

nature

AI-GUIDED INTUITION

Machine learning helps inspire mathematicians to derive fresh conjectures

A taste of the future
The hunt for a healthy and sustainable diet the whole planet can eat

US science policy
Former NIH heads call for department of science and technology

Coronavirus
The hidden role of cryptic transmission in how COVID-19 spread

GOALS
GOALS
GOALS



Nature.2021.12.04

[Sat, 04 Dec 2021]

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

This Week

- **[‘BRICS’ nations are collaborating on science but need a bigger global platform](#)** [01 December 2021]
Editorial • A policy paper published 20 years ago led to expanded research collaboration between emerging economic powers. But its main recommendation was ignored.
- **[Industry scores higher than academia for job satisfaction](#)** [01 December 2021]
Editorial • Nature’s salary survey finds that industry researchers are more positive about their careers. Academia must raise its game.
- **[Victories against AIDS have lessons for COVID-19](#)** [29 November 2021]
World View • Anthony Fauci on four decades of progress against HIV, and what’s needed for the future.
- **[The 3D print job that keeps quake damage at bay](#)** [24 November 2021]
Research Highlight • An easily produced seismic isolator designed to protect buildings from earthquakes mimics the bones of human limbs.
- **[How jellyfish control their lives](#)** [24 November 2021]
Research Highlight • Transgenic jellyfish reveal how these fragile creatures get by without a brain.
- **[Hard times tear coupled seabirds apart](#)** [24 November 2021]
Research Highlight • Even birds get divorced, and, as in humans, resource constraints can drive more couples to the brink.
- **[Tidings from an exploding star make astronomers happy](#)** [24 November 2021]
Research Highlight • Herculean stellar event provides scientific clues to how stars die.
- **[Neutron beam sheds light on medieval faith and superstition](#)** [24 November 2021]
Research Highlight • Hidden inside a leaden amulet, researchers find words of magic and Christian creed side by side.
- **[Mini-machine can chop and channel proteins](#)** [23 November 2021]

Research Highlight • Scientists' creation can cut proteins into pieces and thread them through a molecular tunnel.

- **Earth is headed for well over two degrees of warming** [23 November 2021]

Research Highlight • Modellers look at how climate policies might change with time — and find a wide range of possible outcomes, none of them good.

- EDITORIAL
- 01 December 2021

‘BRICS’ nations are collaborating on science but need a bigger global platform

A policy paper published 20 years ago led to expanded research collaboration between emerging economic powers. But its main recommendation was ignored.



The BRICS leaders' summit in Johannesburg in 2018. This year, there have been more than 100 events, many involving researchers. Credit: Alexey Nikolsky/Sputnik/EPA-EFE/Shutterstock

It's 20 years since economist Jim O'Neil coined the term BRIC to describe Brazil, Russia, India and China. O'Neill, who was head of economics research at the investment bank Goldman Sachs, wrote in an internal policy paper that the four countries were growing faster than the G7 group of large economies (see go.nature.com/3pgtqsd). World policymaking clubs such as the G7 are usually dominated by the United States and Europe. They should invite BRIC representatives, O'Neill advised. The balance of world economic power was tilting and the big institutions of global governance needed to reflect that, he argued.

In one sense, O'Neill's forecast was not new. Since at least the 1950s, economists, notably including Walt Whitman Rostow at the Massachusetts Institute of Technology in Cambridge, had been anticipating that China and India would become economic powers — as they were in pre-industrial times. But what happened next was unprecedented. As a result of O'Neill's analysis, the leaders of the four countries created an intergovernmental organization designed to strengthen the bonds between the nations, with a key focus on boosting research cooperation. They named it BRICS (South Africa joined in 2010).

This year has been a busy one in the BRICS research calendar. India, which holds the rotating chair, organized more than 100 events. These included meetings of astronomers and energy and health researchers, as well as of medical researchers and practitioners.

The year also saw meetings between BRICS agriculture, health and space officials. Last week, the BRICS science ministers ended talks in New Delhi with a plan to connect younger innovators and start-up companies across the five countries, and to set up a centre to facilitate technology transfer. This is in addition to 13 BRICS science working groups that collaborate on fields including polar and ocean technology, astronomy, climate and energy, photonics and biotechnology. There is also a BRICS-wide network of more than 50 universities.



Economist Jim O'Neill conceived of the BRIC nations. But his recommendation that they be invited to top policymaking forums has not been heeded. Credit: Simon Dawson/Bloomberg/Getty

These collaborations are all about creating knowledge and innovation in fields in which an individual country might struggle to act alone, sharing data and recognizing that each country brings different strengths. The astronomy working group, for example, is researching the feasibility of having a network of ground-based telescopes in each country. BRICS space agencies are planning to share remote-sensing data on climate and natural disasters. A group working on climate and energy brings Brazil's expertise on Amazonian weather and climate together with China's researchers in photovoltaic energy systems. Meeting the United Nations Sustainable Development Goals is a running theme in BRICS science cooperation.

However, there's one noticeable gap. The BRICS countries account for almost half of global tuberculosis (TB) cases, and the world's largest number of cases of multi-drug-resistant TB. Intra-BRICS TB research would be an obvious priority, except that it isn't. There is a BRICS TB research network, and in August it announced a pioneering genomic-surveillance network to

study the intersection of TB and COVID-19. But this is an area where funding and collaboration could be at much higher levels.

Conflicts of interest

There is a larger omission that must also be addressed. O'Neill's 2001 call to include BRICS nations at the top policymaking tables went largely unheeded. For a short period, it seemed that the G20 group of the world's largest economies (which includes the BRICS nations) would work together on global challenges, such as ending the pandemic. But this has not happened. And it has not gone unnoticed.



[China creates vast research infrastructure to support ambitious climate goals](#)

In an article for this year's BRICS meetings, researchers Sachin Chaturvedi and Sabyasachi Saha at the think tank Research and Information System for Developing Countries in New Delhi, argue that organizations such as the UN have been unable to defend the interests of the poorest nations and those with the least power. The failure to provide vaccines quickly for LMICs is a case in point. It's no coincidence that India and South Africa are leading an alliance of more than 100 nations asking for intellectual-property protections to be waived during the pandemic so that drugs and vaccines can reach LMICs more quickly.

Similarly, China and India were united in warning November's COP26 climate meeting in Glasgow, UK, that future commitments to decarbonize cannot be made at the expense of denying fossil fuels to communities with no access to electricity. Clearly, LMICs are losing trust in the UN-led system of global governance because, as they see things, higher-income countries are shooting down their ideas, or ignoring their perspectives. And that should concern us all.

It's essential that LMICs work together towards shared goals and collaborate to build their research infrastructure. But at the same time, it is crucial that existing global-cooperation forums — such as the G7, the World Trade Organization and the World Health Organization — regard the BRICS nations and all other LMICs as equals.

If there were a prize for policy analysis that has had world-changing influence, O'Neill's article would be the standout candidate. But the true test of its success is whether the existing powers will heed O'Neill's advice. For global governance to be credible, those who control the levers of power must learn to see LMICs as partners, not as aid recipients.

Nature **600**, 7 (2021)

doi: <https://doi.org/10.1038/d41586-021-03568-2>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03568-2>

- EDITORIAL
- 01 December 2021

Industry scores higher than academia for job satisfaction

Nature's salary survey finds that industry researchers are more positive about their careers. Academia must raise its game.



In many countries, around half of all researchers work for private companies, which often offer better pay and work–life balance than academia. Credit: Ajpek Orsi/Getty

How does being a researcher in industry compare with being an academic? That's a question explored in a series of articles pegged to *Nature's* latest survey of salaries and job satisfaction, which concludes this week. The

results make for sobering reading for academics, revealing a shift towards industry (see [*Nature* 599, 519–521; 2021](#)).

Scientists who work in industry are more satisfied and better paid than are colleagues in academia, according to the self-selected group of respondents, which comprised more than 3,200 working scientists, mostly from high-income countries. Two-thirds of respondents (65%) are in academia; 15% work in industry. Industry employs, on average, half of the researchers in these countries.



Discrimination still plagues science

Another key finding, covered in [this week's piece on workplace diversity](#), is that 30% of respondents in academia reported workplace discrimination, harassment or bullying, compared with 15% of those in industry. Industry respondents (64%) are also much more likely than those in academia (42%) to report feeling positively about their careers. That's a marked shift from the 2016 survey, in which satisfaction levels across the two sectors were neck and neck (63% and 65%, respectively). One research project manager working in the private sector in the United States summed up the latest findings with the words: "I am now an evangelist for all of my friends still in academia to get out and join biotech or any other professional industry."

The latest findings should sound alarm bells for academic employers at a time when morale in the sector is worsening in many countries. As *Nature* went to press, academics at 58 UK universities were set to stage a 3-day strike from 1 December as part of an ongoing dispute about pay, working conditions and planned cuts to their pensions.

A separate survey of more than 1,000 UK faculty and staff members carried out between June and August last year revealed a sense that university leaders are using the pandemic as an excuse to [push through cost-cutting measures \(R. Watermeyer et al. Br. J. Sociol. Educ. 42, 651–666; 2021\)](#). Seven out of ten respondents say that this has created a culture of fear, which has, in turn, led to university leadership becoming more autocratic. Many respondents were also concerned that publicly funded academic institutions are increasingly being run as businesses.



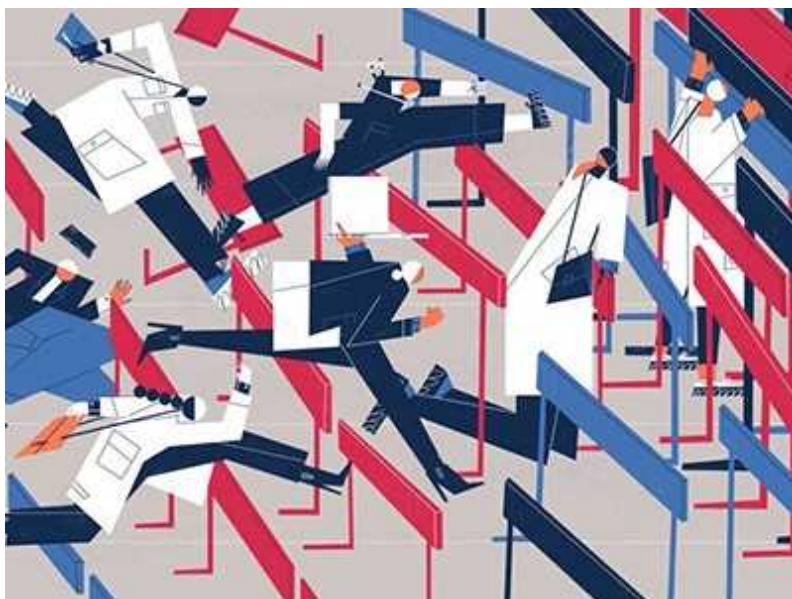
[How burnout and imposter syndrome blight scientific careers](#)

But if universities truly were like businesses, the survey findings suggest, staff would probably be happier — and would not be looking to leave in the numbers they seem to be.

Nature conducted a small number of interviews with research leaders in industry. One research director at a global bioscience company said that about 60% of job applications come from people in academia. A small

proportion of those hired do later return to academia. The director says the company is keen to keep a path for return open, for example by permitting research staff to publish in scientific journals, which is not always an option for researchers in industry.

Labour-market economics offers one explanation for the better pay and greater satisfaction reported by respondents in industry. New companies are popping up every week, and they can struggle to fill vacancies, so will offer higher salaries and additional benefits to attract good candidates. In a number of high-income countries, this is essentially the reverse of the situation in academia, in which postdoctoral researchers greatly outnumber tenure-track positions ([S. C. McConnell et al. *eLife* 7, e40189; 2018](#)). In these countries, industry contributes around two-thirds of all research and development (R&D) funding. All in all, public institutions (including universities) have less money to spend and more researchers chasing every job.



[Stagnating salaries present hurdles to career satisfaction](#)

That said, industry is not at all homogeneous. It ranges from multinational technology and life-sciences companies employing tens of thousands of people to one-person start-ups spun off from universities. And corporate life comes with its own challenges. One head of R&D with experience of working at large pharmaceutical companies said corporate politics and the

slow pace of executive decision-making can be frustrating for a researcher who is used to a more hands-on scientific role in an academic lab. By contrast, junior colleagues at smaller companies can enjoy more varied roles in agile working environments before being head-hunted by competitors offering higher salaries.

Clearly, academic salaries are unlikely to be able to compete with those of industry — at least while there are so many more postdocs than positions available. But these two employment destinations need to learn from each other. Researchers looking to switch from academia to industry often have to contend with a supervisor's disapproval of the move. Such attitudes can discourage researchers from returning to academia, where the perspectives gained in industry could help to re-energize and diversify teams.

Of course, many do find a career in academia hugely rewarding. One academic bioinformatician in the United States who responded to the survey reported earning around 50% of what she was offered for industry positions. She said: "It would be nice if academia could be more competitive with industry, but I love what I do and where I live so I can't really complain."

But as competition for scientific talent increases, academic leaders must not assume that that sentiment will prevail. If they want current and future generations of academics to thrive, they need to learn how other sectors succeed in recruiting, retaining and rewarding staff, and consider how to ensure that they are still attractive to the top talent.

Nature **600**, 8 (2021)

doi: <https://doi.org/10.1038/d41586-021-03567-3>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03567-3>

- WORLD VIEW
- 29 November 2021

Victories against AIDS have lessons for COVID-19



Anthony Fauci on four decades of progress against HIV, and what's needed for the future.

- [Anthony Fauci](#) 9

Amid the COVID-19 pandemic, a milestone of another global scourge has largely escaped the spotlight. It has been 40 years since the earliest reports of what ultimately became known as AIDS, in 1981. I initially dismissed the first report as a curiosity and probably a fluke, but another a month later, from the US Centers for Disease Control and Prevention, changed my mind and the direction of my career. Against the advice of my mentors, I shifted my research focus to understanding why young, healthy men were being beset by unusual conditions. I remember anxiously awaiting results

essential in the fight against a disease that brought on so many seemingly unrelated symptoms: pneumonia, blindness, skin lesions, dementia.

This year's World AIDS Day, on 1 December, marks both incredible progress and the need for more. HIV was shown to be the cause of AIDS in 1983–84. Soon, there were blood tests for diagnosis, and for screening the supply of donated blood. These revealed the enormous scope of the pandemic: between 1984 and 1985, new cases of AIDS in the United States almost doubled. Eventually, progress in understanding — getting the sequence of the virus's genome, and comprehending how it decimated the immune system and how disease progressed — paved the way for dozens of approved therapies.

In 1985, a 25-year-old in the United States diagnosed with AIDS had a life expectancy of less than two years. Today, a person with HIV can expect to die in old age, of other causes. Since the global peak in 2004, AIDS-related deaths are down by 64%. Overall, about 73% of people with HIV are accessing treatment. Still, in 2020, almost 700,000 people died from AIDS-related illnesses and 38 million were living with HIV.



[Achilles heel spotted for promising HIV-prevention drug](#)

Millions of people have also been killed and disabled in the COVID-19 pandemic. The fight against HIV reveals how important it is to make use of

existing treatments and strategies for prevention, to strive for better ones, to reach vulnerable communities and to consider equity, education and outreach. There are many similarities that sadden me: how hard some patient groups need to fight for recognition; how misinformation and denialism can promote illness and death; how slowly interventions reach vulnerable populations; how easy it can be to overlook patient inputs when setting research priorities.

The 40 years of AIDS have been punctuated by periods of both darkness and hope. In the early, frustrating days of the pandemic, physicians had little to offer our patients. Glimmers of hope came with treatments for opportunistic infections and the first partially effective antiretroviral drugs. In the mid-1990s, multidrug antiretroviral regimens transformed the prognoses of many people infected with HIV from dire to excellent. Now, more potent, better-tolerated drug combinations are available as a single daily pill.

A crucial lesson was that antiretroviral therapy both helps the person living with HIV and reduces their risk of transmitting it: treatment cuts the level of the virus in their blood until it can neither be detected nor passed on. Understanding that ‘undetectable equals untransmittable’ was a triumph of research, as well as of HIV activism, which has been a driving influence throughout this journey. Also transformative is pre-exposure prophylaxis: giving people antiretrovirals before they are exposed to HIV can be more than 95% effective in preventing them from acquiring it. I suspect that drugs based on monoclonal antibodies will soon play an important part in treating and preventing HIV, as they do for SARS-CoV-2.

Medical advances mean that HIV/AIDS could, theoretically, now be taken off the table as a major health threat (in the United States and globally), particularly if programmes such as the President’s Emergency Plan for AIDS Relief and the Global Fund to Fight AIDS, Tuberculosis and Malaria continue to get robust support. Sadly, things are never that simple. Uptake of treatment remains suboptimal, in rich and poor communities. Reasons include lack of transport, substandard housing, mental-health issues, substance abuse, pill fatigue, drug toxicity, stigma and discrimination.

And so, a top priority is to ensure that everyone has access to existing treatment and prevention options, while we develop better ones. Earlier this year the first long-acting injectable antiretroviral treatment regimen, cabotegravir + rilpivirine, delivered once a month, was approved by regulatory authorities. An injectable, long-acting form of cabotegravir might soon be approved for prophylactic use.

Exciting work is under way for further HIV treatment and prevention. This includes development of very long-acting drugs that could be taken every six months or even less frequently. And then there are strategies for achieving durable HIV remission. Some people are working on eradicating the replication-competent HIV reservoir (a concept frequently referred to as a cure).

A safe and effective HIV vaccine has been elusive so far, but even a moderately effective vaccine could, together with the rest of the toolkit, bring an end to AIDS as a major health concern. It's likely that one could build on some of the tools and platforms used for COVID-19 vaccines, including messenger RNA and optimizing immunogens to prompt the most effective immune response.

Entering the fifth decade of AIDS, the challenge to researchers is to work with at-risk communities to deliver treatment and prevention to everyone in the world who needs them.

Nature **600**, 9 (2021)

doi: <https://doi.org/10.1038/d41586-021-03569-1>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03569-1>

- RESEARCH HIGHLIGHT
- 24 November 2021

The 3D print job that keeps quake damage at bay

An easily produced seismic isolator designed to protect buildings from earthquakes mimics the bones of human limbs.

 Motion animation showing the telescopic tendons in action.

Motion animation showing the telescopic tendons in action. Credit: F. Fraternali *et al./Nonlinear Dyn.* ([CC BY 4.0](#))

A device inspired by human arms and legs could help to reduce how much a building shakes during an earthquake.

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**, 10 (2021)

doi: <https://doi.org/10.1038/d41586-021-03506-2>

References

1. 1.

Fraternali, F., Singh, N., Amendola, A., Benzoni, G. & Milton, G. W. *Nonlinear Dyn.* <https://doi.org/10.1007/s11071-021-06980-5> (2021).

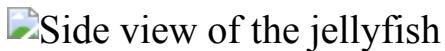
This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03506-2>

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

- RESEARCH HIGHLIGHT
- 24 November 2021
- Correction [30 November 2021](#)

How jellyfish control their lives

Transgenic jellyfish reveal how these fragile creatures get by without a brain.



Distribution of neurons in the jellyfish *Clytia hemisphaerica*. Credit: B. Weissbourd *et al./Cell*

In jellyfish, a complex network of neurons throughout the animals' bodies replaces the need for a centralized brain.

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**, 10 (2021)

doi: <https://doi.org/10.1038/d41586-021-03510-6>

Updates & Corrections

- **Correction 30 November 2021:** An earlier version of this article misstated the role of the neural network under the jellyfish's round 'umbrella'.

References

1. 1.

Weissbourd, B. *et al.* *Cell* **184**, 5854–5868 (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03510-6>

- RESEARCH HIGHLIGHT
- 24 November 2021

Hard times tear coupled seabirds apart

Even birds get divorced, and, as in humans, resource constraints can drive more couples to the brink.



Climate change spells bad news for black-browed albatross couples, with ‘divorce’ more common when sea-surface temperatures are warmer than usual. Credit: Doug Gimesy/Nature Picture Library

Many seabirds form long-term pairings, but do not necessarily mate for life — and are more likely to ‘break up’ in years when environmental conditions are unfavourable, researchers reveal.

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**, 10 (2021)

doi: <https://doi.org/10.1038/d41586-021-03509-z>

References

1. 1.

Ventura, F., Granadeiro, J. P., Lukacs, P. M., Kuepfer, A. & Catry, P.
Proc. R. Soc. B <https://doi.org/10.1098/rspb.2021.2112> (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03509-z>

- RESEARCH HIGHLIGHT
- 24 November 2021

Tidings from an exploding star make astronomers happy

Herculean stellar event provides scientific clues to how stars die.



The NASA Infrared Telescope Facility atop Hawaii's Maunakea was used to gather data on the nova V1674 Herculis. Credit: Ed Darack/RGB Ventures/SuperStock/Alamy

Observations of a record-breaking stellar explosion are shedding light on how some of the most violent astronomical events come to pass.

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**, 10 (2021)

doi: <https://doi.org/10.1038/d41586-021-03508-0>

References

1. 1.

Woodward, C. E. *et al. Astrophys. J.* **922**, L10 (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03508-0>

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

- RESEARCH HIGHLIGHT
- 24 November 2021

Neutron beam sheds light on medieval faith and superstition

Hidden inside a leaden amulet, researchers find words of magic and Christian creed side by side.



The Bispegata amulet contains religious phrases and magic words. Credit: Kirsten Jensen Helgeland

A Norwegian amulet dating back more than 700 years has been hiding a runic inscription that holds religious and magic significance.

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**, 11 (2021)

doi: <https://doi.org/10.1038/d41586-021-03505-3>

References

1. 1.

Wilster-Hansen, B., Mannes, D., Holmqvist, K. L., Ødeby, K. & Kutzke, H. *Archaeometry* <https://doi.org/10.1111/arcm.12734> (2021).

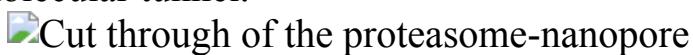
This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03505-3>

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

- RESEARCH HIGHLIGHT
- 23 November 2021

Mini-machine can chop and channel proteins

Scientists' creation can cut proteins into pieces and thread them through a molecular tunnel.



A miniature device for manipulating proteins is built from a tunnel-shaped bacterial protein (cyan) and a 'chopper' (orange and blue). Credit: Cesar Telles de Souza

A molecular machine that threads proteins through a membrane could one day help scientists to identify a protein from a single molecule.

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**, 11 (2021)

doi: <https://doi.org/10.1038/d41586-021-03502-6>

References

1. 1.

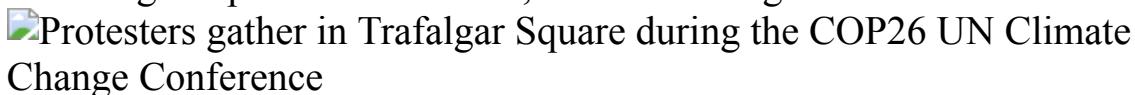
Zhang, S. *et al.* *Nature Chem.* <https://doi.org/10.1038/s41557-021-00824-w> (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03502-6>

- RESEARCH HIGHLIGHT
- 23 November 2021

Earth is headed for well over two degrees of warming

Modellers look at how climate policies might change with time — and find a wide range of possible outcomes, none of them good.



Stronger pledges will be needed from policymakers if dangerous levels of warming are to be prevented. Credit: Niklas Halle'n/AFP/Getty

Humanity is not on track to keep global warming below 2°C, the target of the 2015 Paris climate agreement, according to a study that looks at how climate mitigation efforts might unfold between 2030 and 2050.

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**, 11 (2021)

doi: <https://doi.org/10.1038/d41586-021-03507-1>

References

1. 1.

Sognnaes, I. *et al.* *Nature Clim. Change*
<https://doi.org/10.1038/s41558-021-01206-3> (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03507-1>

News in Focus

- **[Pandemic mental health and Eurasia's oldest jewellery](#)** [01 December 2021]
News Round-Up • The latest science news, in brief.
- **[Cuba's bet on home-grown COVID vaccines is paying off](#)** [22 November 2021]
News • Preprint data show that a three-dose combo of Soberana jabs has 92.4% efficacy in clinical trials.
- **[NASA spacecraft will slam into asteroid in first planetary-defence test](#)** [19 November 2021]
News • The DART mission has launched. Next up, it will try out a manoeuvre that could one day deflect killer asteroids from Earth.
- **[Outcry as men win outsize share of Australian medical-research funding](#)** [26 November 2021]
News • An analysis showing that the National Health and Medical Research Council awards fewer 'investigator grants' to women prompts thousands to sign a petition calling for gender quotas.
- **[China creates vast research infrastructure to support ambitious climate goals](#)** [22 November 2021]
News • Carbon-neutrality institutes, and other initiatives to support a pledge to achieve net zero by 2060, are popping up like mushrooms across China.
- **[Scientists question Max Planck Society's treatment of women leaders](#)** [19 November 2021]
News • An open letter signed by 145 female scientists from across the world expresses concern over recent dismissals of women leaders.
- **[Heavily mutated Omicron variant puts scientists on alert](#)** [25 November 2021]
News • Researchers are racing to determine whether a fast-spreading coronavirus variant poses a threat to COVID vaccines' effectiveness.
- **[What humanity should eat to stay healthy and save the planet](#)** [01 December 2021]

News Feature • What we eat needs to be nutritious and sustainable. Researchers are trying to figure out what that looks like around the world.

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

- NEWS ROUND-UP
- 01 December 2021

Pandemic mental health and Eurasia's oldest jewellery

The latest science news, in brief.



The pendant, seen here from two different angles, features drilled holes and around 50 smaller indentations that create an irregular curve. Credit: Antonino Vazzana/BONES Lab

Mammoth pendant could be Eurasia's oldest jewellery

A 41,500-year-old pendant carved from a piece of a woolly-mammoth tusk could be the [oldest known example of decorated jewellery in Eurasia made by humans](#), according to archaeologists.

The oval-shaped pendant has two drilled holes and is decorated with at least 50 smaller puncture marks that create a looping curve. It was found in the Stajnia Cave, a natural rock shelter in southern Poland. The results of radiocarbon dating, published in *Scientific Reports* ([S. Talamo et al. Sci. Rep. 11, 22078; 2021](#)), suggest that it is 2,000 years older than similarly decorated artefacts from other sites.

But the claim that the pendant is the oldest ornate jewellery in Eurasia could prove controversial. Martin Porr, who studies rock-art archaeology at the University of Western Australia in Perth, thinks that the pendant is “highly significant and intriguing”, but says that the researchers restricted their comparisons to other artefacts that have been decorated with puncture marks, and did not include items of a similar age that could also have been used as jewellery.

“The pendant of Stajnia is the only decorated artefact of its kind [that is] directly radiocarbon dated, and the comparison with the other punctuated ornaments shows it is the oldest,” says Sahra Talamo, a chemist at the University of Bologna in Italy, who led the study.



The COVID-19 lockdowns of 2020 created public space for sunbathing but have been linked to a rise in helpline calls in some countries. Credit: Ollie Millington/Getty

Helpline calls reveal COVID's toll on mental health

A sweeping study of 8 million calls to helplines in 19 countries and regions found that call volumes jumped during the first wave of coronavirus infections. [Loneliness and concerns about the pandemic drove most of the callers](#), rather than imminent threats such as suicidal thoughts or abuse.

The analysis ([M. Brülhart *et al. Nature* https://doi.org/g6fs; 2021](https://doi.org/g6fs)) is one of the largest to address mental-health challenges during the pandemic. The authors collected helpline data from regions including the United States, China, Lebanon and Europe. Analysis showed that six weeks after the initial

wave of coronavirus infections began, the number of calls was 35% higher than before the pandemic.

Despite fears that the pandemic would cause a surge in mental-health crises including suicides and addiction, and in intimate-partner abuse, the authors did not find signs of this in their data. That does not invalidate the suffering of some individuals, or anecdotal or regional studies that might have found different results, says co-author Marius Brülhart, an economist at the University of Lausanne in Switzerland. It just means that the trend did not emerge from his team's large-scale helpline data.

Nature **600**, 13 (2021)

doi: <https://doi.org/10.1038/d41586-021-03570-8>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03570-8>

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

- NEWS
- 22 November 2021

Cuba's bet on home-grown COVID vaccines is paying off

Preprint data show that a three-dose combo of Soberana jabs has 92.4% efficacy in clinical trials.

- [Sara Reardon](#)



A health-care worker vaccinates a woman with the Abdala jab in August. Credit: Yamil Lage/AFP via Getty

When the COVID-19 pandemic began, Cuba decided not to wait on the rest of the world to develop vaccines. The United States' 60-year-old economic embargo against the country, which prevents US-made products from being exported there, would make it difficult for Cuba to acquire vaccines and therapies, researchers and officials knew. "It was best, for protecting our population, to be independent," says Vicente Vérez Bencomo, director-general of the Finlay Institute of Vaccines in Havana.



Can Cuba beat COVID with its homegrown vaccines?

So the Finlay Institute and Cuba's other state-run biotechnology centres started developing their own COVID-19 vaccines in the hope that at least one of them would be effective. Their bet seems to be paying off: in a 6 November preprint published on medRxiv¹, Vérez Bencomo and his colleagues report that one of the institute's vaccines, Soberana 02, is more than 90% effective in protecting against symptomatic COVID-19 infection when used in combination with a related vaccine. Importantly, the combination seems to be effective against the highly transmissible Delta variant of the coronavirus SARS-CoV-2, which has caused surges in hospitalizations and death across the world and now accounts for nearly all COVID-19 cases in Cuba.

As of 18 November, 89% of Cuba's population — including children as young as 2 — has received at least one dose of Soberana 02 or another Cuban vaccine called Abdala, which is produced at the Center for Genetic

Engineering and Biotechnology (CIGB) in Havana. The centre [reported in July](#) that Abdala, a three-dose vaccine, was more than 92% effective in phase III trials that included more than 48,000 participants, but the full results have not yet been published.

Cuba's regulatory agency authorized Abdala and Soberana 02 shots for use in adults in July and August respectively, and health-care workers began immunizing children with both vaccines a few months later. The country has begun exporting the two home-grown vaccines to Venezuela, Vietnam, Iran and Nicaragua. And it has asked the World Health Organization to approve its vaccines — an important step towards making them available throughout the developing world.

Major strides

In developing Soberana 02, Vérez Bencomo's group drew on its existing 'conjugate' vaccine technology. Finlay's conjugate vaccines take a protein or a sugar from a bacterium or virus and chemically link it to a harmless fragment of a neurotoxin protein from the tetanus bacterium. The combination elicits a stronger immune response than either component alone. Conjugate vaccines against meningitis and typhoid are used around the world, and Cuba has been immunizing children with a vaccine of this type for years.

Vérez Bencomo's team adapted conjugate-vaccine technology to tackle COVID-19 by linking the tetanus-toxin protein to a portion, known as the receptor binding domain (RBD), of SARS-CoV-2's spike protein (the spike protein helps the virus to enter cells). After more than 14,000 people received two doses of the vaccine in a phase III trial, recipients' risk of symptomatic COVID-19 was reduced by 71%, compared with that of a placebo group of the same size — an efficacy similar to jabs made by Johnson & Johnson (J&J) in New Brunswick, New Jersey, and AstraZeneca in Cambridge, UK.



Health-care workers participating in a phase III trial wait to receive a Soberana 02 shot in March. Credit: Ramon Espinosa/Reuters

To improve on this protection, the Finlay team also gave participants a third shot. The researchers had previously tested a jab called Soberana Plus on people already sick from COVID-19 and found that it improved their immune response². So they gave Soberana Plus, which is based on the RBD protein alone, to another set of 14,000 participants who had already received two Soberana 02 doses — and discovered that the third dose raised overall efficacy to 92.4%.

Vérez Bencomo says the Finlay Institute can produce 10 million Soberana 02 doses per month. To further test the vaccine, he and his colleagues have partnered with the Pasteur Institute in Tehran to run a similar, 24,000-person trial in Iran and expect to publish those results soon.

CIGB's Abdala vaccine is also making major strides. As with Soberana 02, the technology behind it is adapted from an existing vaccine — one for hepatitis B — that Cuba developed and has used for many years. The researchers engineered yeast cells to produce a part of the RBD different

from that used in Soberana 02, and then purified the protein for use in Abdala. CIGB researcher Merardo Pujol Ferrer says that 24 million doses have been administered to 8 million people in Cuba, giving the researchers a large data set with which to track safety and efficacy. He says the team plans to publish its data later this month.

An expanding toolbox

Protein-based vaccines like Soberana 02 and Abdala [might have some advantages over other vaccine types](#), says Craig Laferrière, head of vaccine development at Novateur Ventures in Toronto, Canada, who has been comparing the safety and efficacy of COVID-19 jabs. Unlike the messenger RNA (mRNA) vaccines produced by Pfizer, based in New York City, and Moderna, based in Cambridge, Massachusetts, protein vaccines do not need to be kept at extremely low temperatures, making them easier to deliver to remote areas.

And they could have fewer side effects than AstraZeneca's and J&J's vaccines, which use an adenovirus to deliver the gene for a different portion of the RBD into cells and [have been linked to blood clots](#). Although Finlay's medRxiv manuscript (which is not peer reviewed) does not contain extensive clinical data, Laferrière expects Soberana 02's side effects will be minimal, because fewer than 1% of participants in the phase III trial developed a fever. Veréz-Bencomo says further data will be published soon.

But Laferrière adds that there are drawbacks to the approach, too. Protein-based vaccines are made using various types of cells to synthesize gobs of protein. Soberana 02 is produced in hamster ovary cells, which is more time-consuming than some other methods for making this type of vaccine. And evidence suggests³ that conjugate vaccines using tetanus-toxin protein are less effective in people who have already received another vaccine of this type, such as the childhood meningitis vaccine.



How protein-based COVID vaccines could change the pandemic

Vérez Bencomo says he has confidence in the vaccine's safety, largely because conjugate-vaccine technology has been used for decades without major problems. Having worked with it to create vaccines for use in children, the Finlay team also knew enough about dosing and side effects to leap into paediatric trials of Soberana 02, which began in June. Nearly 2 million children in Cuba have been vaccinated so far, and Vérez Bencomo says that unpublished data suggest the vaccine is safe and effective.

"I think it'll be a useful addition for the globe," says John Grabenstein, president of the vaccine consultancy Vaccine Dynamics in Easton, Maryland. "Everybody's using a different tool out of the toolbox, and they pretty much are all working." He says that the Soberana 02 data look solid, but it will take time to determine how long immunity conferred by the vaccine will last.

In the meantime, Cuba is pushing ahead with its COVID-19 vaccine-development strategy. Finlay's Soberana 01, which links the spike protein to a sugar from a meningitis-causing bacterium rather than the tetanus-toxin protein, and CIGB's Mambisa, a nasal vaccine that contains the same RBD fragment as is used in Abdala, are still in clinical trials.

doi: <https://doi.org/10.1038/d41586-021-03470-x>

References

1. 1.

Toledo-Romani M. E. *et al.* Preprint at medRxiv
<https://doi.org/10.1101/2021.10.31.21265703> (2021).

2. 2.

Chang-Monteagudo, A. *et al.* *Lancet Reg. Health Am.* **4**, 100079 (2021).

3. 3.

Dagan, R., Eskola, J., Leclerc, C. & Leroy, O. *Infect. Immun.* **66**, 2093–2098 (1998).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03470-x>

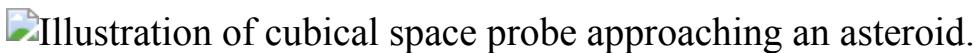
| [Section menu](#) | [Main menu](#) |

- NEWS
- 19 November 2021
- Update [24 November 2021](#)

NASA spacecraft will slam into asteroid in first planetary-defence test

The DART mission has launched. Next up, it will try out a manoeuvre that could one day deflect killer asteroids from Earth.

- [Alexandra Witze](#)



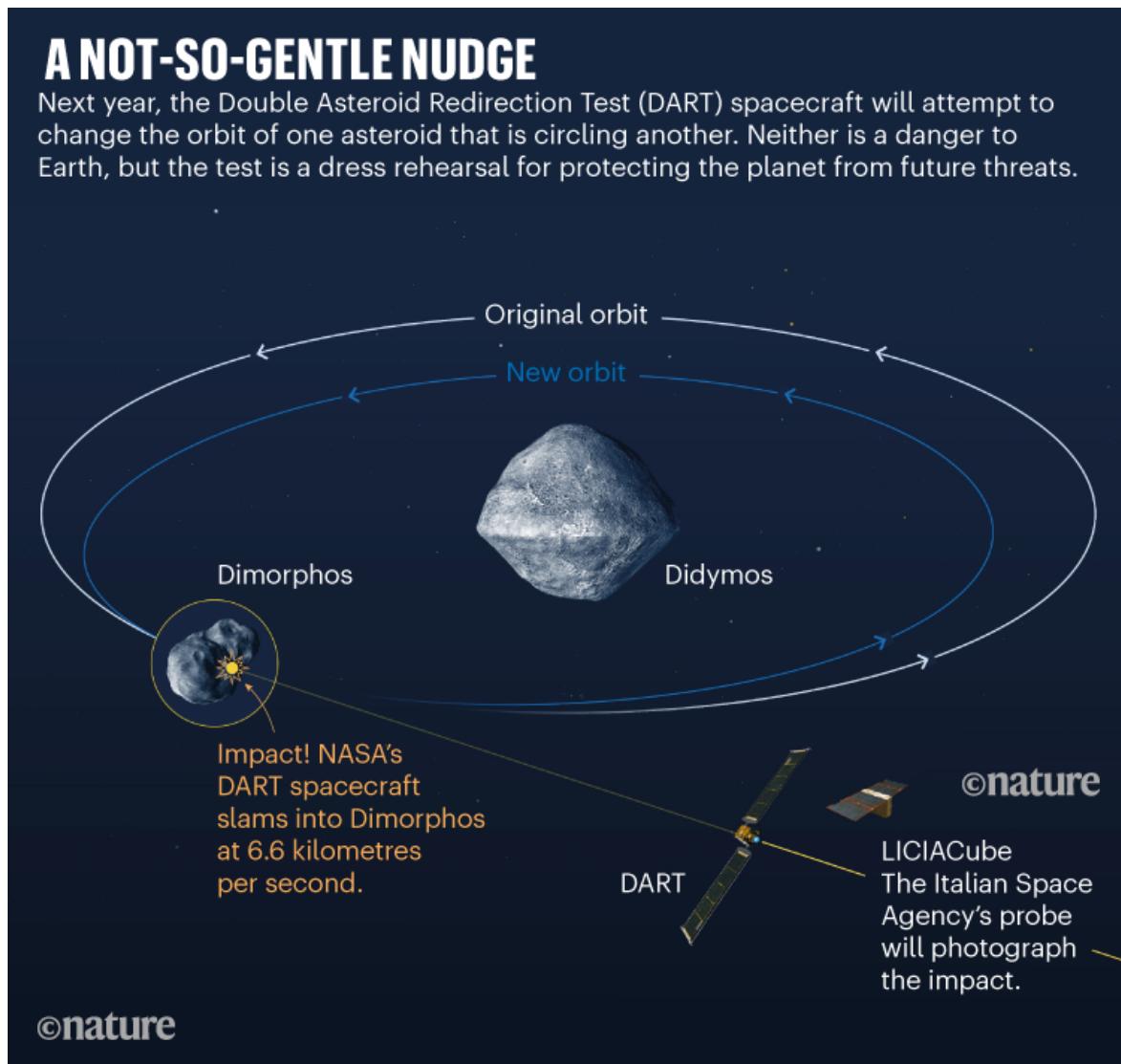
The DART mission aims to smash into the asteroid Dimorphos, as shown in this artist's illustration. Credit: NASA/Johns Hopkins University APL

NASA has just launched a multimillion-dollar spacecraft — to slam into an asteroid. Rather than being a catastrophic error, however, it will be the first test of a way to protect Earth from killer asteroids.

The asteroid that NASA is smashing into, called Dimorphos, is not a threat to Earth. But researchers want to see whether they can change its trajectory, long before they might need to use such a strategy to deflect a truly dangerous asteroid.

“The odds of something large enough to be a problem, that we would have to deflect, are pretty slim in our lifetimes,” says Andy Rivkin, a planetary scientist at the Johns Hopkins University Applied Physics Laboratory (JHU-APL) in Laurel, Maryland, which built the spacecraft for NASA. “But sometimes your number comes up when you don’t expect it, and it’s good to have an insurance policy.”

Launched from California on 23 November, the spacecraft is called the Double Asteroid Redirection Test (DART)¹. Its target is a pair of asteroids that travel together through space, one orbiting the other as they circle the Sun (see ‘A not-so-gentle nudge’). Dimorphos, the smaller of the two at 160 metres wide, orbits Didymos, which is nearly 5 times larger and is named after the Greek word for ‘twin’.



Credit: Adapted from NASA/Johns Hopkins University APL

In late September or early October of next year DART will slam headlong into Dimorphos at 6.6 kilometres per second. The impact should shrink Dimorphos’s orbit so that it circles Didymos at least 73 seconds faster than

before. (Dimorphos is named after the Greek for ‘having two forms’, to signal NASA’s intent to change the asteroid’s orbit.) Astronomers using telescopes on Earth will watch Didymos for signs of that orbital change — which would be evident in the way its brightness changes over time, as Dimorphos passes in front of and behind it.

This complicated choreography is meant to test the idea that smashing into an asteroid can give it enough of a nudge to keep it from hitting Earth, says Nancy Chabot, a planetary scientist at JHU-APL who works on the mission. Using the non-threatening pair Dimorphos and Didymos is “a really smart and a safe way to do this first test”, she says. The impact will occur when the asteroids are 11 million kilometres from Earth.

Battling asteroids

Small asteroids and asteroid fragments hit Earth all the time, but most of them disintegrate in the atmosphere or fall harmlessly to the ground as meteorites. [NASA has identified](#) more than 27,000 asteroids with trajectories that bring them close to Earth. The worry is that some new asteroid could appear, headed directly towards the planet — and that it would be large enough to cause serious consequences when it hits, just as with the asteroid that helped to kill off the dinosaurs and other life on Earth 66 million years ago.

Space scientists have floated all sorts of ideas to battle incoming asteroids, the most dramatic of which involves blasting them with nuclear weapons². Other, less cinematically worthy strategies involve altering the asteroid’s trajectory by flying a spacecraft alongside to tug on it using gravitational forces, or smashing into it as the US\$330-million DART mission will.



Researchers at the Johns Hopkins University Applied Physics Laboratory inspect the DART spacecraft during testing in July. Credit: Ed Whitman/NASA/Johns Hopkins University APL

Depending on the angle at which DART hits the asteroid, it could kick up a small cloud of dust and rubble. The impact will probably leave a crater that could be around 10 metres across. At the same time, bits of the spacecraft's wreckage might scatter across the asteroid's surface, but exactly how DART will break apart remains to be seen. "Just from a pure crime-scene sense, a lot of us are curious about that," Rivkin says.

Researchers will have a chance to get answers, because minutes later, a tiny probe funded by the Italian Space Agency will fly by to photograph the aftermath³. Named LICIACube, it will travel aboard DART and is the agency's first autonomously guided deep-space mission. LICIACube will be released from DART 10 days before impact, and come within 55 kilometres of Dimorphos. As it whizzes past, its cameras should spot the dust cloud, if the impact kicks one up, and possibly the resulting crater. "We might be surprised by the images we collect," says Elisabetta Dotto, an astronomer at

the National Institute for Astrophysics in Rome, which is leading the collaboration of Italian universities and institutions involved in LICIACube.

In 2026, a follow-up spacecraft, the European Space Agency's Hera mission, will visit Dimorphos to take more detailed pictures of the impact site.



[Record number of asteroids seen whizzing past Earth in 2020](#)

Data collected by the DART mission should help scientists to understand how impacts affect asteroids, says Megan Bruck Syal, a physicist at the Lawrence Livermore National Laboratory in California, who will model what happens to Dimorphos. But DART is just one test involving one kind of space rock. There could be scenarios in which planetary defenders want to hit an asteroid with more speed than DART will reach when it hits Dimorphos, or in which they need to pummel an asteroid with several impactors to change its course. “We need to do more experiments like this,” Bruck Syal says.

Although many other spacecraft have been deliberately smashed into celestial objects at the ends of their lives, DART promises to be the first to hit a planetary body in the name of saving Earth.

Nature **600**, 17-18 (2021)

doi: <https://doi.org/10.1038/d41586-021-03471-w>

Updates & Corrections

- **Update 24 November 2021:** This story has been updated to reflect the launch of the DART mission.

References

1. 1.

Rivkin, A. S. *et al.* *Planet. Sci. J.* **2**, 173 (2021).

2. 2.

King, P. K. *et al.* *Acta Astronaut.* **188**, 367–386 (2021).

3. 3.

Dotto, E. *et al.* *Planet. Space Sci.* **199**, 105185 (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03471-w>

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

- NEWS
- 26 November 2021

Outcry as men win outsize share of Australian medical-research funding

An analysis showing that the National Health and Medical Research Council awards fewer ‘investigator grants’ to women prompts thousands to sign a petition calling for gender quotas.

- [Holly Else](#)



Female biomedical researchers in Australia receive less funding than their male counterparts, on average. Credit: Lisa Maree Williams/Getty

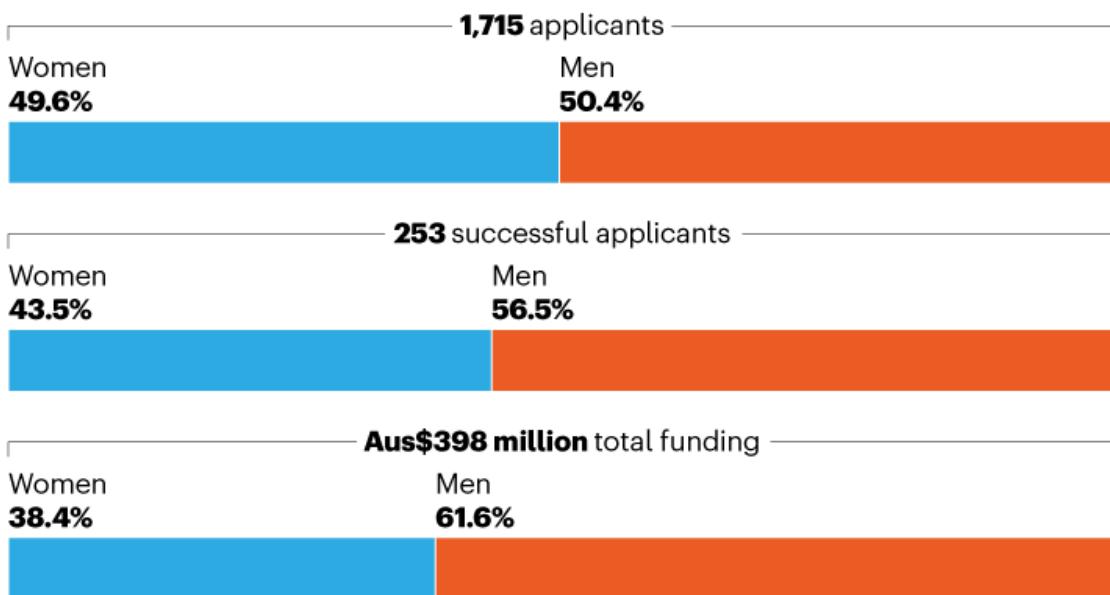
Men secure a greater share of medical-research funding than women in Australia's largest grant-award programme, despite applying at similar rates, according to an analysis. The issue was [first flagged by researchers](#) in 2019; however, this time, nearly 6,000 people have signed a petition calling for the [introduction of gender quotas](#).

"It is soul-destroying to watch a disproportionately higher number of young, bright women stagnate or be pushed out of the system compared to their male counterparts," says Rachael Murray, a biomedical scientist at the Queensland University of Technology in Brisbane.

The findings come after the National Health and Medical Research Council (NHMRC) completely overhauled its funding programmes in 2018–2019, attempting to take gender equity into account.

MORE MONEY FOR MEN

Although the number of male and female applicants for the National Health and Medical Research Council's investigator grants in 2021 were comparable, men received 23% more grants and Aus\$95 million more in funding than women.



©nature

Source: L. Purton & J. Borger [Is Australia's largest medical research funding body doing enough to retain women in STEMM?](#) (Women's Agenda, 2021).

Funding disadvantages women

The awards in question are the NHMRC's investigator grants, comprising Australia's largest research-funding programme, which consolidates salary and project support into one flexible, five-year grant for the best researchers at various stages of their careers. Before 2019, scientists had to apply for a fellowship to fund their salary, and separate grants for their research.

The NHMRC has [previously acknowledged problems](#) with equity, and in 2018 it released a gender-strategy report. The hope was that by combining salary and research funding, the new investigator grants would allow projects to continue if their leaders needed to work part-time because of childcare or other responsibilities.

But the data from the latest round of funding in 2021, released in October, suggest that the new scheme still favours men over women.

Men won more grants and were awarded more money, according to Louise Purton, a stem-cell biologist at St Vincent's Institute of Medical Research in Melbourne, and Jessica Borger, a medical researcher at Monash University in Melbourne, who crunched the numbers and [revealed the disparity in an article](#) for Australian news site Women's Agenda.



Female researchers in Australia less likely to win major medical grants than males

“Men were disproportionately awarded a staggering 23% more grants, corresponding to an extra \$95 million in funding,” they write.

Across the scheme, men and women applied for grants at similar rates — with 865 men applying for funding, alongside 850 women. But 143 men secured funding totalling Aus\$245 million (US\$176 million), compared with 110 women netting just \$153 million (see ‘More money for men’).

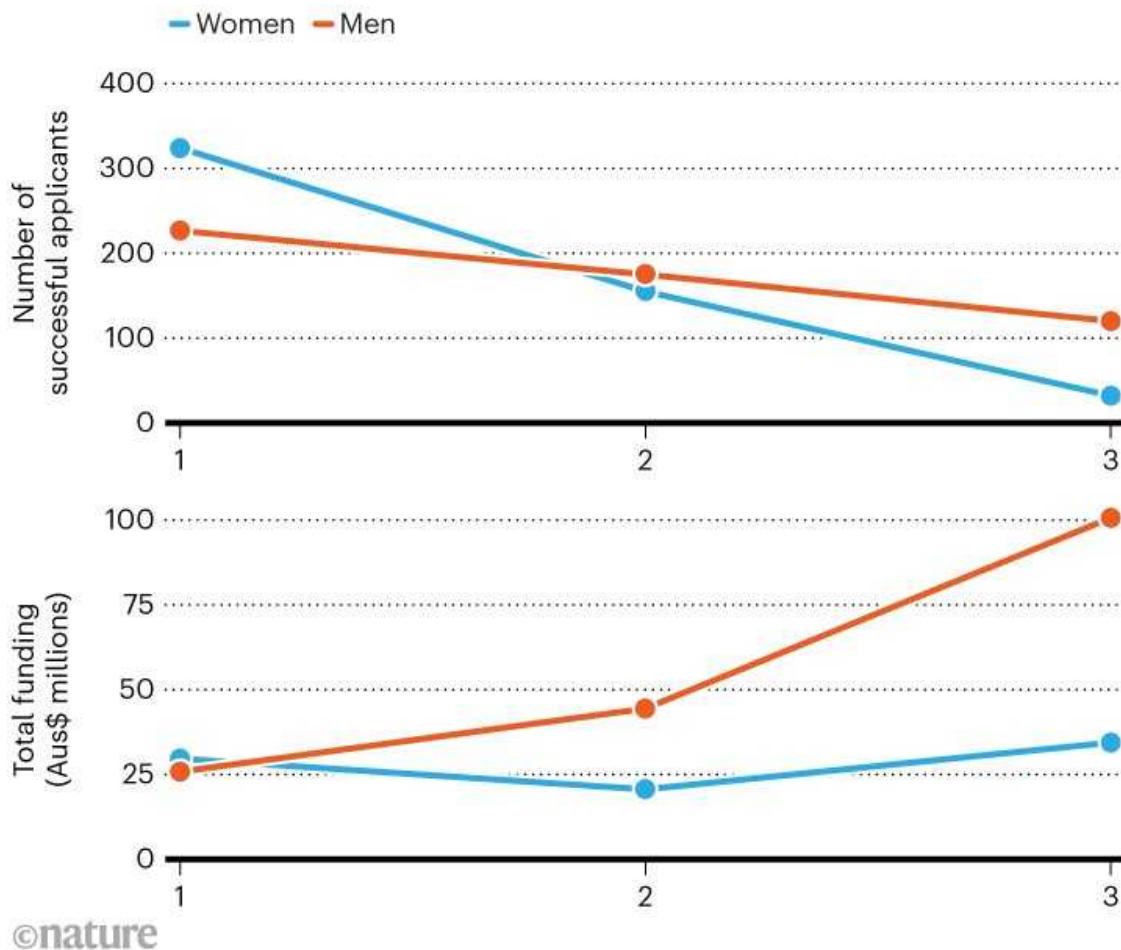
The scheme offers grants at three levels of seniority. At the most junior level, women as a whole secured equal amounts of funding to men, but the distribution of grants for more established scientists was skewed heavily towards men. Only about 20% of the awards for the most experienced scientists went to women, according the analysis.

Anne Kelso, the NHMRC’s chief executive, agreed that there are clear gender disparities, but says that they reflect the disparities in the gender balance of the make-up of scientists at various career stages at Australia’s universities.

“The single biggest contributor to the investigator grant outcomes is the predominance of male applicants at the most senior level of the scheme,” she told *Nature*. At that level, for which the funds awarded are the highest (see ‘Shrinking success rate’), there were about four times more male than female applicants, she says.

SHRINKING SUCCESS RATE

More women at early stages of their careers (1) who applied for the National Health and Medical Research Council's 2021 investigator grants were successful than men, and women at that level were collectively awarded slightly more funding. However, their chances and payouts decreased substantially at the level of junior laboratory head (2) and at the most senior levels of funding (3).



Source: L. Purton & J. Borger [Is Australia's largest medical research funding body doing enough to retain women in STEMM?](#) (Women's Agenda, 2021).

‘Australia has a terrible record’

The petition, created last month in response to Purton and Borger’s analysis, says that the NHMRC “is awarding women significantly less funding than their male counterparts in a broken system”, which it says “requires an urgent strategic overhaul”.

It is calling for the funding body to allocate the same amount of money to men and women, and to include a separate pot for non-binary applicants. It also pushes for set gender quotas for fellowships at each level of seniority.

In response to these calls, Kelso says “all options are on the table”, adding that the “NHMRC schemes are under continuous review to ensure they are meeting their objectives” in terms of gender equity.

Male researchers’ ‘vague’ language more likely to win grants

Megan Head, an evolutionary biologist at the Australian National University in Canberra, agrees with the adoption of gender quotas for funding schemes. “Australia has a terrible record with gender equity in science, technology, engineering and mathematics,” she says.

Similar issues have been reported in other nations. A [2019 survey](#) in the United Kingdom suggested that women starting their own laboratories tended to get a lower salary and have fewer staff than did men.

But an [analysis of data](#) from the US National Institutes of Health, published in 2018, suggests that although women are more likely than men to drop out of academia at early career stages, once a female scientist receives one big grant from the agency, she is almost as likely to remain in research as her male counterparts.

Despite the disparity in grant outcomes this year, Kelso argues that across the whole of the NHMRC — which distributes \$1.1 billion annually — funding rates for men and women are nearly equal. And since 2019, the body has also introduced mechanisms that are designed to improve the diversity of grant holders, including the use of a separate pot of money to fund high-scoring female applicants who just missed out on securing an investigator grant.

Nature **600**, 18 (2021)

doi: <https://doi.org/10.1038/d41586-021-03536-w>

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

- NEWS
- 22 November 2021

China creates vast research infrastructure to support ambitious climate goals

Carbon-neutrality institutes, and other initiatives to support a pledge to achieve net zero by 2060, are popping up like mushrooms across China.

- [Smriti Mallapaty](#) 0



Wind turbine blades are secured and examined before being hoisted into position in China's Jiangsu province. Credit: He Jinghua/VCG via Getty

China, the world's top carbon emitter, has for the first time published plans broadly outlining how it might achieve net-zero carbon emissions by 2060,

and a peak of emissions before 2030 — promises it made [in 2019](#).

Researchers say the documents, released ahead of the COP26 climate talks that [concluded on 15 November](#), send a strong message to industry, government agencies and universities in China to ramp up their efforts to help the country meet its climate goals.



['COP26 hasn't solved the problem': scientists react to UN climate deal](#)

Already this year, more than ten prominent universities and institutions have set up carbon-neutrality-research institutes; the Chinese Academy of Sciences launched a centre last month.

“We start right now,” says Jiang Kejun, a modeller at the Energy Research Institute in Beijing.

The country is experiencing a “national movement”, says Wu Libo, an environmental economist at Fudan University in Shanghai, as companies, regional governments and academia shift gears.

Wu is also director of the Shanghai Research Institute for Energy and Carbon Neutrality Strategy, a collaboration launched earlier this month by the university and the Shanghai city government. Once it secures funding for research grants, the institute will focus on the deregulation of the electricity market and climate finance, she says.

Significant challenge

Achieving carbon neutrality by 2060 “is a big challenge for China”, says Xie Xiaomin, an energy-policy researcher at Shanghai Jiao Tong University (SJTU). She is vice-director of the university’s Research Institute of Carbon Neutrality, which was established in May and has already received about 20 million yuan (US\$3.1 million) in funding to work on a broad range of energy technologies, she says.

From emitting more than 11 gigatonnes of carbon dioxide in 2020, China has to drop to net zero within four decades. This is a scale and speed that no other country has attempted before, says Gang He, an energy-systems modeller at Stony Brook University in New York, who has studied China’s power system.

China’s current emissions are more than double the United States’ and three times as big as those of India, which made a similar pledge to [reach net zero by 2070](#) during COP26.

“There will be a lot of areas needing contribution from researchers,” says Fu Sha, a modeller with the non-profit Energy Foundation China in Beijing. These include low-carbon energy technologies, from hydrogen fuel cells to batteries; market-based mechanisms to control emissions, such as carbon taxes and trading schemes; and modelling that will help local governments and industries set realistic targets for cuts, she says.

Researchers will also need to study which sections of China’s population will be most affected by the transition and learn how to help them cope, says He. Establishing the path to “a more just and inclusive transition would be a very important research topic”, he says.



Hydrogen fuel cells are tested in a lab at Shanghai ReFire Technology, a start-up that manufactures fuel-cell engines. Credit: Qilai Shen/Bloomberg via Getty

Solar, wind and hydro

But many research institutes have a long way to go in terms of aligning their research departments with the carbon-neutrality goals, says Jiang. They will need to drop traditional engineering subjects that focus, for example, on coal-fired boiler technology and internal combustion engines, notes Zhang Xiliang, a climate modeller at Tsinghua University in Beijing, which launched its own Institute for Carbon Neutrality in September.

The two highly anticipated policy documents the central government released in October — a working guidance and a 2030 action plan — outline a path for researchers and flesh out for the first time how China plans to achieve its carbon goals.

The documents talk about “strengthening basic research and research on cutting-edge technologies” — such as [nuclear fusion](#), smart grids and new materials — and the formulation of “an action plan to ensure science and technology better support” achieving carbon neutrality.

Also detailed in the documents are commitments to increase the share of power China produces from renewable and nuclear sources from just 16% today to 80% by 2060. Solar and wind capacity are planned to reach 1,200 gigawatts by 2030 — enough to power hundreds of millions of households — and 80 gigawatts of hydropower will be installed over the next decade.

Carbon capture and sequestration technologies will also be important to achieving China’s goals, according to the documents, and could be another focus for researchers. There is also a plan for electric and hybrid vehicles to make up 40% of those sold by 2030. Much more detail is expected to come in future documents, say researchers.

Global ramifications

Despite this rush of activity towards net zero within China, some researchers are still disappointed with the commitments the country has made so far.

At the COP26 meeting, the nation signed a joint declaration with the United States to find steeper cuts, but frustrated some scientists by not offering more-aggressive pledges for the next decade, and also for pushing for text concerning coal in the final agreement to be edited from ‘phased out’ to ‘phased down’.



Scientists cheer India's ambitious carbon-zero climate pledge

China's promises so far are not likely to keep global warming to below 1.5 °C — the more aspirational of the targets set by world governments at the United Nations Paris climate meeting in 2015 — says Yan Qin, an economist and carbon analyst based in Oslo with Refinitiv, a company that provides data on financial markets. “What has been pledged falls short,” she says.

China has also not yet set any absolute limits on emissions, nor targets for emissions of other greenhouse gases besides carbon, such as methane, but has committed to doing both. Researchers say these measures will be crucial for the world to have a chance at limiting global warming to below 1.5 °C.

China's size means its pledges have global ramifications, says Pep Canadell, chief research scientist at the Australian government's CSIRO Climate Science Centre in Canberra. “When China moves a little to the left or the right, up or down, the whole world feels it.”

And He says that although China's goals are not as ambitious as some would like, they are at least realistic. “What China should do is sometimes not what China can deliver,” he says.

Nature **600**, 19-20 (2021)

doi: <https://doi.org/10.1038/d41586-021-03491-6>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03491-6>

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

- NEWS
- 19 November 2021
- Correction [22 November 2021](#)

Scientists question Max Planck Society's treatment of women leaders

An open letter signed by 145 female scientists from across the world expresses concern over recent dismissals of women leaders.

- [Alison Abbott](#) ⁰



The Max Planck Institute for the Science of Human History in Jena, Germany.Credit: Vladimir Bulgar/Alamy

A group of 145 leading female scientists from across the world has signed an open letter to Germany's Max Planck Society (MPS), expressing concern over "the highly publicized dismissals, demotions, and conflicts involving female directors of Max Planck Institutes".

The MPS is Germany's most powerful basic-research organization, with 86 institutes and research facilities. In the past three years, at least four of the most senior researchers at these institutes — called directors — have been charged with bullying, three of them women.

“Female leaders are judged more harshly, and allegations of leadership shortcomings are more often made against female leaders than male ones,” states the letter. It calls on the society to check its personnel statistics to determine whether women are over-represented among those who have left their positions before retirement or faced sanctions or demotions.

A spokesperson for the MPS rejects the charges of gender bias in the letter, and says that its investigations into research misconduct “are generally conducted by persons who are neutral, objective and independent”.

Archaeologist demoted

The letter was sent to all the MPS’s senate members on 18 November, a day before a senate meeting was due to consider the most recent case — the demotion of archaeologist Nicole Boivin from her position of director at the Max Planck Institute for the Science of Human History in Jena, which she had held for five years.

“This case prompted our letter,” says physicist Ursula Keller at the Swiss Federal Institute of Technology (ETH) in Zurich, who was one of the letter’s instigators. “But we know the problem is much wider than the Max Planck, and we want the issue to be discussed in the community.”

Similar cases involving women in top academic positions at several other major European research institutions have also hit the press since 2018, “indicating that these issues involving senior women extend well beyond the MPG”, the letter says, using the German acronym for the society. It stresses that the signatories do not endorse tolerance of bullying.

The fallout from the many cases could be grave, the letter argues. “Highly publicized failures of women at top level positions in science could have a chilling effect on young women considering careers in science and engineering,” it says.

Boivin was accused of bullying young scientists and appropriating scientific ideas of colleagues. The protracted internal investigation began in October 2018, but was accelerated this autumn. On 22 October, Boivin was demoted from her position without advance notice. She remains in charge of a smaller research group.

Decision challenged

Boivin denies all charges and is challenging the Max Planck decision in court. An injunction ruling is expected on 3 December, which would allow her to continue as director until the court case is settled.

The case has proved controversial. Many Max Planck scientists have also written to the society to express their concern over the procedure. This includes two letters from groups of postdocs and PhD students at Boivin's institute, expressing their support for her and challenging the fairness of the investigation.

Nobel prizewinner Christiane Nüsslein-Volhard, a director at the Max Planck Institute for Developmental Biology in Tübingen, also has written a letter calling on the senate to reverse the decision. Only 54 of the society's 304 directors are women, and such cases will make it hard to encourage more female scientists to join, she writes. "One must conclude that there are still deep-seated, unacknowledged prejudices against women in leadership positions, that the leadership behaviour of women directors is measured by different, [harsher] standards than that of men," she writes.

MPS spokesperson Christina Beck told *Nature* that the confidential investigation was neutral and objective. In 2018, the society introduced leadership training for directors that includes unconscious-bias training, she says.

Nature **600**, 20 (2021)

doi: <https://doi.org/10.1038/d41586-021-03492-5>

Updates & Corrections

- **Correction 22 November 2021:** An earlier version of this article said that groups of postdocs and PhD students had written to the senate. The groups in question wrote to the MPS.

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03492-5>

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

- NEWS
- 25 November 2021
- Update [27 November 2021](#)

Heavily mutated Omicron variant puts scientists on alert

Researchers are racing to determine whether a fast-spreading coronavirus variant poses a threat to COVID vaccines' effectiveness.

- [Ewen Callaway](#)



Medics at an infectious-disease unit in South Africa, where a new strain of COVID is spreading quickly. Credit: Alet Pretorius/Gallo Images/Getty

Researchers in South Africa are racing to track the concerning rise of a new variant of the SARS-CoV-2 coronavirus that causes COVID-19. The variant harbours a large number of the mutations found in other variants, including Delta, and it seems to be spreading quickly across South Africa.

A top priority is to follow the variant more closely as it spreads: it was first identified in Botswana earlier this month and has since turned up in a traveller arriving in Hong Kong from South Africa. Scientists are also trying to understand the variant's properties, such as whether it can evade immune responses triggered by vaccines and whether it causes more or less severe disease than other variants do.

“We’re flying at warp speed,” says Penny Moore, a virologist at the University of the Witwatersrand in Johannesburg, South Africa, whose lab is gauging the variant’s potential to dodge immunity from vaccines and previous infections. There are anecdotal reports of reinfections and of cases in vaccinated individuals, but “at this stage it’s too early to tell anything”, Moore adds.

“There’s a lot we don’t understand about this variant,” Richard Lessells, an infectious-diseases physician at the University of KwaZulu-Natal in Durban, South Africa, said at a press briefing organized by South Africa’s health department on 25 November. “The mutation profile gives us concern, but now we need to do the work to understand the significance of this variant and what it means for the response to the pandemic.”

On 26 November, the World Health Organization (WHO) designated the strain, known as B.1.1.529, as a variant of concern and named it Omicron, on the advice of scientists who are part of the WHO’s Technical Advisory Group on SARS-CoV-2 Virus Evolution. Omicron joins Delta, Alpha, Beta and Gamma on the current WHO list of variants of concern.

Researchers also want to measure the variant’s potential to spread globally — possibly sparking new waves of infection or exacerbating ongoing rises being driven by Delta.

Changes to spike

Researchers spotted B.1.1.529 in genome-sequencing data from Botswana. The variant stood out because it contains more than 30 changes to the spike protein — the SARS-CoV-2 protein that recognizes host cells and is the main target of the body’s immune responses. Many of the changes have been found in variants such as Delta and Alpha, and are linked to heightened infectivity and the ability to evade infection-blocking antibodies.

The apparent sharp rise in cases of the variant in South Africa’s Gauteng province — home to Johannesburg — is also setting off alarm bells. Cases increased rapidly in the province in November, particularly in schools and among young people, according to Lessells. Genome sequencing and other genetic analysis from a team led by Tulio de Oliveira, a bioinformatician at the University of KwaZulu-Natal, found that the B.1.1.529 variant was responsible for all 77 of the virus samples they analysed from Gauteng, collected between 12 and 20 November. Analysis of hundreds more samples are in the works.

The variant harbours a spike mutation that allows it to be detected by genotyping tests that deliver results much more rapidly than genome sequencing does, Lessells said. Preliminary evidence from these tests suggest that B.1.1.529 has spread considerably further than Gauteng. “It gives us concern that this variant may already be circulating quite widely in the country,” Lessells said.

Vaccine effectiveness

To understand the threat B.1.1.529 poses, researchers will be closely tracking its spread in South Africa and beyond. Researchers in South Africa mobilized efforts to quickly study the Beta variant, identified there in late 2020, and a similar effort is starting to study B.1.1.529.

Moore’s team — which provided some of the first data on Beta’s ability to dodge immunity — has already begun work on B.1.1.529. They plan to test the virus’s ability to evade infection-blocking antibodies, as well as other immune responses. The variant harbours a high number of mutations in regions of the spike protein that antibodies recognize, potentially dampening their potency. “Many mutations we know are problematic, but many more

look like they are likely contributing to further evasion,” says Moore. There are even hints from computer modelling that B.1.1.529 could dodge immunity conferred by another component of the immune system called T cells, says Moore. Her team hopes to have its first results in two weeks.

“A burning question is ‘does it reduce vaccine effectiveness, because it has so many changes?’,” says Aris Katzourakis, who studies virus evolution at the University of Oxford, UK. Moore says breakthrough infections have been reported in South Africa among people who have received any of the three kinds of vaccines in use there, from Johnson & Johnson, Pfizer–BioNTech and Oxford–AstraZeneca. Two quarantined travellers in Hong Kong who have tested positive for the variant were vaccinated with the Pfizer jab, according to news reports. One individual had travelled from South Africa; the other was infected during hotel quarantining.

Researchers in South Africa will also study whether B.1.1.529 causes disease that is more severe or milder than that produced by other variants, Lessells said. “The really key question comes around disease severity.”

So far, the threat B.1.1.529 poses beyond South Africa is far from clear, researchers say. It is also unclear whether the variant is more transmissible than Delta, says Moore, because there are currently low numbers of COVID-19 cases in South Africa. “We’re in a lull,” she says. Katzourakis says that countries where Delta is highly prevalent should be watching for signs of B.1.1.529. “We need to see what this virus does in terms of competitive success and whether it will increase in prevalence.”

Nature **600**, 21 (2021)

doi: <https://doi.org/10.1038/d41586-021-03552-w>

Updates & Corrections

- **Update 27 November 2021:** The story now includes that the World Health Organization has named the variant Omicron and designated it as a variant of concern.

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03552-w>

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

- NEWS FEATURE
- 01 December 2021

What humanity should eat to stay healthy and save the planet

What we eat needs to be nutritious and sustainable. Researchers are trying to figure out what that looks like around the world.

- [Gayathri Vaidyanathan](#) ⁰

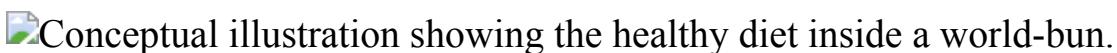


Illustration by Paweł Jońca

A clutch of fishing villages dot the coast near Kilifi, north of Mombasa in Kenya. The waters are home to parrot fish, octopus and other edible species. But despite living on the shores, the children in the villages rarely eat seafood. Their staple meal is ugali, maize (corn) flour mixed with water, and most of their nutrition comes from plants. Almost half the kids here have stunted growth — twice the national rate.

In 2020, Lora Iannotti, a public-health researcher at Washington University in St. Louis, and her Kenyan colleagues asked people in the villages why the children weren't eating seafood, even though all the parents fish for a living; studies show that fish and other animal-source foods can improve growth¹. The parents said it made more financial sense for them to sell their catch than to eat it.

So, Iannotti and her team are running a controlled experiment. They have given fishers modified traps that have small openings that allow young fish to escape. This should improve spawning and the health of the overfished ocean and reef areas over time, and eventually increase incomes, Iannotti says. Then, for half the families, community health workers are using home

visits, cooking demonstrations and messaging to encourage parents to feed their children more fish, especially plentiful and fast-growing local species such as ‘tafi’, or white spotted rabbitfish (*Siganus canaliculatus*) and octopus. The scientists will track whether children from these families eat better and are growing taller than ones who don’t receive the messaging.

The aim of the experiment, says Iannotti, is to understand “which sea foods can we choose that are healthy for the ecosystem as well as healthy in the diet”. The proposed diet should also be culturally acceptable and affordable, she says.

Iannotti is wrestling with questions that are a major focus of researchers, the United Nations, international funders and many nations looking for diets that are good for both people and the planet. More than 2 billion people are overweight or obese, mostly in the Western world. At the same time, 811 million people are not getting enough calories or nutrition, mostly in low- and middle-income nations. Unhealthy diets contributed to more deaths globally in 2017 than any other factor, including smoking². [As the world’s population continues to rise](#) and more people start to eat like Westerners do, the production of meat, dairy and eggs will need to rise by about 44% by 2050, according to the UN Food and Agriculture Organization (FAO).

That poses an environmental problem alongside the health concerns. Our current industrialized food system already emits about one-quarter of the world’s greenhouse-gas emissions. It also accounts for 70% of freshwater use and 40% of land coverage, and relies on fertilizers that disrupt the cycling of nitrogen and phosphorus and are responsible for much of the pollution in rivers and coasts³.



A child is weighed as part of a study into sustainable fishing and child nutrition in a village near Kilifi Creek, Kenya. Credit: Lora Iannotti

In 2019, a consortium of 37 nutritionists, ecologists and other experts from 16 countries— the EAT–*Lancet* Commission on Food, Planet, Health — released a report⁴ that called for a broad dietary change that would take into account both nutrition and the environment. A person following the EAT–*Lancet* reference diet would be ‘flexitarian’, eating plants on most days and occasionally a small amount of meat or fish.

The report provoked a flurry of attention towards sustainable diets, and some criticism about whether it was practical for everyone. Some scientists are now trying to test environmentally sustainable diets in local contexts, without compromising nutrition or damaging livelihoods.

“We need to make progress toward eating diets that have dramatically lower ecological footprints, or it’ll be a matter of a few decades before we start to see global collapses of biodiversity, land use and all of it,” says Sam Myers, director of the Planetary Health Alliance, a global consortium in Boston, Massachusetts, that studies the health impacts of environmental change.

Emissions on the menu

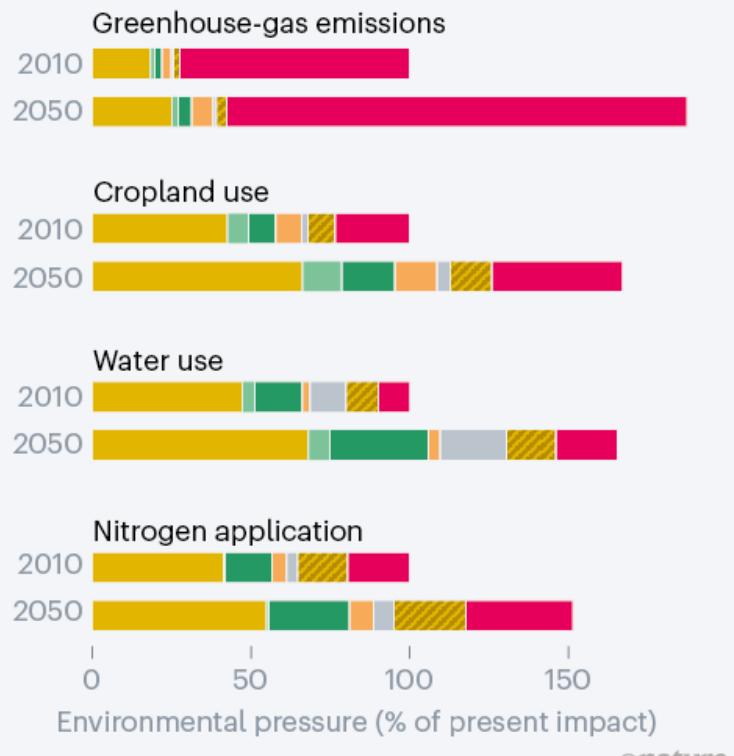
Producing food generates so much greenhouse-gas pollution⁵ that at the current rate, even if nations cut all non-food emissions to zero, they still wouldn’t be able to limit temperature rise to 1.5 °C — the climate target in the Paris agreement. A large proportion of emissions from the food system — 30–50%, according to some estimates — comes from the livestock supply chain, because animals are inefficient at converting feed to food.

In 2014, David Tilman, an ecologist at the University of Minnesota in Saint Paul, and Michael Clark, a food-systems scientist at the University of Oxford, UK, estimated that changes in urbanization and population growth globally between 2010 and 2050 would cause an 80% increase in food-related emissions⁶.

Environmental costs

Between 2010 and 2050, predicted growth in population and income could drive a 50–90% increase in environmental pressures exerted by food systems, such as climate impacts and freshwater use.

- Staple crops
- Plant proteins
- Fruits and vegetables
- Vegetable oils
- Sugars
- Other crops
- Animal products



©nature

Source: M. Springmann *et al. Nature* **562**, 519–525 (2018)

But if everyone, on average, ate a more plant-based diet, and emissions from all other sectors were halted, the world would have a 50% chance of meeting the 1.5 °C climate-change target⁵. And if diets improved alongside broader changes in the food system, such as cutting down waste, the chance of hitting the target would rise to 67%.

Such findings are not popular with the meat industry. For example, when in 2015, the US Department of Agriculture was revising its dietary guidelines, which happens every five years, it briefly considered factoring in the environment after researchers lobbied the advisory committee. But the idea was overruled, allegedly in response to industry pressure, says Timothy Griffin, a food-systems scientist at Tufts University in Boston, who was involved in the lobbying effort⁷. Nonetheless, people took notice of the attempt. “The biggest accomplishment is it brought a lot of attention to the issue of sustainability,” he says.

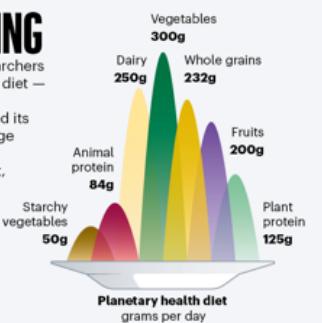
The EAT–*Lancet* Commission, which was funded by Wellcome, a UK-based charity, helped to build a stronger case. Nutritionists reviewed the literature to craft a basic healthy diet composed of whole foods. Then the team set environmental limits for the diet, including carbon emissions, biodiversity loss and the use of fresh water, land, nitrogen and phosphorus. Breaching such environmental limits could make the planet inhospitable to humans.⁸

They ended up with a diverse and mainly plant-based meal plan (see ‘Healthy eating’). The maximum red meat the 2,500-calorie per day diet allows in a week for an average-weight 30-year-old is 100 grams, or one serving of red meat. That’s less than one-quarter of what a typical American consumes. Ultra-processed foods, such as soft drinks, frozen dinners and reconstituted meats, sugars and fats are mostly avoided.

HEALTHY EATING

A commission of food researchers devised a 'planetary health' diet — meant to be nutritious and sustainable — and compared its composition with the average diets in different regions. Further studies showed that, in many regions, following the proposed diet would be prohibitively expensive.

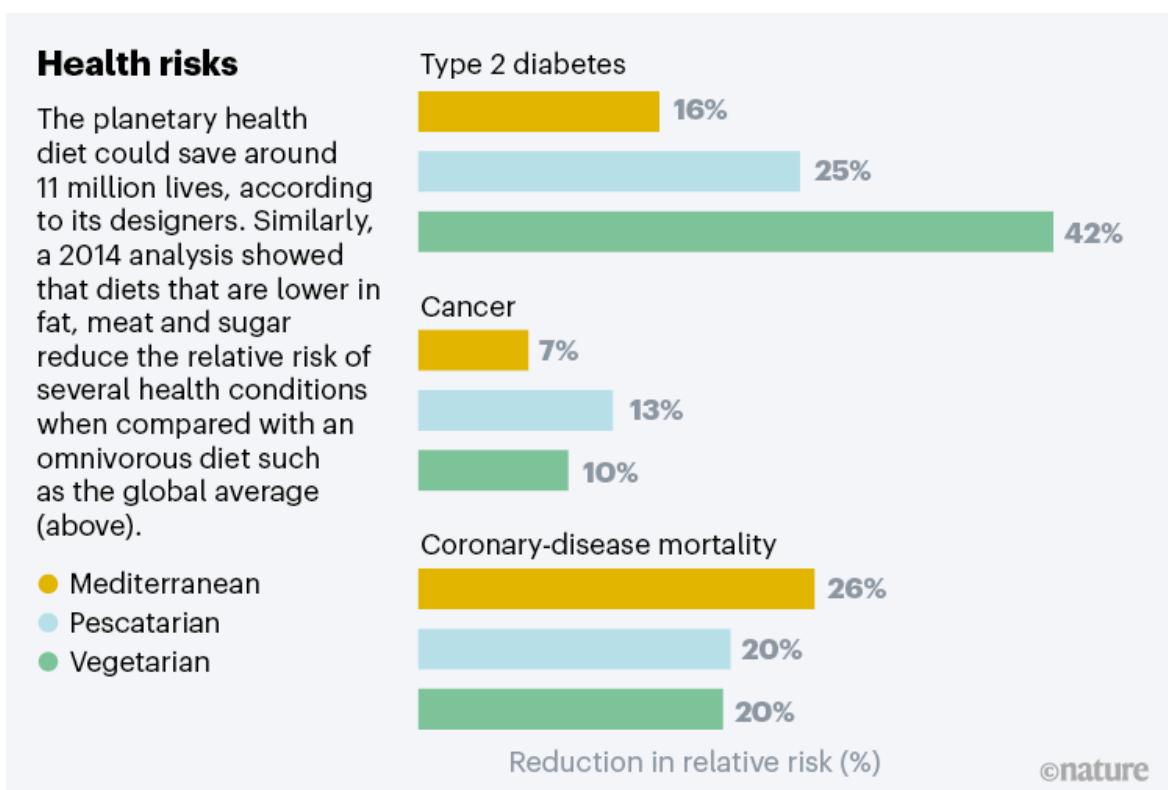
By Kerri Smith
Design by Jasia Krzysztofiak



Sources: Intakes, Ref. 4; Costs, Ref. 12

This diet would save the lives of about 11 million people every year, the commission estimated⁴. “It is possible to feed 10 billion people healthily, without destroying ecosystems further,” says Tim Lang, food-policy researcher at the City University of London and a co-author of the EAT–Lancet report. “Whether the hardliners of the cattle and dairy industry like it or not, they are really on the back foot. Change is now inevitable.”

Many scientists say the EAT–Lancet diet is excellent for wealthy nations, where the average person eats 2.6 times more meat than their counterpart in low-income countries, and whose eating habits are unsustainable. But others question whether the diet is nutritious enough for those in lower-resource settings. Ty Beal, a scientist based in Washington DC with the Global Alliance for Improved Nutrition, has analysed the diet in unpublished calculations and found that it provides 78% of the recommended zinc intake and 86% of calcium for those over 25 years old, and only 55% of the iron requirement for women of reproductive age.



Source: Ref. 6

Despite these critiques, the diet has put environmental concerns front and centre. “Until EAT–*Lancet*, I don’t think it had been at the top of policymakers’ minds that sustainability should be integrated into this global conversation about dietary change,” says Anne Elise Stratton, a food-systems scientist at the University of Michigan in Ann Arbor.

The diet is not a one-size-fits-all recommendation, stresses Marco Springmann, a food scientist at the University of Oxford who was part of the EAT–*Lancet* core modelling team.

Since the report was published, public-health scientists around the world have been studying how to make the diet realistic for people the world over, whether an overweight adult or an under-nourished child.

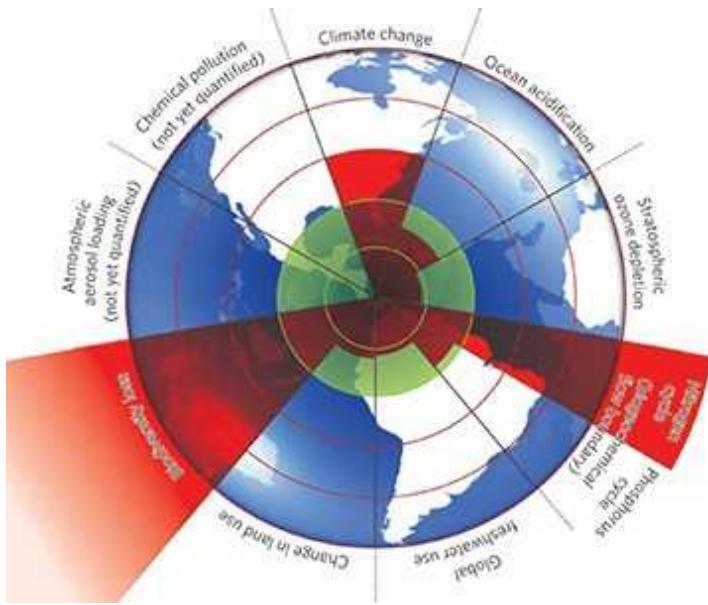
Rich diets

Nutrition researchers know that most consumers do not follow dietary guidelines. So some scientists are exploring ways to convince people to adopt healthy, sustainable diets. In Sweden, Patricia Eustachio Colombo, a nutrition scientist at the Karolinska Institute in Stockholm, and her colleagues are quietly testing a sustainable diet in schools. Their work piggybacks on a social movement that began in Scandinavian countries called the New Nordic Diet, which promotes consumption of traditional, sustainable foods such as seasonal vegetables and free-range meat.

Eustachio Colombo and her colleagues used a computer algorithm to analyse existing school lunches at a primary school with about 2,000 students. The algorithm suggested ways to make them more nutritious and climate-friendly, such as reducing the amount of meat in a typical stew and adding more beans and vegetables. The children and parents were informed that lunches were being improved, but did not know details, Eustachio Colombo says. Most kids did not notice, and there was no more food waste than earlier⁹. The same experiment is now being re-run in 2,800 children.

“School meals are a near unique opportunity to foster sustainable dietary habits. The dietary habits we develop as children, we tend to stick to them into adulthood,” Eustachio Colombo says.

The diet is very different from the EAT–*Lancet* one, she says. It is cheaper and includes more starchy foods such as potatoes, which are a staple of Swedish cuisine. It is also more nutritious and culturally acceptable, she says. “This highlights the importance of tailoring the EAT–*Lancet* diet to the local circumstances in each country or even within countries,” she says.



A safe operating space for humanity

Across the Atlantic, some academics and restaurateurs are trialling the diet in low-income settings. In Baltimore, Maryland, a collaboration between a catering business and a restaurant, both forced to close during the COVID-19 pandemic, started taking donations and providing free meals based on the EAT–*Lancet* diet to families who live in ‘food deserts’ — areas where there is little access to affordable, nutritious food. One meal had salmon cakes with mixed seasonal vegetables, Israeli couscous and creamy pesto sauce.

Researchers at the Johns Hopkins University School of Medicine in Baltimore surveyed 500 people who tried the meals and found that 93% of the 242 people who completed the survey said they either loved or liked it¹⁰. The downside? Each donation-funded meal cost US\$10 — five times the amount currently provided by the US food-stamp programme.

“It’s very clear that if you have a huge shift in diets, you could swing the environment impact for the better, but there’s cultural barriers and economic

barriers to that,” says Griffin.



Workers pack meals at the Alma Cocina Latina restaurant in Baltimore. Credit: Dave Cooper/NYT/Redux/eyevine

Hard to stomach

For researchers exploring future diets in some low- or middle-income nations, one hurdle is finding out what people are eating in the first place. “It’s literally like a black box to me right now,” says Purnima Menon at the International Food Policy Research Institute in Delhi, who has been studying diets in India. The data on what people are eating are a decade old, she says.

Getting that information is crucial, because India ranks 101 out of 116 countries in the Global Hunger Index and has the greatest number of children who are too thin for their height.

Using what’s available, Abhishek Chaudhary, a food-systems scientist at the Indian Institute of Technology Kanpur, who was part of the EAT–*Lancet*

team, and his colleague Vaibhav Krishna at the Swiss Federal Institute of Technology in Zurich used a computer program and local environmental data on water, emissions, land use and phosphorus and nitrogen use to design diets for all of India's states. The algorithm suggested diets that would meet nutritional requirements, cut food-related emissions by 35% and wouldn't stress other environmental resources. But to grow the required amount of food would require 35% more land — which is impractical in the overcrowded nation — or higher yields. And food costs would be 50% higher^{[11](#)}.

Healthy, sustainable diets are expensive elsewhere, too. The dietary diversity advised by EAT–*Lancet* — nuts, fish, eggs, dairy and more — is impossible to access for millions of people, says Iannotti.

In fact, for the average person to eat the diet in 2011 — the most recent data set available on food prices — would have cost a global average of \$2.84 per day, about 1.6 times higher on average than the cost of a basic nutritious meal^{[12](#)}.



Rickshaw drivers in Dhaka, Bangladesh stop for lunch. The cost of a ‘planetary health’ diet is beyond the reach of many in south Asia.Credit: Munir Uz zaman/AFP/Getty

There are other impracticalities. Take restrictions on meat, for instance. In places with nutrient deficiencies and where the diet’s prescribed foods are not available, animal-source products are a crucial source of easily bioavailable nutrients in addition to plants, Iannotti says. In many places in low-income nations, farming systems are small-scale and include both crops and domesticated animals, which can be sold in times of family need, says Jimmy Smith, director-general of the International Livestock Research Institute in Nairobi.

“The farmer in the highlands of Ethiopia doing dairy has three or four animals in his or her backyard, and each of these animals is a member of the family, they have names,” he says.

Menon says that for now, scientists in low- and middle-income regions are more concerned about delivering nutrition than preserving the environment. The FAO has organized a committee to redo the analysis of EAT–*Lancet* to make it more globally inclusive, says Iannotti, who is part of the committee. The global assessment will be published in 2024. “They don’t feel as if it was entirely balanced or holistic in its review of the evidence,” she says. “Let’s go further and make sure we have evidence from around the world.”

The way to find sustainable diets in poor nations is by working closely with communities and farmers, as in Kilifi, scientists say. Clark, having mapped out diet at a global scale using model-based projections, thinks that food-system scientists now need to find the local adjustments and fixes to get people to eat better.

“People working in food sustainability need to go into communities and ask, ‘hey, what’s good for you?’” he says. “And then, given that baseline, how can we start working towards outcomes that those communities are interested in.”

doi: <https://doi.org/10.1038/d41586-021-03565-5>

References

1. 1.
Iannotti, L. L. *et al.* *Pediatrics* **140**, e20163459 (2017).
2. 2.
GBD 2017 Diet Collaborators. *Lancet* **393**, 1958–1972 (2019).
3. 3.
Springmann, M. *et al.* *Lancet Planet. Health* **2**, e451–e461 (2018).
4. 4.
Willett, W. *et al.* *Lancet* **393**, 447–492 (2019).
5. 5.
Clark, M. A. *et al.* *Science* **370**, 705–708 (2020).
6. 6.
Tilman, D. & Clark, M. *Nature* **515**, 518–522 (2014).
7. 7.
Merrigan, K. *et al.* *Science* **350**, 165–166 (2015).
8. 8.
Steffen, W. *et al.* *Science* **347**, 1259855 (2015).
9. 9.

Elinder, L. S., Eustachio Colombo, P., Patterson, E., Parlesak, A. & Lindroos, A. K. *Sustainability* **12**, 8475 (2020).

10. 10.

Semba, R. D., Ramsing, R., Rahman, N. & Bloem, M. *J. Agric. Food Syst. Community Dev.* **10**, 205–213 (2020).

11. 11.

Chaudhary, A. & Krishna, V. *One Earth* **4**, 531–544 (2021).

12. 12.

Hirvonen, K., Bai, Y., Headey, D. & Masters, W. A. *Lancet Glob. Health* **8**, e59–e66 (2020).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03565-5>

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

Books & Arts

- **[When scientists gave 1,000 vulnerable people hepatitis over 30 years](#)** [29 November 2021]

Book Review • What sort of system nurtures a decades-long programme of deliberately infecting children and prisoners with a dangerous disease?

- **[Defining Alzheimer's, and the climate costs of AI: Books in brief](#)** [19 November 2021]

Book Review • Andrew Robinson reviews five of the week's best science picks.

- BOOK REVIEW
- 29 November 2021

When scientists gave 1,000 vulnerable people hepatitis over 30 years

What sort of system nurtures a decades-long programme of deliberately infecting children and prisoners with a dangerous disease?

- [Heidi Ledford](#)



A ward at Willowbrook State School in New York, where some children were given hepatitis.Credit: Eric Aerts

Dangerous Medicine: The Story Behind Human Experiments with Hepatitis *Sydney A. Halpern* Yale Univ. Press (2021)

In 1942, in the grip of the Second World War, the US military faced an existential threat from within. A hepatitis outbreak was suspected to have infected hundreds of thousands of personnel.

There were no animal or cell-culture models for studying the viral liver disease. Desperate to find the source of the outbreak and learn how to contain it, the military joined forces with biomedical researchers, including some from the University of Pennsylvania in Philadelphia and Yale University in New Haven, Connecticut, to launch human experiments that continued for decades after the war.

During the 30-year programme — meticulously chronicled in *Dangerous Medicine* — researchers infected more than 1,000 people, including over 150 children, with viruses that cause hepatitis. The people enrolled were prison inmates, disabled children, people with severe mental illnesses, and conscientious objectors performing community service in lieu of fighting, relates historical sociologist Sydney Halpern. Owing to biases in the US prison and psychiatric-hospital populations, a disproportionate number were Black. The long-term consequences will never be fully reckoned: although rarely fatal in the short term, hepatitis can lead to chronic liver disease and cancer years after the initial infection.



How asylums became the crucible of genetics

These days, horrifying stories of human medical experiments in the mid-twentieth century are well-trodden territory, the most famous being the studies in Tuskegee, Alabama, that withheld treatment from hundreds of Black men with syphilis for decades, starting in the 1930s. “Research abuse narratives are embedded in our collective memory,” Halpern writes. “They serve to affirm the moral priorities we proudly hold.”

But *Dangerous Medicine* does more than add another gruesome narrative to the canon. It offers a thorough exploration of the social, military and scientific context that nurtured what would now be considered repugnant and unethical medical conduct. By avoiding a simple story of a single rogue investigator or team and instead focusing on the system that gave rise to abuse, Halpern amplifies the narrative that she has painstakingly pieced together from hundreds of historical documents.

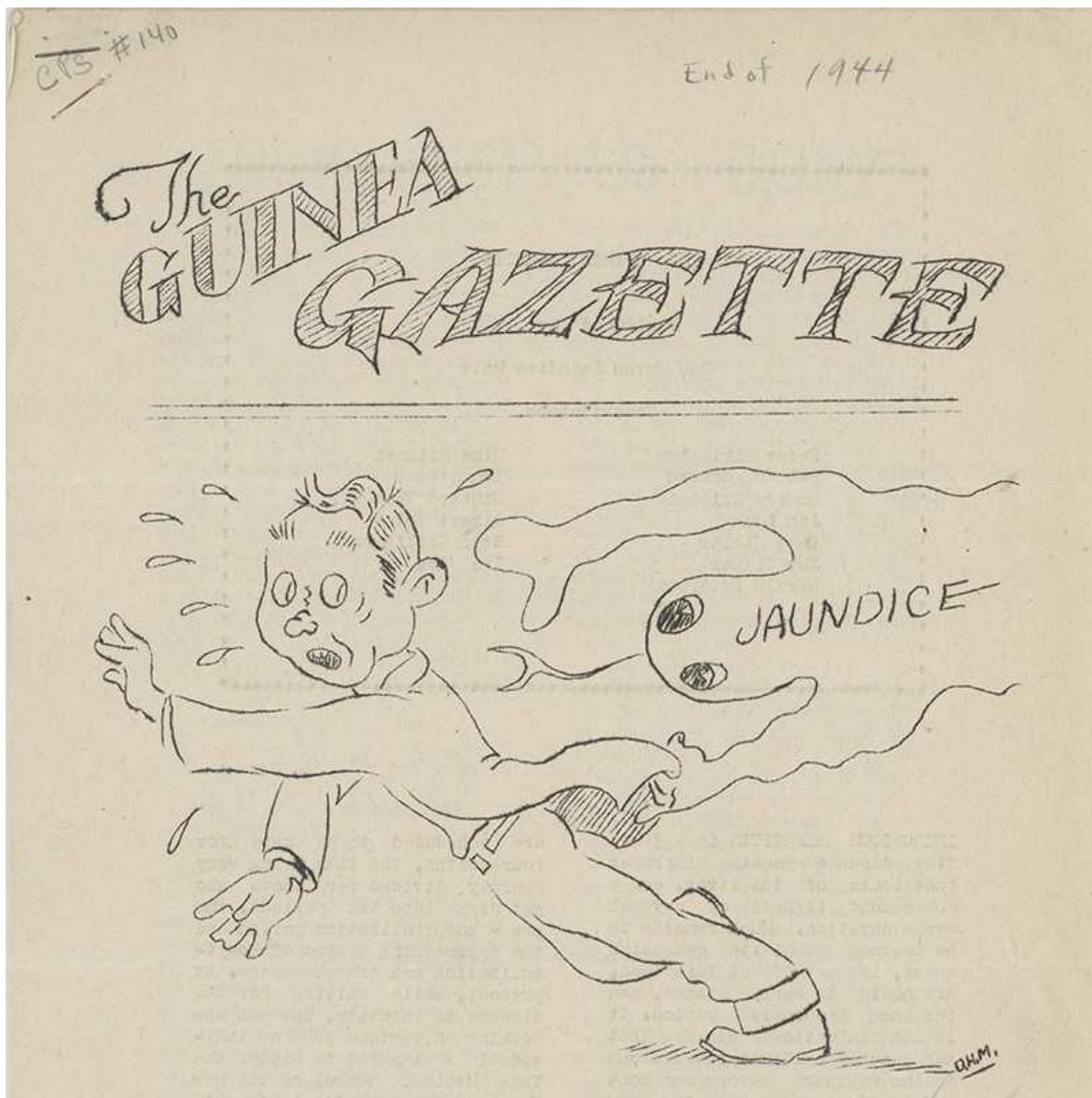
System failure

The hepatitis experiments began in 1942 with the outbreak of the disease among US soldiers and military personnel, which was, researchers eventually determined, caused by a contaminated batch of yellow fever vaccine. Launched under the aegis of the war effort, the studies ended only

in 1972, when public and professional sentiment towards experiments on vulnerable populations shifted.

Researchers studied a wide range of hepatitis biology, distinguishing between hepatitis A, which is transmitted by contaminated food, and hepatitis B, often spread by contaminated blood products. They looked for ways to inactivate the hepatitis B virus in blood supplies, and tested treatments and means of prevention.

Some of the studies involved deliberately exposing people to infected material, either by injection or through ingestion of “milkshakes” containing hepatitis virus in the form of stool samples mixed with chocolate milk. At least four people died from the disease in the course of these experiments. But with no long-term follow-up — monitoring stopped when the individual experiments ended — there are no data on how many became disabled or died years later as a result of their infection.

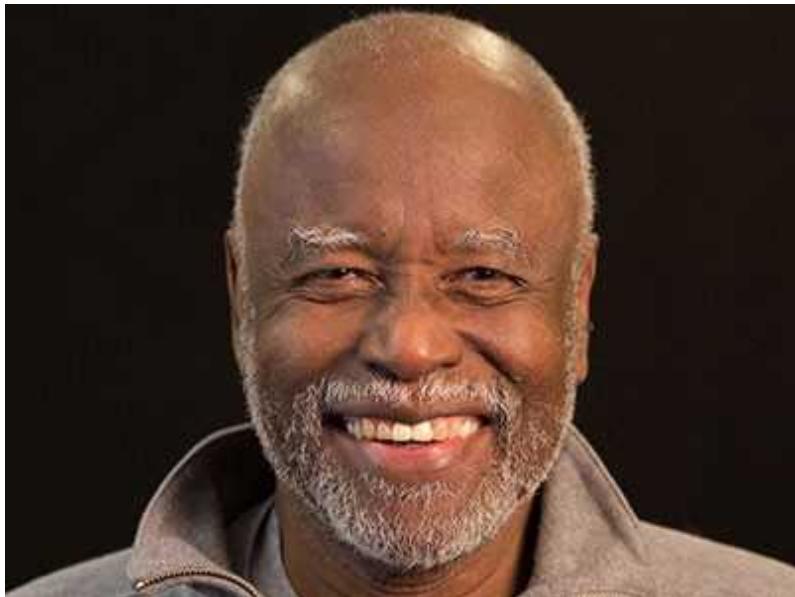


A 1945 newsletter for the Civilian Public Service camp in New Haven, Connecticut, where conscientious objectors were enrolled in hepatitis studies as part of the war effort. Credit: David Hileman Miller/Swarthmore College Peace Collection

For decades, the researchers justified their work to colleagues and the press, reframing it to meet the ethos of the times. Initially, it was portrayed as a necessary sacrifice to support the troops. Later, experiments on people in prison were pitched as paths to rehabilitation through service to society. Studies on mentally ill people were an extension of 'fever therapy' — the idea among some scientists at the time that high body temperature caused by

infectious diseases such as malaria and, perhaps, hepatitis might improve some psychiatric conditions. And experiments on disabled children were rationalized as an attempt to generate immunity against a disease that was already frequent in the crowded and unhygienic institution in which they were housed.

Notwithstanding her attention to context, Halpern does not let individuals off the hook. Some study leaders knew how their work might horrify the public, and tried to control how they were portrayed in the press — first suppressing coverage, and later encouraging narratives that painted participants as heroes. Some journal editors not only published scientific findings from the hepatitis experiments, but also wrote editorials praising the work.



[Bill Carter Jenkins \(1945–2019\), Tuskegee whistle-blower](#)

And, Halpern points out, although the long-term consequences of hepatitis were not fully understood at the time of the experiments, there were signs as early as the 1940s, and the researchers could have acknowledged them. Eventually, by the 1970s and 1980s, epidemiological studies had shown that carriers of hepatitis B were more likely to develop cirrhosis and liver cancer than those who were not carriers.

