’ replisome at 3.9 Å resolution (Fig. 1f, Extended Data Fig. 6). Notably, binding of the Dia2 LRRs across Mcm3, Mcm5 and Mcm7 is indistinguishable from complexes bound around dsDNA (Fig. 1g, Extended Data Fig. 3e). Furthermore, comparison of our dsDNA-bound and off-DNA complexes with a previous structure of a replication fork-associated replisome¹³ revealed no conformational changes in the region of the MCM N-tier to which Dia2 binds (Fig. 1g, Extended Data Fig. 3e). Therefore, we conclude that termination does not induce conformational changes in CMG that are important for the regulation of CMG ubiquitylation by SCF^{Dia2} (ref. 26), either following fork convergence or when CMG dissociates from DNA.

Dia2^{LRR}–MCM interface

The extensive interface between Dia2^{LRR} and MCM is predominantly formed by the Mcm3 N-tier (helices α1 and α5 and the zinc-finger (ZnF) domain), which forms a cradle for the C terminus of Dia2^{LRR} (Fig. 2a, Extended Data Fig. 5l–n). In addition, the N terminus of Mcm7 wraps around the ZnF domain of Mcm3 and becomes sandwiched between Mcm3 and Dia2, while the ZnF domain of Mcm5 interacts with the C-terminal end of Dia2^{LRR}, at the periphery of the Dia2^{LRR}–MCM interface. The details of the residues involved are illustrated in Extended Data Fig. 5n.

Fig. 2: The MCM–Dia2^{LRR} interface is required for replisome disassembly.

 figure2



a, Overview of the MCM–Dia2^{LRR} interface. Leading-strand and lagging-strand template DNA is coloured orange and pink, respectively. Residues altered in Dia2^{LRR} mutants are in yellow. **b**, Reaction scheme to monitor CMG–Mcm7 ubiquitylation after Pif1-stimulated replication fork convergence in vitro¹⁴ (left). Immunoblot of

reactions conducted as indicated is also shown (right). The experiment was repeated three times. IP, immunoprecipitation; Mut, mutant; Ub, ubiquitin; WT, wild type. **c**, SDS-PAGE and immunoblotting of TAP-Sld5 immunoprecipitations from G1-arrested yeast cells with the indicated Dia2 alleles. The experiment was repeated twice. Also see Extended Data Fig. 7*i*. TAP, tandem affinity purification. **d**, Spot-dilution assay (tenfold serial dilutions) with the indicated yeast strains. The experiment was repeated three times. For gel source data, see Supplementary Fig. 1. YPD, yeast extract peptone dextrose.

To examine the significance of the Dia2^{LRR}-MCM interaction for CMG ubiquitylation and replisome disassembly, we generated a series of point mutants targeting the Dia2^{LRR}-MCM interface, in both Dia2 (Fig. 2*a*) and MCM. The majority of MCM mutants exhibited defects in the formation of the Mcm2-7-Cdt1 complex or in MCM loading (data not shown), probably because the Dia2 LRR binding site is positioned at the inter-hexamer interface in the MCM double hexamer²⁷. While this precluded analyses of Mcm7 ubiquitylation after convergence of two replication forks in vitro, we were able to purify a CMG complex containing mutations in Mcm3 and Mcm5, which, while being proficient for DNA replication, was defective for ubiquitylation of Mcm7 (Extended Data Fig. 7*a–d*). Dia2^{LRR} mutants formed stable tetrameric SCF^{Dia2} complexes and supported ubiquitylation of Ctf4 (Extended Data Fig. 7*e–g*).

Importantly, with the exception of Dia2-3A, the Dia2^{LRR} mutants were defective for ubiquitylation of Mcm7, both after replication fork convergence (Fig. 2*b*) and off DNA (Extended Data Fig. 7*h*), with Dia2-13A showing the most penetrant defect. Haploid yeast cells with the *dia2-13A* allele accumulated CMG in the G1 phase of the cell cycle (Fig. 2*c*, Extended Data Fig. 7*i*), reflecting a failure to disassemble CMG during replication termination in the S phase of the previous cell cycle¹. Furthermore, these cells exhibited a profound growth defect at 20 °C, indistinguishable from cells lacking Dia2 (ref. 24) (Fig. 2*d*). Together, these data demonstrate that the Dia2^{LRR}-MCM interface that we describe is essential for CMG ubiquitylation and replisome disassembly, both after fork convergence and when CMG dissociates from DNA.

Human replisome-CUL2^{LRR1} structure

Ubiquitylation of CMG in metazoa is driven by CUL2^{LRR1} (LRR1-CUL2-ELOB-ELOC-RBX1)^{4,5,6,7}. Although LRR1 displays no apparent sequence homology to Dia2, metazoan CUL2^{LRR1} ubiquitylates CMG on its MCM7 subunit^{4,5,6} and is suppressed by the excluded DNA strand^{9,10}, suggesting there might be common features of replisome association that are important for the regulation of both SCF^{Dia2} and CUL2^{LRR1}. To investigate this, we used our approach for human replisome

assembly²⁸ and a DNA substrate lacking a 5' flap, to determine a high-resolution structure of CUL2^{LRR1} in the human replisome (Fig. 3a, b, Extended Data Figs. 8, 9).

Fig. 3: Cryo-EM structures of human replisomes bound by CUL2^{LRR1}.

 figure3

a, Cryo-EM density of the human replisome bound by CUL2^{LRR1}. The density shown is a composite of focused maps (refer to Extended Data Fig. 8). **b**, Atomic models for the human replisome bound by CUL2^{LRR1} displayed using transparent surface rendering, except for CUL2^{LRR1}. Only CUL2^{LRR1}, DNA and the CUL2^{LRR1}-interacting regions of MCM are coloured (left). The model indicating the distance between RBX1 and K28/K29^{MCM7} is coloured according to subunit (right). **c**, LRR1 domain architecture diagram. The primary sequence and LRRs 1–9 are numbered. PH, pleckstrin homology. **d**, Overview of the interface between LRR1 and the replisome. The model is displayed using surface rendering, except for LRR1 and DNA.

The overall architecture of human CMG, AND-1, TIMELESS–TIPIN and Pol-ε are indistinguishable from our previous structure lacking CUL2^{LRR1} (ref. 28) (Fig. 3a, b, Extended Data Fig. 10a). LRR1 is positioned across the MCM N-tier, in close proximity to the parental DNA duplex. In addition, an elongated arm of lower-resolution density, into which the crystal structure of ELOB–ELOC–CUL2–RBX1 could be unambiguously docked²⁹, projects from the MCM N-tier in an analogous manner to yeast Cdc53–Hrt1 (Fig. 3b). Although metazoan CUL2 and yeast Cdc53 are tethered to their respective substrate adaptors (ELOB–ELOC–LRR1 for CUL2, Skp1–Dia2 for Cdc53) via very different interactions, the cullin C terminus and RING-box protein are similarly located in both cases, ~45–70 Å from the primary ubiquitylation sites in Mcm7 (refs. 8,9,22) (Figs. 1d, 3b). Furthermore, like Cdc53, CUL2 displays considerable conformational variability, which is probably important for the conjugation of long polyubiquitin chains onto MCM7 (refs. 8,30) (Extended Data Figs. 5a, 10b).

The majority of LRR1 was well resolved in our cryo-EM map (Extended Data Fig. 9e, k–n), which enabled de novo modelling of an N-terminal pleckstrin homology domain and a C-terminal LRR domain, which are connected by a flexible linker that stretches perpendicularly across the parental dsDNA (Fig. 3c, d). The pleckstrin homology

domain interacts with the ZnF domains of MCM2 and MCM6, parental dsDNA and the N-terminal region of the TIMELESS α -solenoid (Extended Data Fig. 10c, d), consistent with the reported role for TIMELESS–TIPIN in recruiting CUL2^{LRR1} to the replisome in *Caenorhabditis elegans*³⁰. The LRR domain comprises seven canonical and two irregular LRR motifs and forms a shallow arc, reaching from the parental dsDNA to the N-tier face of MCM3 and MCM5 (Fig. 3d, Extended Data Fig. 10e, f). The BC and CUL2 boxes, which link LRR1 to ELOB–ELOC–CUL2–RBX1, are situated between LRR repeats 8 and 9 (Fig. 3c, d, Extended Data Fig. 10e, f), and a two-stranded antiparallel β -sheet caps the LRR domain at its C-terminal end (Extended Data Fig. 10g). In addition, the C-terminal HMG box of AND-1 could be docked into a small region of density alongside ELOC and LRR1 (Extended data Fig. 10h, i), which was absent in 3D classes that lacked AND-1 (Extended Data Fig. 10j, k), indicating that AND-1 interacts with CUL2^{LRR1} in the human replisome.

Remarkably, despite the very different architectures of the LRR1 and Dia2 LRR domains, they bind to the same region of the MCM N-tier, but do so via completely different modes of interaction. The LRR1 LRR domain interacts predominantly with the three-stranded antiparallel β -sheet of the ZnF domain of MCM3, which extends the shallow arc of the LRR1 β -sheet (Extended Data Fig. 10l). This interface is augmented on one side by interactions between the MCM7 N terminus and the tip of the ZnF domain of MCM3 and LRR1 repeats 8 and 9 (Extended Data Fig. 10l, m). On the other side, MCM3 residues 3–8 and 164–174 are significantly rearranged upon CUL2^{LRR1} binding, such that the N terminus of MCM3, now projecting between the ZnF domains of MCM3 and MCM5, stabilizes an interaction between MCM3 residues 164–174 and a loop and short helix preceding LRR1 repeat 9 (Extended Data Fig. 10n). Finally, charged residues immediately preceding the β -strands of LRR1 repeats 4–7 form multiple polar contacts with the tip of the ZnF domain of MCM5 (Extended Data Fig. 10m). Further details are illustrated in Extended Data Fig. 10m, n.

Regulation of CMG ubiquitylation

Ubiquitylation of CMG by both SCF^{Dia2} and CUL2^{LRR1} is suppressed by the excluded DNA strand at replication forks^{8,9,10}; our discovery that Dia2 and LRR1 bind directly to a common site across the ZnF domains of MCM3 and MCM5 suggested that this region of MCM might be important for the regulation of ubiquitylation. In our recent structure of the human replisome bound to a replication fork²⁸, cryo-EM density that we attributed to the excluded strand was positioned in the channel between the ZnF domains of MCM3 and MCM5, consistent with previous structures of *Drosophila* and budding yeast CMG^{13,31,32}. To further validate our assignment of the excluded strand, we identified a subset of particles lacking CUL2^{LRR1} from our dataset of replisomes assembled without an excluded strand (Extended Data Figs. 8, 9g). In the resulting

density map, the MCM N-tier was identical to our previous map of replication fork-associated CMG²⁸, apart from a single region of density, extending from the fork junction between the ZnF domains of MCM3 and MCM5 (Fig. 4a, Extended Data Fig. 10o). This density was present only in the complex associated with the replication fork, thus confirming that it is contributed by the excluded DNA strand.

Fig. 4: A conserved mechanism for regulating replisome disassembly in eukaryotes.



a, Comparison of cryo-EM density maps for human replisome complexes (CMG, TIMELESS, TIPIN, CLASPIN, AND-1 and Pol-ε) bound to DNA substrates either lacking (left) or featuring (right; EMDB: EMD-13375 (ref. 28)) a 15-nucleotide 5' flap, representing the excluded DNA strand. Density is coloured according to chain occupancy using a radius of 5 Å, with the excluded strand coloured manually in UCSF Chimera. ssDNA, single-stranded DNA. **b**, Alternative views of the ZnF domains of MCM3 and MCM5 during replication elongation (red box, excluded strand present²⁸) and termination (green box, excluded strand absent). In the upper panel of the red box, the dashed line shows a putative path for the excluded ssDNA beyond the density observed in **a**, right. In the lower panel of the red box, four sugar-phosphate backbone linkages were built into the excluded strand density (see **a**, right; EMDB: EMD-13375 (ref. 28)). *H. sapiens*, *Homo sapiens*. **c**, Model for the regulation of CMG ubiquitylation. LRR-interacting regions of MCM are occluded in the MCM double hexamer (see Extended Data Fig. 11a) and by the excluded DNA strand at replication forks (see **a**, **b**) (red box). Loss of the excluded strand upon termination allows LRR–MCM engagement, CMG ubiquitylation and replisome disassembly (green box).

Crucially, Fig. 4b shows that the presence of the excluded strand between the ZnF domains of MCM3 and MCM5 sterically blocks the engagement of the Dia2 and LRR1 LRR domains with MCM. As the LRR–MCM interaction is essential for

ubiquitylation of CMG and, in turn, replisome disassembly, the occlusion of this interface by the excluded strand provides an elegant and universal explanation for the regulation of replisome disassembly across yeasts and metazoa. Notably, the LRR domains of Dia2 and LRR1 are not demonstrably homologous in sequence or structure. Thus, we propose that the binding of Dia2^{LRR} and LRR1^{LRR} across the exit channel of the excluded strand reflects convergent evolution, probably indicative of a stringent evolutionary pressure to accurately regulate replisome disassembly, and thereby safeguard replication forks. This evolutionary constraint is not evident in parts of the replisome disassembly machinery that do not contribute to the regulation of CMG ubiquitylation. For example, the Dia2 tetratricopeptide repeat domain binds yeast Mrc1 and Ctf4, whereas the LRR1 pleckstrin homology domain binds human TIMELESS.

On the basis of our results, we propose the model summarized in Fig. 4c. Ubiquitylation of the MCM double hexamer is blocked by the occlusion of the LRR binding site at the inter-hexamer interface²⁷ (Extended Data Fig. 11a). This occlusion probably also suppresses ubiquitylation during the conversion of MCM double hexamers into pairs of active CMG helicases³³, before the lagging-strand template is excluded. Once bidirectional replication forks are established and elongation begins, the spooling of the excluded DNA strand between the ZnF domains of MCM3 and MCM5 sterically blocks LRR engagement on MCM. It is possible that the binding of proteins to the excluded strand may help to block LRR–MCM engagement. However, ubiquitylation of CMG is inhibited at reconstituted budding yeast replication forks in the absence of the lagging-strand machinery (Extended Data Fig. 11b, c), consistent with the excluded DNA alone being sufficient to suppress SCF^{Dia2} during elongation. In principle, the binding of yeast Dia2 to Mrc1 and Ctf4, and human LRR1 to TIMELESS and AND-1, could still occur at replication forks, even when the LRR–MCM interaction is blocked by the excluded strand. Accordingly, Mrc1–Ctf4 can support $\text{SCF}^{\text{Dia2-13A}}$ recruitment to reconstituted replisomes (Extended Data Fig. 11d, e). Critically, however, the essentiality of the LRR–MCM interaction for ubiquitylation of CMG will restrict replisome disassembly to termination, independent of the timing of E3 ligase recruitment, and irrespective of whether a replication fork terminates via fork convergence, or at a telomere.

Finally, we note that if the excluded strand is ever mispositioned, for example, during replication fork stalling or reversal, replisome disassembly could be triggered, due to premature LRR–MCM engagement. As such, the regulatory mechanism that we describe here may have implications for the stability of the replication fork under conditions of replication stress.

Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

Data availability

Cryo-EM density maps of the yeast replisome–SCF^{Dia2} complex on dsDNA have been deposited in the Electron Microscopy Data Bank (EMDB) under the following accession numbers: [EMD-13495](#) (full-complex unsharpened map, conformation I), [EMD-13496](#) (full-complex sharpened map, conformation I), [EMD-13497](#) (multibody refinement (MBR), MCM N-tier, conformation I), [EMD-13498](#) (MBR, MCM C-tier, conformation I), [EMD-13500](#) (full-complex unsharpened map, conformation II), [EMD-13512](#) (full-complex sharpened map, conformation II), [EMD-13513](#) (MBR, MCM N-tier, conformation II), [EMD-13514](#) (MBR, MCM C-tier, conformation II), [EMD-13515](#) (MBR, Dia2–Skp1), [EMD-13516](#) (MBR, Cdc45–GINS–Ctf4–Dpb2^{NTD}), [EMD-13517](#) (MBR, Pol-ε^{non-Cat}–Mcm5^{WH}) and [EMD-13518](#) (full complex enriched for Csm3–Tof1); composite maps produced using Phenix combine_focused_maps have been deposited under accession numbers [EMD-13537](#) (conformation I) and [EMD-13539](#) (conformation II). Cryo-EM density maps of the yeast replisome–SCF^{Dia2} complex in the absence of DNA have been deposited in the EMDB under the following accession numbers: [EMD-13519](#) (full-complex unsharpened map) and [EMD-13540](#) (MBR). Cryo-EM density maps of the human replisome–CUL2^{LRR1} complex used in model building have been deposited in the EMDB under the following accession numbers: [EMD-13494](#) (full complex, consensus refinement), [EMD-13491](#) (MBR, AND-1–CDC45–GINS), [EMD-13490](#) (MBR, ELONGIN–BC–LRR1–CUL2) and [EMD-13492](#) (MBR, CUL2–RBX1). An additional map of the core human replisome not engaged by CUL2^{LRR1} on a DNA substrate lacking a 5' flap has been deposited under the accession number [EMD-13534](#). Atomic coordinates have been deposited in the Protein Data Bank (PDB) with the accession numbers [7PMK](#) for the yeast replisome–SCF^{Dia2} complex on dsDNA (conformation I), [7PMN](#) for the yeast replisome–SCF^{Dia2} complex on dsDNA (conformation II) and [7PLQ](#) for the human replisome–CUL2^{LRR1} complex.

References

1. 1.

Maric, M., Maculins, T., De Piccoli, G. & Labib, K. Cdc48 and a ubiquitin ligase drive disassembly of the CMG helicase at the end of DNA replication. *Science* **346**, 1253596 (2014).

2. 2.

Moreno, S. P., Bailey, R., Campion, N., Herron, S. & Gambus, A. Polyubiquitylation drives replisome disassembly at the termination of DNA replication. *Science* **346**, 477–481 (2014).

3. 3.

Dewar, J. M., Budzowska, M. & Walter, J. C. The mechanism of DNA replication termination in vertebrates. *Nature* **525**, 345–350 (2015).

4. 4.

Dewar, J. M., Low, E., Mann, M., Raschle, M. & Walter, J. C. CRL2^{LRR1} promotes unloading of the vertebrate replisome from chromatin during replication termination. *Genes Dev.* **31**, 275–290 (2017).

5. 5.

Sonneville, R. et al. CUL-2^{LRR-1} and UBXN-3 drive replisome disassembly during DNA replication termination and mitosis. *Nat. Cell Biol.* **19**, 468–479 (2017).

6. 6.

Villa, F. et al. CUL2^{LRR1}, TRAIP and p97 control CMG helicase disassembly in the mammalian cell cycle. *EMBO Rep.* **22**, e52164 (2021).

7. 7.

Fan, Y. et al. LRR1-mediated replisome disassembly promotes DNA replication by recycling replisome components. *J. Cell Biol.* **220**, e202009147 (2021).

8. 8.

Deegan, T. D., Mukherjee, P. P., Fujisawa, R., Polo Rivera, C. & Labib, K. CMG helicase disassembly is controlled by replication fork DNA, replisome components and a ubiquitin threshold. *eLife* **9**, e60371 (2020).

9. 9.

Low, E., Chistol, G., Zaher, M. S., Kochanova, O. V. & Walter, J. C. The DNA replication fork suppresses CMG unloading from chromatin before termination. *Genes Dev.* **34**, 1534–1545 (2020).

10. 10.

Vrtis, K. B. et al. Single-strand DNA breaks cause replisome disassembly. *Mol. Cell* **81**, 1309–1318.e6 (2021).

11. 11.

Labib, K., Tercero, J. A. & Diffley, J. F. Uninterrupted MCM2-7 function required for DNA replication fork progression. *Science* **288**, 1643–1647 (2000).

12. 12.

Fu, Y. V. et al. Selective bypass of a lagging strand roadblock by the eukaryotic replicative DNA helicase. *Cell* **146**, 931–941 (2011).

13. 13.

Baretić, D. et al. Cryo-EM structure of the fork protection complex bound to CMG at a replication fork. *Mol. Cell* **78**, 926–940.e13 (2020).

14. 14.

Deegan, T. D., Baxter, J., Ortiz Bazan, M. A., Yeeles, J. T. P. & Labib, K. P. M. Pif1-family helicases support fork convergence during DNA replication termination in eukaryotes. *Mol. Cell* **74**, 231–244.e9 (2019).

15. 15.

Langston, L. D. & O'Donnell, M. E. An explanation for origin unwinding in eukaryotes. *eLife* **8**, e46515 (2019).

16. 16.

Plechanovova, A., Jaffray, E. G., Tatham, M. H., Naismith, J. H. & Hay, R. T. Structure of a RING E3 ligase and ubiquitin-loaded E2 primed for catalysis. *Nature* **489**, 115–120 (2012).

17. 17.

Kawakami, T. et al. NEDD8 recruits E2-ubiquitin to SCF E3 ligase. *EMBO J.* **20**, 4003–4012 (2001).

18. 18.

Xie, P. et al. The covalent modifier Nedd8 is critical for the activation of Smurf1 ubiquitin ligase in tumorigenesis. *Nat. Commun.* **5**, 3733 (2014).

19. 19.

Goswami, P. et al. Structure of DNA-CMG-Pole elucidates the roles of the non-catalytic polymerase modules in the eukaryotic replisome. *Nat. Commun.* **9**, 5061 (2018).

20. 20.

Zhou, J. C. et al. CMG-Pole dynamics suggests a mechanism for the establishment of leading-strand synthesis in the eukaryotic replisome. *Proc. Natl Acad. Sci. USA* **114**, 4141–4146 (2017).

21. 21.

Sun, J. et al. The architecture of a eukaryotic replisome. *Nat. Struct. Mol. Biol.* **22**, 976–982 (2015).

22. 22.

Maric, M., Mukherjee, P., Tatham, M. H., Hay, R. & Labib, K. Ufd1-Npl4 recruit Cdc48 for disassembly of ubiquitylated CMG helicase at the end of chromosome replication. *Cell Rep.* **18**, 3033–3042 (2017).

23. 23.

Zheng, N. et al. Structure of the Cul1–Rbx1–Skp1–F box^{Skp2} SCF ubiquitin ligase complex. *Nature* **416**, 703–709 (2002).

24. 24.

Morohashi, H., Maculins, T. & Labib, K. The amino-terminal TPR domain of Dia2 tethers SCF^{Dia2} to the replisome progression complex. *Curr. Biol.* **19**, 1943–1949 (2009).

25. 25.

Maculins, T., Nkosi, P. J., Nishikawa, H. & Labib, K. Tethering of SCF^{Dia2} to the replisome promotes efficient ubiquitylation and disassembly of the CMG helicase. *Curr. Biol.* **25**, 2254–2259 (2015).

26. 26.

Dewar, J. M. & Walter, J. C. Mechanisms of DNA replication termination. *Nat. Rev. Mol. Cell Biol.* **18**, 507–516 (2017).

27. 27.

Li, N. et al. Structure of the eukaryotic MCM complex at 3.8 Å. *Nature* **524**, 186–191 (2015).

28. 28.

Jones, M. L., Baris, Y., Taylor, M. R. G. & Yeeles, J. T. P. Structure of a human replisome shows the organisation and interactions of a DNA replication machine. *EMBO J.* <https://doi.org/10.15252/embj.2021108819> (2021).

29. 29.

Cardote, T. A. F., Gadd, M. S. & Ciulli, A. Crystal structure of the Cul2-Rbx1-EloBC-VHL ubiquitin ligase complex. *Structure* **25**, 901–911.e3 (2017).

30. 30.

Xia, Y., Fujisawa, R., Deegan, T. D., Sonneville, R. & Labib, K. P. M. TIMELESS-TIPIN and UBXN-3 promote replisome disassembly during DNA replication termination in *Caenorhabditis elegans*. *EMBO J.* **40**, e108053 (2021).

31. 31.

Eickhoff, P. et al. Molecular basis for ATP-hydrolysis-driven DNA translocation by the CMG helicase of the eukaryotic replisome. *Cell Rep.* **28**, 2673–2688.e8 (2019).

32. 32.

Yuan, Z. et al. DNA unwinding mechanism of a eukaryotic replicative CMG helicase. *Nat. Commun.* **11**, 688 (2020).

33. 33.

Champasa, K., Blank, C., Friedman, L. J., Gelles, J. & Bell, S. P. A conserved Mcm4 motif is required for Mcm2-7 double-hexamer formation and origin DNA unwinding. *eLife* **8**, e45538 (2019).

34. 34.

Hogg, M. et al. Structural basis for processive DNA synthesis by yeast DNA polymerase varepsilon. *Nat. Struct. Mol. Biol.* **21**, 49–55 (2014).

35. 35.

Yuan, Z., Georgescu, R., Schauer, G. D., O'Donnell, M. E. & Li, H. Structure of the polymerase ε holoenzyme and atomic model of the leading strand replisome. *Nat. Commun.* **11**, 3156 (2020).

36. 36.

Martin, E. C. et al. LRRpredictor—a new LRR motif detection method for irregular motifs of plant NLR proteins using an ensemble of classifiers. *Genes* **11**, 286 (2020).

37. 37.

Schulman, B. A. et al. Insights into SCF ubiquitin ligases from the structure of the Skp1–Skp2 complex. *Nature* **408**, 381–386 (2000).

38. 38.

Xing, W. et al. SCF^{FBXL3} ubiquitin ligase targets cryptochromes at their cofactor pocket. *Nature* **496**, 64–68 (2013).

39. 39.

Noguchi, Y. et al. Cryo-EM structure of Mcm2-7 double hexamer on DNA suggests a lagging-strand DNA extrusion model. *Proc. Natl Acad. Sci. USA* **114**, E9529–E9538 (2017).

40. 40.

Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D* **74**, 531–544 (2018).

Acknowledgements

We are grateful to R. T. Hay and J. E. Sale for critical feedback on the manuscript; V. Chandrasekharan, T. Dendooven, C. K. Lau and M. E. Wilkinson for practical advice on data processing and model building; S. Chen, G. Sharov, A. Yeates and B. Ahsan for smooth running of the MRC LMB EM facility; J. Grimmett and T. Darling for maintenance of scientific computing facilities; J. Shi for operation of the baculovirus facility; L. Sanchez-Pulido for discussion relating to the evolutionary conservation of Dia2 and LRR1; and C. Polo Rivera and A. Wood for help with flow cytometry analysis. We acknowledge Diamond for access to and support of the cryo-EM facilities

at the UK National Electron Bio-Imaging Centre (eBIC), proposals BI23268-20 and BI24557-4, funded by the Wellcome Trust, Medical Research Council (MRC), and Biotechnology and Biological Sciences Research Council (BBSRC). We are grateful to Y. Chaban, O. Raschdorf, S. Masiulis and D. Clare for providing support in the use of the facilities at the eBIC. Ubiquitin and yeast Uba1 were provided by A. Knebel (MRC PPU, Dundee) and antibodies were provided by MRC PPU Reagents and Services (<https://mrcppureagents.dundee.ac.uk>) unless otherwise stated. This work was supported by the MRC, as part of UK Research and Innovation (MRC grants MC_UP_1201/12 to J.T.P.Y. and MC_UU_12016/13 to K.P.M.L.) and the Wellcome Trust (reference 204678/Z/16/Z for a Sir Henry Wellcome Postdoctoral Fellowship to T.D.D.).

Author information

Author notes

1. Tom D. Deegan

Present address: MRC Human Genetics Unit, Institute of Genetics and Cancer, University of Edinburgh, Western General Hospital, Edinburgh, UK

2. These authors contributed equally: Michael Jenkyn-Bedford, Morgan L. Jones

Affiliations

1. MRC Laboratory of Molecular Biology, Cambridge, UK

Michael Jenkyn-Bedford, Morgan L. Jones, Yasemin Baris, Giuseppe Cannone & Joseph T. P. Yeeles

2. MRC Protein Phosphorylation and Ubiquitylation Unit, Sir James Black Centre, School of Life Sciences, University of Dundee, Dundee, UK

Karim P. M. Labib & Tom D. Deegan

Contributions

M.J.-B. performed yeast cryo-EM sample preparation, data acquisition and processing, model building, purification of yeast replisome proteins, created figures, and reviewed and edited the manuscript. M.L.J. performed human cryo-EM sample preparation, data acquisition and processing, model building, created figures, and reviewed and edited the manuscript. Y.B. performed purification of human replisome proteins. K.P.M.L.

took part in discussion, acquired funding, and reviewed and edited the manuscript. G.C. acquired the human cryo-EM data. J.T.P.Y. conceptualized and supervised the study, acquired funding, performed yeast biochemistry and optimization of human replisome assembly, purification of yeast and human replisome proteins, wrote the original draft, and reviewed and edited the manuscript. T.D.D. conceptualized and supervised the study, acquired funding, performed yeast biochemistry and genetics, yeast cryo-EM sample preparation, purification of yeast replisome proteins, wrote the original draft, and reviewed and edited the manuscript.

Corresponding authors

Correspondence to [Joseph T. P. Yeeles](#) or [Tom D. Deegan](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Andrew Deans and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer review reports are available.

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 Supporting data for cryo-EM investigation of *S. cerevisiae* dsDNA-bound replisome:SCF^{Dia2} complexes.](#)

a, Schematic of reconstitution approach used for preparation of cryo-EM sample representing terminating SCF^{Dia2}-bound replisome complexes after translocation onto dsDNA. A schematic of the DNA substrate used is shown in orange, with the 20 nt tract of methylphosphonate (MEP) linkages coloured red. **b**, Silver-stained SDS-PAGE gels analysing 100 µL fractions taken across 10–30% glycerol gradients, either lacking (top) or containing (bottom) crosslinking agents. Fractions 13+14 used for cryo-EM sample preparation are indicated. * = Cdc34-Ub; ** = Cdc34. Similar results were

observed for three independent sample preparations. **c**, Representative cryo-EM micrograph. **d**, Representative 2D class averages, 40 nm box width. **e**, Representative angular distribution of particle orientations. A correspondingly oriented model is shown to the right for reference. **f**, Fourier shell correlation graphs for maps used in model building. The resolution of reconstructions calculated at the FSC=0.143 criterion are reported in Extended Data Fig. 2. **g**, Model-to-map correlation graphs. **h**, Cryo-EM density maps relevant to model building, coloured by local resolution. For gel source data, see Supplementary Fig. 1.

Extended Data Fig. 2 Data processing pipeline related to *S. cerevisiae* replisome:SCF_{Dia2} complexes.

The approach used to subclassify complexes based on MCM C-tier conformation and DNA engagement are coloured orange; the approach used to derive cryo-EM reconstructions for model building or adjustment of regions outside MCM are coloured magenta. Reported resolutions are calculated based on the FSC=0.143 criterion (refer to Extended Data Fig. 1f).

Extended Data Fig. 3 DNA engagement by *S. cerevisiae* CMG following translocation onto dsDNA.

a, Comparison of conformations I and II demonstrating the similarity in the overall complex architecture. For clarity, conformation I is rendered as a surface, whilst conformation II is shown as a cartoon. **b**, Overview of MCM C-tier domains bound to dsDNA, highlighting ATPase site occupancy in conformations I and II. **c**, Cryo-EM density (grey) at the MCM C-tier Mcm3-Mcm5 interface for each conformation. Mcm3, cyan; Mcm5, blue; AMP-PNP, red. **d**, Changes in ATPase site occupancy between conformations I and II correspond to movement of Mcm3/5/7 AAA+ domains and their DNA-binding loops (helix-2-insertion, H2I; presensor-1, PS1). Arrows indicate the relative movement of Mcm5 and Mcm3. The outward movement of Mcm3/Mcm7 in conformation II - associated with opening of the Mcm3/5 interface and loss of nucleotide at this ATPase site - leads to loss of the canonical contacts (described in panel g) formed between the Mcm3 H2I/PS1 loops and the leading-strand template DNA phosphate backbone. As such conformation II may reflect a partially disengaged state. **e**, Comparison of MCM:Dia2^{LRR} interface between conformations I and II, demonstrating lack of conformational changes in this region. **f**, Model of DNA in cryo-EM density (mesh) for conformation I demonstrating distortion of the B-form DNA duplex within the MCM N-tier; similar DNA density is observed within the MCM N-tier for conformation II. Approximate trajectories of DNA strands within the MCM N-tier are shown as dotted paths. **g**, Engagement of the leading-strand template DNA phosphate backbone by Mcm H2I/PS1 loops in

complexes that have translocated onto dsDNA is comparable to that previously observed for CMG bound to ssDNA¹³. Contacts shown for the representative Mcm6 subunit (conformation I). Specific contacts with the DNA phosphate moieties indicated by dashed yellow lines.

Extended Data Fig. 4 Insights into Pol ε positioning within the *S. cerevisiae* replisome.

a, Cryo-EM density derived from multi-body refinement (top) and corresponding atomic model (bottom) of the Pol ε non-catalytic module (Pol ε^{non-Cat}), with the exception of the Dpb2 NTD. **b**, Representative cryo-EM density (mesh) allowing de novo model building and adjustment of prior structures. **c**, Pol2-Mcm2 AAA+ domain interface. **d**, Interactions formed by the Mcm5 winged-helix (WH) domain with Pol ε^{non-Cat}. The Mcm5 WH is observed to contact regions of Pol2 (dark green) in addition to the Pol2 CysB and Dpb2 OB-fold domains. Interactions depicted in **c** and **d** have not been characterised previously. **e**, Regions of additional cryo-EM density observed for SCF^{Dia2}-bound replisome complexes on dsDNA, visible at low map contour levels. The crystal structure of the Pol2 catalytic domain (PDB: 4M8O³⁴) has been rigid-body fitted to additional density beside the MCM channel exit. In contrast to previous structures^{19,28,35}, this positions the Pol2 catalytic domain at the C-tier face of CMG, adjacent to the leading-strand template, in a state that may be important for leading-strand synthesis. Additional unassigned density between Ctf4^{SepB}, Tof1 and SCF^{Dia2} is outlined. **f**, Focused view of additional density attributed to the Pol2 catalytic domain (as in **e**, except rotated 180°). The C-terminal residue of the Pol2 catalytic domain (residue 1186), the N-terminal residue of the Pol2 non-catalytic domain (residue 1321), and density linking the two domains are indicated.