Particularly crushing is the naivety about how hepatitis affects children. The immediate symptoms are not as severe in children, so scientists argued that infecting young people would give them immunity that would protect them when they grew older and more vulnerable to severe infections. In fact, children with hepatitis B are much more likely than infected adults to become lifelong carriers, and to experience long-term consequences.

Haunting reminder

There was a time when we could have casually looked down our noses at mid-twentieth-century ignorance about infectious diseases. But with the world still in the throes of a coronavirus pandemic, I was struck by the parallels. Witness how efforts have been focused on the acute impacts of disease (hospitalization, death) without much thought to long-term consequences (disability). Or think of how those with the least agency — children, people in prison, people with severe mental illnesses — have been put at risk by those with the most power.

Formally, our approach to medical ethics has improved since the 1940s, but Halpern reminds us that many clinical trials in healthy individuals still rely on vulnerable populations. Some people move from one to the next in search of food, housing or remuneration in exchange for their participation. People in regions with poor access to health care sometimes have to enrol to get basic medical treatment. And in the United States, there is still no requirement to provide compensation for long-term disability that might arise from participation in clinical trials.

Halpern has created a haunting narrative that forces the reader to confront our modern social and scientific frame of reference. Long after the book is finished, the question remains: what research abuses are we justifying to ourselves today?

Nature **600**, 27-28 (2021)

doi: <https://doi.org/10.1038/d41586-021-03571-7>

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

- BOOK REVIEW
- 19 November 2021

Defining Alzheimer's, and the climate costs of AI: Books in brief

Andrew Robinson reviews five of the week's best science picks.

- [Andrew Robinson](#) 0

How Not to Study a Disease

*The Story
of Alzheimer's*



Karl Herrup

How Not to Study a Disease

Karl Herrup *MIT Press* (2021)

Neurobiologist Karl Herrup argues that an acute problem for research into Alzheimer's disease is the lack of a definition. While writing his important, accessible study, he found that no two experts described the degenerative condition in the same way. Alois Alzheimer's original, post-mortem diagnosis of a single person in 1906 attributed her mental confusion to plaques and tangles in her brain tissue. For Herrup, progress requires a definition "based on the symptoms of the patient, not on the deposits in the patient's brain".

'GROUNDBREAKING'
RUTGER BREGMAN

THE DAWN OF EVERYTHING

'AN INTELLECTUAL FEAST'
NASSIM NICHOLAS TALEB

A NEW HISTORY OF HUMANITY

'THE RADICAL REVISION OF EVERYTHING'
REBECCA SOLNIT

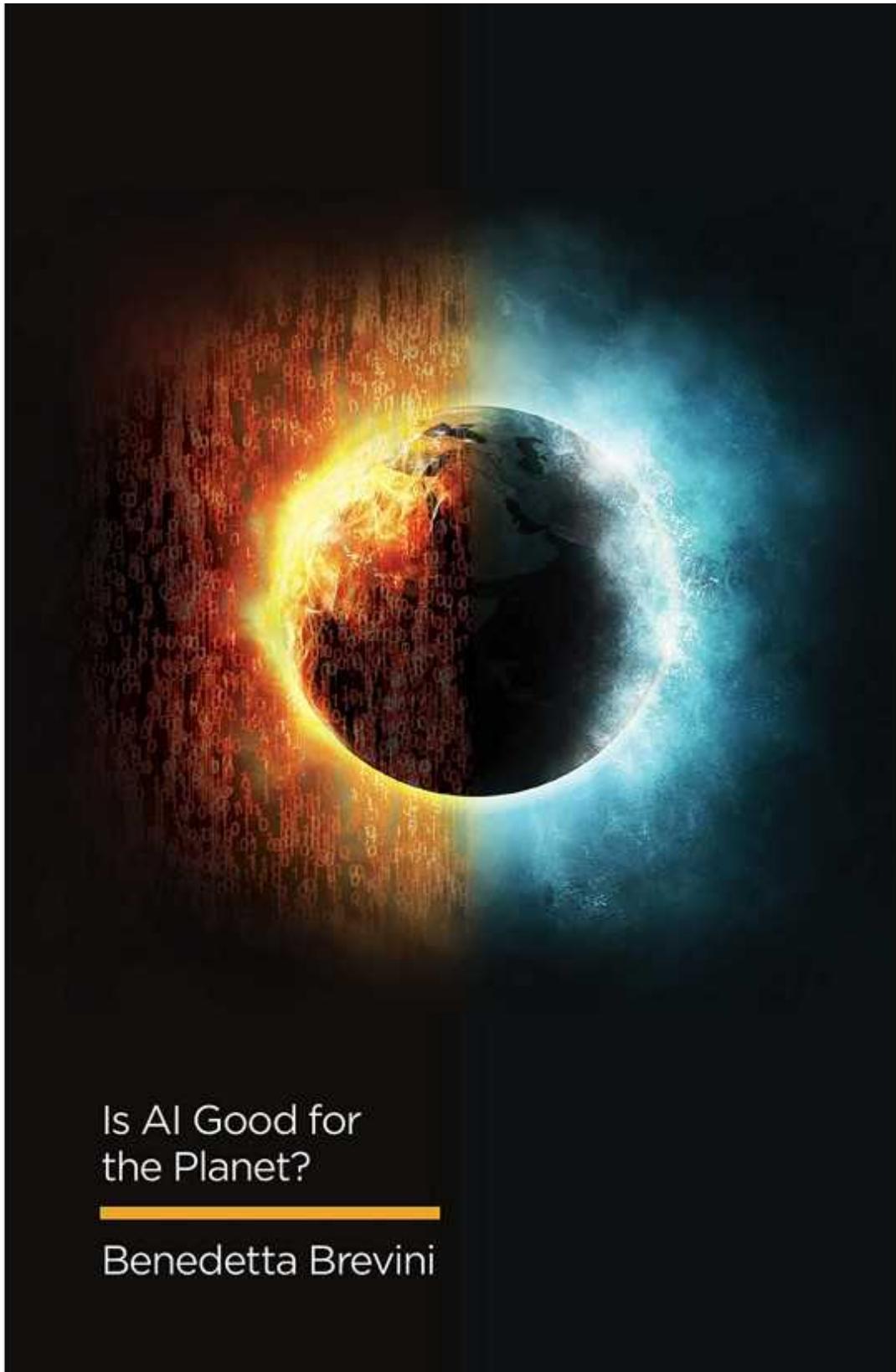
DAVID GRAEBER AND
DAVID WENGROW

allen lane

The Dawn of Everything

David Graeber & David Wengrow *Allen Lane* (2021)

Among the great variety of cultures discussed and compared by anthropologist David Graeber and archaeologist David Wengrow is the Indus civilization, which flourished for seven centuries, starting about 2500 bc, then vanished. It had a complex economy, maritime trade and a chief city with advanced plumbing, but has yielded no sign of palaces, temples, rulers or warriors. This subtle but revolutionary study challenges the consensus that hunter-gatherer egalitarianism inevitably evolved, through cities, into hierarchical, bureaucratic states.



Is AI Good for the Planet?

Benedetta Brevini

Is AI Good for the Planet?

Benedetta Brevini *Polity* (2021)

For all their benefits, artificial intelligence (AI) technologies are damaging Earth, argues journalist Benedetta Brevini in a short but powerful assessment. The effects are most obvious in the ever-increasing energy needed for data centres, but also in AI's boosting of "uberconsumerism", with its proliferation of products, packaging waste and built-in obsolescence — not to mention increasingly efficient fossil-fuel extraction. A 2018 Amazon Web Services report, Brevini notes, was called 'Predicting the next oil field in seconds with machine learning'.

Dennis
Duncan

Index, A History of the

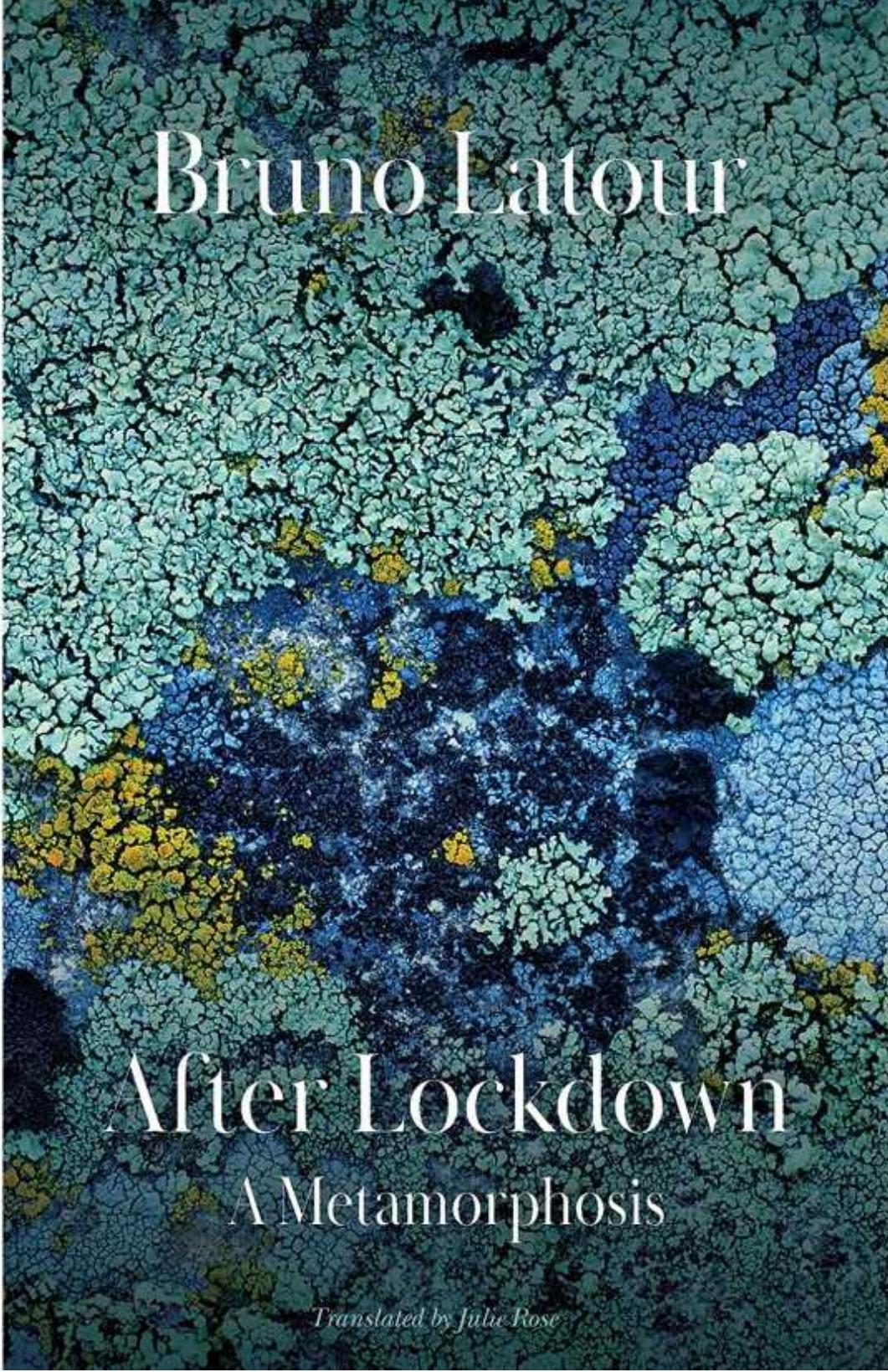
A Bookish
Adventure

allen lane

Index, A History of the

Dennis Duncan *Allen Lane* (2021)

The first web page (launched in 1991) was a subject index. When we Google something, we search not the entire web, but Google's index of it. These observations typify literary historian Dennis Duncan's wide-ranging and entertaining history, beginning in thirteenth-century European monasteries and universities. Despite the existence of indexing software, satisfactory results still require flesh-and-blood indexers. He proves the point with a computer-generated index of his book for 'A', with the absurd final entry: "amusement, 180/ mere, 198".



Bruno Latour

After Lockdown
A Metamorphosis

Translated by Julie Rose

After Lockdown

Bruno Latour *Polity* (2021)

“Let’s celebrate the experience of a pandemic,” writes philosopher Bruno Latour in his brief but dense meditation on COVID-19, inspired by Franz Kafka’s 1915 novella *Metamorphosis*. The pandemic, he says, has made us realize — through social distancing and mask wearing — “to what extent the distinct individual was an illusion”. He is not just being ironic: he argues that a new “global awareness” triggered by lockdowns might help to unite us in facing the even more demanding threat of climate change.

Nature **600**, 28 (2021)

doi: <https://doi.org/10.1038/d41586-021-03484-5>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03484-5>

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

Opinion

- **[The United States needs a department of technology and science policy](#)** [30 November 2021]
Comment • Two former NIH directors call for a cabinet-level department to formulate long-range science policy and oversee technology development.
- **[Africa: tackle HIV and COVID-19 together](#)** [01 December 2021]
Comment • Failure to get COVID-19 vaccines to nations with high rates of uncontrolled advanced HIV puts people living with that virus at even greater risk, and could drive the emergence of coronavirus variants.
- **[Collaborate equitably in ancient DNA research and beyond](#)**
[30 November 2021]
Correspondence •
- **[Ancient-DNA researchers write their own rules](#)** [30 November 2021]
Correspondence •
- **[Climate researchers: consider standing for office — I did](#)**
[30 November 2021]
Correspondence •
- **[Animal experiments: EU is pushing to find substitutes fast](#)**
[30 November 2021]
Correspondence •

- COMMENT
- 30 November 2021

The United States needs a department of technology and science policy

Two former NIH directors call for a cabinet-level department to formulate long-range science policy and oversee technology development.

- [Harold Varmus](#) ⁰ &
- [Elias Zerhouni](#) ¹



The White House in Washington DC. It includes the Office of Science and Technology Policy, which helps to coordinate federal science policy.Credit: Getty

Unlike most other advanced economies, the United States has typically relied on a diversity of decentralized federal science agencies to design and conduct the government's research programmes. With funding decisions made by several congressional committees and with minimal oversight from the Office of Science and Technology Policy (OSTP) in the White House, this approach has been enormously successful.

As former directors of one of those science agencies, the National Institutes of Health (NIH), we remain fans of our distributed federal system. It encourages a broad base of support for the sciences in Congress, and it protects basic research from partisan politics and from narrowly managed planning processes. It has fostered US leadership in scientific discovery since the Second World War.

Yet we have come to realize that the country needs an additional mechanism to promote national and global goals: a Department of Technology and Science Policy (DTSP), the leader of which would join other department heads and administration leaders to form the president's Cabinet.

What should this department do, and why do we propose it now? Over the past decade or more, the US government has been required to do more in science and technology than simply foster discovery and then expect the private sector to apply those discoveries to practical purposes — as Vannevar Bush, who organized the US government's science programmes during the Second World War, advised in his foundational guidance more than 75 years ago¹. Expanding demands have been precipitated by pandemics, climate change, transitions in sources of energy, loss of biological diversity, and other natural and social crises. Furthermore, the growing strength of the scientific enterprise in large parts of Asia and in a mostly unified Europe has sharpened the competition for global leadership, especially in the commercial sectors that are dependent on science and technology.

Governments in nearly all highly developed economies have responded to such pressures by establishing ministries to pursue strategic research goals and to coordinate long-term activities across sectors (see 'Science governance in

research-intensive regions'). China's Ministry of Science and Technology, for example, has designed and funded the Made in China 2025 plan. This will invest in major technologies such as semiconductors, biotechnology, computing and other strategic industries, with the goal of China achieving 70% self-sufficiency by 2025 and 100% by 2049 (see go.nature.com/2zeknx). Similarly, the European Union's 95.5-billion (US\$108-billion) Horizon Europe research programme for 2021–27 is intended to improve industrial competitiveness in all technologies that are crucial to a modern economy (see go.nature.com/2zdjpne).

The United States now seems to be the only research-intensive country without such an instrument, limiting its ability to compete and, equally importantly, to collaborate with other nations.

Science governance in research-intensive regions

Country	Governing agency or programme	Stated responsibilities	Weblink
China	Ministry of Science and Technology	National strategies for innovation-driven development (including Made in China 2025)	http://en.most.gov.cn

Country	Governing agency or programme	Stated responsibilities	Weblink
European Union	Horizon Europe	Long-range research and innovation programme to boost EU competitiveness	https://go.nature.com/2zdjpne
Japan	Ministry of Economy, Trade and Industry	Development of Japan's economy and industry	https://www.meti.go.jp/english
Germany	Federal Ministry of Education and Research	Education policy and research funding, including for “technological sovereignty and innovation”	https://www.bmbf.de/bmbf/en
France	Ministry of Higher Education, Research and Innovation	Strategic research and education activities	https://go.nature.com/3tysjcv

Country	Governing agency or programme	Stated responsibilities	Weblink
Canada	Innovation, Science and Economic Development	Developing a “competitive and knowledge based Canadian economy”	https://go.nature.com/3t7ssqp
India	Department of Science and Technology	Major industrial and science policies and their implementation	https://dst.gov.in
United Kingdom	Office for Science and Technology Strategy (proposed)	Strategic directions and funding for UK science and technology	https://go.nature.com/3hm4tcs
South Korea	Ministry of Science and ICT	All national science and technology strategies and programmes	https://english.mst.go.kr/eng/index.do

Past efforts to consolidate all federal science programmes under one roof have been repeatedly — and, in our view, correctly — rebuffed by the US scientific community, largely because of concerns that such centralization would damage the fertility, flexibility and fiscal support of the nation’s scientific work (see, for example, ref. 2). Still, proposals continue to appear, recommending

consolidation, streamlining or modernization of the governmental apparatus that oversees science and technology^{3,4}.

The department we advocate here, however, would not change the governance or funding of most of the existing federal research agencies. Instead, it would provide a much-needed venue for planning long-term science policies and for developing multidisciplinary, broadly useful technologies.

Presidential support

The administration of President Joe Biden has already signalled its support in several ways for a larger role for science and technology, both to counter global and national threats and to make the country more competitive economically. For example, Biden is the first US president to place the OSTP director in the Cabinet. He has endorsed existing legislative proposals for a drastic expansion of research investment by creating a technology-focused directorate at the National Science Foundation (NSF; go.nature.com/3dktxah). He has proposed the establishment of two further Advanced Research Project Agencies (ARPAs) for mission-driven work in health (ARPA-H; go.nature.com/3cfqjyx) and climate (ARPA-C; go.nature.com/3jwdrt8). And his administration has unveiled extensive, methodologically complex plans for pandemic preparedness that would be managed by the OSTP⁵.



[Memo for President Biden: Five steps to getting more from science](#)

These proposals share defining characteristics that differentiate them from the goals of existing agencies. They are designed to combine public- and private-sector strengths across disciplines, and are intended to provide tangible societal and economic benefits in response to well-documented needs. These objectives require fundamentally different management styles and greater collaboration among existing government agencies, academia and industry, both within and between countries.

Public–private partnerships will be crucial for the success of the proposed technological programmes. Private industry, understandably, requires a sufficient probability of returns on its investments to assume bold but uncertain technological risks. Large companies that once led such advances in their own research laboratories — such as what is now Nokia Bell Labs, headquartered in Murray Hill, New Jersey, or the former Roche Institute of Molecular Biology in Nutley, New Jersey — have largely abandoned long-term efforts of this type in favour of more immediate research and development goals. Academia lacks the means to assume such responsibilities, given current funding mechanisms. US government agencies have not substantially modified this dynamic, despite their relatively large budgets compared with those of similar agencies in competitor nations.

A DTSP could devise new mechanisms of support that bring governmental, academic and commercial labs together to pursue common strategic goals, such as overcoming antibiotic resistance, combating food and water insecurity and reducing greenhouse-gas emissions.

What goes where?

The science community has voiced general support for the initiatives proposed by the Biden administration. But there is less agreement about where these expanded activities should sit in the government’s Executive Branch — as a part of the president’s Executive Office; as a component of a department or of an agency (such as the NIH) within a department; or as a free-standing agency, such as the NSF.

Some scientists have expressed concerns that the NSF’s conventional allegiance to fundamental discovery could be undermined if a large technology directorate were added to it⁶. The plan to host ARPA-H at the NIH⁷ raises

questions about how the new agency would be distinguished from existing NIH institutes and centres with respect to operations, governance, culture and goals. Several government departments might be interested in acquiring ARPA-C, but a new climate agency could be limited in scope if it were assigned to any single existing department. Furthermore, the overburdened OSTP, which has only a modest annual budget (of about \$5 million to \$6 million) and a relatively small staff, might be ill-suited to managing a programme as large and complex as the pandemic preparedness initiative.



Sociologist Alondra Nelson (left) is deputy director for science and society at the Office of Science and Technology Policy.Credit: Kevin Lamarque/Reuters

Our view is that, instead of dispersing the proposed programmes between multiple existing agencies that have widely differing mandates, they should be united under a different kind of governance. We envision a DTSP that is run by a directorate resembling the current OSTP. The secretary of the new department would ideally have broad experience in both technology development and science policy.

Its operational components would include the proposed technology directorate (but presumably not the existing NSF directorates). The proposed department would also include ARPA-H and ARPA-C. But the existing Defense Advanced Research Projects Agency and energy-focused ARPA-E would remain in the Department of Defense and Department of Energy, respectively, where they have pursued goals that are generally unique to those bodies. The DTSP would oversee those components of the pandemic preparedness initiative that are responsible for technology development, and would perhaps administer a few existing agencies that have broad, multidisciplinary objectives, such as the National Institute of Standards and Technology.



The rise of ‘ARPA-everything’ and what it means for science

The creation of a DTSP would strengthen the role of the social sciences, especially in enabling technology to be deployed for the broadest benefit. This is a facet of American life that has engendered much disappointment and controversy during the COVID-19 pandemic, and in efforts to confront climate change. Expanded federal support for multiple approaches united under one departmental roof would be consistent with the concept of disciplinary “convergence”⁸. Methods from a variety of fields are crucial to solving many of today’s most difficult problems, as demonstrated by the rapidly expanding roles of engineering and physics and of data and social sciences in medicine and health care⁸.

In our vision, most of the existing US science agencies that have well-defined commitments to a specific national goal — defence, energy, agriculture, space exploration and health — would remain in their current departments or as free-standing agencies.

Solid foundations

We are not proposing that the United States' successful approach to fundamental science be dismantled, reorganized or subjected to any degree of command and control. Nor do we aim to centralize the funding or planning of US science in any one structure, to introduce redundancies or to add another layer of bureaucracy to existing agencies that have thrived under relatively autonomous leadership.

Could the proposed DTSP make the US science system more vulnerable to politically motivated pressures or arbitrary budgetary reductions in an administration less supportive of science than the current one? In our view, this is very unlikely. Among the advantages of our proposal would be the DTSP's relative freedom from political considerations that often influence departments with more-focused short-term goals and scope. Political influence could be further diminished by providing term appointments of roughly six years for leaders of each of the department's science agencies. Most of the current science agencies would remain in place in the Executive Branch, and are likely to continue to receive bipartisan fiscal support from Congress for their scientific work.

We know that the changes we propose are difficult to achieve in any domain of government. Congress is understandably cautious about creating departments, and long-term changes in the oversight of important federal activities should be undertaken only after thorough evaluation and debate. For these reasons, we urge broad discussion of the ideas we advance here. This could be done by the President's Council of Advisors on Science and Technology; by hearings organized by the House Committee on Science, Space and Technology and the Senate Committee on Commerce, Science, and Transportation; and by the research and public-advocacy communities, such as the National Academies of Sciences, Engineering, and Medicine, and others.

In the long run, we must harness research in an effective way if we are to confront our most dire threats. Achieving that goal could depend on the amount of attention given now to organizing the scientific and technological components of government. This needs to be done in a way that matches the requirements of a world that is both more competitive and in greater need of constructive collaboration.

Nature **600**, 30-32 (2021)

doi: <https://doi.org/10.1038/d41586-021-03543-x>

References

1. 1.

Bush, V. *Science — The Endless Frontier* (US Government Printing Office, 1945).

2. 2.

US National Research Council. *Allocating Federal Funds for Science and Technology* (National Academies Press, 1995).

3. 3.

Kline, R. ‘The U.S. needs a federal department of science and technology.’ *Sci. Am.* (20 February 2021).

4. 4.

Flagg, M. & Garg, A. *Issues Sci. Technol.* **Fall**, 51–59 (2021).

5. 5.

Stolberg, S. G. ‘Biden’s 10-year pandemic preparedness plan includes a full-time ‘Mission Control’ office.’ *The New York Times* (3 September 2021).

6. 6.

Rao, R. ‘The Endless Frontier Act Could Drastically Change NSF.’ *IEEE Spectrum* (1 June 2021).

7. 7.

Collins, F. S., Schwetz, T. A., Tabak, L. A. & Lander, E. S. *Science* **373**, 165–167 (2021).

8. 8.

US National Research Council. *Convergence: Facilitating Transdisciplinary Integration of Life Sciences, Physical Sciences, Engineering, and Beyond* (National Academies Press, 2014).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03543-x>

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

- COMMENT
- 01 December 2021

Africa: tackle HIV and COVID-19 together

Failure to get COVID-19 vaccines to nations with high rates of uncontrolled advanced HIV puts people living with that virus at even greater risk, and could drive the emergence of coronavirus variants.

- [Nokukhanya Msomi](#) ⁰,
- [Richard Lessells](#) ¹,
- [Koleka Mlisana](#) ² &
- [Tulio de Oliveira](#) ³



People in Mombasa, Kenya, hold empty boxes of antiretrovirals in an April protest over drug shortages caused by the COVID-19 pandemic. Credit: AP/Shutterstock

As scientists and clinical practitioners working in hospitals and laboratories across South Africa, we've seen at first hand the devastation that COVID-19 has brought to the country. Many of our colleagues (including leading scientists and health-care professionals) have died. We have also seen how much harder it has become for people with HIV to get treatment and monitoring, which has meant more people being admitted to hospital with serious illnesses such as tuberculosis and cryptococcal meningitis.

As the world shifts gear in the COVID-19 pandemic — from tackling the immediate crisis to a long-term public-health response — we urge governments, health ministries, researchers and other stakeholders worldwide to devote more of their resources and attention to the interactions between COVID-19 and HIV.

Various studies from before the rollout of COVID-19 vaccines show that people with HIV are 30–50% more likely to die from COVID-19. What's more, a failure to tackle the pandemic with sufficient urgency in countries with high rates of uncontrolled advanced HIV could lead to the emergence of variants of the coronavirus SARS-CoV-2 that spread more easily between people or render the vaccines less effective. Conversely, both diseases could be curbed more effectively if they are tackled simultaneously, with public-health responses strengthened by the lessons learnt from both.

Risk analysis

In 2020, the World Health Organization (WHO) created a platform for standardized clinical data obtained from people with suspected or confirmed COVID-19. A preliminary analysis of these data from 37 countries (most of which come from South Africa) revealed that, controlling for other factors, people with HIV were 30% more likely to die from COVID-19 among those hospitalized with the disease¹.

In a more detailed analysis² of hospitalized cases in South Africa, only people with HIV at an advanced stage were more likely to die from COVID-19. People in this group are severely immunocompromised: individuals have a blood count of fewer than 200 CD4 T cells per microlitre, compared with 500–1,500 cells per microlitre in people without HIV. (CD4 T cells are white

blood cells crucial for fighting infection.) In a population-based analysis of around 3.5 million people from the Western Cape Province of South Africa (around 500,000 of whom were HIV-positive), people with HIV were twice as likely to die from COVID-19³. But for those with uncontrolled or advanced HIV, the risk of dying from COVID-19 was almost four times higher³.

Other reports over the past year from various countries indicate that people with weakened immune systems can be infected with SARS-CoV-2 for several weeks or months. (By contrast, those who are otherwise healthy take, on average, around two weeks to clear an infection.) Such prolonged infections have mainly been documented in people receiving cancer chemotherapy and other immunosuppressive agents. But they have also been seen in people with advanced uncontrolled HIV.



A person living with HIV (right) talks to a worker at a mobile clinic in Johannesburg, South Africa, during the COVID-19 pandemic.Credit: Bram

Janssen/AP/Shutterstock

Some researchers have proposed that prolonged infection could explain the emergence of SARS-CoV-2 ‘variants of concern’ (VOCs)⁴. Such variants are more transmissible than the originally identified SARS-CoV-2, or are able to partially evade the immune responses that arise from infection or vaccination. We described a case from South Africa in June in which SARS-CoV-2 persisted in a person with advanced uncontrolled HIV for more than six months⁵. Repeated genomic sequencing revealed significant step changes in the evolution of SARS-CoV-2, and some of the variants emerging had similar mutations to those observed in three of the VOCs.

In short, it is plausible that a prolonged COVID-19 infection in someone who is immunocompromised could lead to the emergence of a variant that is more transmissible even than the Delta variant, or that renders current COVID-19 vaccines less effective.

HIV crisis

Sub-Saharan Africa is home to two out of every three people living with HIV worldwide. Last year, 19.5 million people in sub-Saharan Africa received antiretroviral therapy, a cocktail of drugs that stop HIV from making copies of itself and so lessen the damage the virus causes to the immune system. This is up from 12.1 million people 5 years ago. The increase is largely thanks to a stronger commitment by African governments to tackle the HIV crisis; agreements with antiretroviral-therapy producers to reduce prices; and more funds provided by organizations such as the US President’s Emergency Plan for AIDS Relief (PEPFAR) and the Global Fund to Fight AIDS, Tuberculosis and Malaria.



[COVID vaccines to reach poorest countries in 2023 — despite recent pledges](#)

Nevertheless, 8 million people with HIV in sub-Saharan Africa (around 21% of the 37.7 million people affected worldwide) are not receiving effective antiretroviral therapy. Some people have difficulty accessing clinics to obtain treatment. Others might not have been diagnosed. Yet others might be unaware of the testing and treatment options available, or resist contact with health-care providers because they fear stigmatization, among other reasons.

What's more, the COVID-19 pandemic is continuing to cause significant disruption to HIV treatment and prevention programmes.

Part of the problem is strained health-care systems. [More than 1,300 health-care workers died from COVID-19](#) between March 2020 and August 2021 in South Africa alone. Thousands more could leave the profession because of the impact of COVID-19 on their physical and mental health, according to [a report by the International Council of Nurses](#). Overseas funding for HIV has also declined. For instance, the UK government slashed international aid spending this year, leading to funding cuts of more than 80% for key organizations involved in the global response to HIV, including the [Joint United Nations Programme on HIV/AIDS \(UNAIDS\)](#).

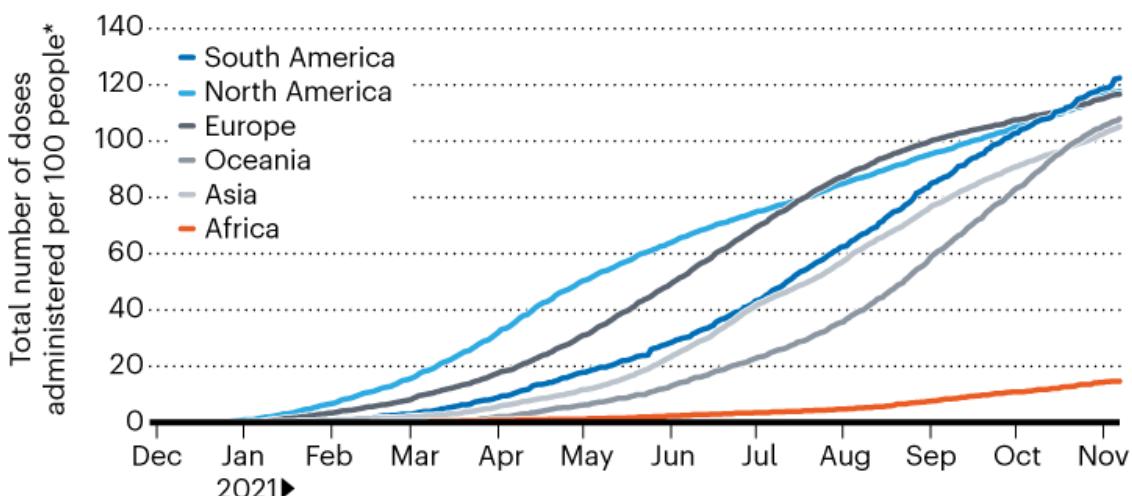
Key testing and prevention services declined in 13 countries supported by the Global Fund between 2019 and 2020. Of these, 12 are in Africa. Compared with 2019, HIV testing, for instance, dropped by 22%; voluntary

medical male circumcision, which reduces female-to-male sexual transmission of HIV by 60%, dropped by 27%; and the number of mothers receiving medicine to prevent the transmission of HIV to their babies dropped by 4.5%⁶.

Meanwhile, Africa is being completely left behind as COVID-19 diagnostics, therapeutics and vaccines are deployed throughout much of the rest of the world. By mid-November, more than 40% of people globally had been fully vaccinated. In Africa, it was less than 7% (see ‘Left behind’). In Africa, as elsewhere, older age groups are being prioritized for vaccination, as is appropriate. But this means further delays to vaccinating Africa’s main immunocompromised population. Around 80% of people with HIV across the continent are under 50 years old⁷.

LEFT BEHIND

On many continents, people who have been fully vaccinated are now receiving an additional dose to boost immunity. In Africa, most people have not had a single dose; less than 7% are fully vaccinated.



*For vaccines that require multiple doses, each individual dose is counted. Because the same person might receive more than one dose, the number of doses per 100 people can be higher than 100.

©nature

Source: Our World In Data

The initial aim of COVAX, an initiative launched in April 2020 by various groups, including the WHO, to distribute vaccines to low- and middle-income countries, was to make two billion vaccine doses available by the

end of 2021. By mid-November, in large part [because of donor countries' failure to deliver on their commitments](#), only [507 million doses had been delivered](#).

All these factors together make sub-Saharan Africa especially vulnerable to continued and potentially worsening devastation from COVID-19 and HIV. They will also negatively affect the rest of Africa and the wider world.

Two birds, one stone

We urge stakeholders to commit to four goals.

Vaccinate Africa. For people living with HIV, protection from COVID-19 requires a renewed commitment to vaccine equity between countries. This must come from high-income nations and multilateral initiatives such as COVAX⁸. Over and above the ethical arguments to address vaccine nationalism and reduce deaths globally, the available data strongly indicate that vaccinating people in Africa will help to reduce transmission rates globally, limit the emergence of new variants and accelerate global control of the pandemic⁹.

Prioritize research. Several gaps in knowledge around the interaction of COVID-19 and HIV urgently need addressing to inform the long-term public-health response.

More clinical and basic studies could reveal whether prolonged SARS-CoV-2 infections are routine in people with advanced uncontrolled HIV or extremely rare — and whether such infections still occur after people have been vaccinated. Studying the interaction between COVID-19 and HIV must not detract from promptly providing individuals with proper care — in particular, antiretroviral therapies, or new types of antiretroviral treatment if previous ones have not worked.



Victories against AIDS have lessons for COVID-19

Although COVID-19 vaccines have been shown to be safe in people living with HIV, there are some uncertainties when it comes to COVID-19 vaccination strategies.

Overall, people living with HIV represent less than 1% of the participants in the phase III COVID-19 vaccine trials conducted by Pfizer–BioNTech, Moderna, AstraZeneca and Johnson & Johnson. (These have all reported safety and efficacy results, but participant follow-up is ongoing.) In fact, our own investigation found that, for some of the vaccines now being rolled out in Africa (such as Sinopharm, Sinovac and Sputnik), people living with HIV were excluded from the phase III trials.

Individuals on antiretroviral therapy who have well-controlled HIV seem to have immune responses to COVID-19 vaccination that are similar to those of people without HIV^{[10,11](#)}. But several observational studies suggest that immune responses to COVID-19 vaccination might be blunted in people with low CD4 T-cell counts^{[12](#)}. This is consistent with what we know about how the immune systems of people living with HIV respond to other vaccines^{[13](#)}.

Clinical-trial data suggest that an additional COVID-19 vaccine dose might enhance the immune response in other immunocompromised groups^{[14](#)}.

These include people who have had an organ transplant and are receiving treatment to suppress their immune systems.

International guidelines, such as those from the [British HIV Association](#) and the [US Centers for Disease Control and Prevention](#), already recommend that health-care providers give an additional dose to people with advanced uncontrolled HIV. But governments of low- and middle-income countries might well demand stronger scientific evidence before they fund the distribution of additional doses to immunocompromised populations. Also, it is not yet clear which subgroups of people living with HIV might benefit, or what the optimal timing of the additional vaccine dose might be, particularly in people starting or restarting antiretroviral therapy.



[African nations missing from coronavirus trials](#)

As vaccines are rolled out in countries that have high rates of HIV, governments, with leadership from organizations such as the WHO and the Africa Centres for Disease Control and Prevention, must support surveillance and monitoring systems for evaluating the effectiveness of COVID-19 vaccines in people living with HIV.

Maintain HIV services. Governments, departments of health and organizations such as PEPFAR and the Global Fund need to take steps now

to ensure that the progress made over the past few years to control the HIV pandemic does not go into reverse.

Across sub-Saharan Africa, health-care providers have used various approaches to try to protect essential health services from disruptions caused by the COVID-19 pandemic. Examples in HIV programmes include providing people with several months' worth of antiretroviral therapy instead of one- or two-months' supply; providing antiretroviral therapies or ways to check CD4 T-cell levels at people's homes or at community centres; and online appointments with physicians¹⁵. Care providers and others can use the experiences of the past two years to map out how best to scale up effective strategies in the future.

In the longer term, stakeholders could also apply some of the innovations in the global COVID-19 response to HIV. Extraordinary achievements, for instance, in the tracking of cases and deaths, in genomic surveillance and in data sharing have demonstrated the degree to which 'knowing your epidemic' in near-real time can help to guide a public-health response at local, national and global scales.

In fact, besides impressive global collaboration, the scientific response to COVID-19 has strengthened local and regional collaboration across Africa, including in clinical research, in the trials of vaccines and therapeutics, and in epidemiology and genomics^{16–18}. The scientific community globally should support these networks and expand their scope to help address HIV as well as the other public-health priorities in Africa — tuberculosis, malaria and neglected tropical diseases.

Integrate health systems. Governments, departments of health and HIV programmes need to find better ways to integrate COVID-19 and HIV services. Providing COVID-19 vaccination at clinics and centres that currently offer testing and antiretroviral therapies for HIV could help to ensure that people with advanced or uncontrolled HIV are prioritized for COVID-19 vaccination, including any additional doses that might be recommended.

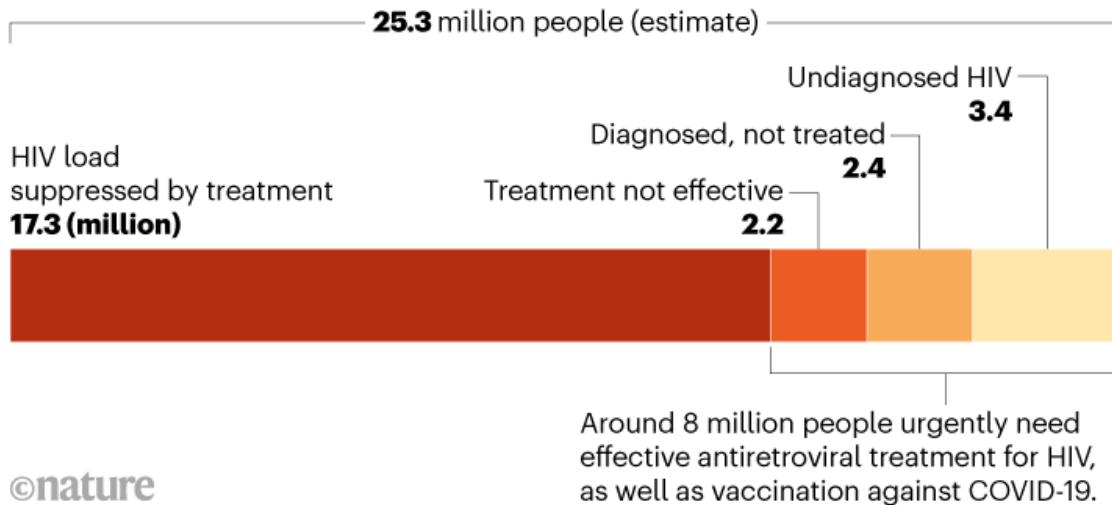
Similarly, people living in countries that have high rates of HIV should be offered HIV testing whenever they receive a COVID-19 vaccine. Currently,

the WHO recommends that people in these countries are offered HIV testing whenever they have any contact with the health system. (In practice, this happens patchily.) Thus, providing people with prevention and care for COVID-19 provides an opportunity to boost care and treatment for HIV.

Ultimately, sustained efforts to ensure that people with HIV are diagnosed, given antiretroviral treatment and then followed up to ensure that their viral load is being suppressed (see ‘Africa’s HIV treatment gap’) are crucial to reducing the burden of advanced HIV in the population — and so limiting the potential for cases of prolonged infection from SARS-CoV-2. Encouragingly, in our case study from South Africa, the individual cleared the SARS-CoV-2 infection three weeks after being given effective HIV treatment⁵.

AFRICA’S HIV TREATMENT GAP

Around one-third of people living with HIV in sub-Saharan Africa in 2020 were not getting effective antiretroviral treatment.



Source: UNAIDS (2020 data)

Avoid stigmatization

What the past year has shown is that early — and ongoing — complacency about how the COVID-19 pandemic is playing out in Africa is misplaced.

Comparing only cases and deaths across countries or continents with different demographics, and obtained using different testing strategies and reporting systems, can give a misleading picture¹⁹. A rolling analysis of excess deaths in South Africa, for instance, indicates that deaths from COVID-19 could be between two and three times higher than reported deaths, and that South Africa has been one of the hardest-hit countries in the world (see go.nature.com/3r6dt3j). Certainly, a 30–50% increased chance of death from COVID-19 for a young sector of the population that makes up most of the workforce could have devastating consequences in the long term.

On the flip side, an obvious risk of drawing attention to the interactions between SARS-CoV-2 and HIV is increasing the stigmatization of people living with HIV. It could also encourage the application of unfair public-health measures (such as travel restrictions) to countries with high rates of HIV.

The United States banned non-citizens who had tested positive for HIV from entering the country for 22 years. (The ban was lifted only in 2010.) And many other countries, such as Australia and New Zealand, still refuse to give citizenship to people with HIV, even though effective antiretroviral treatments are now available.

The best way for governments to protect their citizens is not by further stigmatizing those infected with HIV; it is by quickly providing vaccines to protect the world's most vulnerable. Only global solidarity — not discrimination, blockage or the hoarding of resources — will end the COVID-19 pandemic.

Nature **600**, 33–36 (2021)

doi: <https://doi.org/10.1038/d41586-021-03546-8>

References

1. 1.

World Health Organization. *Clinical Features and Prognostic Factors of COVID-19 in People Living with HIV Hospitalized with Suspected or Confirmed SARS-CoV-2 Infection* (WHO, 2021).

2. 2.

Jassat, W. *et al. Lancet HIV* **8**, e554–e567 (2021).

3. 3.

Boule, A. *et al. Clin. Infect. Dis.* **73**, e2005–e2015 (2021).

4. 4.

Corey, L. *et al. N. Engl. J. Med.* **385**, 562–566 (2021).

5. 5.

Karim, F. *et al. Preprint at medRxiv*
<https://doi.org/10.1101/2021.06.03.21258228> (2021).

6. 6.

The Global Fund. *Results Report 2021* (Global Fund, 2021).

7. 7.

Autenrieth, C. S. *et al. PLoS ONE* **13**, e0207005 (2018).

8. 8.

World Health Organization. *Strategy to Achieve Global Covid-19 Vaccination by Mid-2022* (WHO, 2021).

9. 9.

Wagner, C. E. *et al. Science* **373**, eabj7364 (2021).

10. 10.

Madhi, S. A. *et al.* *Lancet HIV* **8**, E568–E580 (2021).

11. 11.

Frater, J. *et al.* *Lancet HIV* **8**, e474–e485 (2021).

12. 12.

Spinelli, M. A. *et al.* Preprint at SSRN
<https://doi.org/10.2139/ssrn.3909785> (2021).

13. 13.

Kernéis, S. *et al.* *Clin. Infect. Dis.* **58**, 1130–1139 (2014).

14. 14.

Hall, V. G. *et al.* *N. Engl. J. Med.* **385**, 1244–1246 (2021).

15. 15.

Inzaule, S. C. *et al.* *PLoS Med.* **18**, e1003666 (2021).

16. 16.

Salyer, S. J. *et al.* *Lancet* **397**, 1265–1275 (2021).

17. 17.

Wilkinson, E. *et al.* *Science* **374**, 423–431 (2021).

18. 18.

Biccard, B. M. *et al.* *Lancet* **397**, 1885–1894 (2021).

19. 19.

Burki, T. K. *Lancet Respir. Med.* [https://doi.org/10.1016/S2213-2600\(21\)00504-X](https://doi.org/10.1016/S2213-2600(21)00504-X) (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03546-8>

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

- CORRESPONDENCE
- 30 November 2021

Collaborate equitably in ancient DNA research and beyond

- [Mehmet Somel](#) ⁰,
- [N. Ezgi Altınışık](#) ¹,
- [Füsun Özer](#) ² &
- [María C. Ávila-Arcos](#) ³

The expansion of ancient-DNA research has seen a few well-funded genetics groups hoarding archaeological material from all over the world, so comprehensive ethical guidelines are timely (see [S. Alpaslan-Roodenberg *et al. Nature* 599, 41–46; 2021](#)). But well-intended proposals for researchers from prestigious institutions to engage with local stakeholders will need radical reinforcement if contributing researchers in lower-income regions are to be recognized as equal collaborators.

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**, 37 (2021)

doi: <https://doi.org/10.1038/d41586-021-03541-z>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03541-z>

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

- CORRESPONDENCE
- 30 November 2021

Ancient-DNA researchers write their own rules

- [Krystal S. Tsosie](#) ORCID: <http://orcid.org/0000-0002-7291-670X>⁰,
- [Alyssa C. Bader](#) ORCID: <http://orcid.org/0000-0002-0077-9005>¹,
- [Keolu Fox](#) ORCID: <http://orcid.org/0000-0003-4215-5273>²,
- [Deborah A. Bolnick](#) ORCID: <http://orcid.org/0000-0001-8444-5127>³
- ,
- [Nanibaa' A. Garrison](#) ORCID: <http://orcid.org/0000-0002-6228-3216>⁴ &
- [Rick W. A. Smith](#) ORCID: <http://orcid.org/0000-0002-3207-0519>⁵

We challenge the global guidelines proposed for ethical DNA research on human ancestors ([S. Alpaslan-Roodenberg *et al. Nature* 599, 41–46; 2021](#)). In our view, they fail to engage with long-standing critiques from marginalized voices ([D. Harry Chi.-Kent L. Rev. 84, 147–196; 2009](#)) or to advance pre-existing research principles ([J. K. Wagner *et al. Am. J. Hum. Genet.* 107, 183–195; 2020](#)).

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**, 37 (2021)

doi: <https://doi.org/10.1038/d41586-021-03542-y>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03542-y>.

- CORRESPONDENCE
- 30 November 2021

Climate researchers: consider standing for office — I did

- [John Dearing](#) [ORCID: http://orcid.org/0000-0002-1466-9640](http://orcid.org/0000-0002-1466-9640) 0

With the COP26 climate summit receding in the rear-view mirror, in many regions it is the decisions made by local councils that are crucial to meeting national carbon-reduction targets. That's why, five years ago, after a career as an environmental-systems researcher, I decided to devote more time to politics in my home town.

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**, 37 (2021)

doi: <https://doi.org/10.1038/d41586-021-03540-0>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03540-0>

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

- CORRESPONDENCE
- 30 November 2021

Animal experiments: EU is pushing to find substitutes fast

- [Stefan Hippenstiel](#) ORCID: <http://orcid.org/0000-0002-5146-1064>⁰
- ,
• [Christa Thöne-Reineke](#) ORCID: <http://orcid.org/0000-0003-0782-2755>¹ &
- [Jens Kurreck](#) ORCID: <http://orcid.org/0000-0002-1469-0052>²

The European Parliament is once again pushing to accelerate transition to a research system that does not involve testing on animals (see go.nature.com/3hzprhj). In September, its members called for an EU-wide action that is backed by a timetable to phase out non-essential animal experimentation as soon as possible. Meeting this challenge will require extraordinary commitment and dedicated communication by the scientific community.

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**, 37 (2021)

doi: <https://doi.org/10.1038/d41586-021-03539-7>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03539-7>

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

Work

- **Discrimination still plagues science** [29 November 2021]
Career Feature • Employers need to do more to improve workplace diversity, equity and inclusion, a Nature survey finds.
- **‘For a brown invertebrate’: rescuing native UK oysters** [29 November 2021]
Where I Work • Tom Cameron works with local oyster growers to restore native oysters to their natural habitat in the United Kingdom.

- CAREER FEATURE
- 29 November 2021

Discrimination still plagues science

Employers need to do more to improve workplace diversity, equity and inclusion, a *Nature* survey finds.

- [Chris Woolston](#) ⁰

[Find a new job](#)



Illustration by Antonio Rodríguez

Social protest movements such as #MeToo and #BlackInSTEM have shone a light on the need for greater diversity, equity and inclusion at scientific

institutions worldwide. And *Nature*'s 2021 salary and job satisfaction survey, which drew responses from more than 3,200 working scientists around the world, suggests that there's much more work to do.

Just 40% of respondents felt that their employers were doing enough to promote diversity, down from 51% in 2018, when the survey last took place. A substantial minority of respondents said they had witnessed colleagues being subjected to discriminatory behaviour, and another sizeable minority said they had experienced such treatment themselves. The self-selected survey (see '*Nature*'s salary and job survey') included a series of questions that explore attitudes and experiences relating to diversity. Follow-up interviews with selected respondents and free-text comments have helped to fill out the picture.

Nature's salary and job survey

This article is the last of four linked to *Nature*'s global salary and job satisfaction survey. Previous articles looked at the [impact of the COVID-19 pandemic](#) on scientists' careers; at [salary and prospects](#); and at [job satisfaction](#).

The survey runs every three years and [was last conducted in 2018](#). It was created together with Shift Learning, a market-research company based in London, and advertised on nature.com, in Springer Nature digital products and through e-mail campaigns. It was offered in English, Mandarin Chinese, Spanish, French and Portuguese. The full survey data sets are available at go.nature.com/3eqcpk9.

The respondents reflect the relative homogeneity in science in some parts of the world. Eighty-two per cent of respondents in the United Kingdom, 81% in Germany and 74% in the United States identified themselves as white.

The free-text comment section exposed conflicting viewpoints on an often polarizing topic. A late-career Asian woman working in geology and environmental sciences at a European university wrote: "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 ... This is suffocating science and discouraging early-career academics."

But a white male professor of social sciences in the United States offered a different perspective: “When I say I have experienced and seen gender discrimination, it has always been against males. For example, we were directly told during a job search that we could not hire a white male, even though our relative representation of women and minorities is higher than average for our field. White males have long felt there is little likelihood of approval for sabbaticals or positive promotion decisions from the dean and upper administration.”

Some people think that the renewed emphasis on diversity and inclusion in science is overblown and unnecessary, says Zenobia Lewis, an evolutionary ecologist at the University of Liverpool, UK. She’s heard men say that gender-equality initiatives in science, such as the Athena Swan Charter, are no longer necessary because women have made gains. The charter, launched in the United Kingdom in 2005, was introduced in Ireland ten years later, with similar schemes now running in the United States, Canada and Australia. “My response is that it’s about equality for all, not just women,” says Lewis, who identifies as part Persian. “I’m a brown person in ecology, and there aren’t many of us.”

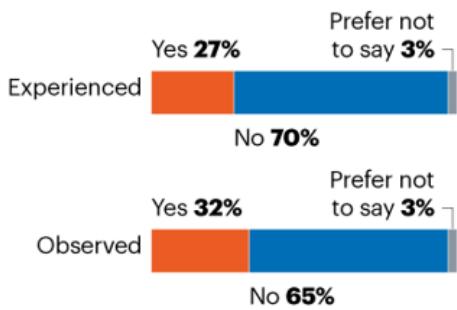
Discrimination

The survey suggests that discrimination remains common in science (see ‘Room for improvement’). Overall, 32% of respondents said they had witnessed discrimination against or harassment of colleagues in their current job. That’s up slightly from the 28% who reported observing such behaviour in 2018. Twenty-seven per cent of respondents said they had personally experienced discrimination, bullying or harassment in their present position. Again, that’s up compared with 2018, when 21% said they had such first-hand experiences.

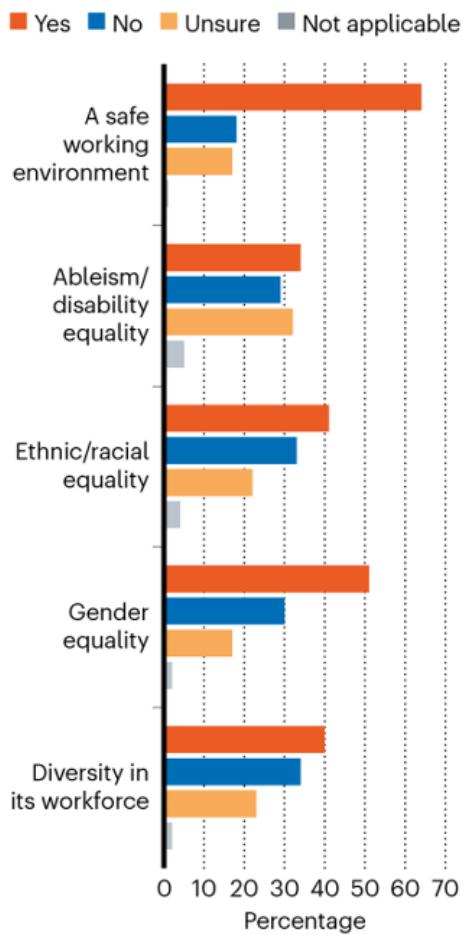
ROOM FOR IMPROVEMENT

Nature's 2021 salary and job satisfaction survey finds that reports of discrimination remain common, despite institutions renewing their focus on diversity and inclusion. Less than half of respondents feel their institution is doing enough to promote a diverse workplace.

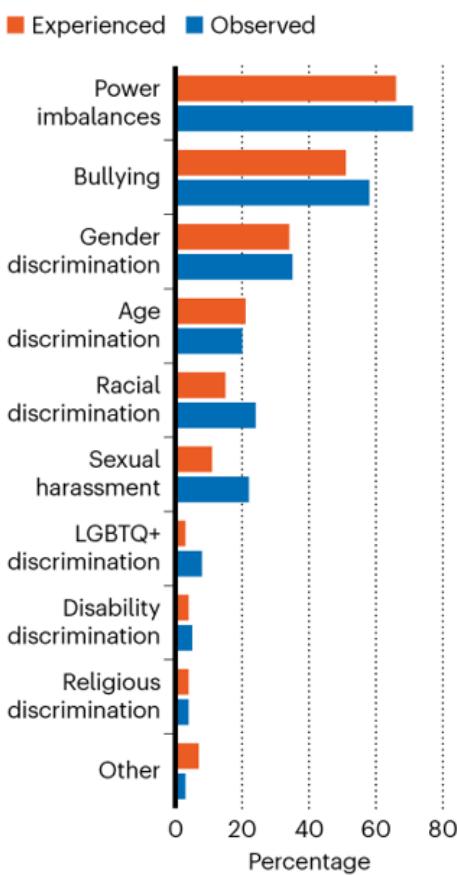
Have you experienced or observed bullying, discrimination or harassment in your current job?



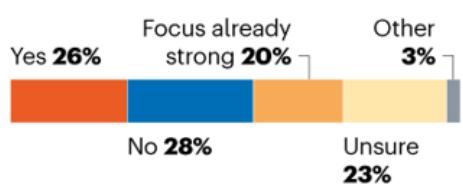
Do you believe that your workplace is doing enough to promote...



If you answered 'yes', which of the following have you experienced/observed?



Did your employer strengthen its focus on diversity as a result of movements like #BlackLivesMatter, #ShutDownSTEM, #MeToo etc?



©nature

Unsurprisingly, some groups are more likely than others to feel targeted. Women reported experiencing mistreatment more often than men: 34% to 21%. Workers in academia were twice as likely as those in industry to report such behaviour: 30% to 15%. A woman who is now a staff scientist at a US biomedical company wrote of her experience in academia: “I was bullied and harassed repeatedly at my previous job, and literally nothing there has changed or will ever change. My current job is much nicer, but I will never ever work in an academic setting again. A postdoc in the lab kept touching my hair and the university did absolutely nothing to protect me or stop it.”

Overall, 17% of the women — and 1% of the men — in the survey reported being the target of gender discrimination. Sexism remains rampant, says survey respondent Fiona Simpson, a cancer researcher at the University of Queensland in Brisbane, Australia. “Everybody talks about equality in science, but it doesn’t actually happen,” she says. “There are so many articles, so much discussion, but over my 30 years it’s gotten worse.”

Simpson says that sexism was more outward and obvious in her early career, but in her view, the more subtle discrimination of today can be just as damaging. “I’ve watched female academics get upset over the way they’re being treated, or the instability of their roles. When they’re emotional, they’re written off as unstable or hormonal. It’s changed from an overt thing to a type of gaslighting,” she says, referring to manipulative behaviour that makes someone question their own sanity.



How burnout and imposter syndrome blight scientific careers

Ethnicity can be an important factor too, especially in countries where the scientific workforce is predominantly white. In the United Kingdom, for example, 27 of the 54 respondents who did not identify as white said they had personally experienced discrimination, bullying or harassment on the job. That's nearly twice the rate reported by 357 white respondents in the United Kingdom. In the United States, the 221 respondents who did not identify as white were also more likely than white colleagues to report such experiences: 33% to 25%.

Among those reporting personal experience of mistreatment, the most commonly cited instances fell under the categories of power imbalances (66%) and bullying (51%). The most commonly reported forms of discrimination were related to gender (34%), age (21%) and race (15%).

Age discrimination was a recurring complaint in the comment section. A government employee in South Africa wrote that "job advancement for older people is also now a huge challenge due to pressure to advance young people". A self-employed US scientist in the field of astronomy and planetary science wrote: "When I lost my previous job as a result of funding cuts, the great majority of people laid off were older than 50, even though that age group did not comprise a majority of employees. Also, I know three people in that age range who were repeatedly rejected for positions, for which they were extremely well qualified, in favour of younger hires. One of them, who has two PhDs, has given up looking and basically retired early."

Less than 1% of respondents reported experiencing discrimination against people from sexual and gender minorities; that rate is essentially unchanged compared with the 2018 survey. A South African researcher in ecology and evolution wrote: "As a lesbian, I think my job prospects are better than they would have been just ten years ago because of cultural shifts and also legal changes in South Africa regarding employment equity."

Only 7% of respondents reported a disability, a reflection of a widely acknowledged lack of representation of this community in scientific fields. Disabled people face particular career challenges in science, says Michelle Moram, a London-based materials scientist who is currently working

remotely for Victoria University of Wellington, New Zealand. In 2010, Moram was diagnosed with a serious autoimmune disorder which would have qualified her for disability protection under the law. Still, she kept the illness hidden for years for fear of jeopardizing her chances of promotion.



Collection: Diversity and scientific careers

A truly diverse research system would employ people from a range of socioeconomic backgrounds, but Lewis says people from poorer families remain at a distinct disadvantage. She co-wrote a paper exploring how socioeconomic background as well ethnicity can affect early-career progression in the fields of ecology and evolution ([K. M. Wanelik et al. Ecol. Evol. 10, 6870–6880; 2020](#)). One finding: early-career researchers from less privileged backgrounds tended to have positions with a teaching component rather than ones solely devoted to research; the latter are often more prestigious for career progression.

Moram, the first member of her family to attend university, says she struggled to find her place in science after moving from University College Cork, Ireland, to take up a PhD at the University of Cambridge, UK, in 2003. “My family is what they call ‘underclass’ in the United Kingdom. Part of my journey was learning to pass as a middle-class person. I had to change my accent, learn different vocabulary, wear different clothes — just fake it to be a different person.”

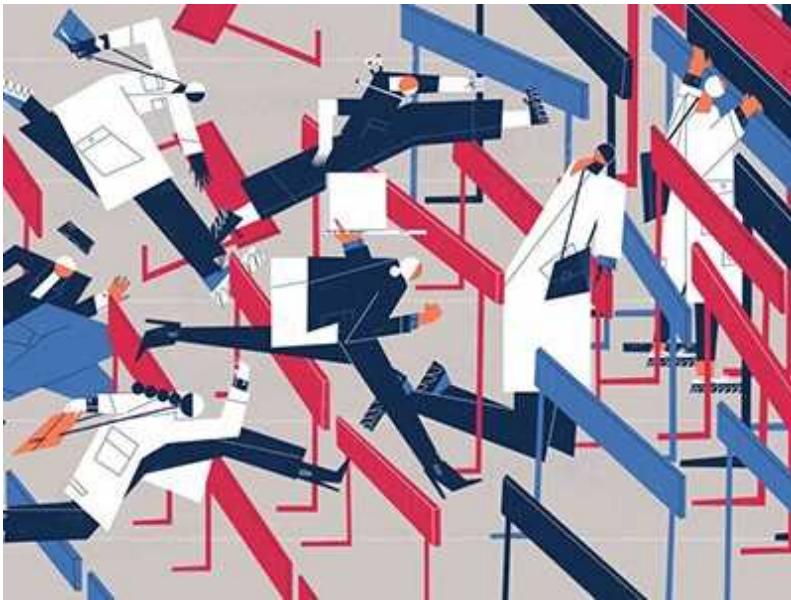
Moram says her background helped motivate her to study especially hard. “I had no idea how anything in the system worked at all. I just decided to work hard at whatever I could,” she says. “It was important for me to get out of home. I was terrified of having to go back home and work in a shop.”

Falling short

A slim majority of survey respondents (51%) felt that their institutions were doing enough to promote gender equity, down from 58% in 2018. Forty-one per cent thought that their place of employment was doing enough to promote ethnic or racial equality, down from 52% in 2018.

It’s not surprising that the number of people who are underwhelmed by their employer’s approach to diversity is growing, says Maria Miriti, a plant ecologist at Ohio State University in Columbus. Miriti wrote a paper ([M. N. Miriti *BioScience* 70, 237–242; 2020](#)) exploring strategies to boost the recruitment and retention of people from minority groups in science, technology, engineering and mathematics (STEM).

“People’s awareness of the breadth of inequality has become more focused in the pandemic,” she says. “There’s pressure to respond. But really getting into the trenches to change the system to promote greater equality and inclusion is hard.” Miriti says that fundamental changes are needed to improve equity in STEM, starting with rethinking how scientists are evaluated and promoted. “We value ‘grantsmanship’ [the ability to secure grants], publications and citations. All three of those factors can be affected by racial and gender discrimination. [Scientists] act like it’s a level playing field. It’s hard for us to accept that recognition can be tied to gender and race.”



Stagnating salaries present hurdles to career satisfaction

Just over one-quarter of respondents felt that their institutions had increased their focus on diversity in response to social justice movements such as #BlackLivesMatter and #MeToo. The former gained momentum following the murder, in May 2020, of George Floyd, an unarmed Black man, by a police officer in Minneapolis, Minnesota. #MeToo protests accelerated in 2017 in response to sexual-abuse allegations relating to film producer Harvey Weinstein.

Lewis is underwhelmed by institutions' responses. "There was suddenly this wave of organizations and institutions putting out statements of solidarity," she says. "It's about a year on, and I'm still waiting to see how much impact it's actually going to have, or whether it's just paying lip service."

Many universities and companies have established diversity committees to improve the recruitment and retention of members of under-represented groups, but their reach remains limited. Just over 12% of respondents said that they had participated in an institution-wide diversity committee, and 21% weren't sure if such committees even existed at their place of work. Women were more likely than men to participate in such committees: 14% to 9%.

Miriti thinks that diversity committees can be important drivers for change, especially in academia. “Universities should absolutely have diversity committees,” she says. “There’s too much change that needs to happen if we’re serious about increasing broad participation in our scientific disciplines.” However, she warns that involvement in such committees isn’t always highly valued when it comes to promotion and tenure. Still, she says, the investment in time and energy can pay off by bringing new people to the table. “Women and minorities should be motivated to do this work, no matter what.”

Miriti adds that more people from majority groups — for example, white men in countries such as the United States and the United Kingdom — should participate in diversity committees, but only if they’re willing to invest the time to truly address the issues. “It’s important to avoid what some refer to as ‘performative allyship’,” she says.

Some researchers have given up fighting bias. A biomedical postdoctoral researcher of Iranian descent in Canada wrote: “I’ve actually identified some perks to being discriminated against. For one, I don’t have to deal with the responsibilities that come with a more senior title. As long as I stay productive, I will likely have a job and have more time for research. I can walk away with the knowledge that I earned every penny and at times gave more than I took. There is comfort in that.”

How do you feel about diversity, equity and inclusion in science?

Free-text comments in *Nature*’s global salary and job satisfaction survey highlight scientists’ thoughts on the issues they have encountered in their careers. Comments have been edited for length and clarity and, when necessary, translated into English.

- I feel there is too much focus on hiring ‘diverse’ candidates, based on easily measurable criteria (gender, race), and not enough focus on diversity of thought and supporting existing employees through challenges such as maternity and disability. *Postdoc in biomedical field, UK.*

- I have recently felt that being white and middle-aged is an impediment to seeking positions such as associate dean or dean. *Professor in social sciences, USA.*
- As a white person I am not offered promotional opportunities so am stuck in same job for over 15 years. *Staff scientist in government, South Africa.*
- Potential employers do not like to hire females due to the possibility that they will have childcare responsibilities, although this is never stated officially. More than half of postdocs in my department are female but at the principal investigator level, 80–90% of them are (white) men. What happens to all those female postdocs? *Postdoc in biomedical field, Denmark.*
- I've never felt like I have white privilege because I've struggled so much and overcome extreme hardships to get where I am. But now I'm still overshadowed because I'm not a person whose race or ethnicity is 'diverse'. *Technician in biomedical field, USA.*
- The people who didn't hire diverse staff/faculty before are still in charge, taking credit for being diverse now. But in fact, they aren't improving anything. They will continue to hire mediocre elites and complement it with some goofy 'diverse candidates only' job openings. Some great researchers will leave academia because of this and history will repeat itself. *Assistant professor in social sciences, USA.*
- Our university has had university-level, college-level and school-level committees to promote Equality, Diversity and Inclusion [EDI] for years now. Although I have been heavily involved at many levels, it has been clear for some time that these have been instituted for the sake of (1) public visibility, and (2) access to research funding (funders have tied funding to EDI initiatives). There is very little real appetite to change what is wrong with the institution. *Associate professor in architecture, Ireland.*

Nature **600**, 177-179 (2021)

doi: <https://doi.org/10.1038/d41586-021-03043-y>

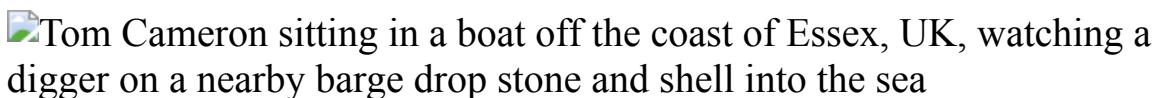
| [Section menu](#) | [Main menu](#) |

- WHERE I WORK
- 29 November 2021

‘For a brown invertebrate’: rescuing native UK oysters

Tom Cameron works with local oyster growers to restore native oysters to their natural habitat in the United Kingdom.

- [Virginia Gewin](#) ⁰



Tom Cameron is a senior lecturer in ecology at the University of Essex, Colchester, UK. Credit: Jessica Hallett/*Nature*

For the past five years, I’ve studied oysters — a commercially and environmentally important species in southeast England. My research is very practical: I help to solve problems by working with oyster growers (known locally as oystermen), regulators and other community members. Resulting papers are evidence of work I’ve already done.

Most oysters in this area are a non-native species (*Crassostrea gigas*). Locally, it’s well established and has been since the 1960s, but allowing it to spread to nearby estuary systems has been controversial: there are concerns that it could become an invasive species.

Working with aquaculture producers, I help to guide efforts to restore the native oyster (*Ostrea edulis*), populations of which declined owing to overfishing, habitat destruction, pollution and disease. *Crassostrea gigas* oysters have provided enough income for oyster growers to spend time and effort restoring the local species. We’ve done some cool things, including

creating one of the largest coastal marine conservation zones in the United Kingdom — more than 284 square kilometres — and all for an unseen brown invertebrate that lacks the charisma of a dolphin.

This picture is from a typical day in the field. During high tides, we go out in a boat to take sonar readings to map potential oyster habitats; at low tide, we put on waders and go out on the mud flats to look for juvenile oysters. We focus our conservation efforts on spots where juvenile oysters are already trying to get established.

Amazingly, these filter feeders don't require feeding by humans, and they clean the water as they grow. Bivalve aquaculture such as this has become a cornerstone of the 'blue economy' — using marine resources sustainably for economic growth while preserving ocean health. It will take more work to determine how the balance can be reached, but oysters will be part of that conversation.

Nature **600**, 182 (2021)

doi: <https://doi.org/10.1038/d41586-021-03573-5>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03573-5>

Research

- [**Robotic sample return reveals lunar secrets**](#) [01 December 2021]
News & Views • A mission to unexplored lunar territory has returned the youngest volcanic samples collected so far. The rocks highlight the need to make revisions to models of the thermal evolution of the Moon.
- [**A reconstruction of early cryptic COVID spread**](#) [01 November 2021]
News & Views • To respond better to future pandemics, we must understand how the SARS-CoV-2 virus dispersed so rapidly. A model of COVID-19 spread sheds light on cryptic transmission, undetected by surveillance efforts, in early 2020.
- [**A seagrass harbours a nitrogen-fixing bacterial partner**](#) [03 November 2021]
News & Views • How underwater seagrasses obtain the nitrogen they need has been unclear. Evidence has now emerged of a partnership with a bacterium that might be analogous to the system used by many land plants to gain nitrogen.
- [**Choreographing water molecules to speed up hydrogen production**](#) [01 December 2021]
News & Views • A technical feat reveals subtle changes in water structure that can accelerate hydrogen production at an electrode interface. The catalytic process could be developed to help boost supply of this clean fuel.
- [**Artificial intelligence aids intuition in mathematical discovery**](#) [01 December 2021]
News & Views • Machine-learning tools have been used to assist the part of mathematical research that usually relies on human intuition and creativity — leading to two fundamental results in different areas of mathematics.
- [**Helpline data used to monitor population distress in a pandemic**](#) [17 November 2021]
News & Views • The initial stages of the COVID-19 pandemic saw an increase in calls to mental-health helplines in 19 countries. Helpline-call data can be used to monitor distress at a population level in near-real time.
- [**A dry lunar mantle reservoir for young mare basalts of Chang'e-5**](#) [19 October 2021]

Article • Water abundance and hydrogen isotope compositions of two-billion-year-old basalt samples returned from the Moon by the Chang'e-5 mission suggest that the samples came from a relatively dry mantle source.

- **Two-billion-year-old volcanism on the Moon from Chang'e-5 basalts** [19 October 2021]

Article • Basalt samples returned from the Moon by the Chang'e-5 mission are revealed to be two billion years old by radioisotopic dating, providing insight on the volcanic history of the Moon.

- **Non-KREEP origin for Chang'e-5 basalts in the Procellarum KREEP Terrane** [19 October 2021]

Article • Isotopic analysis of basalt clasts returned from the Moon by the Chang'e-5 mission indicates that the rocks were derived from a mantle source that lacked potassium, rare-earth elements and phosphorus.

- **Sound emission and annihilations in a programmable quantum vortex collider** [01 December 2021]

Article • By controlling the generation and collision of individual vortices in atomic Fermi superfluids, a study provides a comprehensive view of vortex decay due to mutual friction and vortex–sound interaction.

- **Advancing mathematics by guiding human intuition with AI** [01 December 2021]

Article • A framework through which machine learning can guide mathematicians in discovering new conjectures and theorems is presented and shown to yield mathematical insight on important open problems in different areas of pure mathematics.

- **Optomechanical dissipative solitons** [01 December 2021]

Article • Stable, dissipative optomechanical solitons are realized using optical fields in a whispering gallery mode resonator by balancing the optomechanical nonlinearities with a tailored modal dispersion.

- **In situ Raman spectroscopy reveals the structure and dissociation of interfacial water** [01 December 2021]

Article • Interfacial water consists of hydrogen-bonded water and Na·H₂O, its structure changes at hydrogen evolution reaction (HER) potentials, and when structurally ordered it aids interfacial electron transfer, resulting in higher HER rates.

- **Late Quaternary dynamics of Arctic biota from ancient environmental genomics** [20 October 2021]

Article • A large-scale metagenomic analysis of plant and mammal environmental DNA reveals complex ecological changes across the circumpolar region over the past 50,000 years, as biota responded to changing climates, culminating in the postglacial extinction of large mammals and emergence of modern ecosystems.

- **Temporal transitions in the post-mitotic nervous system of *Caenorhabditis elegans*** [10 November 2021]
Article • In the post-mitotic nervous system of *Caenorhabditis elegans*, the promotion of a mature neuronal transcription program is controlled by the developmental-stage-specific upregulation of the microRNA LIN-4 and the subsequent repression of its target, the transcription factor lin-14.
- **Thalamic circuits for independent control of prefrontal signal and noise** [06 October 2021]
Article • Two different cell types in the mediiodorsal thalamus have complementary roles in decision-making, with one type of mediiodorsal projection amplifying prefrontal activity under low signal levels and one type suppressing it under high noise levels.
- **Terrestrial-type nitrogen-fixing symbiosis between seagrass and a marine bacterium** [03 November 2021]
Article • The N2-fixing symbiont ‘*Candidatus Celerinatantimonas neptuna*’ lives inside the root tissue of the seagrass *Posidonia oceanica*, providing ammonia and amino acids to its host in exchange for sugars and enabling highly productive seagrass meadows to thrive in the nitrogen-limited Mediterranean Sea.
- **The human microbiome encodes resistance to the antidiabetic drug acarbose** [24 November 2021]
Article • Bacteria in the human gut and oral microbiome encode enzymes that selectively phosphorylate the antidiabetic drug acarbose—an inhibitor of both human and bacterial α -glucosidases—resulting in its inactivation and limiting the drug's effects on the ability of the host to metabolize complex carbohydrates.
- **Antiviral activity of bacterial TIR domains via immune signalling molecules** [01 December 2021]
Article • The mechanism of Thoeris—a bacterial anti-phage defence system—is described in detail, revealing that bacterial TIR-domain proteins recognize infection and produce signalling molecules to execute cell death, akin to the roles of these proteins in plants.
- **Mental health concerns during the COVID-19 pandemic as revealed by helpline calls** [17 November 2021]
Article • Data collected from crisis helplines during the COVID-19 pandemic show that pandemic-related issues replaced rather than exacerbated underlying anxieties, and demonstrate that helpline data are useful indicators of public mental health.
- **Cryptic transmission of SARS-CoV-2 and the first COVID-19 wave** [25 October 2021]
Article • Modelling highlights international travel as the main driver of the introduction of SARS-CoV-2 to Europe and the USA, and suggests that introductions and local transmission may have begun in January 2020.

- **Independent infections of porcine deltacoronavirus among Haitian children** [17 November 2021]
Article • The presence of porcine deltacoronavirus has been detected in three children from Haiti that could have originated from zoonotic spillover.
- **Self-guarding of MORC3 enables virulence factor-triggered immunity** [10 November 2021]
Article • MORC3 is revealed as an essential negative regulator of the anti-viral interferon response that functions in an innate immune pathway that detects viral virulence factors.
- **Structural basis of cytokine-mediated activation of ALK family receptors** [13 October 2021]
Article • Structural studies of the complex of anaplastic lymphoma kinase and leukocyte tyrosine kinase and their activating cytokines identify unique architectural features of the complex, and provide a novel mechanistic paradigm among receptor tyrosine kinases.
- **Structural basis for ligand reception by anaplastic lymphoma kinase** [24 November 2021]
Article • Analysis of crystal structures of anaplastic lymphoma kinase elucidate the mechanism by which ligand binding and the glycine-rich domain regulate its activity.
- **Mechanism for the activation of the anaplastic lymphoma kinase receptor** [24 November 2021]
Article • Cryo-electron microscopy, nuclear magnetic resonance and X-ray crystallography are used to provide structural and mechanistic details of the activation of anaplastic lymphoma kinase by the ligands ALKAL1 and ALKAL2.
- **Aldehyde-driven transcriptional stress triggers an anorexic DNA damage response** [24 November 2021]
Article • Endogenous formaldehyde accumulation reveals Cockayne syndrome in mice and stimulates production of the anorexiogenic peptide GDF15 in proximal tubule cells.
- **Structure, function and pharmacology of human itch receptor complexes** [17 November 2021]
Article • Cryo-electron microscopy structures of the MRGPRX2–Gi1 trimer in complex with polycationic compound 48/80 or inflammatory peptides provide insights into the sensing of cationic allergens by MRGPRX2, potentially facilitating the design of therapies to prevent unwanted pseudoallergic reactions.
- **Structure, function and pharmacology of human itch GPCRs** [17 November 2021]
Article • Structural studies of the itch receptors MRGPRX2 and MRGPRX4 in complex with endogenous and synthetic ligands provide a basis for the development of therapeutic compounds for pain, itch and mast cell-mediated hypersensitivity.

- **Non-trivial role of internal climate feedback on interglacial temperature evolution** [01 December 2021]
Matters Arising •
- **Reply to: Non-trivial role of internal climate feedback on interglacial temperature evolution** [01 December 2021]
Matters Arising •

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

- NEWS AND VIEWS
- 01 December 2021

Robotic sample return reveals lunar secrets

A mission to unexplored lunar territory has returned the youngest volcanic samples collected so far. The rocks highlight the need to make revisions to models of the thermal evolution of the Moon.

- [Richard W. Carlson](#) 

The wait is over for more news from the Moon¹. Three studies in this issue, by [Tian et al.](#)², [Hu et al.](#)³ and [Li et al.](#)⁴, together with one in *Science* by [Che et al.](#)⁵, report data on the lunar samples brought back by China's robotic Chang'e-5 mission — the first to return samples since the Soviet Union's Luna 24 mission in 1976. These data shed light on volcanic eruptions that occurred more than one billion years more recently than those known about previously, and provide information on the cause of the volcanism that cannot be obtained from orbit. The results raise questions about the structure and thermal evolution of the lunar interior, and could help to improve methods for estimating the age of planetary surfaces throughout the inner Solar System.

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**, 39–40 (2021)

doi: <https://doi.org/10.1038/d41586-021-03547-7>

References

1. 1.

Spudis, P. *Nature Geosci.* **2**, 234–236 (2009).

2. 2.

Tian, H.-C. *et al.* *Nature* **600**, 59–63 (2021).

3. 3.

Hu, S. *et al.* *Nature* **600**, 49–53 (2021).

4. 4.

Li, Q.-L. *et al.* *Nature* **600**, 54–58 (2021).

5. 5.

Che, X. *et al.* *Science* **374**, 887–890 (2021).

6. 6.

Qian, Y. Q. *et al.* *J. Geophys. Res. Planets* **123**, 1407–1430 (2018).

7. 7.

Jolliff, B. L., Gillis, J. J., Haskin, L. A., Korotev, R. L. & Wieczorek, M. A. *J. Geophys. Res.* **105**, 4197–4216 (2000).

8. 8.

Warren, P. H. *Annu. Rev. Earth Planet. Sci.* **13**, 201–240 (1985).

9. 9.

Borg, L. E., Gaffney, A. M. & Shearer, C. K. *Meteorit. Planet. Sci.* **50**, 715–732 (2015).

10. 10.

Miljković, K. *et al.* *Science* **342**, 724–726 (2013).

11. 11.

Laneuville, M., Wieczorek, M. A., Breuer, D. & Tosi, N. *J. Geophys. Res. Planets* **118**, 1435–1452 (2013).

12. 12.

Saal, A. E., Hauri, E. H., Langmuir, C. H. & Perfit, M. R. *Nature* **419**, 451–455 (2002).

13. 13.

Stevenson, D. J. *Annu. Rev. Earth Planet. Sci.* **15**, 271–315 (1987).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03547-7>

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

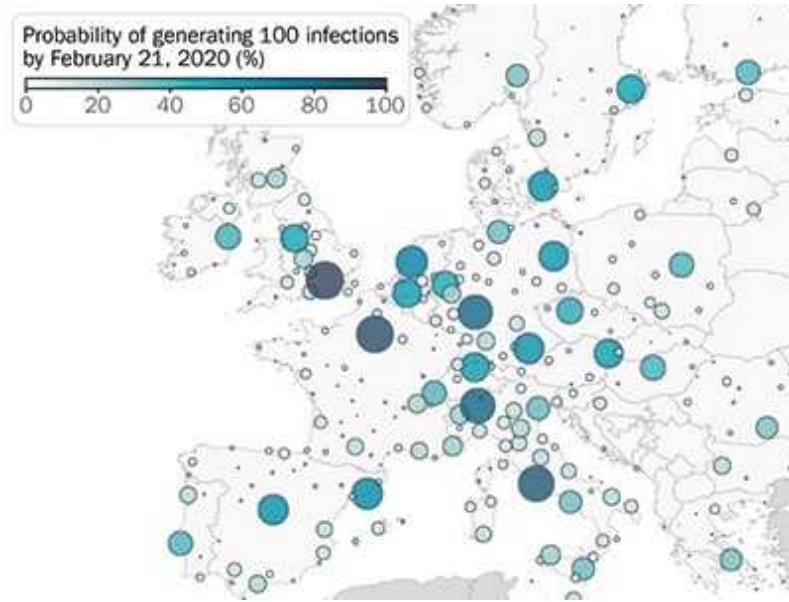
- NEWS AND VIEWS
- 01 November 2021

A reconstruction of early cryptic COVID spread

To respond better to future pandemics, we must understand how the SARS-CoV-2 virus dispersed so rapidly. A model of COVID-19 spread sheds light on cryptic transmission, undetected by surveillance efforts, in early 2020.

- [Simon Cauchemez](#) ⁰ &
- [Paolo Bosetti](#) ¹

Over the past 20 months, the COVID-19 pandemic has caused more than 4.9 million reported deaths (<https://coronavirus.jhu.edu>), and measures to limit the spread of the SARS-CoV-2 virus have affected the lives of people around the world. Although modelling has helped to reconstruct the early dynamics of the epidemic in some countries, we still lack a coherent picture of how the pandemic unfolded globally. [Writing in Nature](#), Davis *et al.*¹ use a worldwide model to assess early, cryptic transmission of SARS-CoV-2 — the spread of the virus that was not detected by initial surveillance efforts — in the United States and Europe.



[Read the paper: Cryptic transmission of SARS-CoV-2 and the first COVID-19 wave](#)

Looking back at the chronology of the first months of the pandemic, it is concerning how fast the virus spread around the world, leading to a massive shutdown of people's social and economic lives. On 10 January 2020, 41 cases of COVID-19 were reported in Wuhan, Hubei province, China. The first reports of infection outside China were made on 13 January (in Thailand) and 16 January (in Japan). Wuhan was locked down on 23 January, followed by lockdowns in Italy (11 March), Spain (14 March), Austria (16 March) and France (17 March). Many countries, caught off-guard by the rapidly changing situation, reported a large death toll. How can we do better next time? To answer this question and improve our preparedness in the face of future pandemics, it is crucial to build a clearer picture of the initial spread of the virus. This is difficult, because the limited capacity to test for virus infections at the time meant that, in many locations, SARS-CoV-2 transmission might have been undetected.

Davis *et al.* used the Global Epidemic and Mobility (GLEAM) model, which has both stochastic (incorporating elements of randomness) and mechanistic (including defined principles about the biological and social mechanisms associated with viral infection and transmission) components to simulate virus spread on a global scale². The model relies on various types of data to capture the multifactorial nature of the epidemic process. This information

includes data describing: the populations in which the virus spread, such as country-specific demographics; the movement of people on international and local scales (for example, airline transport networks and commuting flows); and behaviours, such as information documenting how individuals of different ages mingle with each other. The model also captures biological aspects of SARS-CoV-2 transmission, clinical features (such as lethality for each age group analysed) and the timing of non-pharmaceutical containment measures, such as lockdowns.

Using the model, the authors shed light on how the virus propagated around the world. For example, the model confirms that, at the start (between January and March 2020), many introductions of the virus went unnoticed in both the United States and Europe (Fig. 1). In the United States, California was the first state to be affected by local transmission (week of 26 January, according to the model), even though evidence of local transmission there was reported only a month later (26 February). In Europe, modelling indicated that local transmission started at the end of January in Italy, the United Kingdom, Germany and France, again with substantial delays before local cases were reported. Overall, Davis and colleagues estimate that, by 8 March 2020, only 1–3 out of every 100 SARS-CoV-2 infections were detected in the United States and Europe.

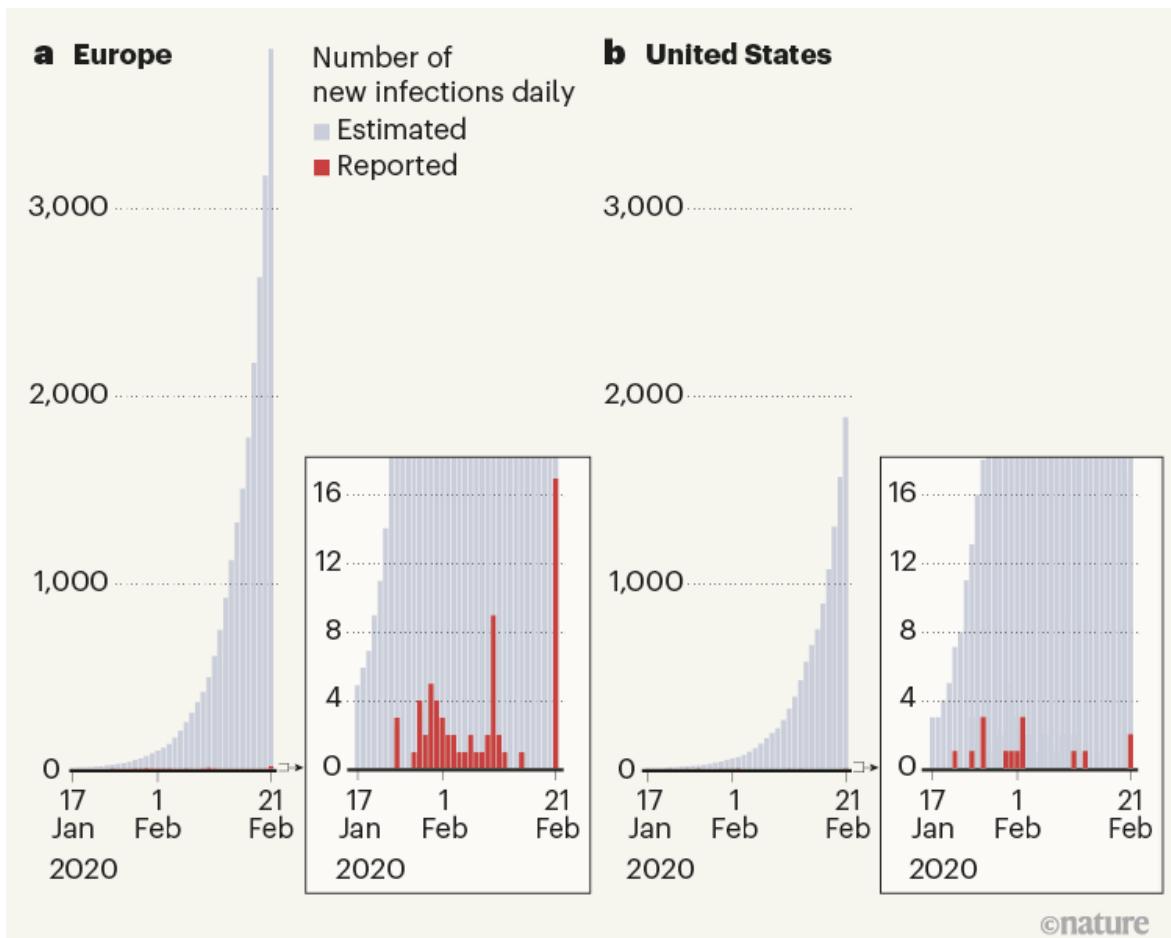


Figure 1 | Reported and estimated numbers of cases of COVID-19 in early 2020. Davis *et al.*¹ used a sophisticated epidemiological model to reconstruct the spread of the SARS-CoV-2 virus in 30 European countries and the United States at the beginning of the COVID-19 pandemic. They estimate that, between 17 January and 21 February 2020, only a tiny proportion of cases were detected by surveillance efforts, and that the rest of the estimated cases resulted from cryptic transmission.

Remarkably, in all 30 European countries studied and in all US states, local transmission started in a relatively narrow time window of about 2 months between mid-January and mid-March 2020. These variations in starting dates — as well as differences in the timing and strength of interventions — led to highly variable patterns of disease spread in different regions. By 4 July 2020, infection attack rates (the proportions of a region's population that contracted the virus) ranged from around 0.2% to about 15% in the analysed regions.

The modelling can help to assess the performance of testing strategies put in place at the time, which mostly affected passengers arriving from China. Although a large proportion of individuals with COVID-19 entering the United States and Europe indeed came from China in January 2020, neighbouring European countries and US states quickly became major contributors to infection rates in the areas analysed.

The lack of robust epidemic information on international locations affected by transmissions means that the testing policies in most countries were too narrow in scope and, probably, ineffective. Davis and co-workers provide a counterfactual scenario: if testing policies had been broader, enabling the detection of 50% of all imported and local infections in the United States and Europe, the start of local transmission might have been delayed by at least a month in many places. This would have given governments more time to prepare — for example, by increasing health-care capacities and acquiring protective equipment. However, considering the optimistic underlying assumption of this counterfactual scenario, it remains difficult to imagine that countries could have scaled up their testing capacity sufficiently fast to detect 50% of all initial SARS-CoV-2 infections.



[A machine-learning algorithm to target COVID testing of travellers](#)

There are other methods for reconstructing the spread of SARS-CoV-2 — for example, studies that sequence the RNA of the virus to determine the

‘family tree’ of circulating strains^{3,4}. In these approaches, the reconstructed history of viral circulation is informed by and consistent with all data available up to the time of analysis. By contrast, the simulations presented by Davis *et al.* are initialized to match data from the start of the pandemic, but are not constrained by what was observed after 21 January 2020. The authors therefore explore a broader range of possible trajectories of viral spread than do the phylogenetic studies, and some of these trajectories might not perfectly match what happened in reality. Despite this feature of the simulations, there is a relatively good correlation between reconstructed incidences of infections and observed numbers of cases.

Compared with other approaches, a key strength of Davis and co-workers’ approach is that their model provides a mechanistic understanding of the spread of a pandemic. This makes it possible to model scenarios that document how the pandemic might have unfolded given different policy options, such as the counterfactual scenarios mentioned above that included more-intensive testing from the start of the pandemic. This capacity of models to capture and anticipate complex nonlinear dynamics and assess the potential effects of different policy options explains in large part why such modelling has increasingly been used to support policymaking, and this trend has been amplified by the pandemic. The creation of a Center for Disease Forecasting by the US Centers for Disease Control and Prevention is only the most recent example of this trend⁵.

Several exciting avenues should be considered for improving these models for future use. Methods that calibrate these models to both epidemiological and virus-sequencing data could improve their performance for ‘now-casting’ and forecasting viral spread, as well as for assessing the effects of policies in real time. A more detailed understanding of the factors affecting spread (such as mobility, mixing patterns, behavioural changes, climate and population immunity) is necessary to improve model assumptions. Models are only as good as the data they rely on. Although COVID-19 surveillance and related data sets have considerably improved during the pandemic, these efforts must be maintained beyond COVID-19, and stable information systems^{6,7} (for example, <https://coronavirus.jhu.edu>) must continue to be made available to assess global disease dynamics in a robust manner.

doi: <https://doi.org/10.1038/d41586-021-02989-3>

References

1. 1.

Davis, J. T. *et al.* *Nature* **600**, 127–132 (2021).

2. 2.

Balcan, D. *et al.* *J. Comput. Sci.* **1**, 132–145 (2010).

3. 3.

Fauver, J. R. *et al.* *Cell* **181**, 990–996 (2020).

4. 4.

du Plessis, L. *et al.* *Science* **371**, 708–712 (2021).

5. 5.

CDC Stands Up New Disease Forecasting Center.

<https://www.cdc.gov/media/releases/2021/p0818-disease-forecasting-center.html> (CDC, 2021).

6. 6.

Hale, T. *et al.* *Nature Hum. Behav.* **5**, 529–538 (2021).

7. 7.

Hadfield, J. *et al.* *Bioinformatics* **34**, 4121–4123 (2018).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02989-3>

- NEWS AND VIEWS
- 03 November 2021

A seagrass harbours a nitrogen-fixing bacterial partner

How underwater seagrasses obtain the nitrogen they need has been unclear. Evidence has now emerged of a partnership with a bacterium that might be analogous to the system used by many land plants to gain nitrogen.

- [Douglas G. Capone](#) ORCID: <http://orcid.org/0000-0002-3968-736X>

Seagrass meadows are a prominent feature of many shallow coastal areas of the temperate through to the tropical ocean. Seagrasses provide a crucial habitat for invertebrates and juvenile fish, stabilize sediments and buffer the shoreline against erosion¹. Moreover, they contribute directly and positively to the ‘blue economy’ of the oceans through their long-term storage of carbon². Lush and highly productive seagrass beds often thrive in nutrient-deficient waters, and attempts to solve the enigma of how they accomplish this feat have driven considerable research over the years. [Writing in *Nature*](#), Mohr *et al.*³ provide crucial evidence indicating that the success of a seagrass called *Posidonia oceanica* (Fig. 1), which proliferates throughout the warm waters of the Mediterranean Sea (and elsewhere), might be attributed to the development of a highly integrated partnership with a bacterium. This system is reminiscent of those found in some terrestrial plants.

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**, 42-43 (2021)

doi: <https://doi.org/10.1038/d41586-021-02956-y>

References

1. 1.

Larkum, A. W. D., Orth, R. J. & Duarte, C. M. (eds) *Seagrasses: Biology, Ecology and Conservation* (Springer, 2006).

2. 2.

Lovelock, C. E. & Duarte, C. M. *Biol. Lett.* **15**, 20180781 (2019).

3. 3.

Mohr, W. *et al.* *Nature* **600**, 105–109 (2021).

4. 4.

Thies, J. E. in *Principles and Applications of Soil Microbiology* 3rd edn (eds Gentry, T. J., Fuhrmann, J. J. & Zuberer, D. A.) 455–487 (Elsevier, 2021).

5. 5.

Zuberer, D. A. in *Principles and Applications of Soil Microbiology* 3rd edn (eds Gentry, T. J., Fuhrmann, J. J. & Zuberer, D. A.) 423–453 (Elsevier, 2021).

6. 6.

Larkum, A. W. D., Waycott, M. & Conran, J. G. in *Seagrasses of Australia: Structure, Ecology and Conservation* (eds Larkum, A. W. D., Kendrick, G. A. & Ralph, P. J.) 3–29 (Springer, 2018).

7. 7.

Welsh, D. T. *Ecol. Lett.* **3**, 58–71 (2000).

8. 8.

Cramer, M. J., Haghshenas, N., Bagwell, C. E., Matsui, G. Y. & Lovell, C. R. *Int. J. Syst. Evol. Microbiol.* **61**, 1053–1060 (2011).

9. 9.

Clúa, J., Roda, C., Zanetti, M. E. & Blanco, F. A. *Genes* **9**, 125 (2018).

10. 10.

Evans, S. M., Griffin, K. J., Blick, R. A. J., Poore, A. G. B. & Vergés, A. *PLoS ONE* **13**, e0190370 (2018).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02956-y>.

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

- NEWS AND VIEWS
- 01 December 2021

Choreographing water molecules to speed up hydrogen production

A technical feat reveals subtle changes in water structure that can accelerate hydrogen production at an electrode interface. The catalytic process could be developed to help boost supply of this clean fuel.

- [Matthias M. Waegle](#) 

Electrocatalysis, which accelerates chemical reactions driven by an electric potential at a solid–liquid interface, could be a key contributor to a sustainable global economy because it can convert electrical energy from renewable power sources into green fuels such as hydrogen gas¹. Writing in *Nature*, Wang and colleagues² describe an important advance in the molecular understanding of how the rate of an electrocatalytic process is rooted in the structure of water at the interface between a solid electrode and an aqueous salt solution (an electrolyte). Their findings could help to improve the reaction selectivity and energy efficiency of electrocatalytic interfaces.

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**, 43–44 (2021)

doi: <https://doi.org/10.1038/d41586-021-03511-5>

References

1. 1.

Katsounaros, I. & Koper, M. T. M. in *Electrochemical Science for a Sustainable Society* (ed. Uosaki, K.) 23–50 (Springer, 2017).

2. 2.

Wang, Y.-H. *et al.* *Nature* **600**, 81–85 (2021).

3. 3.

Kumar, R., Schmidt, J. R. & Skinner, J. L. *J. Chem. Phys.* **126**, 204107 (2007).

4. 4.

Velasco-Velez, J.-J. *et al. Science* **346**, 831–834 (2014).

5. 5.

Tuladhar, A. *et al. J. Am. Chem. Soc.* **142**, 6991–7000 (2020).

6. 6.

Gopalakrishnan, S., Liu, D., Allen, H. C., Kuo, M. & Shultz, M. J. *Chem. Rev.* **106**, 1155–1175 (2006).

7. 7.

Sovago, M. *et al. Phys. Rev. Lett.* **100**, 173901 (2008).

8. 8.

Ataka, K., Yotsuyanagi, T. & Osawa, M. *J. Phys. Chem.* **100**, 10664–10672 (1996).

9. 9.

Schultz, Z. D., Shaw, S. K. & Gewirth, A. A. *J. Am. Chem. Soc.* **127**, 15916–15922 (2005).

10. 10.

Remsing, R. C., McKendry, I. G., Strongin, D. R., Klein, M. L. & Zdilla, M. J. *J. Phys. Chem. Lett.* **6**, 4804–4808 (2015).

11. 11.

Ledezma-Yanez, I. *et al. Nature Energy* **2**, 17031 (2017).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03511-5>

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

- NEWS AND VIEWS
- 01 December 2021

Artificial intelligence aids intuition in mathematical discovery

Machine-learning tools have been used to assist the part of mathematical research that usually relies on human intuition and creativity — leading to two fundamental results in different areas of mathematics.

- [Christian Stump](#) ⁰

Mathematicians have been developing theories by studying examples throughout history. For instance, by looking at a cube and a pyramid, one might realize that the number of vertices, edges and faces are related. A mathematician recognizes such a pattern, extends it to more-general shapes, and then starts to think about why this relationship might hold. Parts of this process involve computations, for which mathematical software has been useful since it first became available in the 1960s. However, human creativity enables mathematicians to instinctively understand where to look for emerging patterns. [Writing in Nature](#), Davies *et al.* now describe a way of using artificial intelligence (AI) techniques to help with the creative core of the mathematical-research process¹.

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**, 44–45 (2021)

doi: <https://doi.org/10.1038/d41586-021-03512-4>

References

1. 1.

Davies, A. *et al.* *Nature* **600**, 70–74 (2021).

2. 2.

Peifer, D., Stillman, M. & Halpern-Leistner, D. *Proc. Mach. Learn. Res.* **119**, 7575–7585 (2020).

3. 3.

Lample, G. & Charton, F. Preprint at <https://arxiv.org/abs/1912.01412> (2019).

4. 4.

He, Y.-H. *The Calabi–Yau Landscape: From Geometry, to Physics, to Machine Learning* (Springer, 2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-03512-4>

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

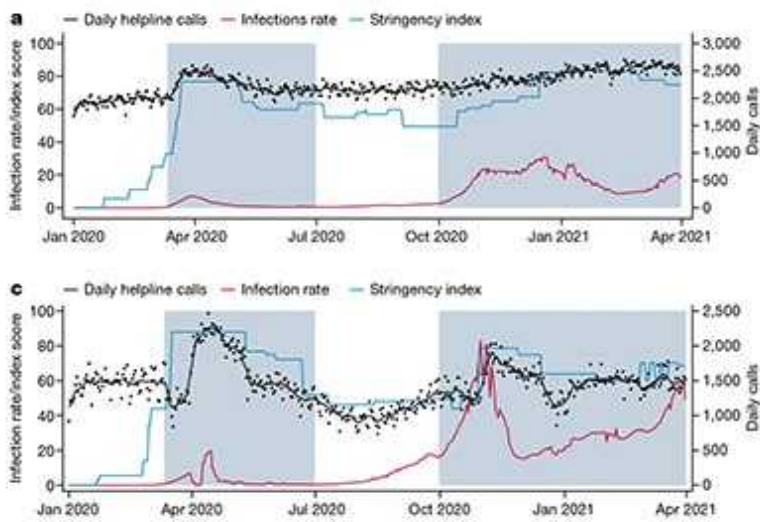
- NEWS AND VIEWS
- 17 November 2021

Helpline data used to monitor population distress in a pandemic

The initial stages of the COVID-19 pandemic saw an increase in calls to mental-health helplines in 19 countries. Helpline-call data can be used to monitor distress at a population level in near-real time.