Extended Data Fig. 5 Supporting information for the Dia2 structure and its interaction with the *S. cerevisiae* replisome.

a, Cryo-EM density for Cdc53-Hrt1 for two 3D classes following signal subtraction/3D subclassification, demonstrating the flexibility observed in the position of replisome-bound SCF^{Dia2}. The position of Hrt1 is shown, derived from rigid-body fitting the crystal structure of homologous CUL1-RBX1 (PDB: 1LDK²³). The approximate distance between Hrt1 and the primary ubiquitylation site (K29^{Mcm7}) is indicated. **b**, Cryo-EM density map for Cdc53-Hrt1 with the crystal structure of homologous CUL1 (PDB: 1LDK²³) rigid-body fitted. **c**, Representative cryo-EM density (mesh) across different regions of Dia2. **d**, Dia2 domain architecture; TPR domain and nuclear localisation signal (NLS) as in ref. ²⁴. **e**, **f**, Alternative views of the Dia2 LRR domain coloured by repeat. The F-box domain and C-terminal tail are shown for context in **e**. **g**, Comparison of Dia2 LRR domain repeats to the LRR

consensus sequence³⁶. L is Leu/Val/Ile/Phe, N is Asn/Thr/Cys, x is any amino acid; we consider L_0 as the first repeat residue. The core $LxxLxL$ motif is highlighted. **h**, Comparison of Dia2 F-box to the consensus sequence³⁷. Exact matches coloured red, conservative differences coloured green. **i**, The Dia2 LRR domain (repeats 1 and 2) closely approaches the parental dsDNA (orange: leading-strand template; pink: lagging-strand template) and Csm3. The region of Csm3 upstream of the DNA-binding motif (DBM) is observed to interact with the Dia2 LRR domain β -sheet and Dia2 C-terminal tail, however cryo-EM density for this region was insufficient to identify details of this interaction. Dia2 residues which are positioned close to DNA are labelled. **j**, Dia2-Skp1 interaction. **k**, Interactions of Skp1 with alternative LRR-domain-containing F-box proteins, to show similarity with **j**. Human SKP1-SKP2: PDB:1FQV³⁷; human SKP1-FBXL3: PDB: 4I6J³⁸. **l**, Overview of interaction between Dia2 LRR domain and Mcm subunits. **m**, The MCM:Dia2^{LRR} interaction does not involve the concave β -sheet surface of Dia2^{LRR}. **n**, Detail of the MCM:Dia2^{LRR} interface. Residues altered in Dia2^{LRR} mutants are yellow. For the Mcm3 ZnF, interaction networks are subdivided into those above (A) and below (B) the position of the Mcm7 N-terminus.

[Extended Data Fig. 6 Supporting data for cryo-EM investigation of *S. cerevisiae* replisome:SCF-Dia2 complexes assembled in the absence of DNA.](#)

a, Silver-stained SDS-PAGE gel analysing 100 μ L fractions taken across a 10-30% GraFix gradient. Fractions 1-17 (of 23) shown. Fractions 11-13 used for cryo-EM sample preparation are indicated. Large-scale sample preparation was performed once; similar results were observed in three independent small-scale gradient preparations. **b**, Representative cryo-EM micrograph. **c**, Representative 2D class averages, 40 nm box width. **d**, Angular distribution of particle orientations contributing to cryo-EM density map (Fig. 1f). **e**, Fourier shell correlation curve for the multi-body refinement map presented in Fig. 1f. **f**, Cryo-EM density map (related to Fig. 1f) coloured by local resolution. For gel source data, see Supplementary Fig. 1.

[Extended Data Fig. 7 Supporting data for functional analyses of Dia2 and MCM mutants.](#)

a, Coomassie-stained SDS-PAGE gel of purified CMG complex containing mutations in Mcm3 and Mcm5 at Dia2^{LRR}-MCM interface **b**, Reaction scheme for in vitro replication of 9.7 kb forked DNA template using CMG and the indicated replication proteins. **c**, Reaction conducted as in panel **b** with wildtype or mutant CMG. Samples were separated on an alkaline agarose gel and visualised by auto-radiography. This

experiment was performed twice. **d**, In vitro CMG ubiquitylation reaction in the absence of DNA. The indicated proteins were incubated in the presence of ubiquitin and ATP and then visualised by SDS-PAGE and immunoblotting. This experiment was repeated three times. **e**, Positions of residues mutated in Dia2 LRR domain. LRR repeats 12-15 are coloured and numbered as in Extended Data Fig. 5e. **f**, Residues are coloured according to the Dia2 mutant in which they are present; all residues shown were mutated in Dia2-13A. Dia2-8A featured the following mutations: D632A, F657A, I662A, Y665A, Q694A, I698A, T699A and Y716A. **f**, Coomassie-stained SDS-PAGE gel of purified SCF^{Dia2} complexes containing Dia2^{LRR} mutants **g**, In vitro Ctf4 ubiquitylation reaction. The indicated proteins were incubated in the presence of ubiquitin and ATP and then visualised by SDS-PAGE and immunoblotting. The Hrt1 immunoblot serves as a loading control for SCF^{Dia2}. This experiment was repeated twice. **h**, Reaction conducted as in panel **d** with the indicated Dia2^{LRR} mutants. This experiment was repeated three times. **i**, DNA content of G1-arrested cells from experiment in Fig. 2c was monitored by flow cytometry after propidium iodide staining. The proportion of G1 cells, expressed as a percentage of the total cells, is given. For details of gating strategy and assignment of the G1 peak see Supplementary Fig. 2. For gel source data, see Supplementary Fig. 1.

Extended Data Fig. 8 Data processing pipeline related to *H. sapiens* replisome:CUL2^{LRR1} complexes.

Each pathway describing the generation of a discrete reconstruction is given its own colour. The final reconstructions are coloured red and boxed.

Extended Data Fig. 9 Supporting data for cryo-EM investigation of *H. sapiens* replisome: CUL2^{LRR1} complexes.

a, Schematic of reconstitution approach used for preparation of replisomes bound to CUL2^{LRR1} for cryo-EM. A schematic of the DNA substrate used is shown with a 39 nucleotide 3' arm and no 5' arm. **b**, Silver-stained SDS-PAGE gels analysing 100 µL fractions taken across 10-30% glycerol gradients, either lacking (top) or containing (bottom) crosslinking agents. Fractions 15+16 used for cryo-EM sample preparation are indicated. This experiment was performed twice. **c**, Representative cryo-EM micrograph. **d**, Representative 2D class averages, 40 nm box width. **e–j**, (Top) cryo-EM reconstructions coloured by local resolution according to inset keys (Bottom) angular distribution of particle orientations. **e**, Consensus refinement for replisome:CUL2^{LRR1} fully engaged. **f**, Consensus refinement for replisome:CUL2^{LRR1} where the LRR1^{PH} domain is bound but the LRRs are disengaged. **g**, Consensus refinement for particles lacking CUL2^{LRR1}. **h**, Multi-body refinement for AND-

1:CDC45:GINS. **i**, Multi-body refinement for LRR1:ELOB:ELOC:CUL2:AND-1-HMG. **j**, Multi-body refinement for CUL2:RBX1. **k**, Cryo-EM density for the LRR1 LRRs. **l**, Cryo-EM density for the LRR1 PH domain. **m**, Representative cryo-EM density for a LRR1 LRR domain β -strand at 3.5 Å resolution. **n**, Representative cryo-EM density for a LRR1 LRR domain α -helix at 3.7 Å resolution. **o**, Fourier-shell correlation (FSC) curves for the various maps used in model building. **p**, Map-to-model FSC curves for the complete model docked into the consensus refinement for replisomes fully engaged by CUL2^{LRR1}. For gel source data, see Supplementary Fig. 1.

Extended Data Fig. 10 Supporting information for the CUL2^{LRR1} structure and its interaction with the *H. sapiens* replisome.

a, Structural overlay of aligned model from replisomes bound to CUL2^{LRR1} (blue) and in the absence of CUL2^{LRR1} (red). **b**, Composite model and map representing the conformational variability of CUL2:RBX1. The model for the replisome, bound to LRR1 and ELOB-ELOC, is displayed using pipes and planks rendering and coloured according to subunit. Three representative 3D classes are displayed encompassing density for CUL2:RBX1 obtained through 3D classification without alignment. The distance between RBX1 and K29_{MCM7} is indicated as a dotted orange line and distances denoted in the inset key. **c**, Overview of the interface between the LRR1 PH domain and the replisome. Subunits interacting with the LRR1 PH domain are displayed using transparent surface rendering. Boxed regions indicate key interaction interfaces expanded in panel **d**. **d**, Detailed structural views of the interface between the LRR1 PH domain and 1: TIMELESS, 2(A): MCM6 ZnF, 2(B): dsDNA and 3: MCM2 ZnF. **e**, Model for the LRR1 LRRs with numbering indicating the order of the leucine-rich repeats. **f**, Consensus motif for the LRR1 LRRs. The sequence of each repeat is indicated with the positions of the key L₀, L₃ and L₅ residues highlighted in red. Repeats 1 and 9 represent irregular LRRs. L is Leu/Val/Ile/Phe, N is Asn/Thr/Cys, x is any amino acidL is Leu/Val/Ile/Phe, N is Asn/Thr/Cys, x is any amino acid. **g**, LRR1 model docked into transparent cryo-EM density with the capping 2-stranded β -sheet highlighted in gold. **h**, Overview of the LRR1:ELOB:ELOC:CUL2:AND-1 interface. Models displayed docked into transparent cryo-EM density with MCM subunits visualised using surface rendering. **i**, Structure of the AND-1 HMG box (PDB:2D7L) docked into the AND-1-dependent cryo-EM density adjacent to ELOC and LRR1. Selected hydrophobic core residues displayed. **j**, Map of the replisome bound to CUL2^{LRR1} in the absence of AND-1 coloured according to subunit. **k**, Cryo-EM density of the LRR1:ELOB:ELOC:CUL2 interface obtained through multi-body refinement from particles lacking AND-1. The density attributed to the AND-1 HMG box is dependent upon AND-1. **l**, Overview of the MCM:LRR1^{LRR} interface. MCM subunits displayed with additional transparent surface rendering and the order of the

$\text{LRR1}^{\text{LRRs}}$ numbered. Red-dashed boxes indicate key interaction sites, expanded in panel **m**. **m**, Detail of the MCM: LRR1^{LRR} interface involving contacts between the $\text{LRR1}^{\text{LRRs}}$ and 1 - MCM3, 2 - MCM5 ZnF and 3 - the MCM7 N-terminus. **n**, Model highlighting local rearrangements of MCM3 upon binding $\text{CUL2}^{\text{LRR1}}$. Structures in the absence (top) and presence (bottom) of $\text{CUL2}^{\text{LRR1}}$, coloured according to inset key, highlight the rearrangement of MCM3₍₁₋₉₎ and MCM3₍₁₆₄₋₁₇₄₎. **o**, Comparison of the $\text{CUL2}^{\text{LRR1}}$ -interacting regions of MCM from complexes assembled on a DNA substrate either lacking a 5'-flap (green) or containing a 15 nucleotide 5'-flap (gold, PDB: 7PFO²⁸). Complexes lacked $\text{CUL2}^{\text{LRR1}}$. The r.m.s.d. between the two structures for the region shown is 0.43 Å.

Extended Data Fig. 11 Supporting data for model for regulation of CMG ubiquitylation.

a, The MCM:Dia2^{LRR} interface is occluded in the inactive Mcm2-7 double hexamer. The structure of the budding yeast Mcm2-7 double hexamer is shown (PDB: 5BK4 (ref. ³⁹)): one Mcm2-7 hexamer is displayed as a cartoon, the other as a surface. Double-stranded DNA is coloured orange. The positions of the N-tier (N) and C-tier (C) are labelled for each hexamer. Inset: focused view of the regions of Mcm2-7 involved in interaction with the Dia2 LRR domain, demonstrating the inaccessibility of these regions to Dia2 in the context of a double hexamer. **b**, Reaction scheme for experiment in panel **c**, to monitor the suppression of CMG ubiquitylation by DNA in the absence of the indicated proteins (top), which are predicted to interact with the excluded DNA strand during lagging strand synthesis. Pif1 was omitted to block fork convergence. DNase was included after the replication step to release the replisome from DNA, which triggers CMG ubiquitylation⁸. **c**, Reaction conducted as in panel **b** and analysed by SDS-PAGE and immunoblot. This experiment was repeated twice. **d**, Reaction scheme for experiment in panel **e**, to monitor the interaction of SCF^{Dia2} with the replisome. **e**, Reaction conducted as in panel **d** and analysed by SDS-PAGE and immunoblot. * is rabbit IgG. This experiment was repeated twice. For gel source data, see Supplementary Fig. 1.

Extended Data Table 1 Cryo-EM statistics

Supplementary information

Supplementary Information

This file contains Supplementary Methods; Supplementary Tables 1 and 2; Supplementary Figures 1 and 2; and Supplementary References.

Reporting Summary

Peer Review File

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

Jenkyn-Bedford, M., Jones, M.L., Baris, Y. *et al.* A conserved mechanism for regulating replisome disassembly in eukaryotes. *Nature* **600**, 743–747 (2021). <https://doi.org/10.1038/s41586-021-04145-3>

- Received: 20 July 2021
- Accepted: 14 October 2021
- Published: 26 October 2021
- Issue Date: 23 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04145-3>

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

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04145-3>

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

- Article
- [Published: 01 December 2021](#)

Activation of homologous recombination in G1 preserves centromeric integrity

- [Duygu Yilmaz](#) ORCID: orcid.org/0000-0002-4009-3147^{1,2,3,4},
- [Audrey Furst](#)^{1,2,3,4},
- [Karen Meaburn](#)⁶,
- [Aleksandra Lezaja](#)⁵,
- [Yanlin Wen](#)⁵,
- [Matthias Altmeyer](#) ORCID: orcid.org/0000-0003-3780-1170⁵,
- [Bernardo Reina-San-Martin](#) ORCID: orcid.org/0000-0003-2083-6166^{1,2,3,4} &
- [Evi Soutoglou](#) ORCID: orcid.org/0000-0002-3724-5613^{1,2,3,4} nAff⁶

Nature volume 600, pages 748–753 (2021)

- 5895 Accesses
- 84 Altmetric
- [Metrics details](#)

Subjects

- [Centromeres](#)
- [DNA](#)
- [Double-strand DNA breaks](#)
- [Histone post-translational modifications](#)

Abstract

Centromeric integrity is key for proper chromosome segregation during cell division¹. Centromeres have unique chromatin features that are essential for centromere maintenance². Although they are intrinsically fragile and represent hotspots for chromosomal rearrangements³, little is known about how centromere integrity in response to DNA damage is preserved. DNA repair by homologous recombination requires the presence of the sister chromatid and is suppressed in the G1 phase of the cell cycle⁴. Here we demonstrate that DNA breaks that occur at centromeres in G1 recruit the homologous recombination machinery, despite the absence of a sister chromatid. Mechanistically, we show that the centromere-specific histone H3 variant CENP-A and its chaperone HJURP, together with dimethylation of lysine 4 in histone 3 (H3K4me2), enable a succession of events leading to the licensing of homologous recombination in G1. H3K4me2 promotes DNA-end resection by allowing DNA damage-induced centromeric transcription and increased formation of DNA–RNA hybrids. CENP-A and HJURP interact with the deubiquitinase USP11, enabling formation of the RAD51–BRCA1–BRCA2 complex⁵ and rendering the centromeres accessible to RAD51 recruitment and homologous recombination in G1. Finally, we show that inhibition of homologous recombination in G1 leads to centromeric instability and chromosomal translocations. Our results support a model in which licensing of homologous recombination at centromeric breaks occurs throughout the cell cycle to prevent the activation of mutagenic DNA repair pathways and preserve centromeric integrity.

This is a preview of subscription content

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Fig. 1: Centromeric DSBs recruit RAD51 throughout the cell cycle.

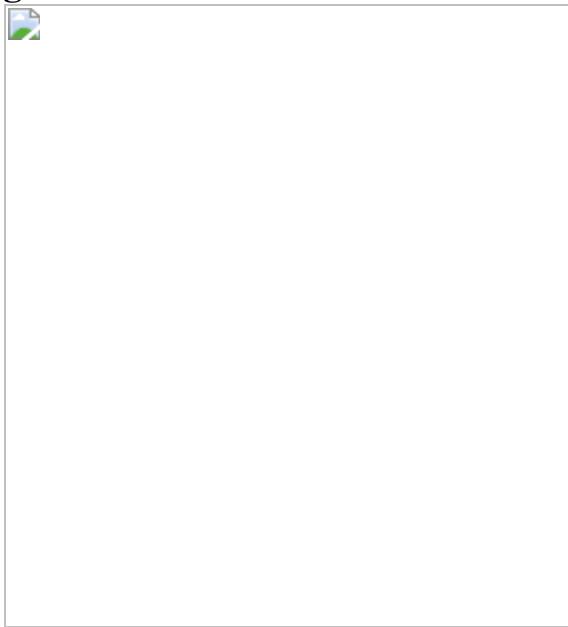


Fig. 2: H3K4me2 supports DNA-end resection by promoting centromeric transcription and increased R-loop formation in response to DSBs.

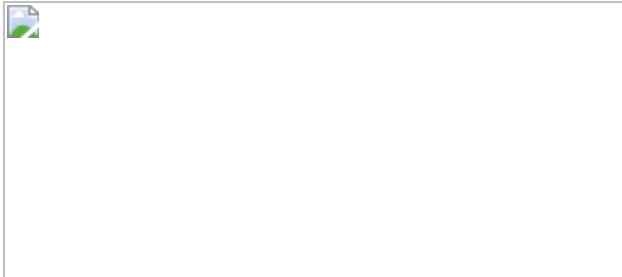


Fig. 3: The USP11–HJURP–CENP-A axis licenses HR at centromeric DSBs in G1.

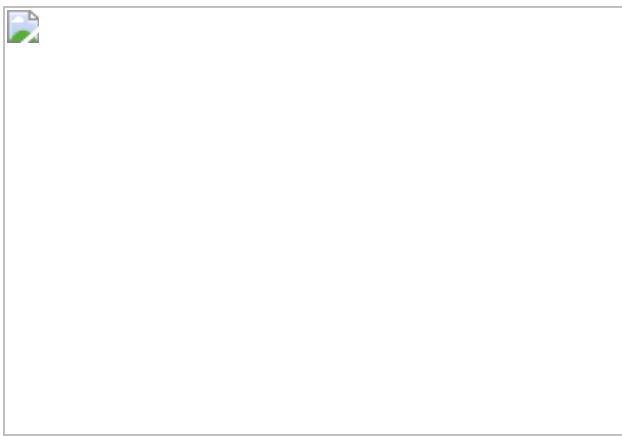


Fig. 4: Licensing of HR at centromeric breaks in G1 promotes centromeric integrity.



Data availability

[Source data](#) are provided with this paper.

References

1. 1.

Barra, V. & Fachinetti, D. The dark side of centromeres: types, causes and consequences of structural abnormalities implicating centromeric DNA. *Nat. Commun.* **9**, 4340 (2018).

2. 2.