- [Cindy H. Liu](#) ORCID: <http://orcid.org/0000-0001-6985-5961>⁰ &
- [Alexander C. Tsai](#) ORCID: <http://orcid.org/0000-0001-5106-1354>¹

An important challenge in addressing mental-health problems is that trends can be difficult to detect because detection relies heavily on self-disclosure. As such, helplines — telephone services that provide crisis intervention to callers seeking help — might serve as a particularly useful source of anonymized data regarding the mental health of a population. This profiling could be especially useful during the COVID-19 pandemic, given the potential emergence or exacerbation of mental-health problems¹. Together, the threat of disease to oneself and others that is associated with a local epidemic², the restrictiveness of local non-pharmaceutical interventions (such as stay-at-home orders) and the potential associated loss of income could have contributed to a decline in the mental health of a population while at the same time inhibiting or delaying people's search for help for problems³. [Writing in Nature](#), Brülhart *et al.*⁴ present evidence suggesting that helpline-call data can be used to monitor real-time changes in the mental health of a population — including over the course of the COVID-19 pandemic.



[Read the paper: Mental health concerns during the COVID-19 pandemic as revealed by helpline calls](#)

More so than in other areas of medicine, the stigma that can be associated with mental illness often prevents people from fully disclosing their experiences and feelings to those in their social networks, or even to licensed mental-health-care professionals. Furthermore, although mental illness contributes immensely to the global disease burden, primary health-care providers are overburdened, mental-health systems are underfunded and access to evidence-based treatment remains poor^{5,6}. For these reasons, helplines have, since their introduction in the United Kingdom by Samaritans in 1953, played a key part in providing low- or no-cost, anonymous support to people with unmet acute and chronic mental-health needs around the world.

Brülhart and colleagues updated and expanded on their previous work looking at helpline calls in one country⁷ by assembling data on more than 7 million helpline calls in 19 countries over the course of 2019, 2020 and part of 2021. They found that, within 6 weeks of the start of a country's initial outbreak (defined as the week in which the cumulative number of reported SARS-CoV-2 infections was higher than 1 in 100,000 inhabitants), call volumes to helplines peaked at 35% higher than pre-pandemic levels (Fig. 1). By examining the changes in the proportion of calls relating to different categories, Brülhart and co-workers attribute these increases to fear,

loneliness and concerns about health. The authors also found that suicide-related calls increased in the wake of more-stringent, non-pharmaceutical interventions, but that such calls decreased when income-support policies were introduced. The latter finding is perhaps unsurprising, but is a welcome addition to the evidence base that supports ongoing appeals for financial and other support to mitigate the adverse effects of non-pharmaceutical interventions on uncertainties over employment, income and housing security.⁸

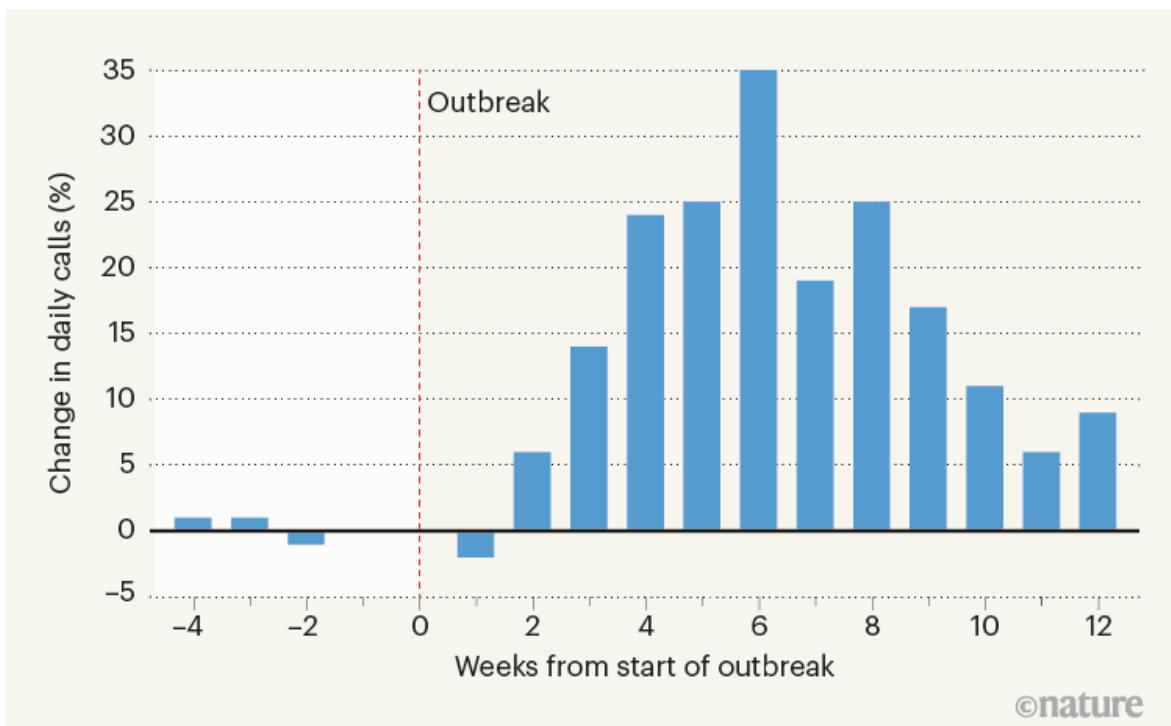
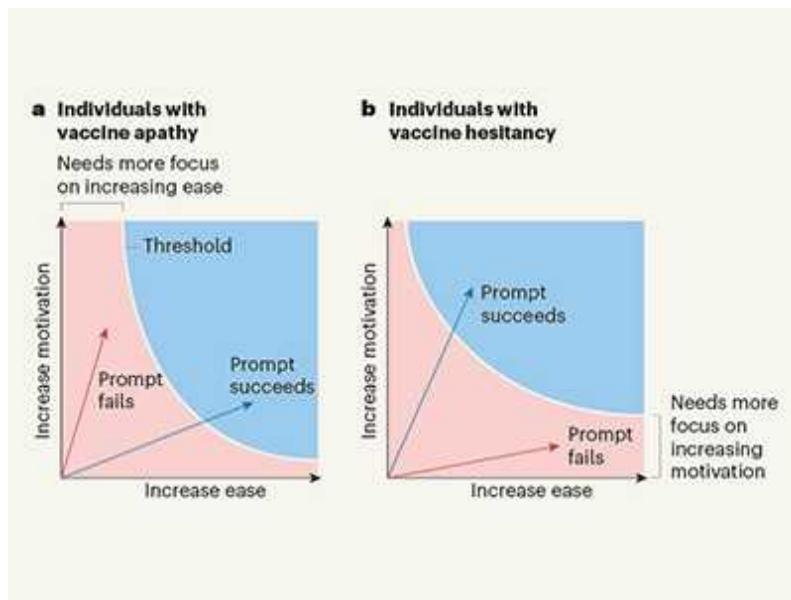


Figure 1 | A surge in helpline calls after local outbreaks of COVID-19. Helpline data can be used to monitor people's mental health in near-real time. Brülhart *et al.*⁴ used helpline data to track changes in mental health at the population level. They gathered anonymized data from 21 helplines across 19 countries. The authors compared the volume of calls in the weeks before and after the start of local COVID-19 outbreaks (defined as the week in which the cumulative number of infections surpassed 1 in 100,000 inhabitants). The number of calls increased by an average of 35% by 6 weeks from the start of the local outbreak, after which the number gradually decreased.

The study by Brülhart *et al.* is among the most comprehensive such analyses of helpline data^{9,10} conducted so far. Other investigators have used data from Internet search-engine queries¹¹ or electronic health records¹² to track mental-health trends in a population. Compared with data on mental-health-related queries in search engines (which might represent information-seeking rather than help-seeking), helpline-call data are more likely to reflect the mental-health concerns of individual callers. And they are not subject to the lag times that might impose limitations on the use of electronic health records for tracking the mental health of a population.

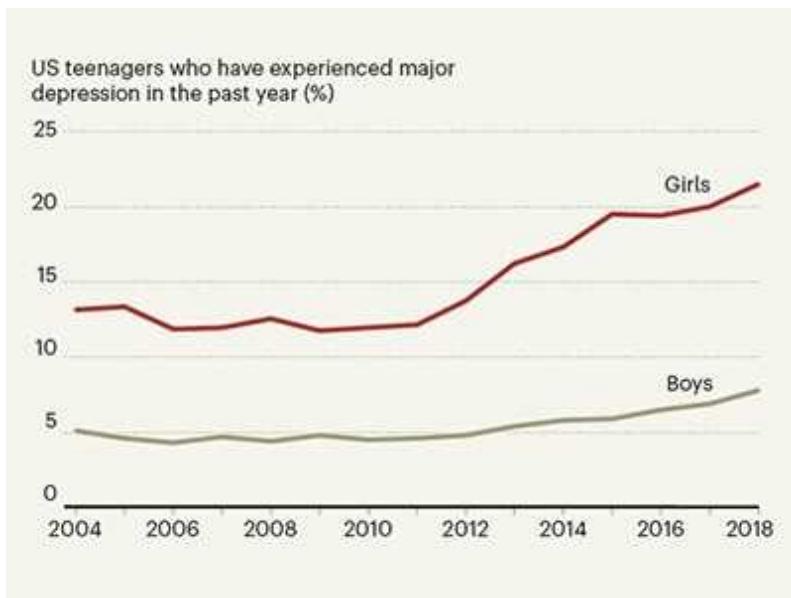
Nevertheless, there are limitations to the use of helpline-call data for monitoring people's mental health. Individuals in emotional distress who choose to call a helpline are unlikely to have a similar mental-health profile to people in emotional distress who do not call helplines or who lack the means to make such a call. Furthermore, for people who do not live alone, stay-at-home restrictions could compromise the privacy that they might need to place a helpline call, particularly in circumstances in which they are at risk of violence from an intimate partner¹³. These factors, which would probably result in underestimates of population mental-health problems, would be likely to reduce the usefulness of helpline-call data for monitoring.



[Text-message nudges encourage COVID vaccination](#)

There are also limitations to relying on the identified topics of the calls, because there might be within- and between-country differences in how the calls are interpreted and categorized. Whether an individual call is categorized as being prompted by ‘a relationship issue’ as opposed to ‘loneliness’ or ‘violence’, for example, is likely to be governed by operator- and culture-specific factors. In addition, the identified topic might not reflect the underlying source of concern, but rather be a topic that the caller wants to discuss with the helpline operator. Furthermore, although helplines are an established service, and call data (such as call topic) can be analysed, less is known about the individuals making the calls, and relatively few studies have examined the characteristics of helpline callers¹⁴. Anonymized data on the characteristics of helpline callers could be correlated with data on the content of helpline calls to further inform mental-health policy, service delivery and programme implementation. In doing this, we would be able to identify the types of individual who might benefit from these intervention or prevention programmes.

Despite these limitations, Brülhart and colleagues’ study improves our understanding of how experiences of the pandemic around the world have led people to reach out for help. Although, for the reasons described above, help-seeking is not a straightforward indicator of distress, the call volumes over time — which showed an initial peak and then a decline to pre-pandemic levels — have a trajectory that is consistent with that seen in longitudinal studies that have assessed distress over time during the pandemic through other methodologies (such as using surveys)^{15,16}. It seems that the convergence of findings from different methods for monitoring mental health in populations can further inform our understanding of how the pandemic has driven both distress and help-seeking behaviour.



Scrutinizing the defects of digital technology on mental health

The rapid changes in the trajectory of the COVID-19 pandemic and its immediate effects on acute distress have revealed the urgent need for real-time monitoring to enable evidence-based, responsive changes in social policies and health-care systems to be implemented and assessed. Public-health officials have relied on the tracking of the spread of the SARS-CoV-2 virus worldwide to guide interventions to prevent disease; however, similar efforts are also urgently needed to address mental-health problems.

Helpline-call data could enable policymakers, mental-health professionals and the general public to determine the best course of action for addressing mental-health concerns and any associated problems. Understanding how policies specifically influence mental health on the basis of location (for example, country) and caller demographic (including age, gender and race) could enable policies to be tailored and optimized.

This study serves as an excellent example of a data resource that it is hoped will spur researchers to assess the promise of applying other data sources (such as Internet applications and text- or Internet-based messaging services) to monitor population mental health in other countries worldwide.

Nature **600**, 46-47 (2021)

doi: <https://doi.org/10.1038/d41586-021-03038-9>

References

1. 1.
McGinty, E. E., Presskreischer, R., Han, H. & Barry, C. L. *JAMA* **324**, 93–94 (2020).
2. 2.
Ellwardt, L. & Präg, P. *Sci. Rep.* **11**, 15958 (2021).
3. 3.
Hoyer, C. *et al.* *Eur. Arch. Psychiatry Clin. Neurosci.* **271**, 377–379 (2021).
4. 4.
Brülhart, M., Klotzbücher, V., Lalive, R. & Reich, S. K. *Nature* **600**, 121–126 (2021).
5. 5.
Saxena, S., Thornicroft, G., Knapp, M. & Whiteford, H. *Lancet* **370**, 878–889 (2007).
6. 6.
Tsai, A. C. & Tomlinson, M. *PLoS Med.* **12**, e1001846 (2015).
7. 7.
Brülhart, M. & Lalive, R. *COVID Econ.* **19**, 143–158 (2020).
8. 8.
Sandoval-Olascoaga, S., Venkataramani, A. S. & Arcaya, M. C. *JAMA Netw. Open* **4**, e2129041 (2021).

9. 9.

Turkington, R. *et al. JMIR Mental Health* **7**, e22984 (2020).

10. 10.

Titov, N. *et al. Internet Interv.* **21**, 100327 (2020).

11. 11.

Arora, V. S., Stuckler, D. & McKee, M. *Publ. Health* **137**, 147–153 (2016).

12. 12.

Holland, K. M. *et al. JAMA Psychiatry* **78**, 372–379 (2021).

13. 13.

Leslie, E. & Wilson, R. *J. Publ. Econ.* **189**, 104241 (2020).

14. 14.

Middleton, A., Gunn, J., Bassilios, B. & Pirkis, J. *J. Telemed. Telecare* **20**, 89–98 (2014).

15. 15.

Pierce, M. *et al. Lancet Psychiatry* **8**, 610–619 (2021).

16. 16.

Riehm, K. E. *et al. Ann. Behav. Med.* **55**, 93–102 (2021).

- Article
- Open Access
- [Published: 19 October 2021](#)

A dry lunar mantle reservoir for young mare basalts of Chang'e-5

- [Sen Hu](#) ORCID: orcid.org/0000-0001-9813-5330¹,
- [Huicun He](#) ORCID: orcid.org/0000-0002-7852-6147¹,
- [Jianglong Ji](#) ORCID: orcid.org/0000-0003-2170-3349¹,
- [Yangting Lin](#) ORCID: orcid.org/0000-0002-3407-4329¹,
- [Hejiu Hui](#) ORCID: orcid.org/0000-0003-2707-4977^{2,3},
- [Mahesh Anand](#) ORCID: orcid.org/0000-0003-4026-4476^{4,5},
- [Romain Tartèse](#) ORCID: orcid.org/0000-0002-3490-9875⁶,
- [Yihong Yan](#) ORCID: orcid.org/0000-0002-0187-6069¹,
- [Jialong Hao](#) ORCID: orcid.org/0000-0003-4898-1023¹,
- [Ruiying Li](#) ORCID: orcid.org/0000-0001-9708-6901¹,
- [Lixin Gu](#) ORCID: orcid.org/0000-0002-1254-8693¹,
- [Qian Guo](#) ORCID: orcid.org/0000-0003-4509-8876⁷,
- [Huaiyu He](#) ORCID: orcid.org/0000-0002-4887-5274⁷ &
- [Ziyuan Ouyang](#)⁸

Nature volume 600, pages 49–53 (2021)

- 5767 Accesses
- 3 Citations
- 144 Altmetric
- [Metrics details](#)

Subjects

- [Geochemistry](#)
- [Petrology](#)

Abstract

The distribution of water in the Moon's interior carries implications for the origin of the Moon¹, the crystallization of the lunar magma ocean² and the duration of lunar volcanism². The Chang'e-5 mission returned some of the youngest mare basalt samples reported so far, dated at 2.0 billion years ago (Ga)³, from the northwestern Procellarum KREEP Terrane, providing a probe into the spatiotemporal evolution of lunar water. Here we report the water abundances and hydrogen isotope compositions of apatite and ilmenite-hosted melt inclusions from the Chang'e-5 basalts. We derive a maximum water abundance of $283 \pm 22 \mu\text{g g}^{-1}$ and a deuterium/hydrogen ratio of $(1.06 \pm 0.25) \times 10^{-4}$ for the parent magma. Accounting for low-degree partial melting of the depleted mantle followed by extensive magma fractional crystallization⁴, we estimate a maximum mantle water abundance of $1\text{--}5 \mu\text{g g}^{-1}$, suggesting that the Moon's youngest volcanism was not driven by abundant water in its mantle source. Such a modest water content for the Chang'e-5 basalt mantle source region is at the low end of the range estimated from mare basalts that erupted from around 4.0 Ga to 2.8 Ga (refs. ^{5,6}), suggesting that the mantle source of the Chang'e-5 basalts had become dehydrated by 2.0 Ga through previous melt extraction from the Procellarum KREEP Terrane mantle during prolonged volcanic activity.

[Download PDF](#)

Main

Water abundance in the lunar mantle places strict constraints on high-temperature processes, including the Moon-forming giant impact¹, the ensuing crystallization of the lunar magma ocean⁷ and the longevity of volcanism on the Moon². On the basis of analyses carried out since the Apollo missions, the Moon was long thought to be anhydrous. Advances in in situ analytical techniques over the past decade have allowed the analysis of water abundances at the microscale in various lunar samples, including in olivine-hosted and pyroxene-hosted melt inclusions in mare basalts^{8,9,10,11,12}, apatite in mare basalts and highlands samples^{13,14,15,16,17,18,19,20}, pyroclastic glass beads^{21,22} and anorthosites^{23,24}. The estimates of water abundances for the mantle source regions of these samples span a wide range, from about $0.3 \mu\text{g g}^{-1}$ to $200 \mu\text{g g}^{-1}$ (ref. ²⁵), which suggests that the lunar interior is not as anhydrous as once thought. Although the large variations in the water abundance estimates for lunar mantle sources could partially be due to the assumptions involved in these calculations, many questions remain regarding the origin(s) and distribution of water in the Moon's interior^{25,26}. Variations in estimated mantle water abundances may be indicative of geographic and/or temporal diversity as these samples were collected

from different regions and crystallized between around 4.0 billion years ago (Ga) to 2.8 Ga (refs. [5,6](#)). Hence, studying additional sample collections of younger basalts from different regions can provide critical additional constraints on the spatiotemporal evolution of water in the Moon. The large range of mantle water abundance estimates could also be affected by the mixing of endogenous water with various exogenic water sources, that is, asteroids, comets and solar wind^{[19,26,27](#)}, and/or by the interplay between many processes, including volatile degassing, partial melting, fractional crystallization, impacting, mixing with potassium (K), rare-earth elements (REE) and phosphorus (P) (KREEP)-rich components, and spallation^{[11,15,25,27,28,29](#)}. It is thus crucial to combine *in situ* analysis of water abundances and hydrogen isotope compositions with detailed contextual petrographic information.

The Chang'e-5 (CE5) mission successfully returned 1.731 kg of lunar soil samples from young mare basalt units dated at 2.0–1.2 Ga using crater-counting chronology^{[30,31](#)}. These CE5 samples have now been precisely dated at $2,030 \pm 4$ million years ago (Ma) using the lead (Pb)–Pb isotope isochron technique^{[3](#)}. The CE5 basalts are thus much younger than the youngest lunar basalt samples dated so far (2.8 Ga (ref. [5](#))). The young basalt unit is located in northwestern Oceanus Procellarum, at the northwestern edge of the Procellarum KREEP Terrane (PKT), which is far from all landing sites of the Apollo and Luna missions (Extended Data Fig. [1](#)). The PKT region is also thought to have enhanced concentrations of two major radioactive heat-producing elements, uranium (U) and thorium (Th), and other incompatible elements^{[32](#)}. Water behaves as a typical incompatible element during magmatic processes^{[33](#)} and thus is expected to be enriched in the PKT as well. Hence, the CE5 basalts provide a unique opportunity to constrain the water inventory of a newly sampled region of the Moon's interior, and thus yields crucial information to account for the prolonged activity of lunar magmatism.

Water in CE5 apatite and melt inclusions

We studied a total of 23 basalt clasts (0.2–1.5 mm in size) from two CE5 soil samples (CE5C0100YJFM00103, about 1 g; CE5C0400YJFM00406, about 2 g; Extended Data Table [1](#)). The basalt clasts show variable textures, including subophitic, poikilitic and equigranular, similar to those observed for other basalt clasts in CE5 soil samples^{[3,4](#)}, and are mainly composed of pyroxene and plagioclase with less abundant olivine and ilmenite (Fig. [1](#), Supplementary Figs. [1, 2](#)). These basalt clasts are probably representative of various portions of the same lava flow, based on their identical mineral chemistry and geochemistry^{[4](#)} and their well defined Pb–Pb isochron^{[3](#)}. The textures of ilmenite in the clasts indicate that this phase began to crystallize early from the melt and continued until the last stages of melt evolution (Supplementary Fig. [1](#)). Ilmenite-hosted melt inclusions are 4–50 μm in diameter and show post-entrainment

crystallization textures (0–52%) (Fig. 1, Extended Data Table 3, Supplementary Fig. 1). Apatite occurs as euhedral grains (mostly less than 10 µm) mainly in the fine-grained interstitial materials, with a few euhedral crystals enclosed in the margins of pyroxene (Fig. 1) and iron(II) oxide (FeO)-rich olivine (Supplementary Fig. 2). Apatite is the main hydroxyl (OH)-bearing phase, fluorine (F) rich and chlorine (Cl) poor, similar to those from Apollo mare basalts (Supplementary Fig. 6), and is an accessory phase in these basalt clasts comprising approximately 0.4 vol% modal abundance (Supplementary Table 1). Eight ilmenite-hosted melt inclusions and several apatite grains were located and selected for in situ analysis (Fig. 1, Extended Data Table 1, Supplementary Figs. 1, 2). The water abundances and hydrogen isotope compositions of the ilmenite-hosted melt inclusions, apatite and nominally anhydrous clinopyroxene were measured using a nanoscale secondary ion mass spectrometer (NanoSIMS) instrument ([Methods](#)).

Fig. 1: Backscattered electron images of ilmenite-hosted melt inclusions and apatite from a CE5 basalt clast.

 [figure1](#)

- a**, The basalt clast (406-010,023) embedded in the metal mount is mainly composed of olivine (Ol), pyroxene (Px), plagioclase (Pl) and ilmenite (Ilm), with minor fayalite (Fa), troilite (Tro), spinel (Sp), apatite (Ap) and silica (Q). The locations of **b**, **c** and **d** are outlined by the rectangles in **a**. **b**, High-resolution BSE image of a melt inclusion hosted in ilmenite. This melt inclusion shows a post-entrainment crystallization texture

with occurrences of submicrometre-sized pyroxene and merrillite (yellow arrows). **c**, High-resolution BSE image of apatite in the interstitial areas. Many euhedral apatite grains, up to 10 µm in length, coexist with fine-grained plagioclase, fayalite and mesostasis (Mesos). **d**, Small apatite grains also occur at the rims of pyroxene, coexisting with mesostasis. The bright pits in cracks are the remains of the gold coating.

The majority of apatite grains contain water abundances ranging from $555 \pm 31 \mu\text{g g}^{-1}$ to $4,856 \pm 217 \mu\text{g g}^{-1}$ (average $1,921 \pm 910 \mu\text{g g}^{-1}$, 1σ) with δD values ranging from $275 \pm 85\text{\textperthousand}$ to $1,022 \pm 87\text{\textperthousand}$ (average $578 \pm 208\text{\textperthousand}$, 1σ) ($\delta\text{D} = 1,000 \times ([\text{D}/\text{H}_{\text{sample}}]/[\text{D}/\text{H}_{\text{standard}}] - 1$, where D is deuterium and H is hydrogen), using Vienna standard mean ocean water as the standard) (Fig. 2, Extended Data Table 2), which overlap with apatite water abundances and δD values measured in high-titanium (Ti) and low-Ti Apollo basalts^{11,15,16,17,28,29,34,35,36,37}. Three apatite analyses yielded lower water abundances ($110 \pm 13 \mu\text{g g}^{-1}$ to $235 \pm 19 \mu\text{g g}^{-1}$; Extended Data Table 2), with corresponding δD values indistinguishable from the majority of other analyses. As apatite is the major water-bearing phase in the CE5 basalts, a water abundance of $7 \pm 3 \mu\text{g g}^{-1}$ for the bulk composition of the CE5 basalts was calculated from the average water content of apatite and its modal abundance of approximately 0.4 vol% (Methods). It is noted that this water abundance is not the original water abundance in the CE5 basaltic magma before eruption, but represents the residual water abundance after magma degassing at the time of apatite crystallization¹⁴. Furthermore, the apatite δD values reflect the signature of the last residual melt after precipitation of most constituent minerals, and the observed large deuterium-enrichment is probably the result of degassing of hydrogen-bearing species from the melt, mostly in the form of H₂ under the reducing conditions that are typical for lunar volcanic products¹⁵.

Fig. 2: Water abundance and δD of apatite and ilmenite-hosted melt inclusions from CE5 basalts.

 figure2

The majority of melt inclusions show a negative correlation between the water abundances and δD values, except for three melt inclusions with higher δD values plotting close to the range of apatite. The dotted lines indicate a three-stage evolutionary path, starting with 2–3% partial melting of the mantle source of CE5 basalts, followed by 43–88% fractional crystallization (stage 1), H₂ degassing from the basaltic melts accompanied by deuterium enrichment (stage 2) and crystallization of apatite from the residual melts, possibly accompanied by further H₂ degassing (stage 3). Apatite and melt inclusion (olivine hosted or pyroxene hosted) data from Apollo samples (Supplementary Table 5) are shown for comparison. The CE5 data have been corrected for a nominal CRE of 50 Ma ([Methods](#)). The error bars are 2σ .

The ilmenite-hosted melt inclusions contain water abundances of $6 \pm 2 \mu\text{g g}^{-1}$ to $370 \pm 21 \mu\text{g g}^{-1}$ after correction of post-entrapment crystallization, and show a wide range of δD values from $-330 \pm 190\text{\textperthousand}$ to $869 \pm 230\text{\textperthousand}$ after correcting for the effects of cosmic ray spallation (Fig. 2, Extended Data Table 3). Cosmic-ray spallation mainly produces deuterium, and can have a large effect on δD values, especially for water-poor (less than $30 \mu\text{g g}^{-1}$) melt inclusions³⁸. Cosmic-ray exposure (CRE) ages determined for

various Apollo samples are mostly younger than around 200 Ma (ref. [39](#)), but have not yet been measured for CE5 samples. We have modelled the spallation effects on δD values of the melt inclusions, using CRE ages of 10 Ma, 50 Ma, 100 Ma and 200 Ma (Extended Data Fig. [4](#), Supplementary Table [4](#)). Using CRE ages of 100 Ma and 200 Ma yields noticeable over-correction of δD values as the resulting values are even more deuterium depleted than the currently accepted hydrogen isotope composition of the lunar mantle (Extended Data Fig. [4](#)). As determined from Apollo samples, lunar regolith from a depth of about 9 mm is thought to overturn at least once in about 10 Myr (ref. [40](#)), which suggests that it is reasonable to assume a CRE age of around 50 Ma for CE5 samples. With a 50-Ma CRE age correction, the melt inclusions with the lowest water abundances yield corrected δD values of $200 \pm 390\text{\textperthousand}$ that overlap with the lowest δD value measured for apatite. Importantly, this correction does not greatly affect the δD values of water-rich melt inclusions nor those of apatite grains (Extended Data Fig. [4](#), Supplementary Table [4](#)). Moreover, spallation by cosmic rays has little effect on water abundances (Extended Data Fig. [4](#), Supplementary Table [4](#)). After correction for spallation, the melt inclusions with $\delta D \leq 200\text{\textperthousand}$ show a negative correlation between water abundances ($6 \pm 2 \mu\text{g g}^{-1}$ to $283 \pm 22 \mu\text{g g}^{-1}$) and δD values ($-330 \pm 190\text{\textperthousand}$ to $200 \pm 390\text{\textperthousand}$). In addition, three melt inclusion analyses with higher δD values ($271 \pm 124\text{\textperthousand}$ to $869 \pm 230\text{\textperthousand}$) overlap with the data for the water-poor apatite grains (Fig. [2](#), Extended Data Table [3](#)). These observations provide convincing evidence that ilmenite-hosted melt inclusions with $\delta D \leq 200\text{\textperthousand}$ have recorded the progressive evolution of melts undergoing H_2 degassing, resulting in considerable deuterium enrichment during crystallization of the CE5 basalts [15,41](#). Hydrogen diffusion out of melt inclusions is another process by which D/H ratios can be fractionated as reported for melt inclusions in olivine and pyroxene from Apollo basalts [11](#). At present, there is no constraint on the diffusion rate of water in ilmenite-hosted melt inclusions. The lowest δD value of approximately $-330\text{\textperthousand}$ measured in ilmenite-hosted melt inclusions suggests that the D/H ratios have been well preserved in the melt inclusions without significant exchange of hydrogen isotopes with the deuterium-enriched residual melt through diffusion.

Magmatic history and source

The lowest δD values ($-330 \pm 190\text{\textperthousand}$) of the melt inclusions are within the ranges of δD values for most types of chondrite and the lunar mantle δD ($0 \pm 200\text{\textperthousand}$) estimated from analyses of various lunar samples [11,16,18,19,20,22,24](#). This similarity suggests that the melt with the lowest δD was trapped in the early stages of magma crystallization before substantial degassing of water in the form of H_2 (refs. [11,15,41](#)). In contrast, the melt inclusions with higher δD (more than $270\text{\textperthousand}$) contain substantial water and overlap with the water-poor apatite values (Fig. [3](#)). This observation can be explained

by late crystallization of these ilmenite grains, when deuterium-enriched water was concentrated in the residual melt and apatite started crystallization.

Fig. 3: Variation of lunar mantle water abundance estimates with time.

figure3

The maximum mantle water abundance at 2.03 Ga, which was estimated in this study using CE5 basalts, plots at the lower end of mantle water abundance estimates for Apollo samples and lunar meteorites formed between around 4.0 Ga and 2.8 Ga. All data are plotted as average values with the error bars representing the ranges of estimates. Estimates for the water abundances in the last dregs of the lunar magma ocean (urKREEP)²⁵ and Earth's primitive mantle⁵⁰ are shown for comparison. The red and dark blue solid lines represent average values estimated from CE5 basalts and previous lunar samples, respectively. It is noted that the vertical axis is a log scale. Literature data are provided in Extended Data Table 5.

The most deuterium-depleted melt inclusions, as discussed, probably captured signatures of the parent magma before notable loss of water by degassing in the form of H₂. Hence, the high water abundance ($283 \pm 22 \mu\text{g g}^{-1}$; Extended Data Table 3) of these melt inclusions can be regarded as the maximum water content of the parent

magma. Alternatively, the water abundance of the parent magma can also be estimated from a water content of $7 \pm 3 \mu\text{g g}^{-1}$ for the bulk CE5 basalts through calibration for degassing loss of 98–99% water in the form of H₂ based on the accompanying δD increasing from −330‰ to the average apatite δD of $578 \pm 208\text{‰}$ (1σ) (Extended Data Table 2, [Methods](#)). This calculation yields an estimate for the water abundance of the parent magma of $600 \pm 400 \mu\text{g g}^{-1}$, which is consistent within errors with that based on the deuterium-depleted melt inclusions. We thus use the better constrained water abundance from the deuterium-depleted melt inclusions ($283 \pm 22 \mu\text{g g}^{-1}$; Extended Data Table 3) as the maximum water abundance of the parent basaltic magma.

The CE5 parental magma was derived from a depleted lunar mantle source not associated with a KREEP component, based on its low initial μ value (²³⁸U/²⁰⁴Pb ratio; 680 ± 20)³, low initial ⁸⁷Sr/⁸⁶Sr ratio (0.69934 to 0.69986) and high positive $\varepsilon_{\text{Nd}}(t)$ value (7.9 to 9.3)⁴. $\varepsilon_{\text{Nd}}(t) = ((^{143}\text{Nd}/^{144}\text{Nd})_{\text{sample}(t)} / (^{143}\text{Nd}/^{144}\text{Nd})_{\text{CHUR}} - 1) \times 10,000$, where $(^{143}\text{Nd}/^{144}\text{Nd})_{\text{sample}(t)}$ and $(^{143}\text{Nd}/^{144}\text{Nd})_{\text{CHUR}}$ are the Nd isotopic compositions of the sample at its formation time (t) and the chondritic uniform reservoir, respectively. The highly elevated abundances of REE and Th, and high FeO and moderate titanium dioxide (TiO₂) concentrations of the CE5 parent magma match a model of low-degree (2–3%) partial melting followed by moderate-to-extensive (43–88%) fractional crystallization⁴. Accordingly, the maximum water concentration in the lunar mantle source beneath the CE5 landing site can be estimated at $1\text{--}5 \mu\text{g g}^{-1}$, corresponding to a maximum water abundance of approximately $280 \mu\text{g g}^{-1}$ in the derived parent magma.

Our analyses of apatite and melt inclusions allow the evolution of CE5 basalts to be divided into three stages. In stage 1, the mantle source region underneath the CE5 landing site with $1\text{--}5 \mu\text{g g}^{-1}$ water experienced low-degree (2–3%) partial melting followed by moderate-to-extensive-degree (43–88%) fractional crystallization⁴, thereby generating a parent magma with approximately $280 \mu\text{g g}^{-1}$ water and preserving its mantle-derived δD of −330‰. This maximum water abundance recorded in the melt inclusions captured by the earliest-formed ilmenite analysed here yields our best estimate for the water abundance of the parent magma. In stage 2, H₂ degassing from the parent magma occurred during its ascent to shallower depths and during eruption, and was accompanied by the crystallization of ilmenite that entrapped melts at various stages of magmatic evolution. Extensive H₂ degassing^{15,40} could have occurred in the reduced lunar environment^{42,43}, resulting in appreciable D/H fractionation from −330‰ up to ~300‰. In stage 3, apatite crystallized from the residual melts that became enriched in water, halogens and other incompatible species, after most nominally anhydrous silicates and ilmenite had formed.

Evolution of lunar mantle water

The maximum water abundance of $1\text{--}5 \mu\text{g g}^{-1}$ estimated for the mantle source of CE5 basalts is notably at the lower end of the mantle water abundances derived from Apollo basalts and lunar meteorites^{9,10,11,15,25} (Fig. 3). These new indications of a water-poor lunar mantle at 2 Ga carry important implications for understanding late volcanism on the Moon. Such a water-poor mantle source for CE5 basalts excludes the possibility that a high water abundance in the lunar mantle reservoir, by lowering its melting point, could be one of the main causes of the prolonged volcanic activity in this part of the PKT.

Our observations indicate that the water abundance in the Moon's interior may have to some extent decreased from 4.0–2.8 Ga to 2.0 Ga (Fig. 3). Such a systematic loss of water over time could be the result of prolonged magmatic activity in the PKT, where multiple water-bearing melt-extraction episodes from the PKT mantle reservoir occurred but did not fractionate D/H significantly⁴⁴. In the northwestern PKT region, in close proximity to the CE5 landing site, up to ten basaltic units ranging in crater-counting age from 3.7 Ga to 1.2 Ga have been identified⁴⁵, although it is difficult to be certain that all these units were derived from the same mantle source region. Nevertheless, such a dehydration partial melting process has also been observed in Earth's mantle^{46,47}.

Alternatively, the wide range of the water abundance estimates for the mantle source regions of all studied lunar basalts may reflect a heterogeneous distribution of water in the Moon's interior. Furthermore, the water abundance estimates could be affected by possible contamination of some volcanic products by KREEP components during either magma transport or in their mantle source regions during convective overturn of the lunar magma ocean^{48,49}. However, it has been demonstrated that CE5 basalts were not contaminated by KREEP components⁴.

Thus, our estimate of the mantle water abundance based on CE5 basalts in the PKT region provides a critical spatiotemporal constraint on the distribution of water in the lunar interior. Nonetheless, it remains an enigma to explain the basaltic volcanism sustained as late as 2.0 Ga at the CE5 landing site, the mantle source of which is depleted in the heat-producing elements, U, Th and K, relative to the bulk silicate Moon⁴ and is water poor.

Methods

Sample preparation

Two CE5 lunar soils (CE5C0100YJFM00103, about 1 g; CE5C0400YJFM00406, about 2 g) allocated by the China National Space Administration were used in this study. Both samples were scooped by the robotic arm of the CE5 lander and separated into different packages in the ultraclean room at the extraterrestrial sample curation centre of the National Astronomical Observatories, Chinese Academy of Sciences. Approximately 240 soil fragments with grain sizes varying from about 100 µm to about 1 mm were sieved and hand-picked under a binocular microscope in the ultraclean room at the Institute of Geology and Geophysics, Chinese Academy of Sciences (IGGCAS). Then, about two-thirds of the picked grains were prepared as eight tin–bismuth metal–alloy mounts following the method of Zhang et al.⁵¹ and the other third was mounted in epoxy and prepared into three polished thin sections (about 100 µm in thickness). The polished metal mounts and thin sections were cleaned using ultrapure water and anhydrous ethanol and then dried at 70 °C in a baking oven overnight. The details of apatite and ilmenite-hosted melt inclusions from 23 CE5 basalt clasts and fragments are summarized in Extended Data Table 1.

Scanning electron microscope observation

Petrographic observations and elemental mapping were carried out using field emission scanning electron microscopes (FE-SEMs) using FEI Nova NanoSEM 450 and ThermoFisher Apreo instruments at the IGGCAS, using electron beam currents of 2 nA to 3.2 nA and an acceleration voltage of 15 kV. Energy dispersive spectroscopy X-ray maps were collected for each basaltic clast to locate P-bearing phases. The phosphates were then observed at higher magnification in back-scattered electron (BSE) images. The modal abundance of apatite from various CE5 basalt clasts were counted by the exposed surface areas (Supplementary Table 1). The prepared sections were initially coated with gold to identify apatite and melt inclusions for in situ NanoSIMS measurement of the water content and hydrogen isotopes. After NanoSIMS measurement, the samples were re-coated with carbon and observed by SEM to confirm the positions of the NanoSIMS spots.

Electron probe microanalysis

After the NanoSIMS analyses, we used a JEOL JXA-8100 electron probe micro-analyser (EPMA) at the IGGCAS to quantify the major and minor elemental abundances in phosphates, melt inclusions in ilmenite and associated mafic minerals (that is, clinopyroxene, olivine, plagioclase and ilmenite). The samples were coated with carbon. The operating acceleration voltage was 15 kV and the beam current was 20 nA. The EPMA analyses were carried out after the NanoSIMS measurements to avoid possible H loss due to bombardment by the electron beam³⁵. The EPMA standards were natural albite (sodium (Na) and aluminium (Al)), bustamite (manganese (Mn)), diopside (calcium (Ca), silicon (Si) and magnesium (Mg)), apatite

(P), K-feldspar (K), tugtupite (Cl), synthetic fluorite (F), rutile (Ti), iron(III) oxide (Fe_2O_3 ; Fe), vanadium pentoxide (V_2O_5 ; V), nickel oxide (NiO ; Ni) and chromium(III) oxide (Cr_2O_3 ; Cr). Na, K, F and Cl were first measured to minimize possible loss of volatiles by electron beam irradiation. The detection limits were (1σ) 0.01 wt% for Cl and sulfur (S), 0.02 wt% for Na, Mg, Al, Cr, K, Si, Mn, Ca and Fe, 0.03 wt% for F, barium (Ba), Ni and Ti, and 0.04 wt% for P. A program based on the ZAF (Z, atomic number; A, absorption of X-rays in the specimen; F, fluorescence caused by other X-rays generated in the specimen) procedure was used for data correction. The EPMA data obtained for apatite, melt inclusions in ilmenite and the coexisting silicates are listed in Supplementary Table 2.

In situ water abundance and hydrogen isotope analysis

Apatite and melt inclusions

The hydrogen isotopes and water contents of apatite and melt inclusions enclosed in ilmenite from the CE5 basaltic clasts were measured with a CAMECA NanoSIMS 50L at IGGCAS. The samples were coated with gold, loaded in sample holders together with the standards and baked overnight at about 60 °C in the NanoSIMS airlock. The holders were then stored in the NanoSIMS sample chamber to improve the vacuum quality and minimize the H background^{52,53,54}. The vacuum pressure in the analysis chamber was 2.8×10^{-10} torr to 3.0×10^{-10} torr during analysis. Each $15 \mu\text{m} \times 15 \mu\text{m}$ analysis area was pre-sputtered for 2 min with a Cs^+ ion beam current of 2 nA to remove the surface coating and potential contamination. During analysis, the secondary anions ${}^1\text{H}^-$, ${}^2\text{D}^-$ and ${}^{12}\text{C}^-$ were simultaneously counted by electron multipliers and ${}^{16}\text{O}^-$ by a Faraday cup from the central $3 \mu\text{m} \times 3 \mu\text{m}$ areas using the NanoSIMS blanking technique. A 44-ns dead time was corrected for all electron multipliers, while the electron multiplier noise (<10⁻² counts per second) was ignored. We used a primary ion beam current of about 0.5 nA for analysis, corresponding to a beam size of about 500 nm in diameter. The charging effect on the sample surface was compensated for by an electron gun during analysis.

A chip of the anhydrous San Carlos olivine reference with a reported water content of $1.4 \mu\text{g g}^{-1}$ (ref. ⁵⁵) was used for H background (bg) corrections, following the relationship: $\text{H/O}_{\text{bg}} = (\text{H}_{\text{counts}} - \text{H}_{\text{bg}})/\text{O}_{\text{counts}}$ and $\text{D/H}_{\text{measured}} = (1 - f) \times \text{D/H}_{\text{true}} + f \times \text{D/H}_{\text{bg}}$, where O is oxygen and f is the proportion of H emitted from the instrumental background⁵⁶. Here D/H_{bg} was $3.36 (\pm 0.55) \times 10^{-4}$ and $\text{H}_{\text{bg}} = 689 \pm 139$ counts per second (2 s.d., N=11, corresponding to a water background abundance of $25 \pm 8 \mu\text{g g}^{-1}$ (2 s.d.)). After background subtraction, the water abundances of apatite grains and melt inclusions were calculated from the background-subtracted H/O ratios multiplied

by the slope of the calibration line (Extended Data Fig. 2), which were determined by measuring two apatite standards, Durango apatite (water (H_2O) = 0.0478 wt% and $\delta\text{D} = -120 \pm 5\%$ ^{29,57}) and Kovdor apatite ($\text{H}_2\text{O} = 0.98 \pm 0.07$ wt% and $\delta\text{D} = -66 \pm 21\%$ ⁵⁸), the southwest Indian French transect mid-ocean ridge basalt (SWIFT MORB) glass ($\text{H}_2\text{O} = 0.258$ wt% and $\delta\text{D} = -73 \pm 2\%$), and two basaltic glasses, 519-4-1 ($\text{H}_2\text{O} = 0.17$ wt%)⁵² and 1833-11 ($\text{H}_2\text{O} = 1.2$ wt%)⁵² (Supplementary Table 3). Corrections for instrumental mass fractionation (IMF) on hydrogen isotopic compositions of both apatite and melt inclusions were conducted using the Kovdor apatite standard, and monitored by analysing both the Durango apatite and SWIFT MORB glass standards during the whole analytical session (Extended Data Fig. 3). The matrix effects on water abundance and IMF on hydrogen isotopic composition are the same between apatite and silicate glass within analytical uncertainties⁵⁴. Hydrogen isotopic compositions are given using the delta notation, $\delta\text{D} = ((\text{D}/\text{H})_{\text{sample}}/(\text{D}/\text{H})_{\text{SMOW}}) - 1) \times 1,000$, where SMOW is the standard mean ocean water with a D/H ratio of 1.5576×10^{-4} . More technical details can be found in Hu et al.^{53,54}. All data are reported with their 2σ uncertainties that include the reproducibility of the D/H measurements on the reference materials, the uncertainty of the H_2O background subtraction and the internal precision on each analysis (Extended Data Tables 2, 3, Supplementary Table 4). The raw measured D/H ratios were corrected for the background, followed by correction for IMF.

Clinopyroxene

The water abundance of clinopyroxene from the CE5 basaltic clasts was measured with the CAMECA NanoSIMS 50L using an identical instrument setup to that described above. We used a higher Cs^+ primary beam current of 7 nA to improve the ${}^1\text{H}^-$ counts on clinopyroxene and reduce the background. Each $25 \mu\text{m} \times 25 \mu\text{m}$ analysis area was pre-sputtered for about 2 min with the same analytical beam current to remove the surface coating and potential contaminations. The secondary ion signals from the central $7 \mu\text{m} \times 7 \mu\text{m}$ areas were counted with 50% blanking of outermost regions. San Carlos olivine ($\text{H}_2\text{O} = 1.4 \mu\text{g g}^{-1}$ (ref. 55)) was used to determine the instrumental background of H_2O , which was $5 \pm 2 \mu\text{g g}^{-1}$, about five times lower than the H_2O background during apatite and melt inclusions analyses ($25 \pm 8 \mu\text{g g}^{-1}$). The analytical results are listed in Extended Data Table 4.

Correction of water abundances and D/H ratios for spallation effects

The measured D/H ratios have been corrected for the potential effects of cosmic-ray spallation, using a deuterium production rate of $2.17 \times 10^{-12} \text{ mol D g}^{-1} \text{ Ma}^{-1}$ (ref. 59)

for melt inclusions and 9.20×10^{-13} mol D g⁻¹ Ma⁻¹ (ref. ⁶⁰) for apatite. The correction errors induced by deuterium spallation are around 50% on δ D and negligible on water content²¹. The CRE ages determined for most Apollo samples are younger than about 200 Ma (ref. ³⁹), although some ages are up to 400 Ma (ref. ⁵⁶). As no CRE age is yet available for the CE5 basaltic clasts, we modelled the effects of corrections for CRE ages of 10 Ma, 50 Ma, 100 Ma and 200 Ma (Extended Data Table 3, Extended Data Fig. 4, Supplementary Table 4). The corrected δ D values for the melt inclusions with low water abundances appear to be overcorrected for CRE ages of 100 Ma and 200 Ma, as indicated by unusually low δ D values. The corrected δ D values for the low-water (less than 50 $\mu\text{g g}^{-1}$) melt inclusions, using a CRE age of 50 Ma, overlap with the lowest δ D values of apatite (approximately 300‰; Fig. 2). This observation is consistent with the late capture of the low-water melt inclusions by ilmenite followed by the crystallization of the high-deuterium apatite. The uncertainty in the CRE age mainly affects the negative correlation between δ D values and water abundances of the low-deuterium melt inclusions, with only minor effects on estimating the maximum water abundances for the CE5 parent magma and the mantle source.

Degassing modelling

The hydrogen isotope fractionation during volatile loss into a vacuum is given by $\alpha^2 = M_1/M_2$, where M₁ and M₂ are the masses of the volatile phase isotopologues. The change of the isotopic composition of H during volatile loss by Rayleigh fractionation is given by $R = R_0 \times f^{\alpha-1}$, where R₀ and R are the initial and final D/H ratios for a fraction f of remaining hydrogen⁴¹. Degassing of H₂ (M₁ = 2 for H₂ and M₂ = 3 for HD) yields an α value of about 0.8165, and degassing of H₂O (M₁ = 18 for H₂O and M₂ = 19 for HDO) yields an α value of about 0.9733 (ref. ⁴¹; Extended Data Fig. 4).

Petrography and mineral chemistry of CE5 basalts

Petrography

Approximately 40% of the lithic clasts on a total of 13 sample mounts used in this study are basalt clasts, consisting mainly of pyroxene, plagioclase, olivine and ilmenite, with minor silica, troilite, Si–K-rich mesostasis, apatite and trace merrillite (Supplementary Figs. 1, 2). The basalt clasts display subophitic, poikilitic, porphyritic and equigranular textures (Supplementary Figs. 1, 2), similar to those reported by ref. ⁴. Most pyroxene gains are compositionally zoned with dark and low FeO cores and bright and high FeO rims in BSE images (Supplementary Fig. 1) and energy dispersive spectroscopy. Ilmenite grains occur as laths partially enclosed by pyroxene, suggestive of early crystallization (Supplementary Fig. 1). Eight melt inclusions in ilmenite were

identified with a diameter of about 4 µm to about 50 µm and circular shapes (Extended Data Table 1, Supplementary Fig. 1). Some of the melt inclusions experienced partial post-entrainment crystallization (0–52%) with pyroxene and merrillite embedded in a glassy matrix (Supplementary Fig. 1). Most grains of apatite occur in the fine-grained interstitial materials, coexisting with fayalite and K–Si-rich mesostasis (Supplementary Fig. 2). A few euhedral grains of apatite are enclosed in the margins of pyroxene and FeO-rich olivine (Supplementary Fig. 2). Most grains of apatite are smaller than 10 µm, and the detailed information is summarized in Extended Data Table 1.

Mineral chemistry

Both pyroxene and olivine from various CE5 basalt clasts are chemically zoned, with higher FeO contents at the rims (Fs about 85.9 mol% and Fa about 98.6 mol%) than in the cores (En about 39.6 mol% and Fa about 43.0 mol%), where $Fs = 100 \times Fe/(Fe + Ca + Mg)$ mol%, $En = 100 \times Mg/(Fe + Ca + Mg)$ mol%, and $Fa = 100 \times Fe/(Fe + Mg)$ mol% (Supplementary Figs. 3, 4, Supplementary Table 1). Plagioclase is relatively homogeneous with a composition of $An_{74.8-92.3}Ab_{7.4-21.6}Or_{0.3-4.7}$, where $An = 100 \times Ca/(Ca + Na + K)$ mol%, $Ab = 100 \times Na/(Ca + Na + K)$ mol%, and $Or = 100 \times K/(Ca + Na + K)$ mol% (Supplementary Fig. 4, Supplementary Table 1). Ilmenite has a homogeneous composition of 52.9 wt% TiO_2 and 44.9 wt% FeO, with minor silicon dioxide (SiO_2 ; <0.45 wt%), Cr_2O_3 (<0.32 wt%), manganese oxide (MnO ; 0.34–0.47 wt%) and magnesium oxide (MgO ; <1.47 wt%) (Supplementary Table 2). Most of the melt inclusions in ilmenite are SiO_2 rich (61.2–77.1 wt%), and have wide ranges of aluminium oxide (Al_2O_3 ; 5.59–16.2 wt%), FeO (2.79–24.4 wt%) and calcium oxide (CaO ; 0.72–15.6 wt%) abundances (Supplementary Table 2). In the diagram of MgO versus SiO_2 , Al_2O_3 , FeO, CaO and sodium oxide (Na_2O ; Supplementary Fig. 5), the melt inclusions plot close to the low- MgO endmember, consistent with the datasets of melt inclusions in olivine and pyroxene from Apollo basalts reported by ref. 11. The apatite grains contain 2.35–3.28 wt% F and 0.11–0.87 wt% Cl (Supplementary Table 2), plotting close to the fluorapatite end-member in the F–Cl–OH ternary diagram (Supplementary Fig. 6). The apatite OH contents calculated by difference, assuming that the volatile site only contains F, Cl and OH, range from about 0 to 0.24 wt%.

Estimate of water abundance for the parent magma of CE5 basalts

The melt inclusions in ilmenite

The water abundances and δD values of melt inclusions were first corrected for the cosmic-ray spallation effects (Extended Data Table 3). After the spallation correction,

the measured water abundances of the melt inclusions were corrected for post-entrapment crystallisation (PEC) based on the percentage of crystallized periphery of the melt inclusions (Extended Data Table 3). The water abundances of the melt inclusions were reduced by 0–52% by PEC correction. The water abundances and δD values of melt inclusions corrected for both spallation and PEC effects are summarized in Extended Data Table 3. Three melt inclusions were analysed twice, with two of them (clast numbers 103-020,013 and 103-020,018) showing larger variation in water content between repeated measurements. This observation could be attributed to partial covering on the re-crystallized materials in the melt inclusions.

The melt inclusions analysed in this work show two distinct populations in terms of water abundance and δD values (Fig. 2). Eight analyses on five individual basalt clasts define a negative correlation (main trend) between the water abundances ($6 \pm 2 \mu\text{g g}^{-1}$ to $283 \pm 22 \mu\text{g g}^{-1}$) and δD values ($-330 \pm 190\text{\textperthousand}$ to $200 \pm 390\text{\textperthousand}$) (Extended Data Table 3), obviously distinct from the analyses of the apatite grains in CE5 basalts (Fig. 2) and Apollo mare basalts^{15,17,29,34,35,37}. The other three analyses of the ilmenite-hosted melt inclusions located in three other basalt clasts are more deuterium enriched ($271 \pm 124\text{\textperthousand}$ to $869 \pm 230\text{\textperthousand}$) and contain $93 \pm 15 \mu\text{g g}^{-1}$ to $370 \pm 21 \mu\text{g g}^{-1}$ water, plotting in the region between the analyses of apatite and the negative trend of the melt inclusions (Fig. 2). The negative correlation between water abundances and δD values of the melt inclusions ($\delta D < 200\text{\textperthousand}$) can be explained by degassing of H₂ in the basaltic magma and deuterium enrichment during this process^{11,15}. The water-rich and deuterium-poor end-member of these melt inclusions have δD values of $-330 \pm 190\text{\textperthousand}$, within the range of the hydrogen isotopic compositions of the lunar mantle (about $0 \pm 200\text{\textperthousand}$) constrained by numerous measurements of Apollo melt inclusions¹¹, anorthosite²⁴, and apatite from KREEP basalts^{16,61}, high-Al basalts^{29,36} and highlands samples^{19,62,63}. The most deuterium-depleted melt inclusions probably captured the parent magma of CE5 basalts before significant degassing of H₂. Therefore, the water abundance of $283 \pm 22 \mu\text{g g}^{-1}$ of the deuterium-depleted melt inclusions can be regarded as the maximum water abundance estimate for the parent magma of CE5 basalts (Extended Data Table 3).

Apatite

Apatite is the major OH-, F- and Cl-bearing phase in lunar and other extraterrestrial samples. It was once widely used to estimate the water contents of the mantle reservoirs of Mars⁶⁴ and the Moon^{13,14}. Recent numerical models have revealed that it is possible to crystallize water-rich apatite from a highly water-depleted magma, because of the fractional crystallization and exchange behaviour of OH, F and Cl in apatite⁶⁵. To calculate the water content of a silicate melt using the composition of coexisting apatite, many parameters are required, including the apatite-based melt

hygrometry, the water content of the apatite, the apatite-melt exchange coefficient, the abundance of F or Cl in the apatite, and that of F or Cl in the melt⁶⁶. However, it is difficult to determine all of these parameters precisely in the case of CE5 apatite.

Instead, we estimated the water abundance of the bulk CE5 basalts from the modal abundance of apatite and its average water content, because apatite is the dominant water-bearing phase in mare basalts. The modal abundance of apatite in the CE5 basalts is determined to be approximately 0.4 vol%, using the surface areas of apatite in all basaltic clasts analysed (Supplementary Table 1). The average water content and δD value of the CE5 apatite measured by NanoSIMS 50L are $1,921 \pm 910 \mu\text{g g}^{-1}$ and $578 \pm 208\text{\textperthousand}$ ($N=40$), respectively, except for the three low water abundance analyses (Extended Data Table 2). Hence, the water abundance of the bulk CE5 basalt is $7 \pm 3 \mu\text{g g}^{-1}$. As discussed above, the parent magma of CE5 basalts has the original δD value of approximately $-330\text{\textperthousand}$ indicated by the most deuterium-depleted melt inclusions, whereas the deuterium enrichment of apatite was probably attributed to degassing of water in the form of H₂. Hence, the water abundance of the parent magma was calibrated to be $600 \pm 400 \mu\text{g g}^{-1}$, with 98–99% water degassing loss required to enhance δD values from about $-330\text{\textperthousand}$ to about $600\text{\textperthousand}$ based on the H₂ degassing modelling⁴¹ (Extended Data Table 3). This estimate is consistent with the water abundance of the most deuterium-depleted melt inclusions in ilmenite.

Clinopyroxene

A higher Cs⁺ primary beam current (7 nA) was used to analyse nominally anhydrous clinopyroxene, and the H₂O background was reduced to $5 \pm 2 \mu\text{g g}^{-1}$. Eighteen analyses on clinopyroxene from 13 individual CE5 basalt clasts yielded an average ¹H/¹⁶O ratio of 8.85×10^{-7} , significantly lower than that of San Carlos olivine (2.30×10^{-6}) measured under the same conditions (Extended Data Table 4). Therefore, all hydrogen emitted during CE5 clinopyroxene analyses can be ascribed to background hydrogen, indicating that the CE5 clinopyroxene contains less than $5 \mu\text{g g}^{-1}$ H₂O. The water abundance of the parent magma equilibrated with the CE5 clinopyroxene could be less than about $170 \mu\text{g g}^{-1}$ using the water partition coefficient of 0.03 that was determined experimentally under reduced lunar conditions³³.

Estimate of water abundance for the mantle source of CE5 basalts

The ilmenite-hosted melt inclusions were used to estimate the water abundance for the mantle source of CE5 basalts because they are the quenched parent melt and have preserved the original δD value of the mantle source. On the basis of the petrogenesis and geochemistry of CE5 basalts, the parent magma was derived from a depleted

mantle source⁴. Furthermore, the REE patterns of the bulk CE5 basalts indicate that these basalts formed through low-degree (2–3%) partial melting followed by moderate-to-high-degree (43–88%) fractional crystallization⁴. Accordingly, a maximum water abundance of $1\text{--}5 \mu\text{g g}^{-1}$ can be estimated for the depleted mantle source, which produced $283 \pm 22 \mu\text{g g}^{-1}$ water in the derived parent magma of the CE5 basalts, assuming that all water partitioned into melt during partial melting of lunar mantle³³.

Data availability

All geochemical data generated in this study are included in Extended Data Tables 1–5 and in Supplementary Tables 1–5, and are available on Zenodo at <https://doi.org/10.5281/zenodo.5341793>.

References

1. 1. Canup, R. M. & Asphaug, E. Origin of the Moon in a giant impact near the end of the Earth's formation. *Nature* **412**, 708–712 (2001).
2. 2. Shearer, C. K. et al. Thermal and magmatic evolution of the moon. *Rev. Mineral. Geochim.* **60**, 365–518 (2006).
3. 3. Li, Q.-L. et al. Two-billion-year-old volcanism on the Moon from Chang'e-5 basalts. *Nature* <https://doi.org/10.1038/s41586-021-04100-2> (2021).
4. 4. Tian, H.-C. et al. Non-KREEP origin for Chang'e-5 basalts in the Procellarum KREEP Terrane. *Nature* <https://doi.org/10.1038/s41586-021-04119-5> (2021).
5. 5. Borg, L. E., Shearer, C. K., Asmerom, Y. & Papike, J. J. Prolonged KREEP magmatism on the Moon indicated by the youngest dated lunar igneous rock. *Nature* **432**, 209–211 (2004).

6. 6.

Compston, W., Vernon, M. J., Berry, H. & Rudowski, R. The age of the Fra Mauro Formation: a radiometric older limit. *Earth Planet. Sci. Lett.* **12**, 55–58 (1971).

7. 7.

Lin, Y., Tronche, E. J., Steenstra, E. S. & van Westrenen, W. Evidence for an early wet Moon from experimental crystallization of the lunar magma ocean. *Nat. Geosci.* **10**, 14–18 (2017).

8. 8.

Hauri, E. H., Weinreich, T., Saal, A. E., Rutherford, M. C. & Van Orman, J. A. High pre-eruptive water contents preserved in lunar melt inclusions. *Science* **333**, 213–215 (2011).

9. 9.

Chen, Y. et al. Water, fluorine, and sulfur concentrations in the lunar mantle. *Earth Planet. Sci. Lett.* **427**, 37–46 (2015).

10. 10.

Ni, P., Zhang, Y. X., Chen, S. & Gagnon, J. A melt inclusion study on volatile abundances in the lunar mantle. *Geochim. Cosmochim. Acta* **249**, 17–41 (2019).

11. 11.

Stephant, A. et al. The hydrogen isotopic composition of lunar melt inclusions: an interplay of complex magmatic and secondary processes. *Geochim. Cosmochim. Acta* **284**, 196–221 (2020).

12. 12.

Stephant, A. et al. The chlorine isotopic composition of the Moon: insights from melt inclusions. *Earth Planet. Sci. Lett.* **523**, 115715 (2019).

13. 13.

Boyce, J. W. et al. Lunar apatite with terrestrial volatile abundances. *Nature* **466**, 466–469 (2010).

14. 14.

McCubbin, F. M. et al. Nominally hydrous magmatism on the Moon. *Proc. Natl Acad. Sci. USA* **107**, 11223–11228 (2010).

15. 15.

Tartèse, R. et al. The abundance, distribution, and isotopic composition of hydrogen in the Moon as revealed by basaltic lunar samples: implications for the volatile inventory of the Moon. *Geochim. Cosmochim. Acta* **122**, 58–74 (2013).

16. 16.

Tartèse, R. et al. Apatites in lunar KREEP basalts: the missing link to understanding the H isotope systematics of the Moon. *Geology* **42**, 363–366 (2014).

17. 17.

Barnes, J. J., Franchi, I. A., McCubbin, F. M. & Anand, M. Multiple reservoirs of volatiles in the Moon revealed by the isotopic composition of chlorine in lunar basalts. *Geochim. Cosmochim. Acta* **266**, 144–162 (2019).

18. 18.

Barnes, J. J. et al. An asteroidal origin for water in the Moon. *Nat. Commun.* **7**, 11684 (2016).

19. 19.

Barnes, J. J. et al. The origin of water in the primitive Moon as revealed by the lunar highlands samples. *Earth Planet. Sci. Lett.* **390**, 244–252 (2014).

20. 20.

Barnes, J. J. et al. Early degassing of lunar urKREEP by crust-breaching impact(s). *Earth Planet. Sci. Lett.* **447**, 84–94 (2016).

21. 21.

Saal, A. E., Hauri, E. H., Van Orman, J. A. & Rutherford, M. J. Hydrogen isotopes in lunar volcanic glasses and melt inclusions reveal a carbonaceous chondrite heritage. *Science* **340**, 1317–1320 (2013).

22. 22.

Hauri, E. H., Saal, A. E., Rutherford, M. J. & Van Orman, J. A. Water in the Moon's interior: truth and consequences. *Earth Planet. Sci. Lett.* **409**, 252–264 (2015).

23. 23.

Hui, H. J., Peslier, A. H., Zhang, Y. X. & Neal, C. R. Water in lunar anorthosites and evidence for a wet early Moon. *Nat. Geosci.* **6**, 177–180 (2013).

24. 24.

Hui, H. J. et al. A heterogeneous lunar interior for hydrogen isotopes as revealed by the lunar highlands samples. *Earth Planet. Sci. Lett.* **473**, 14–23 (2017).

25. 25.

McCubbin, F. M. et al. Magmatic volatiles (H, C, N, F, S, Cl) in the lunar mantle, crust, and regolith: abundances, distributions, processes, and reservoirs. *Am. Mineral.* **100**, 1668–1707 (2015).

26. 26.

Hauri, E. H. et al. Origin and evolution of water in the Moon's interior. *Annu. Rev. Earth. Planet. Sci.* **45**, 89–111 (2017).

27. 27.

Liu, Y. et al. Direct measurement of hydroxyl in the lunar regolith and the origin of lunar surface water. *Nat. Geosci.* **5**, 779–782 (2012).

28. 28.

Treiman, A. H. et al. D-poor hydrogen in lunar mare basalts assimilated from lunar regolith. *Am. Mineral.* **101**, 1596–1603 (2016).

29. 29.

Greenwood, J. P. et al. Hydrogen isotope ratios in lunar rocks indicate delivery of cometary water to the Moon. *Nat. Geosci.* **4**, 79–82 (2011).

30. 30.

Qian, Y. Q. et al. Young lunar mare basalts in the Chang'e-5 sample return region, northern Oceanus Procellarum. *Earth Planet. Sci. Lett.* **555**, 116702 (2021).

31. 31.

Qian, Y. Q. et al. China's Chang'e-5 landing site: geology, stratigraphy, and provenance of materials. *Earth Planet. Sci. Lett.* **561**, 116855 (2021).

32. 32.

Jolliff, B. L., Gillis, J. J., Haskin, L. A., Korotev, R. L. & Wieczorek, M. A. Major lunar crustal terranes: surface expressions and crust-mantle origins. *J. Geophys. Res. Planets* **105**, 4197–4216 (2000).

33. 33.

Potts, N. J., Bromiley, G. D. & Brooker, R. A. An experimental investigation of F, Cl and H₂O mineral-melt partitioning in a reduced, model lunar system. *Geochim. Cosmochim. Acta* **294**, 232–254 (2021).

34. 34.

Boyce, J. W. et al. The chlorine isotope fingerprint of the lunar magma ocean. *Sci. Adv.* **1**, e1500380 (2015).

35. 35.

Barnes, J. J. et al. Accurate and precise measurements of the D/H ratio and hydroxyl content in lunar apatites using NanoSIMS. *Chem. Geol.* **337**, 48–55 (2013).

36. 36.

Pernet-Fisher, J. F., Howarth, G. H., Liu, Y., Chen, Y. & Taylor, L. A. Estimating the lunar mantle water budget from phosphates: complications associated with silicate-liquid-immiscibility. *Geochim. Cosmochim. Acta* **144**, 326–341 (2014).

37. 37.

Singer, J. A., Greenwood, J. P., Itoh, S., Sakamoto, N. & Yurimoto, H. Evidence for the solar wind in lunar magmas: a study of slowly cooled samples of the Apollo 12 olivine basalt suite. *Geochem. J.* **51**, 95–104 (2017).

38. 38.

Robinson, K. L. & Taylor, G. J. Heterogeneous distribution of water in the Moon. *Nat. Geosci.* **7**, 401–408 (2014).

39. 39.

Curran, N. M. et al. A database of noble gases in lunar samples in preparation for mass spectrometry on the Moon. *Planet. Space Sci.* **182**, 104823 (2020).

40. 40.

Hörz, F., Grieve, R., Heiken, G., Spudis, P. & Binder, A. in *Lunar Sourcebook: A User's Guide to the Moon* 61–120 (1974).

41. 41.

Sharp, Z. D., McCubbin, F. M. & Shearer, C. K. A hydrogen-based oxidation mechanism relevant to planetary formation. *Earth Planet. Sci. Lett.* **380**, 88–97 (2013).

42. 42.

Ustunisik, G., Nekvasil, H. & Lindsley, D. Differential degassing of H₂O, Cl, F, and S: potential effects on lunar apatite. *Am. Mineral.* **96**, 1650–1653 (2011).

43. 43.

Ustunisik, G., Nekvasil, H., Lindsley, D. H. & McCubbin, F. M. Degassing pathways of Cl-, F-, H-, and S-bearing magmas near the lunar surface: implications for the composition and Cl isotopic values of lunar apatite. *Am. Mineral.* **100**, 1717–1727 (2015).

44. 44.

Kyser, T. K. in *Stable Isotopes in High Temperature Geological Processes* (eds John, W. V. et al.) 141–164 (De Gruyter, 2018).

45. 45.

Qian, Y. Q. et al. Geology and scientific significance of the Rümker region in Northern Oceanus Procellarum: China's Chang'e-5 landing region. *J. Geophys. Res. Planets* **123**, 1407–1430 (2018).

46. 46.

Ardia, P., Hirschmann, M. M., Withers, A. C. & Tenner, T. J. H₂O storage capacity of olivine at 5–8 GPa and consequences for dehydration partial melting of the upper mantle. *Earth Planet. Sci. Lett.* **345–348**, 104–116 (2012).

47. 47.

Asimow, P. D. & Langmuir, C. H. The importance of water to oceanic mantle melting regimes. *Nature* **421**, 815–820 (2003).

48. 48.

Ringwood, A. & Kesson, S. A dynamic model for mare basalt petrogenesis. *Proc. Lunar Planet. Sci.* **7**, 1697–1722 (1976).

49. 49.

Hess, P. C. & Parmentier, E. M. A model for the thermal and chemical evolution of the Moon’s interior: implications for the onset of mare volcanism. *Earth Planet. Sci. Lett.* **134**, 501–514 (1995).

50. 50.

Saal, A. E., Hauri, E. H., Langmuir, C. H. & Perfit, M. R. Vapour undersaturation in primitive mid-ocean-ridge basalt and the volatile content of Earth’s upper mantle. *Nature* **419**, 451–455 (2002).

51. 51.

Zhang, W. F., Xia, X. P., Zhang, Y. Q., Peng, T. P. & Yang, Q. A novel sample preparation method for ultra-high vacuum (UHV) secondary ion mass spectrometry (SIMS) analysis. *J. Anal. Atom. Spectrom.* **33**, 1559–1563 (2018).

52. 52.

Hauri, E. et al. SIMS analysis of volatiles in silicate glasses 1. Calibration, matrix effects and comparisons with FTIR. *Chem. Geol.* **183**, 99–114 (2002).

53. 53.

Hu, S. et al. NanoSIMS analyses of apatite and melt inclusions in the GRV 020090 Martian meteorite: hydrogen isotope evidence for recent past

underground hydrothermal activity on Mars. *Geochim. Cosmochim. Acta* **140**, 321–333 (2014).

54. 54.

Hu, S. et al. Measurements of water content and D/H ratio in apatite and silicate glasses using a NanoSIMS 50L. *J. Anal. Atom. Spectrom.* **30**, 967–978 (2015).

55. 55.

Zhang, W. F. et al. Optimization of SIMS analytical parameters for water content measurement of olivine. *Surf. Interf. Anal.* **52**, 224–233 (2019).

56. 56.

Tartèse, R., Anand, M. & Franchi, I. A. H and Cl isotope characteristics of indigenous and late hydrothermal fluids on the differentiated asteroidal parent body of Grave Nunataks 06128. *Geochim. Cosmochim. Acta* **266**, 529–543 (2019).

57. 57.

Greenwood, J. P., Itoh, S., Sakamoto, N., Vicenzi, E. P. & Yurimoto, H. Hydrogen isotope evidence for loss of water from Mars through time. *Geophys. Res. Lett.* **35**, L05203 (2008).

58. 58.

Nadeau, S. L., Epstein, S. & Stolper, E. Hydrogen and carbon abundances and isotopic ratios in apatite from alkaline intrusive complexes, with a focus on carbonatites. *Geochim. Cosmochim. Acta* **63**, 1837–1851 (1999).

59. 59.

Füri, E., Deloule, E. & Trappitsch, R. The production rate of cosmogenic deuterium at the Moon's surface. *Earth Planet. Sci. Lett.* **474**, 76–82 (2017).

60. 60.

Merlivat, L., Leiu, M., Neif, G. & Roth, E. Spallation deuterium in rock 70215. *Proc. Lunar Planet. Sci.* 649–658 (1976).

61. 61.

Tartèse, R., Anand, M., Joy, K. H. & Franchi, I. A. H and Cl isotope systematics of apatite in brecciated lunar meteorites Northwest Africa 4472, Northwest Africa 773, Sayh al Uhaymir 169, and Kalahari 009. *Meteorit. Planet. Sci.* **49**, 2266–2289 (2014).

62. 62.

Cernok, A. et al. Preservation of primordial signatures of water in highly-shocked ancient lunar rocks. *Earth Planet. Sci. Lett.* **544**, 116364 (2020).

63. 63.

Robinson, K. L. et al. Water in evolved lunar rocks: evidence for multiple reservoirs. *Geochim. Cosmochim. Acta* **188**, 244–260 (2016).

64. 64.

McCubbin, F. M. et al. Hydrous melting of the martian mantle produced both depleted and enriched shergottites. *Geology* **40**, 683–686 (2012).

65. 65.

Boyce, J. W., Tomlinson, S. M., McCubbin, F. M., Greenwood, J. P. & Treiman, A. H. The lunar apatite paradox. *Science* **344**, 400–402 (2014).

66. 66.

McCubbin, F. M. et al. Experimental investigation of F, Cl, and OH partitioning between apatite and Fe-rich basaltic melt at 1.0–1.2 GPa and 950–1000 °C. *Am. Mineral.* **4779**, 83–89 (2015).

67. 67.

Yang, W. & Lin, Y. New lunar samples returned by Chang'e-5: opportunities for new discoveries and international collaboration. *Innov. J.* **2**, 100070 (2021).

68. 68.

Lawrence, D. J. et al. Global spatial deconvolution of lunar prospector Th abundances. *Geophys. Res. Lett.* **34**, L03201 (2007).

69. 69.

Guggisberg, S. et al. Classification of the Apollo-11 mare basalts according to Ar³⁹–Ar⁴⁰ ages and petrological properties. *Proc. Lunar Planet. Sci.* **10**, 1–39 (1979).

70. 70.

Papanastassiou, D., Wasserburg, G. & Burnett, D. Rb–Sr ages of lunar rocks from the Sea of Tranquillity. *Earth. Planet. Sci. Lett.* **8**, 1–19 (1970).

71. 71.

Nyquist, L., Shih, C.-Y., Wooden, J., Bansal, B. & Wiesmann, H. The Sr and Nd isotopic record of Apollo 12 basalts: implications for lunar geochemical evolution. *Proc. Lunar Planet. Sci.* **10**, 77–114 (1979).

72. 72.

Papanastassiou, D. & Wasserburg, G. Rb–Sr ages of igneous rocks from the Apollo 14 mission and the age of the Fra Mauro Formation. *Earth. Planet. Sci. Lett.* **12**, 36–48 (1971).

73. 73.

Saal, A. E. et al. Volatile content of lunar volcanic glasses and the presence of water in the Moon’s interior. *Nature* **454**, 192–195 (2008).

74. 74.

Snyder, G., Borg, L., Taylor, L., Nyquist, L. & Halliday, A. Volcanism in the Hadley–Apennine region of the Moon: geochronology, Nd–Sr isotopic systematics, and depths of melting. *Proc. Lunar Planet. Sci.* **11**, 141 (1998).

75. 75.

Husain, L. ⁴⁰Ar–³⁹Ar chronology and cosmic ray exposure ages of the Apollo 15 samples. *J. Geophys. Res.* **79**, 2588–2606 (1974).

76. 76.

Tatsumoto, M., Premo, W. R. & Unruh, D. M. Origin of lead from green glass of Apollo 15426: a search for primitive lunar lead. *J. Geophys. Res. Solid Earth* **92**, E361–E371 (1987).

77. 77.

York, D., Kenyon, W. J. & Doyle, R. J. ^{40}Ar - ^{39}Ar ages of Apollo 14 and 15 samples. *Proc. Lunar Planet. Sci.* **3**, 1613–1622 (1972).

78. 78.

Saito, K. & Alexander, E. ^{40}Ar - ^{39}Ar studies of lunar soil 74001. *Proc. Lunar Planet. Sci.* **10**, 1049–1051 (1979).

79. 79.

Füri, E., Deloule, E., Gurenko, A. & Marty, B. New evidence for chondritic lunar water from combined D/H and noble gas analyses of single Apollo 17 volcanic glasses. *Icarus* **229**, 109–120 (2014).

80. 80.

Tera, F. & Wasserburg, G. Lunar ball games and other sports. *Proc. Lunar Planet. Sci.* **107**, (1976).

81. 81.

Tera, F., Papanastassiou, D. A. & Wasserburg, G. J. Isotopic evidence for a terminal lunar cataclysm. *Earth. Planet. Sci. Lett.* **22**, 1–21 (1974).

82. 82.

Fernandes, V. A., Burgess, R. & Morris, A. ^{40}Ar - ^{39}Ar age determinations of lunar basalt meteorites Asuka 881757, Yamato 793169, Miller Range 05035, La Paz Icefield 02205, Northwest Africa 479, and basaltic breccia Elephant Moraine 96008. *Meteorit. Planet. Sci.* **44**, 805–821 (2009).

83. 83.

Anand, M. et al. Petrology and geochemistry of LaPaz Icefield 02205: a new unique low-Ti mare-basalt meteorite. *Geochim. Cosmochim. Acta* **70**, 246–264 (2006).

Acknowledgements

We thank D. Chew for providing Durango and Kovdor apatite; E. Hauri for providing basaltic glass 519-4-1 and 1833-11; R. Francis for providing SWIFT MORB glass; R. Mitchell, X.-H. Li and F.-Y. Wu for constructive comments; Y. Chen and L. Jia for the assistance on EPMA measurement; W. Yang, H. Tian, H. Ma and D. Zhang for hand

picking the CE5 soil fragments; and J. Yuan and X. Tang for the assistance with the SEM observation. This study was funded by the Strategic Priority Research Program of Chinese Academy of Sciences (XDB 41000000), the key research programme of the Chinese Academy of Sciences (ZDBS-SSW-JSC007-15) and the key research programme of the Institute of Geology and Geophysics, Chinese Academy of Sciences (IGGCAS-202101 and 201904), the National Natural Science Foundation of China (41973062) and Pre-research project on Civil Aerospace Technologies by CNSA (D020201, D020203, and D020205). M.A. and R.T. acknowledge funding from the UK Science and Technology Facilities Council (grant numbers ST/P000657/1 and ST/P005225/1, respectively). The CE5 samples were allocated by the China National Space Administration.

Author information

Affiliations

1. Key Laboratory of the Earth and Planetary Physics, Chinese Academy of Sciences, Beijing, China

Sen Hu, Huicun He, Jianglong Ji, Yangting Lin, Yihong Yan, Jialong Hao, Ruiying Li & Lixin Gu

2. State Key Laboratory for Mineral Deposits Research & Lunar and Planetary Science Institute, School of the Earth Sciences and Engineering, Nanjing University, Nanjing, China

Hejiu Hui

3. CAS Center for Excellence in Comparative Planetology, Hefei, China

Hejiu Hui

4. School of Physical Sciences, The Open University, Milton Keynes, UK

Mahesh Anand

5. Department of Earth Sciences, The Natural History Museum, London, UK

Mahesh Anand

6. Department of Earth and Environmental Sciences, The University of Manchester, Manchester, UK

Romain Tartèse

7. State Key Laboratory of Lithospheric Evolution, Chinese Academy of Sciences, Beijing, China

Qian Guo & Huaiyu He

8. Center for Lunar and Planetary Sciences, Institute of Geochemistry, Chinese Academy of Sciences, Guiyang, China

Ziyuan Ouyang

Contributions

S.H., Y.L. and H. Hui designed this research. J.J., Huicun He, Y.Y., L.G., Q.G. and S.H. prepared the sample and characterized the petrography and mineral chemistry of CE5 basalts. Huicun He, J.H., R.L., J.J. and S.H. conducted the NanoSIMS measurements. S.H., Y.L., H. Hui, J.J., Huicun He, M.A. and R.T. wrote the manuscript. All authors contributed to the preparation of the manuscript.

Corresponding authors

Correspondence to [Sen Hu](#) or [Yangting Lin](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Charles Shearer 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 Location of CE5 samples on the Moon.

CE5 landed in a region with a high Th abundance within the northwestern Oceanus PKT (outlined) on 17 December 2020 (refs. [31](#),[67](#)), far from the landing sites of Apollo and Luna missions. The map (LROC WAC Basemap) and the overlapping Th distribution (Lunar Prospector) are adapted from ref. [68](#).

Extended Data Fig. 2 The water content calibration line established on apatite and silicate glass standards.

KOV, Kovdor apatite; DAP, Durango apatite; SWIFT, SWIFT MORB glass; 519-4-1, basaltic glass; 1833-11, basaltic glass; SCOL, San Carlos olivine. The datasets are listed in Supplementary Table [3](#). The analytical uncertainty is 0.6% (2σ).

Extended Data Fig. 3 Reproducibility of hydrogen isotope analysis of the standards and the CE5 apatite and ilmenite-hosted melt inclusions over the analytical session.

The reproducibility of δD analysis throughout the whole analytical sessions was $\pm 43\%$ (2σ), estimated on the Kovdor apatite standard. The average δD values measured for the SWIFT MORB glass and Durango apatite are $-113 \pm 109\text{\textperthousand}$ and $-92 \pm 210\text{\textperthousand}$, respectively, consistent with their recommended values^{[29](#),[57](#)} within analytical errors. The average δD values measured for the basaltic glass standards 1833-11 and 519-4-1 are $-101 \pm 86\text{\textperthousand}$ and $-132 \pm 158\text{\textperthousand}$, respectively. All analytical data are listed in Extended Data Tables [3](#), [4](#), Supplementary Table [3](#).

Extended Data Fig. 4 Spallation correction modelling for ilmenite-hosted melt inclusions after PEC correction.

A cosmogenic deuterium production rate of $2.17 \times 10^{-12} \text{ mol D g}^{-1} \text{ Ma}^{-1}$ (ref. [59](#)) was used for melt inclusion glasses. (a) CRE age = 10 Ma. (b) CRE age = 50 Ma. (c) CRE age = 100 Ma. (d) CRE age = 200 Ma. Open circles correspond to measured data and blue filled circles correspond to CRE-corrected data (Extended Data Table [4](#), Supplementary Table [4](#)). Modelled degassing curves have been calculated assuming Rayleigh fractionation into a vacuum, using masses of 2.016 and 3.022, and a fractionation factor of $(2.016/3.022)^{1/2}$ for H_2 , and masses of 18.015 and 19.021, corresponding to a fractionation factor of $(18.015/19.021)^{1/2}$ for H_2O , following the procedure of ref. [41](#) ([Methods](#)).

Extended Data Table 1 The CE5 basalt clasts

Extended Data Table 2 Water abundance and hydrogen isotopes of CE5 apatite
Extended Data Table 3 Water abundance and hydrogen isotopes of CE5 ilmenite-hosted melt inclusions

Extended Data Table 4 H/O ratios of CE5 clinopyroxene and reference San Carlos olivine measured by NanoSIMS 50L

Extended Data Table 5 Summary of the water abundances estimated for the lunar mantle source regions of basaltic products formed between around 4 Ga and 2 Ga

Supplementary information

Supplementary Information

Supplementary Figs. 1–6 and references.

Supplementary Table 1

Modal abundance of apatite in the CE5 basalt clasts.

Supplementary Table 2

EPMA results of the CE5 basalt clasts.

Supplementary Table 3

NanoSIMS results of the standards and silicates of the CE5 basalt clasts.

Supplementary Table 4

Water abundances and hydrogen isotope compositions of ilmenite-hosted melt inclusions and apatite with correction for spallation effects.

Supplementary Table 5

Summary of water abundances and hydrogen isotope compositions of apatite and melt inclusions from Apollo samples in the literature.

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

Hu, S., He, H., Ji, J. *et al.* A dry lunar mantle reservoir for young mare basalts of Chang'e-5. *Nature* **600**, 49–53 (2021). <https://doi.org/10.1038/s41586-021-04107-9>

- Received: 30 July 2021
- Accepted: 07 October 2021
- Published: 19 October 2021
- Issue Date: 02 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04107-9>

Share this article

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

[Get shareable link](#)

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

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

Further reading

- [Chang'e-5 samples reveal two-billion-year-old volcanic activity on the Moon and its source characteristics](#)
 - Shoujie Liu
 - Qin Zhou
 - Wei Yang

Science China Earth Sciences (2021)

[Robotic sample return reveals lunar secrets](#)

- Richard W. Carlson

News & Views 01 Dec 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04107-9>

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

- Article
- Open Access
- [Published: 19 October 2021](#)

Two-billion-year-old volcanism on the Moon from Chang'e-5 basalts

- [Qiu-Li Li](#) ORCID: [orcid.org/0000-0002-7280-5508¹](https://orcid.org/0000-0002-7280-5508),
- [Qin Zhou](#) ORCID: [orcid.org/0000-0001-8562-9737²](https://orcid.org/0000-0001-8562-9737),
- [Yu Liu](#) ORCID: [orcid.org/0000-0001-7195-7393¹](https://orcid.org/0000-0001-7195-7393),
- [Zhiyong Xiao](#) ORCID: [orcid.org/0000-0002-5026-6937³](https://orcid.org/0000-0002-5026-6937),
- [Yangting Lin](#) ORCID: [orcid.org/0000-0002-3407-4329⁴](https://orcid.org/0000-0002-3407-4329),
- [Jin-Hua Li](#)⁴,
- [Hong-Xia Ma](#) ORCID: [orcid.org/0000-0001-8215-8271¹](https://orcid.org/0000-0001-8215-8271),
- [Guo-Qiang Tang](#) ORCID: [orcid.org/0000-0001-7754-9390¹](https://orcid.org/0000-0001-7754-9390),
- [Shun Guo](#) ORCID: [orcid.org/0000-0002-0439-0507¹](https://orcid.org/0000-0002-0439-0507),
- [Xu Tang](#) ORCID: [orcid.org/0000-0002-2476-1635⁴](https://orcid.org/0000-0002-2476-1635),
- [Jiang-Yan Yuan](#) ORCID: [orcid.org/0000-0002-1845-7565¹](https://orcid.org/0000-0002-1845-7565),
- [Jiao Li](#) ORCID: [orcid.org/0000-0001-9628-9838¹](https://orcid.org/0000-0001-9628-9838),
- [Fu-Yuan Wu](#) ORCID: [orcid.org/0000-0002-2281-7885¹](https://orcid.org/0000-0002-2281-7885),
- [Ziyuan Ouyang](#)⁵,
- [Chunlai Li](#) ORCID: [orcid.org/0000-0002-0817-2742²](https://orcid.org/0000-0002-0817-2742) &
- [Xian-Hua Li](#) ORCID: [orcid.org/0000-0001-5512-7736¹](https://orcid.org/0000-0001-5512-7736)

[Nature](#) volume 600, pages 54–58 (2021)

- 6814 Accesses
- 6 Citations
- 259 Altmetric
- [Metrics details](#)

Subjects

- [Geochemistry](#)

- [Rings and moons](#)

Abstract

The Moon has a magmatic and thermal history that is distinct from that of the terrestrial planets¹. Radioisotope dating of lunar samples suggests that most lunar basaltic magmatism ceased by around 2.9–2.8 billion years ago (Ga)^{2,3}, although younger basalts between 3 Ga and 1 Ga have been suggested by crater-counting chronology, which has large uncertainties owing to the lack of returned samples for calibration^{4,5}. Here we report a precise lead–lead age of $2,030 \pm 4$ million years ago for basalt clasts returned by the Chang'e-5 mission, and a $^{238}\text{U}/^{204}\text{Pb}$ ratio (μ value)⁶ of about 680 for a source that evolved through two stages of differentiation. This is the youngest crystallization age reported so far for lunar basalts by radiometric dating, extending the duration of lunar volcanism by approximately 800–900 million years. The μ value of the Chang'e-5 basalt mantle source is within the range of low-titanium and high-titanium basalts from Apollo sites (μ value of about 300–1,000), but notably lower than those of potassium, rare-earth elements and phosphorus (KREEP) and high-aluminium basalts⁷ (μ value of about 2,600–3,700), indicating that the Chang'e-5 basalts were produced by melting of a KREEP-poor source. This age provides a pivotal calibration point for crater-counting chronology in the inner Solar System and provides insight on the volcanic and thermal history of the Moon.

[Download PDF](#)

Main

Even though mare basalt covers only roughly 17% of the surface of the Moon⁸, its protracted formation record spans more of the lunar magmatic history than any other geological unit. Radioisotope age studies of basaltic samples returned by the Apollo and Luna missions and lunar meteorites have revealed that basaltic magmatism on the Moon occurred between around 4.4 billion years ago (Ga) (ref. ⁹) and around 2.9–2.8 Ga (refs. ^{2,3}), with two major pulses around 3.95–3.58 Ga and 3.38–3.08 Ga (ref. ¹⁰). However, crater-counting chronology suggests a more extended period of basalt volcanism occurring between around 4.0 Ga and around 1.2 Ga (refs. ^{4,11}). Some of the youngest mare basalt units in Oceanus Procellarum have been estimated to be around 2.2–1.2 Ga (refs. ^{4,11,12}), which would putatively expand the duration of mare volcanism to over roughly three billion years. Whether or not mare volcanism continued to around 2.2–1.2 Ga, or perhaps even younger in small eruptions⁵, has long been a major question. These young crater-counting ages, however, have not been confirmed by radioisotopic ages. The actual end of mare volcanism has not yet been

constrained by radioisotopic dating owing to the lack of available samples from these younger volcanic units, which have incurred fewer impacts.

Age of Chang'e-5 basalt

Precise and accurate age determination of these young mare basalts is crucial not only for unravelling the timing and duration of lunar volcanism but also for investigating late-stage basaltic petrogenesis and the melting of lunar mantle sources that are relevant to the thermal and chemical evolution of the Moon. The landing site of Chang'e-5, China's first lunar-sample-return mission, was selected because it is located on one of the youngest mare basalt units northeast of Mons Rümker in northern Oceanus Procellarum¹³. Therefore, basalts returned by Chang'e-5 provide an opportunity to understand the timing and mechanism of one of the youngest units of mare volcanism on the Moon. In addition, precise radioisotopic dating of these newly returned basalts has a critical potential to verify and calibrate lunar impact crater-counting chronology, which is the main basis for dating most geological units of the other inner Solar System bodies.

The Chang'e-5 samples studied in this research were scooped from the lunar regolith surface and include three one-inch epoxy mounts (samples CE5C0000YJYX041GP and CE5C0000YJYX042GP with two basalt clasts larger than 1.5 mm in each, and sample CE5C0800YJFM00102GP with 20 mg of soil) and two aliquots of soils (samples CE5C0100YJFM00103 of 1 g and CE5C0400YJFM00406 of 2 g) allocated by the China National Space Administration. Around 800 lithic clasts (greater than 0.25 mm) were randomly picked from the two soil samples to make additional epoxy mounts. The lithic clasts in the soils are composed of about 45% basalt and about 55% breccia. The breccia clasts are dominated by basalt fragments (more than 80%) with minor impact melt and agglutinate. The majority of the basalt clasts (about 80%) show subophitic and poikilitic textures, with the remainder being porphyritic and equigranular textures. The basalt clasts have various grain sizes (approximately 10–600 µm), but similar major mineral constituents of clinopyroxene, plagioclase, olivine, ilmenite, and rare troilite and cristobalite (Fig. 1). The euhedral-to-subhedral phosphate minerals apatite and merrillite are found in all types of basalt clast (Fig. 1, Extended Data Figs. 1, 2, Supplementary Table 1) and commonly occur along the margins of iron (Fe)-rich olivine, clinopyroxene and ilmenite. The minor zirconium (Zr)-bearing minerals baddeleyite, zirconolite and tranquillityite are fine-grained (3–8 µm), euhedral to subhedral in shape, interstitial-to-major silicate phases, and common in coarse-grained (greater than 100 µm) poikilitic, equigranular and subophitic basalt clasts, but not in porphyritic basalt clasts (Fig. 1, Extended Data Figs. 1, 2, Supplementary Table 2). In some cases, Zr-bearing and phosphate minerals show intergrowth textures with each other (Fig. 1a, c, d, Extended Data Fig. 2), which suggests that they formed during the same final crystallization stage of the magma.

Forty-seven representative basalt clasts with various textures were used for radioisotopic dating.

Fig. 1: Back-scattered electron images of representative dated minerals in the Chang'e-5 basalt clasts.

 figure1



a, Intergrown texture of baddeleyite, zirconolite and apatite in a subophitic basalt clast. **b**, Subhedral baddeleyite and merrillite occur within mesostasis composed of Fe-rich olivine, clinopyroxene, K-feldspar and troilite in a poikilitic clast. **c**, Baddeleyite and merrillite in a subophitic clast. **d**, Euhedral apatite and baddeleyite occur as intergranular phases associated with Fe-rich olivine in an equigranular basalt clast. Pits in Zr-bearing minerals are the in situ analytical spots from SIMS. The areas of dated minerals in the clasts are outlined (red rectangles) in the corresponding insets. Ap, apatite; Bdy, baddeleyite; Cpx, clinopyroxene; Crs, cristobalite; Ilm, ilmenite; Mer, merrillite; Ol, olivine; Pl, plagioclase; Qz, quartz; Tro, troilite; Zrl, zirconolite. Codes

in the top right corner indicate sample name, mount number and grain number on the mount (**a**, **c**, **d**), or mount name and grain number (**b**, note: this is a one-inch epoxy mount sample with two basalt grains on it.).

The uranium (U)–lead (Pb) isotopic compositions of various mineral phases in Chang'e-5 basalt clasts were determined using a CAMECA IMS 1280HR secondary ion mass spectrometer (SIMS) (complete dataset presented in Supplementary Table 3). Zr-bearing minerals, phosphates and others (that is, plagioclase, pyroxene and matrix minerals) were analysed using a primary O⁻ beam with roughly 3-μm, roughly 8-μm and roughly 30-μm spot sizes, respectively ([Methods](#)). Pb isotope analyses on 17 poikilitic clasts, 18 subophitic clasts and 10 equigranular clasts (Extended Data Table 1) were used to construct three Pb–Pb leftmost isochrons¹⁰. The radiogenic $^{207}\text{Pb}/^{206}\text{Pb}$ ratios of y intercepts are translated to Pb–Pb ages of $2,027 \pm 7$ million years ago (Ma) (95% confidence level, and hereafter except where otherwise noted), $2,030 \pm 6$ Ma and $2,034 \pm 8$ Ma, respectively (Extended Data Fig. 3a–f). Forty-four Pb isotope analyses were conducted on plagioclase, pyroxene and matrix from two fine-grained porphyritic clasts (without visible Zr-bearing minerals) and yielded a leftmost isochron Pb–Pb age of $2,027 \pm 54$ Ma (Extended Data Fig. 3g, h). Despite the distinct petrographic textures in Chang'e-5 basalts, these four isochrons have consistent y intercepts and slopes within uncertainties (Extended Data Fig. 3), indicating both their identical age and derivation from most probably the same source. Taken together, a total of 159 analyses with negligible terrestrial Pb contamination for various mineral phases form an integrated isochron yielding a Pb–Pb age of $2,030 \pm 4$ Ma (Fig. 2). The age is interpreted as the best estimate of the crystallization age of the Chang'e-5 basalts given that all the clasts studied here show pristine magmatic textures without evident overprinting from shock metamorphism (Fig. 1). This age represents the youngest crystallization age reported so far for lunar basaltic rocks by the radiometric method and thus extends the range of radioisotopic ages of lunar basalt by about 800–900 Myr. Therefore, this study provides conclusive evidence that magmatic activity on the Moon persisted until at least 2 Ga. This insight into the existence of this youngest-known volcanism provides a critical constraint for understanding the thermal mechanisms behind the longevity of lunar magmatism.

Fig. 2: Pb–Pb isochron for the Chang'e-5 basalts.

 figure2

a, The integrated Pb–Pb isochron showing the mixing line between the y intercept as radiogenic $^{207}\text{Pb}/^{206}\text{Pb}$ and the initial Pb compositions ($^{204}\text{Pb}/^{206}\text{Pb} = 0.00228 \pm 0.00011$, $^{207}\text{Pb}/^{206}\text{Pb} = 0.860 \pm 0.019$). **b**, The enlarged lowest part of the isochron in **a** highlighting the measurements of Zr-bearing minerals. The black line is the best-fitted

isochron with an equation of $y = (323 \pm 7)x + (0.12510 \pm 0.00028)$. Error bars represent 1σ s.e. Mtr, matrix; Pl, plagioclase; Pyx, pyroxene.

Source data

Mantle source signature

One of the leading mechanisms considered for sustaining such young lunar volcanism is potassium, rare-earth elements and phosphorus (KREEP)-related radiogenic heating in the basalt source^{1,3,14}; thus characterizing the geochemistry of the mantle source of this youngest-dated mare basalt reported so far provides a critical test of this hypothesis. The initial Pb isotopic composition of the basalts and corresponding time-integrated μ value ($^{238}\text{U}/^{204}\text{Pb}$ ratio) for their mantle source can fingerprint the Chang'e-5 basalt mantle source and its chemical evolution^{6,7,10}. Five of the 106 U/Pb analyses for rock-forming minerals (mainly plagioclase) show negligible U (with $^{238}\text{UO}^+/^{208}\text{Pb}^+$ ratios of <0.01) and clustered $^{207}\text{Pb}/^{206}\text{Pb}$ values of 0.855–0.872 (Supplementary Table 3). Thus, their weighted mean of 0.860 ± 0.019 (2 s.e., $n = 5$, mean squared weighted deviation = 0.11) provides the best estimate for the initial $^{207}\text{Pb}/^{206}\text{Pb}$ ratio of the Chang'e-5 basalts (Extended Data Fig. 4). The initial $^{204}\text{Pb}/^{206}\text{Pb}$ ratio, however, is difficult to measure precisely and accurately owing to extremely low ^{204}Pb counts (less than 0.05 counts per second (cps)) (Supplementary Table 3). Alternatively, it can be calculated as 0.00228 ± 0.00011 (2 s.e.) based on the best-fit Pb–Pb isochron (Fig. 2a) and the aforementioned best estimate of the initial $^{207}\text{Pb}/^{206}\text{Pb}$ ratio. This calculated initial $^{204}\text{Pb}/^{206}\text{Pb}$ ratio is consistent within errors with the weighted mean of 0.00235 ± 0.00043 (2 s.e.) for the two lowest measured $^{204}\text{Pb}/^{206}\text{Pb}$ ratios (Supplementary Table 3), justifying the rationale of the calculation.

Determination of the μ value of a basalt mantle source is dependent on the lunar Pb-isotope evolution model^{6,7,10}. On the basis of the lunar magma ocean (LMO) model¹⁵ that presumably generated all major lunar silicate reservoirs, including the sources for lunar basalts, a two-stage model for lunar Pb-isotopic evolution is proposed^{6,7}. The evolution of the Pb isotopes of a basalt source starts from t_0 (around 4,500 Ma)¹⁶ for Moon formation with $\mu_1 = 462 \pm 46$ for the LMO⁶, to t_1 (around 4,420–4,300 Ma)^{17,18} for LMO crystallization and the formation of major geochemically distinct reservoirs with different μ_2 values, to t_2 for mare basalt formation with the initial Pb isotopes. Although t_1 remains uncertain, its age range has little effect on the calculated μ_2 values (Extended Data Fig. 5). Thus, for comparison with previous results^{6,7}, a t_1 of $4,376 \pm 18$ Ma was selected to calculate a two-stage μ value of 684 ± 40 for the Chang'e-5 basalt source (Fig. 3). This μ value is within the range ($\mu \approx 300$ –1,000) of the low-titanium (Ti) and high-Ti Apollo basalts, but considerably lower than those ($\mu \approx$

2,600–3,700) of the KREEP and high-aluminium (Al) basalts^{6,7} (Fig. 3). This marked difference suggests that the Chang'e-5 basalts from the Procellarum KREEP Terrane were most probably produced by the melting of a KREEP-poor source. An apparent increase in μ values from around 3.4–3.0 Ga for low-Ti Apollo basalts and low-Ti and very-low-Ti basaltic meteorites (NWA 4734 and NWA 773 clan) suggests a progressive contribution of a KREEP-like component in such rocks^{6,7,10}. However, the Chang'e-5 basalts do not follow this trend, indicating that KREEP-like components were not involved in our samples, neither in the deep source nor during shallow contamination of the KREEP material. Corroborating evidence for a non-KREEP source for the Chang'e-5 basalts is provided by strontium–neodymium (Sr–Nd) isotopes¹⁹, but the results in this study using the radioactive element U offer direct evidence for heat-producing elements not being concentrated in the Chang'e-5 basalt mantle source. Thus, these results strongly suggest that the idea of KREEP-induced heating^{1,14,20} for the generation of these young lunar magmas requires further investigation or the consideration of other mechanisms.

Fig. 3: Initial $^{206}\text{Pb}/^{204}\text{Pb}$ ratios versus crystallization ages of the lunar basalts and meteorites.

 figure3



The lines represent the two-stage Pb isotope evolution of lunar mantle sources at given μ values⁶. The gradient areas are associated with KREEP-poor (yellow) to KREEP-rich (blue) mantle sources according to μ values. The Apollo and meteorite data are from refs. [6,7,10](#). Error bars are 2σ s.e. A11, Apollo 11 high-Ti basalts; A12, Apollo 12 low-Ti basalts; A15, Apollo low-Ti basalts; A17, Apollo 17 high-Ti basalts; High-Al, Apollo 14 high-Al basalts; KREEP, Apollo 15 KREEP basalts; Lunar basalt meteorite, low-Ti and very-low-Ti basaltic meteorites (NWA 4734 and NWA 773 clan).

Source data

Anchor point for cratering chronology

Radioisotope ages provide the yardstick for calibrating the age information obtained from crater statistics. The mare unit on which Chang'e-5 landed features both homogeneous reflectance spectra and surface morphology as seen from orbit²¹, and ballistic sedimentation modelling^{22,23} suggests a dominance of local mare basalts in the surface regolith. The basalt clasts returned show uniform geochemical characteristics¹⁹ and a consistent radioisotope age of $2,030 \pm 4$ Ma (Fig. 2), convincingly pointing to an affinity with the mare unit of the landing site. Therefore, our radioisotope age obtained for the newly returned Chang'e-5 samples offer an opportunity to confirm the first-order reliability of the lunar crater-counting chronology established by the Apollo and Luna missions²⁴. The calculated model ages for the mare unit on which Chang'e-5 landed mostly range between around 2.2 Ga and 1.5 Ga using the prevailing crater-counting chronology of ref. ²⁴, within a difference of about 20% compared with the measured radioisotope age (Fig. 4). This difference is surprisingly consistent with that derived from a comparison of the current impact flux on the Moon from observations²⁵ and prediction by crater-counting chronology²⁴.