McKinley, K. L. & Cheeseman, I. M. The molecular basis for centromere identity and function. *Nat. Rev. Mol. Cell Biol.* **17**, 16–29 (2016).

3. 3.

Martinez-A, C. & van Wely, K. H. M. Centromere fission, not telomere erosion, triggers chromosomal instability in human carcinomas. *Carcinogenesis* **32**, 796–803 (2011).

4. 4.

Takata, M. et al. Homologous recombination and non-homologous end-joining pathways of DNA double-strand break repair have overlapping roles in the maintenance of chromosomal integrity in vertebrate cells. *EMBO J.* **17**, 5497–5508 (1998).

5. 5.

Orthwein, A. et al. A mechanism for the suppression of homologous recombination in G1 cells. *Nature* **528**, 422–426 (2015).

6. 6.

Cheeseman, I. M. The kinetochore. *Cold Spring Harb. Perspect. Biol.* **6**, a015826 (2014).

7. 7.

Thompson, S. L., Bakhoum, S. F. & Compton, D. A. Mechanisms of chromosomal instability. *Curr. Biol.* **20**, R285–R295 (2010).

8. 8.

Beh, T. T. & Kalitsis, P. in *Centromeres and Kinetochores* (ed. Black, B. E.) **vol. 56**, 541–554 (Springer, 2017).

9. 9.

Padilla-Nash, H. M. et al. Jumping translocations are common in solid tumor cell lines and result in recurrent fusions of whole chromosome arms: Jumping Translocations in Solid Tumors. *Genes. Chromosomes Cancer* **30**, 349–363 (2001).

10. 10.

Mitelman, F., Mertens, F. & Johansson, B. A breakpoint map of recurrent chromosomal rearrangements in human neoplasia. *Nat. Genet.* **15**, 417–474 (1997).

11. 11.

Tsouroula, K. et al. Temporal and spatial uncoupling of DNA double strand break repair pathways within mammalian heterochromatin. *Mol. Cell* **63**, 293–305 (2016).

12. 12.

Hédouin, S., Grillo, G., Ivkovic, I., Velasco, G. & Francastel, C. CENP-A chromatin disassembly in stressed and senescent murine cells. *Sci. Rep.* **7**, (2017).

13. 13.

Quénet, D. & Dalal, Y. A long non-coding RNA is required for targeting centromeric protein A to the human centromere. *eLife* **3**, (2014).

14. 14.

Molina, O. et al. Epigenetic engineering reveals a balance between histone modifications and transcription in kinetochore maintenance. *Nat. Commun.* **7**, 13334 (2016).

15. 15.

Bergmann, J. H. et al. Epigenetic engineering shows H3K4me2 is required for HJURP targeting and CENP-A assembly on a synthetic human kinetochore: H3K4me2 and kinetochore maintenance. *EMBO J.* **30**, 328–340 (2011).

16. 16.

Nguyen, H. D. et al. Functions of replication protein A as a sensor of R loops and a regulator of RNaseH1. *Mol. Cell* **65**, 832–847.e4 (2017).

17. 17.

Hatchi, E. et al. BRCA1 recruitment to transcriptional pause sites is required for R-loop-driven DNA damage repair. *Mol. Cell* **57**, 636–647 (2015).

18. 18.

Castellano-Pozo, M. et al. R loops are linked to histone H3 S10 phosphorylation and chromatin condensation. *Mol. Cell* **52**, 583–590 (2013).

19. 19.

Kabeche, L., Nguyen, H. D., Buisson, R. & Zou, L. A mitosis-specific and R loop–driven ATR pathway promotes faithful chromosome segregation. *Science* **359**, 108–114 (2018).

20. 20.

Ouyang, J. et al. RNA transcripts stimulate homologous recombination by forming DR-loops. *Nature* **594**, 283–288 (2021).

21. 21.

Liu, S. et al. RNA polymerase III is required for the repair of DNA double-strand breaks by homologous recombination. *Cell* **184**, 1314–1329.e10 (2021).

22. 22.

Ohle, C. et al. Transient RNA–DNA hybrids are required for efficient double-strand break repair. *Cell* **167**, 1001–1013.e7 (2016).

23. 23.

Wang, Y., Li, X. & Hu, H. H3K4me2 reliably defines transcription factor binding regions in different cells. *Genomics* **103**, 222–228 (2014).

24. 24.

Rothkamm, K., Krüger, I., Thompson, L. H. & Löbrich, M. Pathways of DNA double-strand break repair during the mammalian cell cycle. *Mol. Cell. Biol.* **23**, 5706–5715 (2003).

25. 25.

Hayashi, T. et al. Mis16 and Mis18 are required for CENP-A loading and histone deacetylation at centromeres. *Cell* **118**, 715–729 (2004).

26. 26.

Foltz, D. R. et al. Centromere-specific assembly of CENP-A nucleosomes is mediated by HJURP. *Cell* **137**, 472–484 (2009).

27. 27.

Dunleavy, E. M. et al. HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres. *Cell* **137**, 485–497 (2009).

28. 28.

Soutoglou, E. et al. Positional stability of single double-strand breaks in mammalian cells. *Nat. Cell Biol.* **9**, 675–682 (2007).

29. 29.

Barnhart, M. C. et al. HJURP is a CENP-A chromatin assembly factor sufficient to form a functional de novo kinetochore. *J. Cell Biol.* **194**, 229–243 (2011).

30. 30.

Akimov, V. et al. UbiSite approach for comprehensive mapping of lysine and N-terminal ubiquitination sites. *Nat. Struct. Mol. Biol.* **25**, 631–640 (2018).

31. 31.

Niikura, Y. et al. CENP-A K124 ubiquitylation is required for CENP-A deposition at the centromere. *Dev. Cell* **32**, 589–603 (2015).

32. 32.

Bodor, D. L., Rodríguez, M. G., Moreno, N. & Jansen, L. E. T. in *Current Protocols in Cell Biology* (eds Bonifacino, J. S. et al.) cb0808s55 (John Wiley & Sons, 2012).

33. 33.

Zeitlin, S. G. et al. Double-strand DNA breaks recruit the centromeric histone CENP-A. *Proc. Natl Acad. Sci. USA* **106**, 15762–15767 (2009).

34. 34.

Read, L. R. Gene repeat expansion and contraction by spontaneous intrachromosomal homologous recombination in mammalian cells. *Nucleic Acids Res.* **32**, 1184–1196 (2004).

35. 35.

Khristich, A. N. & Mirkin, S. M. On the wrong DNA track: molecular mechanisms of repeat-mediated genome instability. *J. Biol. Chem.* **295**, 4134–4170 (2020).

36. 36.

Zhang, Y. & Jasin, M. An essential role for CtIP in chromosomal translocation formation through an alternative end-joining pathway. *Nat. Struct. Mol. Biol.* **18**, 80–84 (2011).

37. 37.

Zhou, J. et al. A first-in-class polymerase theta inhibitor selectively targets homologous-recombination-deficient tumors. *Nat. Cancer* **2**, 598–610 (2021).

38. 38.

van Sluis, M. & McStay, B. A localized nucleolar DNA damage response facilitates recruitment of the homology-directed repair machinery independent of cell cycle stage. *Genes Dev.* **29**, 1151–1163 (2015).

39. 39.

Vydzhak, O., Luke, B. & Schindler, N. Non-coding RNAs at the eukaryotic rDNA locus: RNA–DNA hybrids and beyond. *J. Mol. Biol.* **432**, 4287–4304 (2020).

40. 40.

Abraham, K. J. et al. Nucleolar RNA polymerase II drives ribosome biogenesis. *Nature* **585**, 298–302 (2020).

41. 41.

Feretzaki, M. et al. RAD51-dependent recruitment of TERRA lncRNA to telomeres through R-loops. *Nature* **587**, 303–308 (2020).

42. 42.

Giunta, S. et al. CENP-A chromatin prevents replication stress at centromeres to avoid structural aneuploidy. *Proc. Natl Acad. Sci. USA* **118**, e2015634118 (2021).

43. 43.

Mishra, P. K. et al. R-loops at centromeric chromatin contribute to defects in kinetochore integrity and chromosomal instability in budding yeast. *Mol. Biol. Cell* **32**, 74–89 (2021).

44. 44.

Ortega, P., Mérida-Cerro, J. A., Rondón, A. G., Gómez-González, B. & Aguilera, A. DNA–RNA hybrids at DSBs interfere with repair by homologous recombination. *eLife* **10**, e69881 (2021).

45. 45.

Marnef, A. & Legube, G. R-loops as Janus-faced modulators of DNA repair. *Nat. Cell Biol.* **23**, 305–313 (2021).

46. 46.

Nakamura, K. et al. Rad51 suppresses gross chromosomal rearrangement at centromere in *Schizosaccharomyces pombe*. *EMBO J.* **27**, 3036–3046 (2008).

47. 47.

McFarlane, R. J. & Humphrey, T. C. A role for recombination in centromere function. *Trends Genet.* **26**, 209–213 (2010).

48. 48.

Lacoste, N. et al. Mislocalization of the centromeric histone variant CenH3/CENP-A in human cells depends on the chaperone DAXX. *Mol. Cell* **53**, 631–644 (2014).

49. 49.

Jeffery, D. et al. CENP-A overexpression promotes distinct fates in human cells, depending on p53 status. *Commun. Biol.* **4**, 417 (2021).

50. 50.

Giunta, S. & Funabiki, H. Integrity of the human centromere DNA repeats is protected by CENP-A, CENP-C, and CENP-T. *Proc. Natl Acad. Sci. USA* **114**, 1928–1933 (2017).

51. 51.

Mateos-Gomez, P. A. et al. The helicase domain of Polθ counteracts RPA to promote alt-NHEJ. *Nat. Struct. Mol. Biol.* **24**, 1116–1123 (2017).

Acknowledgements

We thank K. Caldecott, D. Durocher, Y. Dalal for critical reading of the manuscript and members of the Soutoglou laboratory for comments; A. Nussenzweig for the BRCA1 antibody; A. Sfeir for the pol $\theta^{-/-}$ MEFs; A. Aguilera and K. Mekhail for RNaseH plasmids; D. Skalnik and G. Stewart for the SETD1A and ΔSET plasmids; Y. Dalal for the GFP–CENPA and GFP–HJURP plasmids; Daniel Foltz for the mCherry–LacI–HJURP plasmid; and the IGBMC Imaging Center for support. D.Y. was supported by the Ministerial fellowship from the UdS doctoral school and the Fondation pour la Recherche Medicale (FRM). The E.S. laboratory was supported by the European Research Council (ERC) under the European Union’s Horizon 2020 Research and Innovation program (ERC-2015-COG-682939) and by ANR-10-LABX-0030-INRT, managed by the Agence

Nationale de la Recherche under the program Investissements d’Avenir labelled ANR-10-IDEX-0002-02 and the Academy of Medical Sciences (AMS). M.A. acknowledges research funding from the Swiss National Science Foundation (PP00P3_179057).

Author information

Author notes

1. Evi Soutoglou

Present address: Genome Damage and Stability Centre, Sussex University, School of Life Sciences, University of Sussex, Brighton, UK

Affiliations

1. Institut de Génétique et de Biologie Moléculaire et Cellulaire, Illkirch, France

Duygu Yilmaz, Audrey Furst, Bernardo Reina-San-Martin & Evi Soutoglou

2. Institut National de la Santé et de la Recherche Médicale, U964, Illkirch, France

Duygu Yilmaz, Audrey Furst, Bernardo Reina-San-Martin & Evi Soutoglou

3. Centre National de Recherche Scientifique, UMR7104, Illkirch, France

Duygu Yilmaz, Audrey Furst, Bernardo Reina-San-Martin & Evi Soutoglou

4. Université de Strasbourg, Strasbourg, France

Duygu Yilmaz, Audrey Furst, Bernardo Reina-San-Martin & Evi Soutoglou

5. Department of Molecular Mechanisms of Disease, University of Zurich, Zurich, Switzerland

Aleksandra Lezaja, Yanlin Wen & Matthias Altmeyer

6. Genome Damage and Stability Centre, Sussex University, School of Life Sciences, University of Sussex, Brighton, UK

Karen Meaburn

Contributions

E.S. and D.Y. conceived the study. Most of the experiments and data analysis were performed by D.Y. Genomic instability in WT and Pol $\theta^{-/-}$ MEF and related experiments were performed by K.M. QIBC experiments and quantifications were performed by A.L. and Y.W. under the supervision of M.A. A.F. provided technical support. B.R.-S.-M provided Cas9 vectors and designed guide RNAs. E.S., D.Y., M.A. and B.R.-S.-M. wrote the manuscript.

Corresponding author

Correspondence to [Evi Soutoglou](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* Nausica Arnoult and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

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 Characterization of HR factors recruitment at centromeric and pericentromeric DSBs in G1 and G2.

a, Immunofluorescence confocal analysis of (upper part) NIH3T3 (EdU⁻ for G1 and RO-3306 arrested for G2) cells expressing GFP-CENP-A and (lower part) U2OS cells (thymidine arrested for G1 and RO-3306-arrested for G2) stained with DAPI and antibodies specific for the DNA damage marker γ-H2AX and the anti-centromere CREST in G1 and G2 phases of the cell cycle. On the right, percentage of cells with γ-H2AX colocalizing with GFP-CENP-A or CREST, corresponding to cells with at least one centromere colocalizing with γ-H2AX. **b**, Western blot analysis of GFP, γ-H2AX, phospho-ATM (pATMS1981) and tubulin in cells expressing dCas9-GFP, Cas9-GFP with a gRNA targeting minor satellite repeats (mi gRNA) or treated with the indicated concentrations of Neocarzinostatin (NCS). **c**, Cell cycle analysis by flow cytometry using propidium iodide and EdU in NIH3T3 (left panel) and U2OS cells (right panel), either untreated or treated with double thymidine or RO-3306. **d, e**, Immunofluorescence confocal analysis of (**d**) NIH3T3 cells expressing Cas9 + mi gRNA and GFP-CENP-A and (**e**) U2OS cells expressing Cas9 + gRNA targeting alpha satellite repeats stained with DAPI and antibodies specific for 53BP1 (used as a damage marker), CREST and RPA in G1 and G2. On the right, percentage of cells with RPA colocalizing with GFP-CENP-A or CREST. **f, g**, Immunofluorescence confocal analysis of (**f**) NIH3T3 cells expressing Cas9 + mi gRNA and GFP-CENP-A and (**g**) U2OS cells expressing Cas9 + gRNA targeting alpha satellite repeats stained with DAPI and antibodies specific for 53BP1, CREST and BRCA1 in G1 and G2. On the right,

percentage of cells with BRCA1 colocalizing with GFP-CENP-A or CREST. **h**, Percentage of cells with RAD51 colocalizing with γ -H2AX and CREST in HeLa and RPE1 cells expressing Cas9 + gRNA targeting alpha satellite repeats in G1. **i**, Percentage of cells with RAD51, RPA and BRCA1 recruitment at pericentromeric DSBs in NIH3T3 cells expressing Cas9 + gRNA targeting the major satellite repeats in G1 and G2. For confocal images, scale bars represent 5 μ m for G1 cells and 10 μ m for G2 cells. For definition of p values, statistic method and sample number see Statistics and Reproducibility section.

[Source data](#)

[**Extended Data Fig. 2 Recruitment of HR factors at centromeres is specific to DSB induction.**](#)

a, b, Immunofluorescence confocal analysis of NIH3T3 cells expressing Cas9 + mi gRNA stained with DAPI and **(a)** RAD51 antibody in 568nm or **(b)** γ -H2AX antibody in 568nm. Images in the other two channels (488nm and 647nm) demonstrate that the RAD51 signal is very specific, and it is not a bleed through signal from different antibody staining. **c**, Immunofluorescence of super resolution analysis of NIH3T3 cells expressing Cas9 + mi gRNA stained with DAPI, CREST, RAD51 and BRCA1 antibodies demonstrate that RAD51 and BRCA1 were detected simultaneously on the same centromere. **d**, Immunofluorescence confocal analysis of NIH3T3 cells expressing dCas9 + mi gRNA stained with DAPI, 53BP1, RPA or BRCA1 antibodies. **e**, Immunofluorescence confocal analysis of U2OS cells expressing dCas9 + gRNA targeting the alpha satellite repeats stained with DAPI, 53BP1, RPA or BRCA1 antibodies. For confocal images, scale bars represent 5 μ m. **f, g**, Immunofluorescence confocal analysis of **(f)** EdU⁻ NIH3T3 cells expressing GFP-CENP-A and **(g)** thymidine treated U2OS cells for G1 and RO-3306-arrested cells for G2 treated with Neocarzinostatin (NCS, 200 ng/ml), stained with DAPI and antibodies specific for γ -H2AX, CREST and RAD51. Percentage of cells with DSBs (γ -H2AX) and RAD51 colocalizing with GFP-CENP-A or CREST, corresponding to cells with at least one centromere colocalizing with γ -H2AX and RAD51, is shown on the right as mean \pm SD. For confocal images, scale bars represent 5 μ m for G1 cells, 10 μ m for G2 cells.

For definition of p values, statistic method and sample number see Statistics and Reproducibility section.

[Source data](#)

Extended Data Fig. 3 The role of H3K4me2 in HR factors recruitment at centromeric DSBs in G1 and G2.

a, Quantification of SETD1A mRNA by RT-qPCR in NIH3T3 cells depleted of SETD1A (siSETD1A), normalized to GAPDH and expressed as relative to control (siSCR). **b**, Western blot analysis of LSD1 and tubulin in NIH3T3 cells expressing GFP-dCas9-LSD1 and control cells. **c**, Chromatin immunoprecipitation analysis of H3K4me2 enrichment at centromeres over the Input and relative to the no-antibody control in siSETD1A vs siSCR NIH3T3 cells. **d**, Chromatin immunoprecipitation analysis of H3K4me2 enrichment at centromeres over the Input and relative to the no-antibody control in cells expressing dCas9 or dCas9-LSD1 and a gRNA targeting the minor satellite repeats (mi gRNA). **e**, Immunofluorescence confocal analysis of NIH3T3 cells expressing Cas9 + mi gRNA and dCas9 or dCas9-LSD1 and stained with DAPI and antibodies specific for γ-H2AX, 53BP1, RPA, BRCA1 and RAD51 in G1. **f**, Quantification of fold change of RPA, BRCA1 and RAD51 recruitment at centromeric DSBs in siSETD1A vs siSCR NIH3T3 cells in G2 and expressing Cas9 + mi gRNA. **g**, Quantification of fold change of RPA, BRCA1 and RAD51 recruitment at centromeric DSBs in NIH3T3 cells in G2, expressing Cas9 + mi gRNA and co-expressing dCas9-LSD1 or dCas9 alone. **h**, Western blot analysis of flag-SETD1A and flag-SETD1A-ΔSET in NIH3T3 cells using flag antibody. **i**, Chromatin immunoprecipitation analysis of H3K4me2 enrichment at centromeres over the Input and relative to the no-antibody control in cells depleted for SETD1A (siSETD1A) and reconstituted with WT SETD1A or with a truncated catalytically inactive mutant (SETD1A-ΔSET) or in control cells (siSCR). **j**, Quantification of fold change of RAD51 recruitment in G2 at pericentromeric DSBs in NIH3T3 cells expressing Cas9 + mi gRNA and siSETD1A relative to siSCR. For definition of p values, statistic method and sample number see Statistics and Reproducibility section.

[Source data](#)

Extended Data Fig. 4 The role of centromeric RNA in HR factors recruitment at centromeric DSBs in G1.

a, b, Quantification by RT-qPCR of fold change of centromeric RNA in NIH3T3 cells expressing dCas9 (left panels) or Cas9 (right panels) + gRNA targeting the minor satellite repeats (mi gRNA) **(a)** depleted of SETD1A (siSETD1A) relative to control (siSCR) and **(b)** expressing dCas9 or dCas9-LSD1 and mi gRNA. **c**, Western blot analysis of H3K4me2, Cas9 and Lamin A in NIH3T3 cells expressing either dCas9 + mi gRNA or Cas9 + mi gRNA. **d**, Chromatin immunoprecipitation analysis of H3K4me2 enrichment at pericentromeric heterochromatin over the input and relative to the no-antibody control in cells expressing Cas9 + mi gRNA. **e**, Immunofluorescence confocal analysis of G1 NIH3T3 cells expressing dRNaseH-GFP and co-expressing dCas9 + mi gRNA or Cas9 + mi gRNA and stained with DAPI and antibodies specific for γ-H2AX and CENP-A. **f**, Quantification of fold change of dRNaseH recruitment in G1 at centromeric DSBs in cells expressing Cas9 + mi gRNA and expressing siSETD1A vs siSCR or expressing dCas9 vs dCas9-LSD1. **g**, Quantification of fold change of RAD51 recruitment at centromeric DSBs in cells expressing Cas9 + mi gRNA and RNaseH vs dRNaseH. **h**, Quantification of fold change of 53BP1 and γ-H2AX intensity at centromeric DSBs co-stained with RPA, BRCA1 and RAD51 presented in Fig. 2h and Extended Data Fig. 4g, in cells expressing RNaseH vs dRNaseH. **i**, Chromatin immunoprecipitation analysis of H3K4me2 enrichment at centromeres over the input and relative to the no-antibody control in cells expressing GFP alone vs GFP-RNaseH. For confocal images, scale bars represent 5 μm. For definition of p values, statistic method and sample number see Statistics and Reproducibility section.

[Source data](#)

Extended Data Fig. 5 The role of HJURP and CENPA in eliciting RAD51 recruitment at centromeric DSBs in G1.

a, Quantification of CENP-A, HJURP and MIS18 mRNA levels by RT-qPCR in NIH3T3 cells expressing siCENP-A, siHJURP or siMIS18, normalized to GAPDH and expressed as relative to siSCR control. **b**, Quantification of fold change of RPA recruitment in G1 at centromeric DSBs in NIH3T3 cells expressing siHJURP, siCENP-A, siHJURP+siCENP-A or siMIS18 relative to siSCR and expressing Cas9 + mi gRNA. **c**, Quantification of fold change of RAD51 recruitment at centromeric DSBs in NIH3T3 cells depleted of CENP-C and CENP-N, components of the CCAN, relative to siSCR and expressing Cas9 + mi gRNA. **d**, Quantification of CENP-C and CENP-N mRNA by RT-qPCR in cells expressing siCENP-C or siCENP-N, normalized to GAPDH and relative to siSCR control. **e**, Quantification of fold change of RAD51 recruitment at lacO locus in NIH3T3 lacO-Isce-I-tet cells expressing mCherry-LacI or mCherry-lacI-HJURP and co-expressing I-SceI. **f**, Immunofluorescence confocal analysis of NIH3T3 lacO-Isce-I-tet cells expressing mCherry-LacI or mCherry-lacI-HJURP and co-expressing I-SceI and stained with DAPI, and antibodies against γ -H2AX and RAD51. **g**, Quantification of fold change of RAD51 recruitment at centromeric DSBs in NIH3T3 cells in G2, expressing Cas9 + mi gRNA and co-expressing siHJURP or siCENP-A vs siSCR. **h**, Western blot analysis of USP11, GFP and GFP-HJURP after GFP-IP in siUSP11 vs siSCR NIH3T3 cells in G1 or G2. The Input was 1 % of the extract used for the IP. **i**, Immunofluorescence confocal analysis of U2OS lacO-Isce-I-tet cells expressing mCherry-LacI or mCherry-lacI-HJURP and co-expressing GFP-USP11 and stained with DAPI. On the right, quantification of percentage of cells with colocalization of GFP-USP11 with LacO locus in G1. **j**, Quantification of percentage of cells with colocalization of GFP-USP11 with LacO locus in G2 arrested NIH3T3 lacO-Isce-I-tet cells expressing mCherry-LacI or mCherry-lacI-HJURP and co-expressing GFP-USP11. **k**, Quantification of USP11 mRNA by RT-qPCR in cells expressing siUSP11, normalized to GAPDH and relative to siSCR control. **l**, Immunofluorescence confocal analysis of NIH3T3 cells expressing Cas9 + mi gRNA and co-expressing dCas9 or dCas9-USP11 and stained with DAPI and antibodies specific for γ -H2AX and RAD51 in G1. On the right, fold change of percentage of cells with colocalization of RAD51 with γ -H2AX positive pericentromeric repeats. **m**, Western blot analysis of BRCA1, RAD51, RPA and tubulin in cells expressing siSETD1A, siUSP11,

siCENPA, or siHJURP vs siSCR. **n**, Quantification of γ -H2AX intensity at centromeres in NIH3T3 cells expressing Cas9 + mi gRNA and siSETD1A, siUSP11, siCENPA, or siHJURP vs siSCR. For confocal images, scale bars represent 5 μ m. For definition of p values, statistic method and sample number see Statistics and Reproducibility section.

[Source data](#)

[Extended Data Fig. 6 Characterization of USP11 role in licensing of HR at centromeric DSBs in G1.](#)

a, Western blot analysis of Ubiquitin (FK2 antibody) after GFP-IP under denaturing conditions in NIH3T3 cells expressing GFP or GFP-CENPA and co-expressing USP11-Flag vs Flag. The input was 1% of the extract used for the IP. **b**, Western blot analysis of ubiquitin (FK2 antibody) after GFP-IP under denaturing conditions in NIH3T3 cells expressing GFP or GFP-HJURP and siUSP11 vs siSCR. The input was 1% of the extract used for the IP. **c**, 2 biological replicates of western blot analysis showing the interaction of CENPA with HJURP in siUSP11 vs siSCR NIH3T3 cells (left panel). Quantification from western blot of CENPA interaction with HJURP in siUSP11 vs siSCR NIH3T3 cells (right panel). **d**, Quantification of fold change of CENP-A recruitment at LacO in cells expressing LacI-HJURP and co-expressing siUSP11 vs siSCR. **e**, Schematic representation of the SNAP technology. **f**, Western blot analysis of USP11, GFP and tubulin in HeLa cells expressing GFP-HJURP or not and co-expressing siUSP11 or siHJURP vs siSCR. **g**, Immunofluorescence confocal analysis of *de novo* deposited CENP-A-SNAP (TMR Star in Red) and total CENPA (Green) in HeLa expressing CENP-A-SNAP and co-expressing siHJURP or siUSP11 vs siSCR. **h**, Cell cycle analysis by flow cytometry using propidium iodide and EdU in HeLa CENP-A-SNAP cells depleted of USP11 relative to siSCR control. **i**, ChIP analysis of USP11 enrichment at pericentromeric heterochromatin over the Input and relative to the no-antibody control, in NIH3T3 cells expressing Cas9 + mi gRNA vs dCas9 + gRNA (left panel). Enrichment of USP11 at GAPDH promoter over the Input and relative to the no-antibody control, in NIH3T3 cells expressing Cas9 + gRNA targeting GAPDH promoter vs dCas9 + gRNA (right panel). **j**, ChIP analysis of

CENPA (left) and Histone H3 (right) enrichment at centromeres over the Input and relative to the no-antibody control, in NIH3T3 cells expressing Cas9 + mi gRNA vs dCas9 + mi gRNA. **k**, IF confocal analysis of *de novo* deposited CENP-A-SNAP in HeLa expressing CENP-A-SNAP and GFP-dCas9 (left) or Cas9 and stained with γ -H2AX antibody (right). **l**, Quantification of fold change of RAD51 recruitment in G1 at pericentromeric DSBs in NIH3T3 cells expressing Cas9 + gRNA targeting major satellite repeats, co-expressing dCas9-CENPA (upper panel) or dCas9-HJURP (lower panel) and depleted of USP11 (siUSP11) vs siSCR. For confocal images, scale bars represent 5 μ m. For definition of p values, statistic method and sample number see Statistics and Reproducibility section.

[Source data](#)

Extended Data Fig. 7 Assessment of genomic instability associated with lack of HR at centromeric DSBs in G1.

a, Immunofluorescence confocal analysis of NIH3T3 cells treated with RAD51 inhibitor (RAD51i) and irradiation (IR) and stained with DAPI and antibodies specific for γ -H2AX and RAD51; RAD51i was washed out or not after IR. **b**, Quantification of the number of translocations per 50 metaphase spreads in cells expressing Cas9 + mi gRNA and depleted of RAD51 (siRAD51), or of RAD51 and RAD52 (siRAD51+siRAD52) and compared to siSCR control. **c**, Quantification of RAD51 and RAD52 mRNA by RT-qPCR in NIH3T3 cells expressing siRAD51 or siRAD52, normalized to GAPDH and relative to siSCR. **d**, Quantification of the number of chromosomes with broken centromeres (lacking or with free centromeres) for 50 metaphase spreads in cells expressing Cas9 + mi gRNA and co-expressing siRAD51, or siRAD51+siRAD52 compared to siSCR. **e**, Quantification of fold enrichment of centromeric DNA after qPCR analysis in U2OS cells expressing Cas9 + gRNA targeting the alpha satellite repeats and siRAD51 alone or together with siRAD52 relative to siSCR control. **f**, Quantification of RAD52 mRNA by RT-qPCR in siRAD52 U2OS cells, normalized to GAPDH and relative to siSCR (left panel). Western blot analysis of RAD51 in siRAD51 U2OS cells compared to siSCR (right panel). **g**, Immunofluorescence confocal analysis of NIH3T3

cells expressing Cas9 + mi gRNA and RAD52-GFP, depleted of RAD51 (siRAD51) compared to siSCR and stained with DAPI and antibodies specific for γ -H2AX and RAD51. **h**, Quantification of the number of translocations in WT and PolQ $^{−/−}$ MEFs expressing Cas9 + mi gRNA and treated with RAD51i or DMSO as a control. For confocal images, scale bars represent 5 μ m. For definition of p values, statistic method and sample number see Statistics and Reproducibility section.

[Source data](#)

[Extended Data Fig. 8 Schematic representation of the step-wise model for the licensing of HR in G1 at centromeres.](#)

For details see text.

Supplementary information

[Supplementary Information](#)

This file contains Supplementary Fig. 1 (the uncropped western blots), Supplementary Fig. 2 (gating strategy for flow cytometry and cell cycle analysis) and Supplementary Tables 1–6 (siRNA references, guide RNA sequences and lists of plasmids, primers, antibodies and FISH probes used).

[Reporting Summary](#)

[Video 1](#)

Z-stack from immunofluorescence confocal analysis of a NIH 3T3 cell expressing GFP-CENP-A and stained with DAPI and antibodies specific for γ -H2AX (Alexa 568) and RAD51 (Alexa 647) in G1 phase of the cell cycle.

[Video 2](#)

Z-stack from immunofluorescence confocal analysis of a NIH 3T3 cell expressing GFP-CENP-A and stained with an antibody specific for RAD51

(Alexa 647) in G1 phase of the cell cycle.

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. 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

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Yilmaz, D., Furst, A., Meaburn, K. *et al.* Activation of homologous recombination in G1 preserves centromeric integrity. *Nature* **600**, 748–753 (2021). <https://doi.org/10.1038/s41586-021-04200-z>

- Received: 08 November 2020
- Accepted: 02 November 2021
- Published: 01 December 2021
- Issue Date: 23 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04200-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/s41586-021-04200-z>

- Article
- [Published: 08 December 2021](#)

β-NAD as a building block in natural product biosynthesis

- [Lena Barra](#) ORCID: [orcid.org/0000-0001-6566-4180](#)^{1 na1},
- [Takayoshi Awakawa](#) ORCID: [orcid.org/0000-0003-2575-3175](#)^{1,2 na1},
- [Kohei Shirai](#)¹,
- [Zhijuan Hu](#)¹,
- [Ghader Bashiri](#) ORCID: [orcid.org/0000-0002-5092-3749](#)³ &
- [Ikuro Abe](#) ORCID: [orcid.org/0000-0002-3640-888X](#)^{1,2}

Nature volume 600, pages 754–758 (2021)

- 6017 Accesses
- 80 Altmetric
- [Metrics details](#)

Subjects

- [Enzyme mechanisms](#)
- [Enzymes](#)

Abstract

β-Nicotinamide adenine dinucleotide (β-NAD) is a pivotal metabolite for all living organisms and functions as a diffusible electron acceptor and carrier in the catabolic arms of metabolism^{1,2}. Furthermore, β-NAD is involved in

diverse epigenetic, immunological and stress-associated processes, where it is known to be sacrificially utilized as an ADP-ribosyl donor for protein and DNA modifications, or the generation of cell-signalling molecules^{3,4}. Here we report the function of β -NAD in secondary metabolite biosynthetic pathways, in which the nicotinamide dinucleotide framework is heavily decorated and serves as a building block for the assembly of a novel class of natural products. The gatekeeping enzyme of the discovered pathway (SbzP) catalyses a pyridoxal phosphate-dependent [3+2]-annulation reaction between β -NAD and *S*-adenosylmethionine, generating a 6-azatetrahydroindane scaffold. Members of this novel family of β -NAD-tailoring enzymes are widely distributed in the bacterial kingdom and are encoded in diverse biosynthetic gene clusters. The findings of this work set the stage for the discovery and exploitation of β -NAD-derived natural products.

This is a preview of subscription content

Access options

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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Fig. 1: The discovered β -NAD-utilizing *sbz* biosynthetic pathway.

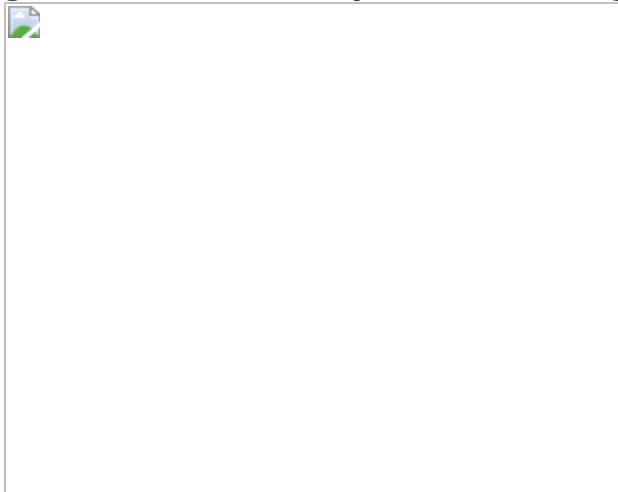


Fig. 2: Proposed reaction mechanism of the SbzP-mediated [3+2]-annulation reaction.

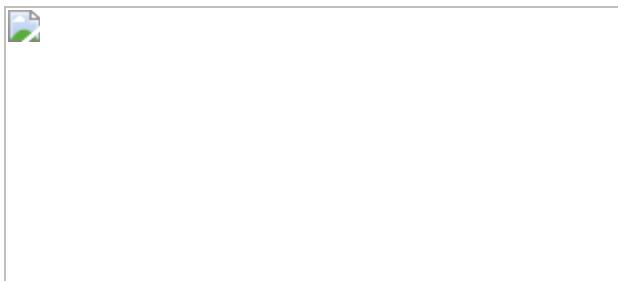
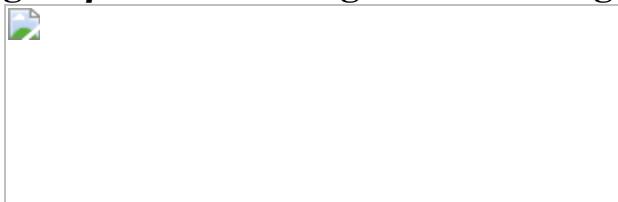


Fig. 3: β -NAD-utilizing SbzP homologues.



Data availability

Data that support the findings of this study are available within the paper and its [Supplementary Information](#), or are available from the corresponding authors upon request.

References

1. 1.