Fig. 4: The critical reference point of the radioisotope age of the Chang'e-5 basalts for the lunar crater-counting chronology.

 figure4

The lines are colour-coded according to different models of the crater-counting chronology of the Moon. The red dot marks the radioisotope age of the Chang'e-5 basalts and the translated crater density based on the crater-counting chronology function of ref. [24](#). The blue squares are the calibration points established from Apollo and Luna samples [27,28,29,30](#). The inset shows the various crater densities and model ages predicted for the mare unit at the Chang'e-5 landing site (blue dots) [4,12,13,21,23,31,32,33](#). Error bars are 1σ s.e. $N(1)$, number of craters with diameter > 1 km. Different models in the main panel are from refs. [24,27,34,35,36,37](#).

Source data

Substantial differences nonetheless exist among reported crater densities for the mare unit on which Chang'e-5 landed (Fig. 4), albeit a relatively simple geological context and the same crater-counting chronology model were used. Therefore, there is much potential to improve the accuracy of predictions by crater statistics. For the existing crater-counting chronology, Apollo and Luna samples have provided an initial database for ages ranging from around 4.0 Ga to 3.1 Ga as well as those younger than around 1 Ga (Fig. 4). The age of 2.03 Ga obtained for the Chang'e-5 basalts resides squarely in the centre of this large gap (Fig. 4), fulfilling the long-sought-after goal to bridge the unanchored middle portion of the lunar crater-counting chronology²⁴ and improving this critical tool for dating unsampled surfaces on the Moon²⁶, as well as for translating the lunar crater-counting chronology to the other planetary bodies²⁴.

Methods

The sample mounting, scanning electron microscope (SEM), electron probe microanalysis (EPMA) and SIMS analyses were performed at the Institute of Geology and Geophysics, Chinese Academy of Sciences (IGGCAS) in Beijing, China.

SEM analysis

The Chang'e-5 basalt clasts studied were first embedded in epoxy mounts and then polished using a grinder. A Thermo Scientific Apreo SEM equipped with an energy dispersive spectroscopy (EDS) detector was used to identify the Zr-bearing and phosphate minerals. High-resolution back-scatter electron (BSE) imaging and semi-quantitative EDS analyses were conducted using a Zeiss Gemini 450 field-emission environmental SEM. For a large-scale BSE imaging of a single basalt clast, the measurement was performed at an acceleration voltage of 15 kV and a current of 2.0 nA, with a working distance of about 8 mm. For dated minerals in a localized area, the analyses were operated with an acceleration voltage of 5 kV, a beam current of 1.0 nA and a working distance of about 7 mm. The phosphate (apatite and merrillite) and Zr-bearing minerals (baddeleyite, zirconolite and tranquillityite) were examined by EDS (Extended Data Fig. 1).

EPMA analyses for dated minerals

Both Zr-bearing minerals and phosphate were analysed using a CAMECA SXFive FE electron probe microanalyser. An acceleration voltage of 20 kV and beam current of 30 nA were used for all analyses, with a spot size of 1 µm. Data were processed with the phi–rho–Z matrix correction using CAMECA PeakSight software (version 6.2).

Synthetic glasses (single REE oxide–calcium oxide–aluminium oxide–silicon dioxide) from P&H Developments were used as the standards for REE. The standards used for the other elements analysed were periclase (magnesium (Mg)), K-feldspar (aluminium (Al)), rhodonite (silicon (Si), calcium (Ca) and manganese (Mn)), rutile (Ti), chromium oxide (chromium (Cr)), specularite (Fe), yttrium Al garnet (yttrium (Y)), zircon (Zr), niobium metal (Nb), tantalum metal (Ta), tungsten metal (W), cubic zirconia (hafnium (Hf)), apatite (P), fluorite (fluoride (F)), halite (chlorine (Cl)) and celestine (sulfur (S) and Sr). The methodology for the analysis of REE followed that of ref. [38](#). The detection limits of Zr, REE and Y vary from about 100 ppm to about 300 ppm, whereas those for major elements are 60–120 ppm. Representative analysis results for phosphate and Zr-bearing minerals are listed in Supplementary Tables [1](#), [2](#), respectively.

SIMS analyses

The target selection strategy was to identify phases that would contain initial Pb and those containing radiogenic Pb generated from in situ decay of U since crystallization of the sample (for example, phosphates and Zr-bearing phases). The distinct Pb isotopic ratios yielded by these two types of phase help to populate the isochron and calculate a precise date. The target areas within these phases were selected to be large enough to accommodate a SIMS analytical spot (in this case, the smallest spot used was less than 3 μm). The Pb isotopic compositions (complete dataset presented in Supplementary Table [3](#)) of the phases were determined over three analytical sessions using a CAMECA IMS 1280HR ion microprobe. The mounts with candidate minerals were cleaned with a fine (0.25 μm) diamond paste and ethanol to remove the carbon coating before adding a roughly 20-nm gold coating.

The first session focused on the Pb-isotope measurement of Zr-bearing minerals. A Gaussian illumination mode was used to focus a primary beam of $^{16}\text{O}^-$ to a size of less than 3 μm (about 2.8 μm) (Fig. [1](#), Extended Data Fig. [2](#)), with an accelerated potential of –13 kV. The beam size can be kept unchanged for a long usage time and intensities were around 250–200 pA. The primary beam setting is described in detail in ref. [39](#). The multi-collector mode with five electron multipliers with low noise (less than 0.001 cps, especially for L2) was used to measure $^{204}\text{Pb}^+$ (L2), $^{206}\text{Pb}^+$ (L1), $^{207}\text{Pb}^+$ (C), $^{208}\text{Pb}^+$ (H1) and $^{96}\text{Zr}_2^{16}\text{O}_2^+$ (H2). The methodology is similar to that outlined in ref. [40](#). Exit slit 3 was used, with a mass resolving power (MRP) of 8,000 (50% peak height). Before analysis, a primary beam of $^{16}\text{O}^-$ with an intensity of 10 nA was used for 120 s of pre-sputtering. The ion images with $^{96}\text{Zr}_2^{16}\text{O}_2^+$ and Pb isotopes on a 25 $\mu\text{m} \times 25\text{-}\mu\text{m}$ area were used to precisely locate the target minerals. The signal of ^{206}Pb was used for peak-centring reference. Each measurement consisted of 4 s \times 80 cycles, with a total analytical time of about 10 min. High-purity oxygen gas was leaked onto

the sample surface to enhance the Pb^+ yield to more than 15 cps $\text{ppm}^{-1} \text{nA}^{-1}$ by using a O^- primary beam according to the M257 zircon standard (561 Ma, 840 ppm U, ref. 41). NIST610 glass and Phalaborwa baddeleyite standard ($^{207}\text{Pb}^*/^{206}\text{Pb}^* = 0.1272$, ref. 42) were used to calibrate the relative yield of different electron multipliers and evaluate the external reproducibility. On the basis of 21 analyses on NIST610 glass under the same analytical conditions, the $^{207}\text{Pb}/^{206}\text{Pb}$ measurements have a relative standard deviation (1 r.s.d.) of 0.66% with ^{207}Pb intensity averaged at 127 cps. The possible SIMS instrumental mass fractionation of Pb isotopes around 0.2% (ref. 43) was propagated to the uncertainty of single-spot $^{207}\text{Pb}/^{206}\text{Pb}$ analysis.

The second session focused on the Pb-isotope measurement of phosphates. A Gaussian illumination mode was used to focus a primary beam of $^{16}\text{O}_2^-$ to a roughly 8- μm size with the intensity kept around 2.5 nA. Ion images with Ca_2PO_3^+ on a 30 $\mu\text{m} \times$ 30- μm area were used to precisely locate the target minerals. The methodology is similar to that outlined in ref. 44. Exit slit 3 was used, with an MRP of 8,000 (50% peak height). The dynamic multi-collector mode was used to measure $^{204}\text{Pb}^+$ (L2), $^{206}\text{Pb}^+$ (L1) and $^{207}\text{Pb}^+$ (C) in the first step with 60-s counting time, and ^{238}U (L1), $^{232}\text{ThO}^+$ (H1) and $^{238}\text{UO}^+$ (H2) in the second step with 4-s counting time. Each measurement consisted of 10 cycles, with a total analytical time of about 15 min including 2 min of pre-sputtering. NIST610 glass and Phalaborwa baddeleyite ($^{207}\text{Pb}^*/^{206}\text{Pb}^* = 0.1272$; ref. 42) were used to calibrate the relative yield of different electron multipliers and evaluate the external reproducibility. On the basis of 20 analyses on Phalaborwa baddeleyite under the same analytical conditions, the $^{207}\text{Pb}/^{206}\text{Pb}$ measurements show 1 r.s.d. of 0.2% with ^{207}Pb intensity averaged at 970 cps.

The third session focused on the initial Pb composition test on the essential minerals (mainly plagioclase and pyroxene) and matrix. A Köhler illumination mode was used to produce a primary beam of about 30 nA O_2^- to a roughly 30- μm size. Before each measurement, an area of 25 μm around the spot location was raster-scanned for 120 s to remove the gold coating and minimize possible surface contamination. The multi-collector mode with five electron multipliers was used to measure $^{204}\text{Pb}^+$ (L2), $^{206}\text{Pb}^+$ (L1), $^{207}\text{Pb}^+$ (C) and $^{208}\text{Pb}^+$ (H1) in the first step with 60-s counting time, and $^{232}\text{ThO}^+$ (H1) and $^{238}\text{UO}^+$ (H2) in the second step with 4-s counting time. Exit slit 3 was used, with an MRP of 8,000 (50% peak height), sufficient to resolve Pb from known molecular interferences. Each measurement consisted of 10 cycles. On the basis of 28 analyses on NIST614 glass under the same analytical conditions, the $^{207}\text{Pb}/^{206}\text{Pb}$ measurements show 1 r.s.d. of 0.68% with ^{207}Pb intensity averaged at 120 cps.

The background counts for each channel were measured at regular intervals during each session by using deflector and aperture settings that effectively blank both

primary and any residual secondary beams. The average background values are reported in Extended Data Table 2. The number of SIMS Pb isotope analyses for each type of basalt clast used in this study are summarized in Extended Data Table 1.

SIMS data processing

The data were processed using in-house SIMS data reduction spreadsheets and the Excel add-in Isoplot (version 4.15; ref. 45). The Pb–Pb isochrons were constructed following the method first established by ref. 46, then applied to SIMS by ref. 6 and further expressed and refined by refs. 7,10. In brief, the Pb isotopic compositions measured in each sample are interpreted as representing a mixture between three main components: (1) initial Pb present in the basaltic melt when it crystallized; (2) radiogenic Pb formed by the decay of U in the basalt after crystallization ;and (3) terrestrial contamination. These end-member components define a triangular array of points on a plot of $^{207}\text{Pb}/^{206}\text{Pb}$ versus $^{204}\text{Pb}/^{206}\text{Pb}$ (Extended Data Fig. 3). The values with the highest $^{207}\text{Pb}/^{206}\text{Pb}$ ratios, those at the top of the triangular array, provide an estimate of the lowest possible value for the initial Pb composition of the sample. The radiogenic Pb component is then located where $^{204}\text{Pb}/^{206}\text{Pb} = 0$. Finally, given the radiogenic Pb isotopic compositions associated with the Moon relative to those found on Earth, the terrestrial contamination end-member will have the highest $^{204}\text{Pb}/^{206}\text{Pb}$ ratios. Any obvious cracks or voids in the sample would be likely places for terrestrial contamination to accumulate during sample polishing and cleaning procedures.

Despite the efforts made to avoid such regions, it was difficult to ensure that the SIMS spots did not overlap with small invisible cracks, or did not depth-profile into such features lying just below the original surface of the sample. Furthermore, given the low Pb concentrations in many of the analysed phases (particularly those with lower radiogenic Pb isotope compositions), these analyses are particularly susceptible to even low levels of terrestrial contamination. On the basis of these assumptions, the bounding edge on the left side of the triangle, between the initial and radiogenic lunar Pb compositions, forms an isochron, which can be determined by iteratively filtering the data to yield the steepest statistically significant weighted regression. The $^{207}\text{Pb}/^{206}\text{Pb}$ of initial Pb was estimated by spots with the highest $^{207}\text{Pb}/^{206}\text{Pb}$ and near-zero $^{238}\text{UO}^+/^{208}\text{Pb}^+$ (Extended Data Fig. 4), following the procedures of refs. 7,10.

The Pb concentrations were estimated based on the Pb yield assuming that Zr-bearing minerals have comparable Pb yield with that in zircon (that is, $15 \text{ cps ppm}^{-1} \text{nA}^{-1}$ with a primary beam of O^-). Baddeleyite, tranquillityite and zirconolite grains have U contents of approximately 100–4,000 ppm, 600–7,000 ppm and 800–4,600 ppm, respectively, whereas the phosphate grains have lower U contents of approximately 12–150 ppm estimated from UO^+ yield based on Durango apatite.

Data availability

All data generated or analysed during this study are available in EarthChem Library at <https://doi.org/10.26022/IEDA/112085>. [Source data](#) are provided with this paper.

Code availability

No code is used in this study.

References

1. 1.
Shearer, C. K. et al. Thermal and magmatic evolution of the Moon. *Rev. Mineral. Geochem.* **60**, 365–518 (2006).
2. 2.
Fernandes, V. A., Burgess, R. & Turner, G. ^{40}Ar – ^{39}Ar chronology of lunar meteorites northwest Africa 032 and 773. *Meteorit. Planet. Sci.* **38**, 555–564 (2003).
3. 3.
Borg, L. E., Shearer, C. K., Asmenron, Y. & Papike, J. J. Prolonged KREEP magmatism on the Moon indicated by the youngest dated lunar igneous rock. *Nature* **432**, 209–211 (2004).
4. 4.
Hiesinger, H., Head, J. W. III, Wolf, U., Jaumann, R. & Neukum, G. Ages and stratigraphy of mare basalts in Oceanus Procellarum, Mare Nubium, Mare Cognitum, and Mare Insularum. *J. Geophys. Res. Planets* **108**, 5065–5091 (2003).
5. 5.
Braden, S. E. et al. Evidence for basaltic volcanism on the Moon within the past 100 million years. *Nat. Geosci.* **7**, 878–791 (2014).
6. 6.

Snape, J. F. et al. Lunar basalt chronology, mantle differentiation and implications for determining the age of the Moon. *Earth Planet. Sci. Lett.* **451**, 149–158 (2016).

7. 7.

Snape, J. F. et al. The timing of basaltic volcanism at the Apollo landing sites. *Geochim. Cosmochim. Acta* **266**, 29–53 (2019).

8. 8.

Head, J. W. Lunar volcanism in space and time. *Rev. Geophys.* **14**, 265–300 (1976).

9. 9.

Snape, J. F. et al. Ancient volcanism on the Moon: insights from Pb isotopes in the MIL 13317 and Kalahari 009 lunar meteorites. *Earth Planet. Sci. Lett.* **502**, 84–95 (2018).

10. 10.

Merle, R. E. et al. Pb–Pb ages and initial Pb isotopic composition of lunar meteorites: NWA 773 clan, NWA 4734, and Dhofar 287. *Meteorit. Planet. Sci.* **55**, 1808–1832 (2020).

11. 11.

Hiesinger, H., Head, J. W. III, Wolf, U., Jaumann, R. & Neukum, G. Ages and stratigraphy of lunar mare basalts in Mare Frigoris and other nearside maria based on crater size–frequency distribution measurements. *J. Geophys. Res. Planets* **115**, E03003 (2010).

12. 12.

Morota, T. et al. Timing and characteristics of the latest mare eruption on the Moon. *Earth Planet. Sci. Lett.* **302**, 255–266 (2011).

13. 13.

Qian, Y. Q. et al. Young lunar mare basalts in the Chang'e-5 sample return region, northern Oceanus Procellarum. *Earth Planet. Sci. Lett.* **555**, 116702 (2021).

14. 14.
Wieczorek, M. A. & Phillips, R. J. The “Procellarum KREEP Terrane”: implications for mare volcanism and lunar evolution. *J. Geophys. Res. Planets* **105**, 20417–20420 (2000).
15. 15.
Wood, J. A., Dickey, J. S. Jr, Marvin, U. B. & Powell, B. N. Lunar anorthosites. *Science* **167**, 602–604 (1970).
16. 16.
Touboul, M., Kleine, T., Bourdon, B., Palme, H. & Wieler, R. Late formation and prolonged differentiation of the Moon inferred from W isotopes in lunar metals. *Nature* **450**, 1206–1209 (2007).
17. 17.
Nyquist, L. E. & Shih, C. Y. The isotopic record of lunar volcanism. *Geochim. Cosmochim. Acta* **56**, 2213–2234 (1992).
18. 18.
Nyquist, L. E. et al. Lunar crustal history recorded in lunar anorthosites. *Lunar Planet. Sci. Conf.* **41**, 1383 (2010).
19. 19.
Tian, H.-C. et al. Non-KREEP origin for Chang'e-5 basalts in the Procellarum KREEP Terrane. *Nature*, <https://doi.org/10.1038/s41586-021-04119-5> (2021).
20. 20.
Hess, P. C. & Parmentier, E. M. Thermal evolution of a thicker KREEP liquid layer. *J. Geophys. Res. Planets* **106**, 28023–28032 (2001).
21. 21.
Qian, Y. Q. et al. Geology and scientific significance of the Rümker region in northern Oceanus Procellarum: China's Chang'e-5 landing region. *J. Geophys. Res. Planets* **123**, 1407–1430 (2018).
22. 22.

Xie, M., Xiao, Z., Zhang, X. & Xu, A. The provenance of regolith at the Chang'e-5 candidate landing region. *J. Geophys. Res. Planets* **125**, e2019JE006112 (2020).

23. 23.

Qian, Y. Q. et al. China's Chang'e-5 landing site: geology, stratigraphy, and provenance of materials. *Earth Planet. Sci. Lett.* **561**, 116855 (2021).

24. 24.

Neukum, G., Ivanov, B. A. & Hartmann, W. K. Cratering records in the inner Solar System in relation to the lunar reference system. *Space Sci. Rev.* **96**, 55–86 (2001).

25. 25.

Speyerer, E. J., Povilaitis, R. Z., Robinson, M. S., Thomas, P. C. & Wagner, R. V. Quantifying crater production and regolith overturn on the Moon with temporal imaging. *Nature* **538**, 215–218 (2016).

26. 26.

Bogert, C. H. & Hiesinger, H. Which samples are needed for improved calibration of the lunar cratering chronology? *Lunar Planet. Sci. Conf.* **51**, 2088 (2020).

27. 27.

Robbins, S. J. New crater calibrations for the lunar crater-age chronology. *Earth Planet. Sci. Lett.* **403**, 188–198 (2014).

28. 28.

Hiesinger, H. et al. How old are young lunar craters? *J. Geophys. Res. Planets* **117**, E00H10 (2012).

29. 29.

Hiesinger, H., Head, J. W. III, Wolf, U., Jaumann, R. & Neukum, G. in *Recent Advances and Current Research Issues in Lunar Stratigraphy (GSA Special Paper 477)* (eds Ambrose, W. A. & Williams, D. A.) 1–51, <https://doi.org/10.1130/SPE477> (GSA, 2011).

30. 30.

Stöffler, D. & Ryder, G. Stratigraphy and isotope ages of lunar geologic units: chronological standard for the inner Solar System. *Space Sci. Rev.* **96**, 9–54 (2001).

31. 31.

Jia, M. et al. A catalogue of impact craters larger than 200 m and surface age analysis in the Chang'e-5 landing area. *Earth Planet. Sci. Lett.* **541**, 116272 (2020).

32. 32.

Xu, Z., Guo, D. & Liu, J. Maria basalts chronology of the Chang'e-5 sampling site. *Remote Sens.* **13**, 1515 (2021).

33. 33.

Wu, B., Huang, J., Li, Y., Wang, Y. & Peng, J. Rock abundance and crater density in the candidate Chang'e-5 landing region on the Moon. *J. Geophys. Res. Planets* **123**, 3256–3272 (2018).

34. 34.

Le Feuvre, M. & Wieczorek, M. A. Nonuniform cratering of the Moon and a revised crater chronology of the inner Solar System. *Icarus*, **214**, 1–20 (2011).

35. 35.

Marchi, S. et al. A new chronology for the Moon and Mercury. *Astronomic. J.* **137**, 4936–4948 (2009).

36. 36.

Wagner, R. et al. Stratigraphic sequence and ages of volcanic units in the Gruithuisen region of the Moon. *J. Geophys. Res.* **107**, 5104 (2002).

37. 37.

Hartmann, W. K., Quantin, C. & Mangold, N. Possible long-term decline in impact rates: 2. Lunar impact-melt data regarding impact history. *Icarus* **186**, 11–23 (2007).

38. 38.

Williams, C. T. in *Rare Earth Minerals: Chemistry, Origin and Ore Deposits* (eds Jones, A. P. et al.) 327–348 (Springer, 1996).

39. 39.

Liu, Y., Li, X. H., Li, Q. L. & Tang, G. Q. Breakthrough of 2- to 3- μm scale U–Pb zircon dating using Cameca IMS-1280HR SIMS. *Surf. Interface Anal.* **52**, 214–223 (2020).

40. 40.

Li, Q. L. et al. Precise U–Pb and Pb–Pb dating of Phanerozoic baddeleyite by SIMS with oxygen flooding technique. *J. Anal. At. Spectrom.* **25**, 1107–1113 (2010).

41. 41.

Nasdala, L. et al. Zircon M257—a homogeneous natural reference material for the ion microprobe U–Pb analysis of zircon. *Geostand. Geoanal. Res.* **32**, 247–265 (2008).

42. 42.

Heaman, L. M. The application of U–Pb geochronology to mafic, ultramafic and alkaline rocks: an evaluation of three mineral standards. *Chem. Geol.* **261**, 43–52 (2009).

43. 43.

Stern, R. et al. Measurement of SIMS instrumental mass fractionation of Pb isotopes during zircon dating. *Geostand. Geoanal. Res.* **33**, 145–168 (2009).

44. 44.

Li, Q. L. et al. In-situ SIMS U–Pb dating of Phanerozoic apatite with low U and high common Pb. *Gondwana Res.* **21**, 745–756 (2012).

45. 45.

Ludwig, K. R. Isoplot 4.15: a geochronological toolkit for Microsoft Excel (Berkeley Geochronology Center, 2008).

46. 46.

Connelly, J. N. et al. The absolute chronology and thermal processing of solids in the solar proto-planetary disk. *Science* **338**, 651–655 (2012).

47. 47.

Steiger, R. H. & Jäger, E. Subcommission on geochronology: convention on the use of decay constants in geo- and cosmochronology. *Earth Planet. Sci. Lett.* **36**, 359–362 (1977).

48. 48.

Hiess, J., Condon, D. J., McLean, N. & Noble, S. R. $^{238}\text{U}/^{235}\text{U}$ systematics in terrestrial uranium-bearing minerals. *Science* **335**, 1610–1614 (2012).

49. 49.

Göpel, C., Manhès, G. & Allègre, C. J. U–Pb systematics in iron meteorites—uniformity of primordial lead. *Geochim. Cosmochim. Acta* **49**, 1681–1695 (1985).

Acknowledgements

We thank all the staff of China's Chang'e Lunar Exploration Project for their hard work in returning lunar samples. The samples studied in this work were allocated by the China National Space Administration. We thank R. Mitchell for constructive comments and editing the manuscript; W. Yang, Y. Chen, H.-J. Hui and H.-C. Tian for helpful discussion; and C. Sun for logistical support. This study was funded by the Key Research Program of the Chinese Academy of Sciences (ZDBS-SSW-JSC007-13), the Institute of Geology and Geophysics, Chinese Academy of Sciences (IGGCAS-202101), the National Natural Science Foundation of China (41773044) and the pre-research project on Civil Aerospace Technologies of China National Space Administration (grant number D020203).

Author information

Author notes

1. These authors contributed equally: Qiu-Li Li, Qin Zhou

Affiliations

1. State Key Laboratory of Lithospheric Evolution, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing, China

Qiu-Li Li, Yu Liu, Hong-Xia Ma, Guo-Qiang Tang, Shun Guo, Jiang-Yan Yuan, Jiao Li, Fu-Yuan Wu & Xian-Hua Li

2. Key Laboratory of Lunar and Deep Space Exploration, National Astronomical Observatories, Chinese Academy of Sciences, Beijing, China

Qin Zhou & Chunlai Li

3. Planetary Environmental and Astrobiological Research Laboratory, School of Atmospheric Sciences, Sun Yat-sen University, Zhuhai, China

Zhiyong Xiao

4. Key Laboratory of Earth and Planetary Physics, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing, China

Yangting Lin, Jin-Hua Li & Xu Tang

5. Center for Lunar and Planetary Sciences, Institute of Geochemistry, Chinese Academy of Sciences, Guiyang, China

Ziyuan Ouyang

Contributions

X.-H.L. and C.L. conceived and supervised this project. Q.-L.L., Q.Z., and Z.X. wrote the manuscript with input of X.-H.L., C.L., Y.L. and F.-Y.W. H.-X.M. and J.L. conducted sample mounting. Q.Z., J.-H.L., X.T., S.G. and J.-Y.Y. performed SEM and EPMA analyses. Q.-L.L., Y.L., G.-Q.T. and Q.Z. conducted SIMS analyses, data processing and interpretation. Z.O. and F.-Y.W. contributed scientific background and geological context.

Corresponding authors

Correspondence to [Chunlai Li](#) or [Xian-Hua Li](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Charles Shearer and Renaud Merle 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 Plots of the representative energy dispersive X-ray spectrum for dated minerals.](#)

Element concentrations detected by EDS are also shown for comparison. Mer, merrillite; Ap, apatite; Bdy, baddeleyite; Zrl, zirconolite; Trq, tranquillityite.

[Extended Data Fig. 2 Microtextures of dated minerals in the Chang'e-5 basalt clasts.](#)

a, b, A porphyritic basalt clast that contains fine-grained ($2 \times 10 \mu\text{m}$) apatite grains. The yellow rectangle in **a** is expanded in **b**. The apatite occurs as inclusions in Fe-rich clinopyroxene and is surrounded by tiny ilmenite aggregates. **c**, In a poikilitic clast, baddeleyite is rimmed by zirconolite. **d**, Square-shaped baddeleyite inclusions in clinopyroxene and ilmenite from an equigranular clast. Hexagonal apatite exhibits an intergranular phase between clinopyroxene and ilmenite. **e**, Euhedral–subhedral apatite and baddeleyite show an equilibrium texture with Fe-rich olivine ($\text{Fo} < 10$), clinopyroxene, and troilite. **f**, Baddeleyite and tranquillityite show intergrowths along the margins of clinopyroxene and Fe-rich olivine. **g**, In an equigranular clast, baddeleyite, tranquillityite and apatite are intergrown with Fe-rich olivine. **h**, Tranquillityite and apatite intergrown crystals in a poikilitic clast. Blurry pits in Zr-bearing minerals are the analytical spots from SIMS. Bdy, baddeleyite; Zrl, zirconolite; Trq, tranquillityite; Mer, merrillite; Ap, Apatite; Cpx, clinopyroxene; Ol, olivine; Pl, plagioclase; Kfs, K-feldspar; Ilm, ilmenite; Crs, cristobalite; Tro, troilite; Spl, spinel.

[Extended Data Fig. 3 Pb–Pb isochrons for the Chang'e-5 basalts with different textures.](#)

The left four plots (**a**, **c**, **e** and **g**) show the data from basalt clasts with poikilitic, subophitic, equigranular and porphyritic textures, respectively. The equations of isochrons (black lines) are shown on the top. The right four plots (**b**, **d**, **f** and **h**) are the enlarged lowest portions of the isochrons highlighting the measurements with low $^{204}\text{Pb}/^{206}\text{Pb}$. The red hexagon represents the initial Pb determined from the integrated isochron, but is not included in each separated isochron. The triangle areas represent the mixing trend among the initial Pb component, the radiogenic Pb, and current terrestrial Pb composition. Outliers excluded from the calculation of the isochron regression are shown in grey while those data used for the leftmost isochron are shown in colour. Error bars represent 1σ standard error. The uncertainties for the isochron dates are quoted at the 95% confidence level.

[Source data](#)

Extended Data Fig. 4 Plot of UO^+/Pb^+ versus $^{207}\text{Pb}/^{206}\text{Pb}$ for points within the analysed main rock-forming minerals.

Five blue dots highlight the points with $\text{UO}^+/\text{Pb}^+ < 0.01$ (the grey line), that most likely represent the best estimate for the initial Pb composition. Note that the measured UO^+/Pb^+ ratios are simply used here to provide an indication of the U/Pb ratios. Error bars represent 1σ standard error.

[Source data](#)

Extended Data Fig. 5 The two-stage Pb evolution model and μ value calculation procedures.

a, The schematic diagram of two-stage Pb evolution model⁶. **b**, The equations used in the two-stage Pb evolution model. **c**, The basic parameters used, including the decay constant of ^{238}U and ^{235}U ⁴⁷ and $^{235}\text{U}/^{238}\text{U}$ ratio⁴⁸. **d**, The summarized parameters from previous work^{6,49} and this study. **e**, The calculated μ value based on different modelled ages.

Extended Data Table 1 Summary of SIMS Pb isotope analyses for each type of Chang'e-5 basalt clast

Extended Data Table 2 Background measurements for electron multiplier collectors

Supplementary information

[Peer Review File](#)

Supplementary Tables 1–3

This file contains Supplementary Tables 1–3. Supplementary Table 1, EMPA data of phosphate minerals; Supplementary Table 2, EMPA data of Zr-bearing minerals; Supplementary Table 3, Pb–Pb isotope data for the Chang'e-5 basalts.

Source data

Source Data Fig. 2

Source Data Fig. 3

Source Data Fig. 4

Source Data Extended Data Fig. 3

Source Data Extended Data Fig. 4

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

Li, QL., Zhou, Q., Liu, Y. *et al.* Two-billion-year-old volcanism on the Moon from Chang'e-5 basalts. *Nature* **600**, 54–58 (2021). <https://doi.org/10.1038/s41586-021-04100-2>

- Received: 28 July 2021
- Accepted: 06 October 2021
- Published: 19 October 2021
- Issue Date: 02 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04100-2>

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

- [**Non-KREEP origin for Chang'e-5 basalts in the Procellarum KREEP Terrane**](#)
 - Heng-Ci Tian
 - Hao Wang
 - Fu-Yuan Wu
- [**A dry lunar mantle reservoir for young mare basalts of Chang'e-5**](#)
 - Sen Hu
 - Huicun He
 - Ziyuan Ouyang

Nature (2021)

- [Chang'e-5 samples reveal two-billion-year-old volcanic activity on the Moon and its source characteristics](#)

- Shoujie Liu
- Qin Zhou
- Wei Yang

Science China Earth Sciences (2021)

[Robotic sample return reveals lunar secrets](#)

- Richard W. Carlson

News & Views 01 Dec 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04100-2>

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

- Article
- Open Access
- [Published: 19 October 2021](#)

Non-KREEP origin for Chang'e-5 basalts in the Procellarum KREEP Terrane

- [Heng-Ci Tian](#) ORCID: orcid.org/0000-0002-6556-3096¹ na1,
- [Hao Wang](#) ORCID: orcid.org/0000-0002-0021-7563² na1,
- [Yi Chen](#) ORCID: orcid.org/0000-0001-6167-8032² na1,
- [Wei Yang](#) ORCID: orcid.org/0000-0001-6267-4981¹,
- [Qin Zhou](#) ORCID: orcid.org/0000-0001-8562-9737³,
- [Chi Zhang](#) ORCID: orcid.org/0000-0001-8538-0505¹,
- [Hong-Lei Lin](#) ORCID: orcid.org/0000-0002-9945-5050¹,
- [Chao Huang](#) ORCID: orcid.org/0000-0003-4921-9612²,
- [Shi-Tou Wu](#) ORCID: orcid.org/0000-0002-8420-9877²,
- [Li-Hui Jia](#) ORCID: orcid.org/0000-0002-4000-4183²,
- [Lei Xu](#) ORCID: orcid.org/0000-0002-5394-3905²,
- [Di Zhang](#) ORCID: orcid.org/0000-0003-1308-4162²,
- [Xiao-Guang Li](#) ORCID: orcid.org/0000-0001-6853-2458²,
- [Rui Chang](#)¹,
- [Yue-Heng Yang](#) ORCID: orcid.org/0000-0002-2504-1111²,
- [Lie-Wen Xie](#)²,
- [Dan-Ping Zhang](#) ORCID: orcid.org/0000-0002-6630-2963²,
- [Guang-Liang Zhang](#) ORCID: orcid.org/0000-0001-8554-2818³,
- [Sai-Hong Yang](#) ORCID: orcid.org/0000-0003-2380-4338³ &
- [Fu-Yuan Wu](#) ORCID: orcid.org/0000-0002-2281-7885²

Nature volume 600, pages 59–63 (2021)

- 5579 Accesses
- 5 Citations
- 131 Altmetric

- [Metrics details](#)

Subjects

- [Geochemistry](#)
- [Petrology](#)

Abstract

Mare volcanics on the Moon are the key record of thermo-chemical evolution throughout most of lunar history^{1,2,3}. Young mare basalts—mainly distributed in a region rich in potassium, rare-earth elements and phosphorus (KREEP) in Oceanus Procellarum, called the Procellarum KREEP Terrane (PKT)⁴—were thought to be formed from KREEP-rich sources at depth^{5,6,7}. However, this hypothesis has not been tested with young basalts from the PKT. Here we present a petrological and geochemical study of the basalt clasts from the PKT returned by the Chang'e-5 mission⁸. These two-billion-year-old basalts are the youngest lunar samples reported so far⁹. Bulk rock compositions have moderate titanium and high iron contents with KREEP-like rare-earth-element and high thorium concentrations. However, strontium-neodymium isotopes indicate that these basalts were derived from a non-KREEP mantle source. To produce the high abundances of rare-earth elements and thorium, low-degree partial melting and extensive fractional crystallization are required. Our results indicate that the KREEP association may not be a prerequisite for young mare volcanism. Absolving the need to invoke heat-producing elements in their source implies a more sustained cooling history of the lunar interior to generate the Moon's youngest melts.

[Download PDF](#)

Main

On 17 December 2020, China's Chang'e-5 mission returned about 1.73 kg of lunar materials from one of the youngest basalt units in northern Oceanus Procellarum^{8,10}. The samples studied here include two epoxy mounts, each containing two basalt clasts and two soils allocated by the China National Space Administration (Fig. 1, Extended Data Table 1). The basalt clasts in epoxy mounts and the soils were scooped from the lunar surface. The lithic clasts (larger than 0.6 mm) were picked from the two soil samples. Approximately 45% of lithic clasts are basalt. The basalt clasts show a range of textures from porphyritic to subophitic, poikilitic and equigranular (Extended Data Fig. 1). Eighteen representative basalt clasts with various textures were selected for petrological and geochemical analyses (Extended Data Table 1). The basalt clasts

analysed are composed of clinopyroxene, plagioclase, olivine and ilmenite, as well as minor amounts of K-feldspar, silica, spinel, apatite, baddeleyite, zirconolite, tranquillityite and merrillite (Fig. 2a, Extended Data Table 2). Uranium (U)-rich minerals (baddeleyite, zirconolite and tranquillityite) of 13 clasts (Extended Data Table 1) were dated by the in situ lead (Pb)–Pb method, which suggested a crystallization age of $2,030 \pm 4$ million years ago (Ma)⁹.

Fig. 1: Chang'e-5 soil samples.

 figure1



Photo of soil samples CE5C0100YJFM00103 and CE5C0400YJFM00406. Photo taken by Hui Ren.

Fig. 2: Petrography and mineral chemistry of the Chang'e-5 basalt clasts.

 **figure2**

a, Backscattered electron image of Chang'e-5 basalt clast 041GP, 001. Cpx, clinopyroxene; Ilm, ilmenite; Pig, pigeonite; Pl, plagioclase; Ol, olivine; Spl, Spinel; Tro, troilite. **b**, Mg# versus TiO₂ diagram of the Chang'e-5 basalt clasts. The Apollo and Luna data are from Clive Neal's Mare Basalt Database

(<https://www3.nd.edu/~cneal/Lunar-L/>). **c**, Quadrilateral diagram of pyroxene in the Chang'e-5 basalt clasts. Temperature contours³³ calculated at 0.5 GPa are shown, where 0.5 GPa was chosen according to the possible pressure range for lunar low-Ti basalts³⁴. The Apollo 12 samples³⁵ are plotted (grey area) for comparison. Di, diopside; En, enstatite; Fs, ferrosilite; Hd, hedenbergite. **d**, Comparison of olivine compositions in Chang'e-5 basalts with those within the Apollo 12 and Apollo 15 basalts³⁶. The data for pyroxene and olivine are provided in Supplementary Table 1.

Source Data

We analysed major and trace elements and strontium–neodymium (Sr–Nd) isotopes for different minerals in the basalt clasts. Our results show that most pyroxene and olivine grains have a low magnesium number (Mg#) (Supplementary Table 1, Fig. 2) and generally show compositional zoning with Mg-rich cores and iron-rich rims (Extended Data Fig. 2). The chemical compositions of pyroxene indicate crystallization temperatures ranging from 1,200 °C to 1,000 °C for the cores and below 800 °C for the rims (Fig. 2c). Plagioclase and K-feldspar in all clasts are homogeneous from the core to the rim within the grains, but show considerable compositional

variations among different grains (anorthite (An)_{76–90} and orthoclase (Or)_{62–93}; Extended Data Fig. [3a](#)). Pyroxene shows parallel rare-earth element (REE) patterns with a large variation from 10 to 50 times that of carbonaceous chondrites. (Supplementary Table [2](#), Extended Data Fig. [4](#)). Fourteen plagioclase and five merrillite grains in the clasts yield homogeneous and low initial $^{87}\text{Sr}/^{86}\text{Sr}$ ratios of 0.69934 to 0.69986 and positive $\varepsilon_{\text{Nd}}(t)$ values of 7.9 to 9.3, respectively (Extended Data Tables [3](#), [4](#)). $\varepsilon_{\text{Nd}}(t) = ((^{143}\text{Nd}/^{144}\text{Nd})_{\text{sample}}(t)/(^{143}\text{Nd}/^{144}\text{Nd})_{\text{CHUR}} - 1) \times 10,000$, where $^{143}\text{Nd}/^{144}\text{Nd})_{\text{sample}}(t)$ and $^{143}\text{Nd}/^{144}\text{Nd})_{\text{CHUR}}$ are the Nd isotopic compositions of sample and Chondritic Uniform Reservoir (CHUR) at time ($t = 2,030$ Ma), respectively. Both the calculated $^{147}\text{Sm}/^{144}\text{Nd}$ and $^{87}\text{Rb}/^{86}\text{Sr}$ of the source do not lie on the array of potassium, rare-earth elements and phosphorus (KREEP)-rich material, but rather indicate a light-REE (LREE)-depleted mantle source (Fig. [3](#)).

Fig. 3: Rb–Sr and Sm–Nd isotopic evolution of lunar materials.

 figure3

a, $^{87}\text{Rb}/^{86}\text{Sr}$ ratios of Chang'e-5 basalt source regions are calculated assuming a single-stage model in which the Moon differentiated at 4.56 Ga with an initial $^{87}\text{Sr}/^{86}\text{Sr}$ ratio of 0.69903 (refs. [37,38](#)). The bulk Moon $^{87}\text{Rb}/^{86}\text{Sr}$ value is from refs. [38,39](#). **b**, The $^{147}\text{Sm}/^{144}\text{Nd}$ ratios of the basalt source regions are calculated assuming a

two-stage growth model following refs. [20,21](#). In this model, the Moon followed a chondritic path until differentiation occurred at 4.42 ± 0.07 Ga, represented by the model age of primeval KREEP formation^{[40,41](#)}. The Sr and Nd isotopic data of the Chang'e-5 basalts were acquired on plagioclase and merrillite, respectively (Extended Data Tables [3, 4](#)). The initial $^{87}\text{Sr}/^{86}\text{Sr}$ and $\varepsilon_{\text{Nd}}(t)$ are calculated using 2.03 Ga (ref. [9](#)). The horizontal solid lines in both panels refer to the primordial reservoir. BABI, Basaltic Achondrite Best Initial; CHUR, Chondrite Uniform Reservoir. The Apollo mare basalts and meteorites data are from ref. [21](#) and references therein.

Source Data

Although the studied clasts have a range of petrographic textures and modal abundances of minerals (Extended Data Table [2](#)), three lines of evidence suggest that they are most likely from a single basaltic lava flow. First, the initial plagioclase $^{87}\text{Sr}/^{86}\text{Sr}$ ratios of 11 clasts and the merrillite $\varepsilon_{\text{Nd}}(t)$ of three clasts are nearly identical (Fig. [3](#)). Second, most of the pyroxene grains in these clasts fall along the 1:2 line on a titanium/aluminium (Ti/Al) diagram (Extended Data Fig. [3b](#)), reflecting the near-simultaneous crystallization of augite and plagioclase during one cooling event^{[11](#)}. Lastly, both the pyroxene and plagioclase grains measured in different clasts show similar REE patterns (Extended Data Fig. [4](#)), like those in the Apollo 12 and Apollo 15 basalts with different cooling rates^{[12](#)}. Therefore, the range in textures of the Chang'e-5 basalt clasts is probably due to different cooling rates^{[13](#)} within different parts of the lava flow. The appreciable variation of mineralogical abundances and inferred whole-rock compositions are probably due to the small sample sizes (smaller than 3 mm) of the basalt clasts.

For comparison with previous results from the Apollo and Luna samples and remote sensing, the bulk-rock major and trace-element compositions of the Chang'e-5 basalts are estimated (Extended Data Tables [5, 6](#)). The Chang'e-5 basalt clasts can be classified as low-Ti/high-Al/low-potassium (K) type^{[14](#)}. However, these basalts have higher iron (FeO ; 22.2 wt%), titanium (TiO_2 ; 5.7 wt%) and aluminium (Al_2O_3 ; 11.6 wt%) contents and a lower Mg# (32.1) relative to the Apollo and Luna low-Ti basalts (Fig. [2b](#), Extended Data Fig. [5](#)). In addition, the Chang'e-5 clasts are highly LREE enriched (50 times heavy REE (HREE) and 150 times LREE enrichment relative to carbonaceous chondrites; Fig. [4a](#)), and show high thorium (Th) contents (approximately 4.5 ppm; Extended Data Table [6](#)). The REE patterns are different from those of the Apollo low-Ti basalts, but parallel to those of the KREEP basalts (Fig. [4a](#)). In addition, the FeO , TiO_2 and Th contents of the Chang'e-5 basalt agree well with the data obtained by the Lunar Prospector Gamma-Ray Spectrometer^{[15,16,17](#)} (FeO about 22.4 wt%; TiO_2 about 4.5 wt%; Th about 5.8 ppm). This agreement suggests that

the Chang'e-5 basalt clasts are representative of the mare basalt unit of the landing site (designated as unit P58; ref. [18](#)).

Fig. 4: Formation of the REE distribution patterns in Chang'e-5 basalts.

 [figure4](#)

a, Comparison of REE distribution patterns of Chang'e-5 and Apollo basalts. Apollo (A) and Luna (L) data are from Clive Neal's Mare Basalt Database (<https://www3.nd.edu/~cneal/Lunar-L/>). The Apollo 14 groups A, B and C were defined by ref. ⁴². **b**, REE modelling of partial melting and fractional crystallization. The blue areas denote the melts produced after 2–3% partial melting of the mantle source (86 PCS + 2% TIRL²²; PCS, per cent crystallized solid; TIRL, trapped instantaneous residual liquid) and those followed by 43–78% fractional crystallization. This source composition is calculated based on the source region $^{87}\text{Rb}/^{86}\text{Sr}$ and $^{147}\text{Sm}/^{144}\text{Nd}$ ratios (Fig. 3). Mineral modes in the source are assumed to be 48% olivine, 23% orthopyroxene, 23% pigeonite, 3% augite and 3% plagioclase. Mineral assemblages of 43–78% crystallization are: 5–10% olivine, 25–59% augite, 2–3% pigeonite and 6–11% plagioclase. Normalization values are from ref. ⁴³. The Chang'e-5 parental melt estimated by the clinopyroxene core with the highest Mg# (sample 406-004, 005) (Extended Data Table 6) requires extensive (43–78%) fractional crystallization after low-degree (2–3%) melting of the mantle source. To match the Chang'e-5 bulk composition, up to 78–88% fractional crystallization is needed. The model parameters are listed in Supplementary Table 3, and details of the batch melting and fractional crystallization model are provided in the Methods.

Source Data

The elevated incompatible trace element (ITE) concentrations and LREE enrichment in the Chang'e-5 basalt clasts are typical characteristics for the KREEP-rich materials. However, the isotopic compositions of the Chang'e-5 basalt are not consistent with the origin of KREEP-rich materials (Fig. 3). Even a small contribution (less than 0.5%; Extended Data Fig. 6) from KREEP materials would result in high $^{87}\text{Rb}/^{86}\text{Sr}$ ratios (greater than 0.19) and low $^{147}\text{Sm}/^{144}\text{Nd}$ ratios (less than 0.173) (Fig. 3), which would shift the Sr–Nd isotopes of the Chang'e-5 basalts considerably. The low initial $^{87}\text{Sr}/^{86}\text{Sr}$ and high $\varepsilon_{\text{Nd}}(t)$ observed in the Chang'e-5 basalts are similar to those of the Apollo 12 low-Ti basalts. This similarity indicates that both the Chang'e-5 basalts and the Apollo 12 low-Ti basalts may originate from a depleted non-KREEP source, which may have crystallized from the early lunar magma ocean cumulates dominated by olivine and pyroxene¹⁹.

The REE and ITE enrichment most likely formed through magmatic processes, such as partial melting and fractional crystallization. Similar characteristics observed in lunar basalt meteorites (for example, NWA 032, NWA 4734 and LAP 02205) were considered to originate from low-degree partial melting of a depleted source^{20,21}. We suggest that a cumulate composition with a crystallized percent solid at 86% and the addition of 2% trapped instantaneous residual liquid in the model of ref. ²² can produce a source with $^{147}\text{Sm}/^{144}\text{Nd} \approx 0.222\text{--}0.227$ and $^{87}\text{Rb}/^{86}\text{Sr} \approx 0.009\text{--}0.022$ (Fig.

[2](#)). Partial melting of such a mantle source alone cannot simultaneously reproduce the LREE and HREE contents even with an unrealistic, low degree of melting (less than 0.3%) (Extended Data Fig. [7](#)). Thus, fractional crystallization must have occurred to elevate the ITE and LREE abundances before the eruption of the Chang'e-5 basalts. Even the parental melt estimated by the clinopyroxene core with the highest Mg# (sample 406-004, 005) (Extended Data Table [6](#)) still requires extensive (43–78%) fractional crystallization after low-degree (2–3%) melting of the mantle source (Fig. [4](#)). This scenario is consistent with the low-Mg# (32.1), high-FeO (22.2 wt%) and high-TiO₂ (5.7 wt%) signatures (Table [1](#)), and compositional zonings in olivine and pyroxene of the samples (Extended Data Fig. [2](#)). Therefore, the Chang'e-5 basalts are highly evolved magmatic products, which implies the presence of a huge magma chamber beneath the PKT at that time.

Table 1 Summary of age and chemical compositions of the Apollo and Chang'e-5 basalts

The youngest mare basalt units in the PKT (for example, P56, P58, P59 and P60; ref. [18](#)) all show moderate-TiO₂ (about 3.8–5.7 wt%) and high-Th (4.9–7.3 ppm) contents [15,23](#). The high-Th materials excavated by impact craters indicate a KREEP layer beneath the PKT [24](#). Melting for the prolonged volcanism in the PKT was thought to be driven by the high abundances of radiogenic heat-producing elements from the KREEP-rich materials [5,6,7](#). However, we find that the elevated Th and other incompatible elements of the Chang'e-5 basalts from the mare unit P58 were not involved with KREEP-rich materials, but rather attributable to the highly evolved basalt produced by low-degree partial melting and extensive fractional crystallization. We deem it unlikely that the KREEP-rich materials provided the heat for partial melting without being involved in the melt itself [7](#). Therefore, the presence of a speculative thick KREEP layer at the base of the crust to generate the young volcanism in the PKT region is not required.

The highly evolved origin of the 2 billion-year-old (Ga) Chang'e-5 basalts implies that the lunar interior was substantially cooler at that time than at around 3.5 Ga when the variety of more primitive basalts sampled by Apollo were formed. In spite of this considerable secular cooling, there must also have been some mechanisms to keep the melt zones in the lunar mantle from solidifying until after 2 Ga (ref. [2](#)). One possible mechanism is that a thick insulating outer layer of the Moon called megaregolith [2,25](#) served as a thermal lid, resulting in a sufficiently slow cooling rate. In addition, lunar cooling can thicken the lithosphere and thus inhibit surface eruption [26](#). Therefore, the thinnest crust of the PKT region (typically less than 30 km; ref. [27](#)) could be a critical factor that facilitated the eruption of young basalts [26](#). Lastly, evidence for the lunar magnetic field persisting until sometime after around 1.92 Ga (ref. [28](#)) is consistent

with sources of lunar interior heat flow such as the latent heat of crystallization of the solid inner core²⁹ being sustained until the age of the Chang'e-5 basalts. Any new model for the thermal evolution of the Moon needs to fit the observation of a non-KREEP origin for the youngest basalts of the PKT.

Methods

Sample preparation

The Chang'e-5 samples studied in this work are two one-inch epoxy mounts (CE5C0000YJYX041GP, 2 mg; CE5C0000YJYX042GP, 44.6 mg) and two soil samples (CE5C0100YJFM00103, about 1,000 mg; CE5C0400YJFM00406, about 2,000 mg) allocated by the China National Space Administration. All these samples were scooped from the lunar surface. The two one-inch epoxy mounts, each containing two basalt clasts, were already polished before allocation. For soil samples, clasts were picked up by a sieve (aperture 600 µm) and hand-picked. Then, the clasts were embedded in epoxy mounts and polished using the grinder.

Scanning electron microscopy analysis and energy dispersive spectrometry mapping

The petrography was carried out on a Zeiss Gemini 450 field emission scanning electron microscope at the Institute of Geology and Geophysics, Chinese Academy of Sciences (IGGCAS). The acceleration voltage was 15.0 KV and the probe current was 2.0 nA. In addition, a Thermo Scientific Apreo scanning electron microscope equipped with an energy dispersive spectrometer was used to identify the phosphorus/zirconium-bearing minerals and to calculate the modal abundance of each mineral based on the elemental mapping.

Electron microprobe analysis of minerals

The major element concentrations of pyroxene, plagioclase, olivine, ilmenite, spinel, quartz, sulfide and phosphates in each sample were analysed by a JEOL JXA8100 electron probe at the IGGCAS. The conditions of the electron microprobe analysis were: acceleration voltage of 15 kV, probe current of 20 nA, focused beam and peak counting time of 10 s. The calibration of the elemental data was done using a series of natural minerals and synthetic materials. On the basis of the analysis of internal laboratory standards, the precision for major (more than 1.0 wt%) and minor (less than 1.0 wt%) elements are better than 1.5% and 5.0%, respectively.

In situ trace-element analysis

The trace-element abundances of pyroxene and plagioclase in basalt fragments were determined by laser ablation–inductively coupled plasma–mass spectrometry (LA–ICP–MS) employing an Element XR HR–ICP–MS instrument coupled to a 193-nm argon-fluoride excimer laser system (Geolas HD) at the IGGCAS. The approach is similar to that outlined in ref. ⁴⁴ with isotopes measured using a peak-hopping mode. The laser diameter is about 32 µm with the repetition rate of 3 Hz. The laser energy density is approximately 3.0 J cm⁻². The Element XR is equipped with a high-capacity interface pump (OnTool Booster 150) in combination with Jet sample and normal H-skimmer cones to achieve a detection efficiency in the range of 1.5% (based on U in a single spot ablation of NIST SRM 612). Helium was employed as the ablation gas to improve the transporting efficiency of ablated aerosols. NIST SRM 610 (ref. ⁴⁵) reference glass was used for external calibration. ARM-1 (ref. ⁴⁶), BCR-2G (ref. ⁴⁷) and BIR-1G (ref. ⁴⁷) glasses were used for quality control monitoring. The bulk normalization as 100 wt% strategy was used for data reduction, which is accomplished using the Iolite software package with an in-house-built data reduction scheme code⁴⁸. For most trace elements (more than 0.05 ppm), the accuracy is better than ±15% with analytical precision (1 relative standard deviation) of ±10%. The spots analysed for trace elements are shown in Supplementary Figs. ^{1–3}.

In situ Sr–Nd isotopic analysis

In situ Sr–Nd isotopic measurements by LA–multiple collector (MC)–ICP–MS followed the method of refs. ^{49,50,51}, hence only a brief description is given below. All analyses were conducted at the IGGCAS.

A Neptune Plus MC–ICP–MS coupled to an Analyte G2 193-nm argon-fluoride excimer laser ablation system was used to determine the Sr isotopic ratios of the plagioclase. A spot size of 85 µm was employed with a repetition rate of 6 Hz and an energy density of approximately 7 J cm⁻². The Sr isotopic data were acquired by static multi-collection in low-resolution mode using nine Faraday collectors. Before laser analysis, the Neptune Plus MC–ICP–MS was tuned using NBS 987 standard solution to obtain maximum sensitivity. A typical data acquisition cycle consisted of a 30 s of measurement of the krypton gas blank with the laser switched off, followed by 60 s of measurement with the laser ablating. Data reduction was conducted offline and the potential isobaric interferences were accounted for in the following order: Kr⁺, Yb²⁺, Er²⁺ and Rb⁺. Finally, the ⁸⁷Sr/⁸⁶Sr ratios were calculated and normalized from the interference-corrected ⁸⁶Sr/⁸⁸Sr ratio using an exponential law. The whole data-reduction procedure was performed using an in-house Excel VBA (Visual Basic for Applications) macro program. The JH56 in-house plagioclase reference material was measured before and after unknown samples for external calibration⁴⁹, which are shown in Extended Data Table ³.

The same LA–MC–ICP–MS system as for the in situ Sr isotopic analysis was used to determine Nd isotopic compositions of the merrillite. Before laser analyses, the Neptune Plus MC–ICP–MS was tuned and optimized for maximum sensitivity using JNd-1 standard solution. A laser spot size of 20 µm was employed with a repetition rate of 3 Hz and an energy density of approximately 6 J cm⁻². Each spot analysis consisted of approximately 60 s of data acquisition with the laser fire on. The SDG in-house apatite reference material was measured before and after the merrillite samples.

To obtain accurate ¹⁴⁷Sm/¹⁴⁴Nd and ¹⁴³Nd/¹⁴⁴Nd data by LA–MC–ICP–MS, care must be taken to adequately correct for the contribution of the isobaric interference of ¹⁴⁴Sm on the ¹⁴⁴Nd signal. The (samarium) Sm interference correction is complicated by the fact that the ¹⁴⁶Nd/¹⁴⁴Nd ratio, which is conventionally used to normalize the other Nd isotope ratios, is also affected by Sm interference. As a result, the mass bias correction of ¹⁴⁴Sm interference on ¹⁴⁴Nd cannot be applied directly from the measured ¹⁴⁶Nd/¹⁴⁴Nd ratio. In this work, we adopted the recently revised Sm isotopic abundances (¹⁴⁷Sm/¹⁴⁹Sm = 1.08680 and ¹⁴⁴Sm/¹⁴⁹Sm = 0.22332)⁵⁰. First, we used the measured ¹⁴⁷Sm/¹⁴⁹Sm ratio to calculate the mass bias factor of Sm and the measured ¹⁴⁷Sm intensity by employing the natural ¹⁴⁷Sm/¹⁴⁴Sm ratio of 4.866559 to estimate the Sm interference on mass 144. Then the interference-corrected ¹⁴⁶Nd/¹⁴⁴Nd ratio can be used to calculate the mass bias factor of Nd. Finally, the ¹⁴³Nd/¹⁴⁴Nd and ¹⁴⁵Nd/¹⁴⁴Nd ratios were normalized using the exponential law. The ¹⁴⁷Sm/¹⁴⁴Nd ratio of unknown samples can also be calculated using the exponential law after correcting for isobaric interference of ¹⁴⁴Sm on ¹⁴⁴Nd as described above. The ¹⁴⁷Sm/¹⁴⁴Nd ratio was then externally further calibrated against the ¹⁴⁷Sm/¹⁴⁴Nd ratio of the SDG reference material during the analytical sessions⁵⁰. The raw data were exported offline and the whole data-reduction procedure was performed using an in-house Excel VBA macro program. The LREE glass analysed in this session gave a mean ¹⁴³Nd/¹⁴⁴Nd ratio of 0.512100 ± 0.000048 (2 s.d., $n = 8$), which is consistent with the recommended value⁵². The data are shown in Extended Data Table 4. The spots analysed for Sr–Nd isotopes are shown in Supplementary Figs. 1–3.

Petrography and mineral chemistry

The lithic clasts from the two lunar soils comprise about 45% basalts, about 35% impact melt breccias and about 20% agglutinates. The basalt clasts can be texturally subdivided into four types: poikilitic, subophitic, porphyritic and equigranular. Poikilitic (about 40%) and subophitic (about 40%) clasts dominate and porphyritic (about 10%) and equigranular (about 10%) clasts are minor.

The poikilitic clasts are mainly composed of clinopyroxene, plagioclase, olivine and accessory Cr-Ti-spinel, ilmenite, troilite and mesostasis including K-feldspar, fayalite,

silica and phosphates. They show various-grain-size clinopyroxene and olivine included in coarse-grained (greater than 200 µm) plagioclase (Extended Data Fig. 1a). Plagioclase is anorthite rich ($An_{76.5-89.0}$). Clinopyroxene shows a large compositional range, with Mg-rich cores (wollastonite ($Wo_{23.8-39.4}$) enstatite ($En_{30.2-45.8}$) and Fe-rich rims ($Wo_{15.7-42.8}En_{2.2-28.9}$; Supplementary Table 1). Olivine occurs as anhedral inclusions (forsterite ($Fo_{29.7-60.1}$) in plagioclase or as mesostasis phase (fayalite, $Fo_{1.5-9.5}$). Euhedral spinel has about 10.3–19.2 wt% chromium(III) oxide (Cr_2O_3), about 49.3–56.5 wt% FeO and about 21.6–28.4 wt% TiO_2 , and can be occasionally observed as inclusions in clinopyroxene and plagioclase (Extended Data Fig. 1a), pointing to an early crystallization phase.

The subophitic clasts show various grain sizes (30–300 µm) and consist mainly of plagioclase, clinopyroxene, olivine and ilmenite, with minor troilite and cristobalite (Extended Data Fig. 1b). Both clinopyroxene and olivine have compositional zoning, with Mg-rich cores and iron-rich rims (Extended Data Fig. 2, Supplementary Table 1). Plagioclase has a euhedral-to-subhedral shape with an anorthite-rich composition ($An_{75.7-90.3}$). Small amounts of Fe-rich olivine ($Fo < 10$) associated with cristobalite and apatite occur as mesostasis phases representing the late-stage crystallization products.

The porphyritic clasts commonly show coarse-grained (100–300 µm) mafic phenocrysts in a fine-grained (smaller than 60 µm) matrix. The mafic phenocrysts include subhedral clinopyroxene, euhedral-to-subhedral olivine and euhedral Cr-spinel (Extended Data Fig. 1c). The clinopyroxene phenocrysts are zoned from Mg-rich cores to iron-rich rims (Supplementary Table 1). The olivine phenocrysts also show compositional zoning ($Fo_{41.4-59.4}$). The Cr-spinel phenocrysts are also compositionally homogeneous, with about 24.5 wt% Cr_2O_3 , about 2.2–2.9 wt% MgO and about 17.3–17.8 wt% TiO_2 contents. The matrix is composed of acicular plagioclase ($An_{76.3-85.2}$), interstitial clinopyroxene and tiny (smaller than 10 µm) Cr-spinel (Extended Data Fig. 1c). Compared with the Cr-spinel phenocrysts, the matrix ones have higher- TiO_2 (about 21.2–28.0 wt%) but lower- Cr_2O_3 contents (8.8–15.7 wt%). Ilmenite needles commonly show three directions cutting the matrix plagioclase and pyroxene, representing a late-stage crystallization phase.

The equigranular clasts are rare and show similar grain sizes (mostly about 100–200 µm) for clinopyroxene and plagioclase. Similar to the poikilitic and subophitic clasts, this type of clast contains clinopyroxene, olivine, plagioclase and ilmenite, with minor troilite and cristobalite (Extended Data Fig. 1d). Clinopyroxene shows significant compositional zoning ($Wo_{13.7-41.1}En_{1.1-35.9}$). The coarse-grained (larger than 100 µm) olivine shows a limited compositional range ($Fo_{35.1-43.2}$); however, the interstitial

olivine grains associated with silica and apatite are systematically fayalite ($\text{Fo} < 5$). Plagioclase has a limited compositional range ($\text{An}_{75.5-83.1}$; Supplementary Table 1).

Except for the above four types of clast, very small amounts of coarse-grained (larger than $100 \mu\text{m}$) fragments of basalt clasts can be found. However, their texture cannot be identified well owing to the limited (about two to five) grains in a single clast. These fragments typically contain coarse-grained clinopyroxene, plagioclase and olivine, with minor fine-grained (mostly smaller than $30 \mu\text{m}$) silica and phosphate minerals.

Estimation of bulk composition

It is assumed that the analysed area proportions are equal to volume proportions and the volume proportions are then converted to mass proportions based the mineral densities reported in previous studies^{53,54,55}. The bulk composition is then calculated by their mass proportions. The oxide concentrations for the bulk composition are normalized to 100%. The average compositions of all samples are calculated based on the contribution for each sample multiplied by its weight, assuming that the weight of each sample is in proportion to its surface area (Supplementary Table 1, Extended Data Table 5). Three clasts (406-002, 002; 406-002, 007; 406-005, 010) have mineral abundances that deviate from other clasts. Clast 406-002, 002 has an extremely high abundance of plagioclase (72.8%). Clast 406-002, 007 has a very high abundances of fayalitic olivine (22.6%) and silica (7.8%). Clast 406-002, 002 has an extremely high abundance of ilmenite (19.1%). These outlier clasts are excluded for the estimation of the bulk composition of the Chang'e-5 basalt (Extended Data Table 5).

The bulk trace elements of the Chang'e-5 basalt were estimated based on the average of all measured pyroxene grains and the partition coefficients (Extended Data Table 6). This estimation assumes that the Chang'e-5 basalt crystallized in a closed system and can be represented by the equilibrium melt of pyroxene. This method may yield large uncertainties owing to the significant variations of the trace-element contents of pyroxene. Therefore, the parental melt estimated by the clinopyroxene core with the highest Mg# (sample 406-004, 005; Extended Data Table 6) is used for the REE modelling of partial melting and fractional crystallization (Fig. 4).

Batch melting model

In this work, we use the batch melting model to calculate the REE concentrations in the parental melts, assuming that each mineral phase melts in proportion to its modal abundance in the source. The batch melting is calculated using the following equation: $C_L/C_0 = 1/[D_0 + F(1 - D_0)]$, where C_L is the weight concentration of a trace element in the melt, C_0 is the weight concentration of a trace element in the original cumulate

source, F is the weight fraction of melt produced and D_0 is the bulk distribution coefficient of the original solid material.

The bulk distribution coefficient is calculated by multiplying each mineral partition coefficient by the fraction of that mineral in the source. The REE partition coefficients for olivine⁵⁶, orthopyroxene⁵⁷, augite⁵⁷, pigeonite⁵⁸ and plagioclase⁵⁹, and the Th partition coefficient for augite⁶⁰ are shown in Supplementary Table 3. Given that the Chang'e-5 basalts have a similar source to the Apollo 12 basalts, the modal mineralogical assemblage calculated for Apollo 12¹⁹ are also adopted here but with a little modification owing to the geochemical differences between Chang'e-5 and Apollo 12 mare basalts. The source materials are compiled in Extended Data Table 6.

Using the bulk distribution coefficients (D_0) and solid cumulate (C_0), the weight concentration of REE in the melt (C_L) is calculated for increasing melt fractions (F).

Fractional crystallization model

The trace-element concentrations in the remaining melt induced by fractional crystallization are calculated using the Rayleigh fractionation equation: $C_L/C_0 = (1 - F)^{D-1}$, where D is bulk distribution coefficient (the same as described in batch melting model), F is the mass fraction of crystals crystallized from the melt, C_0 is the concentration of an element in the initial melt and C_L is the concentration in the final melt. The initial melts during the calculation are assumed to be derived from 2% and 3% batch melting of the mantle source, and the results are shown in Fig. 4.

Data availability

All data generated or analysed during this study are available in EarthChem Library at <https://doi.org/10.26022/IEDA/112076>. Source data are provided with this paper.

References

1. 1.
Shearer, C. K. et al. Thermal and magmatic evolution of the Moon. *Rev. Mineral. Geochem.* **60**, 365–518 (2006).
2. 2.

Ziethe, R., Seiferlin, K. & Hiesinger, H. Duration and extent of lunar volcanism: comparison of 3D convection models to mare basalt ages. *Planet. Space Sci.* **57**, 784–796 (2009).

3. 3.

Laneuville, M., Taylor, J. & Wieczorek, M. A. Distribution of radioactive heat sources and thermal history of the Moon. *J. Geophys. Res. Planets* **123**, 3144–3166 (2018).

4. 4.

Jolliff, B. L., Gillis, J. J., Haskin, L. A., Korotev, R. L. & Wieczorek, M. A. Major lunar crustal terranes: surface expressions and crust-mantle origins. *J. Geophys. Res. Planets* **105**, 4197–4216 (2000).

5. 5.

Haskin, L. A., Gillis, J. J., Korotev, R. L. & Jolliff, B. L. The materials of the lunar Procellarum KREEP Terrane: a synthesis of data from geomorphological mapping, remote sensing, and sample analyses. *J. Geophys. Res. Planets* **105**, 20403–20415 (2000).

6. 6.

Wieczorek, M. A. & Phillips, R. J. The “Procellarum KREEP Terrane”: implications for mare volcanism and lunar evolution. *J. Geophys. Res. Planets* **105**, 20417–20430 (2000).

7. 7.

Borg, L. E., Shearer, C. K., Asmerom, Y. & Papike, J. J. Prolonged KREEP magmatism on the Moon indicated by the youngest dated lunar igneous rock. *Nature* **432**, 209–211 (2004).

8. 8.

Yang, W. & Lin, Y. New lunar samples returned by Chang'e-5: opportunities for new discoveries and international collaboration. *The Innovation* **2**, 100070 (2021).

9. 9.

Li, Q.-L. et al. Two-billion-year-old volcanism on the Moon from Chang'e-5 basalts. *Nature* <https://doi.org/10.1038/s41586-021-04100-2> (2021).

10. 10.

Qian, Y. et al. China's Chang'e-5 landing site: geology, stratigraphy, and provenance of materials. *Earth Planet. Sci. Lett.* **561**, 116855 (2021).

11. 11.

Bence, A. & Papike, J. Pyroxenes as recorders of lunar basalt petrogenesis: chemical trends due to crystal-liquid interaction. *Proc. Lunar Planet. Sci.* **3**, 431–469 (1972).

12. 12.

Shearer, C. K., Papike, J. J., Simon, S. B. & Shimizu, N. An ion microprobe study of the intra-crystalline behavior of REE and selected trace elements in pyroxene from mare basalts with different cooling and crystallization histories. *Geochim. Cosmochim. Acta* **53**, 1041–1054 (1989).

13. 13.

Fagan, T. J. et al. Northwest Africa 032: product of lunar volcanism. *Meteorit. Planet. Sci.* **37**, 371–394 (2002).

14. 14.

Neal, C. R. & Taylor, L. A. Petrogenesis of mare basalts: a record of lunar volcanism. *Geochim. Cosmochim. Acta* **56**, 2177–2211 (1992).

15. 15.

Prettyman, T. H. et al. Elemental composition of the lunar surface: analysis of gamma ray spectroscopy data from Lunar Prospector. *J. Geophys. Res. Planets* **111**, E12007 (2006).

16. 16.

Lawrence, D. J. et al. Global spatial deconvolution of Lunar Prospector Th abundances. *Geophys. Res. Lett.* **34**, L03201 (2007).

17. 17.

Sun, L. & Lucey, P. Mineralogy and Mg# of the Chang'e 5 landing region. *Proc. Lunar Planet. Sci.* **52**, 1796 (2021).

18. 18.

Hiesinger, H. et al. in *Recent Advances and Current Research Issues in Lunar Stratigraphy* (eds Ambrose, W. A. & Williams, D. A.) Vol. 477, 1–51 (Geological Society of America, 2011).

19. 19.

Hallis, L. J., Anand, M. & Strekopytov, S. Trace-element modelling of mare basalt parental melts: implications for a heterogeneous lunar mantle. *Geochim. Cosmochim. Acta* **134**, 289–316 (2014).

20. 20.

Borg, L. E. et al. Mechanisms for incompatible-element enrichment on the Moon deduced from the lunar basaltic meteorite Northwest Africa 032. *Geochim. Cosmochim. Acta* **73**, 3963–3980 (2009).

21. 21.

Elardo, S. M. et al. The origin of young mare basalts inferred from lunar meteorites Northwest Africa 4734, 032, and LaPaz Icefield 02205. *Meteorit. Planet. Sci.* **49**, 261–291 (2014).

22. 22.

Snyder, G. A., Taylor, L. A. & Neal, C. R. A chemical model for generating the sources of mare basalts: combined equilibrium and fractional crystallization of the lunar magmasphere. *Geochim. Cosmochim. Acta* **56**, 3809–3823 (1992).

23. 23.

Sato, H. et al. Lunar mare TiO₂ abundances estimated from UV/vis reflectance. *Icarus* **296**, 216–238 (2017).

24. 24.

Zhang, J., Head, J., Liu, J. & Potter, R. Analysis of thorium concentration anomalies on the lunar surface. *Proc. Lunar Planet. Sci.* **52**, 1744 (2021).

25. 25.

Warren, P. H., Haack, H. & Rasmussen, K. L. Megaregolith insulation and the duration of cooling to isotopic closure within differentiated asteroids and the Moon. *J. Geophys. Res. Solid Earth* **96**, 5909-5923 (1991).

26. 26.

Wilson, L. & Head, J. W. Generation, ascent and eruption of magma on the Moon: new insights into source depths, magma supply, intrusions and effusive/explosive eruptions (part 1: theory). *Icarus* **283**, 146-175 (2017).

27. 27.

Wieczorek, M. A. et al. The crust of the Moon as seen by GRAIL. *Science* **339**, 671-675 (2013).

28. 28.

Mighani, S. et al. The end of the lunar dynamo. *Sci. Adv.* **6**, eaax0883 (2020).

29. 29.

Scheinberg, A., Soderlund, K. M. & Schubert, G. Magnetic field generation in the lunar core: the role of inner core growth. *Icarus* **254**, 62-71 (2015).

30. 30.

Snape, J. F. et al. The timing of basaltic volcanism at the Apollo landing sites. *Geochim. Cosmochim. Acta* **266**, 29-53 (2019).

31. 31.

Warren, P. H. & Taylor, G. J. in *Treatise on Geochemistry* (ed. Turekian, K. K.) 213–250 (Elsevier, 2014).

32. 32.

Wieczorek, M. A. The constitution and structure of the lunar interior. *Rev. Mineral. Geochem.* **60**, 221-364 (2006).

33. 33.

Lindsley, D. H. & Andersen, D. J. A two-pyroxene thermometer. *J. Geophys. Res. Solid Earth* **88**, A887-A906 (1983).

34. 34.

Grove, T. L. & Krawczynski, M. J. Lunar mare volcanism: where did the magmas come from? *Elements* **5**, 29-34 (2009).

35. 35.

Snape, J. F., Joy, K. H., Crawford, I. A. & Alexander, L. Basaltic diversity at the Apollo 12 landing site: inferences from petrologic examinations of the soil sample 12003. *Meteorit. Planet. Sci.* **49**, 842-871 (2014).

36. 36.

Anand, M., Taylor, L. A., Misra, K. C., Demidova, S. I. & Nazarov, M. A. KREEPy lunar meteorite Dhofar 287A: a new lunar mare basalt. *Meteorit. Planet. Sci.* **38**, 485-499 (2003).

37. 37.

Nyquist, L. et al. Rb–Sr systematics for chemically defined Apollo 15 and 16 materials. *Proc. Lunar Planet. Sci.* **4**, 1823 (1973).

38. 38.

Nyquist, L. Lunar Rb–Sr chronology. *Phys. Chem. Earth* **10**, 103-142 (1977).

39. 39.

Neal, C. & Taylor, L. Neal, C. & Taylor, L. Modeling of lunar basalt petrogenesis-SR isotope evidence from Apollo 14 high-alumina basalts. *Proc. Lunar Planet. Sci.* **20**, 101-108 (1990).

40. 40.

Nyquist, L. E. & Shih, C. Y. The isotopic record of lunar volcanism. *Geochim. Cosmochim. Acta* **56**, 2213-2234 (1992).

41. 41.

Nyquist, L. E. et al. ^{146}Sm - ^{142}Nd formation interval for the lunar mantle. *Geochim. Cosmochim. Acta* **59**, 2817-2837 (1995).

42. 42.

Neal, C. R. & Kramer, G. Y. The petrogenesis of the Apollo 14 high-Al mare basalts. *Am. Mineral.* **91**, 1521–1535 (2006).

43. 43.

Anders, E. & Grevesse, N. Abundances of the elements: meteoritic and solar. *Geochim. Cosmochim. Acta* **53**, 197–214 (1989).

44. 44.

Wu, S., Karius, V., Schmidt, B. C., Simon, K. & Wörner, G. Comparison of ultrafine powder pellet and flux-free fusion glass for bulk analysis of granitoids by laser ablation-inductively coupled plasma-mass spectrometry. *Geostand. Geoanal. Res.* **42**, 575–591 (2018).

45. 45.

Jochum, K. P. et al. Determination of reference values for NIST SRM 610–617 glasses following ISO guidelines. *Geostand. Geoanal. Res.* **35**, 397–429 (2011).

46. 46.

Wu, S. et al. The preparation and preliminary characterisation of three synthetic andesite reference glass materials (ARM-1, ARM-2, ARM-3) for in situ microanalysis. *Geostand. Geoanal. Res.* **43**, 567–584 (2019).

47. 47.

Jochum, K. P., Willbold, M., Raczek, I., Stoll, B. & Herwig, K. Chemical characterisation of the USGS reference glasses GSA-1G, GSC-1G, GSD-1G, GSE-1G, BCR-2G, BHVO-2G and BIR-1G using EPMA, ID-TIMS, ID-ICP-MS and LA-ICP-MS. *Geostand. Geoanal. Res.* **29**, 285–sss302 (2005).

48. 48.

Wu, S.-T., Huang, C., Xie, L.-W., Yang, Y.-H. & Yang, J.-H. Iolite based bulk normalization as 100% (*m/m*) quantification strategy for reduction of laser ablation-inductively coupled plasma-mass spectrometry transient signal. *Chin. J. Anal. Chem.* **46**, 1628–1636 (2018).

49. 49.

Yang, Y.-H., Wu, F.-Y., Xie, L.-W., Yang, J.-H. & Zhang, Y.-B. In-situ Sr isotopic measurement of natural geological samples by LA-MC-ICP-MS. *Acta Petrol.*

Sin. **25**, 3431–3441 (2009).

50. 50.

Yang, Y.-H. et al. Sr and Nd isotopic compositions of apatite reference materials used in U–Th–Pb geochronology. *Chem. Geol.* **385**, 35–55 (2014).

51. 51.

Yang, Y.-H. et al. In situ U–Th–Pb dating and Sr–Nd isotope analysis of bastnäsite by LA–(MC)–ICP–MS. *Geostand. Geoanal. Res.* **43**, 543–565 (2019).

52. 52.

Fisher, C. M. et al. Sm–Nd isotope systematics by laser ablation-multicollector-inductively coupled plasma mass spectrometry: methods and potential natural and synthetic reference materials. *Chem. Geol.* **284**, 1–20 (2011).

53. 53.

Robie, R. A. & Hemingway, B. S. *Thermodynamic Properties of Minerals and Related Substances at 298.15 K and 1 bar (105 pascals) Pressure and at Higher Temperatures* No. 2131 (US Government Printing Office, 1995).

54. 54.

Smyth, J. R. & McCormick, T. C. in *Mineral Physics and Crystallography: A Handbook of Physical Constants* (ed. Ahrens, T. J.) Vol. 2, 1–17 (American Geophysical Union, 1995).

55. 55.

Treiman, A. H., Maloy, A. K., Shearer, C. K. & Gross, J. Magnesian anorthositic granulites in lunar meteorites Allan Hills A81005 and Dhofar 309: geochemistry and global significance. *Meteorit. Planet. Sci.* **45**, 163–180 (2010).

56. 56.

McKay, G. A. Crystal/liquid partitioning of REE in basaltic systems: extreme fractionation of REE in olivine. *Geochim. Cosmochim. Acta* **50**, 69–79 (1986).

57. 57.

Yao, L., Sun, C. & Liang, Y. A parameterized model for REE distribution between low-Ca pyroxene and basaltic melts with applications to REE partitioning in low-Ca pyroxene along a mantle adiabat and during pyroxenite-derived melt and peridotite interaction. *Contrib. Mineral. Petrol.* **164**, 261–280 (2012).

58. 58.

McKay, G., Le, L. & Wagstaff, J. Constraints on the origin of the mare basalt europium anomaly: REE partition coefficients for pigeonite. *Proc. Lunar Planet. Sci.* **22**, 883–884 (1991).

59. 59.

Phinney, W. C. & Morrison, D. A. Partition coefficients for calcic plagioclase: implications for Archean anorthosites. *Geochim. Cosmochim. Acta* **54**, 1639–1654 (1990).

60. 60.

Hauri, E. H., Wagner, T. P. & Grove, T. L. Experimental and natural partitioning of Th, U, Pb and other trace elements between garnet, clinopyroxene and basaltic melts. *Chem. Geol.* **117**, 149–166 (1994).

Acknowledgements

The Chang'e-5 lunar samples were provided by the China National Space Administration. We thank R. Mitchell, H. Hui, Y. Li, X.-H. Li, Y. Lin and J. Delano for constructive comments; C. Sun for logistical support; and X. Yan and M. Liang for help with the figures. The photo in Fig. 1 was taken by Hui Ren. This study was funded by the Key Research Program of the Chinese Academy of Sciences (ZDBS-SSW-JSC007-15), the Strategic Priority Research Program of Chinese Academy of Sciences (XDB 41000000), the CAS Interdisciplinary Innovation Team, the pre-research project on Civil Aerospace Technologies of China National Space Administration (grant number D020203), and the key research programme of the Institute of Geology and Geophysics, Chinese Academy of Sciences (IGGCAS-202101). Special thanks goes to the Youth Innovation Promotion Association of Chinese Academy of Sciences for long-term support.

Author information

Author notes

1. These authors contributed equally: Heng-Ci Tian, Hao Wang, Yi Chen

Affiliations

1. Key Laboratory of Earth and Planetary Physics, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing, China

Heng-Ci Tian, Wei Yang, Chi Zhang, Hong-Lei Lin & Rui Chang

2. State Key Laboratory of Lithospheric Evolution, Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing, China

Hao Wang, Yi Chen, Chao Huang, Shi-Tou Wu, Li-Hui Jia, Lei Xu, Di Zhang, Xiao-Guang Li, Yue-Heng Yang, Lie-Wen Xie, Dan-Ping Zhang & Fu-Yuan Wu

3. National Astronomical Observatories, Chinese Academy of Sciences, Beijing, China

Qin Zhou, Guang-Liang Zhang & Sai-Hong Yang

Contributions

W.Y. led the study; W.Y., Q.Z., H.-C.T., C.Z., D.-P.Z., G.-L.Z. and S.-H.Y. prepared the samples; H.-C.T., L.-H.J., D.Z., X.-G.L., S.-T.W., Y.C. and H.W. performed the major and trace-element analyses; C.H., L.X., Y.-H.Y., L.-W.X. and H.W. performed the Sr–Nd isotope analyses; H.-L.L. and R.C. contributed to data sorting and compilation; W.Y., H.-C.T., H.W., Y.C. and F.-Y.W. wrote the manuscript.

Corresponding author

Correspondence to [Wei Yang](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Richard Carlson and Charles Shearer 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 Representative four types of the Chang'e-5 basalt clasts.](#)

a, Poikilitic clast. **b**, Subophitic clast. **c**, Porphyritic clast. **d**, Equigranular clast. Cpx, clinopyroxene; Pl, plagioclase; Ol, olivine; Ilm, ilmenite; Spl, spinel; Tro, troilite; Crs, cristobalite. For detailed sample description, see [Methods](#).

[Extended Data Fig. 2 Compositional zoning profiles of olivine and clinopyroxene in the Chang'e-5 basalts.](#)

a, b, One poikilitic clast and the variation of Fo content across an olivine grain. **c, d**, One subophitic clast and the variations of En and Fs contents across a clinopyroxene grain. Note that both mafic minerals have Mg-rich cores and Fe-rich rims.

[Source Data](#)

[Extended Data Fig. 3 Compositions of feldspar and pyroxene from the Chang'e-5 basalts.](#)

a, Ternary diagram of feldspar in the Chang'e-5 basalt clasts. Different types of basalt clast have similar plagioclase composition. K-feldspar exhibits a large compositional range. **b**, Ti versus Al diagram of the Chang'e-5 pyroxene grains. Most of the data fall along the 1:2 line.

[Source Data](#)

[Extended Data Fig. 4 REE distribution of pyroxene and plagioclase from the Chang'e-5 basalts.](#)

a, Pyroxene. **b**, Plagioclase. The normalized data are from ref. [43](#).

[Source Data](#)

[Extended Data Fig. 5 Compositional diagrams of estimated bulk compositions of the Chang'e-5 basalts.](#)

The symbols of four-type clasts follow those in Extended Data Fig. 3. The red hexagonal symbols refer to the average compositions of the Chang'e-5 basalts (Extended Data Table 5). The Apollo and Luna basalts are shown for comparison (data from Clive Neal's Mare Basalt Database; <https://www3.nd.edu/~cneal/Lunar-L/>).

[Source Data](#)

[Extended Data Fig. 6 Calculated \$^{147}\text{Sm}/^{144}\text{Nd}\$ versus \$^{87}\text{Rb}/^{86}\text{Sr}\$ values for the source regions of the Chang'e-5 basalts.](#)

The two grey areas, KREEP endmember, black and dashed lines (representing binary mixing between the lunar mantle and the KREEP component) are adopted from ref. 20. The crosshair symbols represent 0.01, 0.05, 0.1, 0.2, 0.5 and 2% addition of urKREEP component. The Chang'e-5 basalt falls on the plagioclase-bearing cumulate source.

[Source Data](#)

[Extended Data Fig. 7 Chondrite-normalized REE patterns and different batch melting degrees for the calculated mantle source.](#)

The equilibrium melt of clinopyroxene (core) with the highest MgO is chosen as the parental melt. The normalized data are from ref. 43. PCS, percent crystallized solid; TIRL, trapped instantaneous residual liquid.

[Source Data](#)

Extended Data Table 1 The Chang'e-5 basalt clasts analysed in this study

Extended Data Table 2 Modal mineralogy of the Chang'e-5 basalt clasts

Extended Data Table 3 Rb–Sr isotopic data of plagioclase from the Chang'e-5 basalts

Extended Data Table 4 Sm–Nd isotopic data of merrillite from the Chang'e-5 basalts

Extended Data Table 5 Estimated major element compositions of the Chang'e-5 basalts

Extended Data Table 6 Estimated trace element compositions of the Chang'e-5 basalts

Supplementary information

[Supplementary Figs. 1–3](#)

This file contains Supplementary Figs. 1–3, showing the positions of analytical spots of trace elements and Sr–Nd isotopes for Chang'e-5 basalt clasts.

[Peer Review File](#)

[Supplementary Tables 1–3](#)

This file contains Supplementary Tables 1–3. Supplementary Table 1 contains the major element concentrations of minerals from representative Chang'e-5 basalt clasts. Supplementary Table 2 contains the trace element concentrations of the pyroxene and plagioclase. Supplementary Table 3 contains the REE partition coefficients for minerals used in model calculation.

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. 3](#)

[Source Data Extended Data Fig. 4](#)

[Source Data Extended Data Fig. 5](#)

[Source Data Extended Data Fig. 6](#)

[Source Data Extended Data Fig. 7](#)

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and

reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

Reprints and Permissions

About this article

Cite this article

Tian, HC., Wang, H., Chen, Y. *et al.* Non-KREEP origin for Chang'e-5 basalts in the Procellarum KREEP Terrane. *Nature* **600**, 59–63 (2021).
<https://doi.org/10.1038/s41586-021-04119-5>

- Received: 28 July 2021
- Accepted: 08 October 2021
- Published: 19 October 2021
- Issue Date: 02 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04119-5>

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

Further reading

- **Two-billion-year-old volcanism on the Moon from Chang'e-5 basalts**

- Qiu-Li Li
- Qin Zhou
- Xian-Hua Li

Nature (2021)

- **A dry lunar mantle reservoir for young mare basalts of Chang'e-5**

- Sen Hu
- Huicun He
- Ziyuan Ouyang

Nature (2021)

- **Chang'e-5 samples reveal two-billion-year-old volcanic activity on the Moon and its source characteristics**

- Shoujie Liu
- Qin Zhou
- Wei Yang

Science China Earth Sciences (2021)

Robotic sample return reveals lunar secrets

- Richard W. Carlson

News & Views 01 Dec 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04119-5>

- Article
- [Published: 01 December 2021](#)

Sound emission and annihilations in a programmable quantum vortex collider

- [W. J. Kwon](#) ORCID: [orcid.org/0000-0001-9773-4024^{1,2}](https://orcid.org/0000-0001-9773-4024),
- [G. Del Pace^{1,2}](#),
- [K. Xhani^{1,2}](#),
- [L. Galantucci](#) ORCID: [orcid.org/0000-0002-3435-4259³](https://orcid.org/0000-0002-3435-4259),
- [A. Muzi Falconi^{1,2}](#),
- [M. Inguscio^{1,2,4}](#),
- [F. Scazza](#) ORCID: [orcid.org/0000-0001-5527-1068^{1,2 nAff5}](https://orcid.org/0000-0001-5527-1068) &
- [G. Roati^{1,2}](#)

[Nature](#) volume 600, pages 64–69 (2021)

- 604 Accesses
- 18 Altmetric
- [Metrics details](#)

Subjects

- [Quantum fluids and solids](#)
- [Ultracold gases](#)

Abstract

In quantum fluids, the quantization of circulation forbids the diffusion of a vortex swirling flow seen in classical viscous fluids. Yet, accelerating quantum vortices may lose their energy into acoustic radiations^{1,2}, similar to the way electric charges decelerate on emitting photons. The dissipation of vortex energy underlies central problems in quantum hydrodynamics³, such as the decay of quantum turbulence, highly relevant to systems as varied as neutron stars, superfluid helium and atomic condensates^{4,5}. A deep understanding of the elementary mechanisms behind irreversible vortex dynamics has been a goal for decades^{3,6}, but it is complicated by the shortage of conclusive experimental signatures⁷. Here we address this challenge by realizing a programmable vortex collider in a planar, homogeneous atomic Fermi superfluid with tunable inter-particle interactions. We create on-demand vortex configurations and monitor their evolution, taking advantage of the accessible time and length scales of ultracold Fermi gases^{8,9}. Engineering collisions within and between vortex–antivortex pairs allows us to decouple relaxation of the vortex energy due to sound emission and that due to interactions with normal fluid (that is, mutual friction). We directly visualize how the annihilation of vortex dipoles radiates a sound pulse. Further, our few-vortex experiments extending across different superfluid regimes reveal non-universal dissipative dynamics, suggesting that fermionic quasiparticles localized inside the vortex core contribute significantly to dissipation, thereby opening the route to exploring new pathways for quantum turbulence decay, vortex by vortex.

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: Deterministic generation and frictional dissipation of single vortex dipoles.

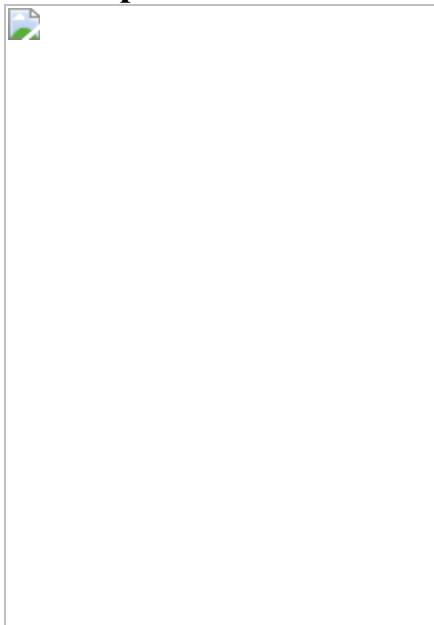


Fig. 2: Arbitrary vortex collisions on demand.

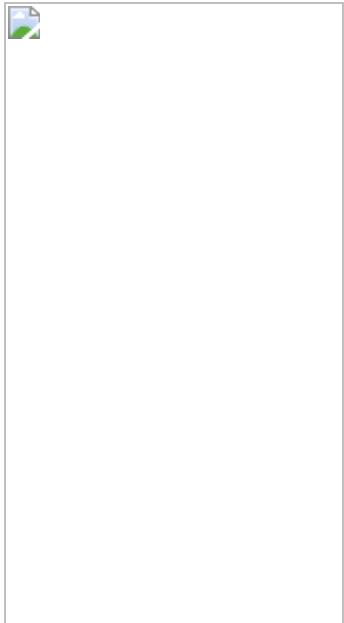


Fig. 3: Dissipation and annihilation in two-dipole collisions.

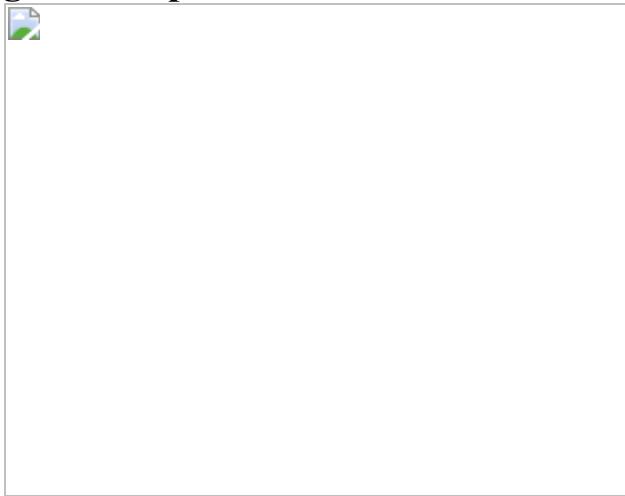
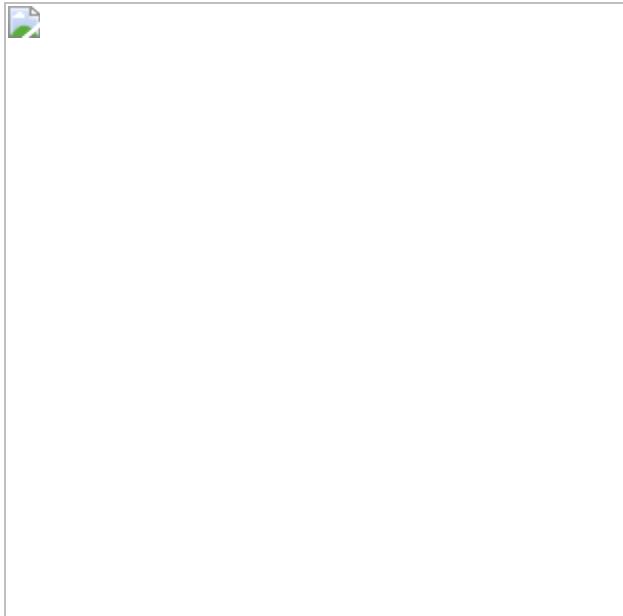


Fig. 4: Direct observation of the sound emission from dipole–dipole annihilation events.



Data availability

The data that support the figures within this paper are available from the corresponding author upon reasonable request.

References

1. 1.

Parker, N. G., Proukakis, N. P., Barenghi, C. F. & Adams, C. S. Controlled vortex-sound interactions in atomic Bose-Einstein condensates. *Phys. Rev. Lett.* **92**, 160403 (2004).

2. 2.

Barenghi, C. F., Parker, N. G., Proukakis, N. P. & Adams, C. S. Decay of quantised vorticity by sound emission. *J. Low Temp. Phys.* **138**, 629–634 (2005).

3. 3.

Tsubota, M., Kobayashi, M. & Takeuchi, H. Quantum hydrodynamics. *Phys. Rep.* **522**, 191–238 (2013).

4. 4.

Vinen, W. F. & Niemela, J. J. Quantum turbulence. *J. Low Temp. Phys.* **128**, 167–231 (2002).

5. 5.

Barenghi, C. F., Skrbek, L. & Sreenivasan, K. R. Introduction to quantum turbulence. *Proc. Natl Acad. Sci. USA* **111**, 4647–4652 (2014).

6. 6.

Feynman, R. P. in *Progress in Low Temperature Physics* Vol. 1, 17–53 (Elsevier, 1955).

7. 7.

Vinen, W. Quantum turbulence: achievements and challenges. *J. Low Temp. Phys.* **161**, 419–444 (2010).

8. 8.

Zwierlein, M. W., Abo-Shaeer, J. R., Schirotzek, A., Schunck, C. H. & Ketterle, W. Vortices and superfluidity in a strongly interacting Fermi gas. *Nature* **435**, 1047–1051 (2005).

9. 9.

Ku, M. J. H., Mukherjee, B., Yefsah, T. & Zwierlein, M. W. Cascade of solitonic excitations in a superfluid Fermi gas: from planar solitons to vortex rings and lines. *Phys. Rev. Lett.* **116**, 045304 (2016).

10. 10.

Bardeen, J. & Stephen, M. J. Theory of the motion of vortices in superconductors. *Phys. Rev.* **140**, A1197–A1207 (1965).

11. 11.

Blatter, G., Feigel'man, M. V., Geshkenbein, V. B., Larkin, A. I. & Vinokur, V. M. Vortices in high-temperature superconductors. *Rev. Mod. Phys.* **66**, 1125–1388 (1994).

12. 12.

Hall, H. E. & Vinen, W. F. The rotation of liquid helium II – II. The theory of mutual friction in uniformly rotating helium II. *Proc. R. Soc. Lond. A* **238**, 215–234 (1956).

13. 13.

Nore, C., Abid, M. & Brachet, M. E. Kolmogorov turbulence in low-temperature superflows. *Phys. Rev. Lett.* **78**, 3896–3899 (1997).

14. 14.

Vinen, W. F. Decay of superfluid turbulence at a very low temperature: the radiation of sound from a Kelvin wave on a quantized vortex. *Phys. Rev. B* **64**, 134520 (2001).

15. 15.

Vinen, W. F., Tsubota, M. & Mitani, A. Kelvin-wave cascade on a vortex in superfluid ^4He at a very low temperature. *Phys. Rev. Lett.* **91**, 135301 (2003).

16. 16.

Leadbeater, M., Winiecki, T., Samuels, D. C., Barenghi, C. F. & Adams, C. S. Sound emission due to superfluid vortex reconnections. *Phys. Rev. Lett.* **86**, 1410–1413 (2001).

17. 17.

Ruostekoski, J. & Dutton, Z. Engineering vortex rings and systems for controlled studies of vortex interactions in Bose-Einstein condensates. *Phys. Rev. A* **72**, 063626 (2005).

18. 18.

Villois, A., Proment, D. & Krstulovic, G. Irreversible dynamics of vortex reconnections in quantum fluids. *Phys. Rev. Lett.* **125**, 164501 (2020).

19. 19.

Kivotides, D., Vassilicos, J. C., Samuels, D. C. & Barenghi, C. F. Kelvin waves cascade in superfluid turbulence. *Phys. Rev. Lett.* **86**, 3080–3083 (2001).

20. 20.

Kozik, E. & Svistunov, B. Kelvin-wave cascade and decay of superfluid turbulence. *Phys. Rev. Lett.* **92**, 035301 (2004).

21. 21.

L'vov, V. S. & Nazarenko, S. Spectrum of Kelvin-wave turbulence in superfluids. *JETP Lett.* **91**, 428–434 (2010).

22. 22.

Popov, V. N. Quantum vortices and phase transitions in Bose systems. *Sov. Phys. JETP* **37**, 341–345 (1973).

23. 23.

Ambegaokar, V., Halperin, B. I., Nelson, D. R. & Siggia, E. D. Dynamics of superfluid films. *Phys. Rev. B* **21**, 1806–1826 (1980).

24. 24.

Arovas, D. P. & Freire, J. Dynamical vortices in superfluid films. *Phys. Rev. B* **55**, 1068–1080 (1997).

25. 25.

Bradley, D. I. et al. Direct measurement of the energy dissipated by quantum turbulence. *Nat. Phys.* **7**, 473–476 (2011).

26. 26.

Fonda, E., Meichle, D. P., Ouellette, N. T., Hormoz, S. & Lathrop, D. P. Direct observation of Kelvin waves excited by quantized vortex reconnection. *Proc. Natl Acad. Sci. USA* **111**, 4707–4710 (2014).

27. 27.

Walmsley, P., Zmeev, D., Pakpour, F. & Golov, A. Dynamics of quantum turbulence of different spectra. *Proc. Natl Acad. Sci. USA* **111**, 4691–4698 (2014).

28. 28.

Caroli, C., De Gennes, P. & Matricon, J. Bound Fermion states on a vortex line in a type II superconductor. *Phys. Lett.* **9**, 307–309 (1964).

29. 29.

Kopnin, N. B. & Salomaa, M. M. Mutual friction in superfluid ^3He : effects of bound states in the vortex core. *Phys. Rev. B* **44**, 9667–9677 (1991).

30. 30.

Nygaard, N., Bruun, G. M., Clark, C. W. & Feder, D. L. Microscopic structure of a vortex line in a dilute superfluid Fermi gas. *Phys. Rev. Lett.* **90**, 210402 (2003).

31. 31.

Sensarma, R., Randeria, M. & Ho, T.-L. Vortices in superfluid Fermi gases through the BEC to BCS crossover. *Phys. Rev. Lett.* **96**, 090403 (2006).

32. 32.

Silaev, M. A. Universal mechanism of dissipation in Fermi superfluids at ultralow temperatures. *Phys. Rev. Lett.* **108**, 045303 (2012).

33. 33.

Kwon, W. J., Moon, G., Choi, J.-Y., Seo, S. W. & Shin, Y.-I. Relaxation of superfluid turbulence in highly oblate Bose-Einstein condensates. *Phys. Rev. A* **90**, 063627 (2014).

34. 34.

Johnstone, S. P. et al. Evolution of large-scale flow from turbulence in a two-dimensional superfluid. *Science* **364**, 1267–1271 (2019).

35. 35.

Liu, X.-P. et al. Universal dynamical scaling of quasi-two-dimensional vortices in a strongly interacting fermionic superfluid. *Phys. Rev. Lett.* **126**, 185302 (2021).

36. 36.

Rooney, S. J., Blakie, P. B., Anderson, B. P. & Bradley, A. S. Suppression of Kelvon-induced decay of quantized vortices in oblate Bose-Einstein condensates. *Phys. Rev. A* **84**, 023637 (2011).

37. 37.

Neely, T. W., Samson, E. C., Bradley, A. S., Davis, M. J. & Anderson, B. P. Observation of vortex dipoles in an oblate Bose-Einstein condensate. *Phys. Rev. Lett.* **104**, 160401 (2010).

38. 38.

Park, J. W., Ko, B. & Shin, Y. Critical vortex shedding in a strongly interacting fermionic superfluid. *Phys. Rev. Lett.* **121**, 225301 (2018).

39. 39.

Samson, E. C., Wilson, K. E., Newman, Z. L. & Anderson, B. P. Deterministic creation, pinning, and manipulation of quantized vortices in a Bose-Einstein condensate. *Phys. Rev. A* **93**, 023603 (2016).

40. 40.

Donnelly, R. J. *Quantized Vortices in Helium II* (Cambridge Univ. Press, 1991).

41. 41.

Parker, N. G. Numerical Studies of Vortices and Dark Solitons in Atomic Bose-Einstein Condensates. PhD thesis, Univ. Durham (2004).

42. 42.

Bulgac, A. & Yu, Y. Vortex state in a strongly coupled dilute atomic fermionic superfluid. *Phys. Rev. Lett.* **91**, 190404 (2003).

43. 43.

Chien, C.-C., He, Y., Chen, Q. & Levin, K. Ground-state description of a single vortex in an atomic Fermi gas: from BCS to Bose-Einstein condensation. *Phys. Rev. A* **73**, 041603 (2006).

44. 44.

Simonucci, S., Pieri, P. & Strinati, G. C. Temperature dependence of a vortex in a superfluid Fermi gas. *Phys. Rev. B* **87**, 214507 (2013).

45. 45.

Jones, C. A. and Roberts, P. H. Motions in a Bose condensate. IV. Axisymmetric solitary waves. *J. Phys. A* **15**, 2599–2619 (1982).

46. 46.

Jackson, B., Proukakis, N. P., Barenghi, C. F. & Zaremba, E. Finite-temperature vortex dynamics in Bose-Einstein condensates. *Phys. Rev. A* **79**, 053615 (2009).

47. 47.

Moon, G., Kwon, W. J., Lee, H. & Shin, Y.-I. Thermal friction on quantum vortices in a Bose-Einstein condensate. *Phys. Rev. A* **92**, 051601 (2015).

48. 48.

Wittmer, P., Schmied, C.-M., Gasenzer, T. & Ewerz, C. Vortex motion quantifies strong dissipation in a holographic superfluid. *Phys. Rev. Lett.* **127**, 101601 (2021).

49. 49.

Aioi, T., Kadokura, T., Kishimoto, T. & Saito, H. Controlled generation and manipulation of vortex dipoles in a Bose-Einstein condensate. *Phys. Rev. X* **1**, 021003 (2011).

50. 50.

Stagg, G. W., Allen, A. J., Parker, N. G. & Barenghi, C. F. Generation and decay of two-dimensional quantum turbulence in a trapped Bose-Einstein condensate. *Phys. Rev. A* **91**, 013612 (2015).

51. 51.

Groszek, A. J., Simula, T. P., Paganin, D. M. & Helmerson, K. Onsager vortex formation in Bose-Einstein condensates in two-dimensional power-law traps. *Phys. Rev. A* **93**, 043614 (2016).

52. 52.

Baggaley, A. W. & Barenghi, C. F. Decay of homogeneous two-dimensional quantum turbulence. *Phys. Rev. A* **97**, 033601 (2018).

53. 53.

Cidrim, A., dos Santos, F. E. A., Galantucci, L., Bagnato, V. S. & Barenghi, C. F. Controlled polarization of two-dimensional quantum turbulence in atomic Bose-Einstein condensates. *Phys. Rev. A* **93**, 033651 (2016).

54. 54.

Karl, M. & Gasenzer, T. Strongly anomalous non-thermal fixed point in a quenched two-dimensional Bose gas. *New J. Phys.* **19**, 093014 (2017).

55. 55.

Yang, G., Zhang, S. & Han, W. Oblique collisions and catching-up phenomena of vortex dipoles in a uniform Bose–Einstein condensate. *Phys. Scr.* **94**, 075006 (2019).

56. 56.

Tylutki, M. & Włazłowski, G. Universal aspects of vortex reconnections across the BCS-BEC crossover. *Phys. Rev. A* **103**, L051302 (2021).

57. 57.

Reeves, M. T., Billam, T. P., Anderson, B. P. & Bradley, A. S. Inverse energy cascade in forced two-dimensional quantum turbulence. *Phys. Rev. Lett.* **110**, 104501 (2013).

58. 58.

White, A. C., Anderson, B. P. & Bagnato, V. S. Vortices and turbulence in trapped atomic condensates. *Proc. Natl Acad. Sci. USA* **111**, 4719–4726 (2014).

59. 59.

Simula, T., Davis, M. J. & Helmerson, K. Emergence of order from turbulence in an isolated planar superfluid. *Phys. Rev. Lett.* **113**, 165302 (2014).

60. 60.

Kwon, W. J., Kim, J. H., Seo, S. W. & Shin, Y. Observation of von Kármán vortex street in an atomic superfluid gas. *Phys. Rev. Lett.* **117**, 245301 (2016).

61. 61.

Seo, S. W., Ko, B., Kim, J. H. & Shin, Y. Observation of vortex-antivortex pairing in decaying 2D turbulence of a superfluid gas. *Sci. Rep.* **7**, 4587 (2017).

62. 62.

Gauthier, G. et al. Giant vortex clusters in a two-dimensional quantum fluid. *Science* **364**, 1264–1267 (2019).

63. 63.

Burchianti, A. et al. Connecting dissipation and phase slips in a Josephson junction between fermionic superfluids. *Phys. Rev. Lett.* **120**, 025302 (2018).

64. 64.

Mukherjee, B. et al. Homogeneous atomic Fermi gases. *Phys. Rev. Lett.* **118**, 123401 (2017).

65. 65.

Hueck, K. et al. Two-dimensional homogeneous Fermi gases. *Phys. Rev. Lett.* **120**, 060402 (2018).

66. 66.

Del Pace, G., Kwon, W. J., Zaccanti, M., Roati, G. & Scazza, F. Tunneling transport of unitary fermions across the superfluid transition. *Phys. Rev. Lett.* **126**, 055301 (2021).

67. 67.

Haussmann, R., Rantner, W., Cerrito, S. & Zwerger, W. Thermodynamics of the BCS-BEC crossover. *Phys. Rev. A* **75**, 023610 (2007).

68. 68.

Pini, M., Pieri, P., Jager, M., Denschlag, J. H. & Strinati, G. C. Pair correlations in the normal phase of an attractive Fermi gas. *New J. Phys.* **22**, 083008 (2020).

69. 69.

Kwon, W. J. et al. Strongly correlated superfluid order parameters from dc Josephson supercurrents. *Science* **369**, 84–88 (2020).

70. 70.

Kwon, W. J., Seo, S. W. & Shin, Y.-I. Periodic shedding of vortex dipoles from a moving penetrable obstacle in a Bose-Einstein condensate. *Phys. Rev. A* **92**, 033613 (2015).

71. 71.

Gertjerenken, B., Kevrekidis, P. G., Carretero-González, R. & Anderson, B. P. Generating and manipulating quantized vortices on-demand in a Bose-Einstein condensate: a numerical study. *Phys. Rev. A* **93**, 023604 (2016).

72. 72.

Ku, M. J. H. et al. Motion of a solitonic vortex in the BEC-BCS crossover. *Phys. Rev. Lett.* **113**, 065301 (2014).

73. 73.

Wilson, K. E., Newman, Z. L., Lowney, J. D. & Anderson, B. P. In situ imaging of vortices in Bose-Einstein condensates. *Phys. Rev. A* **91**, 023621 (2015).

74. 74.

Pethick, C. J. & Smith, H. *Bose-Einstein Condensation in Dilute Gases* (Cambridge Univ. Press, 2002).

75. 75.

Rakonjac, A. et al. Measuring the disorder of vortex lattices in a Bose-Einstein condensate. *Phys. Rev. A* **93**, 013607 (2016).

76. 76.

Iordanskii, S. V. Mutual friction force in a rotating Bose gas. *Sov. Phys. JETP* **22**, 160–167 (1966).

77. 77.

Schwarz, K. W. Three-dimensional vortex dynamics in superfluid ^4He : homogeneous superfluid turbulence. *Phys. Rev. B* **38**, 2398–2417 (1988).

78. 78.

Billam, T. P., Reeves, M. T. & Bradley, A. S. Spectral energy transport in two-dimensional quantum vortex dynamics. *Phys. Rev. A* **91**, 023615 (2015).

79. 79.

Kim, J. H., Kwon, W. J. & Shin, Y. Role of thermal friction in relaxation of turbulent Bose-Einstein condensates. *Phys. Rev. A* **94**, 033612 (2016).

80. 80.

Skaugen, A. & Angheluta, L. Origin of the inverse energy cascade in two-dimensional quantum turbulence. *Phys. Rev. E* **95**, 052144 (2017).

81. 81.

Stockdale, O. R. et al. Universal dynamics in the expansion of vortex clusters in a dissipative two-dimensional superfluid. *Phys. Rev. Res.* **2**, 033138 (2020).

82. 82.

Berloff, N. G. Padé approximations of solitary wave solutions of the Gross–Pitaevskii equation. *J. Phys. A* **37**, 11729 (2004).

83. 83.

Rorai, C., Skipper, J., Kerr, R. M. & Sreenivasan, K. R. Approach and separation of quantised vortices with balanced cores. *J. Fluid Mech.* **808**, 641–667 (2016).

84. 84.

Villois, A., Proment, D. & Krstulovic, G. Universal and nonuniversal aspects of vortex reconnections in superfluids. *Phys. Rev. Fluids* **2**, 044701 (2017).

85. 85.

Galantucci, L., Baggaley, A. W., Parker, N. G. & Barenghi, C. F. Crossover from interaction to driven regimes in quantum vortex reconnections. *Proc. Natl Acad. Sci. USA* **116**, 12204–12211 (2019).

86. 86.

Numasato, R., Tsubota, M. & L'vov, V. S. Direct energy cascade in two-dimensional compressible quantum turbulence. *Phys. Rev. A* **81**, 063630 (2010).

87. 87.

Horng, T.-L., Hsueh, C.-H., Su, S.-W., Kao, Y.-M. & Gou, S.-C. Two-dimensional quantum turbulence in a nonuniform Bose-Einstein condensate. *Phys. Rev. A* **80**, 023618 (2009).

88. 88.

Griffin, A., Nazarenko, S. & Proment, D. Breaking of Josephson junction oscillations and onset of quantum turbulence in Bose–Einstein condensates. *J. Phys. A* **53**, 175701 (2020).

89. 89.

Xhani, K. et al. Critical transport and vortex dynamics in a thin atomic Josephson junction. *Phys. Rev. Lett.* **124**, 045301 (2020).

90. 90.

Xhani, K. et al. Dynamical phase diagram of ultracold Josephson junctions. *New J. Phys.* **22**, 123006 (2020).

Acknowledgements

We thank C. F. Barenghi, A. Bulgac, P. Magierski, F. Marino, M. McNeil Forbes, N. P. Proukakis, G. Wlazłowski and the Quantum Gases group at LENS for fruitful discussions, and N. Cooper for careful reading of the manuscript. This work was supported by the European Research Council under grant agreement no. 307032, the EPSRC under grant no. EP/R005192/1, the Italian Ministry of University and Research under the PRIN2017 project CEnTraL, and the EU's Horizon 2020 research and innovation programme under Marie Skłodowska-Curie grant agreement no. 843303.

Author information

Author notes

1. F. Scazza

Present address: Department of Physics, University of Trieste, Trieste, Italy

Affiliations

1. European Laboratory for Nonlinear Spectroscopy (LENS), Sesto Fiorentino, Italy

W. J. Kwon, G. Del Pace, K. Xhani, A. Muzi Falconi, M. Inguscio, F. Scazza & G. Roati

2. Istituto Nazionale di Ottica del Consiglio Nazionale delle Ricerche (CNR-INO), Sesto Fiorentino, Italy

W. J. Kwon, G. Del Pace, K. Xhani, A. Muzi Falconi, M. Inguscio, F. Scazza & G. Roati

3. Joint Quantum Centre (JQC) Durham–Newcastle, School of Mathematics, Statistics and Physics, Newcastle University, Newcastle upon Tyne, UK

L. Galantucci

4. Department of Engineering, Campus Bio-Medico University of Rome, Rome, Italy

M. Inguscio

Contributions

W.J.K., G.D.P., F.S. and G.R. conceived the study. W.J.K., G.D.P., A.M.F. and F.S. performed the experiments. W.J.K. and G.D.P. analysed the

experimental data. K.X. and L.G. carried out numerical simulations. All authors contributed to the interpretation of the results and to the writing of the manuscript.

Corresponding author

Correspondence to [W. J. Kwon](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Extended data

is available for this paper at <https://doi.org/10.1038/s41586-021-04047-4>.

Peer review information *Nature* thanks Ashton Bradley, Maren Mossman, Davide Proment 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. In situ profiles of homogeneous sample and an individual vortex.](#)

a, In-plane density profile of a UFG sample from a single experimental shot, along with centred vertical and horizontal cuts averaged over 15 different experimental realizations. **b**, *In situ* vortex profile (inset) and its

integrated radial profile (symbols) in a BEC sample. The image consists of the average of 10 experimental realisations. The measured radial profile is fitted with a Lorentzian function (solid line), yielding a width of 0.87(6) μm . This matches the expected value $\xi \approx 0.68 \mu\text{m}$, once the optical resolution of the imaging system is taken into account (see text).

Extended Data Fig. 2 Orbiting motion of a large vortex dipole in UFGs.

a, A single dipole of $d \approx 12 \mu\text{m}$ orbiting the homogeneous unitary Fermi gas of radius $R = 45 \mu\text{m}$. It rectilinearly crosses the cloud and then orbits it immediately adjacent to the boundary, in stark contrast to the observation of a vortex dipole shrinking and expanding during its propagation in a harmonic trap due to density inhomogeneity³⁷. Each image is a single experimental shot. **b**, A trajectory obtained from the identical realisations of **a**. The hold time t varies from 0 to 500 ms with time intervals of 50 ms. The light red \times signs (blue +) indicate single realisations of each vortex (antivortex) for the given t . The red (blue) circles represent the averaged positions of the vortices (antivortices) at the given t . Error bars indicate standard deviation over about 20 experimental measurements. After one orbit $\sim 500 - 550$ ms, a survival probability of a vortex dipole decreases below 50%, probably due to interaction with the boundary.

Extended Data Fig. 3 Decay of short vortex dipoles due to self-annihilation in UFG and BCS regimes.

The dipole half-life τ for each initial d_{12} , i.e. the time required for N_{vd} to drop to half of its initial value, is determined by fitting $N_{\text{vd}}(t)$ with a sigmoid function $1/(1+e^{(t-\tau)/\gamma})$, where the γ is used as the measurement uncertainty. The only exception is the shortest dipole shown in the BCS regime, which is fitted with an exponential function. The initial d_{12} is controlled to range from 3.4 to 6 μm (lighter colours denotes shorter dipoles). See also Fig. 1f. Error bars show the standard error of the mean over ~ 40 experimental realisations.

Extended Data Fig. 4 Time evolution of the number of vortex dipoles N_{vd} during dipole-dipole collisions.

Examples of $N_{\text{vd}}(t)$ for head-on (120°) collisions are shown as orange (purple) symbols in (a) BEC, (b) UFG, and (c) BCS superfluids. Each data point consists of 40 – 60 same experimental realisations and the error bar indicates the standard error of the mean. Data sets are part of those for which P_a is shown in Fig. 3f-h of the main text, and specifically: (a) $d_{\text{in}} \simeq 5\xi$, (b) $d_{\text{in}} \simeq 16/k_F$, and (c) $d_{\text{in}} \simeq 18/k_F$ (head-on) and $d_{\text{in}} \simeq 24/k_F$ (120°). Shaded regions mark the time interval of vortex partner-exchange during a collision, estimated via DPV model imposing the condition $0.9 < d_{13}(t)/d_{12}(t) < 1.1$. The drop of $N_{\text{vd}}(t)$ approximately matches this interval, confirming that the observed annihilations do not stem from single-dipole self-annihilations, but are an outcome of the collision dynamics. Experimental images show typical examples of a partial annihilation for a 120° collisions (b) and a rarely observed annihilation from head-on collisions in BCS superfluids (c). Images consist of single independent experimental shots.

Extended Data Fig. 5 Vortex annihilation images for BECs and UFGs.

Additional images display the clear emission of a density excitation following vortex annihilations in head-on collisions for (a) BECs and (b) UFGs. Two vortex dipoles collide horizontally as in Fig. 4 of the main text. The images in first and second rows of (a) are obtained independently with the same experimental parameters as in Fig. 4. By measuring the ring sizes of the density pulses observed in BECs ($t = 9$ ms and $t = 12$ ms), we find that the propagation speed of the density pulse is around $4.4(3)$ mm/s which coincides with the speed of sound evaluated from the mean density along the tight z -direction of the cloud. Annihilation images observed in UFGs are in general not as clear as in BECs, yet a number of images showing small-amplitude density waves propagating outwards are detected. Each shot is acquired in an independent experimental realisation.

Extended Data Fig. 6 Numerical criterion for selecting the vortex interaction period.

The temporal evolution of the direction β of the velocity of vortex 1 (cf. inset) is displayed for a head-on collision with $d_{\text{in}} = 4.63\xi$. The shaded area indicates the interaction interval $[t_1, t_2]$ during which the the dipole-dipole interaction takes place, with t_1 being the last time instant where $\beta \approx 0^\circ$ and t_2 the first instant where $\beta \approx 90^\circ$. Inset: trajectories of the four vortices in the head-on collision. The dashed blue rectangle denotes the interaction region $[t_1, t_2]$.

Extended Data Fig. 7 Time evolution of the compressible kinetic energy $\langle E \rangle_{\{k\}}^{\{c\}}$ for the head-on collision and $d_{\text{in}} = 4.43\xi$.

The vertical dashed lines indicate times t_1 and t_2 , edges of the interaction interval. The increase of the compressible kinetic energy shown in Fig. 3c (inset) in the main paper is defined as $\langle E \rangle_{\{k\}}^{\{c\}} = \langle E \rangle_{\{k\}}^{\{k\}} + \langle E \rangle_{\{c\}}^{\{c\}}$, $\langle E \rangle_{\{k\}}^{\{k\}} = \langle E \rangle_{\{k\}}^{\{k\}}(t_2) - \langle E \rangle_{\{k\}}^{\{k\}}(t_1)$, $\langle E \rangle_{\{c\}}^{\{c\}} = \langle E \rangle_{\{c\}}^{\{c\}}(t_2) - \langle E \rangle_{\{c\}}^{\{c\}}(t_1)$. More in detail, the initial $\langle E \rangle_{\{k\}}^{\{k\}}(t_1)$ and the final $\langle E \rangle_{\{k\}}^{\{k\}}(t_2)$ values of the compressible energy are extracted by computing an average value on a time interval of width δt centred at t_1 and t_2 , respectively. This is a characteristic time interval defined as $\delta t = d_{\text{in}}/v_d$ corresponding to the shaded areas in the plot, where $v_d = \hbar/Md_{\text{in}}$ is the vortex dipole velocity.

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Kwon, W.J., Del Pace, G., Xhani, K. *et al.* Sound emission and annihilations in a programmable quantum vortex collider. *Nature* **600**, 64–69 (2021). <https://doi.org/10.1038/s41586-021-04047-4>

- Received: 31 May 2021
- Accepted: 22 September 2021
- Published: 01 December 2021
- Issue Date: 02 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04047-4>

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/s41586-021-04047-4>

- Article
- Open Access
- [Published: 01 December 2021](#)

Advancing mathematics by guiding human intuition with AI

- [Alex Davies](#) [ORCID: orcid.org/0000-0003-4917-5234¹](#),
- [Petar Veličković¹](#),
- [Lars Buesing¹](#),
- [Sam Blackwell¹](#),
- [Daniel Zheng¹](#),
- [Nenad Tomasev](#) [ORCID: orcid.org/0000-0003-1624-0220¹](#),
- [Richard Tanburn¹](#),
- [Peter Battaglia¹](#),
- [Charles Blundell¹](#),
- [András Juhász²](#),
- [Marc Lackenby²](#),
- [Geordie Williamson³](#),
- [Demis Hassabis](#) [ORCID: orcid.org/0000-0003-2812-9917¹](#) &
- [Pushmeet Kohli](#) [ORCID: orcid.org/0000-0002-7466-7997¹](#)

[Nature](#) volume 600, pages 70–74 (2021)

- 59k Accesses
- 1205 Altmetric
- [Metrics details](#)

Subjects

- [Computer science](#)
- [Pure mathematics](#)
- [Statistics](#)

Abstract

The practice of mathematics involves discovering patterns and using these to formulate and prove conjectures, resulting in theorems. Since the 1960s, mathematicians have used computers to assist in the discovery of patterns and formulation of conjectures¹, most famously in the Birch and Swinnerton-Dyer conjecture², a Millennium Prize Problem³. Here we provide examples of new fundamental results in pure mathematics that have been discovered with the assistance of machine learning—demonstrating a method by which machine learning can aid mathematicians in discovering new conjectures and theorems. We propose a process of using machine learning to discover potential patterns and relations between mathematical objects, understanding them with attribution techniques and using these observations to guide intuition and propose conjectures. We outline this machine-learning-guided framework and demonstrate its successful application to current research questions in distinct areas of pure mathematics, in each case showing how it led to meaningful mathematical contributions on important open problems: a new connection between the algebraic and geometric structure of knots, and a candidate algorithm predicted by the combinatorial invariance conjecture for symmetric groups⁴. Our work may serve as a model for collaboration between the fields of mathematics and artificial intelligence (AI) that can achieve surprising results by leveraging the respective strengths of mathematicians and machine learning.

[Download PDF](#)

Main

One of the central drivers of mathematical progress is the discovery of patterns and formulation of useful conjectures: statements that are suspected to be true but have not been proven to hold in all cases. Mathematicians have always used data to help in this process—from the early hand-calculated prime tables used by Gauss and others that led to the prime number theorem⁵, to modern computer-generated data^{1,5} in cases such as the Birch and Swinnerton-Dyer conjecture². The introduction of computers to generate data and test conjectures afforded mathematicians a new understanding of problems that were previously inaccessible⁶, but while computational techniques have become consistently useful in other parts of the mathematical process^{7,8}, artificial intelligence (AI) systems have not yet established a similar place. Prior systems for generating conjectures have either contributed genuinely useful research conjectures⁹ via methods that do not easily generalize to other mathematical areas¹⁰, or have demonstrated novel, general methods for finding conjectures¹¹ that have not yet yielded mathematically valuable results.

AI, in particular the field of machine learning^{12,13,14}, offers a collection of techniques that can effectively detect patterns in data and has increasingly demonstrated utility in

scientific disciplines¹⁵. In mathematics, it has been shown that AI can be used as a valuable tool by finding counterexamples to existing conjectures¹⁶, accelerating calculations¹⁷, generating symbolic solutions¹⁸ and detecting the existence of structure in mathematical objects¹⁹. In this work, we demonstrate that AI can also be used to assist in the discovery of theorems and conjectures at the forefront of mathematical research. This extends work using supervised learning to find patterns^{20,21,22,23,24} by focusing on enabling mathematicians to understand the learned functions and derive useful mathematical insight. We propose a framework for augmenting the standard mathematician’s toolkit with powerful pattern recognition and interpretation methods from machine learning and demonstrate its value and generality by showing how it led us to two fundamental new discoveries, one in topology and another in representation theory. Our contribution shows how mature machine learning methodologies can be adapted and integrated into existing mathematical workflows to achieve novel results.

Guiding mathematical intuition with AI

A mathematician’s intuition plays an enormously important role in mathematical discovery—“It is only with a combination of both rigorous formalism and good intuition that one can tackle complex mathematical problems”²⁵. The following framework, illustrated in Fig. 1, describes a general method by which mathematicians can use tools from machine learning to guide their intuitions concerning complex mathematical objects, verifying their hypotheses about the existence of relationships and helping them understand those relationships. We propose that this is a natural and empirically productive way that these well-understood techniques in statistics and machine learning can be used as part of a mathematician’s work.

Fig. 1: Flowchart of the framework.



The process helps guide a mathematician’s intuition about a hypothesized function f , by training a machine learning model to estimate that function over a particular distribution of data P_Z . The insights from the accuracy of the learned function \hat{f} and attribution techniques applied to it can aid in the understanding of the

problem and the construction of a closed-form f . The process is iterative and interactive, rather than a single series of steps.

Concretely, it helps guide a mathematician's intuition about the relationship between two mathematical objects $X(z)$ and $Y(z)$ associated with z by identifying a function \hat{f} such that $\hat{f}(X(z)) \approx Y(z)$ and analysing it to allow the mathematician to understand properties of the relationship. As an illustrative example: let z be convex polyhedra, $X(z) \in \{\{\mathbb{Z}\}\}^2 \times \{\mathbb{R}\}^2$ be the number of vertices and edges of z , as well as the volume and surface area, and $Y(z) \in \mathbb{Z}$ be the number of faces of z . Euler's formula states that there is an exact relationship between $X(z)$ and $Y(z)$ in this case: $X(z) \cdot (-1, 1, 0, 0) + 2 = Y(z)$. In this simple example, among many other ways, the relationship could be rediscovered by the traditional methods of data-driven conjecture generation¹. However, for $X(z)$ and $Y(z)$ in higher-dimensional spaces, or of more complex types, such as graphs, and for more complicated, nonlinear \hat{f} , this approach is either less useful or entirely infeasible.

The framework helps guide the intuition of mathematicians in two ways: by verifying the hypothesized existence of structure/patterns in mathematical objects through the use of supervised machine learning; and by helping in the understanding of these patterns through the use of attribution techniques.

In the supervised learning stage, the mathematician proposes a hypothesis that there exists a relationship between $X(z)$ and $Y(z)$. By generating a dataset of $X(z)$ and $Y(z)$ pairs, we can use supervised learning to train a function \hat{f} that predicts $Y(z)$, using only $X(z)$ as input. The key contributions of machine learning in this regression process are the broad set of possible nonlinear functions that can be learned given a sufficient amount of data. If \hat{f} is more accurate than would be expected by chance, it indicates that there may be such a relationship to explore. If so, attribution techniques can help in the understanding of the learned function \hat{f} sufficiently for the mathematician to conjecture a candidate f . Attribution techniques can be used to understand which aspects of \hat{f} are relevant for predictions of $Y(z)$. For example, many attribution techniques aim to quantify which component of $X(z)$ the function \hat{f} is sensitive to. The attribution technique we use in our work, gradient saliency, does this by calculating the derivative of outputs of \hat{f} , with respect to the inputs. This allows a mathematician to identify and prioritize aspects of the problem that are most likely to be relevant for the relationship. This iterative process might need to be repeated several times before a viable conjecture is settled on. In this process, the mathematician can guide the choice of conjectures to those that not just fit the data but also seem interesting, plausibly true and, ideally, suggestive of a proof strategy.

Conceptually, this framework provides a 'test bed for intuition'—quickly verifying whether an intuition about the relationship between two quantities may be worth

pursuing and, if so, guidance as to how they may be related. We have used the above framework to help mathematicians to obtain impactful mathematical results in two cases—discovering and proving one of the first relationships between algebraic and geometric invariants in knot theory and conjecturing a resolution to the combinatorial invariance conjecture for symmetric groups⁴, a well-known conjecture in representation theory. In each area, we demonstrate how the framework has successfully helped guide the mathematician to achieve the result. In each of these cases, the necessary models can be trained within several hours on a machine with a single graphics processing unit.

Topology

Low-dimensional topology is an active and influential area of mathematics. Knots, which are simple closed curves in $\{\{\mathbb{R}\}\}^3$, are one of the key objects that are studied, and some of the subject's main goals are to classify them, to understand their properties and to establish connections with other fields. One of the principal ways that this is carried out is through invariants, which are algebraic, geometric or numerical quantities that are the same for any two equivalent knots. These invariants are derived in many different ways, but we focus on two of the main categories: hyperbolic invariants and algebraic invariants. These two types of invariants are derived from quite different mathematical disciplines, and so it is of considerable interest to establish connections between them. Some examples of these invariants for small knots are shown in Fig. 2. A notable example of a conjectured connection is the volume conjecture²⁶, which proposes that the hyperbolic volume of a knot (a geometric invariant) should be encoded within the asymptotic behaviour of its coloured Jones polynomials (which are algebraic invariants).

Fig. 2: Examples of invariants for three hyperbolic knots.



We hypothesized that there was a previously undiscovered relationship between the geometric and algebraic invariants.

Our hypothesis was that there exists an undiscovered relationship between the hyperbolic and algebraic invariants of a knot. A supervised learning model was able to detect the existence of a pattern between a large set of geometric invariants and the signature $\sigma(K)$, which is known to encode important information about a knot K , but was not previously known to be related to the hyperbolic geometry. The most relevant features identified by the attribution technique, shown in Fig. 3a, were three invariants of the cusp geometry, with the relationship visualized partly in Fig. 3b. Training a second model with $X(z)$ consisting of only these measurements achieved a very similar accuracy, suggesting that they are a sufficient set of features to capture almost all of the effect of the geometry on the signature. These three invariants were the real and imaginary parts of the meridional translation μ and the longitudinal translation λ . There is a nonlinear, multivariate relationship between these quantities and the signature. Having been guided to focus on these invariants, we discovered that this relationship is best understood by means of a new quantity, which is linearly related to the signature. We introduce the ‘natural slope’, defined to be $\text{slope}(K) = \text{Re}(\lambda/\mu)$, where Re denotes the real part. It has the following geometric interpretation. One can realize the meridian curve as a geodesic γ on the Euclidean torus. If one fires off a geodesic γ^\perp from this orthogonally, it will eventually return and hit γ at some point. In doing so, it will have travelled along a longitude minus some multiple of the meridian. This multiple is the natural slope. It need not be an integer, because the endpoint of γ^\perp might not be the same as its starting point. Our initial conjecture relating natural slope and signature was as follows.

Fig. 3: Knot theory attribution.



- a**, Attribution values for each of the input $X(z)$. The features with high values are those that the learned function is most sensitive to and are probably relevant for further exploration. The 95% confidence interval error bars are across 10 retrainings of the model. **b**, Example visualization of relevant features—the real part of the meridional translation against signature, coloured by the longitudinal translation.

Conjecture: There exist constants c_1 and c_2 such that, for every hyperbolic knot K ,

$$\text{(1)} \quad \$\$ |\sigma(K) - \text{slope}(K)| < c_1 \text{vol}(K) + c_2 \$\$$$

While this conjecture was supported by an analysis of several large datasets sampled from different distributions, we were able to construct counterexamples using braids of a specific form. Subsequently, we were able to establish a relationship between $\text{slope}(K)$, signature $\sigma(K)$, volume $\text{vol}(K)$ and one of the next most salient geometric invariants, the injectivity radius $\text{inj}(K)$ (ref. [27](#)).

Theorem: There exists a constant c such that, for any hyperbolic knot K ,

$$\text{(2)} \quad \$\$ |\sigma(K) - \text{slope}(K)| \leq c \text{vol}(K) \text{inj}(K)^{-3} \$\$$$

It turns out that the injectivity radius tends not to get very small, even for knots of large volume. Hence, the term $\text{inj}(K)^{-3}$ tends not to get very large in practice. However, it would clearly be desirable to have a theorem that avoided the dependence on $\text{inj}(K)^{-3}$, and we give such a result that instead relies on short geodesics, another of the most salient features, in the Supplementary Information. Further details and a full proof of the above theorem are available in ref. [27](#). Across the datasets we generated, we can place a lower bound of $c \geq 0.23392$, and it would be reasonable to conjecture that c is at most 0.3, which gives a tight relationship in the regions in which we have calculated.

The above theorem is one of the first results that connect the algebraic and geometric invariants of knots and has various interesting applications. It directly implies that the signature controls the non-hyperbolic Dehn surgeries on the knot and that the natural slope controls the genus of surfaces in \mathbb{R}^4 whose boundary is the knot. We expect that this newly discovered relationship between natural slope and signature will have many other applications in low-dimensional topology. It is surprising that a simple yet profound connection such as this has been overlooked in an area that has been extensively studied.

Representation theory

Representation theory is the theory of linear symmetry. The building blocks of all representations are the irreducible ones, and understanding them is one of the most important goals of representation theory. Irreducible representations generalize the fundamental frequencies of Fourier analysis [28](#). In several important examples, the structure of irreducible representations is governed by Kazhdan–Lusztig (KL)

polynomials, which have deep connections to combinatorics, algebraic geometry and singularity theory. KL polynomials are polynomials attached to pairs of elements in symmetric groups (or more generally, pairs of elements in Coxeter groups). The combinatorial invariance conjecture is a fascinating open conjecture concerning KL polynomials that has stood for 40 years, with only partial progress²⁹. It states that the KL polynomial of two elements in a symmetric group S_N can be calculated from their unlabelled Bruhat interval³⁰, a directed graph. One barrier to progress in understanding the relationship between these objects is that the Bruhat intervals for non-trivial KL polynomials (those that are not equal to 1) are very large graphs that are difficult to develop intuition about. Some examples of small Bruhat intervals and their KL polynomials are shown in Fig. 4.

Fig. 4: Two example dataset elements, one from S_5 and one from S_6 .



The combinatorial invariance conjecture states that the KL polynomial of a pair of permutations should be computable from their unlabelled Bruhat interval, but no such function was previously known.

We took the conjecture as our initial hypothesis, and found that a supervised learning model was able to predict the KL polynomial from the Bruhat interval with reasonably high accuracy. By experimenting on the way in which we input the Bruhat interval to the network, it became apparent that some choices of graphs and features were particularly conducive to accurate predictions. In particular, we found that a subgraph inspired by prior work³¹ may be sufficient to calculate the KL polynomial, and this was supported by a much more accurate estimated function.

Further structural evidence was found by calculating salient subgraphs that attribution techniques determined were most relevant and analysing the edge distribution in these graphs compared to the original graphs. In Fig. 5a, we aggregate the relative frequency of the edges in the salient subgraphs by the reflection that they represent. It shows that extremal reflections (those of the form $(0, i)$ or $(i, N-1)$ for S_N) appear more commonly in salient subgraphs than one would expect, at the expense of simple reflections (those of the form $(i, i+1)$), which is confirmed over many retrainings of

the model in Fig. 5b. This is notable because the edge labels are not given to the network and are not recoverable from the unlabelled Bruhat interval. From the definition of KL polynomials, it is intuitive that the distinction between simple and non-simple reflections is relevant for calculating it; however, it was not initially obvious why extremal reflections would be overrepresented in salient subgraphs. Considering this observation led us to the discovery that there is a natural decomposition of an interval into two parts—a hypercube induced by one set of extremal edges and a graph isomorphic to an interval in S_{N-1} .

Fig. 5: Representation theory attribution.



- a**, An example heatmap of the percentage increase in reflections present in the salient subgraphs compared with the average across intervals in the dataset when predicting q^4 . **b**, The percentage of observed edges of each type in the salient subgraph for 10 retrainings of the model compared to 10 bootstrapped samples of the same size from the dataset. The error bars are 95% confidence intervals, and the significance level shown was determined using a two-sided two-sample t -test. * $p < 0.05$; *** $p < 0.0001$. **c**, Illustration for the interval $021435-240513 \in S_6$ of the interesting substructures that were discovered through the iterative process of hypothesis,

supervised learning and attribution. The subgraph inspired by previous work³¹ is highlighted in red, the hypercube in green and the decomposition component isomorphic to an interval in S_{N-1} in blue.

The importance of these two structures, illustrated in Fig. 5c, led to a proof that the KL polynomial can be computed directly from the hypercube and S_{N-1} components through a beautiful formula that is summarized in the [Supplementary Information](#). A further detailed treatment of the mathematical results is given in ref. ³².

Theorem: Every Bruhat interval admits a canonical hypercube decomposition along its extremal reflections, from which the KL polynomial is directly computable.

Remarkably, further tests suggested that all hypercube decompositions correctly determine the KL polynomial. This has been computationally verified for all of the $\sim 3 \times 10^6$ intervals in the symmetric groups up to S_7 and more than 1.3×10^5 non-isomorphic intervals sampled from the symmetric groups S_8 and S_9 .

Conjecture: The KL polynomial of an unlabelled Bruhat interval can be calculated using the previous formula with any hypercube decomposition.

This conjectured solution, if proven true, would settle the combinatorial invariance conjecture for symmetric groups. This is a promising direction as not only is the conjecture empirically verified up to quite large examples, but it also has a particularly nice form that suggests potential avenues for attacking the conjecture. This case demonstrates how non-trivial insights about the behaviour of large mathematical objects can be obtained from trained models, such that new structure can be discovered.

Conclusion

In this work we have demonstrated a framework for mathematicians to use machine learning that has led to mathematical insight across two distinct disciplines: one of the first connections between the algebraic and geometric structure of knots and a proposed resolution to a long-standing open conjecture in representation theory. Rather than use machine learning to directly generate conjectures, we focus on helping guide the highly tuned intuition of expert mathematicians, yielding results that are both interesting and deep. It is clear that intuition plays an important role in elite performance in many human pursuits. For example, it is critical for top Go players and the success of AlphaGo (ref. ³³) came in part from its ability to use machine learning to learn elements of play that humans perform intuitively. It is similarly seen as critical for top mathematicians—Ramanujan was dubbed the Prince of Intuition³⁴ and it has

inspired reflections by famous mathematicians on its place in their field^{35,36}. As mathematics is a very different, more cooperative endeavour than Go, the role of AI in assisting intuition is far more natural. Here we show that there is indeed fruitful space to assist mathematicians in this aspect of their work.

Our case studies demonstrate how a foundational connection in a well-studied and mathematically interesting area can go unnoticed, and how the framework allows mathematicians to better understand the behaviour of objects that are too large for them to otherwise observe patterns in. There are limitations to where this framework will be useful—it requires the ability to generate large datasets of the representations of objects and for the patterns to be detectable in examples that are calculable. Further, in some domains the functions of interest may be difficult to learn in this paradigm. However, we believe there are many areas that could benefit from our methodology. More broadly, it is our hope that this framework is an effective mechanism to allow for the introduction of machine learning into mathematicians' work, and encourage further collaboration between the two fields.

Methods

Framework

Supervised learning

In the supervised learning stage, the mathematician proposes a hypothesis that there exists a relationship between $X(z)$ and $Y(z)$. In this work we assume that there is no known function mapping from $X(z)$ to $Y(z)$, which in turn implies that X is not invertible (otherwise there would exist a known function $Y \circ X^{-1}$). While there may still be value to this process when the function is known, we leave this for future work. To test the hypothesis that X and Y are related, we generate a dataset of $X(z)$, $Y(z)$ pairs, where z is sampled from a distribution P_Z . The results of the subsequent stages will hold true only for the distribution P_Z , and not the whole space Z . Initially, sensible choices for P_Z would be, for example, uniformly over the first N items for Z with a notion of ordering, or uniformly at random where possible. In subsequent iterations, P_Z may be chosen to understand the behaviour on different parts of the space Z (for example, regions of Z that may be more likely to provide counterexamples to a particular hypothesis). To first test whether a relation between $X(z)$ and $Y(z)$ can be found, we use supervised learning to train a function \hat{f} that approximately maps $X(z)$ to $Y(z)$. In this work we use neural networks as the supervised learning method, in part because they can be easily adapted to many different types of X and Y and knowledge of any inherent geometry (in terms of invariances and symmetries) of

the input domain X can be incorporated into the architecture of the network³⁷. We consider a relationship between $X(z)$ and $Y(z)$ to be found if the accuracy of the learned function \hat{f} is statistically above chance on further samples from P_Z on which the model was not trained. The converse is not true; namely, if the model cannot predict the relationship better than chance, it may mean that a pattern exists, but is sufficiently complicated that it cannot be captured by the given model and training procedure. If it does indeed exist, this can give a mathematician confidence to pursue a particular line of enquiry in a problem that may otherwise be only speculative.

Attribution techniques

If a relationship is found, the attribution stage is to probe the learned function \hat{f} with attribution techniques to further understand the nature of the relationship. These techniques attempt to explain what features or structures are relevant to the predictions made by \hat{f} , which can be used to understand what parts of the problem are relevant to explore further. There are many attribution techniques in the body of literature on machine learning and statistics, including stepwise forward feature selection³⁸, feature occlusion and attention weights³⁹. In this work we use gradient-based techniques⁴⁰, broadly similar to sensitivity analysis in classical statistics and sometimes referred to as saliency maps. These techniques attribute importance to the elements of $X(z)$, by calculating how much \hat{f} changes in predictions of $Y(z)$ given small changes in $X(z)$. We believe these are a particularly useful class of attribution techniques as they are conceptually simple, flexible and easy to calculate with machine learning libraries that support automatic differentiation^{41,42,43}. Information extracted via attribution techniques can then be useful to guide the next steps of mathematical reasoning, such as conjecturing closed-form candidates f , altering the sampling distribution P_Z or generating new hypotheses about the object of interest z , as shown in Fig. 1. This can then lead to an improved or corrected version of the conjectured relationship between these quantities.

Topology

Problem framing

Not all knots admit a hyperbolic geometry; however, most do, and all knots can be constructed from hyperbolic and torus knots using satellite operations⁴⁴. In this work we focus only on hyperbolic knots. We characterize the hyperbolic structure of the knot complement by a number of easily computable invariants. These invariants do not fully define the hyperbolic structure, but they are representative of the most commonly interesting properties of the geometry. Our initial general hypothesis was that the hyperbolic invariants would be predictive of algebraic invariants. The specific

hypothesis we investigated was that the geometry is predictive of the signature. The signature is an ideal candidate as it is a well-understood and common invariant, it is easy to calculate for large knots and it is an integer, which makes the prediction task particularly straightforward (compared to, for example, a polynomial).

Data generation

We generated a number of datasets from different distributions P_Z on the set of knots using the SnapPy software package^{[45](#)}, as follows.

1. 1.

All knots up to 16 crossings ($\sim 1.7 \times 10^6$ knots), taken from the Regina census^{[46](#)}.

2. 2.

Random knot diagrams of 80 crossings generated by SnapPy's random_link function ($\sim 10^6$ knots). As random knot generation can potentially lead to duplicates, we calculate a large number of invariants for each knot diagram and remove any samples that have identical invariants to a previous sample, as they are likely to represent that same knot with very high probability.

3. 3.

Knots obtained as the closures of certain braids. Unlike the previous two datasets, the knots that were produced here are not, in any sense, generic. Instead, they were specifically constructed to disprove Conjecture 1. The braids that we used were 4-braids ($n = 11,756$), 5-braids ($n = 13,217$) and 6-braids ($n = 10,897$). In terms of the standard generators σ_i for these braid groups, the braids were chosen to be $\langle (\{\sigma_i\}_{j=1}^{n_j})^{k_j} \rangle$. The integers i_j were chosen uniformly at random for the appropriate braid group. The powers n_j were chosen uniformly at random in the ranges $[-10, -3]$ and $[3, 10]$. The final power N was chosen uniformly between 1 and 10. The quantity $\sum |n_j|$ was restricted to be at most 15 for 5-braids and 6-braids and 12 for 4-braids, and the total number of crossings $N \sum |n_j|$ was restricted to lie in the range between 10 and 60. The rationale for these restrictions was to ensure a rich set of examples that were small enough to avoid an excessive number of failures in the invariant computations.

For the above datasets, we computed a number of algebraic and geometric knot invariants. Different datasets involved computing different subsets of these, depending

on their role in forming and examining the main conjecture. Each of the datasets contains a subset of the following list of invariants: signature, slope, volume, meridional translation, longitudinal translation, injectivity radius, positivity, Chern–Simons invariant, symmetry group, hyperbolic torsion, hyperbolic adjoint torsion, invariant trace field, normal boundary slopes and length spectrum including the linking numbers of the short geodesics.

The computation of the canonical triangulation of randomly generated knots fails in SnapPy in our data generation process in between 0.6% and 1.7% of the cases, across datasets. The computation of the injectivity radius fails between 2.8% of the time on smaller knots up to 7.8% of the time on datasets of knots with a higher number of crossings. On knots up to 16 crossings from the Regina dataset, the injectivity radius computation failed in 5.2% of the cases. Occasional failures can occur in most of the invariant computations, in which case the computations continue for the knot in question for the remaining invariants in the requested set. Additionally, as the computational complexity of some invariants is high, operations can time out if they take more than 5 min for an invariant. This is a flexible bound and ultimately a trade-off that we have used only for the invariants that were not critical for our analysis, to avoid biasing the results.

Data encoding

The following encoding scheme was used for converting the different types of features into real valued inputs for the network: reals directly encoded; complex numbers as two reals corresponding to the real and imaginary parts; categoricals as one-hot vectors.

All features are normalized by subtracting the mean and dividing by the variance. For simplicity, in Fig. 3a, the salience values of categoricals are aggregated by taking the maximum value of the saliences of their encoded features.

Model and training procedure

The model architecture used for the experiments was a fully connected, feed-forward neural network, with hidden unit sizes [300, 300, 300] and sigmoid activations. The task was framed as a multi-class classification problem, with the distinct values of the signature as classes, cross-entropy loss as an optimizable loss function and test classification accuracy as a metric of performance. It is trained for a fixed number of steps using a standard optimizer (Adam). All settings were chosen as a priori reasonable values and did not need to be optimized.

Process

First, to assess whether there may be a relationship between the geometry and algebra of a knot, we trained a feed-forward neural network to predict the signature from measurements of the geometry on a dataset of randomly sampled knots. The model was able to achieve an accuracy of 78% on a held-out test set, with no errors larger than ± 2 . This is substantially higher than chance (a baseline accuracy of 25%), which gave us strong confidence that a relationship may exist.

To understand how this prediction is being made by the network, we used gradient-based attribution to determine which measurements of the geometry are most relevant to the signature. We do this using a simple sensitivity measure \mathbf{r}^i that averages the gradient of the loss L with respect to a given input feature \mathbf{x}_i over all of the examples \mathbf{x} in a dataset $\{\mathbf{X}\}$:

$$\frac{1}{|\mathbf{X}|} \sum_{\mathbf{x} \in \mathbf{X}} \nabla L / \nabla \mathbf{x}_i \quad (3)$$

This quantity for each input feature is shown in Fig. 3a, where we can determine that the relevant measurements of the geometry appear to be what is known as the cusp shape: the meridional translation, which we will denote μ , and the longitudinal translation, which we will denote λ . This was confirmed by training a new model to predict the signature from only these three measurements, which was able to achieve the same level of performance as the original model.

To confirm that the slope is a sufficient aspect of the geometry to focus on, we trained a model to predict the signature from the slope alone. Visual inspection of the slope and signature in Extended Data Fig. 1a,b shows a clear linear trend, and training a linear model on this data results in a test accuracy of 78%, which is equivalent to the predictive power of the original model. This implies that the slope linearly captures all of the information about the signature that the original model had extracted from the geometry.

Evaluation

The confidence intervals on the feature saliences were calculated by retraining the model 10 times with a different train/test split and a different random seed initializing both the network weights and training procedure.

Representation theory

Data generation

For our main dataset we consider the symmetric groups up to S_9 . The first symmetric group that contains a non-trivial Bruhat interval whose KL polynomial is not simply 1 is S_5 , and the largest interval in S_9 contains $9! \approx 3.6 \times 10^5$ nodes, which starts to pose computational issues when used as inputs to networks. The number of intervals in a symmetric group S_N is $O(N^2)$, which results in many billions of intervals in S_9 . The distribution of coefficients of the KL polynomials uniformly across intervals is very imbalanced, as higher coefficients are especially rare and associated with unknown complex structure. To adjust for this and simplify the learning problem, we take advantage of equivalence classes of Bruhat intervals that eliminate many redundant small polynomials^{[47](#)}. This has the added benefit of reducing the number of intervals per symmetric group (for example, to ~ 2.9 million intervals in S_9). We further reduce the dataset by including a single interval for each distinct KL polynomial for all graphs with the same number of nodes, resulting in 24,322 non-isomorphic graphs for S_9 . We split the intervals randomly into train/test partitions at 80%/20%.

Data encoding

The Bruhat interval of a pair of permutations is a partially ordered set of the elements of the group, and it can be represented as a directed acyclic graph where each node is labelled by a permutation, and each edge is labelled by a reflection. We add two features at each node representing the in-degree and out-degree of that node.

Model and training procedure

For modelling the Bruhat intervals, we used a particular GraphNet architecture called a message-passing neural network (MPNN)^{[48](#)}. The design of the model architecture (in terms of activation functions and directionality) was motivated by the algorithms for computing KL polynomials from labelled Bruhat intervals. While labelled Bruhat intervals contain privileged information, these algorithms hinted at the kind of computation that may be useful for computing KL polynomial coefficients.

Accordingly, we designed our MPNN to algorithmically align to this computation^{[49](#)}. The model is bi-directional, with a hidden layer width of 128, four propagation steps and skip connections. We treat the prediction of each coefficient of the KL polynomial as a separate classification problem.

Process

First, to gain confidence that the conjecture is correct, we trained a model to predict coefficients of the KL polynomial from the unlabelled Bruhat interval. We were able to do so across the different coefficients with reasonable accuracy (Extended Data

Table 1) giving some evidence that a general function may exist, as a four-step MPNN is a relatively simple function class. We trained a GraphNet model on the basis of a newly hypothesized representation and could achieve significantly better performance, lending evidence that it is a sufficient and helpful representation to understand the KL polynomial.

To understand how the predictions were being made by the learned function \hat{f} , we used gradient-based attribution to define a salient subgraph S_G for each example interval G , induced by a subset of nodes in that interval, where L is the loss and x_v is the feature for vertex v :

$$\text{(4)} \quad \$\$ \{ S \}_{G} = \{ v \in G \mid \frac{\partial L}{\partial x_v} | > C_k \} \$\$$$

We then aggregated the edges by their edge type (each is a reflection) and compared the frequency of their occurrence to the overall dataset. The effect on extremal edges was present in the salient subgraphs for predictions of the higher-order terms (q^3, q^4), which are the more complicated and less well-understood terms.

Evaluation

The threshold C_k for salient nodes was chosen a priori as the 99th percentile of attribution values across the dataset, although the results are present for different values of C_k in the range [95, 99.5]. In Fig. 5a, we visualize a measure of edge attribution for a particular snapshot of a trained model for expository purposes. This view will change across time and random seeds, but we can confirm that the pattern remains by looking at aggregate statistics over many runs of training the model, as in Fig. 5b. In this diagram, the two-sample two-sided t -test statistics are as follows—simple edges: $t = 25.7, P = 4.0 \times 10^{-10}$; extremal edges: $t = -13.8, P = 1.1 \times 10^{-7}$; other edges: $t = -3.2, P = 0.01$. These significance results are robust to different settings of the hyper-parameters of the model.

Code availability

Interactive notebooks to regenerate the results for both knot theory and representation theory have been made available for download at <https://github.com/deepmind>.

Data availability

The generated datasets used in the experiments have been made available for download at <https://github.com/deepmind>.

References

1. 1.

Borwein, J. & Bailey, D. *Mathematics by Experiment* (CRC, 2008).

2. 2.

Birch, B. J. & Swinnerton-Dyer, H. P. F. Notes on elliptic curves. II. *J. Reine Angew. Math.* **1965**, 79–108 (1965).

3. 3.

Carlson, J. et al. *The Millennium Prize Problems* (American Mathematical Soc., 2006).

4. 4.

Brenti, F. Kazhdan-Lusztig polynomials: history, problems, and combinatorial invariance. *Sémin. Lothar. Combin.* **49**, B49b (2002).

5. 5.

Hoche, R. *Nicomachi Geraseni Pythagorei Introductionis Arithmeticae Libri 2* (In aedibus BG Teubneri, 1866).

6. 6.

Khovanov, M. Patterns in knot cohomology, I. *Exp. Math.* **12**, 365–374 (2003).

7. 7.

Appel, K. I. & Haken, W. *Every Planar Map Is Four Colorable* Vol. 98 (American Mathematical Soc., 1989).

8. 8.

Scholze, P. Half a year of the Liquid Tensor Experiment: amazing developments Xena <https://xenaproject.wordpress.com/2021/06/05/half-a-year-of-the-liquid-tensor-experiment-amazing-developments/> (2021).

9. 9.

Fajtlowicz, S. in *Annals of Discrete Mathematics* Vol. 38 113–118 (Elsevier, 1988).

10. 10.

Larson, C. E. in *DIMACS Series in Discrete Mathematics and Theoretical Computer Science* Vol. 69 (eds Fajtlowicz, S. et al.) 297–318 (AMS & DIMACS, 2005).

11. 11.

Raayoni, G. et al. Generating conjectures on fundamental constants with the Ramanujan machine. *Nature* **590**, 67–73 (2021).

12. 12.

MacKay, D. J. C. *Information Theory, Inference and Learning Algorithms* (Cambridge Univ. Press, 2003).

13. 13.

Bishop, C. M. *Pattern Recognition and Machine Learning* (Springer, 2006).

14. 14.

LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).

15. 15.

Raghu, M. & Schmidt, E. A survey of deep learning for scientific discovery. Preprint at <https://arxiv.org/abs/2003.11755> (2020).

16. 16.

Wagner, A. Z. Constructions in combinatorics via neural networks. Preprint at <https://arxiv.org/abs/2104.14516> (2021).

17. 17.

Peifer, D., Stillman, M. & Halpern-Leistner, D. Learning selection strategies in Buchberger’s algorithm. Preprint at <https://arxiv.org/abs/2005.01917> (2020).

18. 18.

Lample, G. & Charton, F. Deep learning for symbolic mathematics. Preprint at <https://arxiv.org/abs/1912.01412> (2019).

19. 19.

He, Y.-H. Machine-learning mathematical structures. Preprint at <https://arxiv.org/abs/2101.06317> (2021).

20. 20.

Carifio, J., Halverson, J., Krioukov, D. & Nelson, B. D. Machine learning in the string landscape. *J. High Energy Phys.* **2017**, 157 (2017).

21. 21.

Heal, K., Kulkarni, A. & Sertöz, E. C. Deep learning Gauss-Manin connections. Preprint at <https://arxiv.org/abs/2007.13786> (2020).

22. 22.

Hughes, M. C. A neural network approach to predicting and computing knot invariants. Preprint at <https://arxiv.org/abs/1610.05744> (2016).

23. 23.

Levitt, J. S. F., Hajij, M. & Sazdanovic, R. Big data approaches to knot theory: understanding the structure of the Jones polynomial. Preprint at <https://arxiv.org/abs/1912.10086> (2019).

24. 24.

Jejjala, V., Kar, A. & Parrikar, O. Deep learning the hyperbolic volume of a knot. *Phys. Lett. B* **799**, 135033 (2019).

25. 25.

Tao, T. There's more to mathematics than rigour and proofs *Blog* <https://terrytao.wordpress.com/career-advice/theres-more-to-mathematics-than-rigour-and-proofs/> (2016).

26. 26.

Kashaev, R. M. The hyperbolic volume of knots from the quantum dilogarithm. *Lett. Math. Phys.* **39**, 269–275 (1997).

27. 27.

Davies, A., Lackenby, M., Juhasz, A. & Tomašev, N. The signature and cusp geometry of hyperbolic knots. Preprint at arxiv.org (in the press).

28. 28.

Curtis, C. W. & Reiner, I. *Representation Theory of Finite Groups and Associative Algebras* Vol. 356 (American Mathematical Soc., 1966).

29. 29.

Brenti, F., Caselli, F. & Marietti, M. Special matchings and Kazhdan–Lusztig polynomials. *Adv. Math.* **202**, 555–601 (2006).

30. 30.

Verma, D.-N. Structure of certain induced representations of complex semisimple Lie algebras. *Bull. Am. Math. Soc.* **74**, 160–166 (1968).

31. 31.

Braden, T. & MacPherson, R. From moment graphs to intersection cohomology. *Math. Ann.* **321**, 533–551 (2001).

32. 32.

Blundell, C., Buesing, L., Davies, A., Veličković, P. & Williamson, G. Towards combinatorial invariance for Kazhdan-Lusztig polynomials. Preprint at arxiv.org (in the press).

33. 33.

Silver, D. et al. Mastering the game of Go with deep neural networks and tree search. *Nature* **529**, 484–489 (2016).

34. 34.

Kanigel, R. *The Man Who Knew Infinity: a Life of the Genius Ramanujan* (Simon and Schuster, 2016).

35. 35.

Poincaré, H. *The Value of Science: Essential Writings of Henri Poincaré* (Modern Library, 1907).

36. 36.

Hadamard, J. *The Mathematician's Mind* (Princeton Univ. Press, 1997).

37. 37.

Bronstein, M. M., Bruna, J., Cohen, T. & Veličković, P. Geometric deep learning: grids, groups, graphs, geodesics, and gauges. Preprint at <https://arxiv.org/abs/2104.13478> (2021).

38. 38.

Efroymson, M. A. in *Mathematical Methods for Digital Computers* 191–203 (John Wiley, 1960).

39. 39.

Xu, K. et al. Show, attend and tell: neural image caption generation with visual attention. In *Proc. International Conference on Machine Learning* 2048–2057 (PMLR, 2015).

40. 40.

Sundararajan, M., Taly, A. & Yan, Q. Axiomatic attribution for deep networks. In *Proc. International Conference on Machine Learning* 3319–3328 (PMLR, 2017).

41. 41.

Bradbury, J. et al. JAX: composable transformations of Python+NumPy programs (2018); <https://github.com/google/jax>

42. 42.

Martín A. B. A. D. I. et al. TensorFlow: large-scale machine learning on heterogeneous systems (2015); <https://doi.org/10.5281/zenodo.4724125>.

43. 43.

Paszke, A. et al. in *Advances in Neural Information Processing Systems* 32 (eds Wallach, H. et al.) 8024–8035 (Curran Associates, 2019).

44. 44.

Thurston, W. P. Three dimensional manifolds, Kleinian groups and hyperbolic geometry. *Bull. Am. Math. Soc.* **6**, 357–381 (1982).

45. 45.

Culler, M., Dunfield, N. M., Goerner, M. & Weeks, J. R. SnapPy, a computer program for studying the geometry and topology of 3-manifolds (2020); <http://snappy.computop.org>.

46. 46.

Burton, B. A. The next 350 million knots. In *Proc. 36th International Symposium on Computational Geometry (SoCG 2020)* (Schloss Dagstuhl-Leibniz-Zentrum für Informatik, 2020).

47. 47.

Warrington, G. S. Equivalence classes for the μ -coefficient of Kazhdan–Lusztig polynomials in S_n . *Exp. Math.* **20**, 457–466 (2011).

48. 48.

Gilmer, J., Schoenholz, S. S., Riley, P. F., Vinyals, O. & Dahl, G. E. Neural message passing for quantum chemistry. Preprint at <https://arxiv.org/abs/1704.01212> (2017).

49. 49.

Veličković, P., Ying, R., Padovano, M., Hadsell, R. & Blundell, C. Neural execution of graph algorithms. Preprint at <https://arxiv.org/abs/1910.10593> (2019).

Acknowledgements

We thank J. Ellenberg, S. Mohamed, O. Vinyals, A. Gaunt, A. Fawzi and D. Saxton for advice and comments on early drafts; J. Vonk for contemporary supporting work; X. Glorot and M. Overlan for insight and assistance; and A. Pierce, N. Lambert, G. Holland, R. Ahamed and C. Meyer for assistance coordinating the research. This research was funded by DeepMind.

Author information

Affiliations

1. DeepMind, London, UK

Alex Davies, Petar Veličković, Lars Buesing, Sam Blackwell, Daniel Zheng, Nenad Tomašev, Richard Tanburn, Peter Battaglia, Charles Blundell, Demis Hassabis & Pushmeet Kohli

2. University of Oxford, Oxford, UK

András Juhász & Marc Lackenby

3. University of Sydney, Sydney, New South Wales, Australia

Geordie Williamson

Contributions

A.D., D.H. and P.K. conceived of the project. A.D., A.J. and M.L. discovered the knot theory results, with D.Z. and N.T. running additional experiments. A.D., P.V. and G.W. discovered the representation theory results, with P.V. designing the model, L.B. running additional experiments, and C.B. providing advice and ideas. S.B. and R.T. provided additional support, experiments and infrastructure. A.D., D.H. and P.K. directed and managed the project. A.D. and P.V. wrote the paper with help and feedback from P.B., C.B., M.L., A.J., G.W., P.K. and D.H.

Corresponding authors

Correspondence to [Alex Davies](#) or [Pushmeet Kohli](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Sanjeev Arora, Christian Stump 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 Empirical relationship between slope and signature.](#)

a Signature vs slope for random dataset. **b** Signature vs slope for Regina dataset.

Extended Data Table 1 Model accuracies at predicting KL coefficients from Bruhat intervals in S_9

Supplementary information

[Supplementary Information](#)

This file contains Supplementary text, Figs. 1–6, hyperbolic knots and references.

[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

Davies, A., Veličković, P., Buesing, L. *et al.* Advancing mathematics by guiding human intuition with AI. *Nature* **600**, 70–74 (2021). <https://doi.org/10.1038/s41586-021-04086-x>

- Received: 10 July 2021
- Accepted: 30 September 2021
- Published: 01 December 2021
- Issue Date: 02 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04086-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

[Artificial intelligence aids intuition in mathematical discovery](#)

- Christian Stump

News & Views 01 Dec 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04086-x>

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

- Article
- [Published: 01 December 2021](#)

Optomechanical dissipative solitons

- [Jing Zhang^{1,2}](#),
- [Bo Peng¹](#),
- [Seunghwi Kim³](#),
- [Faraz Monifi¹](#),
- [Xuefeng Jiang¹](#),
- [Yihang Li¹](#),
- [Peng Yu⁴](#),
- [Lianqing Liu⁴](#),
- [Yu-xi Liu⁵](#),
- [Andrea Alù ORCID: orcid.org/0000-0002-4297-5274^{3,6}](#) &
- [Lan Yang ORCID: orcid.org/0000-0002-9052-0450¹](#)

[Nature](#) volume 600, pages 75–80 (2021)

- 2238 Accesses
- 26 Altmetric
- [Metrics details](#)

Subjects

- [Nonlinear optics](#)
- [Silicon photonics](#)
- [Solitons](#)

Abstract

Nonlinear wave–matter interactions may give rise to solitons, phenomena that feature inherent stability in wave propagation and unusual spectral characteristics. Solitons have been created in a variety of physical systems and have had important roles in a broad range of applications, including communications, spectroscopy and metrology^{1,2,3,4}. In recent years, the realization of dissipative Kerr optical solitons in microcavities has led to the generation of frequency combs in a chip-scale platform^{5,6,7,8,9,10}. Within a cavity, photons can interact with mechanical modes. Cavity optomechanics has found applications for frequency conversion, such as microwave-to-optical or radio-frequency-to-optical^{11,12,13}, of interest for communications and interfacing quantum systems operating at different frequencies. Here we report the observation of mechanical micro-solitons excited by optical fields in an optomechanical microresonator, expanding soliton generation in optical resonators to a different spectral window. The optical field circulating along the circumference of a whispering gallery mode resonator triggers a mechanical nonlinearity through optomechanical coupling, which in turn induces a time-varying periodic modulation on the propagating mechanical mode, leading to a tailored modal dispersion. Stable localized mechanical wave packets—mechanical solitons—can be realized when the mechanical loss is compensated by phonon gain and the optomechanical nonlinearity is balanced by the tailored modal dispersion. The realization of mechanical micro-solitons driven by light opens up new avenues for optomechanical technologies¹⁴ and may find applications in acoustic sensing, information processing, energy storage, communications and surface acoustic wave technology.

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: Mechanism of acoustic-wave propagation in an optomechanical resonator.



Fig. 2: Generation of an optomechanical soliton.

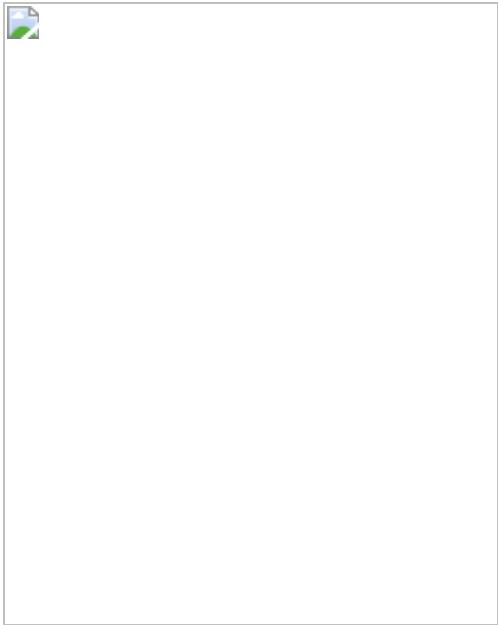


Fig. 3: Localized periodic phonon pulses in the cnoidal wave regime and soliton regime.

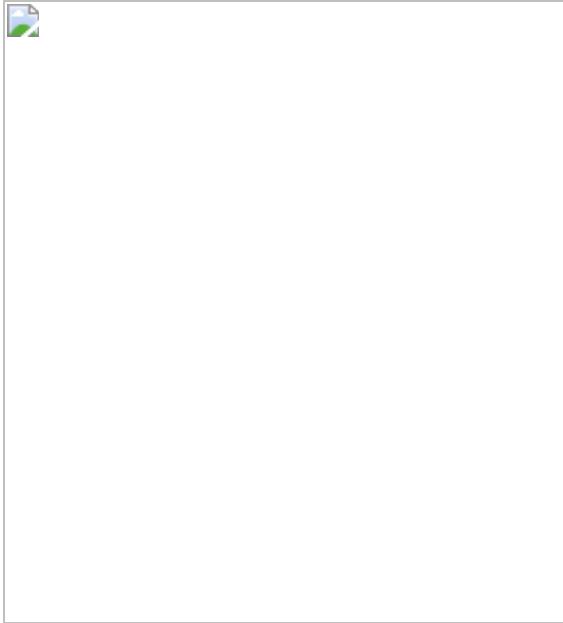
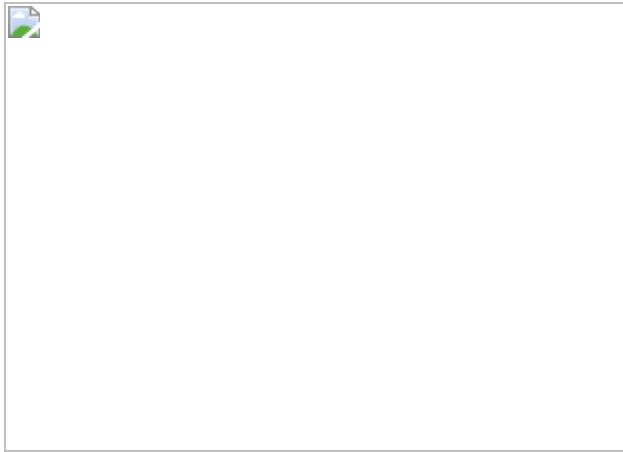


Fig. 4: Detection of a low-frequency vibration of a cantilever tip by an optomechanical soliton.



Data availability

The datasets generated during and/or analysed in this study are available from the corresponding author upon reasonable request.

References

1. 1.

Holzwarth, R. et al. Optical frequency synthesizer for precision spectroscopy. *Phys. Rev. Lett.* **85**, 2264–2267 (2000).

2. 2.

Udem, T., Holzwarth, R. & Hänsch, T. W. Optical frequency metrology. *Nature* **416**, 233–237 (2002).

3. 3.

Haus, H. A. & Wong, W. S. Solitons in optical communications. *Rev. Mod. Phys.* **68**, 423–444 (1996).

4. 4.

Kippenberg, T. J., Holzwarth, R. & Diddams, S. A. Microresonator-based optical frequency combs. *Science* **332**, 555–559 (2011).

5. 5.

Herr, T. et al. Temporal solitons in optical microresonators. *Nat. Photon.* **8**, 145–152 (2014).

6. 6.

Kippenberg, T. J., Gaeta, A. L., Lipson, M. & Gorodetsky, M. L. Dissipative Kerr solitons in optical microresonators. *Science* **361**, eaan8083 (2018).

7. 7.

Brasch, V. et al. Photonic chip-based optical frequency comb using soliton Cherenkov radiation. *Science* **351**, 357–360 (2016).

8. 8.

Stern, B., Ji, X., Okawachi, Y., Gaeta, A. L. & Lipson, M. Battery-operated integrated frequency comb generator. *Nature* **562**, 401–405 (2018).

9. 9.

Suh, M.-G., Yang, Q.-F., Yang, K. Y., Yi, X. & Vahala, K. J. Microresonator soliton dual-comb spectroscopy. *Science* **354**, 600–603 (2016).

10. 10.

Yi, X., Yang, Q.-F., Yang, K. Y., Suh, M.-G. & Vahala, K. J. Soliton frequency comb at microwave rates in a high-*Q* silica microresonator. *Optica* **2**, 1078–1085 (2015).

11. 11.

Shao, L. et al. Microwave-to-optical conversion using lithium niobate thin-film acoustic resonators. *Optica* **6**, 1498–1505 (2019).

12. 12.

Forsch, M. et al. Microwave-to-optics conversion using a mechanical oscillator in its quantum ground state. *Nat. Phys.* **16**, 69–74 (2020).

13. 13.

Yamazaki, R. et al. Radio-frequency-to-optical conversion using acoustic and optical whispering-gallery modes. *Phys. Rev. A* **101**, 053839 (2020).

14. 14.

Aspelmeyer, M., Kippenberg, T. J. & Marquardt, F. Cavity optomechanics. *Rev. Mod. Phys.* **86**, 1391–1452 (2014).

15. 15.

Chan, J. et al. Laser cooling of a nanomechanical oscillator into its quantum ground state. *Nature* **478**, 89–92 (2011).

16. 16.

LIGO Scientific Collaboration and Virgo Collaboration. Observation of gravitational waves from a binary black hole merger. *Phys. Rev. Lett.* **116**, 061102 (2016).

17. 17.

Grudinin, I. S., Lee, H., Painter, O. & Vahala, K. J. Phonon laser action in a tunable two-level system. *Phys. Rev. Lett.* **104**, 083901 (2010).

18. 18.

Jing, H. et al. PT -symmetric phonon laser. *Phys. Rev. Lett.* **113**, 053604 (2014).

19. 19.

Zhang, J. et al. A phonon laser operating at an exceptional point. *Nat. Photon.* **12**, 479–484 (2018).

20. 20.

Carmon, T., Cross, M. C. & Vahala, K. J. Chaotic quivering of micron-scaled onchip resonators excited by centrifugal optical pressure. *Phys. Rev. Lett.* **98**, 167203 (2007).

21. 21.

Monifi, F. et al. Optomechanically induced stochastic resonance and chaos transfer between optical fields. *Nat. Photon.* **10**, 399–405 (2016).

22. 22.

Gan, J.-H., Xiong, H., Si, L.-G., Lü, X.-Y. & Wu, Y. Solitons in optomechanical arrays. *Opt. Lett.* **41**, 2676–2679 (2016).

23. 23.

Xiong, H., Gan, J. H. & Wu, Y. Kuznetsov–Ma soliton dynamics based on the mechanical effect of light. *Phys. Rev. Lett.* **119**, 153901 (2017).

24. 24.

Xiong, H. & Wu, Y. Optomechanical Akhmediev breathers. *Laser Photon. Rev.* **12**, 1700305 (2018).

25. 25.

Ganesan, A., Do, C. & Seshia, A. Phononic frequency comb via intrinsic three-wave mixing. *Phys. Rev. Lett.* **118**, 033903 (2017).

26. 26.

Butsch, A., Koehler, J. R., Noskov, R. E. & Russell, P. St. J. CW-pumped single-pass frequency comb generation by resonant

optomechanical nonlinearity in dual-nanoweb fiber. *Optica* **1**, 158–163 (2014).

27. 27.

Savchenkov, A. A., Matsko, A. B., Ilchenko, V. S., Seidel, D. & Maleki, L. Surface acoustic wave opto-mechanical oscillator and frequency comb generator. *Opt. Lett.* **36**, 3338–3340 (2011).

28. 28.

Miri, M.-A., D’Aguanno, G. & Alù, A. Optomechanical frequency combs. *New J. Phys.* **20**, 043013 (2018).

29. 29.

Del’Haye, P. et al. Optical frequency comb generation from a monolithic microresonator. *Nature* **450**, 1214–1217 (2007).

30. 30.

Savchenkov, A. A. et al. Tunable optical frequency comb with a crystalline whispering gallery mode resonator. *Phys. Rev. Lett.* **101**, 093902 (2008).

31. 31.

Rueda, A., Sedlmeir, F., Kumari, M., Leuchs, G. & Schwefel, H. G. L. Resonant electro-optic frequency comb. *Nature* **568**, 378–381 (2019); correction **569**, E11 (2019).

32. 32.

Zhang, M. et al. Broadband electro-optic frequency comb generation in a lithium niobate microring resonator. *Nature* **568**, 373–377 (2019).

33. 33.

Li, Q. et al. Stably accessing octave-spanning microresonator frequency combs in the soliton regime. *Optica* **4**, 193–203 (2017).

34. 34.

Cao, L. S., Qi, D. X., Peng, R. W., Wang, M. & Schmelcher, P. Phononic frequency combs through nonlinear resonances. *Phys. Rev. Lett.* **112**, 075505 (2014).

35. 35.

Czaplewski, D. A. et al. Bifurcation generated mechanical frequency comb. *Phys. Rev. Lett.* **121**, 244302 (2018).

36. 36.

Hao, H. Y. & Maris, H. J. Experiments with acoustic solitons in crystalline solids. *Phys. Rev. B* **64**, 064302 (2001).

37. 37.

Hereman, W. Shallow water waves and solitary waves. In *Encyclopedia of Complexity and Systems Science* (ed. Meyers, R. A.) 480 (Springer, 2009); https://doi.org/10.1007/978-0-387-30440-3_480.

38. 38.

Barland, S. et al. Temporal localized structures in optical resonators. *Adv. Phys. X* **2**, 496–517 (2017).

39. 39.

Lugiato, L., Prati, F. & Brambilla, M. *Nonlinear Optical Systems* (Cambridge Univ. Press, 2015).

40. 40.

Jang, J. K., Erkintalo, M., Murdoch, S. G. & Coen, S. Ultraweak long-range interactions of solitons observed over astronomical distances.

Nat. Photon. **7**, 657–663 (2013).

41. 41.

Barland, S. et al. Cavity solitons as pixels in semiconductor microcavities. *Nature* **419**, 699–702 (2002).

42. 42.

Leo, F. et al. Temporal cavity solitons in one-dimensional Kerr media as bits in an all-optical buffer. *Nat. Photon.* **4**, 471–476 (2010).

43. 43.

Grelu, P. & Akhmediev, N. Dissipative solitons for mode-locked lasers. *Nat. Photon.* **6**, 84–92 (2012).

44. 44.

Korteweg, D. J. & de Vries, G. On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves. *Philos. Mag.* **39**, 422–443 (1895).

45. 45.

Boyd, J. P. The double cnoidal wave of the Korteweg–de Vries equation: an overview. *J. Math. Phys.* **25**, 3390–3401 (1984).

46. 46.

Nayanov, V. I. Surface acoustic cnoidal waves and solitons in a LiNbO₃-(SiO film) structure. *JETP Lett.* **44**, 314–317 (1986); translated from *Pis'ma Zh. Eksp. Teor. Fiz.* **44**, 245–247 (1986).

47. 47.

Fiore, V. et al. Storing optical information as a mechanical excitation in a silica optomechanical resonator. *Phys. Rev. Lett.* **107**, 133601 (2011).

48. 48.

Carmon, T., Rokhsari, H., Yang, L., Kippenberg, T. J. & Vahala, K. J. Temporal behavior of radiation-pressure-induced vibrations of an optical microcavity phonon mode. *Phys. Rev. Lett.* **94**, 223902 (2005).

Acknowledgements

The project is supported by the NSF grant number EFMA1641109 and ARO grant numbers W911NF1710189 and W911NF1210026. J.Z. is supported by the NSFC under grant numbers 61622306 and 11674194. Y.-x.L. is supported by the NSFC under grant number 61025022. Y.-x.L. and J.Z. are supported by the National Basic Research Program of China (973 Program) under grant number 2014CB921401, the Tsinghua University Initiative Scientific Research Program, and the Tsinghua National Laboratory for Information Science and Technology (TNList) Cross-discipline Foundation. L.L. is supported by the NSFC under grant number 61925307. S.K. and A.A. are supported by the Office of Naval Research and the Air Force Office of Scientific Research.

Author information

Affiliations

1. Department of Electrical and Systems Engineering, Washington University, St. Louis, MO, USA

Jing Zhang, Bo Peng, Faraz Monifi, Xuefeng Jiang, Yihang Li & Lan Yang

2. Department of Automation, Tsinghua University, Beijing, P. R. China

Jing Zhang

3. Photonics Initiative, Advanced Science Research Center, City University of New York, New York, NY, USA

Seunghwi Kim & Andrea Alù

4. State Key Laboratory of Robotics, Shenyang Institute of Automation,
Chinese Academy of Sciences, Shenyang, P. R. China

Peng Yu & Lianqing Liu

5. Institute of Microelectronics, Tsinghua University, Beijing, P. R. China

Yu-xi Liu

6. Physics Program, Graduate Center, City University of New York, New
York, NY, USA

Andrea Alù

Contributions

J.Z., B.P., F.M. and L.Y. conceived the idea. L.Y. designed the experiments.
J.Z. performed the experiments and processed the data with the help of X.J.,
Y.L., P.Y. and L.L. J.Z. and S.K. provided theoretical analysis under the
guidance of Y.-x.L. and A.A. J.Z. and L.Y. wrote the manuscript with
contributions from all authors. L.Y. supervised the project.

Corresponding author

Correspondence to [Lan Yang](#).

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.

Supplementary information

Supplementary Information

This file contains supplementary text, supplementary equations, supplementary sections S1–S8 and supplementary references.

Peer Review File

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Zhang, J., Peng, B., Kim, S. *et al.* Optomechanical dissipative solitons. *Nature* **600**, 75–80 (2021). <https://doi.org/10.1038/s41586-021-04012-1>

- Received: 08 January 2020
- Accepted: 09 September 2021
- Published: 01 December 2021
- Issue Date: 02 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04012-1>

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

Optomechanics joins the soliton club

- Alessia Pasquazi

News & Views 01 Dec 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04012-1>

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

- Article
- [Published: 01 December 2021](#)

In situ Raman spectroscopy reveals the structure and dissociation of interfacial water

- [Yao-Hui Wang¹](#)^{na1},
- [Shisheng Zheng²](#)^{na1},
- [Wei-Min Yang¹](#),
- [Ru-Yu Zhou¹](#),
- [Quan-Feng He¹](#),
- [Petar Radjenovic¹](#),
- [Jin-Chao Dong¹](#),
- [Shunning Li²](#),
- [Jiaxin Zheng²](#),
- [Zhi-Lin Yang¹](#),
- [Gary Attard³](#),
- [Feng Pan](#) [ORCID: orcid.org/0000-0002-8216-1339²](#),
- [Zhong-Qun Tian](#) [ORCID: orcid.org/0000-0002-9775-8189^{1,4}](#) &
- [Jian-Feng Li](#) [ORCID: orcid.org/0000-0003-1598-6856^{1,4,5}](#)

[Nature](#) volume 600, pages 81–85 (2021)

- 6201 Accesses
- 14 Altmetric
- [Metrics details](#)

Subjects

- [Electrocatalysis](#)
- [SERS](#)

Abstract

Understanding the structure and dynamic process of water at the solid–liquid interface is an extremely important topic in surface science, energy science and catalysis^{1,2,3}. As model catalysts, atomically flat single-crystal electrodes exhibit well-defined surface and electric field properties, and therefore may be used to elucidate the relationship between structure and electrocatalytic activity at the atomic level^{4,5}. Hence, studying interfacial water behaviour on single-crystal surfaces provides a framework for understanding electrocatalysis^{6,7}. However, interfacial water is notoriously difficult to probe owing to interference from bulk water and the complexity of interfacial environments⁸. Here, we use electrochemical, *in situ* Raman spectroscopic and computational techniques to investigate the interfacial water on atomically flat Pd single-crystal surfaces. Direct spectral evidence reveals that interfacial water consists of hydrogen-bonded and hydrated Na⁺ ion water. At hydrogen evolution reaction (HER) potentials, dynamic changes in the structure of interfacial water were observed from a random distribution to an ordered structure due to bias potential and Na⁺ ion cooperation. Structurally ordered interfacial water facilitated high-efficiency electron transfer across the interface, resulting in higher HER rates. The electrolytes and electrode surface effects on interfacial water were also probed and found to affect water structure. Therefore, through local cation tuning strategies, we anticipate that these results may be generalized to enable ordered interfacial water to improve electrocatalytic reaction rates.

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: Probing interfacial water on Pd(hkl) surfaces.



Fig. 2: Raman spectra of interfacial water.

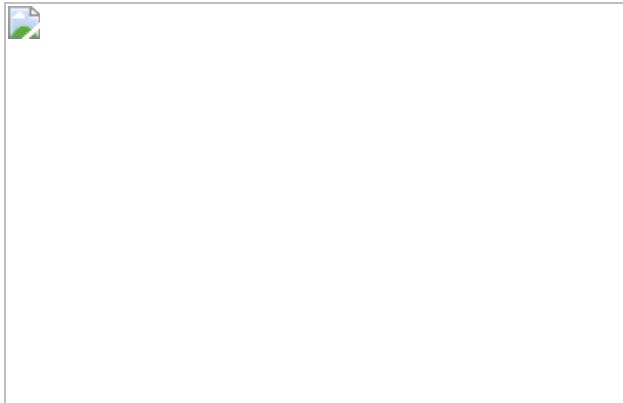


Fig. 3: Water dissociation.

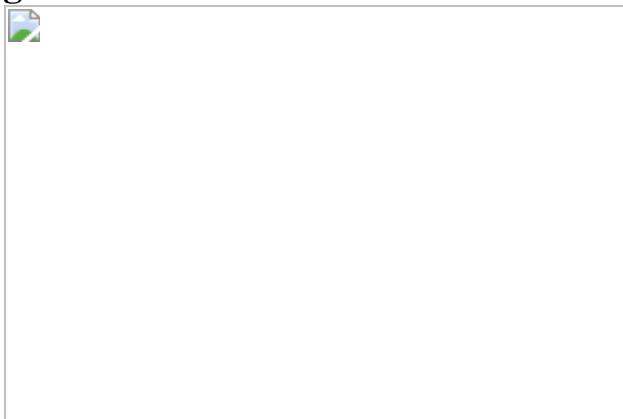


Fig. 4: HER profiles and Raman spectra of interfacial water.

Data availability

The data generated or analysed during this study are included in this published article and its Supplementary information files. [Source data](#) are provided with this paper.

Code availability

The code that supports the findings of this research is available from the corresponding authors upon reasonable request.

References

1. 1.

Seh, Z. W. et al. Combining theory and experiment in electrocatalysis: insights into materials design. *Science* **355**, eaad4998 (2017).

2. 2.

Mubeen, S. et al. An autonomous photosynthetic device in which all charge carriers derive from surface plasmons. *Nat. Nanotechnol.* **8**, 247–251 (2013).

3. 3.

Guo, J. et al. Real-space imaging of interfacial water with submolecular resolution. *Nat. Mater.* **13**, 184–189 (2014).

4. 4.

Vidal-Iglesias, F. J., Solla-Gullon, J., Herrero, E., Aldaz, A. & Feliu, J. M. Pd adatom decorated (100) preferentially oriented Pt nanoparticles for formic acid electrooxidation. *Angew. Chem. Int. Ed.* **49**, 6998–7001 (2010).

5. 5.

Strmcnik, D. et al. Enhanced electrocatalysis of the oxygen reduction reaction based on patterning of platinum surfaces with cyanide. *Nat. Chem.* **2**, 880–885 (2010).

6. 6.

Kendrick, E., Kendrick, J., Knight, K. S., Islam, M. S. & Slater, P. R. Cooperative mechanisms of fast-ion conduction in gallium-based oxides with tetrahedral moieties. *Nat. Mater.* **6**, 871–875 (2007).

7. 7.

Mesa, C. A. et al. Multihole water oxidation catalysis on haematite photoanodes revealed by operando spectroelectrochemistry and DFT. *Nat. Chem.* **12**, 82–89 (2020).

8. 8.

Velasco-Velez, J. J. et al, The structure of interfacial water on gold electrodes studied by x-ray absorption spectroscopy. *Science* **346**, 831–834 (2014).

9. 9.

Subbaraman, R. et al. Trends in activity for the water electrolyser reactions on 3d M(Ni,Co,Fe,Mn) hydr(oxy)oxide catalysts. *Nat. Mater.* **11**, 550–557 (2012).

10. 10.

Wang, X., Xu, C., Jaroniec, M., Zheng, Y. & Qiao, S. Z. Anomalous hydrogen evolution behavior in high-pH environment induced by locally generated hydronium ions. *Nat. Commun.* **10**, 4876 (2019).

11. 11.

Ledeza-Yanez, I. et al. Interfacial water reorganization as a pH-dependent descriptor of the hydrogen evolution rate on platinum electrodes. *Nat. Energy* **2**, 17031 (2017).

12. 12.

Guha, A., Narayananar, S. & Narayanan, T. N. Tuning the hydrogen evolution reaction on metals by lithium salt. *ACS Appl. Energy Mater.* **1**, 7116–7122 (2018).

13. 13.

Guha, A., Kaley, N. M., Mondal, J. & Narayanan, T. N. Engineering the hydrogen evolution reaction of transition metals: effect of Li ions. *J. Mater. Chem. A* **8**, 15795–15808 (2020).

14. 14.

Ataka, K., Yotsuyanagi, T. & Osawa, M. Potential-dependent reorientation of water molecules at an electrode/electrolyte interface studied by surface-enhanced infrared absorption spectroscopy. *J. Phys. Chem.* **100**, 10664–10672, (1996).

15. 15.

Yamakata, A. & Osawa, M. Destruction of the hydration shell around tetraalkylammonium ions at the electrochemical interface. *J. Am. Chem. Soc.* **131**, 6892–6893, (2009).

16. 16.

Tong, Y., Lapointe, F., Thamer, M., Wolf, M. & Campen, R. K. Hydrophobic water probed experimentally at the gold electrode/aqueous interface. *Angew. Chem. Int. Ed.* **56**, 4211–4214 (2017).

17. 17.

Liu, W. T. & Shen, Y. R. In situ sum-frequency vibrational spectroscopy of electrochemical interfaces with surface plasmon resonance. *Proc. Natl Acad. Sci. USA* **111**, 1293–1297 (2014).

18. 18.

Li, J. F. et al. Shell-isolated nanoparticle-enhanced Raman spectroscopy. *Nature* **464**, 392–395 (2010).

19. 19.

Dong, J. C. et al. In situ Raman spectroscopic evidence for oxygen reduction reaction intermediates at platinum single-crystal surfaces. *Nat. Energy* **4**, 60–67 (2019).

20. 20.

Davis, J. G., Gierszal, K. P., Wang, P. & Ben-Amotz, D. Water structural transformation at molecular hydrophobic interfaces. *Nature* **491**, 582–585 (2012).

21. 21.

Chen, Y. X., Zou, S. Z., Huang, K. Q. & Tian, Z. Q. SERS studies of electrode/electrolyte interfacial water part II—librations of water correlated to hydrogen evolution reaction. *J. Raman Spectrosc.* **29**, 749–756 (1998).

22. 22.

Toney, M. F. et al. Voltage-dependent ordering of water molecules at an electrode-electrolyte interface. *Nature* **368**, 444–446 (1994).

23. 23.

Senior, W. A. & Thompson, W. K. Assignment of the infra-red and Raman bands of liquid water. *Nature* **205**, 170 (1965).

24. 24.

Chen, H. C. et al. Active and stable liquid water innovatively prepared using resonantly illuminated gold nanoparticles. *ACS Nano* **8**, 2704–2713 (2014).

25. 25.

Scatena, L. F., Brown, M. G. & Richmond, G. L. Water at hydrophobic surfaces: weak hydrogen bonding and strong orientation effects. *Science* **292**, 908–912 (2001).

26. 26.

Lambert, D. K. Vibrational Stark effect of adsorbates at electrochemical interfaces. *Electrochim. Acta* **41**, 623–630 (1996).

27. 27.

Li, J. F. et al. SERS and DFT study of water on metal cathodes of silver, gold and platinum nanoparticles. *Phys. Chem. Chem. Phys.* **12**, 2493–2502 (2010).

28. 28.

Subbaraman, R. et al. Enhancing hydrogen evolution activity in water splitting by tailoring Li^+ - Ni(OH)_2 -Pt interfaces. *Science* **334**, 1256–1260 (2011).

29. 29.

Kibler, L. A. Hydrogen electrocatalysis. *ChemPhysChem* **7**, 985–991 (2006).

30. 30.

Kresse, G. & Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **59**, 1758–1775 (1999).

31. 31.

Hammer, B., Hansen, L. B. & Nørskov, J. K. Improved adsorption energetics within density-functional theory using revised Perdew-Burke-Ernzerhof functionals. *Phys. Rev. B* **59**, 7413–7421 (1999).

32. 32.

Tonigold, K. & Gross, A. Dispersive interactions in water bilayers at metallic surfaces: a comparison of the PBE and RPBE functional

including semiempirical dispersion corrections. *J. Comput. Chem.* **33**, 695–701 (2012).

33. 33.

Sakong, S., Forster-Tonigold, K. & Gross, A. The structure of water at a Pt(111) electrode and the potential of zero charge studied from first principles. *J. Chem. Phys.* **144**, 194701 (2016).

34. 34.

Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **132**, 154104 (2010).

Acknowledgements

This work was financially supported by the National Key Research and Development Program of China (2019YFA0705400, 2020YFB1505800, 2020YFB0704500 and 2019YFD0901100), the National Natural Science Foundation of China (21925404, 22021001, 21521004, 21775127, 21991151 and 21902137), the Shenzhen Science and Technology Research Grant (JCYJ20200109140416788) and ‘111’ Project (B17027). We thank Y. X. Chen, B. W. Mao, D. P. Zhan, Z. Wei and J. B. Le for fruitful discussions, and Q. Q. Zhao, M. F. Cao and Y. Hui for technical assistance.

Author information

Author notes

1. These authors contributed equally: Yao-Hui Wang, Shisheng Zheng

Affiliations

1. State Key Laboratory of Physical Chemistry of Solid Surfaces, MOE Key Laboratory of Spectrochemical Analysis and Instrumentation,

iChEM, College of Chemistry and Chemical Engineering, College of Energy, College of Materials, College of Physical Science and Technology, Xiamen University, Xiamen, China

Yao-Hui Wang, Wei-Min Yang, Ru-Yu Zhou, Quan-Feng He, Petar Radjenovic, Jin-Chao Dong, Zhi-Lin Yang, Zhong-Qun Tian & Jian-Feng Li

2. School of Advanced Materials, Peking University, Shenzhen Graduate School, Shenzhen, China

Shisheng Zheng, Shunning Li, Jiaxin Zheng & Feng Pan

3. Department of Physics, University of Liverpool, Liverpool, UK

Gary Attard

4. Innovation Laboratory for Sciences and Technologies of Energy Materials of Fujian Province (IKKEM), Xiamen, China

Zhong-Qun Tian & Jian-Feng Li

5. College of Optical and Electronic Technology, China Jiliang University, Hangzhou, China

Jian-Feng Li

Contributions

Y.-H.W., S.Z., F.P. and J.-F.L. conceived the project; Y.-H.W., R.-Y.Z. and Q.-F.H. conducted the experiments; F.P., S.Z., S.L. and J.Z. performed the mechanism study and AIMD simulation; W.-M.Y. and Z.-L.Y. performed the 3D-FDTD simulation; Y.-H.W., S.Z., P.R., G.A., F.P., Z.-Q.T. and J.-F.L. wrote the paper. All authors participated in the analysis and discussion.

Corresponding authors

Correspondence to [Feng Pan](#) or [Jian-Feng Li](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Angel Cuesta 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 file includes Supplementary Notes, Figs. 1–21, Table 1 and Refs. 1–29.

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

Wang, YH., Zheng, S., Yang, WM. *et al.* In situ Raman spectroscopy reveals the structure and dissociation of interfacial water. *Nature* **600**, 81–85 (2021). <https://doi.org/10.1038/s41586-021-04068-z>

- Received: 07 July 2020
- Accepted: 29 September 2021
- Published: 01 December 2021
- Issue Date: 02 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04068-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

[Choreographing water molecules to speed up hydrogen production](#)

- Matthias M. Waegle

News & Views 01 Dec 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04068-z>

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

- Article
- Open Access
- [Published: 20 October 2021](#)

Late Quaternary dynamics of Arctic biota from ancient environmental genomics

- [Yucheng Wang](#) ORCID: orcid.org/0000-0002-7838-226X^{1,2 na1},
- [Mikkel Winther Pedersen](#) ORCID: orcid.org/0000-0002-7291-8887^{2 na1},
- [Inger Greve Alsos](#) ORCID: orcid.org/0000-0002-8610-1085^{3 na1},
- [Bianca De Sanctis](#)^{1,4},
- [Fernando Racimo](#)²,
- [Ana Prohaska](#) ORCID: orcid.org/0000-0001-5459-6186¹,
- [Eric Coissac](#)^{3,5},
- [Hannah Lois Owens](#)⁶,
- [Marie Kristine Føreid Merkel](#)³,
- [Antonio Fernandez-Guerra](#)²,
- [Alexandra Rouillard](#) ORCID: orcid.org/0000-0001-5778-6620^{2,7},
- [Youri Lammers](#)³,
- [Adriana Alberti](#) ORCID: orcid.org/0000-0003-3372-9423^{8,9},
- [France Denoeud](#) ORCID: orcid.org/0000-0001-8819-7634⁹,
- [Daniel Money](#) ORCID: orcid.org/0000-0001-5151-3648¹,
- [Anthony H. Ruter](#)²,
- [Hugh McColl](#) ORCID: orcid.org/0000-0002-7568-4270²,
- [Nicolaj Krog Larsen](#)²,
- [Anna A. Cherezova](#) ORCID: orcid.org/0000-0002-6199-8164^{10,11},
- [Mary E. Edwards](#)^{12,13},
- [Grigory B. Fedorov](#)^{10,11},
- [James Haile](#) ORCID: orcid.org/0000-0002-8521-8337²,
- [Ludovic Orlando](#) ORCID: orcid.org/0000-0003-3936-1850¹⁴,
- [Lasse Vinner](#)²,
- [Thorfinn Sand Korneliussen](#)^{2,15},
- [David W. Beilman](#) ORCID: orcid.org/0000-0002-2625-6747¹⁶,

- [Anders A. Bjørk](#)¹⁷,
- [Jialu Cao](#)²,
- [Christoph Dockter](#) ORCID: orcid.org/0000-0001-5923-3667¹⁸,
- [Julie Esdale](#)¹⁹,
- [Galina Gusarova](#)^{3,20},
- [Kristian K. Kjeldsen](#) ORCID: orcid.org/0000-0002-8557-5131²¹,
- [Jan Mangerud](#) ORCID: orcid.org/0000-0003-4793-7557^{22,23},
- [Jeffrey T. Rasic](#) ORCID: orcid.org/0000-0002-3549-6590²⁴,
- [Birgitte Skadhauge](#) ORCID: orcid.org/0000-0001-7317-4376¹⁸,
- [John Inge Svendsen](#)^{22,23},
- [Alexei Tikhonov](#)²⁵,
- [Patrick Wincker](#) ORCID: orcid.org/0000-0001-7562-3454⁹,
- [Yingchun Xing](#)²⁶,
- [Yubin Zhang](#) ORCID: orcid.org/0000-0003-4920-3100²⁷,
- [Duane G. Froese](#) ORCID: orcid.org/0000-0003-1032-5944²⁸,
- [Carsten Rahbek](#)^{6,29},
- [David Bravo Nogues](#)⁶,
- [Philip B. Holden](#) ORCID: orcid.org/0000-0002-2369-0062³⁰,
- [Neil R. Edwards](#) ORCID: orcid.org/0000-0001-6045-8804³⁰,
- [Richard Durbin](#) ORCID: orcid.org/0000-0002-9130-1006⁴,
- [David J. Meltzer](#) ORCID: orcid.org/0000-0001-8084-9802^{2,31},
- [Kurt H. Kjær](#)²,
- [Per Möller](#)³² &
- [Eske Willerslev](#) ORCID: orcid.org/0000-0002-7081-6748^{1,2,33,34}

Nature volume 600, pages 86–92 (2021)

- 22k Accesses
- 1 Citations
- 819 Altmetric
- [Metrics details](#)

Subjects

- [Climate-change ecology](#)
- [Ecological networks](#)
- [Metagenomics](#)
- [Next-generation sequencing](#)

- [Palaeoecology](#)

Abstract

During the last glacial–interglacial cycle, Arctic biotas experienced substantial climatic changes, yet the nature, extent and rate of their responses are not fully understood^{1,2,3,4,5,6,7,8}. Here we report a large-scale environmental DNA metagenomic study of ancient plant and mammal communities, analysing 535 permafrost and lake sediment samples from across the Arctic spanning the past 50,000 years. Furthermore, we present 1,541 contemporary plant genome assemblies that were generated as reference sequences. Our study provides several insights into the long-term dynamics of the Arctic biota at the circumpolar and regional scales. Our key findings include: (1) a relatively homogeneous steppe–tundra flora dominated the Arctic during the Last Glacial Maximum, followed by regional divergence of vegetation during the Holocene epoch; (2) certain grazing animals consistently co-occurred in space and time; (3) humans appear to have been a minor factor in driving animal distributions; (4) higher effective precipitation, as well as an increase in the proportion of wetland plants, show negative effects on animal diversity; (5) the persistence of the steppe–tundra vegetation in northern Siberia enabled the late survival of several now-extinct megafauna species, including the woolly mammoth until 3.9 ± 0.2 thousand years ago (ka) and the woolly rhinoceros until 9.8 ± 0.2 ka; and (6) phylogenetic analysis of mammoth environmental DNA reveals a previously unsampled mitochondrial lineage. Our findings highlight the power of ancient environmental metagenomics analyses to advance understanding of population histories and long-term ecological dynamics.

[Download PDF](#)

Main

Climate changes are amplified at high latitudes and have pronounced effects on Arctic ecosystems¹. Their effects on Arctic plant and animal communities, as well as the human populations who are dependent on them, would have been especially pronounced during the extremely cold and arid Last Glacial Maximum (LGM) (26.5–19 ka)² and later during the rapid warming that preceded the Holocene. However, precisely what those effects were, and how they played out across the Arctic, are not fully understood. These dynamics were further complicated by differences in the timing and extent of glaciation in different regions across this vast and topographically complex landscape. Previous studies based on pollen and plant macrofossils have documented substantial spatiotemporal variations in Arctic vegetation over the past 50,000 years (50 kyr)^{1,3}, yet it continues to be debated how climatic changes during this period affected plant communities in different regions of the Arctic, and how

changes in climate and vegetation may have affected large mammals (that is, megafauna)^{4,5,6}. Skeletal remains show that several megafaunal species, including woolly mammoth (*Mammuthus primigenius*), woolly rhinoceros (*Coelodonta antiquitatis*), steppe bison (*Bison priscus*) and horse (*Equus* spp.), were abundant in the Arctic during the Pleistocene epoch, but are thought to have become regionally or globally extinct by the onset of the Holocene^{4,5}. However, the precise timing of megafaunal extinctions, and whether and to what extent some of these taxa survived into the Holocene, is uncertain. Similarly, the contribution of various abiotic and biotic drivers to the extinction process of different taxa remains an open question^{7,8}.

To address these knowledge gaps, we performed a metagenomics analysis of ancient environmental DNA (eDNA) of plants and animals recovered from sediments from sites distributed across much of the Arctic covering the past 50 kyr. Relative to other palaeoecological proxies (such as pollen and macrofossils), ancient eDNA offers distinct advantages—including greater taxonomic resolution across the full tree of life⁹ and higher spatial and temporal precision than pollen—as eDNA mainly derives from the local community¹⁰. We used metagenomic analysis rather than the widely used metabarcoding approach because it enables the sequencing of DNA fragments from entire genomes without taxon-specific amplifications, therefore improving the specificity and sensitivity of taxonomic identification, as well as facilitating the authentication of endogenous ancient DNA from modern contaminants⁹. However, metagenomic analysis requires genome-scale reference data, which are limited for most regions of the world, including the Arctic. Thus, a key component of our study is the generation of a substantial corpus of plant reference sequences.

Metagenomic dataset and database

We generated the eDNA metagenomic dataset from 535 sediment samples obtained at 74 circumpolar sites (Fig. 1). Samples come from lake sediments and stratigraphic exposures (unconsolidated permafrost). For the purpose of understanding regional variability, we grouped sites into four regions: North Atlantic; northwest and central Siberia; northeast Siberia; and North America (Fig. 1). Sample ages span the past 50 kyr, albeit in varying numbers, from all regions with the notable exception of the North Atlantic, which was largely covered by ice sheets that often erased pre-LGM deposits^{2,11}.

Fig. 1: Site distribution (North Pole-centred view).

figure1

Samples ($n = 535$) from a total of 74 circumpolar sites were grouped into four geographical regions (Supplementary Information 2). The grey dashed circle indicates the Arctic Circle (66.5° N). Site IDs are labelled on the map. The corresponding information is provided in Supplementary Data 1.

From the 535 samples, we generated 10.2 billion sequencing reads that passed the filtering criteria and were used for analysis (Methods). We created a comprehensive reference database for taxonomic identification by merging the NCBI-nt and NCBI-

RefSeq databases, and supplemented the limited genomic-scale public reference data for Arctic species with 12 Arctic animals and an extensive sequencing effort of 1,541 modern Holarctic plant genome skims (PhyloNorway; [Methods](#)). These new sequences comprise 311.3 million whole-genome contigs and provide a broader and more reliable plant reference database than previously available. The merged reference database contains a total of 380.4 million entries and covers about 1.47 million organisms. We developed a *k*-mer-based method to evaluate the availability and coverage of our combined reference database for different taxa ([Methods](#)) and found that it covers a wide range of both Arctic and non-Arctic species (Supplementary Information [9.2.3](#)). Accordingly, the addition of our new reference genomes did not cause bias towards Arctic taxa, providing confidence in our identifications. We used robust approaches to identify taxa from individual reads and collated the resulting taxonomic composition at the generic or familial level ([Methods](#)). We applied several methods to authenticate the plant and animal taxonomic profiles; the identifications were reliably classified despite the short DNA sequences that were preserved in these samples ([Methods](#)).

Moreover, 131 samples in this dataset were processed for metabarcoding, targeting the short DNA barcodes of plants^{[12](#)}, enabling a comparison between the two approaches ([Methods](#)). The results showed that the metagenomic analysis captured greater floristic and faunal diversity and achieved better taxonomic resolution (Supplementary Information [11.2](#)). We also found that only about 1.26% of the plant DNA reads are of ribosomal and chloroplast origin (Supplementary Information [9.2.5](#)), suggesting that the metabarcoding approach—which relies on organelle DNA—makes use of only a small fraction of preserved DNA. However, we acknowledge that these comparisons are sample- and method-specific; more studies are needed before broader conclusions about the relative merits of the two approaches can be reached.