Walsh, C. T. & Tang, Y. *The Chemical Biology of Human Vitamins* (RSC, 2019).

2. 2.

Cantó, C., Menzies, K. J. & Auwerx, J. NAD⁺ metabolism and the control of energy homeostasis: a balancing act between mitochondria and the nucleus. *Cell Metab.* **22**, 31–53 (2015).

3. 3.

Depaix, A. & Kowalska, J. NAD analogs in aid of chemical biology and medicinal chemistry. *Molecules* **24**, 4187 (2019).

4. 4.

Schuller, M. et al. Molecular basis for DarT ADP-ribosylation of a DNA base. *Nature* **596**, 597–602 (2021).

5. 5.

Newman, D. J. & Cragg, G. M. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J. Nat. Prod.* **83**, 770–803 (2020).

6. 6.

Walsh, C. T. & Tang, Y. *Natural Product Biosynthesis: Chemical Logic and Enzymatic Machinery* (RSC, 2017).

7. 7.

Blin, K., Kim, H. U., Medema, M. H. & Weber, T. Recent development of antiSMASH and other computational approaches to mine secondary metabolite biosynthetic gene clusters. *Brief. Bioinform.* **20**, 1103–1113 (2018).

8. 8.

Houge-Frydrych, C. S. V., Gilpin, M. L., Skett, P. W. & Tyler, J. W. SB-203207 and SB-203208, two novel isoleucyl tRNA synthetase inhibitors from a *Streptomyces* sp. II. Structure determination. *J. Antibiot.* **53**, 364–372 (2000).

9. 9.

Stefanska, A. L., Cassels, R., Ready, S. J. & Warr, S. R. SB-203207 and SB-203208, two novel isoleucyl tRNA synthetase inhibitors from a *Streptomyces* sp. I. Fermentation, isolation and properties. *J. Antibiot.* **53**, 357–363 (2000).

10. 10.

Takahashi, A., Kurasawa, S., Ikeda, D., Okami, Y. & Takeuchi, T. Altemicidin, a new acaricidal and antitumor substance. I. Taxonomy, fermentation, isolation and physico-chemical and biological properties. *J. Antibiot.* **42**, 1556–1561 (1989).

11. 11.

Takahashi, A., Kurasawa, S., Ikeda, D., Okami, Y. & Takeuchi, T. Altemicidin, a new acaricidal and antitumor substance. II. Structure determination. *J. Antibiot.* **42**, 1562–1566 (1989).

12. 12.

Yan, Y., Liu, N. & Tang, Y. Recent developments in self-resistance gene directed natural product discovery. *Nat. Prod. Rep.* **37**, 879–892 (2020).

13. 13.

Hu, Z., Awakawa, T., Ma, Z. & Abe, I. Aminoacyl sulfonamide assembly in SB-203208 biosynthesis. *Nat. Commun.* **10**, 184 (2019).

14. 14.

Bull, J. A., Mousseau, J. J., Pelletier, G. & Charette, A. B. Synthesis of pyridine and dihydropyridine derivatives by regio- and stereoselective addition to n-activated pyridines. *Chem. Rev.* **112**, 2642–2713 (2012).

15. 15.

Eliot, A. C. & Kirsch, J. F. Pyridoxal phosphate enzymes: mechanistic, structural, and evolutionary considerations. *Annu. Rev. Biochem.* **73**, 383–415 (2004).

16. 16.

Du, Y. L. & Ryan, K. S. Pyridoxal phosphate-dependent reactions in the biosynthesis of natural products. *Nat. Prod. Rep.* **36**, 430–457 (2019).

17. 17.

Cook, P. D. & Holden, H. M. A structural study of GDP-4-keto-6-deoxy-d-mannose-3-dehydratase: caught in the act of geminal diamine formation. *Biochemistry* **46**, 14215–14224 (2007).

18. 18.

Hirayama, A., Miyanaga, A., Kudo, F. & Eguchi, T. Mechanism-based trapping of the quinonoid intermediate by using the K276R mutant of PLP-dependent 3-aminobenzoate synthase PctV in the biosynthesis of pactamycin. *ChemBioChem* **16**, 2484–2490 (2015).

19. 19.

Harmange Magnani, C. S. & Maimone, T. J. Dearomatic synthetic entry into the altemicidin alkaloids. *J. Am. Chem. Soc.* **143**, 7935–7939 (2021).

20. 20.

Chen, M., Liu, C. T. & Tang, Y. Discovery and biocatalytic application of a PLP-dependent amino acid α -substitution enzyme that catalyzes C-C bond formation. *J. Am. Chem. Soc.* **142**, 10506–10515 (2020).

21. 21.

Hai, Y., Chen, M., Huang, A. & Tang, Y. Biosynthesis of mycotoxin fusaric acid and application of a PLP-dependent enzyme for chemoenzymatic synthesis of substituted L-pipecolic acids. *J. Am. Chem. Soc.* **142**, 19668–19677 (2020).

22. 22.

Spiering, M. J., Moon, C. D., Wilkinson, H. H. & Schardl, C. L. Gene clusters for insecticidal lolaine alkaloids in the grass-endophytic fungus *Neotyphodium uncinatum*. *Genetics* **169**, 1403–1414 (2005).

23. 23.

Cui, Z. et al. Pyridoxal-5'-phosphate-dependent alkyl transfer in nucleoside antibiotic biosynthesis. *Nat. Chem. Biol.* **16**, 904–911 (2020).

24. 24.

Brzovic, P. et al. Reaction mechanism of *Escherichia coli* cystathione γ -synthase: direct evidence for a pyridoxamine derivative of vinylglyoxylate as a key intermediate in pyridoxal phosphate dependent γ -elimination and γ -replacement reactions. *Biochemistry* **29**, 442–451 (1990).

25. 25.

Bailey, H. J. et al. Human aminolevulinate synthase structure reveals a eukaryotic-specific autoinhibitory loop regulating substrate binding and product release. *Nat. Commun.* **11**, 1–12 (2020).

26. 26.

Raboni, S. et al. in *Comprehensive Natural Products II: Chemistry and Biology* Vol. 7 (eds Mander, L. & Liu, H.-W.) 273–315 (Elsevier, 2010).

27. 27.

Poulos, T. L. Heme enzyme structure and function. *Chem. Rev.* **114**, 3919–3962 (2014).

28. 28.

Ruszczycky, M. W., Zhong, A. & Liu, H. W. Following the electrons: peculiarities in the catalytic cycles of radical SAM enzymes. *Nat. Prod. Rep.* **35**, 615–621 (2018).

29. 29.

Mehta, A. P. et al. Radical *S*-adenosylmethionine (SAM) enzymes in cofactor biosynthesis: a treasure trove of complex organic radical rearrangement reactions. *J. Biol. Chem.* **290**, 3980–3986 (2015).

30. 30.

Walsh, C. T. Insights into the chemical logic and enzymatic machinery of NRPS assembly lines. *Nat. Prod. Rep.* **33**, 127–135 (2016).

31. 31.

Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* **9**, 1–8 (2008).

32. 32.

Roy, A., Kucukural, A. & Zhang, Y. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat. Protoc.* **5**, 725–738 (2010).

33. 33.

Roy, A., Yang, J. & Zhang, Y. COFACTOR: an accurate comparative algorithm for structure-based protein function annotation. *Nucleic Acids Res.* **40**, 471–477 (2012).

34. 34.

Bashiri, G., Rehan, A. M., Greenwood, D. R., Dickson, J. M. J. & Baker, E. N. Metabolic engineering of cofactor F420 production in *Mycobacterium smegmatis*. *PLoS ONE* **5**, 1–10 (2010).

35. 35.

Bashiri, G., Squire, C. J., Baker, E. N. & Moreland, N. J. Expression, purification and crystallization of native and selenomethionine labeled *Mycobacterium tuberculosis* FGD1 (Rv0407) using a *Mycobacterium smegmatis* expression system. *Protein Expr. Purif.* **54**, 38–44 (2007).

36. 36.

Oyugi, M. A., Bashiri, G., Baker, E. N. & Johnson-winters, K. Investigating the reaction mechanism of F 420-dependent glucose-6-phosphate dehydrogenase from *Mycobacterium tuberculosis*: kinetic analysis of the wild-type and mutant enzymes. *Biochemistry* **55**, 5566–5577 (2016).

37. 37.

McCarty, R. M., Krebs, C. & Bandarian, V. Spectroscopic, steady-state kinetic, and mechanistic characterization of the radical SAM enzyme QueE, which catalyzes a complex cyclization reaction in the biosynthesis of 7-deazapurines. *Biochemistry* **52**, 188–198 (2013).

Acknowledgements

This work was supported by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology, Japan (JSPS KAKENHI grant numbers JP16H06443, JP17H04763, JP19H04641, JP20H00490, JP20KK0173 and JP21K18246), the New Energy and Industrial Technology Development Organization (NEDO, Grant Number JPNP20011) and AMED (Grant Number JP21ak0101164), Japan Science and Technology Agency (JST SICORP grant number JPMJSC1701), UTEC-UTokyo FSI Research Grant Program, Mochida Memorial Foundation for Medical and Pharmaceutical Research, and Takeda Science Foundation, Kato Memorial Bioscience Foundation, The Asahi Glass Foundation, and a Sir Charles Hercus Fellowship (Health Research Council of New Zealand) to G.B. (17/058). We thank GlaxoSmithKline for providing *Streptomyces* sp. NCIMB 40513; G. Cao for providing *S. gilvosporeus* F607; M. Kobayashi for providing pHSA81; Deutsche Forschungsgemeinschaft (DFG BA 6870/1-1) for a postdoctoral research fellowship to L.B.; T. Mori and R. Ushimaru for critical reading of the manuscript; F. Kudo, T. Eguchi and T. Ueda for instruction on the stopped flow apparatus; and H. Zhang, F. Hayashi and Y. Ishii for measurement of NMR spectra at RIKEN Yokohama.

Author information

Author notes

1. These authors contributed equally: Lena Barra, Takayoshi Awakawa

Affiliations

1. Graduate School of Pharmaceutical Sciences, The University of Tokyo, Tokyo, Japan

Lena Barra, Takayoshi Awakawa, Kohei Shirai, Zhijuan Hu & Ikuro Abe

2. Collaborative Research Institute for Innovative Microbiology, The University of Tokyo, Tokyo, Japan

Takayoshi Awakawa & Ikuro Abe

3. Laboratory of Molecular and Microbial Biochemistry, School of Biological Sciences, The University of Auckland, Auckland, New Zealand

Ghader Bashiri

Contributions

T.A. and I.A. conceived the idea for the study. L.B., T.A. and I.A. developed the hypothesis and designed the experiments. L.B. and T.A. performed the *in vivo* and *in vitro* experiments. L.B. and K.S. performed compound isolation and characterization. L.B. and T.A. performed bioinformatic analysis to identify biosynthetic gene clusters and protein functions. L.B., T.A. and K.S. performed protein purification. G.B. expressed SbzF and provided the F420 cofactor regeneration system. T.A. and Z.H. performed the feeding experiments. All authors analysed and discussed the results. L.B., T.A. and I.A. prepared the manuscript.

Corresponding authors

Correspondence to [Takayoshi Awakawa](#) or [Ikuro Abe](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

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 Identification of the gatekeeping enzyme SbzP.](#)

a, LC-MS analysis of single gene expression culture extracts (UV trace 340 nm). **b**, Summary of isotope feeding experiment results and retrobiosynthetic considerations. **c**, LC-MS analysis of the productive SbzP substrate combination of β -NAD (**5**) and SAM (**6**) with concomitant appearance of dinucleotide product **7** (UV trace 258 nm). “Control” indicates incubation of substrates without enzyme.

[Extended Data Fig. 2 Gene deletion-guided *in vitro* reconstitution of the sbz pathway.](#)

a, LC-MS analysis of gene deletion strain culture extracts (UV trace 280 nm). **b**, LC-MS analysis of SbzQ-mediated conversion of **7** to **8** (EIC traces of substrate and product). **c**, LC-MS analysis of SbzI reaction of **8** to **9** (UV trace 340 nm). **d**, LC-MS analysis of deadenosylribosylation catalyzed by SbzN+O+H (EIC traces of substrate and product). **e**, LC-MS analysis of F₄₂₀-dependent reduction of **11** to **12** (EIC traces of substrate and products). **f**, LC-MS analysis of SbzE-mediated methylation towards **1** (UV trace 287 nm). **g**, Proposed steps in the deadenosylribosylation reaction. “Control” indicates incubation of substrates without enzyme(s).

[Extended Data Fig. 3 SbzP as a functionally and phylogenetically distinct PLP-dependent enzyme family.](#)

a, Time-resolved photospectroscopic changes of SbzP after addition of SAM (6) (see Supplementary Note 5 for further details). **b**, Phylogenetic analysis of SbzP and reported subfamily members of the AAT-I protein superfamily.

Supplementary information

Supplementary Information

This file contains Supplementary Text; Supplementary Figs. 1 – 81; Supplementary Table 1; Supplementary Notes 1–8 and Supplementary References

Reporting Summary

Peer Review File

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Barra, L., Awakawa, T., Shirai, K. *et al.* β -NAD as a building block in natural product biosynthesis. *Nature* **600**, 754–758 (2021).
<https://doi.org/10.1038/s41586-021-04214-7>

- Received: 29 March 2021
- Accepted: 04 November 2021
- Published: 08 December 2021
- Issue Date: 23 December 2021

- DOI: <https://doi.org/10.1038/s41586-021-04214-7>

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/s41586-021-04214-7>

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

- Article
- [Published: 08 December 2021](#)

Structures of the σ_2 receptor enable docking for bioactive ligand discovery

- [Assaf Alon](#) [ORCID: orcid.org/0000-0001-8102-5290](#)^{1 na1},
- [Jiankun Lyu](#) [ORCID: orcid.org/0000-0002-0461-5454](#)^{2 na1},
- [Joao M. Braz](#) [ORCID: orcid.org/0000-0001-8955-0735](#)^{3 na1},
- [Tia A. Tummino](#) [ORCID: orcid.org/0000-0001-9155-820X](#)^{2,4},
- [Veronica Craik](#) [ORCID: orcid.org/0000-0001-5463-1199](#)³,
- [Matthew J. O'Meara](#) [ORCID: orcid.org/0000-0002-3128-5331](#)⁵,
- [Chase M. Webb](#)^{2,4},
- [Dmytro S. Radchenko](#) [ORCID: orcid.org/0000-0001-5444-7754](#)^{6,7},
- [Yuri S. Moroz](#) [ORCID: orcid.org/0000-0001-6073-002X](#)⁸,
- [Xi-Ping Huang](#)^{9,10},
- [Yongfeng Liu](#) [ORCID: orcid.org/0000-0002-8649-7892](#)^{9,10},
- [Bryan L. Roth](#) [ORCID: orcid.org/0000-0002-0561-6520](#)^{9,10,11},
- [John J. Irwin](#)²,
- [Allan I. Basbaum](#) [ORCID: orcid.org/0000-0002-1710-6333](#)³,
- [Brian K. Shoichet](#) [ORCID: orcid.org/0000-0002-6098-7367](#)² &
- [Andrew C. Kruse](#) [ORCID: orcid.org/0000-0002-1467-1222](#)¹

Nature volume **600**, pages 759–764 (2021)

- 6375 Accesses
- 35 Altmetric

- [Metrics details](#)

Subjects

- [Neuropathic pain](#)
- [Receptor pharmacology](#)
- [Virtual drug screening](#)
- [X-ray crystallography](#)

Abstract

The σ_2 receptor has attracted intense interest in cancer imaging¹, psychiatric disease², neuropathic pain^{3,4,5} and other areas of biology^{6,7}. Here we determined the crystal structure of this receptor in complex with the clinical candidate roluperidone² and the tool compound PB28⁸. These structures templated a large-scale docking screen of 490 million virtual molecules, of which 484 compounds were synthesized and tested. We identified 127 new chemotypes with affinities superior to 1 μM , 31 of which had affinities superior to 50 nM. The hit rate fell smoothly and monotonically with docking score. We optimized three hits for potency and selectivity, and achieved affinities that ranged from 3 to 48 nM, with up to 250-fold selectivity versus the σ_1 receptor. Crystal structures of two ligands bound to the σ_2 receptor confirmed the docked poses. To investigate the contribution of the σ_2 receptor in pain, two potent σ_2 -selective ligands and one potent σ_1/σ_2 non-selective ligand were tested for efficacy in a mouse model of neuropathic pain. All three ligands showed time-dependent decreases in mechanical hypersensitivity in the spared nerve injury model⁹, suggesting that the σ_2 receptor has a role in nociception. This study illustrates the opportunities for rapid discovery of *in vivo* probes through structure-based screens of ultra large libraries, enabling study of underexplored areas of biology.

This is a preview of subscription content

Access options

[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.

Rent or Buy article

Get time limited or full article access on ReadCube.

from \$8.99

[Rent or Buy](#)

All prices are NET prices.

Additional access options:

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

Fig. 1: Structure of the σ_2 receptor and binding site ligand recognition.

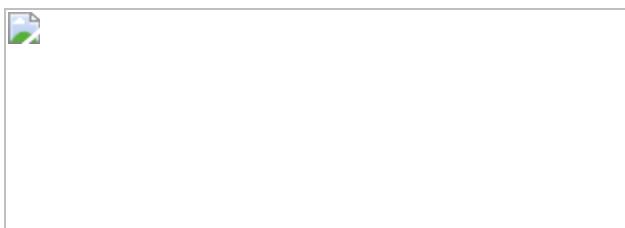


Fig. 2: Docking 490 million molecules against the σ_2 receptor.