Circum-Arctic vegetation dynamics

We combined plant assemblages that were reconstructed from all of the samples to describe the temporal changes in floristic composition, diversity and community structure across the Arctic (Fig. [2a](#) and Extended Data Fig. [1](#)). Our results show substantial and repeated responses of Arctic vegetation to changing climates over the past 50 kyr.

Fig. 2: Climate and vegetation changes over the past 50 kyr.

 **figure2**

a, Pan-Arctic climate changes and vegetation variations. LGM (26.5–19 ka) and Younger Dryas (YD) (12.9–11.7 ka) are indicated by grey bars. The six time intervals are indicated by light blue bars (Supplementary Information 2). The error bands denote s.e. From top to bottom (see [Methods](#) for detailed calculations): the Greenlandic ice-core $\delta^{18}\text{O}$ ratio and snow accumulation rate; the plant Shannon diversity and the Greenlandic ice-core calcium concentration; the average modelled annual temperature and precipitation for all eDNA sampling sites; the proportion of plant growth forms; the proportion of the herbaceous plant growth forms; and the vegetation turnover rates. **b**, The number of observed genera in different regions. **c**, Regional vegetation turnovers. **d**, Regional vegetation morphological compositions. The sample sizes for each region and time interval are provided in Supplementary Information 2. Calculations are supplied in the Methods.

The overall floristic diversity increased steadily from 50 ka and reached its highest levels at the onset of the LGM (about 26.5 ka), when the climate reached its coldest and driest point at many locations^{2,11} (Fig. 2a). Vegetation turnover was high before about 38 ka, and the identified shrubs, forbs and grasses suggest a shifting mosaic of steppe–tundra vegetation. Herbaceous plants were the dominant plant group until about 19 ka, with forbs more abundant than graminoids (Fig. 2a), but not as dominant as suggested by a previous metabarcoding study¹². Trees and aquatic plants were limited in distribution to lower-latitude sites—consistent with overall dry and cold climate conditions during this period. The scarcity of cold-tolerant trees such as *Pinus* and *Picea*, and absence of *Larix*, reflect low precipitation and strong winds (Fig. 2a and Extended Data Fig. 1a).

The transition into the LGM featured declining temperature and precipitation (Fig. 2a). Across the Arctic, trees remained absent, and there was a sharp decrease in floristic diversity, mainly caused by the decline in herbaceous taxa. Overall, vegetation turnover was consistently high during this decline in diversity, suggesting that cold and dry extremes caused the loss of taxa from all plant communities, although the taxa that were dominant in the pre-LGM period remained (Extended Data Fig. 1a). LGM vegetation dissimilarity was the lowest of all time periods (Extended Data Fig. 1b,c), indicating considerable homogeneity across much of the unglaciated Arctic.

After the LGM, warming towards the Bølling–Allerød interstadial (approximately 14.6–12.9 ka)¹³ led to vegetation divergence among sites (Extended Data Fig. 1b,c). There was a substantial increase in the abundance of woody plants (such as *Salix* and *Betula*), whereas the herbaceous diversity continued to decline, causing the overall diversity to reach its lowest point at the beginning of the cold Younger Dryas stadial (approximately 12.9–11.7 ka)¹⁴ (Fig. 2a). The abundance of woody taxa and vegetation turnover rate reached the highest point during the Younger Dryas; the latter is consistent with the intensive climate changes that mark the transition from the Pleistocene to the Holocene.

Shortly after the Younger Dryas, summer insolation peaked and atmospheric CO₂ reached Holocene levels¹⁵. Previously abundant plant taxa such as *Artemisia* and *Poa* rapidly declined or vanished locally. Other plant taxa, particularly boreal trees and prostrate shrubs (such as *Vaccinium*), appeared and later became abundant (Extended Data Fig. 1a), suggesting that there was a shift from open, cold-adapted tundra–steppe to a mosaic of herbaceous and woody plant communities. The floristic diversity of this more mesophilic vegetation increased during the Early Holocene as climate continued to warm and effective precipitation increased, but then declined during the middle Holocene (Fig. 2a).

Owing to dating uncertainties and limits on the temporal resolution of palaeoclimatic simulations, our results captured only broader changes in vegetation dynamics under climate change. During much of the past 50 kyr, overall plant diversity decreased when the proportion of trees and shrubs increased, as they outcompete herbaceous taxa through shading¹⁶. By contrast, when climate became more suitable for herbaceous taxa, diverse taxa expanded to share the landscape, and the overall diversity therefore increased.

Regional vegetation dynamics

Underlying the generalized pattern of Holarctic vegetation changes are significant geographical differences. Early in postglacial times, the North Atlantic experienced the sharpest rises in taxonomic richness (Fig. 2b), along with the steepest temperature increase (Extended Data Fig. 2b). The increase in postglacial richness was probably driven by species dispersals coupled with habitat diversification¹⁷, that is, gynomorphically dynamic substrates that were exposed by glacial retreat and shaped by meltwater. The resultant vegetation initially had low diversity but was rich in aquatic taxa (Fig. 2b, d). The abundance of aquatic taxa relates in part to the prevalence of samples from lakes in the North Atlantic (Supplementary Information 10), but nonetheless highlights the ability of aquatic plants to disperse rapidly into newly deglaciated terrain containing abundant streams and lake basins¹⁸. As the postglacial climate continued to warm, the overall proportion of aquatic taxa declined as trees and shrubs (for example, *Betula*, *Salix* and *Vaccinium*) became abundant in this region (Fig. 2d and Extended Data Fig. 3).

Northeast Siberia and North America experienced less radical postglacial changes in vegetation type (Fig. 2c, d). During the Late Glacial, trees and shrubs became more widely distributed, and floristic diversity started to decline—a trend that was especially pronounced in North America (Fig. 2b, d). By about 12 ka, rising sea levels had flooded the Bering Strait, and the vegetation on each side started to diverge (Extended Data Fig. 2a). In northeast Siberia, greater effective moisture within the Holocene led to the expansion of aquatic plants (such as *Hippuris* and *Menyanthes*). The previously dominant steppe taxa (for example, *Poa* and *Artemisia*) declined, although sedges, of which many species are hygrophilous, continued to be abundant (Extended Data Fig. 3). The vegetation of this region became a mosaic of steppe and tundra elements. In North America, trees such as *Populus* and *Picea* became more widespread during the Early Holocene and previously widespread steppe species declined (Fig. 2d and Extended Data Fig. 3). A broad, southern swath of eastern Beringia became boreal forest.

In contrast to the changes observed in these regions, vegetation in northwest and central Siberia remained relatively unchanged through the Pleistocene–Holocene

transition (Fig. 2c, d). However, some cold- and/or dry-adapted taxa (such as *Artemisia* and *Poa*) were replaced by forbs that were better adapted to warmer climates, and *Salix* was partially replaced by *Betula* and *Alnus* (Extended Data Fig. 3). The vegetation in this region persisted as a steppe–tundra mosaic through much of the Holocene, probably due to central Siberia’s extreme climatic continentality caused by the Siberian anticyclone¹⁹, which created largely ice-free conditions during the LGM and fostered dry hydrogeological conditions in postglacial times that mitigated the effects of rising global temperatures on vegetation¹¹.

Overall, these results show that postglacial plant communities regionally diverged in response to warming temperatures, increasing moisture, retreating ice sheets and marine transgressions. Although regions that were once overridden by continental ice sheets experienced extreme vegetation changes, the vegetation in unglaciated interior regions remained rather stable. This maritime–continental contrast highlights the importance of moisture in driving ecosystem changes in the Arctic^{7,20}. We next incorporate these insights into vegetation dynamics, together with other potential drivers, into a model to identify the factors influencing animal distributions.

Animal distribution drivers

We developed a model using reconstructed animal distributions and floristic compositions, modelled palaeoclimate variables and inferred human occurrences ([Methods](#)) to examine the relative effects of abiotic and biotic factors on Arctic mammal distributions over the past 50 kyr.

We found that certain herbivores tend to co-occur in time and space. For example, the eDNA presences of caribou, hare and vole are statistically strong co-indicators for the presence of horse and mammoth eDNA (Fig. 3). This suggests that co-existence was more common among Arctic herbivores than interspecies exclusion²¹. By contrast, the distribution of humans over time was almost entirely unrelated to the presence of most herbivores (apart from hares) (Fig. 3). Given that the model purposefully overestimated the presence of humans ([Methods](#)), their largely independent distributions from megafauna, their sparseness in the high Arctic before 4 ka (Supplementary Data 7) and the scarcity of kill sites in archaeological records, the notion of human overkill as the cause of Arctic megafaunal extinction is highly improbable^{6,8}. Interestingly, the only predator–prey relationship of note in the model is the significant positive effect of caribou on the distribution of wolves (Fig. 3), probably reflecting that the wolf is well-adapted to hunt caribou.

Fig. 3: Spatiotemporal models to retrodict the explanatory factors for animal distribution.

 figure3

The values indicate posterior parameter estimates of covariate effects for the models explaining the presence–absence of each animal’s eDNA. Only covariates included in the model with lowest Watanabe–Akaike information criterion are shown ([Methods](#)). The dots represent the posterior means, and the whiskers represent the posterior 2.5%

and 97.5% quantiles. Covariate effects of which the 2.5% and 97.5% quantiles are both negative (red), and effects of which the 2.5% and 97.5% quantiles are both positive (blue) are indicated.

To better gauge the explanatory power of environmental variables, we removed the effects of the presence of the eDNA of other animals (Extended Data Fig. 4a and [Methods](#)). The most consistent and widely prevalent patterns are the generally negative effects of plant NMDS1 and NMDS3—the first and third components of the non-metric multidimensional scaling (NMDS) of the vegetation compositions ([Methods](#))—on the presence of animal eDNA. Plant NMDS1 reflects an aquatic-to-terrestrial plant gradient, and plant NMDS3 reflects a graminoids-to-woody plant gradient, particularly sedges within the graminoids, which include species that are prominent in present-day wetland communities (Extended Data Fig. 4b). These two negative covariates apply to the distribution of both small (vole and hare) and large (horse and mammoth) mammals, indicating that a wetter environment with a high proportion of hygrophilous plants (that is, moisture-loving plants) was a key factor restricting animal distributions. The distribution of mammoths tends to be positively affected by plant NMDS2, which mainly reflects the proportion of woody plants (particularly shrubs and subshrubs) as opposed to herbaceous plants, whereas the reverse is true for horses (Fig. 3). We also found that horses are more sensitive to vegetation composition compared with other herbivores (Supplementary Information [13.3](#)). These findings support the hypothesis that horses were more restricted to a grassland environment and may also indicate a greater dietary flexibility in mammoths.

When each herbivore species is considered individually, the only climate variable that is consistently and positively associated with the presence of their eDNA is temperature seasonality (Fig. 3 and Extended Data Fig. 4a), consistent with expectations based on the continental climate associated with the Mammoth Steppe, a biome that is associated with extremely cold and dry conditions that supported abundant large mammal grazers¹⁹. The importance of climatic variables becomes more evident when herbivores are considered as a group. Precipitation—in greater amounts and seasonality—is a principal negative factor in the distribution of Arctic herbivores (Fig. 3), presumably because increased snow cover during winter limited the food access of grazers, and a wetter substrate is more difficult for them to exploit, in contrast to the firm and dry ground of the steppe–tundra^{7,19}.

Late-surviving megafauna

The timing of Arctic megafaunal extinction is a matter of debate, not least because last appearance dates (LADs) are repeatedly revised as younger fossils are reported^{5,6}, and also because discovering the remains of the last surviving individuals of a species is

extremely unlikely²². As a result, LADs systematically underestimate when a species disappeared, raising the possibility that populations persisted longer than is now evident^{4,23}. The extinction timing can be better gauged with eDNA; an animal leaves behind only a single skeleton, which is much less likely to be preserved, recovered and dated, when compared with the amount of DNA it continuously spread into the environment while it was alive.

Our data indicate that mammoths survived into the Early Holocene in present-day continental northeast Siberia until 7.3 ± 0.2 ka (seven samples younger than 10 ka) and North America until 8.6 ± 0.3 ka. Notably, we recovered mammoth DNA from a series of samples from the Taimyr Peninsula that indicate the presence of mammoths in north central Siberia as late as 3.9 ± 0.2 ka (site LUR10) (Fig. 4 and Supplementary Information 3.3). The survival of mammoths into the Holocene in these regions is probably attributable to the persistence of the steppe–tundra vegetation of dry- and cold-adapted herbaceous plants that was present during the Pleistocene (Fig. 2d). This vegetation would have provided a suitable habitat for mammoths and possibly other dryland grazers such as horses (Extended Data Fig. 5), which are known to have survived in the region until at least 5 ka (ref. 24). Together, these eDNA results indicate that mammoths survived much longer than previously thought—which, on the basis of skeletal remains, was around 10.7 ka on continental Eurasia²⁵ and around 13.8 ka in Alaska⁸. Given that humans occupied northern Eurasia sporadically from at least 40 ka and continuously after 16 ka (refs. 26,27), the late-surviving Taimyr mammoths potentially encountered and co-existed with humans over at least a 20-kyr interval, therefore giving no support to the human overkill (blitzkrieg) model that postulates the mammoth extinction occurred within centuries after the first human contact⁶.

Fig. 4: Mammoth distribution and mitochondrial haplotypes.

 **figure4**

A total of 78 mammoth mitochondrial genomes and 159 eDNA-identified mammoths (79 among them were assigned to mitochondrial haplotypes) are shown. Records of dated mammoth fossils⁶² are also plotted. All samples older than 26.5 ka were combined into the pre-LGM interval.

We also detected woolly rhinoceros DNA as late as 9.8 ± 0.2 ka in northeast Kolyma, horse DNA in Alaska and the Yukon as late as 7.9 ± 0.2 ka, and bison as late as 6.4 ± 0.6 ka in high-latitude localities of northeast Siberia (Extended Data Fig. 5). All of these instances represent substantially later LADs than fossil-based dates (that is, for woolly rhinoceros in Eurasia, about 14 ka (ref. 28); and for horses and steppe bison in Alaska, 12.5 ka (refs. 5,8)). Collectively, these findings highlight the value of eDNA in improving megafauna extinction chronologies.

Population diversity of megafauna

Megafaunal eDNA from across the Arctic also enables us to resolve population-level patterns, which is crucial for uncovering species-specific demographic and evolutionary responses to past climatic and environmental changes. We applied a method for phylogenetically assigning the identified eDNA to mitochondrial

haplogroups of mammoth and horse, the two most abundant species detected in our dataset ([Methods](#)).

A mammoth phylogeny composed of four previously described major mitochondrial clades (clade 1, including 1C and 1DE, and clades 2 and 3)²⁹ was reconstructed from 78 mammoth mitochondrial genomes. The recovered mammoth eDNA was then assigned to a best-fit node on the tree based on single-nucleotide polymorphism (SNP) support/conflict, enabling clade assignment for 79 eDNA samples (Extended Data Fig. [6](#)).

The mammoth haplogroups that we identified are consistent with those that were previously identified from fossil remains and have comparable biogeographical and biostratigraphic distributions (Fig. [4](#)). Overall, clade 3 was present mainly in Europe and northwest Siberia, whereas clade 2 occurred mostly in central and northeast Siberia. Clade 1 was widely scattered across North America and the Asian Arctic, with 1DE occurring throughout Siberia and 1C in North America. Temporally, clades 2 and 3 were the older lineages, and disappeared between 40 ka and 30 ka. Only clade 1 survived past the LGM, with the last 1C individual dating to 10.35 ka. Like the late-surviving mammoths on Wrangel Island³⁰, the late-surviving mammoths on mainland Siberia were also members of 1DE, the only clade detected to date that postdates the Early Holocene (that is, after 8.2 ka). However, despite belonging to the same clade, none of the mainland late-surviving populations is placed in the Wrangel Island haplogroup (Extended Data Fig. [6](#)). Furthermore, we note that two mammoth eDNA samples (cr5_11 and tm4_13) attach to the existing tree at the shared root of clades 2 and 3 (Extended Data Fig. [6](#)), with cr5_11 containing many sequence variants not found in previously sequenced samples (Supplementary Information [14.1.2](#)), suggesting that they represent a separate and previously unrecorded mitochondrial lineage. The distinctive mitochondrial genome haplogroups, together with the shrinking and increasingly isolated occurrences of mammoths (Fig. [4](#)), hint that Siberian mainland mammoths experienced a similar fate to those on Wrangel and St Paul Islands. However, whether the precise causes of their disappearance were the same^{4,30}, and whether the mainland mammoth also accumulated detrimental mutations consistent with genetic decline³¹, will require further data to resolve.

The reliability of our method was further corroborated on the horse phylogeny (Supplementary Information [14.2](#)). Successful assignment of ancient eDNA data to mitochondrial haplogroups, even when the DNA is highly degraded, highlights the potential for applying eDNA analysis to uncover population histories in regions in which fossils are rare or absent.

Concluding remarks

Controversy has persisted for decades over the nature of the Mammoth Steppe, a distinctive, now-vanished biome dominated by large mammal grazers^{1,19,32}. Some studies, emphasizing the abundance of grazers (and the absence of large browsers), suggest that broad swaths of the unglaciated Late Pleistocene Arctic were covered by an extensive steppe dominated by low-sward herbaceous plants that were well-suited for megafaunal grazers^{19,32}. Others, on the basis of pollen and plant macrofossil records, suggest that Arctic vegetation during this period was regionally diverse and included both tundra and steppe taxa^{3,33}. Our results suggest the nature of the Mammoth Steppe lies in between these two seemingly conflicting interpretations. Consistent with the view of the Mammoth Steppe as a biome of intercontinental extent, our data show that various regions of the Arctic supported a more homogenous vegetation cover before and during the LGM (Extended Data Fig. 1b, c). We also found evidence of an elevated and episodic turnover of plant taxa during the Late Pleistocene compared with during the Holocene (Fig. 2a), consistent with inferences about changeable vegetation types during the glacial age based on the network of palaeobotanical (and fossil insect) sites presently available^{3,12}. Jointly, our results suggest that the Mammoth Steppe was a regionally complex cryo-arid steppe, composed of forbs, graminoids and willow shrubs.

Our findings relating to the late survival of megafauna have important implications for the debate over the causes of Late Quaternary extinctions. Megafaunal survival into the Holocene indicates that, at least in certain parts of the Arctic and Subarctic, humans coexisted with these species for tens of thousands of years, which implies that human hunting was not an important factor in their extinction^{6,25}. Instead, our results suggest that their extinction came when the last pockets of the steppe–tundra vegetation finally disappeared, when the Arctic-wide paludification was brought on by warmer and wetter climates^{7,20}.

What we have mined from this substantial dataset does not exploit its full potential. For example, we detected DNA of Camelidae (most probably the Arctic camel³⁴) and *Panthera* (possibly the steppe lion). However, due to a lack of reference genomes for these species, we could not confirm these identifications. This constraint also applies to other species because our reference database—large as it is—is far from complete, despite our extensive sequencing efforts. With more species sequenced and new bioinformatics methods developed, this dataset can be reanalysed to explore more questions of Arctic biotic history.

Our study demonstrates how metagenomic analysis of eDNA extracted from ancient sediments can provide diverse insights, from detailed records of past flora and fauna to reconstructions of population histories and biotic interactions, to a greatly expanded spatiotemporal network of palaeoecological records. These advances are important in the context of continuous efforts to elucidate the past 50 kyr of Arctic biotic dynamics,

especially given that the coevolution of plant and animal species, and their responses to the past climatic changes across this vast region, have previously been challenging to address at this resolution and at this scale using classical palaeobotanical and palaeontological data.

Methods

Sampling, chronology and eDNA taphonomy

Sampling and subsampling methods are described in Supplementary Information 1. Sample ages were determined through conventional or accelerator mass spectrometer radiocarbon (^{14}C) as well as optically stimulated luminescence. In total, 631 radiocarbon ages and 81 optically stimulated luminescence dates were used. For sedimentary sections with multiple contiguous dates without stratigraphic inversions, age–depth models were built to calculate sedimentation rates and estimate the ages of undated samples within these sections. All radiocarbon ages are in calibrated years before present, calibrated using IntCal20 (ref. 35). Chronological information is provided in Supplementary Information 2 and Supplementary Data 1 and 2.

To determine whether DNA was *in situ*, control samples were obtained from modern surfaces, from water in adjacent rivers and lakes, and from stratigraphic layers bracketing the samples. Consistent with previous eDNA studies in the Arctic^{12,23,36}, we found no evidence of DNA leaching or redeposition in either terrestrial or lake sediment samples (Supplementary Information 5).

DNA extraction and sequencing

We tested the performance of different operations included in the widely used ancient eDNA extraction protocols^{36,37,38} and a variety of purification methods on different sediment sample types. On the basis of these tests, we developed two new eDNA-extraction protocols that were optimized for isolating and purifying eDNA from our sediment samples (Supplementary Information 6.1 and 6.2). The InhibitEx-based protocol was then applied for extracting DNA from all samples. DNA extracts were thereafter converted into sequencing libraries according to the standard protocol³⁹, and sequenced using Illumina platforms after quality controls (Supplementary Information 6.3). All DNA extractions and pre-index analyses were performed in the dedicated ancient DNA laboratories at the Centre for GeoGenetics, University of Copenhagen, according to established ancient DNA protocols⁴⁰.

PhyloNorway plant genome database construction

The PhyloNorway plant genome database was constructed by sequencing 1,541 Arctic and boreal plant specimens collected from herbaria. DNA was extracted from the selected specimens using a modified Macherey–Nagel Nucleospin 96 Plant II protocol. Two different library preparation protocols were applied depending on DNA yields. All of the libraries were then sequenced. Nuclear ribosomal DNA and chloroplast genome from each plant were assembled to evaluate the data quality. Whole-genome contigs for each plant were assembled and annotated as the final reference database. A list of plant species, herbarium information, DNA extraction, sequencing and database statistics are supplied in Supplementary Data [3](#). Data for three standard barcodes skimmed from this database were also used in ref. [41](#). Details are provided in Supplementary Information [7](#).

Taxonomic identification, authentication and quantification

We performed taxonomic classification by mapping reads against a comprehensive genomic database that was annotated with taxonomic information according to the principle of the Holi pipeline^{[36](#)}. Details of the composition of the reference database are provided in Supplementary Information [9.2.1](#).

All reads were first quality-controlled, and each read was then offered an equal chance to be aligned against all entries in the database after duplicate removal (Supplementary Information [9.1](#) and [9.2](#)). No limitation to specific taxonomic group, geography or environment was applied for the alignment. The lowest common ancestor of all of the hits with 100% similarity was assigned to each read that had been aligned to multiple taxa. The taxonomic coverage of different database compositions and their effects on taxa identification were evaluated using a *k*-mer-based method (Supplementary Information [9.2.2](#)). We found that using a proper reference database is important for eDNA metagenomics-based taxa identification, particularly for ancient datasets in which the DNA is highly fragmented. Even reference genome availability across taxa can improve the sensitivity and specificity of the identification by increasing the identified reads and correcting the misidentifications (Supplementary Information [9.2.5](#)). Taxa that were detected in the laboratory controls were combined into a list, and all of the listed taxa were subtracted from samples (Supplementary Information [9.3](#)). The resulting plant and animal taxonomic profiles were thereafter parsed for additional authentication using a series of conservative thresholds (Supplementary Information [9.4](#) and [9.6](#)), on the basis of an Arctic flora and faunal checklist (Supplementary Information [8](#)). Plant taxa that passed these filters all have Arctic or boreal distributions (Supplementary Information [9.4](#)). All eDNA reads aligned to an animal were further confirmed as exclusive alignments, by requiring perfect alignment to that animal, and no alignment to any other organisms when allowing for 1 or 2 mismatches (Supplementary Information [9.6.3](#)). The two extinct

animals—mammoth and woolly rhinoceros—were also confirmed by the DNA-damage patterns (Supplementary Information [9.6.2](#)).

Relative abundances for plants were estimated on the basis of the number of the assigned reads, by excluding the effects of DNA degradation in different samples, and eliminating the effects of the sequencing depth among different samples and the efficiency of the taxa-identification pipeline among different taxa (Supplementary Information [9.5](#)).

Vegetation diversity and dissimilarity

The Shannon diversity index was calculated according to the method in ref. [42](#). Plant morphological forms were assigned at the genus level on the basis of the plant trait database of eFloras (<http://www.efloras.org>). Beta-diversity (dissimilarity) between every two plant assemblages was calculated according to the method in ref. [43](#). For the pan-Arctic vegetation turnover (Fig. [2a](#)), plant genera identified in all samples in each 2,000-year interval were combined as an assemblage; beta-diversity between each two consecutive intervals was calculated. Regional vegetation turnover (Fig. [2c](#)) was calculated at 5,000-year intervals. NMDS ($k = 3$; Extended Data Fig. [1c](#)) was performed using the R package vegan^{[44](#)}, allowing 100,000 iterations of random starting to find the best convergent solution. Correlations between the abundance of each plant genus (or proportion of each morphological form) and the values of each of the three NMDS components (Extended Data Fig. [4b](#)) were assessed using the Pearson product-moment correlation and *t*-test ($P < 0.05$).

Comparison of eDNA shotgun metagenomics and metabarcoding

We applied two modules for comparing the metabarcoding and shotgun metagenomics in taxa identifications. (1) We conducted the two sequencing techniques in parallel on 14 DNA extracts to directly compare the retrieved taxonomic profiles. (2) We compared the floristic profiles reconstructed by this study and a previous metabarcoding study^{[12](#)} on 131 overlapping samples of the two datasets. The results show that metagenomics performed better on our samples in both captured floristic and faunal diversity. Details are provided in Supplementary Information [11](#).

Palaeoclimate panels and human distribution niche modelling

For the ice-core data from Greenland (Fig. [2a](#)), we rescaled the available $\delta^{18}\text{O}$ ratios (20-year slices) retrieved from NGRIP1 (ref. [45](#)), NGRIP2 (ref. [46](#)) and GISP2 (ref. [47](#)) to the range of the corresponding ratio of GRIP^{[48](#)}, for which there are valid values for all age slices, using the rescale function in the R package scales. The mean of the

available ratios for each time slice from the four datasets was calculated and used. Calcium concentrations were calculated from refs. [49,50](#) using the same method as for $\delta^{18}\text{O}$. Snow-accumulation rates were based on GISP2 (ref. [51](#)).

We also modelled monthly palaeoclimate anomalies at 1,000-year time steps using an emulator [52](#) and downscaled them onto a modern baseline climatology (CHELSA) [53](#) at a spatial resolution of 1°. From these data, we calculated four environmental variables—annual mean temperature, temperature seasonality, annual precipitation and precipitation seasonality—that were used to represent the climate for each of our eDNA sites. Details are provided in Supplementary Information [12.1](#).

We developed distribution models to map environmentally suitable conditions for Palaeolithic human occurrence in steps of 1,000 years from 5 ka to 31 ka and steps of 2,000 years from 32 ka to 47 ka. First, geo-references for human remains in the Arctic were collected and dates from ^{14}C calibrations inferred from two databases CARD2.0 (ref. [54](#)) and the Palaeolithic of Europe [55](#). These data were filtered for quality, resulting in a final set of 6,497 occurrences. From 32 ka to 47 ka, we calculated 2,000-year averages of the four environmental variables. We then generated five-algorithm ensemble models at each time step to characterize the climatic niche of Palaeolithic humans. We validated all of the models by assessing the area under the receiver operating characteristic curve (AUC) and true skill statistic; we also used model AUCs to generate weighted ensemble models at each time step. Finally, we projected the ensemble models into geographic space to map climatic suitability for humans, expressed as the potential presence or absence at each time step at each of the eDNA sites. Details are provided in Supplementary Information [12.2](#).

Spatiotemporal models for animal eDNA

We combined our animal eDNA data with the modelled climate variables, projected human occurrence and the NMDS ordinations of vegetation to examine the relative impacts of climate, human activity and vegetation on the geographical distributions of a selected group of Arctic mammals. We developed a method to spatiotemporally model animal eDNA presence, using these three sets of variables, while accounting for auto-correlation in time and space. The method uses a hierarchical Bayesian model that includes a spatiotemporal Gaussian random field, and was implemented in R-INLA [56,57](#). We used the Watanabe–Akaike information criterion to assess the model fit using different sets of covariates. Detailed methods are provided in Supplementary Information [13](#).

Mammoth and horse mitochondrial haplotyping

We placed eDNA mitochondrial reads for mammoth and horse into their respective mitochondrial reference phylogenies using recently developed software⁵⁸. We used existing variation to assign informative markers onto branches of a mitochondrial phylogeny, then determined the number of supporting and conflicting single-nucleotide polymorphisms for each eDNA sample on each branch of the tree to place the sample onto the most likely branch. Detailed methods are provided in Supplementary Information [14](#).

Statistics and data visualization

Changing trends are illustrated against time (Fig. [2a–c](#) and Extended Data Fig. [2b,c](#)) or distance (Extended Data Fig. [1b](#)) via the Loess Smooth (span = 4) function in the R package ggplot2 (ref. [59](#)), with original data points or confidence intervals (s.e.) shown when other curves are not obstructed. The heat maps showing the mean of a genus' proportions across all samples within an age interval were generated using the R package ComplexHeatmap⁶⁰. The mammoth phylogenetic tree was illustrated using ggtree, which is included in the R package ggplot2. The base map source for Fig. [1](#) and Extended Data Fig. [5](#) was Arctic SDI and, for Fig. [4](#), was the R package maptools.

Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

Data availability

Adapter-removed plant or animal eDNA data were deposited at EMBL-ENA under project accession [ERP127790](#). The raw data of PhyloNorway plant genome database are available at EMBL-ENA under project accession [PRJEB43865](#). Assembled plant genome contigs of the PhyloNorway database are available at DataverseNO⁶¹. NCBI databases are available at the NCBI ftp server (<https://ftp.ncbi.nlm.nih.gov>). The Canadian Archaeological Radiocarbon Database (CARD2.0) is available online (<https://www.canadianarchaeology.ca>). The Radiocarbon Palaeolithic Europe Database is available online (<https://ees.kuleuven.be/geography/projects/14c-palaeolithic>). All other data are provided in the [Supplementary Information](#) and Supplementary Data [1–9](#).

Code availability

Scripts are archived at GitHub (https://github.com/wyc661217/Arctic_eDNA_2021).

References

1. 1.
Binney, H. et al. Vegetation of Eurasia from the last glacial maximum to present: key biogeographic patterns. *Quat. Sci. Rev.* **157**, 80–97 (2017).
2. 2.
Clark, P. U. et al. The Last Glacial Maximum. *Science* **325**, 710–714 (2009).
3. 3.
Bigelow, N. H. Climate change and Arctic ecosystems: 1. Vegetation changes north of 55°N between the last glacial maximum, mid-Holocene, and present. *J. Geophys. Res.* **108**, <https://doi.org/10.1029/2002jd002558> (2003).
4. 4.
Graham, R. W. et al. Timing and causes of mid-Holocene mammoth extinction on St. Paul Island, Alaska. *Proc. Natl Acad. Sci. USA* **113**, 9310–9314 (2016).
5. 5.
Stuart, A. J. Late Quaternary megafaunal extinctions on the continents: a short review. *Geol. J.* **50**, 338–363 (2015).
6. 6.
Koch, P. L. & Barnosky, A. D. Late Quaternary extinctions: state of the debate. *Ann. Rev. Ecol. Evol. Syst.* **37**, 215–250 (2006).
7. 7.
Rabanus-Wallace, M. T. et al. Megafaunal isotopes reveal role of increased moisture on rangeland during late Pleistocene extinctions. *Nat. Ecol. Evol.* **1**, 0125 (2017).
8. 8.
Mann, D. H., Groves, P., Kunz, M. L., Reanier, R. E. & Gaglioti, B. V. Ice-age megafauna in Arctic Alaska: extinction, invasion, survival. *Quat. Sci. Rev.* **70**, 91–108 (2013).

9. 9.

Capo, E. et al. Lake sedimentary DNA research on past terrestrial and aquatic biodiversity: overview and recommendations. *Quaternary* **4**, <https://doi.org/10.3390/quat4010006> (2021).

10. 10.

Edwards, M. E. et al. Metabarcoding of modern soil DNA gives a highly local vegetation signal in Svalbard tundra. *Holocene* **28**, 2006–2016 (2018).

11. 11.

Hughes, P. D., Gibbard, P. L. & Ehlers, J. Timing of glaciation during the last glacial cycle: evaluating the concept of a global ‘Last Glacial Maximum’ (LGM). *Earth Sci. Rev.* **125**, 171–198 (2013).

12. 12.

Willerslev, E. et al. Fifty thousand years of Arctic vegetation and megafaunal diet. *Nature* **506**, 47–51 (2014).

13. 13.

Rasmussen, S. O. et al. A new Greenland ice core chronology for the last glacial termination. *J. Geophys. Res.* **111**, <https://doi.org/10.1029/2005jd006079> (2006).

14. 14.

Mangerud, J. The discovery of the Younger Dryas, and comments on the current meaning and usage of the term. *Boreas* **50**, 1–5 (2020).

15. 15.

Bauska, T. K. et al. Carbon isotopes characterize rapid changes in atmospheric carbon dioxide during the last deglaciation. *Proc. Natl Acad. Sci. USA* **113**, 3465–3470 (2016).

16. 16.

Wesser, S. D. & Armbruster, W. S. Species distribution controls across a forest-steppe transition: a causal model and experimental test. *Ecol. Monogr.* **61**, 323–342 (1991).

17. 17.

Rijal, D. P. et al. Sedimentary ancient DNA shows terrestrial plant richness continuously increased over the Holocene in northern Fennoscandia. *Sci. Adv.* **7**, eabf9557 (2021).

18. 18.

Birks, H. H. Aquatic macrophyte vegetation development in Kråkenes Lake, western Norway, during the late-glacial and early-Holocene. *J. Paleolimnol.* **23**, 7–19 (2000).

19. 19.

Guthrie, R. D. Origin and causes of the mammoth steppe: a story of cloud cover, woolly mammal tooth pits, buckles, and inside-out Beringia. *Quat. Sci. Rev.* **20**, 549–574 (2001).

20. 20.

Mann, D. H., Peteet, D. M., Reanier, R. E. & Kunz, M. L. Responses of an Arctic landscape to Lateglacial and early Holocene climatic changes: the importance of moisture. *Quat. Sci. Rev.* **21**, 997–1021 (2002).

21. 21.

Ritchie, M. in *Competition and Coexistence* (eds Sommer, U. & Worm, B.) 109–131 (Springer, 2002).

22. 22.

Signor, P. W., Lipps, J. H., Silver, L. & Schultz, P. in *Geological Implications of Impacts of Large Asteroids and Comets on the Earth* vol. 190 (eds Silver, L. T. & Schultz, P. H.) 291–296 (1982).

23. 23.

Haile, J. et al. Ancient DNA reveals late survival of mammoth and horse in interior Alaska. *Proc. Natl Acad. Sci. USA* **106**, 22352–22357 (2009).

24. 24.

Librado, P. et al. Tracking the origins of Yakutian horses and the genetic basis for their fast adaptation to subarctic environments. *Proc. Natl Acad. Sci. USA* **112**,

E6889–E6897 (2015).

25. 25.

Nikolskiy, P. A., Sulerzhitsky, L. D. & Pitulko, V. V. Last straw versus Blitzkrieg overkill: climate-driven changes in the Arctic Siberian mammoth population and the Late Pleistocene extinction problem. *Quat. Sci. Rev.* **30**, 2309–2328 (2011).

26. 26.

Pavlov, P., Svendsen, J. I. & Indrelid, S. Human presence in the European Arctic nearly 40,000 years ago. *Nature* **413**, 64–67 (2001).

27. 27.

Kuzmin, Y. V. & Keates, S. G. Siberia and neighboring regions in the Last Glacial Maximum: did people occupy northern Eurasia at that time? *Archaeol. Anthropol. Sci.* **10**, 111–124 (2016).

28. 28.

Stuart, A. J. & Lister, A. M. Extinction chronology of the woolly rhinoceros *Coelodonta antiquitatis* in the context of late Quaternary megafaunal extinctions in northern Eurasia. *Quat. Sci. Rev.* **51**, 1–17 (2012).

29. 29.

Chang, D. et al. The evolutionary and phylogeographic history of woolly mammoths: a comprehensive mitogenomic analysis. *Sci. Rep.* **7**, 44585 (2017).

30. 30.

Vartanyan, S. L., Arslanov, K. A., Karhu, J. A., Possnert, G. & Sulerzhitsky, L. D. Collection of radiocarbon dates on the mammoths (*Mammuthus primigenius*) and other genera of Wrangel Island, northeast Siberia, Russia. *Quat. Res.* **70**, 51–59 (2017).

31. 31.

Rogers, R. L. & Slatkin, M. Excess of genomic defects in a woolly mammoth on Wrangel island. *PLoS Genet.* **13**, e1006601 (2017).

32. 32.

Zimov, S. A., Zimov, N. S., Tikhonov, A. N. & Chapin, F. S. Mammoth steppe: a high-productivity phenomenon. *Quat. Sci. Rev.* **57**, 26–45 (2012).

33. 33.

Yurtsev, B. A. The Pleistocene “Tundra-Steppe” and the productivity paradox: the landscape approach. *Quat. Sci. Rev.* **20**, 165–174 (2001).

34. 34.

Rybczynski, N. et al. Mid-Pliocene warm-period deposits in the High Arctic yield insight into camel evolution. *Nat. Commun.* **4**, 1550 (2013).

35. 35.

Reimer, P. J. et al. The IntCal20 Northern Hemisphere radiocarbon age calibration curve (0–55 cal kBP). *Radiocarbon* **62**, 725–757 (2020).

36. 36.

Pedersen, M. W. et al. Postglacial viability and colonization in North America’s ice-free corridor. *Nature* **537**, 45–49 (2016).

37. 37.

Slon, V. et al. Neandertal and Denisovan DNA from Pleistocene sediments. *Science* **356**, 605–608 (2017).

38. 38.

Lorenz, M. G. & Wackernagel, W. Adsorption of DNA to sand and variable degradation rates of adsorbed DNA. *Appl. Environ. Microb.* **53**, 2948–2952 (1987).

39. 39.

Meyer, M. & Kircher, M. Illumina sequencing library preparation for highly multiplexed target capture and sequencing. *Cold Spring Harb. Protoc.* **2010**, pdb.prot5448 (2010).

40. 40.

Willerslev, E., Hansen, A. J. & Poinar, H. N. Isolation of nucleic acids and cultures from fossil ice and permafrost. *Trends Ecol. Evol.* **19**, 141–147 (2004).

41. 41.

Alsos, I. G. et al. The treasure vault can be opened: large-scale genome skimming works well using herbarium and silica gel dried material. *Plants* **9**, <https://doi.org/10.3390/plants9040432> (2020).

42. 42.

Hill, M. O. Diversity and evenness: a unifying notation and its consequences. *Ecology* **54**, 427–432 (1973).

43. 43.

Koleff, P., Gaston, K. J. & Lennon, J. J. Measuring beta diversity for presence-absence data. *J. Anim. Ecol.* **72**, 367–382 (2003).

44. 44.

Dixon, P. VEGAN, a package of R functions for community ecology. *J. Veg. Sci.* **14**, 927–930 (2003).

45. 45.

Grootes, P. M. & Stuiver, M. Oxygen 18/16 variability in Greenland snow and ice with 10^{-3} - to 10^5 -year time resolution. *J. Geophys. Res. Oceans* **102**, 26455–26470 (1997).

46. 46.

Andersen, K. K. et al. High-resolution record of Northern Hemisphere climate extending into the last interglacial period. *Nature* **431**, 147–151 (2004).

47. 47.

Stuiver, M. & Grootes, P. M. GISP2 oxygen isotope ratios. *Quat. Res.* **53**, 277–284 (2017).

48. 48.

Johnsen, S. J. et al. The $\delta^{18}\text{O}$ record along the Greenland Ice Core Project deep ice core and the problem of possible Eemian climatic instability. *J. Geophys. Res. Oceans* **102**, 26397–26410 (1997).

49. 49.

Fuhrer, K., Neftel, A., Anklin, M. & Maggi, V. Continuous measurements of hydrogen peroxide, formaldehyde, calcium and ammonium concentrations along the new grip ice core from summit, Central Greenland. *Atmos. Environ. A* **27**, 1873–1880 (1993).

50. 50.

Mayewski, P. A. et al. Major features and forcing of high-latitude northern hemisphere atmospheric circulation using a 110,000-year-long glaciochemical series. *J. Geophys. Res. Oceans* **102**, 26345–26366 (1997).

51. 51.

Alley, R. B. et al. Abrupt increase in Greenland snow accumulation at the end of the Younger Dryas event. *Nature* **362**, 527–529 (1993).

52. 52.

Holden, P. B. et al. PALEO-PGEM v1.0: a statistical emulator of Pliocene–Pleistocene climate. *Geosci. Model Dev.* **12**, 5137–5155 (2019).

53. 53.

Karger, D. N. et al. Climatologies at high resolution for the earth’s land surface areas. *Sci. Data* **4**, 170122 (2017).

54. 54.

Martindale, A. et al. Canadian Archaeological Radiocarbon Database (CARD 2.1) (Laboratory of Archaeology at the University of British Columbia, and the Canadian Museum of History, accessed 6 February 2020).

55. 55.

Vermeersch, P. M. Radiocarbon Palaeolithic Europe database: a regularly updated dataset of the radiometric data regarding the Palaeolithic of Europe, Siberia included. *Data Brief* **31**, 105793 (2020).

56. 56.

Rue, H., Martino, S. & Chopin, N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J. R. Stat. Soc. B* **71**, 319–392 (2009).

57. 57.

Lindgren, F. & Rue, H. Bayesian spatial modelling with R-INLA. *J. Stat. Softw.* **63**, 1–25 (2015).

58. 58.

Martiniano, R., De Sanctis, B., Hallast, P. & Durbin, R. Placing ancient DNA sequences into reference phylogenies. Preprint at <https://doi.org/10.1101/2020.12.19.423614> (2020).

59. 59.

Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2016).

60. 60.

Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).

61. 61.

Wang, Y. et al. *Supporting Data for: Late Quaternary Dynamics of Arctic Biota from Ancient Environmental Metagenomics* <https://dataverse.no/privateurl.xhtml?token=86979109-5605-43b5-b3fb-f470d85b114c> (2021).

62. 62.

Theodoridis, S. et al. Climate and genetic diversity change in mammals during the Late Quaternary. Preprint at <https://doi.org/10.1101/2021.03.05.433883> (2021).

Acknowledgements

We thank D. H. Mann for his detailed and constructive comments; and T. Ager, J. Austin, T. B. Brand, A. Cooper, S. Funder, M. T. P. Gilbert, T. Jørgensen, N. J. Korsgaard, S. Liu, M. Meldgaard, P. V. S. Olsen, M. L. Siggaard-Andersen, J. Stenderup, S. A. Woodroffe and staff at the GeoGenetics Sequencing Core and National Park Service-Western Arctic National Parklands for help and support. E.W. and D.J.M. thank the staff at St. John's College, Cambridge, for providing a stimulating environment for scientific discussion of the project. E.W. thanks Illumina for collaboration. The Lundbeck Foundation GeoGenetics Centre is supported by the

Carlsberg Foundation (CF18-0024), the Lundbeck Foundation (R302-2018-2155), the Novo Nordisk Foundation (NNF18SA0035006), the Wellcome Trust (UNS69906) and GRF EXC CRS Chair (44113220)—Cluster of Excellence. The PhyloNorway plant genome database is part of the Norwegian Barcode of Life Network (<https://www.norbol.org>) funded by the Research Council of Norway (226134/F50), the Norwegian Biodiversity Information Centre (14-14, 70184209) and The Arctic University Museum of Norway. Metabarcoding sequencing was funded by the Central Public-Interest Scientific Institution Basal Research Fund, CAFS (2017B001 and 2020A001). B.D.S. is supported by the Wellcome Trust programme in Mathematical Genomics and Medicine (WT220023); F.R. by a Villum Fonden Young Investigator award (no. 00025300); D.J.M. by the Quest Archaeological Research Fund; P.M. by the Swedish Research Council (VR); R.D. by the Wellcome Trust (WT207492); and A.R. by a Marie Skłodowska-Curie Actions Individual Fellowship (MSCA-IF, 703542) and the Research Council of Norway (KLIMAFORSK, 294929). L.O. has received funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (no. 681605); I.G.A. and Y.L. from the ERC under the European Union’s Horizon 2020 research and innovation programme (no. 819192). J.I.S. and J.M. are supported by the Research Council of Norway. P.B.H. and N.R.E. acknowledge NERC funding (grant NE/P015093/1). D.W.B. was supported by a Marie Skłodowska-Curie Actions Incoming International Fellowship (MCIIF-40974). T.S.K. is funded by a Carlsberg Foundation Young Researcher Fellowship (CF19-0712).

Author information

Author notes

1. These authors contributed equally: Yucheng Wang, Mikkel Winther Pedersen, Inger Greve Alsos

Affiliations

1. Department of Zoology, University of Cambridge, Cambridge, UK

Yucheng Wang, Bianca De Sanctis, Ana Prohaska, Daniel Money & Eske Willerslev

2. Lundbeck Foundation GeoGenetics Centre, GLOBE Institute, University of Copenhagen, Copenhagen, Denmark

Yucheng Wang, Mikkel Winther Pedersen, Fernando Racimo, Antonio Fernandez-Guerra, Alexandra Rouillard, Anthony H. Ruter, Hugh

McColl, Nicolaj Krog Larsen, James Haile, Lasse Vinner, Thorfinn Sand
Korneliussen, Jialu Cao, David J. Meltzer, Kurt H. Kjær & Eske Willerslev

3. The Arctic University Museum of Norway, UiT—The Arctic University of Norway, Tromsø, Norway

Inger Greve Alsos, Eric Coissac, Marie Kristine Føreid Merkel, Youri Lammers & Galina Gusarova

4. Department of Genetics, University of Cambridge, Cambridge, UK

Bianca De Sanctis & Richard Durbin

5. Université Grenoble Alpes, Université Savoie Mont Blanc, CNRS, LECA, Grenoble, France

Eric Coissac

6. Center for Macroecology, Evolution and Climate, GLOBE Institute, University of Copenhagen, Copenhagen, Denmark

Hannah Lois Owens, Carsten Rahbek & David Bravo Nogues

7. Department of Geosciences, UiT—The Arctic University of Norway, Tromsø, Norway

Alexandra Rouillard

8. Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), Gif-sur-Yvette, France

Adriana Alberti

9. Génomique Métabolique, Genoscope, Institut François Jacob, CEA, CNRS, Université Evry, Université Paris-Saclay, Evry, France

Adriana Alberti, France Denoeud & Patrick Wincker

10. Institute of Earth Sciences, St Petersburg State University, St Petersburg, Russia

Anna A. Cherezova & Grigory B. Fedorov

11. Arctic and Antarctic Research Institute, St Petersburg, Russia

Anna A. Cherezova & Grigory B. Fedorov

12. School of Geography and Environmental Science, University of Southampton,
Southampton, UK

Mary E. Edwards

13. Alaska Quaternary Center, University of Alaska Fairbanks, Fairbanks, AK, USA

Mary E. Edwards

14. Centre d'Anthropobiologie et de Génomique de Toulouse, Université Paul
Sabatier, Faculté de Médecine Purpan, Toulouse, France

Ludovic Orlando

15. National Research University, Higher School of Economics, Moscow, Russia

Thorfinn Sand Korneliussen

16. Department of Geography and Environment, University of Hawaii, Honolulu, HI,
USA

David W. Beilman

17. Department of Geosciences and Natural Resource Management, University of
Copenhagen, Copenhagen, Denmark

Anders A. Bjørk

18. Carlsberg Research Laboratory, Copenhagen, Denmark

Christoph Dockter & Birgitte Skadhauge

19. Center for Environmental Management of Military Lands, Colorado State
University, Fort Collins, CO, USA

Julie Esdale

20. Faculty of Biology, St Petersburg State University, St Petersburg, Russia

Galina Gusarova

21. Department of Glaciology and Climate, Geological Survey of Denmark and
Greenland, Copenhagen, Denmark

Kristian K. Kjeldsen

22. Department of Earth Science, University of Bergen, Bergen, Norway

Jan Mangerud & John Inge Svendsen

23. Bjerknes Centre for Climate Research, Bergen, Norway

Jan Mangerud & John Inge Svendsen

24. US National Park Service, Gates of the Arctic National Park and Preserve, Fairbanks, AK, USA

Jeffrey T. Rasic

25. Zoological Institute, , Russian Academy of Sciences, St Petersburg, Russia

Alexei Tikhonov

26. Resource and Environmental Research Center, Chinese Academy of Fishery Sciences, Beijing, China

Yingchun Xing

27. College of Plant Science, Jilin University, Changchun, China

Yubin Zhang

28. Department of Earth and Atmospheric Sciences, University of Alberta, Edmonton, Alberta, Canada

Duane G. Froese

29. Center for Global Mountain Biodiversity, GLOBE Institute, University of Copenhagen, Copenhagen, Denmark

Carsten Rahbek

30. School of Environment, Earth and Ecosystem Sciences, The Open University, Milton Keynes, UK

Philip B. Holden & Neil R. Edwards

31. Department of Anthropology, Southern Methodist University, Dallas, TX, USA

David J. Meltzer

32. Department of Geology, Quaternary Sciences, Lund University, Lund, Sweden

Per Möller

33. Wellcome Trust Sanger Institute, Wellcome Genome Campus, Cambridge, UK

Eske Willerslev

34. MARUM, University of Bremen, Bremen, Germany

Eske Willerslev

Contributions

B.D.S., F.R., A.P. and E.C. contributed equally to this work. E.W. and Y.W. initiated and led the study. E.W., Y.W., M.W.P. and K.H.K. designed the study. P.M., K.H.K., E.W., I.G.A., Y.W., J.I.S., D.G.F., J.M., A.R., M.E.E., J.H., D.W.B., A.A.B., J.E., K.K.K., J.T.R., A.T., A.A.C. and G.B.F. provided samples. Y.W., N.K.L. and A.R. revised the age of samples. Y.W. performed the DNA laboratory work. Y.W., B.D.S., A.F.-G., R.D., D.M. and T.S.K. did the bioinformatics and statistical analyses. I.G.A., Y.W. and A.R. evaluated the plant taxonomic profiles. For the PhyloNorway plant genome database: I.G.A. and E.C. designed the study and led the work; M.K.F.M. and A.A. did the laboratory work; P.W. led the sequencing; E.C., Y.W., Y.L. and F.D. did the bioinformatics analyses. H.L.O., D.B.N., Y.W., C.R. and D.J.M. modelled the human distribution. P.B.H. and N.R.E. modelled the climate. F.R. developed the animal spatiotemporal model. Y.W., M.W.P., D.J.M., I.G.A., E.W. and A.P. drafted the manuscript which all of the co-authors commented on.

Corresponding author

Correspondence to [Eske Willerslev](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Patricia Fall, Brian Huntley, Paul Valdes 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

Circum-Arctic plant abundance variations and vegetation similarity clustering. **a**, Pan-Arctic plant abundance heatmap. **b**, Spatial vegetation dissimilarities. Pairwise spatial beta-diversities (dissimilarities between every two plant communities) against the geographical distances between the two communities. **c**, Non-metric Multidimensional Scaling (NMDS, k=3) on vegetation communities.

Extended Data Fig. 2

Regional vegetation differences and climate changes. **a**, Vegetation similarities between each two regions. All identified plant genera across sites in a region during a time interval were merged as a plant assemblage. Spatial beta-diversity between every two assemblages were calculated and illustrated. NAt, North Atlantic; WcS, Northwest and central Siberia; ES, Northeast Siberia; Nam, North America. **b** and **c**, Modelled annual temperature and precipitation in different regions. Means of the modelled annual temperature and precipitation values ([Methods](#)) at all eDNA sampling sites within a region at each 1,000-year time step were calculated. The changing trends are illustrated.

Extended Data Fig. 3

Regional plant abundance heatmaps. Heatmaps show the relative abundances of the 40 abundant plant genera in each region.

Extended Data Fig. 4 Environmental explanatory factors for animal distribution, and plant NMDS components.

a, Posterior parameter estimates of covariate effects for the models explaining the presence/absence of each animal's eDNA using climate, human presence and plant NMDS as explanatory variables. The dots represent the posterior means, and the whiskers represent the posterior 2.5% and 97.5% quantiles. The colour red denotes covariate effects whose 2.5% and 97.5% quantiles are both negative, while the colour blue denotes covariate effects 2.5% and 97.5% quantiles are both positive. **b**, The plant

genera and morphological forms correlated to the 3 components of plant NMDS. Plant genera (morphological forms) are ranked by the p-value of t-test, and only the top 20 Pearson correlations are shown. The colour red denotes negative correlations while the colour blue denotes positive correlations.

Extended Data Fig. 5 Distribution chronologies for woolly rhinoceros, bison, horse, caribou, hare, wolf, and vole.

We combined our DNA results and the fossil records⁶² (available for woolly rhinoceros, bison, and caribou). Samples older than 26.5 ka were combined into Pre-LGM; samples younger than 4.2 ka were combined into the Late Holocene.

Extended Data Fig. 6 Mammoth mitochondrial phylogenetic tree.

For placed eDNA samples the number of supporting single-nucleotide polymorphisms is given in braces ([Methods](#)). IDs for the Wrangel Island population are underlined.

Supplementary information

Supplementary Information

See SI guide for full description of contents.

Reporting Summary

Supplementary Data 1 Site description

An Excel table covering the site metadata for all 74 eDNA sampling sites.

Supplementary Data 2 Sample metadata & age-depth models.xlsx

An Excel table covering the sample metadata for all 535 eDNA samples, and the age–depth models.

Supplementary Data 3 PhyloNorway metadata

An Excel table covering the metadata for all 1,541 PhyloNorway herbarium specimens.

Supplementary Data 4 Arctic flora and fauna checklist

An Excel table covering the assembled Arctic plant and animal checklist.

Supplementary Data 5 Plant abundance

An Excel table covering the reconstructed plant genera abundance.

Supplementary Data 6 Animal distribution & DNA damage

An Excel table covering the animal existence/absence matrix and the animal ancient DNA damage rate.

Supplementary Data 7 Human Presence/Absence

A PDF file showing the modelled environmentally suitable conditions for Palaeolithic human occurrence.

Supplementary Data 8 Human Presence/Absence for eDNA sites

An Excel table covering the modelled human existence/absence matrix for eDNA sampling sites.

Supplementary Data 9 Horse mitochondrial genomes

An Excel table covering the metadata for the horse mitochondrial genomes used for constructing the horse phylogeny.

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

Wang, Y., Pedersen, M.W., Alsos, I.G. *et al.* Late Quaternary dynamics of Arctic biota from ancient environmental genomics. *Nature* **600**, 86–92 (2021).
<https://doi.org/10.1038/s41586-021-04016-x>

- Received: 21 March 2021
- Accepted: 13 September 2021
- Published: 20 October 2021
- Issue Date: 02 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04016-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/s41586-021-04016-x>

- Article
- [Published: 10 November 2021](#)

Temporal transitions in the post-mitotic nervous system of *Caenorhabditis elegans*

- [HaoSheng Sun](#) [ORCID: orcid.org/0000-0003-3919-559X¹](#) nAff2 &
- [Oliver Hobert](#) [ORCID: orcid.org/0000-0002-7634-2854¹](#)

[Nature](#) volume **600**, pages 93–99 (2021)

- 3118 Accesses
- 25 Altmetric
- [Metrics details](#)

Subjects

- [Developmental biology](#)
- [Neuroscience](#)

Abstract

In most animals, the majority of the nervous system is generated and assembled into neuronal circuits during embryonic development¹. However, during juvenile stages, nervous systems still undergo extensive anatomical and functional changes to eventually form a fully mature nervous system by the adult stage^{2,3}. The molecular changes in post-mitotic neurons across post-embryonic development and the genetic programs that control these

temporal transitions are not well understood^{4,5}. Here, using the model system *Caenorhabditis elegans*, we comprehensively characterized the distinct functional states (locomotor behaviour) and the corresponding distinct molecular states (transcriptome) of the post-mitotic nervous system across temporal transitions during post-embryonic development. We observed pervasive, neuron-type-specific changes in gene expression, many of which are controlled by the developmental upregulation of the conserved heterochronic microRNA LIN-4 and the subsequent promotion of a mature neuronal transcriptional program through the repression of its target, the transcription factor *lin-14*. The functional relevance of these molecular transitions are exemplified by a temporally regulated target gene of the LIN-14 transcription factor, *nlp-45*, a neuropeptide-encoding gene, which we find is required for several distinct temporal transitions in exploratory activity during post-embryonic development. Our study provides insights into regulatory strategies that control neuron-type-specific gene batteries to modulate distinct behavioural states across temporal, sexual and environmental dimensions of post-embryonic development.

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 transitions in locomotor behaviour and the neuronal transcriptome.



Fig. 2: *lin-4* and *lin-14* control temporal transitions in exploratory behaviour and neuronal transcriptome.

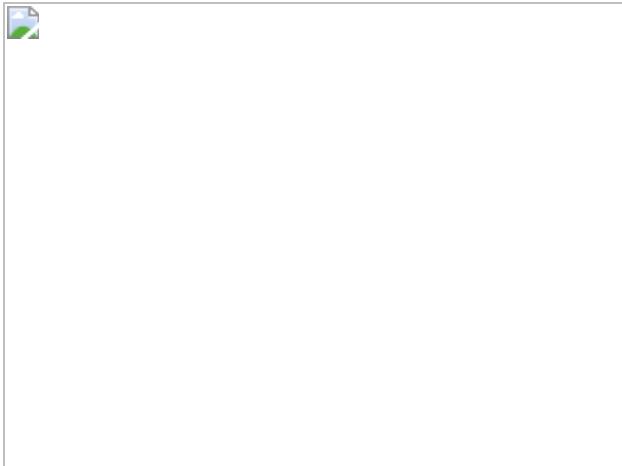


Fig. 3: *lin-4* and *lin-14* control developmental changes in exploratory behaviour through regulation of *nlp-45*.

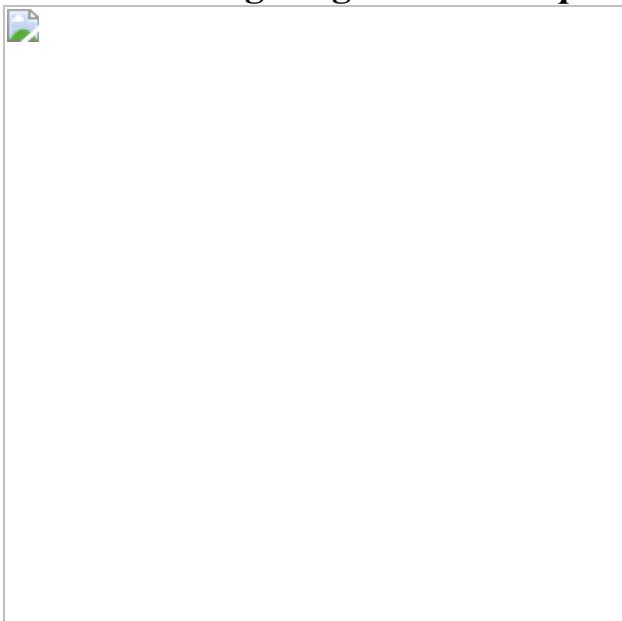
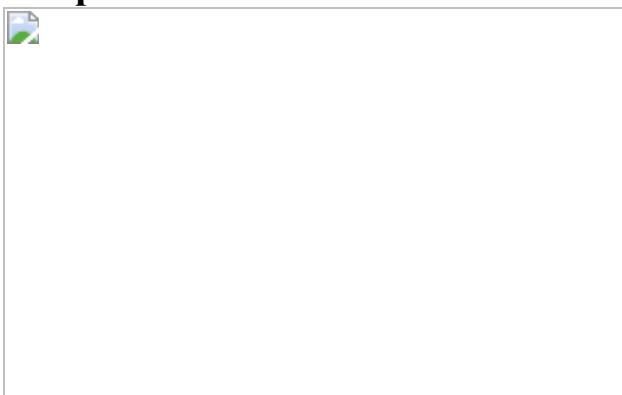


Fig. 4: Mechanism of *nlp-45* gene expression regulation across spatial, temporal, sexual and environmental dimensions of post-embryonic development.



Data availability

Raw and processed RNA-seq data are available under Gene Expression Omnibus (GEO) accession [GSE158274](#). Raw and processed ChIP-seq data are available under GEO accession [GSE181288](#). Raw microscope images of gene expression reporters with NeuroPAL for neuronal cell ID are available at <https://zenodo.org/communities/neuropal>.

References

1. 1.

Cadwell, C. R., Bhaduri, A., Mostajo-Radji, M. A., Keefe, M. G. & Nowakowski, T. J. Development and arealization of the cerebral cortex. *Neuron* **103**, 980–1004 (2019).

2. 2.

Gogtay, N. et al. Dynamic mapping of human cortical development during childhood through early adulthood. *Proc. Natl Acad. Sci. USA* **101**, 8174–8179 (2004).

3. 3.

Okaty, B. W., Miller, M. N., Sugino, K., Hempel, C. M. & Nelson, S. B. Transcriptional and electrophysiological maturation of neocortical fast-spiking GABAergic interneurons. *J. Neurosci.* **29**, 7040–7052 (2009).

4. 4.

Bakken, T. E. et al. A comprehensive transcriptional map of primate brain development. *Nature* **535**, 367–375 (2016).

5. 5.

Kang, H. J. et al. Spatio-temporal transcriptome of the human brain. *Nature* **478**, 483–489 (2011).

6. 6.

Spitzer, N. C. Electrical activity in early neuronal development. *Nature* **444**, 707–712 (2006).

7. 7.

Stroud, H. et al. An activity-mediated transition in transcription in early postnatal neurons. *Neuron* **107**, 874–890 (2020).

8. 8.

Cepko, C. L. The roles of intrinsic and extrinsic cues and bHLH genes in the determination of retinal cell fates. *Curr. Opin. Neurobiol.* **9**, 37–46 (1999).

9. 9.

Holguera, I. & Desplan, C. Neuronal specification in space and time. *Science* **362**, 176–180 (2018).

10. 10.

Miyares, R. L. & Lee, T. Temporal control of *Drosophila* central nervous system development. *Curr. Opin. Neurobiol.* **56**, 24–32 (2019).

11. 11.

Pearson, B. J. & Doe, C. Q. Specification of temporal identity in the developing nervous system. *Annu. Rev. Cell Dev. Biol.* **20**, 619–647 (2004).

12. 12.

Witvliet, D. et al. Connectomes across development reveal principles of brain maturation. *Nature* **596**, 257–261 (2021).

13. 13.

Yemini, E., Jucikas, T., Grundy, L. J., Brown, A. E. & Schafer, W. R. A database of *Caenorhabditis elegans* behavioral phenotypes. *Nat. Methods* **10**, 877–879 (2013).

14. 14.

Steiner, F. A., Talbert, P. B., Kasinathan, S., Deal, R. B. & Henikoff, S. Cell-type-specific nuclei purification from whole animals for genome-wide expression and chromatin profiling. *Genome Res.* **22**, 766–777 (2012).

15. 15.

Deal, R. B. & Henikoff, S. A simple method for gene expression and chromatin profiling of individual cell types within a tissue. *Dev. Cell* **18**, 1030–1040 (2010).

16. 16.

Ambros, V. & Horvitz, H. R. Heterochronic mutants of the nematode *Caenorhabditis elegans*. *Science* **226**, 409–416 (1984).

17. 17.

Rougvie, A. E. & Moss, E. G. Developmental transitions in *C. elegans* larval stages. *Curr. Top. Dev. Biol.* **105**, 153–180 (2013).

18. 18.

Feinbaum, R. & Ambros, V. The timing of *lin-4* RNA accumulation controls the timing of postembryonic developmental events in *Caenorhabditis elegans*. *Dev. Biol.* **210**, 87–95 (1999).

19. 19.

Ruvkun, G. & Giusto, J. The *Caenorhabditis elegans* heterochronic gene *lin-14* encodes a nuclear protein that forms a temporal developmental switch. *Nature* **338**, 313–319 (1989).

20. 20.

Marder, E. Neuromodulation of neuronal circuits: back to the future. *Neuron* **76**, 1–11 (2012).

21. 21.

Schoofs, L. & Beets, I. Neuropeptides control life-phase transitions. *Proc. Natl Acad. Sci. USA* **110**, 7973–7974 (2013).

22. 22.

Barrios, A., Ghosh, R., Fang, C., Emmons, S. W. & Barr, M. M. PDF-1 neuropeptide signaling modulates a neural circuit for mate-searching behavior in *C. elegans*. *Nat. Neurosci.* **15**, 1675–1682 (2012).

23. 23.

Lipton, J., Kleemann, G., Ghosh, R., Lints, R. & Emmons, S. W. Mate searching in *Caenorhabditis elegans*: a genetic model for sex drive in a simple invertebrate. *J. Neurosci.* **24**, 7427–7434 (2004).

24. 24.

Berkseth, M., Ikegami, K., Arur, S., Lieb, J. D. & Zarkower, D. TRA-1 ChIP-seq reveals regulators of sexual differentiation and multilevel feedback in nematode sex determination. *Proc. Natl Acad. Sci. USA* **110**, 16033–16038 (2013).

25. 25.

Oren-Suissa, M., Bayer, E. A. & Hobert, O. Sex-specific pruning of neuronal synapses in *Caenorhabditis elegans*. *Nature* **533**, 206–211 (2016).

26. 26.

Lee, K. & Portman, D. S. Neural sex modifies the function of a *C. elegans* sensory circuit. *Curr. Biol.* **17**, 1858–1863 (2007).

27. 27.

Fielenbach, N. & Antebi, A. *C. elegans* dauer formation and the molecular basis of plasticity. *Genes Dev.* **22**, 2149–2165 (2008).

28. 28.

Kumar, N. et al. Genome-wide endogenous DAF-16/FOXO recruitment dynamics during lowered insulin signalling in *C. elegans*. *Oncotarget* **6**, 41418–41433 (2015).

29. 29.

Hobert, O. Terminal selectors of neuronal identity. *Curr. Top. Dev. Biol.* **116**, 455–475 (2016).

30. 30.

Berghoff, E. G. et al. The Prop1-like homeobox gene *unc-42* specifies the identity of synaptically connected neurons. *eLife* **10**, e64903 (2021).

31. 31.

Yemini, E. et al. NeuroPAL: a multicolor atlas for whole-brain neuronal identification in *C. elegans*. *Cell* **184**, 272–288 (2021).

32. 32.

Tursun, B., Cochella, L., Carrera, I. & Hobert, O. A toolkit and robust pipeline for the generation of fosmid-based reporter genes in *C. elegans*. *PLoS ONE* **4**, e4625 (2009).

33. 33.

Dokshin, G. A., Ghanta, K. S., Piscopo, K. M. & Mello, C. C. Robust genome editing with short single-stranded and long, partially single-stranded dna donors in *Caenorhabditis elegans*. *Genetics* **210**, 781–787 (2018).

34. 34.

Dickinson, D. J., Pani, A. M., Heppert, J. K., Higgins, C. D. & Goldstein, B. Streamlined genome engineering with a self-excising drug selection cassette. *Genetics* **200**, 1035–1049 (2015).

35. 35.

Ahier, A. & Jarriault, S. Simultaneous expression of multiple proteins under a single promoter in *Caenorhabditis elegans* via a versatile 2A-based toolkit. *Genetics* **196**, 605–613 (2014).

36. 36.

Frokjaer-Jensen, C. et al. Random and targeted transgene insertion in *Caenorhabditis elegans* using a modified Mos1 transposon. *Nat. Methods* **11**, 529–534 (2014).

37. 37.

Zhang, F. et al. The LIM and POU homeobox genes *tx-3* and *unc-86* act as terminal selectors in distinct cholinergic and serotonergic neuron types. *Development* **141**, 422–435 (2014).

38. 38.

Brockie, P. J., Madsen, D. M., Zheng, Y., Mellem, J. & Maricq, A. V. Differential expression of glutamate receptor subunits in the nervous system of *Caenorhabditis elegans* and their regulation by the homeodomain protein UNC-42. *J. Neurosci.* **21**, 1510–1522 (2001).

39. 39.

Bhattacharya, A., Aghayeva, U., Berghoff, E. G. & Hobert, O. Plasticity of the electrical connectome of *C. elegans*. *Cell* **176**, 1174–1189 (2019).

40. 40.

Flavell, S. W. et al. Serotonin and the neuropeptide PDF initiate and extend opposing behavioral states in *C. elegans*. *Cell* **154**, 1023–1035 (2013).

41. 41.

Chalfie, M., Horvitz, H. R. & Sulston, J. E. Mutations that lead to reiterations in the cell lineages of *C. elegans*. *Cell* **24**, 59–69 (1981).

42. 42.

Liao, Y., Smyth, G. K. & Shi, W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res.* **47**, e47 (2019).

43. 43.

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).

44. 44.

Li, Y. et al. Establishment and maintenance of motor neuron identity via temporal modularity in terminal selector function. *eLife* **9**, e59464 (2020).

45. 45.

Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

46. 46.

Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

47. 47.

Feng, J., Liu, T., Qin, B., Zhang, Y. & Liu, X. S. Identifying ChIP-seq enrichment using MACS. *Nat. Protoc.* **7**, 1728–1740 (2012).

48. 48.

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).

49. 49.

Machanick, P. & Bailey, T. L. MEME-ChIP: motif analysis of large DNA datasets. *Bioinformatics* **27**, 1696–1697 (2011).

50. 50.

Ross-Innes, C. S. et al. Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature* **481**, 389–393 (2012).

51. 51.

Zhang, L., Ward, J. D., Cheng, Z. & Dernburg, A. F. The auxin-inducible degradation (AID) system enables versatile conditional protein depletion in *C. elegans*. *Development* **142**, 4374–4384 (2015).

52. 52.

Aghayeva, U., Bhattacharya, A. & Hobert, O. A panel of fluorophore-tagged *daf-16* alleles. *MicroPubl. Biol.* **2020**, <https://doi.org/10.17912/micropub.biology.000210> (2020).

53. 53.

Harris, D. T. & Horvitz, H. R. MAB-10/NAB acts with LIN-29/EGR to regulate terminal differentiation and the transition from larva to adult in *C. elegans*. *Development* **138**, 4051–4062 (2011).

54. 54.

Aeschimann, F. et al. LIN41 post-transcriptionally silences mRNAs by two distinct and position-dependent mechanisms. *Mol. Cell* **65**, 476–489 (2017).

55. 55.

Ramirez, F. et al. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–W165 (2016).

Acknowledgements

We thank Q. Chen and A. Bhattacharya for generating transgenic lines; D. Rahe for help with INTACT optimization; A. Romero and E. Yemini for help with worm tracking; L. Cochella, M. P. Hart, I. Beets and members of the Hobert laboratory for comments on the manuscript; and staff at Wormbase and the CGC for providing resources and reagents. This work was funded by the NIH K99 HD098371, National Research Council of Canada (Holmes Award) and by the Howard Hughes Medical Institute. Some strains were provided by the CGC, which is funded by the NIH Office of Research Infrastructure Program (P40 OLD010440).

Author information

Author notes

1. HaoSheng Sun

Present address: Department of Cell, Developmental and Integrative Biology, Heersink School of Medicine, University of Alabama at Birmingham, Birmingham, AL, USA

Affiliations

1. Department of Biological Sciences, Howard Hughes Medical Institute, Columbia University, New York, NY, USA

HaoSheng Sun & Oliver Hobert

Contributions

Conceptualization: H.S. and O.H. Methodology: H.S. Validation: H.S. Formal analysis: H.S. Investigation: H.S. Resources: H.S. Writing—original draft: H.S. Writing—review and editing: H.S. and O.H. Visualization: H.S. Supervision: H.S. and O.H. Funding acquisition: H.S. and O.H.

Corresponding authors

Correspondence to [HaoSheng Sun](#) or [Oliver Hobert](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Douglas Portman 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 Developmental transitions in neuronal transcriptome across post-embryonic life stages.

a, Schematic and experimental design for INTACT sample collection, protocol, and data analysis for neuronal transcriptome profiling across development. Representative images of the pan-neuronal INTACT strain as well as neuronal nuclei after immunoprecipitation (IP) are shown in bottom left panels. Representative tracks from IGV are shown for input and neuronal IP samples to demonstrate IP enrichment for pan-neuronally expressed gene, *rab-3*. **b–d**, Principal component analysis (PCA) of neuronal transcriptome across post-embryonic development was conducted using DESeq2 in R studio⁴³. Both batch as well as developmental stage were taken as factors for analysis. Each dot represents a replicate in the RNA-seq analysis. **b**, PC2 vs PC1. PC1 and 2 delineated the transitions between early larval (L1 and L2) stages and late larval (L4)/adult stages, and between all larval (L1 through L4) stages and the adult stage, respectively. **c**, PC3 vs PC1. PC3 largely accounted for variation as a result of batch. **d**, PC4 vs PC1. PC4 largely accounted for L2 specific changes. **e**, The numbers of significant ($P_{adj} < 0.01$) increases/decreases in gene expression are shown for each stage transition. **f**, Venn diagram of developmental changes in neuronal gene expression across different stage transitions, showing some overlaps but also distinct developmental changes across each stage transition. **g**, Gene ontology analysis of the 2639 developmentally regulated genes using the Enrichment Tool from Wormbase. **h**, Top: heat map of the 249 developmentally regulated genes between L1 and L2 stages across post-embryonic development. In addition to developmentally upregulated and downregulated genes, there was a small subset of genes that showed specific upregulation at the L2 stage. Bottom: gene ontology analysis of these genes using the Enrichment Tool from Wormbase. **i**, Top: heat map of the 448 developmentally regulated genes between L2 and L4 stages across post-embryonic development. Bottom: gene ontology analysis of these genes using the Enrichment Tool from Wormbase. **j**, Top: heat map of the 510 developmentally regulated genes between L4 and adult stages across post-embryonic development. Bottom: gene ontology analysis of these genes using the Enrichment Tool from Wormbase.

Extended Data Fig. 2 Temporal transitions in nervous system gene expression across *C. elegans* post-embryonic development.

For all panels, validations of developmentally regulated genes with expression reporters are shown. On top are the scattered dot plots (each point represents a single replicate, n in bracket) of the normalized read counts across all developmental stages from the neuronal INTACT/RNA-seq profiling. Mean +/- SEM are shown for each stage. Adjusted P values (P_{adj}), as calculated by DESeq2, for each developmental comparison are below. Below the RNA-seq read count plots are the schematics and allele names of the expression reporters. Below that are representative confocal microscopy images of the expression reporters across development. Specific regions/neurons are labelled with dotted lines: those labelled with black dotted lines/names are not altered developmentally while those labelled with green and red lines/names demonstrate, respectively, decreases and increases in expression across development. Those labelled with brown lines/names demonstrate both increases and decreases in expression in the same neurons across development. Red scale bars (10 μ m) are on the bottom right of all representative images. For all panels, L1 through L4 represent the first through the fourth larval stage animals. For **a**, additional quantification of fluorescence intensity is also shown at the bottom. Two-sided t-test with post hoc correction P values and n (in bracket) are shown. Additional details are included in Supplementary Table 6. **a**, Metabotropic glutamate receptor *gbb-2*, as validated with a translational fosmid reporter (*gfp*), shows expression in the same set of neurons across development³¹, although the intensity of expression is decreased across development, including that in the nerve ring as measured with fluorescence intensity. **b**, Gap junction molecule *inx-19*, as validated with a transcriptional fosmid reporter (*sl2::yfp::h2b*), loses expression in sixteen neuronal classes across development and gains expression in the AVA neuron upon entry into adulthood. **c**, Transcription cofactor *mab-10*, as validated with an endogenous translational reporter (*3xflag::mcherry*) engineered with CRISPR/Cas9, gains expression across the nervous system amongst other tissue during transition into the L4 stage that is further upregulated in adulthood. The RNA prediction matches well with previous RNA FISH analysis⁵³. The difference between RNA data and protein

reporter expression is consistent with previous characterized post-transcriptional regulation by LIN-41⁵⁴. **d**, Nuclear hormone receptor *daf-12*, as validated with an endogenous translational reporter (*gfp*) engineered with CRISPR-Cas9, shows increased expression broadly across the nervous system during early/mid-larval stage and then decreased expression upon transition into late larval/adult stage. **e**, Homeodomain transcription factor *tab-1*, as validated with an endogenous translational reporter (*gfp*) engineered with CRISPR-Cas9, loses expression in five classes of neuron during early larval development. **f**, Immunoglobulin-like domain molecule *oig-8*, as validated with a translational fosmid reporter (*gfp*), loses expression in two classes of neuron during early larval development.

Extended Data Fig. 3 Temporal transitions in nervous system gene expression across *C. elegans* post-embryonic development for the neuropeptide family.

a, Heat map of all neuronally enriched neuropeptides across post-embryonic development. Values were z-score normalized and plotted using pheatmap in R studio. Each row represents a single gene, and each column represents a single RNA-seq replicate. For **b–k**, validations of developmentally regulated genes with expression reporters are shown. On top (left for **b**) are the scattered dot plots (each point represents a single replicate, n in bracket) of the normalized read counts across all developmental stages from the neuronal INTACT/RNA-seq profiling. Mean +/- SEM are shown for each stage. Adjusted P values (P_{adj}), as calculated by DESeq2, for each developmental comparison are below. Below the RNA-seq read count plots are the schematics and allele names of the expression reporters. Below that (to the right for **b, g**) are representative confocal microscopy images of the expression reporters across development. Specific regions/neurons are labelled with dotted lines: those labelled with black dotted lines/names are not altered developmentally while those labelled with green and red lines/names demonstrate, respectively, decreases and increases in expression across development. Those labelled with brown lines/names demonstrate both increases and decreases in expression in the same neurons across development. **d** and **v** in brackets denotes dim and variable expression, respectively. Red scale bars

(10 μ m) are on the bottom right of all representative images. For all panels, L1 through L4 represent the first through the fourth larval stage animals. Additional details are included in Supplementary Table 6. **b**, Other than some remnant expression from embryo in early L1 animals, *nlp-45* gains expression progressively in a number of neurons across development. **c**, Neuropeptide-encoding gene *ins-6*, as validated with an endogenous reporter (*t2a::3xnlsgfp*) engineered with CRISPR-Cas9, gains expression in ASJ across the L2->L3 transition. Variable AWA expression (not shown) is detected in L3 animals onwards. **d**, Neuropeptide-encoding gene *ins-9*, as validated with an endogenous reporter (*t2a::3xnlsgfp*) engineered with CRISPR-Cas9, gains expression in a number of neurons as it enters L3/L4 stages and loses expression in a subset of these neurons upon entry into adulthood. Extremely variable and dim RVG neuron expression (VB1/2, not shown) is detected in L3/L4 animals. **e**, Neuropeptide-encoding gene *flp-26*, as validated with an endogenous reporter (*t2a::3xnlsgfp*) engineered with CRISPR-Cas9, loses expression in M4 and I6 pharyngeal neurons across early larval development and loses expression in PVQ as it enters late larval/adult stages. **f**, Neuropeptide-encoding gene *flp-28*, as validated with an endogenous reporter (*t2a::3xnlsgfp*) engineered with CRISPR-Cas9, gains expression in hermaphrodite specific neurons (VC, HSN) as it enters late larval/adult stages. Head/tail neurons do not appear to be developmentally regulated in this reporter (images not shown). **g**, Neuropeptide-encoding gene *nlp-50*, as validated with an endogenous reporter (*t2a::3xnlsgfp*) engineered with CRISPR-Cas9, gains expression in a number of neurons across development. It also loses expression in the RID neuron during early larval development. Additional quantification of RIP and BAG fluorescence intensity across development is also shown at the bottom. Two-sided t-test with post hoc correction P values and n in brackets are shown. **h**, Neuropeptide-encoding gene *flp-17*, as validated with a promoter fusion reporter (*gfp*), loses and gains expression in nine and two classes of neurons, respectively, across post-embryonic development. **i**, Neuropeptide-encoding gene *flp-14*, as validated with an endogenous reporter (*t2a::3xnlsgfp*) engineered with CRISPR-Cas9, loses and gains expression in one (LUA) and two (AVB and AVG) classes of neurons as it enters late larval/adult stages, respectively. **j**, Neuropeptide-encoding gene *nlp-13*, as validated with an endogenous reporter (*t2a::3xnlsgfp*) engineered with CRISPR-Cas9, gains expression in the ventral nerve cord

neurons (DA, VA, VD, VC) across development. Head/tail neurons do not appear to be developmentally regulated in this reporter (images not shown). **k**, Neuropeptide receptor gene *npr-17*, as validated with an endogenous translational reporter (*gfp*) engineered with CRISPR-Cas9, shows decreased and increased expression in four and one classes of neurons, respectively.

Extended Data Fig. 4 Temporal transitions in nervous system gene expression across *C. elegans* post-embryonic development for the receptor-type guanylyl cyclases (rGCs) family.

a, Heat map of all neuronally enriched receptor-type guanylyl cyclases (rGCs) across post-embryonic development. Values were z-score normalized and plotted using pheatmap in R studio. Each row represents a single gene, and each column represents a single RNA-seq replicate. For **b–d**, validations of developmentally regulated genes with expression reporters are shown. On top are the scattered dot plots (each point represents a single replicate, n in bracket) of the normalized read counts across all developmental stages from the neuronal INTACT/RNA-seq profiling. Mean +/- SEM are shown for each stage. Adjusted P values (P_{adj}), as calculated by DESeq2, for each developmental comparison are below. Below the RNA-seq read count plots are the schematics and allele names of the expression reporters. Below that are representative confocal microscopy images of the expression reporters across development. Specific regions/neurons labelled with dotted lines: those labelled with black dotted lines/names are not altered developmentally while those labelled with green and red lines/names demonstrate, respectively, decreases and increases in expression across development. Red scale bars (10 μ m) are on the bottom right of all representative images. For all panels, L1 through L4 represent the first through the fourth larval stage animals. Additional details are included in Supplementary Table 6. **b**, rGC *gcy-5*, as validated with a transcriptional fosmid reporter (*sl2::1xnls::gfp*), shows decreased expression in ASER and increased expression in RIG, as measured with fluorescence intensity, across development. Quantifications of fluorescence intensity are shown at the bottom. Two-sided t-test with post hoc correction P values and n in brackets are shown. **c**, rGC *gcy-21*, as validated with a promoter fusion reporter (*gfp*), loses expression in three neuronal

classes across the L1->L2 transition. **d**, rGC *gcy-12*, as validated with a promoter fusion reporter (*gfp*), gains expression in A and B type motor neurons across mid/late larval development.

Extended Data Fig. 5 Expression of *lin-14* is downregulated in the nervous system amongst other tissues across post-embryonic development in a *lin-4* dependent manner in hermaphrodite animals.

a, Expression of *lin-4* is turned on in the nervous system amongst other tissues during the L1->L2 transition. Schematic of the *lin-4* fosmid expression reagent, in which the *lin-4* pre-miRNA sequence was replaced with YFP, is shown on the left. Representative images of *lin-4* expression in L1 and L2 animals are shown on the right. Ellipse and polygon outline the anterior and lateral/ventral neuronal ganglions respectively. **b**, Expression of *lin-14* is downregulated in the nervous system amongst other tissues across post-embryonic development in a *lin-4* dependent manner in hermaphrodite animals. Schematic of the *lin-14* translational GFP allele, as engineered by CRISPR-Cas9, as well as *lin-14* gain-of-function (gf) alleles (*ot1087/ot1149/ot1150/ot1151*), where a 466bp region containing all seven *lin-4* repressive binding sites is deleted, are shown on the upper left. All 4 gf alleles represent the same molecular lesion but resulted from independent CRISPR-Cas9 mediated deletions. Quantification of LIN-14::GFP expression in the lateral/ventral ganglion is shown on the upper right. Two-sided t-test with post hoc correction *P* values and n (in bracket) are shown. On the bottom are the representative images of the *lin-14*::GFP allele across post-embryonic development in control, *lin-4(e912)*-null, and *lin-14(ot1087)* gf animals. Ellipse and polygon outline the anterior and lateral/ventral neuronal ganglions respectively. Red scale bars (10 μ m) are on the bottom right of all representative images. Expression of *lin-14* is still detectable in the adult hermaphrodite. *lin-14* expression is upregulated in *lin-4(e912)* null and *lin-14(ot1087)* gf animals across development. The incomplete juvenization of *lin-14* expression across development in the *lin-4* null mutant suggests additional mechanisms beyond *lin-4* that downregulate *lin-14* across post-embryonic development.

Extended Data Fig. 6 *lin-4* controls a subset of the developmentally regulated gene battery through direct repression of *lin-14*, and not *lin-28*.

a, *lin-4(e912)* null mutation juvenizes a subset of the adult control(Ctl) neuronal transcriptome to resemble that of the L1 Ctl neuronal transcriptome through direct de-repression of *lin-14* and not *lin-28*. Principal component analysis (PCA) of the neuronal transcriptomes across post-embryonic development and across genotypes was conducted using DESeq2 in R studio. Each dot represents a replicate in the RNA-seq analysis. **b**, Correlation between developmentally gene expression changes ($\log_2\text{FoldChange}[\text{Adult Expression/L1 Expression}]$) with gene expression changes in *lin-4(e912)* null mutation ($\log_2\text{FoldChange}[\text{Adult } \textit{lin-4(e912)} \text{ null expression /Adult control expression}]$, left), in *lin-14(ot1149)* gain-of-function (gf) mutation ($\log_2\text{FoldChange}[\text{Adult } \textit{lin-14(ot1149)} \text{ gf expression/Adult control expression}]$, middle), and in *lin-28(ot1154)* gf mutation ($\log_2\text{FoldChange}[\text{Adult } \textit{lin-28(ot1154)} \text{ gf expression/Adult control expression}]$, right). Linear regression was fitted through each set of data points, and the equation and R^2 values are shown for each. *lin-4* null and *lin-14* gf mutations accounted for some of the developmentally gene expression changes between L1 and adult, while *lin-28* gf mutation did not. **c**, Correlation between gene expression changes in *lin-4(e912)* null mutation ($\log_2\text{FoldChange}[\text{Adult } \textit{lin-4(e912)} \text{ null expression /Adult control expression}]$) with gene expression changes in *lin-14(ot1149)* gf mutation ($\log_2\text{FoldChange}[\text{Adult } \textit{lin-14(ot1149)} \text{ gf expression/Adult control expression}]$, left), and in *lin-28(ot1154)* gf mutation ($\log_2\text{FoldChange}[\text{Adult } \textit{lin-28(ot1154)} \text{ gf expression/Adult control expression}]$, right). Linear regression was fitted through each set of data points, and the equation and R^2 values are shown for each. *lin-14* gain of function mutation accounted for most of the changes observed in the *lin-4* null mutation, but *lin-28* gain of function mutation did not. **d**, Top Venn diagram showing that the difference between the adult *lin-4* null neuronal transcriptome compared to the adult control(Ctl) neuronal transcriptome is largely recapitulated in the transcriptome of adult *lin-14(ot1149)* gf mutants. Only one gene is significantly different in the adult *lin-28(ot1154)* gf vs adult control

comparison and does not overlap with the genes regulated by *lin-4*/*lin-14*. Bottom left Venn diagram showing that 48% of genes that demonstrate developmental upregulation (adult control(Ctl)>L1 Ctl) are juvenized in the adult *lin-4* null and/or *lin-14(ot1149)* gf animals. Bottom right Venn diagram showing that 33% of genes that demonstrate developmental downregulation (adult control(Ctl)<L1 Ctl) are juvenized in the adult *lin-4* null and/or *lin-14(ot1149)* gf animals. **e**, *lin-4* regulates *lin-28* mainly through *lin-14* and not through direct repression of *lin-28*. On the top left is the schematic of the regulation between *lin-4*, *lin-14* and *lin-28* based upon previous studies. On the top right is the schematic of the *lin-28* translational GFP allele, as engineered by CRISPR-Cas9, as well as *lin-28(ot1153/54/55)* gf and *lin-28(n719)* loss-of-function (lf) alleles. These three gf alleles represent the same molecular lesion (deletion of single *lin-4* binding site in the *lin-28* 3'UTR) but independent CRISPR-Cas9 mediated deletion events. On the bottom are the representative images of the *lin-28* translational GFP allele across post-embryonic development in control, *lin-4(e912)* null, *lin-14(ot1151)* gf, and *lin-28(ot1153)* gf animals. The signal is diffuse and cytoplasmic but can be observed in all tissues including the nervous system in early larval animals. Red scale bars (10 μ m) are on the bottom right of all representative images. LIN-28 is downregulated across post-embryonic development. *lin-4(e912)* null and *lin-14(ot1151)* gf mutations delay the downregulation of LIN-28, particularly during the L2->L3 transition, while the *lin-28(ot1153)* gf mutation does not. **f**, *lin-28* does not regulate developmental expression pattern of *nlp-45*. Representative images of the *nlp-45* expression reporter in control, *lin-28(n719)* lf and *lin-28(1155)* gf animals. Neurons that are labelled in black are not developmentally regulated while those that are labelled in red are developmentally upregulated. Red scale bars (10 μ m) are on the bottom right of all representative images.

Extended Data Fig. 7 Developmentally regulated genes not controlled by *lin-4*/*lin-14*.

a, *ins-6* developmental expression is not regulated by *lin-4* nor *lin-14*. Representative images of the *ins-6* expression reporter in control, *lin-4* null and *lin-14* null animals are shown on the left while quantification of the *ins-6* expression in ASJ is shown on the right (number of animals for each

condition is shown in red brackets). **b**, *ins-9* developmental expression is not regulated by *lin-4*. Representative images of the *ins-9* expression reporter in control and *lin-4* null animals are shown on the left while quantification of the *ins-9* expression is shown on the right (number of animals for each condition is shown in red brackets). **c**, *inx-2* developmental expression is not regulated by *lin-4*. Representative images of the *inx-2* expression reporter in control and *lin-4* null animals are shown on the left while quantification of neuronal *inx-2* expression is shown on the right (number of animals for each condition is shown in red brackets). **d**, *flp-26* developmental expression is not regulated by *lin-4*. Representative images of the *flp-26* expression reporter in control and *lin-4* null animals are shown. Control images for A-D are taken from Extended Data Fig. 3. Red scale bars (10 μ m) are on the bottom right of all representative images.

Extended Data Fig. 8 Decreased LIN-14 binding across L1->L2 transition controls *lin-4*/*lin-14* dependent transcriptomic changes.

a, Decreased LIN-14 binding at promoters of target genes during the L1 -> L2 transition. Normalized datasets at each developmental stage against respective inputs are generated using the bamCompare function of deepTools2⁵⁵. LIN-14 enrichment around TSS +/- 2kb is then computed using the computeMatrix function and plotted using the plotHeatmap function of deepTools2⁵⁵. **b**, LIN-14 ChIP-seq peak distribution and motif across L1 and L2 animals. The ChIP-seq peak distribution on the left is plotted using ChIPseeker⁴⁸. The consensus binding motif on the right is obtained using MEME-ChIP⁴⁹. **c**, Amalgamation of different methods of assessing differential LIN-14 binding across the L1->L2 transition. Outer three Venn Diagrams on the top are the 32990 LIN-14 peaks in L1 animals within 3kb of the TSS of 18245 genes (green, Supplementary Table 8), the 7240 LIN-14 peaks in L2 animals within 3kb of the TSS of 5818 genes (red, Supplementary Table 9), and the 5267 differential L1 vs L2 LIN-14 binding within 3kb of the TSS of 4532 genes (blue, as determined by DiffBind, Supplementary Table 10), each overlapped with the 7974 neuronally enriched genes. These neuronally enriched peaks in L1 (green), peaks in L2 (red) and differential L1 vs L2 binding (blue) were overlapped

in the middle Venn Diagram to result in the 3466 genes (Supplementary Table 11) that show differential (mostly decreased) L1 vs L2 binding. These 3466 genes overlapped with 339 (60%) of the *lin-4/lin-14* controlled developmentally upregulated genes and 138 (49%) of the *lin-4/lin-14* controlled developmentally downregulated genes. **d**, The 339 genes that showed developmental upregulation (LIN-14 as a repressor) had increased LIN-14 peak enrichment and number of LIN-14 peaks within 3kb of TSS as compared to the 138 genes that showed developmental downregulation (LIN-14 as an activator) or the 2989 genes that show no developmental regulation. Box Whisker plots (10-90 percentile) are used, and two-sided t-test with post hoc correction *P* values are shown for the comparisons. **e**, Heat map of the developmental up-/down-regulated genes that also had differential LIN-14 binding from **c** across development (L1 and adult) and genotype (Control[Ctl] and *lin-4/lin-14* mutants). Values were z-score normalized and plotted using pheatmap in R studio. Each row represents a single gene, and each column represents a single RNA-seq replicate. The rows are clustered according to gene expression patterns.

Extended Data Fig. 9 *lin-4/lin-14* control developmental regulation of neuropeptide-encoding and receptor-type guanylyl cyclase (rGC) genes.

For all panels, the plots with normalized RNA-seq read counts for the L1/adult control/heterochronic mutant animals are plotted, with each point representing a replicate (*n* in bracket) and the mean +/- SEM shown for each stage, on the upper left. Adjusted *P* values (P_{adj}), as calculated by DESeq2, for each comparison are below. *lin-4* null/*lin-14* gain-of-function mutations juvenize the expression of all four genes. LIN-14 ChIP-seq binding at each gene is shown on the upper right. All samples from their respective experimental conditions are merged for their respective tracks in IGV. Decrease in LIN-14 binding during the L1->L2 transition is observed for all 4 genes. Representative confocal images under control and mutant conditions are shown in the bottom left. d in brackets denote dim expression, while v in brackets denote variable expression. Red scale bars (10 μ m) are on the bottom right of all representative images. Quantification of the images are shown on the bottom right (number of animals for each

condition is shown in red brackets). **a**, Juvenization of *nlp-45* expression by *lin-4/lin-14* across development as predicted from the neuronal INTACT/RNA-seq profiling and LIN-14 ChIP-seq binding at the *nlp-45* gene. Representative images are shown in Fig. 3a. **b**, *fhp-28* expression is gained in hermaphrodite specific neurons (HSN, VC) during transition into late larval/adult stages while the same developmental upregulation is not observed in *lin-4* null mutants. *fhp-28* expression in head/tail neurons is not regulated in *lin-4* null animals (images not shown). Control images are taken from Extended Data Fig. 3f. **c**, *fhp-14* loses and gains expression in one (LUA, outlined in green) and two (AVB and AVG, outlined in red) classes of neurons as it enters late larval/adult stages, respectively. *fhp-14* expression is de-repressed in the AVB and AVG neurons in L1 *lin-14* null animals while *fhp-14* expression is repressed in the LUA neurons in L1 *lin-14* null animals. Consistently, *fhp-14* expression in the AVB and AVG neurons are repressed in adult *lin-4* null animals while *fhp-14* expression is increased in the LUA neuron as compared to adult control animals. *fhp-14* expression is also weakly de-repressed in two classes of neurons that express *fhp-14* in the adult male (RMDD/V, outlined in blue) and one additional class of neuron (AVE, outlined in purple) in L1 *lin-14* null animals. Control images are taken from Extended Data Fig. 3i. **d**, *gcy-12* gains expression in the A and B type motor neurons across mid/late larval development. *gcy-12* expression in the A and B type motor neurons is de-repressed in *lin-14* null L1 animals while *gcy-12* expression in the A and B type motor neurons is repressed in the *lin-4* null adult animals as compared to respective control animals. *gcy-12* expression in head/tail neurons is not regulated in *lin-4* null animals (images not shown). **e**, *nlp-13* gains expression in the ventral nerve cord neurons (DA, VA, VD, VC) across development. *nlp-13* expression in the DA neurons is de-repressed in *lin-14* null L1 animals as compared to L1 control animals. Control L1 and adult images are taken from Extended Data Fig. 3j.

Extended Data Fig. 10 Regulation and function of *nlp-45* across temporal, sexual, and environmental dimensions of post-embryonic development.

a, Schematic of *nlp-45* deletion mutants and exploratory assay. **b**, Increased dwelling during L1->L2 transition is partially juvenized in *lin-4(e912)*

animals. Mean +/- SEM and n (in bracket) are shown for each condition, and each point of the scatter dot plot represents a single animal. Wilcoxon rank-sum tests and false-discovery rate q values for each comparison shown below. **c**, Schematic showing *lin-4/lin-14* regulation of *nlp-45* to alter exploratory behaviour during the L1>L2 transition. **d**, *nlp-45* deletion mutants do not significantly affect the food leaving behaviours of juvenile males/hermaphrodites. Values were plotted as mean +/- SEM of three independent experiments (n=6 animals per independent experiments). **e**, Leaving assay for adult hermaphrodite in *nlp-45* and *pdf-1* mutant animals. Values were plotted as mean +/- SEM of three independent experiments (n= 8 animals per independent experiments). Statistical analysis (two-sided t-test with post hoc correction) is only shown for the comparison to respective controls (colour coded respectively). **f**, Leaving assay for adult male in *nlp-45* and *pdf-1* mutant animals. Values plotted as mean +/- SEM of three independent experiments (n= 8 animals per independent experiments). Statistical analysis (two-sided t-test with post hoc correction) is only shown for the comparison to respective controls (colour coded respectively). **g**, Schematic of the opposing role of *nlp-45* and *pdf-1* on male food leaving behaviour. **h**, Developmental expression of *lin-14* in hermaphrodites and males. Representative images for the *lin-14(cc2841[lin-14::gfp])* reporter are shown across all developmental stages for both sexes. LIN-14 expression was similarly downregulated in both sexes at early larval stages, its expression in the late larval and particularly in the adult stage was significantly more reduced in the male nervous system compared to that of the hermaphrodite. Ellipse and polygon outline anterior and lateral/ventral neuronal ganglia. Representative images for hermaphrodite are re-used here from Extended Data Fig. [5b](#) for direct side by side comparison with male animals across development. **i**, Pan-neuronal depletion of sex determination master regulator TRA-1, through overexpression of FEM-3, decreases nervous system LIN-14 expression in adult hermaphrodites to mimic that of adult males. Representative microscope images, shown above, are overexposed in comparison to previous *lin-14* reporter images to better show the dim expression in adult males and FEM-3 overexpressed hermaphrodites. The quantifications of head neuron numbers across the three conditions are shown below. The mean +/- SEM and n (in bracket) shown for each condition, and each point of the scatter dot plot represents a single animal. Two-sided t-test with post

hoc correction *P* values are shown for each comparison. **j**, Pan-neuronal depletion of TRA-1, through overexpression of FEM-3, masculinizes *nlp-45* expression in adult hermaphrodite VNC. Representative images are shown on the left. Quantifications of VNC neuron numbers are shown on the right. The mean +/- SEM and n (in bracket) are shown for each condition, and each point of the scatter dot plot represents a single animal. Two-sided t-test with post hoc correction *P* values are shown for each comparison. **k**, Sexually dimorphic expression of *flp-14* in adult hermaphrodites and males. In addition to stronger *flp-14* expression in the AVB neuron (red) as compared to adult hermaphrodites, adult males gain *flp-14* expression in the RMDD/V and SIA neurons (blue). Adult hermaphrodite images are re-used from Extended Data Fig. [3i](#). **l**, Sexually dimorphic expression of *flp-28* in adult hermaphrodites and males. As compared to adult hermaphrodites, adult males gain *flp-28* expression in the IL1D/V, URB, and AIM neurons (blue). Adult hermaphrodite images are re-used from Extended Data Fig. [3f](#). **m**, Sexually dimorphic expression of *nlp-13* in adult hermaphrodites and males. In addition to stronger *nlp-13* expression in the VD neurons (red) as compared to adult hermaphrodites, adult males gain *nlp-13* expression in the male specific CA/CP motor neurons (blue) and lose *nlp-13* expression in the DA, VA and hermaphrodite specific VC motor neurons. Adult hermaphrodite images are re-used from Extended Data Fig. [3j](#). **n**, Pan-neuronal degradation of DAF-16 in auxin-treated dauers leads to pan-neuronal de-repression of *lin-14* in dauer animals. Representative images are on the left while binary quantifications of pan-neuronal expression are shown on the right (number of animals for each condition is shown in red brackets). Animals were grown (from embryo onward) on NGM plates supplemented with OP50 and 4mM auxin in EtOH (indole-3 acetic acid, IAA, Alfa Aesar) at 25 °C to degrade DAF-16 pan-neuronally and to induce dauer formation. As controls, plates were supplemented with the solvent EtOH instead of auxin. Additional control animals without pan-neuronal TIR-1 expression grown on EtOH and auxin were also included for comparison. **o**, Pan-neuronal degradation of DAF-16 in auxin-treated dauers leads to a loss or reduced *nlp-45* expression in several neuronal classes. Representative images are shown in Fig. [4f](#). Binary quantifications are shown for the SAAD/V neurons (number of animals for each condition is shown in red brackets) while fluorescence quantifications are shown for the RMED/V, RMEL/R, IL1D/V, RIV, RIM,

and ADE neurons. The mean +/- SEM and n (in bracket) are shown for each condition, and each point of the scatter dot plot represents a single animal. Two-sided t-test with post hoc correction P values are shown for each comparison. **p**, Expression of *flp-14* in dauer animals as compared to L3 animals. Upon entry into dauer, similar to expression pattern in the adult hermaphrodite/male, animals gain *flp-14* expression in the AVB neurons (red) and lose *flp-14* expression in the LUA neurons (green). Additional dauer specific *flp-14* expression is gained in the ASE neurons (orange). Expression in the PVR neurons is also lost in dauer animals. L3 hermaphrodite images are re-used from Extended Data Fig. [3i](#). **q**, Expression of *flp-28* in dauer animals as compared to L3 animals. Upon entry into dauer, animals gain *flp-28* expression in the ALA, AVH, AIN, ADA, and DVA neurons (labelled in orange) and lose *flp-28* expression in the DVC neurons. L3 hermaphrodite images are re-used from Extended Data Fig. [3f](#). d and v in brackets denote dim and variable expression, respectively. Scale bars = 10 μ m.

[Extended Data Fig 11 Terminal selector provides spatial specificity to *nlp-45* expression pattern.](#)

a, Regulation of *nlp-45* by cell specific regulator, *unc-42*. Representative images of adult *unc-42(e419)* hermaphrodite and male animals are shown on the left while binary quantifications of *nlp-45* expression in the RMDD/V, RIV, SAAD/V and AVA neurons are shown on the right (number of animals for each condition is shown in red brackets). *nlp-45* expression was lost in all *unc-42* expressing neurons with the exception of AVA in *unc-42* mutant animals. **b**, Regulation of *nlp-45* by cell specific regulator, *unc-3*. Representative images of adult *unc-3(e151)* hermaphrodite and male animals are shown on the left while binary quantifications of *nlp-45* expression in the SAAD/V and AVA neurons are shown in the middle (number of animals for each condition is shown in red brackets). The quantification of ventral nerve cord (VNC) motor neurons is shown on the right. n (in bracket) and two-sided t-test with post hoc correction P values are shown. Each point of the scatter dot plot represents a single animal. *nlp-45* expression was lost in *unc-3* expressing head neurons (i.e. SAAD/V, AVA) while severely affected in the VNC in *unc-3* mutant animals. **c**, Regulation of *nlp-45* by cell specific regulator, *ast-1*. Representative images

of adult *ast-1(ot417)* hypomorph hermaphrodite and male animals are shown on the left while quantifications of *nlp-45* expression in the CEPD and CEPV neurons are shown on the right (number of animals for each condition is shown in red brackets). *nlp-45* expression was severely affected in the CEPD neurons and slightly affected in the CEPV neurons in the *ast-1* mutant animals. **d**, Regulation of *nlp-45* by cell specific regulator, *ceh-8*. Representative images of adult *ceh-8(gk116531)* hermaphrodite and male animals are shown on the left while binary quantifications of *nlp-45* expression in the RIA and AVE neurons are shown on the right (number of animals for each condition is shown in red brackets). *nlp-45* expression is lost in the RIA neurons and ectopically gained in the AVE neurons in *ceh-8* mutant animals. **e**, Terminal selector (*unc-42*) does not regulate heterochronic pathway (*lin-14*). On the left are representative images of L1 *lin-14* translational GFP allele worms in control and *unc-42(e419)* backgrounds. On the right is the quantification of fluorescence intensity in the lateral/ventral ganglion of control vs *unc-42(e419)* L1 animals. The mean +/- SEM and n (in bracket) are shown for each condition, and each point of the scatter dot plot represents a single animal. **f**, Heterochronic pathway (*lin-14*) does not regulate terminal selector (*unc-42*). Representative images of L1 *unc-42* translational GFP allele worms in control and *lin-14(mal35)* null backgrounds are shown. **g**, On top are representative images of the *nlp-45* expression reporter in control, *lin-14(0)*, and *lin-14(0); unc-42(0)* L1 animals. On the bottom are the binary quantifications of *nlp-45* expression in different neuronal subtypes in control, *lin-14(0)*, and *lin-14(0); unc-42(0)* L1 animals (number of animals for each condition is shown in red brackets). *nlp-45* showed precocious expression in *lin-14, unc-42* double mutants, similar to the *lin-14* null mutant alone, except in the neurons (i.e. RMDD/V, SAAD/V) where *unc-42* acts as a terminal selector.

Supplementary information

Supplementary Information

Supplementary Discussion and References, Supplementary Fig. 1 and the legends for Supplementary Tables 1–13.

Reporting Summary

Peer Review File

Supplementary Table 1

Worm tracking summary. The mean, s.d. and s.e.m. for each of the 726 parameters at each developmental stage are shown. The statistics ($q(\text{Wilcoxon})$) for each parameter and each comparison are also shown.

Supplementary Table 2

Neuronally enriched genes. Comparison of neuronal nuclei IP samples to input (total nuclei) samples was conducted using DESeq2 (ref. ⁴²) in R Studio to determine the enrichment; 7,974 genes have a $\log_2[\text{fold change}] > 0$ (neuronally enriched over input) and an adjusted $P (P_{\text{adj}}) < 0.05$. The genes are sorted by P_{adj} from the smallest to largest. baseMean is the average normalized read counts of all IP and input samples. $\log_2[\text{fold change}]$ was calculated using the formula $\log_2[\text{average read counts of IP samples}/\text{average read counts of input samples}]$.

Supplementary Table 3

Normalized read counts of all 7,974 neuronally enriched genes across post-embryonic development. Raw read count for the 7,974 neuronally enriched genes are extracted for the neuronal IP samples and adjusted for library size. These are then used in DESeq2 (ref. ⁴²) to conduct comparisons between developmental stages, as shown in Supplementary Tables 4 and 5. The final read counts displayed are as a result of normalization performed using DESeq2. There are 4, 5, 6 and 7 replicates shown for the L1, L2, L4 and adult stages, respectively.

Supplementary Table 4

Comparison of all temporal transitions among the 7,974 neuronally enriched genes across post-embryonic development. Pairwise comparisons between all stages are conducted using DESeq2 (ref. ⁴²) as described in Supplementary Table 2. The $\log_2[\text{fold change}]$ and P_{adj} are shown for each comparison. All 7,974 neuronally enriched genes are shown regardless of P_{adj} . The table is sorted alphabetically by gene name.

Supplementary Table 5

Normalized read counts of 2,639 developmentally regulated genes. The same as Supplementary Table 3, but with a subset of only 2,639 genes with $P_{\text{adj}} < 0.01$ in any pairwise comparisons between developmental stages.

Supplementary Table 6

Developmental change summary. Expression patterns, at the single-neuron resolution, for all of the validated reporters from Extended Data Figs. 2–4 are displayed here. The neurons that show expression for each gene at each post-embryonic developmental stage (L1 to adult) are listed. Neurons in black do not show altered developmental expression patterns, whereas those labelled in green and red show decreases and increases in gene expression across development, respectively. Those labelled in brown demonstrate both increases and decreases in expression across development. Dim expression (d), moderate level of expression (m), and variable expression (v) are indicated.

Supplementary Table 7

Genes that are juvenized in adult heterochronic mutants. Pairwise comparisons between adult *lin-4(e912)* null, *lin-14(ot1149)* gain-of-function (gf) and *lin-28(ot1154)* gain-of-function mutant neuronal IPs are conducted against adult control neuronal IP data, using only the read counts of the 7,974 neuronally enriched genes. The $\log_2[\text{fold change}]$ and P_{adj} are shown for each comparison. All 7,974 neuronally enriched genes are shown regardless of P_{adj} . The table is sorted alphabetically by gene name. $\log_2[\text{fold}$

change] is calculated using the formula \log_2 [average read counts of mutant adult samples/average read counts of control adult samples].

Supplementary Table 8

LIN14 ChIP peaks in L1 animals. The five replicates from L1 animals are merged and the peaks were called using MACS2 (ref. ⁴⁶). The ChIP-seq peaks were annotated to the nearest gene using ChIPseeker (ref. ⁴⁷). Only the peaks within 3 kb of the closest genes were retained for this table. The peaks are arranged alphabetically by the name of the closest associated gene.

Supplementary Table 9

LIN14 ChIP peaks in L2 animals. The five replicates from L2 animals are merged and the peaks were called using MACS2 (ref. ⁴⁶). The ChIP-seq peaks were annotated to the nearest gene using ChIPseeker⁴⁷. Only the peaks within 3 kb of the closest genes were retained for this table. The peaks were arranged alphabetically by the name of the closest associated gene.

Supplementary Table 10

Differential LIN14 ChIP peaks between L1 and L2 animals as determined using DiffBind. Differential binding analysis between L1 and L2 samples was performed using Diffbind⁴⁹. All of the differential binding sites are annotated and assigned to the nearest gene using ChIPseeker⁴⁷. Only the peaks within 3 kb of the closest genes were retained for this table. The peaks were arranged alphabetically by the name of the closest associated gene.

Supplementary Table 11

List of genes that exhibit differential LIN14 binding during the L1-to-L2 transition. This table contains the 3,466 genes that show differential (mostly decreased) L1 versus L2 binding. This list was obtained using an

amalgamation of different methods of assessing differential LIN-14 binding across the L1-to-L2 transition, as described in Extended Data Fig. 8c. The genes are sorted alphabetically.

Supplementary Table 12

Sexually dimorphic and dauer-induced expression patterns of LIN-4/LIN-14 controlled developmentally regulated genes. The sexually dimorphic and dauer-induced expression patterns of five genes (*nlp-45*, *flp-28*, *flp-14*, *gcy-12* and *nlp-13*) that demonstrate LIN-4/LIN-14 controlled developmental regulation are summarized in this table (for representative images, see Fig. 3a and Extended Data Fig. 9). For each gene, the neurons of which the developmental regulation is controlled by LIN-4/LIN-14, are listed in the second column. Those labelled in green and red show decreases and increases in gene expression across post-embryonic development in hermaphrodite animals, respectively. Those labelled in blue gain expression in adult male animals, whereas those in orange show additional expression after entry into dauer. Those labelled in purple are regulated by LIN-4/LIN-14 but are not observed in any conditions tested in control animals. The neurons in which LIN-14 acts as a repressor or an activator are listed in the third and fourth column, respectively. The last three columns show the neurons for each gene of which expression is regulated across post-embryonic development in hermaphrodites, between adult males and hermaphrodites, and between dauer and the comparable L3 mid-larval animals, respectively. For most *nlp-45* expressing neurons and several *flp-14* expressing neurons, the sexually dimorphic and dauer-specific expression patterns are consistent with regulation through LIN-14. For *flp-28*, the hash (#) symbol indicates the observation that these neurons, which show increased expression in male and dauer animals, are consistent with the model that these expression patterns are regulated through *lin-14*. However, due to the proximity of *flp-28* and *lin-14* locus, the *flp-28* expression pattern could not be examined in *lin-14* null animals to determine the full battery of neurons that demonstrate *lin-14* regulated *flp-28* expression. Asterisks indicate that sexually dimorphic or dauer-specific expression patterns of *flp-14* for these neurons are not regulated through *lin-14*. The single plus (+) symbols indicate that no obvious sexually dimorphic and dauer-specific expression patterns were observed for *gcy-12*. This could

be due to the type of reporter (promoter fusion) and the diffuse (cytoplasmic) signal of the reporter. Additional regulatory mechanisms that antagonize *lin-14* regulation of *gcy-12* could also explain the lack of sexually dimorphic or dauer specific expression pattern of *gcy-12* in the ventral nerve cord motor neurons or elsewhere. The double-plus (++) symbols indicate that, for *nlp-13*, the sexually dimorphic and dauer-specific (no difference) expression patterns are not consistent with *lin-14* regulation alone. This suggests that there are additional *lin-14*-independent regulatory mechanisms in male and dauer animals. Together, these data suggest that for some genes, regulation through *lin-14* can largely explain the developmental, sexually dimorphic and dauer-specific expression patterns. However, for other genes regulated by LIN-14, there are probably additional *lin-14*-independent mechanisms that either synergize with or antagonize the regulation by LIN-14, leading to the complex expression patterns observed across temporal, sexual and environmental dimensions of post-embryonic development.

Supplementary Table 13

Strains Used in this manuscript.

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Sun, H., Hobert, O. Temporal transitions in the post-mitotic nervous system of *Caenorhabditis elegans*. *Nature* **600**, 93–99 (2021).
<https://doi.org/10.1038/s41586-021-04071-4>

- Received: 09 October 2020
- Accepted: 29 September 2021

- Published: 10 November 2021
- Issue Date: 02 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04071-4>

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/s41586-021-04071-4>

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

- Article
- Open Access
- [Published: 06 October 2021](#)

Thalamic circuits for independent control of prefrontal signal and noise

- [Arghya Mukherjee ORCID: orcid.org/0000-0002-3341-4408^{1,2}](#)
- [Norman H. Lam^{1,2}](#)
- [Ralf D. Wimmer^{1,2}](#) &
- [Michael M. Halassa ORCID: orcid.org/0000-0003-1386-0336^{1,2}](#)

Nature volume 600, pages 100–104 (2021)

- 8590 Accesses
- 148 Altmetric
- [Metrics details](#)

Subjects

- [Cognitive control](#)
- [Neurophysiology](#)

Abstract

Interactions between the mediodorsal thalamus and the prefrontal cortex are critical for cognition. Studies in humans indicate that these interactions may resolve uncertainty in decision-making¹, but the precise mechanisms are unknown. Here we identify two distinct mediodorsal projections to the prefrontal cortex that have complementary mechanistic roles in decision-making under uncertainty. Specifically, we found that a dopamine receptor (D2)-expressing projection amplifies prefrontal signals when task inputs are sparse and a kainate receptor (GRIK4) expressing-projection suppresses prefrontal noise when task inputs are dense but conflicting. Collectively, our data suggest that there are distinct brain mechanisms for handling uncertainty due to low signals versus uncertainty due to high noise, and provide a

mechanistic entry point for correcting decision-making abnormalities in disorders that have a prominent prefrontal component^{2,3,4,5,6}.

[Download PDF](#)

Main

Activating the mediodorsal thalamus (MD) in mice has two distinct effects on neural activity in the prefrontal cortex (PFC): amplification of local functional connectivity⁷ and suppression of spike rates⁸. To ask what the circuit mechanisms of these effects were, we first replicated them (Extended Data Fig. 1, Fig. 1), and confirmed that they were specific to this associative thalamocortical loop⁸ (Extended Data Fig. 1a–g). We noted that, in contrast to sensory systems⁹, the MD heavily targets cortical interneurons that are positive for vasoactive intestinal peptide (VIP^+)¹⁰ and known to be important for input amplification through disinhibition¹¹. Therefore, we asked whether MD-dependent amplification of PFC functional connectivity ([Methods](#)) was dependent on VIP^+ interneurons. Indeed, suppressing VIP^+ interneurons eliminated this MD effect (Fig. 1a–c, Extended Data Fig. 1i,j), but, notably, did not affect basal cortical spike rates (Fig. 1d,e, Extended Data Fig. 1h,k). The two MD effects were uncorrelated, suggesting mechanistic independence (Fig. 1f). As such, we hypothesized that the MD may contain two projections that differentially target prefrontal interneurons for independent control over input amplification and suppression (Fig. 1g). We also hypothesized that suppression may be carried out by parvalbumin positive (PV^+) prefrontal interneurons, as several studies have shown robust activation of these interneurons by the MD¹². The specific subdivision of the PFC that we focus on in this study is the prelimbic cortex (PL).

Fig. 1: MD amplifies functional PFC connectivity through cortical VIP^+ interneurons.

 **figure1**

a, Top, cartoon of experimental set-up. Bottom left, stabilized step function opsin (SSFO) MD expression. Bottom middle, PL tetrode location (white arrow); eNHpR3.0-expressing VIP⁺ neuron (inset). Bottom right, somatic ChR2 in contralateral PL. MDl, lateral MD; MDc, central MD; MDm, medial MD. Scale bars, 200 μ m; 20 μ m (inset). **b**, Top, putative excitatory PL neuron showing amplification of its response to intracortical stimulation (blue tick) when the MD is activated. Bottom, this effect is eliminated by inactivation of local VIP⁺ interneurons. **c**, Population quantification of effect in **b** ($n = 151$ excitatory PL neurons). **d**, Baseline spike rate suppression in another PL neuron after MD activation is unaffected by VIP⁺ interneuron inactivation. **e**, Population quantification of effect in **d** ($n = 373$ neurons). **f**, The two MD effects are uncorrelated ($n = 151$ neurons). **g**, Hypothesized MD projections target prefrontal interneurons for independent control over amplification and suppression of cortical activity patterns. Data from 4 VIP-cre mice. For **c**, **e** Mann-Whitney U for comparisons to baseline; Wilcoxon signed-rank for group comparisons. All statistical tests are two-tailed. For box plots in **c**, **e**, boundaries, 25–75th percentiles; midline, median; whiskers, minimum–maximum.

[Source data](#).

Identifying genetic MD cell types

To investigate the anatomical circuitry for these two hypothesized MD projections, we performed monosynaptic rabies tracing from either VIP⁺ or PV⁺ interneurons in the PL. Notably, we found that MD neurons projecting to these two prefrontal interneuron types occupied distinct anatomical territories (Fig. 2a–d, Extended Data Fig. 3a–e). Given genetic variation across the mediolateral axis of the thalamus¹³, we reasoned that these thalamic projections may be genetically distinct.

Fig. 2: Two MD circuits for amplification and suppression of PFC activity.

 figure2

a, Prefrontal PV⁺ and VIP⁺ input mapping. **b**, MD neurons targeting VIP⁺ (left) and PV⁺ (right) interneurons occupy distinct MDI domains. Scale bars, 200 μm. **c**, Group summary for location of VIP- and PV-projecting MDI neurons ($n = 73$ VIP-projecting (7 mice) and $n = 117$ PV-projecting (4 mice)). **d**, KNN clustering and representational similarity analysis show robust separation. **e**, Labelling MD_{D2} and MD_{GRIK4} neurons using the corresponding Cre lines. **f**, MD_{D2} and MD_{GRIK4} neurons also occupy distinct anatomical locations. Scale bars, 200 μm. **g**, Group summary of MD_{D2} and MD_{GRIK4} neurons in MDI ($n = 177$ MD_{D2} neurons and 194 MD_{GRIK4} neurons from 3 mice each). **h**, MD_{D2} and MD_{GRIK4} locations show additional high representational similarity to VIP- and PV-projecting neurons, respectively. **i, j**, mGRASP labelling shows higher innervation of VIP⁺ neurons by MD_{D2} (**i**; $n = 21$ neurons from 3 D2-cre, $n = 25$ neurons from 3 GRIK4-cre mice, respectively; Kolmogorov–Smirnov) and higher

innervation of PV⁺ neurons by MD_{GRIK4} (**j**; $n = 27$ neurons from 3 D2-cre, $n = 32$ neurons from 3 GRIK4-cre mice, respectively; Kolmogorov–Smirnov). Scale bars, 3 μm. **k**, Hypothesized circuit. **l**, Selective MD cell-type activation set-up. **m**, MD_{D2} but not MD_{GRIK4} amplify functional PL connectivity ($n = 100$ and $n = 68$ PL neural responses from 3 mice each for MD_{D2} and MD_{GRIK4}, respectively; left to right: MD_{D2} $P = 1.0 \times 10^{-5}$ for all; MD_{GRIK4} $P = 0.0599, 0.0789, 0.0575, 0.1311$ (NS) for laser powers displayed; Mann-Whitney U , compared to baseline). **n**, MD_{GRIK4} but not MD_{D2} suppress PL neural spike rates ($n = 1,257$ and $n = 697$ putative excitatory PL neurons from 3 mice each; MD_{D2} $P = 0.184, 0.605, 0.579, 0.739$ (NS); MD_{GRIK4} $P = 0.298, P = 0.067$, * $P = 0.033$, *** $P = 1.61 \times 10^{-5}$, respectively, for laser powers displayed; Mann-Whitney U compared to baseline). All statistical tests are two-tailed. Box plot parameters as in Fig. 1. Data are mean ± s.e.m. for **m**, **n**.

Source data.

A recent study in the paraventricular thalamus showed that the dopamine type 2 receptor (D2) distinguishes two subpopulations of functionally distinct thalamic projection neurons¹⁴. The MD is known to receive dopaminergic inputs¹⁵, and we found the mRNA expression of the D2 receptor to be reminiscent of the anatomical location of VIP-projecting MD neurons (Extended Data Fig. 2). Indeed, MD labelling in the D2-cre mice indicated that the D2⁺ genotype and the VIP-projecting one may be related (Fig. 2e,f).

To identify a potential genotype for the PV-projecting MD neurons, we took note of a previous study that used the kainate receptor, GRIK4, to label a population of MD neurons that drove feedforward inhibition¹⁶, mediated through PV⁺ interneurons¹⁷. MD labelling in GRIK4-cre mice (Fig. 2e) resulted in a pattern resembling the PV projection identified earlier (Fig. 2f). In addition, the D2⁺ (MD_{D2}) and GRIK4⁺ (MD_{GRIK4}) neurons could reliably be anatomically separated across mice, in a manner similar to VIP- and PV-projecting neurons (Fig. 2g,h). We confirmed the correspondence between this anatomical connectivity phenotype and its genetic identity through cross-validation (Fig. 2h, Extended Data Fig. 3f).

To further test the hypothesis that the two thalamic projections map onto distinct genetic identities, we used a synaptic labelling technique: mammalian GFP reconstitution across synaptic partners (mGRASP)¹⁸ (Extended Data Fig. 4). After Cre-dependent presynaptic mGRASP injection into the MD of either D2-cre or GRIK4-cre mice, and pan-neuronal postsynaptic mGRASP in the PL, we quantified the pattern of synaptic innervation of PV⁺ and VIP⁺ neurons (identified by immunohistochemistry) across these preparations (Extended Data Fig. 4a–c). We

found that the MD_{D2} population preferentially targeted VIP⁺ neurons (Fig. 2*i*, Extended Data Fig. 4*d*), whereas the MD_{GRIK4} preferentially targeted PV⁺ neurons (Fig. 2*j*, Extended Data Fig. 4*e*). This finding was independently supported by synaptophysin-based labelling; MD_{D2} neurons preferentially targeted layer I (Extended Data Fig. 4*f,g*), where VIP⁺ neurons are known to be enriched¹⁰. Collectively, these experiments indicated that PL amplification and suppression may indeed be under the control of genetically distinct MD thalamic cell types (Fig. 2*k*).

To directly test this idea, we selectively activated either MD_{D2} or MD_{GRIK4} neurons (Fig. 2*l*), and found that the former—but not the latter—resulted in amplification of functional PL connectivity (Fig. 2*m*, Extended Data Fig. 3*g*), whereas the opposite dependence was true for spike rate suppression (Fig. 2*n*, Extended Data Fig. 3*h,i*). These experiments definitively show that the MD contains two genetically distinct projections that independently control PL activation and suppression. Of note, MD_{D2} and MD_{GRIK4} segregation was independently verified using a viral strategy (Extended Data Fig. 3*j-l*), and GRIK4 immunohistochemistry allowed us to estimate their overlap to be 5–15% (Extended Data Fig. 3*m-o*).

To test whether these two cell types differentially engage in MD–PL-dependent behaviour, we leveraged an attentional control task that can distinguish MD enhancement of PL activity to maintain attentional control signals^{7,19}, and MD suppression of PL activity to enable task switching^{12,20}, or engagement (Extended Data Fig. 5*a-f*). Selective MD_{D2} inactivation diminished the former, whereas selective MD_{GRIK4} inactivation diminished the latter (Extended Data Fig. 5*g-n*, Supplementary Note 1).

Mouse MD tracks task uncertainty

We next turned our attention to asking whether these cell types contribute to a domain that may generalize to human cognition. Studies of the human brain have indicated a particular role for the MD in decision-making that scales with the degree of task input uncertainty^{1,21}. Therefore, we reasoned that incorporating input uncertainty into a task requiring MD–PL interaction in mice could achieve this goal. Consequently, we modified an attentional control task⁷ by parametrizing its cueing component (Fig. 3*a*, [Methods](#)). Specifically, on each trial a mouse was presented with a sequence of sixteen sound pulses (different mixtures of high-pass (HP, ‘attend to audition’), low-pass (LP, ‘attend to vision’) or broadband white noise (‘blank’)). Target selection was tied to the rule with the highest number of corresponding pulses on each sequence, and the ambiguity was mainly controlled by the conflict between HP and LP pulses. Multiple controls were incorporated to ensure that mice were adopting an attentional selection

strategy (Extended Data Fig. [6a](#)) and that they interpreted broadband white noise pulses as ‘blanks’ (Extended Data Fig. [6b](#)). Finally, regression analysis further validated that the mice were weighing evidence in the early and late halves of the cueing period equivalently (Extended Data Fig. [6c](#)).

Fig. 3: Task input uncertainty engages the mouse MD.



a, Task schematic (see text). **b**, PL inactivation (blue) diminishes performance regardless of uncertainty level ($n = 17$ sessions, 5 mice; $***P < 7.72 \times 10^{-4}$; chi-squared). **c**, Putative excitatory PL neurons showing responses during the cueing period, with the later neuron showing selectivity to the attentional choice ($*P =$

0.0157; Mann-Whitney *U*). **d**, Population decoding (top) and mutual information (bottom) show choice selectivity ($n = 1,112$ neurons from 7 mice). Both measures are modulated by uncertainty. **e**, MD inactivation (yellow) diminishes performance as a function of uncertainty ($n = 56$ sessions, 6 mice; *** $P < 1.43 \times 10^{-13}$; chi-squared). **f**, Example task-relevant MD neurons, one conflict-preferring and one conflict-non-preferring, both exhibiting little choice selectivity. **g**, Relative fraction of the MD neural functional types ($n = 2,669$ neurons from 7 mice). All statistical tests are two-tailed. Box plot parameters as in Fig. 1. Data are mean \pm s.e.m. for **b**, **e** and mean \pm 95% confidence interval (CI) for **d**.

Source data.

Inactivation of the PL during the cueing period diminished performance regardless of cueing uncertainty (Fig. 3b, Extended Data Fig. 6d). Electrophysiological recordings provided a putative explanation; PL neurons showed activity patterns consistent with transforming the task inputs to an attentional choice (Fig. 3c, d, Extended Data Fig. 7a–c). Notably, the rates of rise of attentional choice signals were modulated by uncertainty (Fig. 3d), indicating that PL ensembles may be integrating incoming cues into an attentional choice at a rate commensurate with input reliability. Consistent with this notion, putative inhibitory prefrontal fast spiking neurons showed modulation of spike rate by input uncertainty (Extended Data Fig. 7d–f). This finding gives rise to the notion that input uncertainty (which here we control through cueing conflict), engages prefrontal inhibition to modulate the speed of the cue-to-choice transformation.

Given the role of the MD in driving prefrontal inhibition^{12,20}, and the human findings about its activity scaling with task input uncertainty¹, we asked whether the MD was causally involved in the task. In contrast to the PL, MD inactivation during the cueing period did not cause a uniform detrimental effect in behavioural performance.

Specifically, its effect scaled with the level of input uncertainty (Fig. 3e, Extended Data Fig. 6d). The effect of optical MD inactivation was not simply a weaker form of PL inactivation (Extended Data Fig. 7i). Multi-electrode recordings provided insight into its causal engagement; MD neurons showed a high degree of specialization for input uncertainty, with some neurons showing a preference to trials with high conflict, and others to low conflict (Fig. 3f, g, Extended Data Fig. 7g). Critically, although relative conflict could be decoded from the PL, that signal was carried by the same neurons that encoded the attentional choice, standing in sharp contrast to the specialization seen in the MD (Extended Data Fig. 7g, h).

We asked whether this specialized encoding of input uncertainty could be causal to scaling prefrontal inhibition (Extended Data Fig. 7d). Because associative thalamic areas like the MD may integrate their cortical inputs to generate such ‘summary

statistic' type signals^{22,23}, we first tested whether optical deafferentiation of the MD by inhibiting PL terminals would diminish the encoding of conflict or uncertainty signals. Indeed, MD deafferentiation diminished conflict MD encoding (Extended Data Fig. 7k). Although this manipulation diminished behavioural performance (Extended Data Fig. 7j,q) and choice encoding in the PL (Extended Data Fig. 7k), it resulted in an overall increase in spike rates (Extended Data Fig. 7l), consistent with the MD primarily influencing PL cue-to-choice transformation through cortical inhibition in the current version of the task (Extended Data Fig. 7m).

To gain formal computational insight into this process, we built a neural model to study MD–PL interaction when inputs are conflicting (Extended Data Fig. 7o, Methods). This model was able to reproduce experimental data (Extended Data Fig. 7n,p–r), and provided insight into how the choice signal may be accumulated over time, and how MD-mediated suppression may slow it down when task inputs are conflicting and thereby unreliable (Extended Data Fig. 8a,b).

MD types engage differently if inputs conflict

Our results showed that MD_{GRIK4} neurons preferentially innervate PL PV⁺ neurons and that their activation inhibits baseline PL activity (Fig. 2n). Also, our neural model suggested that conflict-tracking in the MD drives PL inhibition to slow down cue integration when the inputs are less reliable (uncertain; Extended Data Fig. 7p). Thus, we hypothesized that conflict-tracking (or preferring) neurons may be GRIK4⁺. Indeed, optical tagging of MD_{GRIK4} neurons (Extended Data Fig. 9a) revealed that they were primarily conflict-preferring (Fig. 4a,b). By contrast, optically tagged MD_{D2} neurons showed the opposite functionality (Fig. 4c,d). Notably, non-tagged neurons in both of these preparations showed selectivity patterns consistent with generic MD recordings (Extended Data Fig. 9b,c). In addition, tagged MD neurons showed a spatial localization that is predicted by their anatomy (Extended Data Fig. 9f,g).

Fig. 4: The two thalamic cell types are engaged by different task input statistics.

 **figure4**

a, Example tagged MD_{GRIK4} neuron recorded in the task. **b**, Tagged MD_{GRIK4} neurons are more likely to be conflict-preferring ($n = 17$ neurons from 3 mice; $P = 0.0042$; binomial). **c**, Example tagged MD_{D2} neuron recorded in the task. **d**, Tagged MD_{D2} neurons are more likely to be conflict-non-preferring ($n = 20$ neurons from 3 mice; $P = 4.0 \times 10^{-5}$; binomial). **e**, MD_{GRIK4} suppression recapitulates generic MD suppression ($n = 20$ sessions from 4 GRIK4-cre mice; Wilcoxon signed-rank). **f**, MD_{D2} inactivation enhances performance accuracy on trials with high cueing conflict ($n = 20$ sessions from 4 D2-cre mice; Wilcoxon signed-rank). **g**, Expanded neural model with two MD cell types. **h**, The two-cell-type model captures experimental data ($n = 2,000$ trials, chi-

squared). MD+, MD intact; G-, without GRIK4; D-, without D2. **i**, Stimulus configuration for sparseness-driven uncertainty. **j**, Performance accuracy is modulated by cueing sparseness, and optical MD deafferentiation diminishes performance on trials with higher cueing sparseness ($n = 25$ sessions, 4 mice; $*P = 0.0222$, $***P = 1.02 \times 10^{-4}$; chi-squared). **k**, MD_{GRIK4} inactivation improves performance accuracy on both high and low signal trials ($n = 20$ sessions from 4 GRIK4-cre mice; Wilcoxon signed-rank). **l**, Optical MD_{D2} inactivation recapitulates optical generic MD deafferentiation ($n = 20$ sessions from 4 D2-cre mice; Wilcoxon signed-rank). All statistical tests are two-tailed. Box plot parameters as in Fig. 1. Data are mean \pm s.e.m. for **h, j**.

Source data

To examine whether these selectivity patterns translate to effects on behaviour, we performed optical inactivation of MD_{GRIK4} neurons or their terminals in the PL. Both manipulations reproduced generic MD inactivation (Fig. 4e, Extended Data Fig. 9d), confirming that this specific neural population suppresses the PL when its inputs are uncertain due to conflict.

Given that MD suppression did not affect behaviour in trials in which cueing uncertainty was low, we reasoned that the PL can maintain these task inputs without requiring thalamic amplification. As such, we predicted that inactivation of MD_{D2} neurons (or their terminals in the PL) would have no effect on task performance (Supplementary Note 2). Although this prediction was validated for trials with low conflict, it resulted in performance improvement on trials with high conflict (Fig. 4f, Extended Data Fig. 9e). This finding raised the hypothesis that MD_{D2} neurons must be engaged during the cueing period, and under our current task conditions they would be amplifying prefrontal signals in a manner that is detrimental to behaviour. We tested this idea first by modifying our neural model to incorporate the hypothesized function of the two thalamic cell types (Fig. 4g), which reproduced the data on one end (Fig. 4h) and provided computational insight into the idea that MD_{D2}-dependent amplification would increase the likelihood of non-preferred prefrontal inputs generating an erroneous choice (Extended Data Fig. 8c, d).

MD_{D2} neurons are required when inputs are sparse

If MD_{D2} neurons were amplifying functional cortical connectivity underlying the generation of a choice signal, we sought to ascertain whether there are conditions under which eliminating this MD_{D2} neural function would be detrimental to performance. We reasoned that if the task uncertainty was not due to input conflict and

instead due to input sparseness, this thalamic function may be required for optimal task performance. We first explored this conjecture in the model ([Methods](#)) and found it plausible (Extended Data Fig. [8e](#)). Therefore, we designed a task in which we controlled input uncertainty by varying the degree of informative pulse sparseness within each sequence, rather than conflict (Fig. [4i](#)). Because our earlier data indicated that MD inactivation was not required for such sequences when they included seven informative pulses, we varied their number between one and five in this new task design (Fig. [4j](#), Extended Data Fig. [10](#)). We found that the MD is also causally required for this task in a manner that scales with uncertainty due to input sparseness (Fig. [4j](#)). In other words, MD inactivation has a stronger effect on performance in trials with a low compared to a high input signal. Performing optical inactivation in cell-type specific Cre mice revealed that optimizing performance in this type of uncertainty condition was also segregated across these two thalamic populations: MD_{D2} neurons were required for performance in the low signal trials (Fig. [4l](#)), whereas optical inhibition of MD_{GRIK4} neurons resulted in enhanced performance in both high and low signal trials under this task condition (Fig. [4k](#)). Collectively, our experiments reveal a functional dissociation within MD–PFC loops in decision-making when inputs are uncertain. Specifically, MD_{D2} neurons that target disinhibitory VIP⁺ interneurons in the PL are required when task inputs are sparse (low signal), whereas MD_{GRIK4} neurons that target inhibitory PV neurons in the PL are required when inputs are dense but conflicting (high noise).

Discussion

Although studies in humans have shown that MD thalamic activity tracks task input uncertainty, our ability to capture this process in mice has revealed, first, that these responses are heterogeneous at the single-cell level; and, second, that they effectively break down input uncertainty into two categories: low signal and high noise. Notably, these different neural signals are carried by two genetically distinct thalamic projections.

Our data may be relevant for identifying interventions in schizophrenia. Several studies have indicated a heightened susceptibility of patients to uncertainty during decision-making^{[24](#)}, a process that may result in an unstable belief-updating process^{[25](#)}. As such, examining how the MD–PFC network responds to different types of uncertainty and in the context of hierarchical decisions is likely to be of value ([Supplementary Discussion](#)). On the more mechanistic end, given that some of the leading aetiological hypotheses are related to PV interneurons (a target of MD_{GRIK4} neurons)^{[26,27](#)} and D2 receptors^{[28](#)} (a marker for MD_{D2} neurons), we are optimistic that our findings will provide key details to link recently discovered thalamocortical

abnormalities^{5,29} to these classical ideas, opening up fresh avenues for therapeutic intervention.

Methods

Mice

A total of 94 mice were used in this study. Adult C57Bl/6 (wild-type) mice, of both sexes, aged 8–12 weeks old were purchased from Taconic Biosciences. GRIK4-cre, PV-cre, VIP-cre and SST-cre mice, of both sexes and aged between 8–12 weeks, were obtained from The Jackson Laboratory. D2-cre mice (GENSAT, line ER44), of both sexes and aged between 8 and 12 weeks, were a gift from M. Heiman. Cre mice were backcrossed to C57Bl/6 mice for at least six generations. All mice were kept in rooms with controlled temperature and ventilation (20–22 °C; 40–60% humidity) on a constant 12-h light–dark cycle. Mice were group housed with ad libidum access to food and water. All mouse experiments were performed according to the guidelines of the US National Institutes of Health and the Institutional Animal Care and Use Committee at the Massachusetts Institute of Technology.

Viruses

For retrograde monosynaptic tracing, EnvA-RVdG expressing mCherry (titre: 1.9×10^{11} vp ml $^{-1}$) was provided by I. Wickersham. Helper viruses AAV1-syn-FLEX-TA-TVA-GFP and AAV1-TREtight-B19G for monosynaptic tracing were also provided by I. Wickersham (titre: 1.0×10^{13} vp ml $^{-1}$). Retrograde AAV expressing Cre (AAVrg-hSyn-Cre-WPRE-hGH) was sourced from Addgene vector core (Addgene, lot 105553, titre: 7.0×10^{12} vp ml $^{-1}$). For optogenetic manipulation experiments, AAV1-CamKIIa-SSFO-eYFP (titre: 1.0×10^{13} vp ml $^{-1}$), AAV1-syn-ChR2-eYFP-Kv (titre: 4.6×10^{12} vp ml $^{-1}$), AAV2-CamkII-eNPHR3.0-eYFP (titre: 1.5×10^{12} vp ml $^{-1}$) and AAV2-EF1a-DIO-eNpHR3.0-eYFP (titre: 4.1×10^{12} vp ml $^{-1}$) were sourced from UNC vector. AAV8-EF1a-DiO-iC++-eYFP (titre: 1.5×10^{13} vp ml $^{-1}$) and AAV8-CamKIIa-iC++-eYFP (titre: 1.5×10^{13} vp ml $^{-1}$) were sourced from the Stanford Vector core. mGRASP labelling studies were performed using viruses AAV2/8-CAG-JxON-pre-mGRASP (titre: 2.0×10^{13} vp ml $^{-1}$) and AAV2/8-CAG-post-mGRASP-2A-dTomato (titre: 1.0×10^{13} vp ml $^{-1}$) sourced from Neurophotonics, University of Laval. For our intersectional approach to label MD_{D2} neurons in wild-type mice we used an AAV-8-D2SP-Cre-P2A-mCherry (titre: 1.50×10^{12} vp ml $^{-1}$) that drove mCherry and Cre expression under a D2-neuron-specific promoter. Simultaneous injections of another AAV-DJ hSyn Coff/Fon eYFP-WPREF (titre: 1.0×10^{12} vp ml $^{-1}$) allowed expression of YFP in Cre-negative (CreOFF) neurons. For cell-type-specific MD→PL

Kolmogorov–Smirnov labelling experiments, an AAV-DJ-hSyn-FLEX-mGFP-2A-Synaptophysin-mRuby (titre: 1.5×10^{13} vp ml $^{-1}$) virus was used.

Surgeries for anatomical tracing studies

Mice were first anaesthetized in an induction chamber receiving a continuous supply of oxygen and 5% isoflurane and then placed on a heating pad within a stereotaxic frame (Kopf Instruments). Throughout the surgery, anaesthesia was maintained through continuous delivery of 1–2% isoflurane via a nose cone at a rate of 1 l min $^{-1}$ and analgesia was provided by dual subcutaneous injections of slow-release buprenorphine (0.1 mg kg $^{-1}$) and Meloxicam (1 mg kg $^{-1}$). The midline of the scalp was sectioned and retracted, and a small craniotomy was made over the target region. After levelling the head, a small burr hole was made over each target region using coordinates based on the mouse brain atlas of Paxinos and Franklin³⁰. The coordinates are as follows (in mm from bregma): PL: antero-posterior (AP) 2.6, medio-lateral (ML) ± 0.3 , dorso-ventral (DV) –1.9; MD: AP –1.1, ML ± 0.6 , DV –3.0; A1: AP –2.92, ML ± 4 , DV –2.6; medial geniculate body (MGB), AP –3.0, ML ± 2.05 , DV –2.9 (from brain surface). For monosynaptic retrograde tracing experiments 300 nl of helper AAVs (1:1 mix of AAV1-syn-FLEX-TA-TVA-GFP and AAV1-TREtight-B19G) were injected into the PL of PV-cre, VIP-cre or SST-cre mice. Two weeks later, 100 nl of RVdG-mBFP2 (envA) was injected into the PL. Five days later the mice were euthanized to visualize monosynaptically labelled PL projection neurons in the MD (Fig. 2, Extended Data Fig. 3) and starter populations in the PL (Extended Data Fig. 3). To label cell-type-specific thalamocortical synapses, from MD neurons onto cortical PV $^+$ and VIP $^+$ interneurons with mGRASP (Extended Data Fig. 4), 75 nl of AAV2/8-CAG-JxON-pre-mGRASP (Cre-dependent) was injected into the MD and 200 nl of AAV2/8-CAG-post-mGRASP-2A-dTomato was injected into the PL of GRIK4-cre and D2-cre mice. Mice were given two weeks for expression of fluorescent proteins before being perfused as described in ‘Histology and immunohistochemistry’ below. To label MD→PL synaptic terminal densities across layers of the PL (Extended Data Fig. 4) we injected 75 nl of AAV-DJ-hSyn-FLEX-mGFP-2A-Synaptophysin-mRuby into the MD of GRIK4-cre and D2-cre mice. Mice were given two weeks for expression of fluorescent proteins before being perfused as described in ‘Histology and immunohistochemistry’ below.

Viruses were injected through a glass micropipette (Drummond Scientific) using a quintessential stereotactic injector (QSI, Stoelting) at a flow rate of 50 nl min $^{-1}$ and given 10 min to spread after injection. After the injection micropipettes were slowly retracted followed by closure of the incision.

Alternative strategy to target MD_{D2} neurons

Here we use a viral strategy to target D2⁺ neurons in the MD, independent of transgenic Cre lines. To this end, we injected an AAV with a short promoter that was previously established to express in D2 neurons only (AAV-8-D2SP-Cre-P2A-mCherry)³¹. This allowed us to examine neurons that are both D2⁺ (with a Cre-ON fluorophore) and D2⁻ through a simultaneous injection of another virus (AAV-DJ hSyn Coff/Fon eYFP-WPRE) that expresses only in the absence of Cre (Cre-OFF fluorophore³²). Fourteen days after injection of a 1:1 mixture of the two viruses into the MD we found a substantial overlap between neurons that were D2⁺ with this approach and neurons that are D2⁺ in the Cre line as well as neurons that project to VIP interneurons on the basis of rabies tracing.

Histology and immunohistochemistry

Mice were transcardially perfused with 30 ml of 0.1 M phosphate-buffered saline (PBS) followed by 20 ml of 4% paraformaldehyde (PFA) prepared in PBS. Brains were allowed to post-fix in the same fixative, overnight at 4 °C, then cryoprotected in 30% sucrose prepared in PBS for 24 h. Serial 50-μm-thick coronal sections were prepared using a Thermo HM550 cryotome. The GFP signal from the TVA helper constructs as well as the EYFP signal fused to opsins were enhanced with immunohistochemistry. In brief, sections were permeabilized and blocked in 10% bovine serum albumin (BSA, Sigma-Millipore) in PBS with 0.3% Triton X-100 (PBSTx) for 1 h. Then, sections were incubated overnight at 4 °C in primary chicken anti-GFP antibody (1:1,000, Aves Labs, GFP1011) prepared in PBSTx with 3% BSA. After two further washes, sections were incubated in an Alexa Fluor 488 goat anti-chicken secondary antibody (1:500, Thermo Fisher Scientific, A32931) for 2 h at room temperature, washed again and mounted for imaging. For mGRASP experiments, a similar protocol was followed to immunostain alternatePL sections (50 μm thick) from each brain for PL PV⁺ and VIP⁺ interneurons. We used rabbit anti-PV (1:1,000, Swant, PV-27) and rabbit anti-VIP (1:200, Immunostar, 20077) primary antibodies and an Alexa Fluor 647 donkey anti-rabbit secondary antibody (1:200, Thermo Fisher Scientific, A31573). GRIK4 protein was detected by an anti-rabbit primary GRIK4 antibody (1: 100, Alomone labs, AGC-041). For all viral injections, specificity of injection sites was verified using virally expressed fluorescent proteins (GFP, EYFP, mCherry). Mice in which injection sites missed the target location were discarded from further analysis.

In situ hybridization

Fresh-frozen brains from adult C57BL/6NJ mice (8–12 weeks) were sectioned at a thickness of 20 μm using a cryostat (Thermo Fisher Scientific). Sections were collected onto Superfrost Plus slides, immediately stored in a –20 °C freezer for 1 h

for tissue adherence and subsequently transferred to a -80°C freezer until staining. The D2 receptor mRNA signal was detected using the RNAscope fluorescent kit (Advanced Cell Diagnostics). Specifically, slides with sections corresponding to the MD were removed from the freezer, fixed with fresh and chilled 4% PFA for 15 min at 4°C and then dehydrated using a series of ethanol solutions of increasing concentrations (5 min each, room temperature): once 50%, once 70% and twice 100%. Next, sections were treated with hydrogen peroxide for 10 min followed by Protease IV (Advanced Cell Diagnostics) at room temperature for 30 min. Hybridization was performed on a HybEZ (Advanced Cell Diagnostics) oven for 2 h at 40°C using a mouse-specific D2 probe (Advanced Cell Diagnostics). After this, the slides were washed twice with a washing buffer (2 min each), then incubated with Hybridize Amp 1-FL for 30 min, Hybridize Amp 2-FL for 15 min and Hybridize Amp 3-FL for 30 min. Next, slides were incubated in horseradish peroxidase followed by TSA Plus Cyanine 3 fluorescent dye (1:750, Akoya Biosciences) for 30 min each at 40°C . Next, HRP blocker was added for 10 min at 40°C followed by counterstaining with DAPI for 30 s. The slides were washed twice with washing buffer (2 min each) and coverslips added using Prolong antifade mounting medium (Thermo Fisher Scientific). For negative controls the D2R probe was substituted with a probe against the *dapB* gene from the soil bacterium *Bacillus subtilis* while keeping all other steps the same.

Image analysis

For monosynaptic input tracing experiments, images were acquired on a confocal microscope (LSM 710, Zeiss) with a $20\times/0.80$ numerical aperture objective (Zeiss) and analysed using Imaris Image analysis software (Imaris 9.3.2, Oxford Instruments). Images were manually overlaid with vectorized outlines from a modified version of the Reference atlas from the Allen Brain Atlas (Unified anatomical atlas)³³ using anatomical landmarks as guides. Co-expression of GFP from the TVA-expressing helper virus and mBFP2 from the rabies virus were used to find bona fide starter neurons in the PL. Only those brains in which the starter neuron location was confined to the PL were processed for further analysis.

Monosynaptically labelled input cells, expressing mBFP2, were counted and their anatomical locations within the lateral MD recorded as follows. We measured the perpendicular distance of a candidate neuron from the lateral (medio-lateral distance axis) and ventral (dorso-ventral distance axis) boundaries of the MDI using the distance measure tool within Imaris. Their antero-posterior distance was measured from the anteriormost bregma location (AP -1.2 mm) where the MD is distinguishable into its three subdivisions—lateral, central and medial. The same method described above was used to image and record the anatomical location of MD_{GRIK4} and MD_{D2} neurons expressing mCherry in GRIK4-cre and D2-cre lines, respectively, as well as MD_{D2SP} neurons (Fig. 2, Extended Data Fig. 3).

These three distance measures (dorso-ventral, medio-lateral and antero-posterior distances) were used to perform a k -nearest neighbours (KNN) algorithm-based classification and cross-validation to examine anatomical separability of the prefrontal PV- and VIP-projecting MD neurons as well as the anatomical separability of MD_{GRIK4} and MD_{D2} neurons. In brief, each neuron is classified on the basis of a popularity vote of the identity of its five nearest neighbours, categorized as the most common identity among the five. The algorithm is repeated 100 times using 10-fold cross-validation. Neurons outside the 2.5% to 97.5% percentile, in any of the three distance axes, were excluded from further analysis as outliers.

As a second independent measure to validate the KNN-based classification, we performed representational similarity analysis. For each MD population (PV-projecting, VIP-projecting, MD_{GRIK4}, MD_{D2} and MD_{D2SP}), neuronal density is constructed along a three-dimensional (3D) space. The boundaries of the 3D space on each axis are placed at the minimum and maximum of location coordinates across all neurons. The 3D space is subsequently filled with evenly distributed nodes, with 10 each across the medio-lateral and dorso-ventral axis and 3 across antero-posterior axis (a total of 300 nodes), and the neuronal density is computed at each node. The representational similarity is computed as the Pearson correlation of the densities between different MD populations. When comparing within a population the comparison is performed across densities from 2 randomly separated halves, and the process is repeated 100 times.

To determine the laminar distribution of cell-type-specific MD terminal innervations, PL sections were imaged on a confocal microscope (LSM 710, Zeiss) with a 20 \times /0.80 numerical aperture objective (Zeiss). Multiple optical sections (1 $\mu\text{-m}$ thickness) were imaged to cover the entire z axis of the section and reconstructed in 3D using Imaris. The acquired image was subdivided into 50- $\mu\text{-m}$ -wide bins starting from the pial surface and the volume of fluorescent signal, from the synaptically tagged GFP within a bin, was quantified normalized to the total volume of GFP fluorescence across all the bins. Laminar layers within were delineated using ‘unified anatomical atlas’ demarcations³³.

For analysis of synapses labelled by mGRASP (Extended Data Fig. 4), PL sections were imaged using a confocal microscope (LSM 710, Zeiss) and 63 \times /1.40 numerical aperture objectives (Zeiss). Appropriate excitation wavelengths were used for different fluorescent protein markers: 488 nm for GFP (mGRASP-labelled synapses), 561 nm for TdTomato (post-mGRASP-labelled postsynaptic neurons) and 633 nm to detect anti PV or anti VIP immunohistochemistry fluorescent signal. Multiple optical sections (1- $\mu\text{-m}$ thickness) were imaged to cover the entire z axis of the section. Thereafter images were reconstructed in 3D and analysed using Imaris Image analysis software (Imaris 9.3.2, Oxford Instruments). Three-dimensional isosurfaces

(smoothness, 0.2 mm; quality level, 5) were created for each PV or VIP neuron identified by the co-expression of the post mGRASP TdTomato signal and immunohistochemistry for PV or VIP. A mask was then created to isolate the fluorescent signals within and surrounding the cell body to eliminate fluorescent signals from outside the cell boundaries. For each masked cell, a second round of 3D isosurfaces were created (smoothness, 0.1 mm; quality level, 7) for the mGRASP signal. Care was taken to ensure that the entire mGRASP signal was covered by the isosurfaces created. The number of such isosurfaces created was used to quantify the number of synapses per cell.

A similar approach was used to quantify GRIK4 expression in MD_{GRIK4} and MD_{D2} neurons. In brief, after acquisition, the images were reconstructed in 3D and analysed using Imaris. Three-dimensional isosurfaces (smoothness, 0.2 mm; quality level, 5) were created for each MD_{GRIK4} or MD_{D2} neuron identified by reporter fluorescence. Subsequently the mean intensity of GRIK4 immunolabelled fluorescent signal within each isosurface is used to quantify GRIK4 expression in the corresponding neuron.

To quantify D2 receptor mRNA expression in the MD using *in situ* hybridization as described above, stained slides were imaged in an LSM710 laser-scanning confocal microscope at 40 \times magnification. The lateral MD region from each section was isolated using ImageJ and individual images were merged into a stack. Then a maximum intensity projection of the stack in the *z* plane was generated using the ‘stacks’ plug-in in ImageJ and visualized as a heat map using the ‘EzColocalization’ plug-in³⁴.

Multi-electrode array construction and implantation

Custom multi-electrode array scaffolds (drive bodies) were designed using 3D CAD software (SolidWorks) and printed in Accura 55 plastic (American Precision Prototyping) as described in previous studies³⁵. Before implantation, each array scaffold was loaded with 16–24 independently movable micro-drives carrying 12.5- μ m nichrome (California Fine Wire) tetrodes. Electrodes were pinned to custom-designed, 64- or 96-channel electrode interface boards (EIB, Sunstone Circuits) along with a common reference wire (A-M Systems). For combined optogenetic manipulations and electrophysiological recordings, optic fibres (Doric Lenses) were embedded above or adjacent (for fibres equipped with a 45-degree mirror tip) to the electrodes. For analgesia, mice were injected with slow-release buprenorphine (1 mg kg⁻¹) before surgery. Then mice were deeply anaesthetized with 1% isofluorane and mounted on a stereotactic frame. The mouse head was shaved, and remaining hair removed with Nair. Body temperature was measured through a rectal probe and maintained using an electrical heating pad. An incision in the skin allowed access to the skull. An approximately 1.2 \times 1.6-mm craniotomy was drilled centred at (in mm

from bregma) AP 2, ML 0.6 for PL; at AP -1, ML 0.5 for MD; at AP -2.8, ML 4 for A1; and at AP -3.0, ML 2.0, DL 3.3 for MGB recordings. The dura was carefully removed, and the drive implant was lowered into the craniotomy using a stereotactic arm until the shortest tetrodes touched the cortical surface. Surgilube (Savage Laboratories) was applied around electrodes to guard against fixation through dental cement. Stainless steel screws were implanted into the skull to provide electrical and mechanical stability and the entire array was secured to the skull using dental cement. The skin was subsequently closed with Vetbond and the mouse was allowed to recover on a heating blanket.

Head fixation recordings

Simultaneous recordings from MD and PL or MGB and A1 were conducted in a custom-built set-up. The head-fixation system consisted of a pair of custom 3D printed plastic fixation clamps (MakerBot Replicator) used to lock the implanted plastic crown at the base of the implant into place during recordings. These were fixed to an acrylic plastic frame which also supported a platform on which the mouse stood. The platform was composed of low-friction acrylic and was adjusted based on the height of the mouse and spring-loaded to minimize torque on the implant.

Electrophysiological recordings

Signals from tetrodes (thalamic recordings) were acquired using a Neuralynx multiplexing digital recording system (Neuralynx) via a combination of 64- and 96-channel digital multiplexing head stages plugged to the 64–96 channel EIB of the implant. Signals from each electrode were amplified, filtered between 0.1 kHz and 9 kHz and digitized at 30 kHz. For thalamic recordings, tetrodes were lowered from the cortex into MD -2.8 to -3.2 mm DV and into the MGB -2.8 to -3.2 mm DV. For PL recordings, adjustments accounted for the change of depth of PL across the anterior-posterior axis. Thus, in anterior regions, unit recordings were obtained between -1.2 to -1.7 mm DV whereas for more posterior recordings electrodes were lowered -2 to -2.4 mm DV. For A1 unit recordings were obtained between -2.5 to -3.0 mm DV. Following acquisition, spike sorting was performed offline on the basis of relative spike amplitude and energy within electrode pairs using the MClust toolbox (<http://redishlab.neuroscience.umn.edu/mclust/MClust.html>).

Identification of fast spiking and regular spiking cells

After initial spike sorting, PL units were divided into fast spiking (FS) and regular spiking (RS) according to waveform characteristics and spike rate as described previously⁷. Basic features of spike waveforms, including peak to trough time, half trough time, and trough depth, were measured for each unit across all spike

waveforms. We also incorporated a measure of spike timing that has previously been used to identify FS neurons (spike rate)³⁶. Recorded neurons were then separated using a clustering method for the four feature dimensions: (1) half trough time; (2) peak to trough time; (3) trough depth; and (4) spike rate. Clustering across the four dimensions were assessed using *k*-means clustering as described previously.

Connectivity assay

To assess the effect of changes in thalamic excitability on cortical connection strength, we measured intra-cortical responses evoked by ChR2-mediated activation of the contralateral cortex for A1–MGB and PL–MD. Responses to either cortical stimulation alone (10 ms ChR2 activation to the contralateral cortex), thalamic activation alone (500 ms stabilized step function opsin (SSFO) activation in ipsilateral MGB or MD) or the combination were recorded in A1 and PL (50 interleaved trials per condition). For the combined condition, thalamic activation preceded cortical stimulation by 100 ms. To test the role of PL VIP neurons on MD-driven amplification of cortical connection strength, we also measured the responses of contralateral cortical stimulation alone, ipsilateral MD stimulation alone or combined stimulation with concurrent suppression of PL VIP neurons (1,000 ms NpHR3.0 activation) with an onset 500 ms before ChR2 activation).

For all cortical neurons, changes in baseline and evoked spike rates were assessed using peri-stimulus time histograms (PSTHs). PSTHs were computed using a 1 ms bin width for individual neurons in each recording session convolved with a Gaussian kernel (20 ms full width at half maximum) to create a spike density function (SDF). Evoked response through intracortical stimulation was measured as the baseline rate normalized delta between the maximum firing rate in a window 100 ms after ChR2 onset and the baseline rate measured over 500 ms before any laser stimulation. Proportional spike rate changes in the absence of contralateral cortical stimulation were calculated relative to the baseline rate.

Behaviour

Set-up

Behavioural training and testing took place in custom-built enclosures as previously described³⁷. All enclosures contained custom-designed operant ports, each equipped with an IR LED/IR phototransistor pair (Digikey) for nose-poke detection. An additional port for trial initiation was mounted on the floor 6 cm away from the ‘response ports’ located at the front of the chamber. Auditory cues and targets were presented with millisecond precision through a ceiling mounted speaker controlled by an RX8 Multi I/O processing system (Tucker-Davis Technologies). Visual stimuli

were presented via two dimmable, white light emitting diodes (Mouser) mounted on each side of the initiation port. Two response ports were mounted at the angled front wall and a milk reward (10 µl evaporated milk) was directly delivered into the ports via a syringe pump (New Era Pump Systems) to reward correct choices. Access to the response ports was restricted by vertical sliding gates controlled through a servo motor (Tower Hobbies). The TDT Rx8 sound production system (Tucker Davis Technologies) was triggered through MATLAB (MathWorks), interfacing with a custom written software running on an Arduino Mega (Ivrea) for trial logic control. Across experiments, mice were randomly selected for training and all mice trained to criteria were included in testing. For optogenetic studies and physiological recording, mice were randomly selected from the overall cohort for inclusion in each type of manipulation or recording.

Training for the PL-dependent task

Training was largely similar to a previously described approach^{7,37}. First, 10 µl of evaporated milk (reward) was delivered randomly to each reward port for shaping and reward habituation. Making response ports accessible signalled reward availability. Illumination of the LED at the spatially congruent side was used to establish the association with the visual targets on half of the trials while a similar presentation of a 100-ms tone cloud on the other half of the trials was used to build the association with the auditory target. An individual trial was terminated 15 s after reward collection, and a new trial became available 5 s later.

Second, mice learned to poke to receive a reward. All other parameters remained constant. An incorrect poke had no negative consequence. By the end of this training phase, all mice collected at least 20 rewards per 30-min session.

Third, mice were trained to initiate trials in which mice had to briefly (50 ms) break the infrared beam in the initiation port to trigger target stimulus presentation and render reward ports accessible. Trial rule ('attend to vision' or 'attend to audition') was indicated by 4 to 8 kHz low-pass (LP)-filtered white noise (vision) or 12 to 40 kHz high-pass (HP)-filtered white noise (audition) sound cues. Stimuli were presented in blocks of six trials consisting of single-modality stimulus presentation (no conflict). An incorrect response immediately rendered the response port inaccessible. Rewards were available for 15 s after correct poking, followed by a 5-s inter-trial interval (ITI). Incorrect poking was punished with a time-out, which consisted of a 30-s ITI. During an ITI, mice could not initiate new trials.

Fourth, conflict trials were introduced, in which auditory and visual targets were co-presented indicating reward at opposing locations. Trial types were presented in blocks

of visual or auditory trials. The time that mice had to break the infrared barrier in the initiation port was continuously increased until it reached 0.8 s.

Fifth, trial availability and task rule were dissociated. Broadband white noise indicated trial availability, which prompted a mouse to initiate a trial. After successful initiation, the white noise was immediately replaced by either low-pass- or high-pass-filtered noise for 0.1 s to indicate the rule. This was followed by a delay period (variable, but for most experiments it was 0.4 s) before target stimuli presentation. All block structure was removed, and trial type was randomized. Mice were trained on this discrete cueing version of the task until mean performance plateaued and remained stable over 4–5 consecutive sessions (mean accuracy of $69 \pm 3\%$ correct). On a subset of trials, the two targets were shown on congruent sides to ensure that mice did not develop a pro-anti strategy for a single cue.

Mice were implanted with optic fibres in the PL and MD at this stage and retrained for testing with optogenetic manipulation (described below) for experiments involving a single HP or LP cueing pulse (Extended Data Fig. 5).

Sixth, single HP or LP pulses were replaced by sequences of several 50-ms-long pulses of either HP or LP, separated by a 25-ms gap of silence. In parallel, snout fixation duration was increased until a total of 16 pulses could fit within the cueing period (1,200 ms). Finally, unlike the single-pulse version of the task, the noise-free delay between the end of the cueing pulses and the presentation of choice targets was intentionally kept below 250 ms to focus our study on uncertainty in sensory inputs. Once the mice performed on these ‘pure’ sequences equivalent to the single-pulse trials, input uncertainty trial types were introduced in which the evidence varied for attend to vision versus attend to audition. Conflict-driven input uncertainty trials were generated by incorporating different mixtures of HP, LP, and broadband white noise (conflict mediated uncertainty). Out of the 16 pulses, only 9 conveyed rule information (either HP or LP). The remaining seven pulses consisted of broadband white noise. Low-signal-driven input uncertainty trials only contained one type of meaningful pulses (either HP or LP) embedded in broadband white noise pulses. Out of the 16 pulses, only 1 to 5 pulses were meaningful to make those cueing sequences sparse in signal. Mice were required to select the appropriate target stimulus based on the rule with the highest number of corresponding pulses on a trial-by-trial basis. Trial types were presented in random order.

Training for the PL-independent task

The first two training steps were similar to the PL-dependent 2AFC task except the target modality was restricted to the visual domain where an LED was illuminated for 10 ms at a spatially congruent side to indicate rewarded response port. In the next

stage of training mice were trained to initiate trials in which they had to briefly (50 ms) break the infrared beam in the initiation port to trigger target stimulus presentation and render reward ports accessible. Target stimuli were presented in blocks of six trials consisting of single-modality stimulus presentation (no conflict). An incorrect response immediately rendered the response port inaccessible. Rewards were available for 15 s after correct poking, followed by a 5-s ITI. Incorrect poking was punished with a time-out, which consisted of a 30-s ITI. During an ITI, mice could not initiate new trials. On the final stage of the task trial availability and target presentation were dissociated. Broadband white noise indicated trial availability, which prompted a mouse to initiate a trial. After successful initiation, the white noise was immediately replaced by illumination of a LED light on the left or right to indicate the response port where reward was available. All block structure was removed, and trial type was randomized. Mice were trained on this version of the task until performance plateaued and remained stable over 4–5 consecutive sessions.

Optogenetic manipulation

We used a dual wavelength optical silencing method to independently suppress neurons in the PL and MD. Specifically, we virally expressed halorhodopsin (AAV2-CamkII-eNPHR3.0-eYFP) in the PL and a Cre-dependent (in GRIK4-cre and D2-cre mice; AAV8-EF1a-DiO-iC++-eYFP) or Cre-independent (in wild type mice; AAV8-CamKIIa-iC++-eYFP) inhibitory channelrhodopsin iC++ in the MD. As the peak spectrum of NpHR3.0 is red-shifted (peak around 550 nm), we could independently inactivate both populations or their terminals in either structure, through implanted optic fibres, using a 473-nm and a 556-nm laser (OptoEngine) to activate iC++ and NpHR3.0 respectively. For all optogenetic experiments (Figs. 3, 4, Extended Data Figs. 6, 7, 9, 10), optogenetic trials were randomly interleaved among other trial types and investigators were blinded to trial type; longitudinal comparisons were then used within individuals between trial types. This is true except for experiments in which the role of MD in task engagement was evaluated (Extended Data Fig. 5), or in the optotagging experiments (Fig. 4). In the former experiments, optogenetic inactivation of the MD was done on trial number 1 to 30 of the session (Extended Data Fig. 5). Laser duration varied depending on the trial type between 100 ms (during single-pulse cueing period; Extended Data Fig. 5), 400 ms (single-pulse delay period; Extended Data Fig. 5) and 1,200 ms (entire cueing period of a 16-pulse cueing sequence). In the latter experiments, optogenetic tagging was performed after the behaviour session (see below). During a session, only one condition was tested with optogenetic manipulation.

Firing rate analysis

For all thalamic and cortical neurons, changes in spike rates associated with task performance were assessed using PSTHs. PSTHs were computed using a 1 ms bin width for individual neurons in each recording session convolved with a Gaussian kernel (20 ms full width at half maximum) to create an SDF. Proportional firing rate change was calculated relative to a 500-ms-long baseline before event onset. Notably, all task-related rasters and PSTHs (and neural analysis such as decoding analysis) are aligned to cue onset ($t = 0$).

Classification of thalamic neurons into conflict-preferring versus conflict-non-preferring

Conflict-preferring and conflict-non-preferring neurons were identified using the area under receiver operating characteristics (auROC) method. In brief, auROC provides an aggregate measure of the association between single-trial firing rates and trial type, across levels of response. For each neuron, the proportional response for each trial was computed over the 300–1,200 ms window after cue onset (the beginning of the cueing period, when the conflict signal had just began to emerge, was omitted). The fraction of trials for which the proportional response exceeds a threshold, as a function of varying threshold, was computed over two trial types (for example, low conflict trials and high conflict trials). The ROC curves are pairs of fractions for the two trial types (f_1, f_2) over each shared threshold value, plotted with one trial type over one axis. As such, the ROC curve goes from (0,0) (when the threshold is higher than the response in all trials) to (1,1) (when the threshold is lower than the response in all trials). The auROC computes the area below the ROC curve between (0,0) and (1,1). All neurons from the population of interest were pooled together and their auROC was computed as above. Neurons with auROC significantly above 0.5 (that is, > 1.5 standard deviation (SD)) for high versus low conflict trials are defined as conflict-preferring. Neurons with auROC significantly above 0.5 (that is, > 1.5 SD) for low versus high conflict trials are defined as conflict-non-preferring.

Decoding analysis

Trial-by-trial classification analysis was performed using a support vector machine (SVM) implemented through LIBSVM and MATLAB (Mathworks) neural decoding toolbox³⁸, similar to previously reported³⁹. To perform decoding on cue, choice or conflict, the firing rates of neurons on each trial from the entire population (pooled across sessions) were first smoothed using a Gaussian filter of 20 ms width. The SVM classifier with a Gaussian radial basis function kernel was then trained on (randomly selected) half of the data and tested on the other half of the data, with a sliding window of 300 ms and time step of 100 ms. The classes were balanced during training, such that an equal number of trials were (randomly) selected for each class. This classifier works by first constructing an optimal hyperplane based on labelled training data and

then generating predictions of the labels on testing data. Accuracy of the decoding was assessed by comparing the predicted labels to the actual labels. Classification accuracy was also quantified by computing the mutual information via the following equation:

$$\text{MI} = \sum_{i=1}^s \sum_{j=1}^s p_{ij} \log \frac{p_{ij}}{p_i p_j}$$

where p_{ij} is the probability of observing label i (cue, choice, or conflict) given that the original label is j . This classification process was repeated 100 times to obtain and accurately estimate the error of the classification accuracy.

To analyse the separability of conflict and choice information in MD and PL, 50 of the most conflict-selective MD neurons, and 50 of the most choice-selective PL putative excitatory neurons, are pooled. Decoding is performed as described above, and the maximum classification accuracy is computed.

Optogenetic tagging and identification of cell-type-specific MD neurons

GRIK4-cre and D2-cre mice trained on the cueing uncertainty version of the attention control task were injected with AAV2-EF1a-DIO-eNpHR3.0-eYFP in the MD and implanted with multi electrode arrays and optic fibres targeted to the MD. After every behaviour session, and in a separate box outside of task context, each mouse received 50 trials of 10-ms-long pulses of eNpHR3.0 stimulation. Three features of the response of an MD neuron to eNpHR3.0 stimulation were measured for each neuron in a 50-ms window after eNpHR3.0 stimulation: (1) change in mean proportional spike rates; (2) fraction of trials with spike rate suppression; and (3) recovery half-time (Extended Data Fig. 9a). Tagged neurons were identified using k -means clustering across the three dimensions. Optotagged clusters of MD_{GRIK4} or MD_{D2} neurons so identified demonstrated a strong decrease in proportional spike rates and high fraction of trials with rate suppression. Subsequently, the tagged MD_{GRIK4} or MD_{D2} neurons were classified into conflict-preferring versus conflict-non-preferring from the responses recorded in the preceding behaviour session.

Neural model for decision-making circuit

To study how MD may optimize PL computation in generating choice signal under input conflict, we constructed a neural mean field model (reduced form of a spiking circuit model) of the PL circuit executing a 2AFC decision-making task⁴⁰. Whereas a spiking circuit model describes the temporal evolution of hundreds or thousands of neural units (under a defined circuit architecture), a mean-field model averages over homogeneous populations, smearing over interactions and resulting in a low-dimensional system with key dynamics of interest. Similar models were used in the

literature to capture key features of human and primate behavioural and neural data³⁹. Variants of the model regime had also shed light on the decision-making neural circuitry in mice⁴¹.

Specifically, our model (custom Python code) described two excitatory populations within the PL that received inputs corresponding to high-pass and low-pass pulses respectively, and the outputs of which would be read out to form the attentional choices. Each excitatory population had recurrent connections onto itself that allowed integration of the input pulses. The two populations also project to an inhibitory population that symmetrically suppresses both populations, resulting in competition between the two populations. We also incorporated MD→PL projections into the model as constrained by experimental data. We considered two different implementations of the MD module. In the first implementation (Extended Data Fig. 7o), MD dynamically computes cueing conflict to activate the PL inhibitory population and suppress both PL excitatory populations accordingly. The second implementation incorporated the two thalamic cell types, with MD_{GRIK4} dynamically activated under cueing conflict to suppress PL, whereas MD_{D2} was conflict-suppressed and amplified recurrence in PL.

The mean-field model described the temporal evolution of NMDA receptor (NMDA-R) gating variables of the two excitatory populations (S_1, S_2), which were also the decision variable representing the integrated evidence for the two choices. The model also included firing rates and other synaptic gating variables of the two populations. However, they were treated as steady states owing to their much shorter timescales than NMDA-R gating variables.

The two NMDA-R gating variables evolved according to:

$$\frac{dS_i}{dt} = -\frac{S_i}{\tau_{NMDA}} + (1 - S_i)\gamma r_i \quad (1)$$

for $i = 1, 2$. $\tau_{NMDA} = 100$ ms and $\gamma = 0.641$ were the synaptic time constant and saturation factor for NMDA-R. r_1, r_2 were the firing rates of the two excitatory populations. These rates were computed from the transfer function based on the total input currents I_1, I_2 . The input currents:

$$I_1 = \alpha_1 S_1 + \alpha_2 S_2 + \beta_1 r_1 + \beta_2 r_2 + I_{ext} \quad (2)$$

$$I_2 = \alpha_1 S_2 + \alpha_2 S_1 + \beta_1 r_2 + \beta_2 r_1 + I_{ext}$$

(3)

arose from the NMDA-Rs of the same population (for example, $\alpha_1 S_1$ in equation (1)) and competing population (for example, $\alpha_2 S_2$ in equation (2)), the AMPA receptor gating variables of the same population (for example, $\beta_1 r_1$ in equation (2)) and competing population (for example, $\beta_2 r_2$ in equation (2)), and external inputs (for example, $\{I\}_{1,2}^{\rm ext}$ in equation (2)). GABA receptor gating variables were also expressed in α_i and β_i to account for lateral inhibition. The synaptic parameter values are $\alpha_1 = 0.164$ nA, $\alpha_2 = -0.022$ nA, $\beta_1 = 9.9 \times 10^{-4}$ nC, $\beta_2 = -6.5 \times 10^{-5}$ nC. The external input $\{I\}_{1,2}^{\rm ext}$ is due to a constant but noisy input $\{I\}_{1,2}^{\eta}$, and a stimulus input $\{I\}_{1,2}^{\rm stim}$. $\{I\}_{1,2}^{\eta} = \{I\}_{1,2}^{\eta} + \{I\}_{1,2}^{\rm stim}$. $\{I\}_{1,2}^{\eta}$ is described by an Ornstein-Uhlenbeck process with mean $I_{\text{OU}} = 0.350$ nA, noise $\sigma_{\text{OU}} = 0.015$ nA, and time constant $\tau_{\text{OU}} = 2$ ms. $\{I\}_{1,2}^{\rm stim}$ is 0.017 nA under the presence of favoured input pulses, but 0 otherwise. Using change of variables $x_1 = \alpha_1 S_1 + \alpha_2 S_2$, $x_2 = \alpha_1 S_2 + \alpha_2 S_1$, the transfer function can be written as

$$\frac{r_1}{r_2} = \frac{a(x_1) - f(x_2) - b}{a(x_2) - f(x_1) - b} \quad (4)$$

$$\frac{r_2}{r_1} = \frac{a(x_2) - f(x_1) - b}{a(x_1) - f(x_2) - b} \quad (5)$$

where a, b, d were constants that depended on β_1 , and f was a function of x_i that depended on β_2 . The expression of $\{\alpha_i, \beta_i, a, b, d, f, \{I\}_i^{\rm ext}\}$ are detailed in a previous study⁴⁰, but in brief, the transfer function results in a smooth and thresholded input–output response (Extended Data Fig. 8c, bottom). A choice was selected at the end of stimulus presentation, based on the population with higher decision variable (S_1, S_2). Stimulus inputs in general drove categorical, winner-take-all competitions such that the two decision variables were largely separated (with the loser decision variable near 0; Extended Data Fig. 8).

In the model with a generic MD (Extended Data Fig. 7o) inactivation, the effect of MD was incorporated as inhibitory inputs to the two PL populations in the presence of conflict $\{I\}_{1,2}^{\rm ext} = \{I\}_{1,2}^{\eta} + \{I\}_{1,2}^{\rm stim} + \{I\}^{\rm MD}$. Conflict was dynamically computed by considering the current pulse and the last non-white-noise pulse (that is, if one pulse was HP and the other LP), although other implementations of conflict computation yielded consistent results. In addition, a baseline suppression to PL was added to dissociate the effects of

MD inactivation versus MD deafferentation (Extended Data Fig. 7p,r) ($I_{MD} = -0.1$ nA under conflict, $= -0.01$ nA without conflict). In particular, MD inactivation removed all effect of MD, whereas the baseline suppression to PL remained under optical inhibition of PL \rightarrow MD terminals.

In the model with two MD cell types (Fig. 4g), the effect of MD_{GRIK4} was similarly incorporated similarly as inhibitory inputs to the PL in the presence of conflict \((\{\{ I \}_{1,2}^{\wedge}\{\{\backslash rm\{ ext\} \}\}}=\{ I \}_{1,2}^{\wedge}\{\backslash eta \}+\{ I \}_{1,2}^{\wedge}\{\{\backslash rm\{ stim\} \}\})+\{ I \}^{\wedge}\{\{\backslash rm\{ GRIK4\} \}\})\). However, the baseline suppression was removed for simplicity ($I_{GRIK4} = -0.1$ nA under conflict, 0 without conflict) considering similar effects of MD inactivation and optical inhibition of PL \rightarrow MD terminals in the previous model. The effect of MD_{D2} was incorporated as an augmentation to the recurrent synaptic connections (8% increase to β_1 and β_2 , equations (2) and (3)), resulting in a gain increase of the transfer function (equations (4) and (5); Extended Data Fig. 8c, bottom). In the models without MD_{GRIK4} or MD_{D2} (Fig. 4h), the corresponding module was removed. Finally, a slightly altered circuit model was used to demonstrate the viability that MD_{D2} may contribute to decision-making under input uncertainty due to cueing sparseness (Fig. 4j, Extended Data Fig. 8e). We reduced I_{OU} to slow down the rate for which decision variables approach attractor states. This corresponded to a slower integration process, allowing the model circuit to accumulate sparse evidence distributed across the cueing period, early or late. We note that this altered model was only used to generate example traces (Extended Data Fig. 8e) and was not used in any analysis.

Regression analysis

Regression analysis was used to ensure mice used the entire cue sequence to inform their choice behaviour. In particular, a logistic regression model on choice (correct or error) was performed with the evidence in the first (early) and second (late) half of the cue sequence as regressors:

$$\begin{aligned} \$\$ \ln \left(\frac{P}{1-P} \right) &= \beta_0 + \beta_e \sum_{i=1}^{8} C_i + \beta_l \sum_{i=9}^{16} C_i \\ (6) \end{aligned}$$

where P is the probability to be correct, C_i is the i th pulse in the trial (= 1 for a low-pass pulse, $= -1$ for a high-pass pulse, $= 0$ for a white noise pulse), β_0 is the bias term, and β_e and β_l reflect the degree the magnitude of momentary cues in the early and late half, respectively, contribute to animal choice behaviour.

Statistical analysis

Statistical analysis was performed in MATLAB (Mathworks) and GraphPad Prism software (v.8.0, Prism). We did not assume normality in the distribution of our datasets and hence used two-sided non-parametric statistics to test for significance. For each statement of statistical difference included in the manuscript, a corresponding statistical comparison was performed, as mentioned in the figure legends. In brief, we used a Mann-Whitney *U* test for all comparisons between two groups comprising independent samples and a Wilcoxon signed-rank test when the samples were dependent. For comparison of cumulative distributions, the Kolmogorov–Smirnov test was used. For comparisons of observed proportions of binary (categorical) variables, we used a binomial test to compare to chance, and a chi-squared test to compare across two groups. For comparisons of decoding accuracies, we used permutation tests, rerunning the decoding analysis with shuffled trial labels, computing the fraction of trials exceeding the reported value. When comparing across conditions (laser off versus laser on), the shuffling is performed on neurons across conditions. For logistic regression, a two-sided Student’s *t*-test was used, as part of the output of MATLAB function *glmfit*. All *P* values are listed in the figure legends. Values are expressed as medians \pm 95% range in box-and-whisker plots and mean \pm s.e.m. for bar graphs.

Power analysis

For behavioural studies, power analyses were performed to determine the number of mice needed to establish an effect. Specifically, the MATLAB function *sampsizewr* was used to estimate the number of mice. For the single-cue tasks, we performed a priori power analysis based on previously published data of the same task with MD manipulation^{7,20}. The expected value and standard deviation of the null hypothesis (that optical manipulation has no effect), respectively, were 0.64 and 0.025, and the expected value of the alternative hypothesis (that optical manipulation abolishes performance) is 0.5, resulting in an effect size of Cohen’s *d* = 5.6. With a significance value of 0.05 and a power of 0.7, we estimated a number of 3 mice to be appropriate. We used 3–4 mice across experiments. Number of mice in each panel: Extended Data Fig. 5*i,l*: 4 mice; rest of Extended Data Fig. 5*d–n*: 3 mice.

For the conflict and sparseness tasks, we assumed similar variability in the data and effect size, thus resulting in the same estimated number of 3 mice. However, to be cautious with variability of the effect size we collected data from 4–6 mice for distinct optical manipulation experiments. Number of mice in each panel: Fig 3, Extended Data Fig. 7*i*: 5 mice; Fig. 4, Extended Data Fig. 9*d,e*: 4 mice of each genotype; Extended Data Figs. 6*a–c*, 7*j,q*: 6 mice. Also see Supplementary Table 1.

Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request. [Source data](#) are provided with this paper.

Code availability

Custom codes for analysis and modelling were written in MATLAB and are available from the corresponding author upon request.

References

1. 1. Kosciessa, J. Q., Lindenberger, U. & Garrett, D. D. Thalamocortical excitability modulation guides human perception under uncertainty. *Nat. Commun.* **12**, 2430 (2021).
2. 2. Krug, A. et al. Attenuated prefrontal activation during decision-making under uncertainty in schizophrenia: a multi-center fMRI study. *Schizophr. Res.* **152**, 176–183 (2014).
3. 3. Culbreth, A. J., Gold, J. M., Cools, R. & Barch, D. M. Impaired activation in cognitive control regions predicts reversal learning in schizophrenia. *Schizophr. Bull.* **42**, 484–493 (2016).
4. 4. Barbalat, G., Chambon, V., Franck, N., Koechlin, E. & Farrer, C. Organization of cognitive control within the lateral prefrontal cortex in schizophrenia. *Arch. Gen. Psychiatry* **66**, 377–386 (2009).
5. 5.

Giraldo-Chica, M., Rogers, B. P., Damon, S. M., Landman, B. A. & Woodward, N. D. Prefrontal–thalamic anatomical connectivity and executive cognitive function in schizophrenia. *Biol. Psychiatry* **83**, 509–517 (2018).

6. 6.

Pinault, D. A neurophysiological perspective on a preventive treatment against schizophrenia using transcranial electric stimulation of the corticothalamic pathway. *Brain Sci.* **7**, 34 (2017).

7. 7.

Schmitt, L. I. et al. Thalamic amplification of cortical connectivity sustains attentional control. *Nature* **545**, 219–223 (2017).

8. 8.

Mukherjee, A. et al. Variation of connectivity across exemplar sensory and associative thalamocortical loops in the mouse. *eLife* **9**, e62554 (2020).

9. 9.

Usrey, W. M. & Alitto, H. J. Visual functions of the thalamus. *Annu. Rev. Vis. Sci.* **1**, 351–371 (2015).

10. 10.

Anastasiades, P. G., Collins, D. P. & Carter, A. G. Mediodorsal and ventromedial thalamus engage distinct L1 circuits in the prefrontal cortex. *Neuron* **109**, 314–330 (2021).

11. 11.

Williams, L. E. & Holtmaat, A. Higher-order thalamocortical inputs gate synaptic long-term potentiation via disinhibition. *Neuron* **101**, 91–102 (2019).

12. 12.

Ferguson, B. R. & Gao, W. J. Thalamic control of cognition and social behavior via regulation of gamma-aminobutyric acidergic signaling and excitation/inhibition balance in the medial prefrontal cortex. *Biol. Psychiatry* **83**, 657–669 (2018).

13. 13.

Phillips, J. W. et al. A repeated molecular architecture across thalamic pathways. *Nat. Neurosci.* **22**, 1925–1935 (2019).

14. 14.

Gao, C. et al. Two genetically, anatomically and functionally distinct cell types segregate across anteroposterior axis of paraventricular thalamus. *Nat. Neurosci.* **23**, 217–228 (2020).

15. 15.

García-Cabezas, M. Á., Martínez-Sánchez, P., Sánchez-González, M. Á., Garzón, M. & Cavada, C. Dopamine innervation in the thalamus: monkey versus rat. *Cereb. Cortex* **19**, 424–434 (2009).

16. 16.

Baek, J. et al. Neural circuits underlying a psychotherapeutic regimen for fear disorders. *Nature* **566**, 339–343 (2019).

17. 17.

Hu, H., Gan, J. & Jonas, P. Fast-spiking, parvalbumin⁺ GABAergic interneurons: from cellular design to microcircuit function. *Science* **345**, 1255263–1255263 (2014).

18. 18.

Feng, L., Kwon, O., Lee, B., Oh, W. C. & Kim, J. Using mammalian GFP reconstitution across synaptic partners (mGRASP) to map synaptic connectivity in the mouse brain. *Nat. Protoc.* **9**, 2425–2437 (2014).

19. 19.

Bolkan, S. S. et al. Thalamic projections sustain prefrontal activity during working memory maintenance. *Nat. Neurosci.* **20**, 987–996 (2017).

20. 20.

Rikhye, R. V., Gilra, A. & Halassa, M. M. Thalamic regulation of switching between cortical representations enables cognitive flexibility. *Nat. Neurosci.* **21**, 1753–1763 (2018).

21. 21.

Grinband, J., Hirsch, J. & Ferrera, V. P. A neural representation of categorization uncertainty in the human brain. *Neuron* **49**, 757–763 (2006).

22. 22.

Hayden, B. Y., Pearson, J. M. & Platt, M. L. Neuronal basis of sequential foraging decisions in a patchy environment. *Nat. Neurosci.* **14**, 933–939 (2011).

23. 23.

Jaramillo, J., Mejias, J. F. & Wang, X.-J. Engagement of Pulvino-cortical feedforward and feedback pathways in cognitive computations. *Neuron* **101**, 321–336 (2019).

24. 24.

Cole, D. M. et al. Atypical processing of uncertainty in individuals at risk for psychosis. *NeuroImage Clin.* **26**, 102239 (2020).

25. 25.

Nassar, M., Waltz, J., Albrecht, M., Gold, J. & Frank, M. All or nothing belief updating in patients with schizophrenia reduces precision and flexibility of beliefs. *Brain* **144**, 1013–1029 (2021).

26. 26.

Mukherjee, A., Carvalho, F., Eliez, S. & Caroni, P. Long-lasting rescue of network and cognitive dysfunction in a genetic schizophrenia model. *Cell* **178**, 1387–1402 (2019).

27. 27.

Lewis, D. A., Curley, A. A., Glausier, J. R. & Volk, D. W. Cortical parvalbumin interneurons and cognitive dysfunction in schizophrenia. *Trends Neurosci.* **35**, 57–67 (2012).

28. 28.

Brisch, R. et al. The role of dopamine in schizophrenia from a neurobiological and evolutionary perspective: old fashioned, but still in vogue. *Front. Psychiatry* **5**, 47 (2014).

29. 29.

Chen, P., Ye, E., Jin, X., Zhu, Y. & Wang, L. Association between thalamocortical functional connectivity abnormalities and cognitive deficits in schizophrenia. *Sci. Rep.* **9**, 2952 (2019).

30. 30.

Franklin, K. B. J., & Paxinos, G. *The Mouse Brain in Stereotaxic Coordinates*, 3rd edn (Academic, 2008).

31. 31.

Zalocusky, K. A. et al. Nucleus accumbens D2R cells signal prior outcomes and control risky decision-making. *Nature* **531**, 642–646 (2016).

32. 32.

Fenno, L. E. et al. Targeting cells with single vectors using multiple-feature Boolean logic. *Nat. Methods* **11**, 763–772 (2014).

33. 33.

Chon, U., Vanselow, D. J., Cheng, K. C. & Kim, Y. Enhanced and unified anatomical labeling for a common mouse brain atlas. *Nat. Commun.* **10**, 5067 (2019).

34. 34.

Stauffer, W., Sheng, H. & Lim, H. N. EzColocalization: an ImageJ plug-in for visualizing and measuring colocalization in cells and organisms. *Sci. Rep.* **8**, 15764 (2018).

35. 35.

Brunetti, M. et al. Design and fabrication of ultralight weight, adjustable multi-electrode probes for electrophysiological recordings in mice. *J. Vis. Exp.* **91**, e51675 (2014).

36. 36.

English, D. F. et al. Pyramidal cell–interneuron circuit architecture and dynamics in hippocampal networks. *Neuron* **96**, 505–520 (2017).

37. 37.

Wimmer, R. D. et al. Thalamic control of sensory selection in divided attention. *Nature* **526**, 705–709 (2015).

38. 38.

Meyers, E. M. The neural decoding toolbox. *Front. Neuroinform.* **7**, 8 (2013).

39. 39.

Gold, J. I. & Shadlen, M. N. The neural basis of decision making. *Annu. Rev. Neurosci.* **30**, 535–574 (2007).

40. 40.

Wang, X. J. Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* **36**, 955–968 (2002).

41. 41.

Najafi, F. et al. Excitatory and inhibitory subnetworks are equally selective during decision-making and emerge simultaneously during learning. *Neuron* **105**, 165–179 (2020).

Acknowledgements

We thank all members of the Halassa laboratory for discussions, advice and support; I. Wickersham for providing viral tools for retrograde monosynaptic tracing; and M. Heiman for providing us with D2-cre mice. M.M.H. is supported by grants from the US National Institute of Mental Health (R01MH120118 and R01MH107680) and Pew Foundations. A.M. is supported by the Y. Eva Tan Fellowship.

Author information

Author notes

1. These authors contributed equally: Arghya Mukherjee, Norman H. Lam

Affiliations

1. McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA, USA

Arghya Mukherjee, Norman H. Lam, Ralf D. Wimmer & Michael M. Halassa

2. Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA

Arghya Mukherjee, Norman H. Lam, Ralf D. Wimmer & Michael M. Halassa

Contributions

A.M. collected and analysed anatomical, electrophysiological and behavioural data. R.D.W. collected electrophysiological data from behaving mice and analysed behavioural data. N.H.L. analysed anatomical, electrophysiological and behavioural data and also performed simulations with the spiking neural model. M.M.H. supervised the project and wrote the manuscript with contribution from A.M., N.H.L. and R.D.W.

Corresponding author

Correspondence to [Michael M. Halassa](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Laura Bradfield, Mathieu Wolff 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 Distinct effect of MD activation on PL activity compared to MGB on A1 and controls relevant to VIP \$\pm\$ mediation of MD-driven amplification of PL connectivity.](#)

a, Left: Cartoon of setup testing the role of MD thalamus activation on intra-PL activity. Right: Representative histology showing the expression of somatic ChR2 in PL contralateral to the recording site (top) and SSFO expression in the MD (bottom). Scale bar in μm : 200. **b**, Example rasters and PSTHs of a putative excitatory PL neuron showing an evoked response to intra-PL activation alone (left) and no change with concurrent MD activation (right). Blue ticks mark the period of contralateral PL stimulation. **c**, Population quantification of effect in **b** ($n = 151$ excitatory PL neurons from 4 mice, $***p = 1 \times 10^{-15}$, compared across groups, Wilcoxon signed-rank). **d**, Left: Same as in **a** except for stimulation of auditory thalamus (MGB) and measuring evoked responses in the auditory cortex (A1). Right: Representative histology of somatic ChR2 expression in A1 contralateral to the recording site (top) and SSFO expression in the MGB (bottom). Scale bar in μm : 200. **e**, An excitatory A1 neuron showing response to intra-A1 activation alone (left) and an amplification of its response with concurrent MGB activation (right). Blue ticks mark the period of contralateral A1 stimulation. **f**, Population quantification of effect in **e** ($n = 196$ neurons from 3 mice, $p = 0.6802$ (NS), compared across groups, Wilcoxon signed-rank test). **g**, ChR2 stimulation in PL and A1 respectively evoke comparable responses in the contralateral PL and A1 ($n = 151$ and 196 excitatory units recorded from the PL of 4 animals and the A1 of 3 animals respectively; $p = 1 \times 10^{-5}$, compared to baseline, $p = 0.2532$ (NS), across groups, Mann-Whitney U test). **h**, MD activation induced increase in baseline spike rates of PL inhibitory neurons is unaffected by concurrent suppression of PL VIP+ neurons ($n = 48$ neurons; $^+p = 0.0243$, $^{++}p = 0.0086$, compared to baseline, Mann-Whitney U test, $p = 0.7555$ (NS), compared across groups, Wilcoxon signed-rank test). **i**, Example of a putative PL excitatory neuron showing a response to intra-PL activation alone (top), which remains unaffected by concurrent suppression of PL VIP+ neurons (bottom). Blue tick marks the period of contralateral PL stimulation and yellow bar marks the duration of VIP+ inactivation. **j**, Quantification of effect in **i** ($n = 151$ neurons from 4 mice, $^{+++}p = 1.0 \times 10^{-5}$ (VIP int.), $^{+++}p = 1.0 \times 10^{-5}$ (VIP sil.) compared to baseline; Mann-Whitney U test; $p = 0.5956$ (NS), compared across groups; Wilcoxon sign ranked test). **k**, Optical inactivation of PL VIP+ interneurons do not affect baseline spike rates of putative excitatory or inhibitory neurons in the PL ($n = 385$ excitatory (RS) and $n = 98$ inhibitory (FS) neurons from 4 mice; $p=0.0955$ (NS, RS); $p=0.4933$ (NS, FS), compared to baseline, Mann-Whitney U test). All statistical tests are two-tailed. For box plots **g**, **h**, **j**, **k** boundaries, 25–75th percentiles; midline, median; whiskers, minimum–maximum

[Source data](#).

[Extended Data Fig. 2 D2 receptor mRNA expression in the mouse MD.](#)

a, Representative histology showing the expression of D2 receptor mRNA in the lateral MD using fluorescent in situ hybridization. **b**, Negative control probe does not result in staining. Scale bar = 300 μ m **c**, Heat map quantifying expression of D2 receptor mRNA across all MD sections (n = 8 sections for each condition from 2 mice).

Extended Data Fig. 3 Supportive evidence for anatomical and functional segregation of the two MD cell types.

a-c, Left: Starter neurons (arrowheads) in the PL of VIP-cre (**a**), PV-cre (**b**) and SST-cre (**c**) mice for monosynaptic retrograde tracing using rabies viruses. Right: Starter neurons identified by co-expression of TVA fused to GFP (top) and blue fluorescent protein from rabies viruses (bottom) in VIP, PV, and SST neurons, respectively. Scale bars in μ m: 200 μ m(left), 30 μ m (right). **d**, Left: Representative images of MD neurons that monosynaptically target PL SST+ interneurons. Scale Bar in μ m: 200. (Note a lack of preferential localization within the MDI) Right: 3D plot of the anatomical location of SST-projecting MDI neurons (n = 86 SST projecting neurons from 3 mice). **e**, Anatomical separation between SST-projecting MD neurons and VIP-/PV- projecting MD neurons quantified as high misclassification using KNN clustering. **f**, MD_{D2} and MD_{GRIK4} neuronal locations show low misclassification compared to VIP- and PV-projecting MD neurons respectively. **g**, Example of a PL neuron showing amplification of evoked responses through concurrent intra-PL and MD_{D2} optical stimulation (left), but not when intra-PL stimulation is combined with MD_{GRIK4} stimulation (right). **h**, Examples of excitatory (RS) and inhibitory (FS) PL neurons showing, respectively, suppression and increase in spike rates with optical activation of MD_{GRIK4} neurons but not with activation of MD_{D2} neurons. **i**, Parametric activation of MD_{GRIK4} , but not MD_{D2} , neurons increase spike rates of PL inhibitory neurons (n = 68 and n = 78 PL-FS neurons from 3 animals each of MD_{D2} and MD_{GRIK4} respectively; p = 0.874, For MD_{GRIK4} p = 0.556, *p = 0.0387, ***p = 9.36 x 10⁻⁶, *p = 0.0387 respectively for laser powers 0.65, 1.3, 3.5 and 7.0 mW/mm²; Mann-Whitney U test compared to baseline). **j**, left: D2 specific promoter (D2SP) driven expression of mCherry + (CreON) and co-expression of eYFP (CreOFF) in Cre-negative neurons using a Cre - Out intersectional strategy labels two populations similar to D2-cre and Grik4-cre, but in WT animal. Right: Magnified images showing mCherry (D2SP+) and eYFP (Cre negative) neurons. Scale bar = 200 μ m **k**, Consistent anatomical similarity between MD_{D2SP} and MD_{D2} populations and a corresponding segregation between MD_{D2SP} and MD_{GRIK4} neurons, quantified using representational similarity analysis (n = 95 cells from 2 animals for MD_{D2SP}). **l**, A comparable similarity and segregation as shown in (**k**) is found when comparing MD_{D2SP} neurons to VIP-projecting and PV projecting neurons. **m**, top row: MD_{D2}

Cre-expressing neurons (MD_{D2+}) labelled with GFP have extremely sparse Grik4 protein expression (IHC) compared to MD_{D2} Cre-negative (MD_{D2-} neurons (middle row) or MD_{GRIK4} expressing neurons (bottom row). Scale bar = 3 μ m **n**, Quantification of data ($n = 116$ MD_{D2+} , 106 MD_{D2-} and 124 MD_{GRIK4} neurons from 2 animals, demonstrating substantial Grik4 immunolabelling overlap between D2- and Grik4+ neural populations (not significantly different), but both being different from the D2+ population. *** $p = 0.0001$ for both comparisons, Kruskal Wallis test). **o**, Direct comparison of Grik4 immunolabelling across D2+ and D2- neurons (thresholded by the lowest 10th percentile of this analysis puts an upper bound estimate of 15% overlap between the D2+ and Grik4+ population. ‘positive control’ Grik4+ neurons). All statistical tests are two-tailed. For box plot **n** boundaries, 25–75th percentiles; midline, median; whiskers, minimum–maximum. Data are presented as mean \pm SEM for **i**

[Source data.](#)

Extended Data Fig. 4 mGRASP and synaptophysin labelling provide evidence for output segregation of the two MD cell types.

a, Cartoon depicting strategy to label cell type specific MD \rightarrow PL thalamocortical synapses using mGRASP. The pre mGRASP component is virally expressed in MD_{D2} or MD_{GRIK4} neurons in the respective Cre lines while the post mGRASP component is ubiquitously expressed in the PL. MD_{D2} or MD_{GRIK4} specific mGRASP synapses onto VIP vs PV neurons in the PL are identified by immunohistochemistry guided detection of PV and VIP neurons expressing post mGRASP in the PL. **b–c**, Left: Representative images of MD_{D2} (**b**) and MD_{GRIK4} (**c**) neurons expressing pre mGRASP in the MD of D2-cre and GRIK4-cre mice respectively. Right: Ubiquitous expression of post mGRASP+ neurons detected by TdTomato fluorescence in the PL of D2-cre (**b**) and GRIK4-cre (**c**) mice. Scale bar in μ m: 200. **d**, Left to right: Examples of PL VIP+ neurons showing post mGRASP expression (magenta), VIP expression detected via immunohistochemistry (yellow) and mGRASP+ synapses from MD_{D2} (cyan dots, top row) or MD_{GRIK4} (cyan, dots, bottom row) neurons. **e**, Same as in **d**, for PL PV+ neurons. Scale bars in μ m: 3 μ m. **f**, Representative images showing layer-wise termination of synapses from MD_{D2} (left) and MD_{GRIK4} (right) neurons in the PL, labelled with virally expressed GFP fused to synaptic protein (synaptophysin). Scale bar in μ m: 100. **g**, MD_{D2} neurons terminate in L1 of the PL with a higher frequency compared to MD_{GRIK4} neurons ($n = 12$ sections each from 3 D2-cre and 3 GRIK4-cre mice, * $p = 0.0253$, * $p = 0.039$, two-tailed Mann-Whitney U test comparing 50 μ m bins from the pial surface across groups). All statistical tests are two-tailed. Data are presented as mean \pm SEM for **g**

[Source data](#).

Extended Data Fig. 5 Differential engagement of the two MD cell types in a PL-dependent behaviour.

a, Schematic illustration of PL dependent attention control task (see [Methods](#)). In brief, on each trial, animals have to remember a 100ms HP or a LP auditory cue over a delay period to execute the corresponding rule (HP - attend to audition vs LP – attend to vision) and make a choice to either follow a target stimulus (auditory vs visual) to collect a milk reward. Orange bars highlight the two epochs where MD is optically inhibited across the experiment types **b**, Illustration of PL independent task with a 2AFC design, where animals have to respond to the side a LED light target was presented (without a distractor). **c**, Example sessions plotting the performance of an animal in the PL dependent (black) and PL independent (grey) versions of the task. **d**, Well trained animals show a daily ramp-up of performance in the first 30 trials of the PL dependent (but not PL independent) task, starting from chance (trials 1 to 10) and progressing onto performance > 0.7 proportion correct ('task engagement', trials 20 to 30) ($n = 12$ sessions over 4 mice, *** $p = 0.0002$). **e, f**, optical MD inactivation (yellow) during the cueing period of the first 30 trials in a session prolongs the ramp up to task engagement in the PL dependent task (**e**) but not the PL independent task ($n = 12$ sessions over 4 mice, $p = 0.0008$, Kolmogorov-Smirnov test). **g**, Optical inactivation of MD_{D2} neurons have no effect on number of trials taken to task engagement in PL dependent attention control task ($n = 12$ sessions over 4 D2-cre mice, $p = 0.466$ (NS), comparing across laser ON vs laser OFF sessions). **h**, Optical inactivation of MD_{GRIK4} neurons increase the number of trials required to reach task engagement in the attention control task ($n = 12$ sessions over 4 GRIK4-cre mice, *** $p = 2.2 \times 10^{-6}$, comparing across laser ON vs laser OFF sessions). **i**, Optical MD inactivation during the cueing period of the first 30 trials delays task engagement (performance at > 0.7 proportion correct; $n = 12$ sessions over 4 animals, *** $p = 1.5 \times 10^{-6}$). **j-k**, Optical MD_{GRIK4} (but not MD_{D2}) inactivation recapitulates the effect in **b** ($n = 12$ sessions each from 4 D2-cre and 4 GRIK4-cre mice; *** $p = 2.9 \times 10^{-5}$, $p = 0.7657$ (NS);). **l**, Optical MD inactivation in the delay period of <20% of trials diminish performance on those trials ($n = 12$ sessions over 4 mice, *** $p = 7.4 \times 10^{-7}$). **m-n**, Optical inactivation of MD_{D2} (but not MD_{GRIK4}) recapitulates the effect in **e**. ($n = 12$ sessions each from 4 D2-cre and 4 GRIK4-cre mice respectively; $p = 0.3394$ (NS), *** $p = 2.95 \times 10^{-6}$). two-tailed Mann-Whitney U test. All statistical tests are two-tailed. For box plots **i-n** boundaries, 25–75th percentiles; midline, median; whiskers, minimum–maximum. Data are presented as mean \pm SEM for **d-h**

[Source data](#).