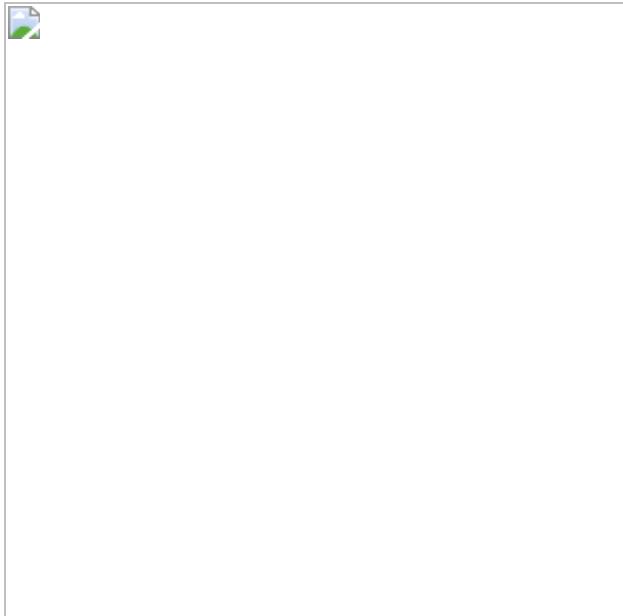


Fig. 3: High structural fidelity between docked and crystallographic poses of σ_2 receptor ligands.

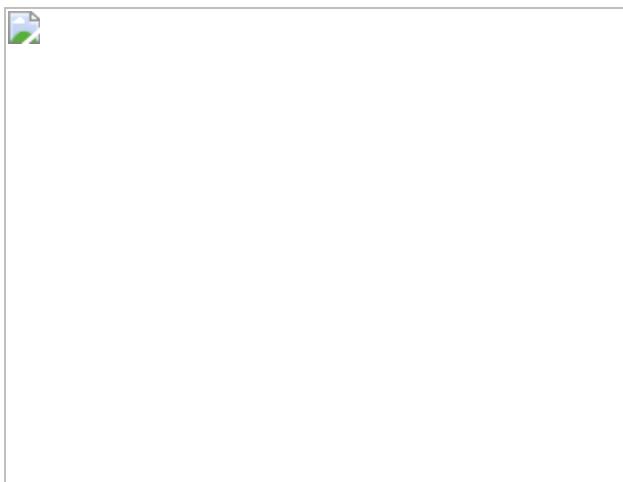
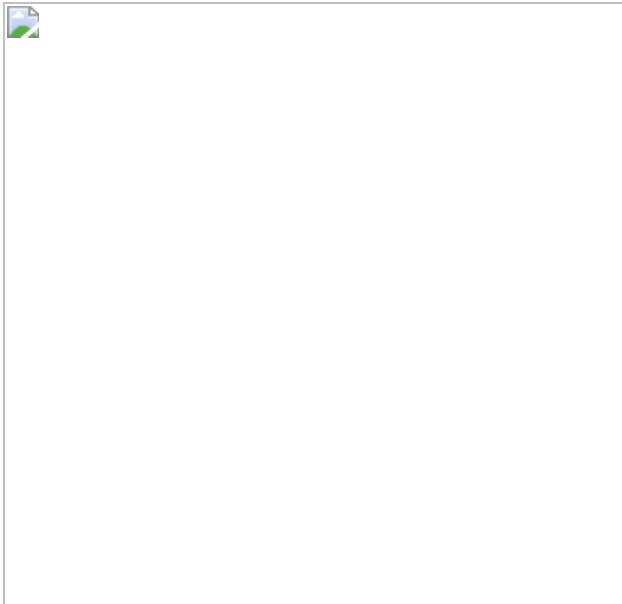


Fig. 4: $\sigma_{1/2}$ ligands are anti-allodynic in a model of neuropathic pain.



Data availability

The coordinates and structure factors for PB28-bound σ_2 , roluperidone-bound σ_2 , Z1241145220-bound σ_2 , Z4857158944-bound σ_2 and cholesterol-bound σ_2 have been deposited in the PDB with accession codes [7M93](#), [7M94](#), [7M95](#), [7M96](#) and [7MFI](#), respectively. The identities of the compounds docked in this study are freely available from the ZINC database (<http://zinc15.docking.org>) and active compounds may be purchased from Enamine. Any other data relating to this study are available from the corresponding authors on reasonable request. [Source data](#) are provided with this paper.

Code availability

DOCK 3.7 is freely available for non-commercial research <http://dock.compbio.ucsf.edu/DOCK3.7/>. A web-based version is available at <http://blaster.docking.org/>.

References

1. 1.

Waarde, A. V. et al. Potential applications for sigma receptor ligands in cancer diagnosis and therapy. *Biochim. Biophys. Acta* **1848**, 2703–2714 (2015).

2. 2.

Harvey, P. D. et al. Effects of roluperidone (MIN-101) on two dimensions of the negative symptoms factor score: reduced emotional experience and reduced emotional expression. *Schizophr. Res.* **215**, 352–356 (2020).

3. 3.

Sahn, J. J., Mejia, G. L., Ray, P. R., Martin, S. F. & Price, T. J. Sigma 2 receptor/Tmem97 agonists produce long lasting antineuropathic pain effects in mice. *ACS Chem. Neurosci.* **8**, 1801–1811 (2017).

4. 4.

Intagliata, S. et al. Discovery of a highly selective sigma-2 receptor ligand, 1-(4-(6,7-dimethoxy-3,4-dihydroisoquinolin-2(1H)-yl)butyl)-3-methyl-1H-benzo[d]imidazol-2(3H)-one (CM398), with drug-like properties and antinociceptive effects in vivo. *AAPS J.* **22**, 94 (2020).

5. 5.

Quadir, S. G. et al. The Sigma-2 receptor / transmembrane protein 97 (σ 2R/TMEM97) modulator JVW-1034 reduces heavy alcohol drinking and associated pain states in male mice. *Neuropharmacology* **184**, 108409 (2021).

6. 6.

Grundman, M. et al. A phase 1 clinical trial of the sigma-2 receptor complex allosteric antagonist CT1812, a novel therapeutic candidate for Alzheimer's disease. *Alzheimers Dement.* **5**, 20–26 (2018).

7. 7.

Riad, A. et al. Sigma-2 receptor/TMEM97 and PGRMC-1 increase the rate of internalization of LDL by LDL receptor through the formation of a ternary complex. *Sci. Rep.* **8**, 16845 (2018).

8. 8.

Abate, C. et al. PB28, the sigma-1 and sigma-2 receptors modulator with potent anti-SARS-CoV-2 activity: a review about its pharmacological properties and structure affinity relationships. *Front. Pharmacol.* **11**, 589810 (2020).

9. 9.

Shields, S. D., Eckert, W. A. & Basbaum, A. I. Spared nerve injury model of neuropathic pain in the mouse: a behavioral and anatomic analysis. *J. Pain* **4**, 465–470 (2003).

10. 10.

Hellewell, S. B. et al. Rat liver and kidney contain high densities of σ_1 and σ_2 receptors: characterization by ligand binding and photoaffinity labeling. *Eur. J. Pharmacol.* **268**, 9–18 (1994).

11. 11.

Hellewell, S. B. & Bowen, W. D. A sigma-like binding site in rat pheochromocytoma (PC12) cells: decreased affinity for (+)-benzomorphans and lower molecular weight suggest a different sigma receptor form from that of guinea pig brain. *Brain Res.* **527**, 244–253 (1990).

12. 12.

Hanner, M. et al. Purification, molecular cloning, and expression of the mammalian sigma₁-binding site. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 8072–8077 (1996).

13. 13.

Langa, F. et al. Generation and phenotypic analysis of sigma receptor type I (sigma1) knockout mice. *Eur. J. Neurosci.* **18**, 2188–2196 (2003).

14. 14.

Alon, A. et al. Identification of the gene that codes for the σ_2 receptor. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 7160–7165 (2017).

15. 15.

Ebrahimi-Fakhari, D. et al. Reduction of TMEM97 increases NPC1 protein levels and restores cholesterol trafficking in Niemann-pick type C1 disease cells. *Hum. Mol. Genet.* **25**, 3588–3599 (2016).

16. 16.

Bartz, F. et al. Identification of cholesterol-regulating genes by targeted RNAi screening. *Cell Metab.* **10**, 75 (2009).

17. 17.

Sanchez-Pulido, L. & Ponting, C. P. TM6SF2 and MAC30, new enzyme homologs in sterol metabolism and common metabolic disease. *Front. Genet.* **5**, 439 (2014).

18. 18.

Mahdessian, H. et al. TM6SF2 is a regulator of liver fat metabolism influencing triglyceride secretion and hepatic lipid droplet content. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 8918 (2014).

19. 19.

Vilner, B. J., John, C. S. & Bowen, W. D. Sigma-1 and sigma-2 receptors are expressed in a wide variety of human and rodent tumor cell lines. *Cancer Res.* **55**, 408–413 (1995).

20. 20.

Scott, L. L. et al. Small molecule modulators of σ_2 R/Tmem97 reduce alcohol withdrawal-induced behaviors. *Neuropsychopharmacology* **43**, 1867–1875 (2018).

21. 21.

Vázquez-Rosa, E. et al. Neuroprotective efficacy of a sigma 2 receptor/TMEM97 modulator (DKR-1677) after traumatic brain injury. *ACS Chem. Neurosci.* **10**, 1595–1602 (2019).

22. 22.

Stein, R. M. et al. Virtual discovery of melatonin receptor ligands to modulate circadian rhythms. *Nature* **579**, 609–614 (2020).

23. 23.

Schuller, M. et al. Fragment binding to the Nsp3 macrodomain of SARS-CoV-2 identified through crystallographic screening and computational docking. *Sci. Adv.* **7**, eabf8711 (2021).

24. 24.

Schmidt, H. R. & Kruse, A. C. The molecular function of σ receptors: past, present, and future. *Trends Pharmacol. Sci.* **40**, 636–654 (2019).

25. 25.

Caffrey, M. & Cherezov, V. Crystallizing membrane proteins using lipidic mesophases. *Nat. Protoc.* **4**, 706–731 (2009).

26. 26.

Long, T. et al. Structural basis for human sterol isomerase in cholesterol biosynthesis and multidrug recognition. *Nat. Commun.* **10**, 2452 (2019).

27. 27.

Audet, M. & Stevens, R. C. Emerging structural biology of lipid G protein-coupled receptors. *Protein Sci.* **28**, 292–304 (2019).

28. 28.

Schmidt, H. R. et al. Crystal structure of the human σ_1 receptor. *Nature* **532**, 527–530 (2016).

29. 29.

Hubler, Z. et al. Accumulation of 8,9-unsaturated sterols drives oligodendrocyte formation and remyelination. *Nature* **560**, 372–376 (2018).

30. 30.

Lyu, J. et al. Ultra-large library docking for discovering new chemotypes. *Nature* **566**, 224–229 (2019).

31. 31.

Fischer, A., Smieško, M., Sellner, M. & Lill, M. A. Decision making in structure-based drug discovery: visual inspection of docking results. *J. Med. Chem.* **64**, 2489–2500 (2021).

32. 32.

Cendán, C. M., Pujalte, J. M., Portillo-Salido, E., Montoliu, L. & Baeyens, J. M. Formalin-induced pain is reduced in σ_1 receptor knockout mice. *Eur. J. Pharmacol.* **511**, 73–74 (2005).

33. 33.

Puente, B. D. L. et al. Sigma-1 receptors regulate activity-induced spinal sensitization and neuropathic pain after peripheral nerve injury. *Pain* **145**, 294–303 (2009).

34. 34.

Cendán, C. M., Pujalte, J. M., Portillo-Salido, E. & Baeyens, J. M. Antinociceptive effects of haloperidol and its metabolites in the formalin test in mice. *Psychopharmacology* **182**, 485–493 (2005).

35. 35.

Romero, L. et al. Pharmacological properties of S1RA, a new sigma-1 receptor antagonist that inhibits neuropathic pain and activity-induced spinal sensitization. *Br. J. Pharmacol.* **166**, 2289–2306 (2012).

36. 36.

Bruna, J. et al. Efficacy of a novel sigma-1 receptor antagonist for oxaliplatin-induced neuropathy: a randomized, double-blind, placebo-controlled phase IIa clinical trial. *Neurotherapeutics* **15**, 178–189 (2018).

37. 37.

Vela, J. M., Merlos, M. & Almansa, C. Investigational sigma-1 receptor antagonists for the treatment of pain. *Expert Opin. Investig. Drugs* **24**, 883–896 (2015).

38. 38.

Kooistra, A. J. et al. GPCRdb in 2021: integrating GPCR sequence, structure and function. *Nucleic Acids Res.* **49**, D335–D343 (2021).

39. 39.

Hauser, A. S. et al. Pharmacogenomics of GPCR drug targets. *Cell* **172**, 41–54 (2018).

40. 40.

Kroeze, W. K. et al. PRESTO-Tango as an open-source resource for interrogation of the druggable human GPCRome. *Nat. Struct. Mol. Biol.* **22**, 362–369 (2015).

41. 41.

Nastasi, G. et al. S2RSLDB: a comprehensive manually curated, internet-accessible database of the sigma-2 receptor selective ligands. *J. Cheminformatics* **9**, 3 (2017).

42. 42.

Huang, X.-P. et al. Allosteric ligands for the pharmacologically dark receptors GPR68 and GPR65. *Nature* **527**, 477–483 (2015).

43. 43.

Wang, S. et al. Structure of the D2 dopamine receptor bound to the atypical antipsychotic drug risperidone. *Nature* **555**, 269–273 (2018).

44. 44.

Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

45. 45.

Lomize, M. A., Pogozheva, I. D., Joo, H., Mosberg, H. I. & Lomize, A. L. OPM database and PPM web server: resources for positioning of proteins in membranes. *Nucleic Acids Res.* **40**, D370–D376 (2012).

46. 46.

Gordon, D. E. et al. A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **583**, 459–468 (2020).

47. 47.

Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).

48. 48.

Kabsch, W. XDS. *Acta Crystallogr. D* **66**, 132 (2010).

49. 49.

Alford, R. F. et al. An integrated framework advancing membrane protein modeling and design. *PLoS Comput. Biol.* **11**, e1004398 (2015).

50. 50.

Mccoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 674 (2007).

51. 51.

Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D* **75**, 861–877 (2019).

52. 52.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).

53. 53.

The PyMOL Molecular Graphics System v.2.5 (Schrödinger).

54. 54.

Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

55. 55.

Chu, U. B. & Ruoho, A. E. Sigma receptor binding assays. *Curr. Protoc. Pharmacol.* **71**, 1.34.1–1.34.21 (2015).

56. 56.

Weiner, S. J. et al. A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* **106**, 765–784 (1984).

57. 57.

Meng, E. C., Shoichet, B. K. & Kuntz, I. D. Automated docking with grid-based energy evaluation. *J. Comput. Chem.* **13**, 505–524 (1992).

58. 58.

Gallagher, K. & Sharp, K. Electrostatic contributions to heat capacity changes of DNA–ligand binding. *Biophys. J.* **75**, 769–776 (1998).

59. 59.

Mysinger, M. M. & Shoichet, B. K. Rapid context-dependent ligand desolvation in molecular docking. *J. Chem. Inf. Model.* **50**, 1561–1573 (2010).

60. 60.

Stein, R. M. et al. Property-unmatched decoys in docking benchmarks. *J. Chem. Inf. Model.* **61**, 699–714 (2021).

61. 61.

Gu, S., Smith, M. S., Yang, Y., Irwin, J. J. & Shoichet, B. K. Ligand strain energy in large library docking. *J. Chem. Inf. Model.* **61**, 4331–4341 (2021).

62. 62.

Carpenter, B. et al. Stan: a probabilistic programming language. *J. Stat. Softw.* **76**, 1–32 (2017).

63. 63.

Bürkner, P.-C. brms: an R package for Bayesian multilevel models using Stan. *J. Stat. Softw.* **80**, 1–28 (2017).

64. 64.

Gelman, A., Goodrich, B., Gabry, J. & Vehtari, A. R-squared for Bayesian regression models. *Am. Statistician* **73**, 307–309 (2018).

65. 65.

Kay, M. *tidybayes: tidy data and geoms for Bayesian models*. <https://doi.org/10.5281/zenodo.1308151> (Zenodo, 2020).

66. 66.

Wickham, H. *ggplot2: elegant graphics for data analysis* (Springer, 2016).

67. 67.

Wickham, H. et al. Welcome to the tidyverse. *J. Open Source Softw.* **4**, 1686 (2019).

68. 68.

R Core Team. *R: a Language and Environment for Statistical Computing*. <http://www.R-project.org> (R Foundation for Statistical Computing, 2018).

69. 69.

Skubic, C., Vovk, I., Rozman, D. & Križman, M. Simplified LC-MS method for analysis of sterols in biological samples. *Molecules* **25**, 4116 (2020).

70. 70.

Longo, P. A., Kavran, J. M., Kim, M.-S. & Leahy, D. J. Transient mammalian cell transfection with polyethylenimine (PEI). *Methods*

Enzymol. **529**, 227–240 (2013).

71. 71.

Besnard, J. et al. Automated design of ligands to polypharmacological profiles. *Nature* **492**, 215–220 (2012).

72. 72.

Scherrer, G. et al. Dissociation of the opioid receptor mechanisms that control mechanical and heat pain. *Cell* **137**, 1148–1159 (2009).

73. 73.

Muralidharan, A. et al. Identification and characterization of novel candidate compounds targeting 6- and 7-transmembrane μ -opioid receptor isoforms. *Br. J. Pharmacol.* **178**, 2709–2726 (2021).

74. 74.

Chaplan, S. R., Bach, F. W., Pogrel, J. W., Chung, J. M. & Yaksh, T. L. Quantitative assessment of tactile allodynia in the rat paw. *J. Neurosci. Methods* **53**, 55–63 (1994).

75. 75.

Solorzano, C. et al. Primary afferent and spinal cord expression of gastrin-releasing peptide: message, protein, and antibody concerns. *J. Neurosci.* **35**, 648–657 (2015).

76. 76.

Notredame, C., Higgins, D. G. & Heringa, J. T-coffee: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* **302**, 205–217 (2000).

77. 77.

Liebschner, D. et al. Polder maps: improving OMIT maps by excluding bulk solvent. *Acta Crystallogr. D* **73**, 148–157 (2017).

Acknowledgements

Funding to support this research was provided by NIH grant R01GM119185, the Vallee Foundation and the Sanofi iAwards program (A.C.K.); by DARPA grant HR0011-19-2-0020 and NIH grant R35GM122481 (B.K.S.); and by grant GM133836 (J.J.I.). GM/CA @ APS has been funded by the National Cancer Institute (ACB-12002) and the National Institute of General Medical Sciences (AGM-12006 and P30GM138396). This research used resources of the Advanced Photon Source, a U.S. 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. The Eiger 16M detector at GM/CA-XSD was funded by NIH grant S10 OD012289. We thank K. Arnett and the Harvard Center for Macromolecular Interactions for support with biophysical experiments including circular dichroism and SEC–MALS, and C. Vidoudez and the Harvard Center for Mass Spectrometry for performing mass spectrometry analysis of sterols. Molecular graphics and analyses were performed with UCSF Chimera, which was developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311.

Author information

Author notes

1. These authors contributed equally: Assaf Alon, Jiankun Lyu, Joao M. Braz

Affiliations

1. Department of Biological Chemistry and Molecular Pharmacology, Blavatnik Institute, Harvard Medical School, Boston, MA, USA

Assaf Alon & Andrew C. Kruse

2. Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, CA, USA

Jiankun Lyu, Tia A. Tummino, Chase M. Webb, John J. Irwin & Brian K. Shoichet

3. Department of Anatomy, University of California, San Francisco, San Francisco, CA, USA

Joao M. Braz, Veronica Craik & Allan I. Basbaum

4. Graduate Program in Pharmaceutical Sciences and Pharmacogenomics, University of California, San Francisco, San Francisco, CA, USA

Tia A. Tummino & Chase M. Webb

5. Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI, USA

Matthew J. O'Meara

6. Enamine, Kyiv, Ukraine

Dmytro S. Radchenko

7. Taras Shevchenko National University of Kyiv, Kyiv, Ukraine

Dmytro S. Radchenko

8. Chemspace, Kyiv, Ukraine

Yurii S. Moroz

9. Department of Pharmacology, University of North Carolina at Chapel Hill School of Medicine, Chapel Hill, NC, USA

Xi-Ping Huang, Yongfeng Liu & Bryan L. Roth

10. National Institute of Mental Health Psychoactive Drug Screening Program (NIMH PDSP), University of North Carolina at Chapel Hill School of Medicine, Chapel Hill, NC, USA

Xi-Ping Huang, Yongfeng Liu & Bryan L. Roth

11. Division of Chemical Biology and Medicinal Chemistry, Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Bryan L. Roth

Contributions

A.A. performed cloning, mutagenesis, protein purification, SEC–MALS experiments, circular dichroism measurements, crystallization, X-ray data collection and processing, structure determination and refinement, radioligand binding, yeast complementation experiments and the sterol isomerization enzymatic assay. J.L. conducted the docking, chemoinformatics analyses, docking energy analysis and ligand picking, assisted in the latter by T.A.T. and B.K.S. J.M.B. conducted and analysed the mouse allodynia experiments with assistance from V.C., as well as the receptor expression experiments, which were supervised and co-analysed by A.I.B. M.J.O. conducted the Bayesian analysis of docking scores versus hit rates. C.M.W. tested molecules for activity against the μ OR. X.-P.H. and Y.L. tested compounds against the GPCRome and other off-targets, with supervision from B.L.R. Y.S.M. supervised the synthesis of molecules from the virtual library. D.S.R. participated in the creation of Enamine REAL library. J.J.I. was responsible for the building of the version of the ZINC library that was docked. A.C.K., B.K.S. and A.I.B. supervised the project. The manuscript was written by A.A., J.L., B.K.S. and A.C.K. with input from other authors.

Corresponding authors

Correspondence to [Allan I. Basbaum](#), [Brian K. Shoichet](#) or [Andrew C. Kruse](#).

Ethics declarations

Competing interests

A.C.K. is a founder and consultant for biotechnology companies Tectonic Therapeutic and Seismic Therapeutic, as well as the Institute for Protein Innovation, a non-profit research institute. B.K.S. is a founder of Epiodyne, a company active in analgesia, and of Blue Dolphin, which undertakes fee-for-service ligand-discovery.

Additional information

Peer review information *Nature* thanks Kevin Jones and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer review reports are available.

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 Characterization of the \$\sigma_2\$ receptor.](#)

a, SEC–MALS of the human σ_2 receptor. The σ_2 receptor was run either without ligand or with 1 μM of the indicated ligand. Lines indicate calculated total mass (grey), detergent micelle (purple), and protein (blue). **b**, Sequence alignment between the human and bovine σ_2 protein sequences performed using T-coffee⁷⁶. Residues that line the binding pocket are marked in red. **c**, Circular dichroism analysis of the bovine σ_2 receptor alone (black) or with the indicated ligand. Data is representative of multiple experiments. **d**, Circular dichroism melting curves of the bovine σ_2

receptor. Temperature was raised from 20 °C to 90 °C and molar ellipticity was measured at 222 nm. Protein was incubated either with or without indicated ligand at 12 μM. Melting temperatures for each measurement are indicated with a circle. Data is representative of multiple experiments **e**, SEC of the bovine σ_2 receptor. Blue trace is after proteolytic tag removal. Red trace is protein applied on size exclusion after reapplying the tag-free protein on affinity resin to remove proteins with intact tags. The trace presented is representative of multiple purifications. **f**, Analysis of receptor purity after the second SEC using SDS-PAGE. Grey rectangle in **e** represents fractions chosen for analysis. The SDS-PAGE presented here is representative of multiple purifications. See Source Data for uncropped version. **g**, Crystals of bovine σ_2 receptor in the lipidic cubic phase. **h**, Aspartate 56 (D56) is important for receptor structure but not for ligand binding. A tight network of hydrogen bonds that bridges extracellular loop 1 to TM4 is depicted with black dashed lines. **i**, Electron density maps for the various ligands. Polder maps⁷⁷ were calculated in PHENIX. Maps are contoured at a level of 3 σ . **j**, View of cholesterol-binding pose, showing contacts with other binding pocket residues. Hydrogen bonds are marked with black dashed lines. **k**, Yeast complementation assay. A Δ ERG2 yeast strain was transformed with plasmids harbouring the indicated genes. Yeast cells were grown to logarithmic phase and diluted to OD600 of 0.1, and then further diluted in a five-fold serial dilution series. Two microliters of each dilution were spotted on plates. Yeast cells were grown either in permissive conditions of no cycloheximide or in the restrictive conditions of 50 ng/ml cycloheximide, which requires functional Δ8-9 sterol isomerase activity for viability. ERG2 and EBP can act as sterol isomerases and rescue the growth of Δ ERG2 *S. cerevisiae* whereas the σ_2 receptor, the σ_1 receptor, or any other member of the EXPERA family cannot. **l**, EBP can catalyse the conversion of zymostenol to lathosterol whereas σ_2 cannot. Standards are in dark grey. EBP converts zymostenol to lathosterol (apricot) but does not convert lathosterol to zymostenol (dark red). The σ_2 receptor does not convert lathosterol to zymostenol (dark blue) or zymostenol to lathosterol (light purple). Structures of zymostenol and lathosterol are depicted below the traces

[Source data](#).

Extended Data Fig. 2 Comparisons of the distribution of docking scores.

a–d, The distribution of docking scores of tested molecules for hit rate curves against σ_2 (left column) and D_4 (right column) receptors. All tested molecules are grouped based on docking score bins. The distributions are shown in box plots for **a**, net electrostatic energy, **b**, ligand desolvation energy, **c**, van der Waals (vdW) energy and **d**, delta ligand desolvation energy after recalculating atomic desolvation energy based on the docked pose. **e–h**, Comparison of hit rates and affinities achieved by combined docking score and human inspection and these achieved by docking score alone. **e**, Overall hit rates for selecting compounds from the first 3 scoring bins by each strategy: human prioritization and docking score (orange), or docking score alone (blue). Hit rate is the ratio of active compounds/tested compounds; the raw numbers appear at the top of each bar. **f**, Hit rates for selecting compounds at different scoring ranges by each strategy: human prioritization and docking score (orange) or docking score alone (blue). **g**, Distribution of the binding affinity level among the hits from **e** (top panel). We measured competition binding curves for 14 docking hits from human prioritization and docking score, and 7 hits from the docking score alone. These are divided into three affinity ranges: <5 nM; 5 nM–50 nM; >50 nM; Distribution of the binding affinity level among the hits from all different scoring ranges (bottom panel). We measured competition binding curves for 14 docking hits from human prioritization and docking score, and 17 hits from the docking score alone. **h**, Hit-rate curve comparison with/without human picks. The hit rate without human picks at the top plateau is 39% and at the bottom plateau is 0%, and the docking score ($dock_{50}$) and slope at the maximum ($slope_{50}$) are -46.5 kcal mol⁻¹ and -3.5% per kcal mol⁻¹, respectively

[Source data](#).

Extended Data Fig. 3 Analogues of σ_2 receptor ligands and the effect of a structural water molecule.

a–c, Initial hits and selected analogues of σ_2 receptor ligands. Competition binding curves on the top panel, 2D drawings of compounds are on the bottom panel. Parent compound is indicated by grey background. Points shown as mean \pm s.e.m. from three technical replicates. **a**, Parent compound ZINC548355486 and its three potent analogues. **b**, Parent compound ZINC895657866 and its three potent analogues. **c**, Parent compound ZINC450573233 and its three potent analogues. **d–f**, The binding site of the σ_2 receptor contains a structural water. **d**, Water coordination at the binding site of the σ_2 receptor. Water molecule is depicted as a red sphere. Hydrogen bonds are indicated by black dashed lines. **e**, Saturation binding curve to measure the dissociation constant (K_d) of [3 H]DTG for the various mutants of σ_2 receptor meant to disrupt water coordination. Residues proximal to the structural water were chosen for mutation. Residues were mutated to the indicated amino acid. Points shown as mean \pm s.e.m. from three technical replicates. **f**, Competition binding measurement of affinity of Z1241145220 in various mutants of σ_2 . Points shown as mean \pm s.e.m. from three technical replicates

[Source data.](#)

Extended Data Fig. 4 Effect of systemic σ receptor ligands on motor behaviour.

a, Response of mice to a von Frey filament after spared nerve injury (SNI). All five ligands are compared to their respective vehicles (PD-144418 10 mg/kg ($n = 5$) and 30 mg/kg ($n = 5$) vs. kolliphor ($n = 5$), one-way ANOVA, $F(2, 12) = 7.49, p = 0.008$; Z4446724338 10 mg/kg ($n = 10$) and 20 mg/kg ($n = 5$) vs cyclodextrin ($n = 10$), one-way ANOVA, $F(2, 22) = 25.12, p < 0.001$; Z4857158944 10 mg/kg ($n = 5$) and 20 mg/kg ($n = 5$) vs cyclodextrin ($n = 10$), one-way ANOVA, $F(2, 17) = 5.10, p = 0.02$; Z1665845742 10 mg/kg ($n = 10$) and 20 mg/kg ($n = 5$) and PB28 30 mg/kg ($n = 10$) vs saline ($n = 10$), one-way ANOVA, $F(3, 31) = 6.18, p = 0.002$; asterisks define individual group differences to respective vehicle control using Dunnett's multiple comparisons Post-hoc test; ns = not significant, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). Data shown are mean \pm s.e.m. Data for higher doses and vehicles is replotted from Fig. 4. **b**, No sedation or

motor impairment on the rotarod was observed after drug treatments compared to vehicle at 1 h (Z1665845742 10 mg/kg ($n = 5$) and Z4857158944 20 mg/kg ($n = 5$) vs saline ($n = 5$), one-way ANOVA, $F(2, 12) = 1.04, p = 0.38$; Z4446724338 10 mg/kg ($n = 5$) vs kolliphor ($n = 5$), unpaired two-tailed Student's t -test, $t(8) = 0.47, p = 0.65$) or 24 h post-injection (Z1665845742 10 mg/kg ($n = 5$) and Z4857158944 20 mg/kg ($n = 5$) vs saline ($n = 5$), one-way ANOVA, $F(2, 12) = 0.45, p = 0.65$; Z4446724338 10 mg/kg ($n = 5$) vs kolliphor ($n = 5$), unpaired two-tailed Student's t -test, $t(8) = 0.72, p = 0.49$); ns = not significant. Data shown are means \pm s.e.m. **c**, Response of SNI mice to a von Frey filament after repeated injections of Z4446724338 10 mg/kg ($n = 5$). Mechanical thresholds were assessed 1 h and 24 h after four separate injections. Data shown are means \pm s.e.m. normalized to each mouse's SNI baseline. **d**, Response of SNI mice to a von Frey filament after repeated injections of Z4857158944 10 mg/kg ($n = 5$). Mechanical thresholds were assessed 1 h and 24 h after four separate injections. Data shown are means \pm s.e.m. normalized to each mouse's SNI baseline. **e**, Quantification of the expression levels of *Sigmar1* (σ_1) and *Tmem97* (σ_2) in wildtype (WT) and SNI mice detected by *in situ* hybridization ($n = 3$ mice per group). Representative images can be found in panel **f**. Data shown are mean \pm s.e.m.; unpaired two-tailed Student's t -test—*Sigmar1*: $t(4) = 0.5, p = 0.64$; *Tmem97*: $t(4) = 1.0, p = 0.37$; ns = not significant. AU = arbitrary units. **f**, *in situ* hybridization of mouse dorsal root ganglion (DRG) sections for *Sigmar1* (σ_1) and *Tmem97* (σ_2) genes illustrates expression in myelinated (Nefh-positive; blue) and unmyelinated (Acpp-positive; red) subsets of sensory neurons and no change after SNI.

Source data.

Extended Data Fig. 5 Off-target profiling of Z4446724338, Z1665845742 and Z4857158944.

a–c, TANGO screens against a panel of 320 GPCRs for the indicated σ_2 ligand. **a**, Z4446724338, **b**, Z1665845742, **c**, Z4857158944. **d**, GloSensor μ OR-mediated cAMP inhibition (G_i activation) by DAMGO, Z4446724338, Z1665845742, and Z4857158944. **e, f**, Follow-up does-

response curves for pain-related receptors that showed activation in **a–c**. **e**, Z4446724338 and Z1665845742 against 5HT1A. **f**, Z4857158944 against κOR. Data shown are means ± s.e.m

[Source data.](#)

Extended Data Fig. 6 Paw withdrawal thresholds.

a, Paw withdrawal thresholds (PWT) before (blue bar) and after (red bar) SNI, as well as after SNI + treatment (purple bar). For easier visualization of individual data points, data was also plotted without the pre-SNI baseline. Data are the same as in Fig. [4b](#) and Extended Data Fig. [4a](#), but without the normalization to the individual post-SNI baselines and are expressed as mean ± s.e.m.; mice per group: saline ($n = 10$); cyclodextrin ($n = 10$); kolliphor ($n = 5$); PB28 30 mg/kg ($n = 10$); PD-144418 10 mg/kg ($n = 5$) and 30 mg/kg ($n = 5$); Z4446724338 10 mg/kg ($n = 10$) and 20 mg/kg ($n = 5$); Z1665845742 10 mg/kg ($n = 5$) and 20 mg/kg ($n = 5$); Z4857158944 10 mg/kg ($n = 5$) and 20 mg/kg ($n = 5$); unpaired two-tailed Student's t-test. **b**, PWTs 1 h, 24 h, and 48 h after saline or drug treatment. Data are the same as in Fig. [4c](#), but without the normalization to the individual post-SNI baselines, and are expressed as mean ± s.e.m. Significance levels determined using Dunnett's multiple comparisons Post-hoc test reflect the difference between Z4446724338 and saline for simplicity (two-way ANOVA; time x treatment interaction: $F(8, 80) = 2.4$, $p = 0.02$; time: $F(2, 74) = 5.2$, $p = 0.009$; treatment: $F(4, 40) = 3.3$, $p = 0.02$; four treatment groups ($n = 10$) except PD-144418 ($n = 5$); ns = not significant. **c**, Response of SNI mice to a von Frey filament after repeated injections of Z4446724338 10 mg/kg ($n = 5$). Mechanical thresholds were assessed 1 h and 24 h after four separate injections. Data shown are paw withdrawal thresholds in grams, expressed as mean ± s.e.m. **d**, Response of SNI mice to a von Frey filament after repeated injections of Z4857158944 10 mg/kg ($n = 5$). Mechanical thresholds were assessed 1 h and 24 h after four separate injections. Data shown are paw withdrawal thresholds in grams, expressed as mean ± s.e.m

[Source data.](#)

Extended Data Table 1 Data collection and refinement statistics

Extended Data Table 2 14 of the highest-affinity direct docking hits for the σ_2 receptor

Extended Data Table 3 Measured pharmacokinetic parameters for PB28, Z1665845742, Z4446724338 and Z4857158944 in male CD-1 mice by 10 mg kg⁻¹ subcutaneous administration

Supplementary information

Supplementary Information

This file contains Supplementary Tables 2–5, Supplementary Figs. 1–8 and details regarding the Synthetic Procedures and Chemical Characterization Data.

Reporting Summary

Supplementary Table 1

Supplementary Table listing all the compounds tested in the manuscript.

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. 6**](#)

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Alon, A., Lyu, J., Braz, J.M. *et al.* Structures of the σ_2 receptor enable docking for bioactive ligand discovery. *Nature* **600**, 759–764 (2021).
<https://doi.org/10.1038/s41586-021-04175-x>

- Received: 20 April 2021
- Accepted: 19 October 2021
- Published: 08 December 2021
- Issue Date: 23 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04175-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/s41586-021-04175-x>

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

Career Guide