Extended Data Fig. 6 Controls that clarify behavioural strategy and weighing of evidence in the attentional control task with input uncertainty.

a, Behavioural validation of animals using correct task execution strategies (see Fig. 3a). Omitting the distractor on a subset of interleaved trials (15%, valid target only) during choice 2 (Fig. 3a) had no effect on behaviour ($n = 41$ sessions over 6 mice; $p = 0.859$ (NS, visual); $p = 0.728$ (NS, auditory); Mann-Whitney U test). Omitting the target on a similar subset of trials (invalid target only) reduced performance accuracy down to chance level (** $p = 0.00649$ (visual); ** $p = 0.00216$ (auditory); Mann-Whitney U test). Combined, these data indicate that animals did not adopt a pro-anti strategy based on a single target (vision or audition). **b**, Average performance on uninformative trials is comparable when the underlying sequences are composed of only broadband white noise pulses ($n = 11$ sessions over 6 mice; pure 0, $p = 0.353$ (NS), binomial test) or informative cues with zero overall net evidence (net 0, $p = 0.690$ (NS), binomial test). **c**, Regression analysis shows that evidence in the early half and the late half of the cueing sequence contribute equally to animal choice behaviour ($n = 54$ sessions over 6 mice; early *** $p = 2.98 \times 10^{-7}$, $t = 5.12$ compared to 0; late *** $p = 5.03 \times 10^{-10}$, $t = 6.22$ compared to 0; early vs late $p = 0.368$ (NS), $t = 0.901$; degree of freedom=3946; student's t-test). **d**, Full psychometric functions of individual mice in the conflict-driven input uncertainty task. Performance accuracy in the distributed cue task with input uncertainty due to cueing conflict, separated by animals. For each animal, performance accuracy consistently diminishes with increased cueing conflict (black traces, top and bottom row), while optical PL inactivation (blue traces, top row) during the cueing period strongly suppresses performance regardless of input uncertainty (M1: *** $p = 1.60 \times 10^{-10}$ (relative conflict = 0), *** $p = 1.89 \times 10^{-6}$ (relative conflict = 0.28); M2: *** $p = 6.40 \times 10^{-8}$ (relative conflict = 0), ** $p = 0.00366$ (relative conflict = 0.28); M3: ** $p = 0.00149$ (relative conflict = 0), *** $p = 2.38 \times 10^{-6}$ (relative conflict = 0.28), *** $p = 2.85 \times 10^{-4}$ (relative conflict = 0.5); M4: *** $p = 1.19 \times 10^{-4}$ (relative conflict = 0), *** $p = 4.83 \times 10^{-4}$ (relative conflict = 0.28), * $p = 0.0361$ (relative conflict = 0.5); M5: *** $p = 4.18 \times 10^{-6}$ (relative conflict = 0.28), *** $p = 8.86 \times 10^{-4}$ (relative conflict = 0.5), * $p = 0.0130$ (relative conflict = 0.67); chi-squared test). In contrast, Optical MD inactivation (yellow traces, bottom row) during the cueing period reduces performance more strongly on high conflict trials than on low conflict trials, consistently across animals (M1: ** $p = 0.00345$ (relative conflict = 0.28), *** $p = 2.27 \times 10^{-4}$ (relative conflict = 0.5); M2: $p = 0.111$ (NS; relative conflict = 0.28), ** $p = 0.00676$ (relative conflict = 0.5); M3: * $p = 0.0208$ (relative conflict = 0.28), ** $p = 0.00556$ (relative conflict = 0.5); M4: $p = 0.426$ (NS; relative conflict = 0.28), ** $p = 0.00651$ (relative conflict = 0.5); M5: * $p = 0.0486$ (relative conflict = 0.28), ** $p = 0.00107$ (relative conflict = 0.5), ** $p = 0.00227$ (relative conflict = 0.67); chi-squared test). Inset in each panel

highlights the effect of PL/MD inactivation on trials with low (0.28) and high (0.5) conflict. All statistical tests are two-tailed. For box plots **a-c** and insets in **d**, boundaries, 25–75th percentiles; midline, median; whiskers, minimum–maximum. Data are presented as mean \pm SEM for **d**

[Source data.](#)

Extended Data Fig. 7 Extended analysis and relevant controls of PL RS and FS cells, and differential encoding of task relevant variables across the MD and PL.

a, Two example excitatory PL neurons shown in Fig. 3c, sorted by momentary cue (cue-sorted) and attentional choice (choice-sorted). The earlier-responding neuron (left) shows selectivity to momentary cue and the later-responding neuron (right) shows selectivity to the attentional choice ($***p = 6.17 \times 10^{-4}$; $*p = 0.0157$; Mann-Whitney U test). In contrast, there are weak choice selectivity for the earlier-responding neuron and weak cue selectivity for the later-responding neuron. **b**, Quantification of PL population selectivity to momentary cue (top) and attentional choice (bottom) using linear decoding (n=1112 neurons from 7 mice). Note that population cue selectivity is strong early on but gradually decreases, while population choice selectivity peaks late in the cueing period. **c**, Quantification of PL population selectivity to momentary cue (top) and attentional choice (bottom) using mutual information. **d**, Example putative inhibitory fast spiking neuron, showing higher firing rate for trials with high conflict, and little attentional choice selectivity. This neuron shows similar selectivity to the example conflict-preferring MD neuron (Fig. 3f). **e**, Quantification of selectivity of putative inhibitory fast spiking neuron population in PL to momentary cue (top) and attentional choice (bottom) using linear decoding (n = 104 neurons from 7 mice). The selectivity for both cue and choice are weak compared to the putative excitatory neuron population (**b**). **f**, Quantification of conflict selectivity of putative inhibitory fast spiking neuron population in PL using linear decoding, showing strong conflict selectivity. **g**, Quantification of PL and MD population selectivity to conflict (top) and attentional choice (bottom) using linear decoding. MD population demonstrates strong conflict and weak choice selectivity (n = 2669 neurons from 7 mice), while PL population demonstrates strong choice and weak conflict selectivity. **h**, Choice modulated PL neurons demonstrate moderate conflict selectivity ($*p = 0.046$ choice, $*p = 0.02$ choice; permutation test). In contrast, conflict modulated MD neurons have no choice selectivity ($***p = 0.0005$ conflict, $p=0.695$ choice (NS); permutation test). n=50 most modulated neurons each. **i**, Optical inactivation of PL and MD result in distinct impairments in task performance across different levels of conflict driven input uncertainty. The magnitude of optical PL inactivation is titrated to match the task performance on low conflict trials with optical MD inactivation. PL inactivation results in comparable impairments in performance

accuracy across low and high conflict trials, while MD inactivation has a stronger effect on high conflict trials compared to low conflict trials ($n = 13$ sessions from 5 mice; *** $p < 0.001$, Mann-Whitney U test). **j**, MD deafferentiation during the cueing period impairs performance more strongly on high conflict trials than on low conflict trials ($n = 37$ sessions over 6 mice; *** $p = 4.84 \times 10^{-6}$ (relative conflict = 0.28); *** $p = 1 \times 10^{-15}$ (relative conflict=0.5); chi-squared test), similar to optical MD inactivation (Fig. 3e). **k**, Quantification of MD population conflict selectivity and PL population choice selectivity. MD deafferentiation annihilates MD conflict classification accuracy and weakens PL choice classification accuracy ($n = 386$ putative excitatory neurons and $n = 666$ MD neurons from 3 mice; ** $p = 0.005$ (MD conflict, Laser OFF); $p = 0.96$ (NS, MD conflict, Laser ON); ** $p = 0.0042$ (MD conflict, Laser OFF vs ON); ** $p = 0.005$ (PL choice, Laser OFF); * $p = 0.048$ (PL choice, Laser ON); * $p = 0.012$ (PL choice, Laser OFF vs ON); permutation test). **l**, MD deafferentation result in lowered firing rate in MD (** $p = 9.32 \times 10^{-11}$) and higher firing rate in PL excitatory neurons (* $p = 0.0231$; Wilcoxon signed-rank test). Data is pooled over conflict-preferring MD neurons ($n = 201$ neurons), and choice-selective PL neurons ($n = 85$ neurons.). **m**, MD neurons respond to conflict earlier in time compared to PL excitatory neurons (* $p = 0.0289$; Mann-Whitney U test). Shown are the latency to reach maximum regression coefficient after the conflict signal emerges. **n**, Data in Fig. 3e reorganized, highlighting the effect of MD inhibition on trials with low (0.28) and high (0.5) conflict. **o, p**, A mean-field neural model, which describes choice accumulation in the PL recaptures experimental data in **n** ($n = 2,000$ trials, * $p = 0.0137$; *** $p = 1.18 \times 10^{-6}$; chi-squared test). **q**, Data in **j** reorganized, highlighting the effect of optical inhibition of PL→MD terminals on trials with low and high conflict. **r**, Mean-field neural model (see Extended Data Fig. 7o) captures the effect of inhibition of PL→MD terminals on task performance ($n = 2,000$ trials, * $p = 0.0189$; *** $p = 4.90 \times 10^{-6}$; chi-squared test). All statistical tests are two-tailed. For box plots **h, k-n, q** boundaries, 25–75th percentiles; midline, median; whiskers, minimum–maximum. Data are presented as mean ± SEM for **i, j, p, r**, and mean ± CI for **b, c, e-g**

[Source data](#).

[Extended Data Fig. 8 Basic and extended mean-field models.](#)

a, Schematic of the mean-field neural model that describes generic MD inactivation results (see Extended Data Fig. 7o). The model describes two PL populations that receive separate inputs corresponding to the cues in favour of the two attentional rules (HP - attend to vision or LP – attend to audition). Each population has strong recurrent self-excitation and net inhibition on the other population. The MD component of the model receives inputs from the PL (see Extended Data Fig. 7) and is activated by conflict to inhibit the two PL populations. **b**, Example model decision variables in a trial early biased to the wrong attentional choice, demonstrating how MD-mediated

suppression may improve performance of the model. When MD is intact (left), strong early evidence to the wrong choice (high-pass in this example; cueing sequence in inset) increases the decision variable of the non-preferred population early on, but the preferred population prevails when the preferred stimulus dominates in the latter half of the cueing sequence. On the other hand, in the absence of MD conflict-driven suppression of cue integration in the PL (right), the early non-preferred inputs drive the non-preferred population to maintain high activity, suppressing the preferred population's response to late inputs. **c**, Schematic of the mean-field neural model incorporating the two cell types, where MD_{GRIK4} is conflict-activated and suppresses PL, and MD_{D2} is conflict-suppressed and amplifies PL recurrence. MD_{D2} results in enhanced gain of the PL input-output function (bottom). **d**, Example model decision variables for high conflict trials, with (left) and without (right) MD_{D2} . Increased PL recurrence due to MD_{D2} results in larger response to input cues. However, the effect is less pronounced for preferred cues as the population activity and decision variable saturate with inputs. As a result, the larger response to input cues asymmetrically favours the non-preferred population, and the separation between preferred and non-preferred activity is larger without MD_{D2} (shown are median over 1,000 trials). **e**, Example model decision variables for low signal sparse trials (Fig. 4), with (left) and without (right) MD_{D2} module. Increased PL recurrence due to MD_{D2} allows amplified response of the preferred population to sparse input cues, but minimally affects the non-preferred population which receives no input cues. As such, MD_{D2} results in a larger separation between preferred and non-preferred activity (shown are median over 1,000 trials).

Extended Data Fig. 9 Untagged neurons in the tagging experiments are no different than generic recordings, and optical inhibition of terminals of the two cell types replicates cell body inactivation.

a, (Top Left) Schematic of optogenetic tagging and identification of MD_{D2} and MD_{GRIK4} neurons. MD_{D2} or MD_{GRIK4} neurons are tagged with NpHR3.0 and identified via light activated spike rate suppression. (Bottom) Example tagged neuronal response to NpHR3.0 activation. (Right) Tagged neurons from one mouse (red) are identified using k-means clustering (features: change in firing rate, proportion of trials suppressed, and half-time to recover from suppression ($n = 262$ total number of neurons)). **b**, Relative fraction of all MD neurons from GRIK4-cre mice that are conflict-preferring vs. non-preferring are comparable to that of wild-type animals (Fig. 3g) ($n = 91$ neurons from 3 mice; $p = 0.429$ (NS), chi-squared test). Note that tagged MD_{GRIK4} neurons are significantly more conflict-preferring compared to the whole population (Fig. 4b) ($p = 0.0175$; chi-squared test). **c**, Relative fraction of all MD neurons from D2-cre mice that are conflict-preferring vs. non-preferring, are also

comparable to that of wild-type animals (Fig. 3g) ($n = 95$ neurons from 3 mice; $p = 0.166$ (NS), chi-squared test). Note, that tagged MD_{D2} neurons are significantly more conflict-non-preferring (Fig. 4d) ($p=1.34 \times 10^{-4}$; chi-squared test). **d**, Optical inhibition of MD_{GRIK4} terminals in the PL recapitulates the loss in task accuracy across low and high conflict trials as seen with optical MD_{GRIK4} inactivation (Fig. 4e; $n = 20$ sessions over 4 GRIK4-cre mice, $*p = 0.0199$, $***p = 0.0002$; Mann-Whitney U test). **e**, Optical inhibition of MD_{D2} terminals in the PL enhances performance accuracy on trials with high cueing conflict, similar to the effect of optical MD_{D2} inactivation (Fig. 4f; $n = 20$ sessions over 4 D2-cre mice, $p = 0.3941$ (NS), $**p = 0.0023$; Mann-Whitney U test. **f**, Schematic of micro-drive bottom piece and the 3x3 grid organization of the tetrode array for MD recordings. **g**, Summary of the density of tagged neurons on the medial-lateral axis separated by animal. We show the result for 2 Grik4-cre (top and bottom) and 2 D2-cre animals ((top and bottom rows) that have enough numbers of tagged neurons. All statistical tests are two-tailed. For box plots **b** – **e** boundaries, 25–75th percentiles; midline, median; whiskers, minimum–maximum

[Source data](#).

[Extended Data Fig. 10 Full psychometric functions of individual mice in the sparseness task.](#)

Performance accuracy in the distributed cue task with input uncertainty due to cueing sparseness, separated by animals. For each animal, performance accuracy consistently diminishes with decreasing signal (black traces), while optical inhibition of PL→MD terminals (yellow traces) during the cueing period generally reduces performance more strongly on low signal trials than on high signal trials (M1: $p = 0.644$ (NS; relative signal = 0.25), $*p = 0.0348$ (relative signal = 0.13); M2: $p = 0.676$ (NS; relative signal = 0.25), $*p = 0.0426$ (relative signal = 0.13); M3: $p = 0.139$ (NS, relative signal = 0.25), $p = 0.0604$ (NS; relative signal = 0.13); M4: $p = 0.343$ (NS; relative signal = 0.25), $**p = 0.0251$ (relative signal = 0.13); chi-squared test). Inset in each panel highlights the inactivation effect on trials with high (0.25) and low (0.13) signal. All statistical tests are two-tailed. For inset box plots, boundaries, 25–75th percentiles; midline, median; whiskers, minimum–maximum. Data are presented as mean ± SEM

[Source data](#).

Supplementary information

[Supplementary Information](#)

This file contains Supplementary Table 1, a Supplementary Introduction, two Supplementary Notes, a Supplementary Discussion, and Supplementary References.

[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. 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

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

Mukherjee, A., Lam, N.H., Wimmer, R.D. *et al.* Thalamic circuits for independent control of prefrontal signal and noise. *Nature* **600**, 100–104 (2021).
<https://doi.org/10.1038/s41586-021-04056-3>

- Received: 21 March 2021
- Accepted: 27 September 2021
- Published: 06 October 2021
- Issue Date: 02 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04056-3>

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/s41586-021-04056-3>

- Article
- Open Access
- [Published: 03 November 2021](#)

Terrestrial-type nitrogen-fixing symbiosis between seagrass and a marine bacterium

- [Wiebke Mohr](#) [ORCID: orcid.org/0000-0002-1126-1455¹](#),
- [Nadine Lehnen¹](#),
- [Soeren Ahmerkamp¹](#),
- [Hannah K. Marchant](#) [ORCID: orcid.org/0000-0002-1482-9165¹](#),
- [Jon S. Graf](#) [ORCID: orcid.org/0000-0003-1058-9480¹](#),
- [Bernhard Tschitschko¹](#),
- [Pelin Yilmaz^{1 nAff4}](#),
- [Sten Littmann¹](#),
- [Harald Gruber-Vodicka](#) [ORCID: orcid.org/0000-0001-5819-1549¹](#),
- [Nikolaus Leisch¹](#),
- [Miriam Weber²](#),
- [Christian Lott](#) [ORCID: orcid.org/0000-0003-4799-6012²](#),
- [Carsten J. Schubert](#) [ORCID: orcid.org/0000-0003-1668-5967³](#),
- [Jana Milucka](#) [ORCID: orcid.org/0000-0003-4013-230X¹](#) &
- [Marcel M. M. Kuypers](#) [ORCID: orcid.org/0000-0001-7991-5091¹](#)

Nature volume 600, pages 105–109 (2021)

- 11k Accesses
- 1 Citations
- 315 Altmetric
- [Metrics details](#)

Subjects

- [Biogeochemistry](#)
- [Environmental microbiology](#)
- [Microbial ecology](#)

Abstract

Symbiotic N₂-fixing microorganisms have a crucial role in the assimilation of nitrogen by eukaryotes in nitrogen-limited environments^{1,2,3}. Particularly among land plants, N₂-fixing symbionts occur in a variety of distantly related plant lineages and often involve an intimate association between host and symbiont^{2,4}. Descriptions of such intimate symbioses are lacking for seagrasses, which evolved around 100 million years ago from terrestrial flowering plants that migrated back to the sea⁵. Here we describe an N₂-fixing symbiont, ‘*Candidatus Celerinatantimonas neptuna*’, that lives inside seagrass root tissue, where it provides ammonia and amino acids to its host in exchange for sugars. As such, this symbiosis is reminiscent of terrestrial N₂-fixing plant symbioses. The symbiosis between *Ca. C. neptuna* and its host *Posidonia oceanica* enables highly productive seagrass meadows to thrive in the nitrogen-limited Mediterranean Sea. Relatives of *Ca. C. neptuna* occur worldwide in coastal ecosystems, in which they may form similar symbioses with other seagrasses and saltmarsh plants. Just like N₂-fixing microorganisms might have aided the colonization of nitrogen-poor soils by early land plants⁶, the ancestors of *Ca. C. neptuna* and its relatives probably enabled flowering plants to invade nitrogen-poor marine habitats, where they formed extremely efficient blue carbon ecosystems⁷.

[Download PDF](#)

Main

Seagrasses form vast meadows in coastal environments around the globe, providing a breeding ground and food for fish and protection from coastal erosion^{8,9,10}. Furthermore, seagrass meadows have a major role in the drawdown of carbon dioxide (CO₂) due to their large biomass production, which matches that of the most prolific terrestrial ecosystems¹¹. The nitrogen (N) that is required for this biomass production is generally believed to be taken up by the seagrasses through leaves and roots from the surrounding environment¹². As many seagrasses are found in oligotrophic, N-depleted environments, the seagrass productivity is thought to be at least partially supported by N originating from microbial N₂ fixation^{13,14}. The N₂ fixation is generally assumed to take place in the surrounding sediment by microorganisms that are associated with either the rhizosphere/rhizoplane^{14,15} or with animals residing in

the seagrass meadows¹⁶. By contrast, terrestrial plants that thrive in N-poor habitats often enter more intimate, mutually beneficial interactions with N₂-fixing bacteria^{2,17,18}. The bacteria usually reside within the plant tissue, and the interaction between these symbionts and their plant hosts is genetically complex¹⁹, relying on a sophisticated communication and metabolite exchange²⁰. Here, we report the discovery of a marine N₂-fixing bacterium that lives inside the root tissue of the seagrass *P. oceanica*, exhibiting features that are reminiscent of terrestrial plant N₂-fixing symbionts.

Growth and N₂ fixation in seagrass meadows

P. oceanica from the oligotrophic Mediterranean Sea is one of the most productive seagrasses²¹. At our study site, *P. oceanica* forms dense meadows with around 600 shoots per m² (Fig. 1a). In situ measurements taken during summer 2019 revealed that these meadows had high rates of photosynthesis, resulting in a net primary production of around 50 mmol m⁻² d⁻¹ CO₂ fixed (Fig. 1b and Extended Data Fig. 1). The primary production was comparable to that reported for other *P. oceanica* meadows in the Mediterranean Sea¹¹. By contrast, the non-vegetated sandy sediments surrounding the meadows were a net source of CO₂, releasing ~6 mmol m⁻² d⁻¹ CO₂, despite the presence of benthic algal biofilms (Fig. 1b and Extended Data Fig. 1). The high primary production associated with the *P. oceanica* meadows occurred in the absence of detectable nutrient-N in the water column (Extended Data Table 1a).

Fig. 1: Productivity, root-associated N₂ fixation and N transfer to leaves.

 **figure1**

a, *P. oceanica* meadow in the sandy sediments of Fetovaia Bay, Elba (Italy). **b**, Areal net O₂ fluxes in the seagrass meadow (green) and neighbouring non-vegetated sandy sediments (light brown). The four individual bars represent the daily mean values of averaged hourly fluxes, and the error bars indicate the variability of night-time and daytime fluxes (propagated s.d.). The four measurements, two seagrass (June 2019) and two sand (June 2019 and September 2018), were performed on four different days

(see Extended Data Fig. 1 for examples of a daily cycle). **c**, Illustration of the incubation set-up enabling the detection of the transfer of freshly fixed (^{15}N -enriched) N from the roots to the leaves (orange arrow). **d, e**, Root-associated N_2 fixation (e) and N transfer to the leaves (d) of *P. oceanica* from April to September. The bars, boxes and error bars represent the mean values, 25th and 75th percentiles, and s.e.m. The number of measured plant pieces (*n*) is indicated in parentheses above the data points with the number of incubated plants indicated below each month. DW, dry weight. Note the tenfold difference in scale between d and e.

Source Data

Our $^{15}\text{N}_2$ -labelling experiments with roots and rhizomes of whole plants (Fig. 1c) revealed seasonal N_2 fixation activity that was mainly associated with the roots of *P. oceanica* (Fig. 1e). The rates of root-associated N_2 fixation were highest during summer, when inorganic-N concentrations in the water column were below the detection limit (Figs. 1e and 2a and Extended Data Table 1a). In spring, when inorganic N was detectable, N_2 fixation rates were low to non-detectable (Fig. 1e and Extended Data Table 1a). Although the $^{15}\text{N}_2$ -labelling experiments were restricted to roots and rhizomes, leaf biomass was also enriched in ^{15}N in summer, indicating that newly fixed N was transferred from the roots to the leaves (Fig. 1d). This transfer was rapid, with up to around 20% of the freshly fixed N being assimilated into leaf biomass already within 24 h (Fig. 1d). Such rapid transfer to the leaves was previously reported for *Zostera marina*²². Taking into account all fixed N recovered in the different plant organs, root-associated N_2 fixation could fully support the measured in situ net plant biomass production in summer (Fig. 1b and [Supplementary Information](#)), which is the main growth season of *P. oceanica*²³. Furthermore, root-associated N_2 fixation is probably a source of N for the wider seagrass ecosystem, as indicated by elevated inorganic N concentrations in seagrass sediments relative to the surrounding, non-vegetated sediments (Extended Data Table 1b).

Fig. 2: Relative abundance and phylogeny of *Ca. C. neptuna*.

 **figure2**

a, The root-associated N₂ fixation rates of individual plants. Each symbol represents an individually measured root piece. **b**, The relative abundance of *Ca. C. neptuna* 16S rRNA gene sequence reads (magenta) in roots of individually analysed plants. One column represents one plant and each column corresponds to the measured N₂ fixation rates in **a**. **c**, Phylogeny of *Ca. C. neptuna* (bold, magenta) within the Celerinatimonadaceae based on 16S rRNA gene sequences (consensus tree; the scale bar is substitutions per site; the black dots indicate strong bootstrap support). Macrophytes and metazoa from which sequences were recovered are indicated on the right. Blue and magenta circles indicate the acquisition of abilities to fix N₂ and to

degrade pectin (Pec), respectively. *C. yamalensis*, *Celerinatantimonas yamalensis*; *T. hemprichii*, *Thalassia hemprichii*. Accession numbers and references are provided in Supplementary Data [1](#) and [2](#).

[Source Data](#)

An N₂-fixing root endophyte

Metagenomic sequencing revealed substantial differences between the microbiome of *P. oceanica* roots and the surrounding sediments (Extended Data Fig. [2](#)), indicating that, like other seagrasses^{[24](#)}, *P. oceanica* also has a specialized root microbiome. Moreover, on the basis of 16S rRNA amplicon data, the root microbial communities of N₂-fixing plants ([Methods](#)) differed substantially from plants with non-detectable N₂ fixation rates (Fig. [2b](#) and Extended Data Fig. [3](#)). The difference was largely driven by a single gammaproteobacterium belonging to the genus *Celerinatantimonas*, which was abundant in the roots of N₂-fixing plants and relatively rare in non-N₂-fixing plants (Fig. [2b](#) and Extended Data Fig. [3](#)). The closest cultured relative was *Celerinatantimonas diazotrophica* (~95% 16S rRNA gene similarity), an N₂-fixing bacterium isolated from saltmarsh grasses^{[25](#)} (Fig. [2c](#)). On the basis of thresholds for genus discernment^{[26](#)}, the bacterium recovered from *P. oceanica* roots represents a new species within the genus *Celerinatantimonas* ([Supplementary Information](#)), which we named *Candidatus Celerinatantimonas neptuna* (*Ca. C. neptuna*).

Specific 16S rRNA-targeted probes were designed to visualize *Ca. C. neptuna* cells in root sections using fluorescence in situ hybridization (FISH). Few *Ca. C. neptuna* cells were found inside the roots of non-N₂-fixing plants from spring (Extended Data Fig. [4](#)). By contrast, endophytic *Ca. C. neptuna* cells were abundant (~80% of 4',6-diamidino-2-phenylindole (DAPI) counts) throughout the root cortex and stele of N₂-fixing plants in summer (Fig. [3a–c](#) and Extended Data Figs. [4](#) and [5](#)). *Ca. C. neptuna* cells resided in the intercellular spaces as well as inside plant root cells (Fig. [3b,c](#) and Extended Data Fig. [5](#)).

Fig. 3: Distribution and N₂ fixation activity of *Ca. C. neptuna*.

 **figure3**

a, Epifluorescence image (stitched) of a root cross-section (overlay image of DAPI (blue) and autofluorescence (green/orange)). **b, c**, Images of *Ca. C. neptuna* cells and clusters (pinkish colour due to the overlap of the DAPI and FISH probe (orange) signals) in the middle cortex (**b**) and inner cortex (**c**). **d, e**, Correlative imaging (stitched images) of *Ca. C. neptuna* cells in the inner and middle cortex (**d**) and the corresponding nanoSIMS image showing ^{15}N enrichment (**e**). H, hypodermis; IC, inner cortex; MC, middle cortex; NA, natural abundance; OC, outer cortex; R, rhizoplane; S, stele. Scale bars, 150 μm (**a**) and 10 μm (**b–e**).

Single-cell measurements using nanoscale secondary ion mass spectrometry (nanoSIMS) provided direct evidence that *Ca. C. neptuna* fixed $^{15}\text{N}_2$ in the roots of *P. oceanica* in summer (Fig. 3d, e and Extended Data Fig. 4). The *P. oceanica* root tissue was also substantially enriched in ^{15}N , indicating that a substantial amount of freshly fixed N (up to ~98%) was transferred to the seagrass (Extended Data Fig. 4 and [Supplementary Information](#)).

Metabolic capacity of the seagrass endophyte

To gain insights into the metabolic interaction between *Ca. C. neptuna* and *P. oceanica*, we obtained the genome and transcriptome of *Ca. C. neptuna* from N₂-fixing plants. The ~4.3 Mb metagenome-assembled genome of *Ca. C. neptuna* comprised a single circular chromosome encoding all of the enzymes necessary for N₂ fixation (Extended Data Fig. 6). Genes coding for nitrogenase, the enzyme that catalyses the reduction of N₂ to ammonium, as well as proteins that transfer electrons to the nitrogenase (that is, *nifHDK/F*) were highly transcribed under N₂-fixing conditions (Fig. 4a and Extended Data Figs. 6 and 7). Some of the ammonium produced by *Ca. C. neptuna* was probably directly transferred to the seagrass (Fig. 4b). Furthermore, in analogy to terrestrial N₂-fixing plant symbioses²⁷, fixed N was also transferred in the form of amino acids. Glutamate, phenylalanine and leucine were probably transferred from *Ca. C. neptuna* to the seagrass roots, as indicated by the incorporation of ¹⁵N into these protein-bound amino acids (Extended Data Fig. 8 and [Supplementary Information](#)). In return, the seagrass may provide the amino acid GABA (4-aminobutyrate) or precursors (such as arginine or ornithine), analogous to some rhizobia-legume symbioses^{27,28,29}. Correspondingly, the *gabT* gene, which encodes an aminotransferase that catalyses the amino-group transfer from GABA to 2-oxoglutarate to yield glutamate, was among the most highly transcribed genes (Fig. 4a and Extended Data Fig. 6).

Fig. 4: Highly transcribed genes in *Ca. C. neptuna* and the proposed metabolic interaction between *Ca. C. neptuna* and *P. oceanica*.

 figure4

a, The circular *Ca. C. neptuna* genome with GC content (grey), GC skew (purple/green) and the average transcription of protein-coding genes plotted as transcripts per million (TPM) (orange; TPM values above 1,000 were cut off). Note that most of the highlighted genes related to key metabolic functions have average

TPM values of >1,000. A list of gene names and corresponding enzymes is provided in the [Supplementary Information](#). **b**, Schematic of the symbiotic interaction between *Ca. C. neptuna* (magenta) and *P. oceanica* indicating the transfer of fixed N from N₂ fixation (dark red arrows) and plant-derived sugars (black arrows); the potential for further plant growth promotion and defence mechanisms is also indicated. ROS, reactive oxygen species.

In addition to GABA, the seagrass probably provides sugars, based on the high transcription of genes encoding extracellular sucrose degradation (*sacB*), sugar-transport proteins (*ptsIH/crr*) and enzymes of the glycolysis pathway (*gapA1*, *pgk*, *eno1*, *eno2*, *pykF*, *fbaA*, *tpiA*) (Fig. 4a and Extended Data Fig. 7). Although genes involved in the uptake of dicarboxylic acids (*dctPQM*) were only moderately transcribed, *Ca. C. neptuna* might also receive C₄-dicarboxylates from its host, analogous to Rhizobia²⁰. Sugar oxidation in *Ca. C. neptuna* might proceed under microoxic and partly anoxic conditions as indicated by the low transcription of genes encoding the low-O₂-affinity *bo*-type terminal oxidase (*cyoABCDE*), high transcription of genes encoding the high-O₂-affinity *bdI*-type terminal oxidase (*cydAB*) and proteins involved in fermentation (*adh*, *pflB*) (Fig. 4a and Extended Data Fig. 7). Microoxic/anoxic conditions might easily develop in *P. oceanica* roots, which reside in anoxic sediments at our study site (Extended Data Fig. 1c) and other sites throughout the Mediterranean³⁰. Under such conditions, root endophytes depend on their host for oxygen supply, which may enable *P. oceanica* to regulate the proliferation of *Ca. C. neptuna* similarly to the manner in which legume hosts control Rhizobial growth³¹. Microoxic conditions would be favourable for the activity of the oxygen-sensitive nitrogenase of *Ca. C. neptuna*.

In many aspects, the genome of *Ca. C. neptuna* exhibits hallmarks of a facultative endophytic symbiont. Just like many terrestrial plant endophytes, *Ca. C. neptuna* might switch between free-living and host-associated stages³². Genes related to motility and attachment (*flaA* and *flp*) were highly transcribed (Fig. 4a and Extended Data Fig. 6), indicating active invasion and colonization of seagrass root tissue. High transcription of the quorum-sensing master regulator *luxR* indicates cell-to-cell communication and orchestration of *Ca. C. neptuna* population activity, which is also important for the establishment of rhizobia-legume symbioses³³. Furthermore, high transcription of genes related to peroxide detoxification (*dps*, *ahpC/F*) by *Ca. C. neptuna* (Fig. 4a and Extended Data Fig. 6) indicates that reactive oxygen species are produced by seagrasses as a defence mechanism that is comparable to responses by legume hosts³⁴. The genome also contains genes that are commonly found in endophytes³⁵, which are used in host–symbiont recognition, chemotaxis, degradation of plant cell-wall components, plant growth promotion and effector secretion ([Supplementary Information](#)). Although many of these traits are not unique to plant-

beneficial and/or endophytic microorganisms³⁶, many are deemed to be vital for establishing a beneficial association^{37,38}. On the basis of our combined results, *Ca. C. neptuna* is a plant-beneficial N₂-fixing endophyte (Fig. 4b and Extended Data Fig. 9) that is strikingly similar to those in terrestrial plants^{2,17,18}.

Acquisition of a marine N₂-fixing symbiont

Seagrasses evolved around 100 million years ago⁵ from terrestrial flowering plants that migrated back to the sea, where they had to adjust their physiology to a fully submerged lifestyle in salt water³⁹. During the transition to a marine plant, the root-associated microbiome of terrestrial origin was probably replaced by a marine one. Accordingly, many of the root-associated microorganisms of *P. oceanica* are typical marine benthic bacteria, such as sulfate reducers and sulfide oxidizers ([Supplementary Information](#)).

We can only speculate when and where *P. oceanica* acquired its marine N₂-fixing symbiont *Ca. C. neptuna*. The phylogeny of *Ca. C. neptuna* suggests that its ancestor was obtained in a coastal marine environment. Interestingly, the closest relative of *Ca. C. neptuna*—*C. diazotrophica*—was isolated from the roots of the saltmarsh grasses *Juncus* and *Spartina* (Fig. 2c). Moreover, 16S rRNA gene sequences belonging to another member of the genus *Celerinatantimonas* were recovered from the seagrass *Thalassia*. These two plant-associated members of the genus *Celerinatantimonas* may form symbioses with their hosts similar to the symbiosis that *Ca. C. neptuna* forms with its host, but a confirmation of their lifestyle is so far lacking. Members of the more deeply branching genera *Agarivorans*, *Aliagarivorans* and *Alginatibacterium*, which also belong to the family Celerinatantimonadaceae ([Supplementary Information](#)), have been found to be associated with macroalgae (Fig. 2c). Interestingly, the potential to fix N₂ is prevalent among the members of the family Celerinatantimonadaceae, whereas this potential is missing from known representatives of the neighbouring families Idiomarinaceae and Colwelliaceae (Supplementary Data 3), which typically do not associate with macrophytes. Thus, the last common ancestor of the family Celerinatantimonadaceae was probably associated with macroalgae, and the ability to fix N₂ may have been acquired as an adaptation to living on N-poor carbohydrates. This N₂-fixing ancestor diversified and formed a new lineage—the genus *Celerinatantimonas*—that was adapted to living in association with marine flowering plants such as seagrasses and saltmarsh grasses.

It appears that the colonization of seagrass may have required comparably few genomic adaptations as many traits that have been deemed to be vital for the establishment of a beneficial association with flowering plants^{35,38} were also found in

the genomes of related organisms (Supplementary Data 3). However, a crucial step enabling members of the genus *Celerinatantimonas* to invade marine flowering plants was probably the acquisition of a pathway to extracellularly degrade pectin, a polysaccharide that is typically found in terrestrial plant cell walls but is rare in marine algae⁴⁰. The comparison of 34 genomes of related genera and families revealed that only members of the genus *Celerinatantimonas* have the ability to degrade pectin extracellularly (Supplementary Data 3).

Our results show that seagrasses have independently evolved a mechanism to cope with N limitation that is similar to a variety of terrestrial plants^{2,17,18}. Just like N₂-fixing microorganisms presumably facilitated early land plants to successfully colonize N-poor soil⁶, the ancestors of *Ca. C. neptuna* and its relatives probably enabled marine flowering plants to invade and thrive in N-poor marine habitats, where their descendants form the basis of extremely efficient blue carbon ecosystems⁷.

Methods

Etymology

‘*Candidatus Celerinatantimonas neptuna*’ (nep.tu’na L. fem. n.), pertaining to *Neptunus* (L. masc. n. Neptune), the Roman god of the seas and the Neptune grass, *Posidonia oceanica*.

Sampling

A *P. oceanica* meadow at 8 m water depth and nearby sandy sediments in Fetovaia Bay, Elba, Italy¹³ were sampled between June 2014 and September 2019; individual sampling months and years are indicated in the sections below and/or in the figures and tables. In May 2017, a *P. oceanica* meadow at the island of Pianosa, Italy was also sampled. All of the samples were obtained via SCUBA diving.

Complete plants of *P. oceanica* were carefully separated from the meadow by hand and stored in seawater-filled containers until arrival at the shore-based laboratory. Sediment for use in the laboratory-based aquaria was scooped into containers from nearby sandy patches. Seawater was pumped through a hose (placed at about 0.5 m above the *P. oceanica* meadow) into several 50 l barrels onboard the boat and was later used in the laboratory for the aquarium and the incubation experiments.

The sediment within the seagrass meadow was sampled with stainless steel core tubes (length, 50 cm), which were drilled into the sediment by divers, and the cores were

briefly stored at 22 °C (ambient temperature, September 2019) in a seawater-filled barrel until further processing at the shore-based laboratory.

Porewater nutrient samples were obtained using stainless steel lances⁴¹ at intervals of around 10 cm. Water column nutrient samples were obtained from above the seagrass meadow at the start or end of sampling. Nutrient samples were collected in 15 ml or 50 ml centrifuge tubes and were stored in a cooler box until further processing.

Nutrient measurements

Water column nutrients were measured during several sampling campaigns as indicated in Extended Data Table [1a](#). Ammonium (NH_4^+) concentrations were measured fluorometrically⁴² in the nearby shore-based laboratory, and the remaining water was frozen (-20°C) for later analyses of nitrate (NO_3^-), nitrite (NO_2^-), phosphate (PO_4^{3-}) and silicate (SiO_4^{4-}) using an autoanalyser (QuAAstro, Seal Analytical). Porewater samples were obtained in June 2019 and were processed the same as the water column nutrient samples with the exception that ammonium was not measured on site but at the home laboratory at the same time as the other nutrients. Dissolved inorganic nitrogen (ammonium plus NO_x^-) concentrations in the porewater were averaged for the upper 20 cm (Extended Data Table [1b](#)).

Net primary production measurements using the EC method

Net carbon dioxide (CO_2) fluxes were calculated on the basis of oxygen (O_2) fluxes determined using the aquatic eddy covariance (EC) method. In this non-invasive approach, turbulence-induced transport is resolved using high-frequency current meters combined with fast O_2 microsensors. Under the assumption of stationarity, the instantaneous turbulent flux contributions are calculated by correlating vertical current fluctuations to oxygen fluctuations. Our EC system was equipped with an acoustic Doppler velocimeter (ADV, Nortek) and ultra-fast responding optode microsensors with a tip diameter of 430 μm ($t_{90} < 0.3$ s, Pyroscience). The microsensor tip was aligned vertically to the centre of the ADV measuring volume and shifted 2.5 cm horizontally to avoid current disturbances. Inside the seagrass meadow, the EC system was installed upside-down, such that the measuring volume was approximately 40 cm above the seagrass canopy height, while an upright installation was chosen for sandy sediments with the measuring volume about 20 cm above the sea floor. Additional sensors were used to monitor long-term O_2 changes (Aanderaa, 4831), temperature variations (PT100, custom made) and photosynthetic active radiation sensor (Biospherical Instruments, QCP-2000). All of the instruments were attached to an aluminium frame, which enabled SCUBA-diver-operated deployments. Current data

were recorded at 16 Hz and O₂ data at 1–4 Hz, limited by the response time of the sensor. All of the other instruments recorded at 0.1 Hz.

The data were processed according to standard procedures for aquatic eddy correlation measurements^{43,44} in MATLAB 2018b (Mathworks). First, current data were downsampled to the frequency of O₂ measurements, despiked and corrected for the tilt of the ADV. Subsequently, current and O₂ data were decomposed into the steady and fluctuating component using a low-pass filter. The fluctuating O₂ time series was shifted until a maximum correlation with the vertical velocity fluctuations was achieved; this was typically in the range of 2–3 s resulting from low horizontal velocities. Instantaneous fluxes are usually highly variable; thus, two averaging procedures were applied: (1) instantaneous fluxes were temporally integrated to determine cumulative fluxes, (2) if the cumulative fluxes showed sudden jumps indicating erroneous measurements, the time series was truncated at that point. Instantaneous fluxes were then averaged for a 60 min burst and were subsequently averaged over the entire measurement time. Negative fluxes during the night represent respiration, whereas positive fluxes during the day represent the sum of respiration and gross primary production. In June 2019, the EC system was deployed twice for 22–24 h in the centre of the sampled seagrass meadow. Seagrass metabolism rates were referenced by a sandy sediment station (24 h) at a similar water depth at a distance of approximately 50 m to the meadow sampling site. One additional measurement in sandy sediments was performed for 13 h in September 2018. Net O₂ fluxes were converted to net CO₂ fluxes using a ratio of 1 mol O₂:1 mol CO₂. Custom codes for data processing are available (see the ‘Code availability’ section).

Oxygen concentrations within seagrass sediments were measured using microsensors⁴⁵ mounted to the same frame as the EC system.

Sediment-free incubation of complete plants with the ¹⁵N₂ tracer

During six campaigns (June 2014, May 2015, April 2016, August 2016, May 2017 and September 2018), complete plants of *P. oceanica* were collected by SCUBA diving and were transported to the shore-based laboratory in seawater-filled containers to prevent desiccation. On arrival at the laboratory, the plants were embedded into an aquarium containing sediment that was collected from close to the meadow such that the rhizomes and roots were sediment-covered. The aquarium was equipped with two lamps (OSRAM L58W/77, FLUORA 2250 lm; OSRAM L30W/77, FLUORA 1000 lm) and the site seawater in the aquarium was cooled to ambient temperatures. The plants were kept in the aquarium for up to one day before starting tracer incubations for N₂ fixation. To prepare ¹⁵N₂-enriched water for the incubations, site seawater was

filtered through a filter (pore size, 0.2 µm) and placed into 0.5 l wide-neck Duran bottles. The filtered seawater was deoxygenated by bubbling with a mixture of nitrogen (N₂) and argon (Ar) gas. Concentrations of oxygen (O₂) were checked using needle optodes (PreSens, Precision Sensing), and bubbling continued until O₂ concentrations were close to zero but at least <10 µmol l⁻¹. The nearly anoxic seawater was carefully filled into 60 ml or 120 ml serum bottles, which were crimp-sealed. Then, 10 ml or 20 ml of ¹⁵N₂ gas was injected into the serum bottles in exchange for 5 ml or 10 ml of seawater to produce a slight gas overpressure. The serum bottles were vortexed for around 1 min and left overnight to equilibrate. The ¹⁵N₂ gas (\geq 99 atomic percentage (at%) ¹⁵N; lot numbers 19197 and 16727, Cambridge Isotopes; purchased from Eurisotop) was tested for ¹⁵N-ammonium contamination⁴⁶ using the hypobromite method⁴⁷ before use, and no contamination was detected.

To incubate only the roots and rhizomes of complete plants (that is, leaves, rhizome and roots connected) with the ¹⁵N₂ tracer, the water surrounding the roots and rhizomes was separated from that surrounding the leaves (Fig. 1c). We therefore fitted latex gloves (rinsed three times with filtered seawater) with a sampling port and a little hole through which the leaves of each individual plant were carefully threaded. The latex glove was secured with a clamp on the leaves as close to the meristem as possible. The roots and rhizomes were placed into 0.5 l wide-neck Duran bottles that were filled with deoxygenated water (prepared as described above), and the leaves remained outside the bottle. The glove was used to seal the Duran bottle (together with cable ties and rubber bands). The sampling port enabled the addition of the ¹⁵N₂-enriched water. During all campaigns, except for June 2014, 110 ml of ¹⁵N₂-enriched water was added to each Duran bottle containing roots and rhizomes, without leaving a headspace. Before closing the sampling port, a subsample of the mixed incubation water was taken to measure the enrichment of ¹⁵N in the N₂ pool at the start of the incubation. In June 2014, about 40% of the incubation water was replaced with ¹⁵N₂-enriched water and the enrichment of ¹⁵N in the N₂ pool was measured in the batch of enriched water. Final enrichment was calculated on the basis of the addition of enriched water to the final incubation¹³. The Duran bottle was covered with a black plastic bag to prevent light from reaching the roots and rhizomes. The incubation bottles with the plants were carefully placed into the aquarium. For every experiment, three replicate plants were prepared with the ¹⁵N₂ tracer while one plant was prepared the same way but without the added ¹⁵N₂ tracer (that is, an incubated control for background natural abundance values) except for June 2014 and May 2015, for which natural abundance values were obtained from non-incubated plants at the start of the experiments. Plants were incubated for 4–96 h with most incubations lasting 24 h or 48 h. Incubations were performed as light–dark cycles if incubated for \geq 24 h. In June

2014, one additional set of plants was also incubated with 24 h of light. At the end of the incubation, a second subsample of the incubation water was taken through the sampling port to measure the final enrichment of ^{15}N in the N_2 pool (except for June 2014, see above). The incubation set-up was disassembled, and the plant was dissected into root, rhizome and leaf tissues. Pieces from each tissue were preserved for the determination of microbial N_2 fixation or N transfer rates (frozen at $-20\text{ }^\circ\text{C}$), microbial community analyses/sequencing (frozen at $-20\text{ }^\circ\text{C}$) and microscopy/single-cell analyses (see below). The remaining plant tissues were frozen at $-20\text{ }^\circ\text{C}$ and were later used, for example, for amino acid analyses. At least one experiment was performed for each sampling campaign.

N₂ fixation and N transfer rates

The elemental and isotopic composition of 1–22 individual root, rhizome and leaf pieces from each incubated plant (for June 2014 and May 2015, also the non-incubated plant) was measured using an elemental analyser (Thermo Fisher Scientific, Flash EA, 1112 Series) coupled to a continuous-flow isotope ratio mass spectrometer (Delta Plus Advantage, Thermo Finnigan) (EA-IRMS) as described by Lehnert et al.¹³. The enrichment of ^{15}N in the N_2 pool was measured using membrane inlet mass spectrometry (MIMS, GAM200, IPI). The enrichments of ^{15}N in the N_2 pool at the beginning and at the end of the incubation were averaged (except for June 2014) for the rate calculation (June 2014: 5.4 at% ^{15}N ; May 2015: 29–46 at% ^{15}N ; April 2016: 12–17 at% ^{15}N ; August 2016: 18–32 at% ^{15}N ; May 2017: 32–40 at% ^{15}N ; September 2018: 22–32 at% ^{15}N). Detection limits were set as a minimum change in $\delta^{15}\text{N}$ from natural abundance values within tissue types (that is, three times the s.d. of natural abundance measurements within each set of plants). This approach resulted in minimum changes in $\delta^{15}\text{N}$ values of 0.1–6.1‰ with an average of 1.8‰. When natural abundance measurements were not available (for example, failed measurements), the natural abundance values of plants and tissues closest to the same incubation conditions were used. Negative rates and rates below the detection limit were set to zero for plotting and further analysis. Root-associated microbial N_2 fixation rates were calculated according to Lehnert et al.¹³ and are presented as μmol or nmol N fixed per gram dry weight (DW) of tissue per day ($\mu\text{mol g}^{-1}\text{DW}^{-1}\text{d}^{-1}$ or $\text{nmol g}^{-1}\text{DW}^{-1}\text{d}^{-1}$ N).

Any significant enrichment of ^{15}N in leaf pieces can originate only from root-associated fixation of $^{15}\text{N}_2$ and the subsequent transfer of ^{15}N -labelled, freshly fixed N to the leaves as leaves were outside the incubation bottle with $^{15}\text{N}_2$ and rhizomes do not have a substantial role in N_2 fixation¹³. The measured isotopic composition of the leaf pieces was therefore used to calculate transfer rates of freshly fixed N from roots

to leaves using the same rate and detection limit calculations. Transfer rates are presented as μmol or nmol N fixed per gram dry weight of tissue per day ($\mu\text{mol g}_{\text{DW}}^{-1} \text{d}^{-1}$ or $\text{nmol g}_{\text{DW}}^{-1} \text{d}^{-1}$ N).

On the basis of average rate detection limits of root-associated N_2 fixation rates (values obtained from propagating the minimum change through the rate equations; $0.01 \mu\text{mol g}_{\text{DW}}^{-1} \text{d}^{-1}$ N), plants were classified as non- N_2 -fixing (below the average detection limit) or N_2 -fixing (above the average detection limit) for subsequent microbial community analyses (see below).

The amount of primary production that can be sustained by root-associated N_2 fixation was calculated on the basis of (1) average N_2 fixation rates (roots and rhizomes) and N transfer rates (leaves); (2) the biomass of each tissue per incubated shoot (using an empiric conversion between dry weight and wet weight); (3) tissue-specific carbon-to-nitrogen ratios (obtained from EA-IRMS measurements); and (4) the number of shoots (counts obtained during SCUBA diving in June 2019) ([Supplementary Information](#)).

Amino acid quantification and ^{15}N enrichment

During the extraction of total acid-hydrolysable amino acids and downstream processing, precautions were taken to avoid contamination by combusting all laboratory glassware before use (450°C for 12 h). Frozen root material (0.1–1.1 g wet weight) from the $^{15}\text{N}_2$ incubations in August 2016 was freeze-dried for 2 d (Christ).

Total acid-hydrolysable amino acids were extracted as follows: 20–50 mg of the freeze-dried sample were added to 3 ml of 6 M hydrochloric acid (HCl). Vials were closed with an N_2 -flushed headspace and kept at 110°C for 20 h. Then, 0.1–0.2 ml of the internal standard norleucine ($11.1 \mu\text{mol ml}^{-1}$) was added after the hydrolysis. After centrifuging the samples at 3,000 r.p.m. for 4 min, the supernatant was decanted and the pellet was dissolved in 1 ml nanopure water (MilliQ) by vortexing for 10 s and again centrifuged. The supernatant was added to the previous one, and was heated to 95°C while flushing with N_2 gas until completely dried. The samples were derivatized according to a modified method by Corr et al.⁴⁸ to transform amino acids to *N*-acetyl *i*-propyl ester derivatives. In brief, amino acids were propylated with 0.63 ml of a 1:4 acetylchloride:isopropanol solution, while flushed with argon, and then kept at 100°C for 1 h. Each vial was then cooled down to room temperature and flushed with N_2 until dried. Then, 0.75 ml of a derivatization solution (7.2 ml acetic anhydride, 14.4 ml triethylamine and 36 ml acetone) was added to each vial, flushed with N_2 while closing, vortexed and kept at 60°C for 10 min. The samples were carefully flushed with N_2 until just dry. To each sample, 2 ml ethylacetate and 1 ml of saturated sodium chloride (NaCl) solution were then added and the sample was centrifuged at 2,400

r.p.m. for 3 min. The (top) organic phase was separated and carefully dried down with N₂. The derivatized amino acids were redissolved in 200 µl ethylacetate from which 1.5–3 µl was used for concentration and ¹⁵N/¹⁴N isotope ratio measurements. Amino acid concentrations were quantified using a gas chromatography (GC) system equipped with a flame-ionization detector (Agilent 6890N GC/7683 ALS Autosampler) and an InertCap 35 GC column (GL Sciences, 60 m × 0.32 mm × 0.50 µm). The isotope ratios (¹⁵N/¹⁴N) of individual amino acids were determined using a TRACE 1310 GC (equipped with the same column) coupled to an isotope ratio mass spectrometer (Delta V/GC IsoLink II IRMS System, Thermo Fisher Scientific).

Nucleic acid extractions

Nucleic acid extractions of *P. oceanica* root, rhizome and leaf pieces were started by submerging several different root pieces of a plant into liquid nitrogen and homogenizing the frozen pieces with a mortar and pestle. The powdered root material was divided into two aliquots—one aliquot was extracted using the DNeasy Plant Mini Kit (Qiagen) according to the manufacturer’s instructions but excluding the RNase step. The other aliquot was extracted according to the protocol by Pjevac et al.⁴⁹. The two extracts were pooled for each sample. The nucleic acids were then concentrated in a Speedvac (Eppendorf) at 30 °C for 50 min. The concentrate was cleaned using the Wizard DNA clean up Kit (Promega), eluted in PCR-grade water and stored at –20 °C. These nucleic acid extracts were used for Illumina-based 16S rRNA gene amplicon sequencing, Illumina-based shotgun metagenomes and for metatranscriptomes. For the PacBio-based metagenome, nucleic acids were extracted from frozen root tissue at the Max-Planck Genome Centre Cologne using the NucleoBond HMW DNA Kit (Macherey and Nagel). DNA was quality- and quantity-assessed by capillary electrophoresis (Agilent Femtopulse) and Quantus (Promega), respectively. DNA was not fragmented further and was directly used for PacBio library preparation.

Sediment for nucleic acid extractions was retrieved from cores in September 2019. Cores were sectioned and the sandy surface layer from 2–10 cm (D1) was frozen at –20 °C until further processing. DNA was subsequently extracted using the DNeasy PowerSoil Kit (Qiagen) according to the manufacturer’s instructions and quantified using the Qubit dsDNA HS Assay Kit on a Qubit 2.0 Fluorometer (Invitrogen).

16S rRNA gene amplicon sequencing and analyses

Microbial community analyses were performed for root pieces from the ¹⁵N₂ fixation experiments to determine differences between N₂-fixing and non-N₂-fixing plants (see above). Nucleic acid extracts were sent to the Max Planck-Genome-Centre Cologne, Germany (<http://mpgc.mpi-pz.mpg.de/home/>) for barcoding PCR, library preparation

and sequencing. The barcoding PCR was performed using the bacterial primers Bact341F (barcoded) and Bact805R (ref. ⁵⁰) and the DreamTaq DNA Polymerase (5 U μl^{-1} , Thermo Fisher Scientific). PCR started with an initial denaturation step at 98 °C for 30 s; followed by 30 cycles of 98 °C for 10 s, 55 °C for 30 s and 72 °C for 30 s; and one final elongation step at 72 °C for 5 min. The 16S rRNA gene amplicons were sequenced using the Illumina HiSeq2500 sequencing platform with 2 \times 250 bp paired-end reads.

Microbial community analysis of the 16S rRNA gene amplicon data was carried out using the QIIME2 environment with a number of available plugins⁵¹. In brief, after importing demultiplexed reads into QIIME2, primer sequences were removed using cutadapt⁵² and read pairs were joined using vsearch⁵³. Error correction, trimming (to a length of 400 nucleotides) and operational taxonomic unit (OTU) clustering at the 100% similarity level was performed using deblur⁵⁴. Taxonomy was assigned to OTUs with a sklearn-based classifier⁵⁵ through the feature-classifier plugin⁵⁶ using the full-length 16S SILVA-SSU-132 database (QIIME-compatible release from April 2018; <https://www.arb-silva.de/documentation/release-132/>).

As an initial assessment of whether the 16S rRNA gene amplicon datasets were representative of the sequenced root material, we calculated the ratio of bacterial to organellar reads. Some samples had a very high ratio, suggesting that the root material (and therefore also the endophytic microbial community) was not well represented. We therefore chose a cut-off of a minimum of 10% of organellar reads (out of the total reads) and, on the basis of this cut-off, three samples (all from the largest group of samples in May) were subsequently excluded from further analyses. After this initial assessment and before removing OTUs representing plastids and mitochondria, the ratio of *Celerinatantimonas*-related reads (later renamed *Ca. C. neptuna*) to organellar reads was calculated to assess whether *Ca. C. neptuna* had increased in absolute abundance in N₂-fixing plants relative to non-N₂-fixing plants⁵⁷. After removing OTUs representing plastids and mitochondria, the final OTU table comprised 13,886 OTUs (from a total of 31 samples that had a corresponding N₂ fixation rate). In a separate analysis, 16S rRNA gene amplicons sequenced from roots of a *P. oceanica* plant sampled from a meadow at the island of Pianosa were analysed equivalently.

For alpha diversity analysis, the OTU table containing the 31 samples was rarefied to a total count of 2,200 (therefore excluding 6 samples with a total OTU count <2,200; indicated in Extended Data Fig. 3) and statistical differences in alpha diversity indices between N₂-fixing and non-N₂-fixing plants were assessed with the Kruskal–Wallis pairwise test⁵⁸ using QIIME2 diversity alpha-group-significance. Beta diversity was assessed on the basis of the non-rarefied OTU table, including all 31 samples, using Aitchison principal-component analysis (PCA) through the DEICODE plugin⁵⁹ and

visualized with EMPeror⁶⁰. DEICODE also identified the OTU that contributed most to the clustering of samples in the PCA. Statistical differences in beta diversity clustering between N₂-fixing and non-N₂-fixing plants were assessed by permutational analysis of variance using QIIME2 diversity beta-group-significance testing⁶¹. Differential abundance testing of OTUs between N₂-fixing and non-N₂-fixing plants was performed using Songbird⁶² and visualized with Qurro⁶³ in QIIME2. Relative abundances of bacterial OTUs were visualized with phyloseq⁶⁴.

Metagenome sequencing and analysis

N₂-fixing plants (from June 2014 and August 2016) were selected for metagenome sequencing of nucleic acids extracted from root (two plants), rhizome (one plant) and leaf (one plant) tissues as well as meadow sediment (three cores). Nucleic acid extracts were sent to the Max Planck-Genome-Centre Cologne and sequencing was performed using the Illumina MiSeq platform with 2 × 250 bp paired-end reads (0.6–7.3 Gb and 8.7–10.6 Gb for plant tissue and sediment, respectively). Taxonomic assignment of raw metagenomic reads was performed using phyloFlash v3.3b3 (ref. ⁶⁵) and the parameters --tophit with the SILVA 138 database. Before analysis, reads assigned to mitochondria, chloroplasts and Eukarya were removed. Bar plots were generated using the phyloFlash_compare.pl script included in phyloFlash.

Raw metagenomic reads were later mapped onto the genome of *Ca. C. neptuna* using bbmap v.38.75 and the following parameters: minid=0.99, maxindel=1000. To reduce false positives, the rRNA operons were removed from the genome of *Ca. C. neptuna* before the mapping. For each metagenome, the number of mapped reads was normalized to the total number of reads (per million).

One of the plants with high relative abundances of *Ca. C. neptuna* (from August 2016) was also sequenced using PacBio technology (at the Max Planck Genome Centre Cologne) to obtain the genome of *Ca. C. neptuna*. In brief, the PacBio library was prepared using the SMRTbell Express Kit 2.0 (Pacific Biosciences). The library was size-selected to remove fragments smaller than 9 kb. The resulting fraction was sequenced on a single SMRT Cell (8M ZMWs) on the Sequel II system with sequencing chemistry 2.0 and binding kit 2.0 in continuous long read mode for 30 h with a total yield of 318.45 Gb (continuous long read mode). High-quality PacBio circular consensus sequencing (CCS) reads were assembled using metaFlye v.2.7 (ref. ⁶⁶). The assembly contained a circular 4.26 Mb contig with a coverage of 85, encoding 16S rRNA sequences with 100% identity to the *Celerinatantimonas*-related OTU associated with N₂-fixing plants and 95% identity to the 16S rRNA sequence of *C. diazotrophica*. For polishing of the *Ca. C. neptuna* metagenome-assembled genome (MAG), the 2 × 250 bp reads of the Illumina metagenome were mapped onto the

metaFlye assembly with the BWA-MEM short read aligner⁶⁷ using the default settings. The resulting SAM mapping file was converted into the BAM format, sorted and indexed using SAMtools v.1.10 (ref. ⁶⁸), and subsequently used for polishing using Pilon v.1.23 (ref. ⁶⁹). The polished MAG had an estimated completeness of 100% with 0.81% contamination (CheckM (v.1.0.18)⁷⁰) and was annotated using Prokka⁷¹. Mapping of CCS reads onto the *Ca. C. neptuna*-MAG was performed using minimap2 (ref. ⁷²).

At the time of our analysis, the genome of the closest relative *C. diazotrophica* was not available for comparison. We therefore obtained the isolate (DSM18577) from the DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen), grew the culture according to Cramer et al.²⁵ and sequenced the genome using the Illumina HiSeq 2500 platform at the MP-GC in Cologne for comparison (Extended Data Fig. 6 and [Supplementary Information](#)).

To determine the presence/absence of selected genes/pathways in the genomes of species, genera and families closely related to *Ca. C. neptuna*, we used the RAST annotation webserver⁷³ to annotate *Ca. C. neptuna* and 34 other genomes of other genera of the Celerinatantimonadaceae, Idiomarinaceae and Colwelliaceae. The genome accession codes and the presence/absence of selected pathways is summarized in Supplementary Data 3. Moreover, carbohydrate-active enzymes were annotated in genomes belonging to genera *Celerinatantimonas*, *Agarivorans*, *Aliagarivorans* and *Alginatibacterium* using the dbCAN meta server^{74,75} using predicted protein sequences of the RAST annotation. Carbohydrate-active enzymes were predicted using HMMER, DIAMOND and Hotpep and only annotations made by ≥2 tools were retained.

16S rRNA phylogenetic tree reconstruction

The full-length 16S rRNA gene sequences of the closed, Prokka-annotated *Ca. C. neptuna* genome and the *Ca. C. neptuna*-related OTUs obtained from *P. oceanica* roots off the island of Pianosa, Italy (identical to those recovered at the island of Elba) were analysed phylogenetically to infer evolutionary relationships. The 16S rRNA gene sequences were added to the Silva database SSURef NR 99 release 138 (released on 11 November 2019)⁷⁶, automatically aligned using SINA⁷⁷ and the alignment was refined manually in ARB. Phylogenetic trees were calculated using distance matrix neighbour joining, maximum parsimony and maximum likelihood (FastDNAML) algorithms in ARB without position variability filters, and a consensus tree was constructed.

Metatranscriptomic sequencing and analysis

Metatranscriptomic analyses were performed on nucleic acid extracts from June 2014. DNA was degraded using TURBO DNase ($2\text{ U }\mu\text{l}^{-1}$; Thermo Fisher Scientific), and RNA-sequencing libraries were constructed using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (New England Biolabs). Sequencing-by-synthesis was performed on the Illumina HiSeq3000 sequencer (Illumina) with the 1×150 bp read mode. Library preparation and sequencing were performed by the Max Planck-Genome-Centre Cologne, Germany.

Raw transcriptomic reads were trimmed using Trimmomatic v.0.32 (MAXINFO:100:0.2, MINLEN:75)⁷⁸ after rRNA removal using SortMeRNA v.2.1 (ref. ⁷⁹) on the basis of both bacterial and archaeal rRNA databases. Non-rRNA reads were then mapped onto the genome of *Ca. C. neptuna* using Bowtie2 v.2.1.0 with the default settings⁸⁰. Indexed BAM files were generated using samtools v.0.1.19 (ref. ⁶⁸) and transcripts per feature were quantified using featureCounts v.1.4.6 (ref. ⁸¹) with a minimum read overlap of 75 bp (--minReadOverlap). Normalized gene transcription was subsequently quantified as transcripts per million (TPM)⁸² using the formula:

$$\text{\$}\$\{\{\backslash rm\{ TPM\}\}\}_{\{ i\}}=\backslash frac\{ \{ c\}_{\{ i\}}\}\{\{ l\}_{\{ i\}}\}\backslash times \backslash frac\{ 1\}\{\{\backslash sum\}_{\{ j\}}\backslash frac\{ \{ c\}_{\{ i\}}\}\{\{ l\}_{\{ i\}}\}\}\backslash times \{ 10\}^{\{ 6\}}\$\$ \\ (1)$$

to assign each feature i a TPM value where c is the feature count, l is the length in kilobases and j is all features. TPM values were visualized together with the mapped Illumina short reads, the mapped PacBio CCS reads in circular genome figures using BRIG⁸³. Furthermore, TPM values were normalized to the average TPM of a set of housekeeping genes (*rpoA*, *rpoB*, *ftsZ*, *rho*, *recN*, *gyrB*, *recA* and *gyrA*)^{84,85,86,87,88,89} for each of the five samples from June 2014 (one sample did not return enough *Ca. C. neptuna* reads to be mapped) to visualize the selected pathways (Extended Data Fig. ⁷).

Fixation, embedding and sectioning of *P. oceanica* root material

Root pieces were preserved at the end of the $^{15}\text{N}_2$ tracer incubations (see above) in paraformaldehyde solution (4% (w/v) final concentration in filtered seawater) at ambient/room temperature for 1 h. Root pieces were then washed with phosphate-buffered saline (PBS) solution and in nanopure water (MilliQ) for 15 min and 10 min, respectively. The pieces were then dehydrated in 96% ethanol for 2 min and air dried for 30 min. The fixed root material was stored at -20°C until further processing.

Before resin infiltration, the formaldehyde-fixed root material was dehydrated using an ethanol series of 30%, 50%, 70%, 80% and 90% (once), and 100% (twice) for 10 min each. Pieces were then infiltrated with resin by stepwise increases of London

Resin White (LRW; Sigma-Aldrich) with concentrations of 25%, 50% and 75% (each once), and 100% (twice) LRW in ethanol (modified from McDonald⁹⁰). Each infiltration step was performed for 15 min, and the root pieces were then centrifuged for 5–10 min using a benchtop centrifuge. For polymerization of the resin, the individual root pieces were submerged in 100% LRW resin inside gelatin capsules or Eppendorf tubes. Capsules or tubes were placed inside a gas-tight bag, which was flushed with N₂ gas for 1 h. The gas-tight bags were subsequently kept at 65 °C for 4–5 d. Semi-thin (thickness, 0.5–1 µm) sections of root pieces were cut with glass knives using the Leica UC7 Ultramicrotome (Leica Microsystems). The sections were placed onto Polysine Adhesion Slides (Thermo Fisher Scientific) or indium tin oxide (Präzisions Glas & Optik) glass slides and were dried on a heating plate at 60 °C for 5 min. The semi-thin sections were stored at 4 °C until further processing.

FISH analysis of semi-thin root sections

To visualize *Ca. C. neptuna* cells in *P. oceanica* roots, we designed four FISH-probes targeting the 16S rRNA gene of *Celerinatantimonas* spp. (that is, the 16S rRNA genes of the MAG, *C. diazotrophica* and *C. yamalensis*). The probe set had some matches outside *Celerinatantimonas*, which were, however, not present in our 16S rRNA gene amplicon dataset. All individual FISH probes, the probe set (all four probes together) and the EUB338-I (positive control) and the NON338 (negative control) probes^{91,92} were used to determine melting curves using the *C. diazotrophica* DSM18577 culture. All four probes (5'–3': Cel_442 (ACCCTTCCTCACAAAC), Cel_186 (TCCCCTGCTTGGTCCGTAG), Cel_660 (AAATTCTACCTCCCTCTACA) and Cel_227 (TAATCTCACTGGGTGCATC)) were then used in combination for all hybridizations using a formamide concentration of 25%. All FISH probes were obtained from Biomers and were labelled with two Atto550 molecules⁹³.

Hybridizations were performed on semi-thin root sections (see above) using standard hybridization protocols with the following modifications. Root sections were encircled with a water-repellent barrier oil layer using an oil PAP PEN (G. Kisker) to ensure that the sections were always submerged in the respective solutions. Hybridization using the *Celerinatantimonas* probe mix was performed at 35 °C for 2 h (hybridization solution with 0.9 M NaCl, 0.02 M Tris-HCl, 25% formamide and 0.01% (w/v) SDS). After hybridization, thin sections were sequentially washed in washing buffer (0.01% (w/v) SDS, 20 mM Tris-HCl, 5 mM EDTA and 149 µM NaCl) for 45 min at 37 °C, in 4 °C cold PBS for 15 min and in nanopure water (MilliQ) for 5 min. Sections were placed in ethanol (96%) for 2 min and then air-dried. DNA was counterstained with 4',6-diamidino-2-phenylindole (DAPI). The root sections were covered with a mixture of Citifluor AF1 (Citifluor) and Vectashield (Vector Laboratories) (ratio of 1:4) and a coverslip for microscopy. FISH was performed on sections prepared from root material from June 2014, April 2016 and August 2016 for cell counts, quantitative

visualization and images. Additional material was sampled and preserved in June 2019 and processed for FISH imaging.

Microscopy and nanoSIMS analysis

Bacterial cells hybridized with the *Celerinatantimonas* FISH probe (that is, *Ca. C. neptuna* cells) were counted manually using epifluorescence microscopy on several root cross-sections (thickness, 0.5–1 µm) from April 2016 (non-N₂-fixing), June 2014 (N₂-fixing) and August 2016 (N₂-fixing). Individual representative images were taken from these root cross-sections for illustration. Additional FISH images were obtained from root material collected in June 2019.

To obtain a proxy for the contribution of *Ca. C. neptuna* to total biomass within the complete root cross-sections, the distribution of *Ca. C. neptuna* cells was mapped using epifluorescence images acquired using the Zeiss Axio Imager. An M2 microscope at 100-fold magnification equipped with an automated XYZ stage (Märzhäuser Wetzlar, SCAN IM, 130 × 85, 2 mm). Approximately 10 × 10 images in horizontal directions and 30 images in the vertical direction were taken to cover a full cross-section. Each image was composed of three channels: DAPI (blue), autofluorescence (green) and FISH (red/orange). The raw image stack was processed using the Zeiss microscope software ZEN (ZEN 3.2 blue edition). In brief, images were first stitched and corrected for shading. A deconvolution algorithm and orthogonal projection was applied to correct for noise and light scattering into the image from planes above and below the focal plane. The channels were then merged to optimize the contrast of positively hybridized cells attached to the root tissue. In the resulting RGB image, overlapping signals of DAPI and FISH appear pink and autofluorescence in green. Subsequently, the images were processed in MATLAB (Mathworks 2018b) to determine the area occupied by FISH-positive cells and root tissue. The area surrounding the root was masked, and the image was decomposed into the red, green and blue channels. Root tissue was determined based on the green channel, which was binarized with a threshold of 1% of its maximum intensity. In the binarized image, root tissue appears white (1) while the remaining pixels are black (0). To determine the total root area, all white pixels were integrated (1 px represents 0.1 µm). For the positively hybridized cells, a similar procedure was applied on the basis of the combined red and blue channel. However, the root tissue, namely the rhizoplane, epidermis, hypodermis and the innermost areas of the stele, were masked due to strong autofluorescence signals along the whole spectrum and were excluded from the processing of the FISH-positive cellular area. Manual cell counts had confirmed that these root tissues did not contain any *Ca. C. neptuna* cells, and the exclusion of these areas therefore did not bias our automated analyses. Areas of root tissue and *Ca. C. neptuna* cells were later used in the nanoSIMS-based mass balance assuming that the occupied area is representative of the biomass contribution. This

quantification was performed on semi-thin root sections from April 2016 (non-N₂-fixing), June 2014 (N₂-fixing) and August 2016 (N₂-fixing).

For visualization purposes (Extended Data Fig. 4), the black and white image of the root image was inverted, such that the root tissue appears in black. The root image was then overlain with the positively hybridized cells in red (Extended Data Fig. 4). To better visualize the location of cells in the root cross-section, cells were artificially blurred by applying a Gauss filter at increasing kernel sizes (3–20 pixel).

After microscopy, the Citifluor–Vectashield-mix was washed off the sections with nanopure water (MilliQ) three times, and the sections were air-dried on their slides. Before nanoSIMS measurements, the root sections were sputter-coated with 10 nm gold (Au) using the Leica EM ACE600 (Leica Microsystems) sputter coater to ensure conductivity. For nanoSIMS measurements, the area of interest was presputtered for 2 min with a positively charged caesium (Cs⁺) primary ion beam to implant Cs⁺ on the sample surface. Sample surfaces were rastered with a Cs⁺ primary ion beam with a current of 1.5 pA. Primary ions were focused into a nominal \leq 100 nm spot diameter. The image resolution was 256 px \times 256 px with a dwelling time of 1 ms per pixel. Analysed areas were 20 μm \times 20 μm . Secondary ion counts of carbon (¹²C⁻), nitrogen (as ¹²C¹⁴N⁻ and ¹²C¹⁵N⁻), phosphorus (³¹P⁻) and sulfur (³²S⁻) were recorded simultaneously by the electron multiplier detectors of the multicollection system of the instrument.

To have a better statistical representation of the ¹⁵N enrichment in root tissue and *Ca. C. neptuna* cells, a total of 167 NanoSIMS images (37 images for June 2014 section and 130 images for August 2016) were processed using a semi-automated algorithm. First, all nanoSIMS images were processed using look@nanoSIMS⁹⁴. All planes (40 for each image) were drift-corrected and accumulated. NanoSIMS (¹²C¹⁴N⁻) images were manually aligned and overlapped with their corresponding epifluorescence microscopic images. Next, the NanoSIMS and epifluorescence microscopic images were exported and further processed using a custom-developed MATLAB algorithm. On the basis of the epifluorescence images, binary matrices for positively hybridized cells and root material were determined as described above. These binary matrices were pixel-wise multiplied with the ¹⁵N/¹⁴N isotope ratio matrices, yielding an isotope ratio matrix for *Ca. C. neptuna* cells and an isotope ratio matrix for root tissue. The isotope ratios within the two matrices were then averaged yielding one isotope ratio value for *Ca. C. neptuna* cells and one for root tissue per nanoSIMS image. Isotope ratios ($r = {}^{15}\text{N}/{}^{14}\text{N}$ from ¹²C¹⁵N⁻/¹²C¹⁴N⁻) were then converted to atomic percentage using ${}^{15}\text{N}$ at% = $r/(r + 1) \times 100$ (at%). Of all 167 images, 78 images were of root tissue only whereas 89 contained both root tissue and *Ca. C. neptuna* cells.

To account for the different labelling percentages of the N₂ pool in June 2014 and August 2016, relative incorporation (per day) was calculated using the ¹⁵N at% (at%_{cell}) of (1) *Ca. C. neptuna* cells; (2) root tissue with *Ca. C. neptuna* cells close by (that is, nanoSIMS image with root tissue and *Ca. C. neptuna* cells present) as well as root tissue alone (that is, nanoSIMS images without *Ca. C. neptuna* cells present), natural abundance background ¹⁵N (June 2014: 0.367194; and August 2016: 0.367456; at%_{NA}) from the bulk biomass, the enrichment of ¹⁵N in the N₂ pool (at%_{N2}, measured by MIMS; see above) and incubation time (*t*) as follows:

$$\text{Relative incorporation} (\%) = \left(\frac{\text{at}(\text{cell})\% - \text{at}(\text{NA})\%}{\text{at}(\text{N2})\% - \text{at}(\text{NA})\%} \right) \times 1/t \times 100\% \quad (2)$$

As the application of FISH procedures can lead to underestimates of the isotopic ratio⁹⁵, the calculated relative incorporation represents a minimum estimate.

Finally, to visualize a larger area of the root cross-section (Fig. 3d,e), 11 nanoSIMS images that overlapped by 5 px were stitched based on the position of the nanoSIMS XYZ stage. Due to 3D effects, the raster areas were not always aligned perfectly. Thus, a custom developed cross-correlation algorithm (MATLAB, Mathworks 2018b) was applied with a maximum allowed shift of 5 px to improve stitching. In case offsets were larger, the position was manually corrected. Look@nanosims was modified to allow for reading of the stitched images and subsequent processing according to the same procedure as described above. As nanoSIMS measurements were performed on root cross-sections in the embedding medium, regions without plant material had low count-statistics, which can lead to false ¹⁵N/¹⁴N ratios. To not overemphasize these regions with low count-statistics, we used the thresholding method implemented in look@nanoSIMS for Fig. 3e (for comparison, an example raw image is presented in Extended Data Fig. 4d). Custom codes for processing of nanoSIMS data are available (see the ‘Code availability’ section).

STEM analysis

Paraformaldehyde-fixed root pieces (from June 2014) that were previously used for FISH (see above) were also used for scanning transmission electron microscopy (STEM) imaging. Thin sections (~70 nm) were prepared with the Leica UC7 Ultramicrotome (Ultracut UC7, Leica Microsystems) using a diamond knife, mounted on formvar-coated slot-grids (Agar Scientific). Sections were stained with osmium tetroxide (OsO₄), followed by 0.5% aqueous uranyl acetate (Science Services) for 20 min and 2% Reynold’s lead citrate for 6 min, with three washing steps between each

step. Sections were imaged at 20–30 kV using the Quanta 250 FEG scanning electron microscope (FEI) equipped with a STEM detector using the xT microscope control software (v.6.2.6).

Statistics and reproducibility

No statistical methods were used to predetermine sample size and experiments were not randomized. The investigators were not blinded to allocation during experiments and outcome assessment.

For Fig. [3a–c](#), respectively, the fluorescence images are representative of $n = 3$ images from 1 sample; $n = 9$ images from 4 sections of 1 sample; and $n = 23$ images from 4 sections of 1 sample. In Fig. [3d, e](#), the stitched images are representative of $n = 2$ images from 2 samples.

In Extended Data Fig. [4d](#), the correlative images (FISH and nanoSIMS) are representative of $n = 167$ measurements from 2 samples; 11 of the 167 images were merged and illustrated in Fig. [3d, e](#). Of the 167 measurements, 89 measurements contained both root tissue and *Ca. C. neptuna* cells while 78 measurements contained root tissue only.

In Extended Data Fig. [5a, b](#), the STEM images are representative of $n = 10$ images from 1 sample. In Extended Data Fig. [5c](#), the epifluorescence image is representative of $n = 11$ from 4 sections of 1 sample.

Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

Data availability

Raw reads of the 16S rRNA gene amplicon sequencing, the MAGs of *Ca. C. neptuna* and *C. diazotrophica* (DSM18577), and the mapped reads of the transcriptomes are available under Bioproject number [PRJEB37438](#) at the European Nucleotide Archive (ENA). Sequences that were included in the phylogenetic tree are available in Supplementary Data [1](#) (with accession numbers and references) and as a tree file (Supplementary Data [2](#)). The comparison of 34 genomes for presence/absence of specific genes and/or pathways is available in Supplementary Data [3](#) including their accession numbers. The PhyloFlash results (as presented in Extended Data Fig. [2](#)) are available in Supplementary Data [4](#). Publicly available sequences used for phylogenetic tree construction and genome comparison can be found under their respective

accession numbers at NCBI (<https://www.ncbi.nlm.nih.gov/>) or ENA (<https://www.ebi.ac.uk/ena/browser/home>). Ribosomal subunit databases used for taxonomic classification can be found at the SILVA rRNA database (<https://www.arb-silva.de/>). [Source data](#) are provided with this paper.

Code availability

MATLAB codes used for processing of eddy correlation data and nanoSIMS data can be found at GitHub (<https://github.com/SoerenAhmerkamp/EddyCorrelation> and <https://github.com/SoerenAhmerkamp/NanoSIMS/>).

References

1. 1.

Lilburn, T. C. et al. Nitrogen fixation by symbiotic and free-living spirochetes. *Science* **292**, 2495–2498 (2001).

2. 2.

Poole, P., Ramachandran, V. & Terpolilli, J. Rhizobia: from saprophytes to endosymbionts. *Nat. Rev. Microbiol.* **16**, 291–303 (2018).

3. 3.

Thompson, A. W. et al. Unicellular cyanobacterium symbiotic with a single-celled eukaryotic alga. *Science* **337**, 1546–1550 (2012).

4. 4.

Santi, C., Bogusz, D. & Franche, C. Biological nitrogen fixation in non-legume plants. *Ann. Bot.* **111**, 743–767 (2013).

5. 5.

Waycott, M., Biffin, E. & Les, D. H. in *Seagrasses of Australia* Ch. 5, 129–154 (Springer, 2018).

6. 6.

Knack, J. et al. Microbiomes of streptophyte algae and bryophytes suggest that a functional suite of microbiota fostered plant colonization of land. *Int. J. Plant Sci.* **176**, 405–420 (2015).

7. 7.

McLeod, E. et al. A blueprint for blue carbon: toward an improved understanding of the role of vegetated coastal habitats in sequestering CO₂. *Front. Ecol. Environ.* **9**, 552–560 (2011).

8. 8.

Barbier, E. B. et al. The value of estuarine and coastal ecosystem services. *Ecol. Monogr.* **81**, 169–193 (2011).

9. 9.

Costanza, R. et al. The value of the world's ecosystem services and natural capital. *Nature* **387**, 253–260 (1997).

10. 10.

Waycott, M. et al. Accelerating loss of seagrasses across the globe threatens coastal ecosystems. *Proc. Natl Acad. Sci. USA* **106**, 12377–12381 (2009).

11. 11.

Duarte, C. M. et al. Seagrass community metabolism: assessing the carbon sink capacity of seagrass meadows. *Glob. Biogeochem. Cycles* **24**, <https://doi.org/10.1029/2010gb003793> (2010).

12. 12.

Hemminga, M. A., Harrison, P. G. & van Lent, F. The balance of nutrient losses and gains in seagrass meadows. *Mar. Ecol. Prog. Ser.* **71**, 85–96 (1991).

13. 13.

Lehnen, N. et al. High rates of microbial dinitrogen fixation and sulfate reduction associated with the Mediterranean seagrass *Posidonia oceanica*. *Syst. Appl. Microbiol.* **39**, 476–483 (2016).

14. 14.

Welsh, D. T. Nitrogen fixation in seagrass meadows: Regulation, plant-bacteria interactions and significance to primary productivity. *Ecol. Lett.* **3**, 58–71 (2000).

15. 15.

Seymour, J., Laverock, B., Nielsen, D., Trevathan-Tackett, S. M. & Macreadie, P. I. in *Seagrasses of Australia* (eds Larkum, A. W. D., Kendrick, G. A. & Ralph, P. J.) Ch. 12, 343–392 (Springer, 2018).

16. 16.

Petersen, J. M. et al. Chemosynthetic symbionts of marine invertebrate animals are capable of nitrogen fixation. *Nat. Microbiol.* **2**, <https://doi.org/10.1038/nmicrobiol.2016.195> (2017).

17. 17.

Hurek, T. & Reinhold-Hurek, B. *Azoarcus* sp. strain BH72 as a model for nitrogen-fixing grass endophytes. *J. Biotechnol.* **106**, 169–178 (2003).

18. 18.

Iniguez, A. L., Dong, Y. M. & Triplett, E. W. Nitrogen fixation in wheat provided by *Klebsiella pneumoniae* 342. *Mol. Plant Microbe Interact.* **17**, 1078–1085 (2004).

19. 19.

Wheatley, R. M. et al. Lifestyle adaptations of *Rhizobium* from rhizosphere to symbiosis. *Proc. Natl Acad. Sci. USA* **117**, 23823–23834 (2020).

20. 20.

Udvardi, M. & Poole, P. S. Transport and metabolism in legume-rhizobia symbioses. *Ann. Rev. Plant Biol.* **64**, 781–805 (2013).

21. 21.

Fourqurean, J. W. et al. Seagrass ecosystems as a globally significant carbon stock. *Nat. Geosci.* **5**, 505–509 (2012).

22. 22.

Capone, D. G. in *Nitrogen Cycling in Coastal Marine Environments* (eds Blackburn, T. H. & Sørensen, J.) Ch. 5, 85–123 (John Wiley & Sons Ltd, 1988).

23. 23.

Alcoverro, T., Manzanera, M. & Romero, J. Annual metabolic carbon balance of the seagrass *Posidonia oceanica*: the importance of carbohydrate reserves. *Mar. Ecol. Prog. Ser.* **211**, 105–116 (2001).

24. 24.

Fahimipour, A. K. et al. Global-scale structure of the eelgrass microbiome. *Appl. Environ. Microbiol.* **83**, <https://doi.org/10.1128/aem.03391-16> (2017).

25. 25.

Cramer, M. J., Haghshenas, N., Bagwell, C. E., Matsui, G. Y. & Lovell, C. R. *Celerinatantimonas diazotrophica* gen. nov., sp nov., a nitrogen-fixing bacterium representing a new family in the Gammaproteobacteria, Celerinatantimonadaceae fam. nov. *Int. J. Syst. Evol. Microbiol.* **61**, 1053–1060 (2011).

26. 26.

Yarza, P. et al. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat. Rev. Microbiol.* **12**, 635–645 (2014).

27. 27.

Lodwig, E. M. et al. Amino-acid cycling drives nitrogen fixation in the legume–*Rhizobium* symbiosis. *Nature* **422**, 722–726 (2003).

28. 28.

Flores-Tinoco, C. E. et al. Co-catabolism of arginine and succinate drives symbiotic nitrogen fixation. *Mol. Syst. Biol.* **16**, e9419 (2020).

29. 29.

Prell, J., Bourdès, A., Karunakaran, R., Lopez-Gomez, M. & Poole, P. Pathway of gamma-aminobutyrate metabolism in *Rhizobium leguminosarum* 3841 and its role in symbiosis. *J. Bacteriol.* **191**, 2177–2186 (2009).

30. 30.

Holmer, M., Duarte, C. M. & Marbá, N. Sulfur cycling and seagrass (*Posidonia oceanica*) status in carbonate sediments. *Biogeochemistry* **66**, 223–239 (2003).

31. 31.

Kiers, E. T., Rousseau, R. A., West, S. A. & Denison, R. F. Host sanctions and the legume–*Rhizobium* mutualism. *Nature* **425**, 78–81 (2003).

32. 32.

Hardoim, P. R., van Overbeek, L. S. & van Elsas, J. D. Properties of bacterial endophytes and their proposed role in plant growth. *Trends Microbiol.* **16**, 463–471 (2008).

33. 33.

González, J. E. & Marketon, M. M. Quorum sensing in nitrogen-fixing rhizobia. *Microbiol. Mol. Biol. Rev.* **67**, 574–592 (2003).

34. 34.

Tavares, F., Santos, C. L. & Sellstedt, A. Reactive oxygen species in legume and actinorhizal nitrogen-fixing symbioses: the microsymbiont's responses to an unfriendly reception. *Physiol. Plant.* **130**, 344–356 (2007).

35. 35.

Liu, H. W. et al. Inner plant values: diversity, colonization and benefits from endophytic bacteria. *Front. Microbiol.* **8**, <https://doi.org/10.3389/fmicb.2017.02552> (2017).

36. 36.

Lòpez-Fernàndez, S. et al. Whole-genome comparative analysis of virulence genes unveils similarities and differences between endophytes and other symbiotic bacteria. *Front. Microbiol.* **6**, <https://doi.org/10.3389/fmicb.2015.00419> (2015).

37. 37.

Bulgarelli, D., Schlaeppi, K., Spaepen, S., van Themaat, E. V. L. & Schulze-Lefert, P. Structure and functions of the bacterial microbiota of plants. *Ann. Rev. Plant Biol.* **64**, 807–838 (2013).

38. 38.

Santoyo, G., Moreno-Hagelsieb, G., Orozco-Mosqueda, M. D. & Glick, B. R. Plant growth-promoting bacterial endophytes. *Microbiol. Res.* **183**, 92–99 (2016).

39. 39.

Olsen, J. L. et al. The genome of the seagrass *Zostera marina* reveals angiosperm adaptation to the sea. *Nature* **530**, 331–335 (2016).

40. 40.

Mohnen, D. Pectin structure and biosynthesis. *Curr. Opin. Plant Biol.* **11**, 266–277 (2008).

41. 41.

Ruff, S. E. et al. Methane seep in shallow-water permeable sediment harbors high diversity of anaerobic methanotrophic communities, Elba, Italy. *Front. Microbiol.* **7**, <https://doi.org/10.3389/fmicb.2016.00374> (2016).

42. 42.

Holmes, R. M., Aminot, A., Kérouel, R., Hooker, B. A. & Peterson, B. J. A simple and precise method for measuring ammonium in marine and freshwater ecosystems. *Can. J. Fish. Aquat. Sci.* **56**, 1801–1808 (1999).

43. 43.

Holtappels, M. et al. Effects of transient bottom water currents and oxygen concentrations on benthic exchange rates as assessed by eddy correlation measurements. *J. Geophys. Res. Oceans* **118**, 1157–1169 (2013).

44. 44.

Lorrai, C., McGinnis, D. F., Berg, P., Brand, A. & Wüest, A. Application of oxygen eddy correlation in aquatic systems. *J. Atmos. Oceanic Technol.* **27**, 1533–1546 (2010).

45. 45.

Ahmerkamp, S. et al. Regulation of benthic oxygen fluxes in permeable sediments of the coastal ocean. *Limnol. Oceanogr.* **62**, 1935–1954 (2017).

46. 46.

Dabundo, R. et al. The contamination of commercial $^{15}\text{N}_2$ gas stocks with ^{15}N -labeled nitrate and ammonium and consequences for nitrogen fixation

measurements. *PloS ONE* **9**, <https://doi.org/10.1371/journal.pone.0110335> (2014).

47. 47.

Warembourg, F. R. in *Nitrogen Isotope techniques* (eds Knowles, R. & Blackburn, T. H.) 127–156 (Academic, 1993).

48. 48.

Corr, L. T., Berstan, R. & Evershed, R. P. Optimisation of derivatisation procedures for the determination of delta ^{13}C values of amino acids by gas chromatography/combustion/isotope ratio mass spectrometry. *Rapid Commun. Mass Spectrom.* **21**, 3759–3771 (2007).

49. 49.

Pjevac, P. et al. Metaproteogenomic profiling of microbial communities colonizing actively venting hydrothermal chimneys. *Front. Microbiol.* **9**, <https://doi.org/10.3389/fmicb.2018.00680> (2018).

50. 50.

Herlemann, D. P. R. et al. Transitions in bacterial communities along the 2000 km salinity gradient of the Baltic Sea. *ISME J.* **5**, 1571–1579 (2011).

51. 51.

Bolyen, E. et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**, 852–857 (2019).

52. 52.

Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **17**, 10–12 (2011).

53. 53.

Rognes, T., Flouri, T., Nichols, B., Quince, C. & Mahé, F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ* **4**, <https://doi.org/10.7717/peerj.2584> (2016).

54. 54.

Amir, A. et al. Deblur rapidly resolves single-nucleotide community sequence patterns. *mSystems* **2**, <https://doi.org/10.1128/mSystems.00191-16> (2017).

55. 55.

Pedregosa, F. et al. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

56. 56.

Bokulich, N. A. et al. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2’s q2-feature-classifier plugin. *Microbiome* **6**, <https://doi.org/10.1186/s40168-018-0470-z> (2018).

57. 57.

Edwards, J. et al. Structure, variation, and assembly of the root-associated microbiomes of rice. *Proc. Natl Acad. Sci. USA* **112**, E911–E920 (2015).

58. 58.

Kruskal, W. H. & Wallis, W. A. Use of ranks in one-criterion variance analysis. *J. Am. Stat. Assoc.* **47**, 583–621 (1952).

59. 59.

Martino, C. et al. A novel sparse compositional technique reveals microbial perturbations. *mSystems* **4**, <https://doi.org/10.1128/mSystems.00016-19> (2019).

60. 60.

Vázquez-Baeza, Y., Pirrung, M., Gonzalez, A. & Knight, R. EMPeror: a tool for visualizing high-throughput microbial community data. *Gigascience* **2**, <https://doi.org/10.1186/2047-217x-2-16> (2013).

61. 61.

Anderson, M. J. A new method for non-parametric multivariate analysis of variance. *Austral Ecol.* **26**, 32–46 (2001).

62. 62.

Morton, J. T. et al. Establishing microbial composition measurement standards with reference frames. *Nat. Commun.* **10**, <https://doi.org/10.1038/s41467-019-1038-s> (2019).

[10656-5](#) (2019).

63. 63.

Fedarko, M. W. et al. Visualizing 'omic feature rankings and log-ratios using Qurro. *NAR Genom. Bioinform.* **2**, lqaa023 (2020).

64. 64.

McMurdie, P. J. & Holmes, S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8**, e61217 (2013).

65. 65.

Gruber-Vodicka, H. R., Seah, B. K. & Pruesse, E. phyloFlash: rapid small-subunit rRNA profiling and targeted assembly from metagenomes. *mSystems* **5**, e00920-20 (2020).

66. 66.

Kolmogorov, M. et al. metaFlye: scalable long-read metagenome assembly using repeat graphs. *Nat. Methods* **17**, 1103–1110 (2020).

67. 67.

Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

68. 68.

Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

69. 69.

Walker, B. J. et al. Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE* **9**, <https://doi.org/10.1371/journal.pone.0112963> (2014).

70. 70.

Parks, D. H., Imelfort, M., Skennerton, C. T., Hugenholtz, P. & Tyson, G. W. CheckM: assessing the quality of microbial genomes recovered from isolates,

- single cells, and metagenomes. *Genome Res.* **25**, 1043–1055 (2015).
71. 71.
- Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).
72. 72.
- Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).
73. 73.
- Aziz, R. K. et al. The RAST server: rapid annotations using subsystems technology. *BMC Genom.* **9**, <https://doi.org/10.1186/1471-2164-9-75> (2008).
74. 74.
- Zhang, H. et al. dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* **46**, W95–W101 (2018).
75. 75.
- Yin, Y. B. et al. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* **40**, W445–W451 (2012).
76. 76.
- Quast, C. et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* **41**, D590–D596 (2013).
77. 77.
- Pruesse, E., Peplies, J. & Glöckner, F. O. SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* **28**, 1823–1829 (2012).
78. 78.
- Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
79. 79.

Kopylova, E., Noé, L. & Touzet, H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28**, 3211–3217 (2012).

80. 80.

Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).

81. 81.

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).

82. 82.

Wagner, G. P., Kin, K. & Lynch, V. J. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory Biosci.* **131**, 281–285 (2012).

83. 83.

Alikhan, N. F., Petty, N. K., Ben Zakour, N. L. & Beatson, S. A. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genom.* **12**,<https://doi.org/10.1186/1471-2164-12-402> (2011).

84. 84.

Boudvillain, M., Figueroa-Bossi, N. & Bossi, L. Terminator still moving forward: expanding roles for Rho factor. *Curr. Opin. Microbiol.* **16**, 118–124 (2013).

85. 85.

Martens, M. et al. Advantages of multilocus sequence analysis for taxonomic studies: a case study using 10 housekeeping genes in the genus *Ensifer* (including former *Sinorhizobium*). *Int. J. System. Evol. Microbiol.* **58**, 200–214 (2008).

86. 86.

Metcalf, D., Sharif, S. & Weese, J. S. Evaluation of candidate reference genes in *Clostridium difficile* for gene expression normalization. *Anaerobe* **16**, 439–443 (2010).

87. 87.

Naser, S. M. et al. Application of multilocus sequence analysis (MLSA) for rapid identification of *Enterococcus* species based on *rpoA* and *pheS* genes. *Microbiology* **151**, 2141–2150 (2005).

88. 88.

Rocha, D. J. P., Santos, C. S. & Pacheco, L. G. C. Bacterial reference genes for gene expression studies by RT-qPCR: survey and analysis. *Antonie Van Leeuwenhoek* **108**, 685–693 (2015).

89. 89.

Zeigler, D. R. Gene sequences useful for predicting relatedness of whole genomes in bacteria. *Int. J. Syst. Evol. Microbiol.* **53**, 1893–1900 (2003).

90. 90.

McDonald, K. L. Out with the old and in with the new: rapid specimen preparation procedures for electron microscopy of sectioned biological material. *Protoplasma* **251**, 429–448 (2014).

91. 91.

Amann, R. I., Krumholz, L. & Stahl, D. A. Fluorescent oligonucleotide probing of whole cells for determinative, phylogenetic, and environmental studies in microbiology. *J. Bacteriol.* **172**, 762–770 (1990).

92. 92.

Wallner, G., Amann, R. & Beisker, W. Optimizing fluorescent in situ hybridization with ribosomal-RNA-targeted oligonucleotide probes for flow cytometric identification of microorganisms. *Cytometry* **14**, 136–143 (1993).

93. 93.

Stoecker, K., Dorninger, C., Daims, H. & Wagner, M. Double labeling of oligonucleotide probes for fluorescence in situ hybridization (DOPE-FISH) improves signal intensity and increases rRNA accessibility. *Appl. Environ. Microbiol.* **76**, 922–926 (2010).

94. 94.

Polerecky, L. et al. Look@ NanoSIMS—a tool for the analysis of nanoSIMS data in environmental microbiology. *Environ. Microbiol.* **14**, 1009–1023 (2012).

95. 95.

Musat, N. et al. The effect of FISH and CARD-FISH on the isotopic composition of ^{13}C - and ^{15}N -labeled *Pseudomonas putida* cells measured by nanoSIMS. *Syst. Appl. Microbiol.* **37**, 267–276 (2014).

Acknowledgements

We thank the staff at the HYDRA Marine Sciences and HYDRA Fieldwork for field sampling and assistance as well as scientific discussions during our study; staff at the National Park Tuscan Archipelago, Portoferraio, Italy for granting access to the protected waters of the Island of Pianosa (permit no. 2930/2017); N. Dubilier for valuable discussions and the organization of two workshops on symbiotic ecology at the Elba Field Station; J.-H. Hehemann, B. Reinhold-Hurek, T. Hurek, L. van Niftrik, M. Marín and B. Kartal for discussions; and T. Alarcon Schumacher, P. Bourceau, C. Cornet, J. Dekaezemacker, P. Downes, A. Frayssinet, B. Fuchs, P. Hach, K. Imhoff, F. Moin Jalaluddin, A. Kidane, K. Kitzinger, G. Klockgether, S. Lilienthal, M. Maeke, S. Murugan, W. Neweshy, S. Piasek, S. Robert, N. Rujanski, S. Schorn, A. Schwedt, P. Stücheli, D. Tienken and B. Vekeman for technical assistance and help with sampling and sample preparations. C.J.S. was funded by internal Eawag funds. This study was funded by the Max Planck Society.

Funding

Open access funding provided by Max Planck Society.

Author information

Author notes

1. Pelin Yilmaz

Present address: Data Science Research Group, Institute for Artificial Intelligence in Medicine, University Hospital Essen, Essen, Germany

Affiliations

1. Max Planck Institute for Marine Microbiology, Bremen, Germany

Wiebke Mohr, Nadine Lehnens, Soeren Ahmerkamp, Hannah K. Marchant, Jon S. Graf, Bernhard Tschitschko, Pelin Yilmaz, Sten Littmann, Harald Gruber-Vodicka, Nikolaus Leisch, Jana Milucka & Marcel M. M. Kuypers

2. HYDRA Marine Sciences GmbH, Bühl, Germany

Miriam Weber & Christian Lott

3. Swiss Federal Institute of Aquatic Science and Technology (Eawag), Department of Surface Waters-Research and Management, Kastanienbaum, Switzerland

Carsten J. Schubert

Contributions

W.M., N. Lehnens, H.K.M. and M.M.M.K. performed N₂ fixation experiments, and W.M. and N. Lehnens analysed N₂ fixation data. N. Lehnens carried out embedding, FISH and microscopy on root cross-sections as well as nucleic acid and amino acid extractions of root pieces. S.A. performed EC measurements and analysed and processed microscopic and nanoSIMS images. N. Lehnens, J.S.G. and B.T. analysed amplicon-sequencing data, metagenomes, MAGs and metatranscriptomes, and performed genome comparisons. S.L. carried out nanoSIMS measurements. P.Y. and H.G.-V. carried out initial bioinformatics analyses. C.J.S. provided amino acid data. M.W. and C.L. carried out field sampling and in situ field measurements. N. Leisch carried out STEM and provided help with embedding, sectioning and FISH. M.M.M.K. conceived the project. W.M., N. Lehnens, H.K.M., J.M. and M.M.M.K. designed the study. W.M., J.M. and M.M.M.K. wrote the manuscript with contributions from all of the co-authors.

Corresponding author

Correspondence to [Wiebke Mohr](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Douglas Capone, Susannah Tringe 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 Productivity in seagrass sediments and non-vegetated sediments.](#)

Hourly-averaged oxygen fluxes (green) and photosynthetically active radiation (PAR; grey) measured over a daily cycle in June 2019 for the *P. oceanica* meadow (**a**) and unvegetated sandy sediments (**b**). Negative fluxes during night indicate net respiration, whereas positive fluxes during the day indicate net photosynthesis. The small positive oxygen fluxes in the lower panel indicate the presence of benthic algae. (**c**) Oxygen (O₂) concentrations at the rhizome zone-sediment interface showing the depletion of O₂ at about 4 mm depth.

[Source Data](#)

[Extended Data Fig. 2 Composition of microbial communities associated with *P. oceanica* plant tissues and seagrass sediment.](#)

((a)–(c)) 16S rRNA gene-based community composition of the top ten taxa on domain-level (a), class-level (only Bacteria) (b), and families within the order Alteromonadales (c) in metagenomes associated with *P. oceanica* plant tissues (Leaf, Rhizome, Root 1, Root 2) and the sediment within the *P. oceanica* seagrass meadow (replicate cores C2, C3, C6; 2–10 cm depth horizon). Note that while reads associated with Alteromonadales were detected in all samples, reads assigned to Celerinatantimonadaceae were only detected in plant-associated samples (mainly roots and rhizome). Reads taxonomically assigned to either mitochondria, chloroplasts or Eukarya were not included in the analysis. (d) Read recruitment to *Ca. C. neptuna* from metagenomes sampled from different *P. oceanica* plant tissues and seagrass bed sediment. Metagenomic reads mapped onto the genome of *Ca. C. neptuna* and the number of mapped reads normalized to one million total reads is shown above each bar. Reads mapping to rRNA genes were not counted in this analysis. (e) Non-metric multidimensional scaling (NMDS) ordination plot showing changes in the prokaryotic community composition (Class-level) associated with different *P. oceanica* plant tissues and seagrass bed sediment. Community composition on Class-level (see panel b) was derived from metagenomic 16S rRNA gene sequences sampled from *P. oceanica* plant tissues (Leaf, Rhizome, Root 1, Root 2) and the sediment within the *P. oceanica* seagrass meadow (replicate cores C2, C3, C6; 2–10 cm depth horizon). Ordinations are based on Bray-Curtis dissimilarity.

[Source Data](#)

[Extended Data Fig. 3 Root-associated microbial community analyses of N₂-fixing and non-N₂-fixing *P. oceanica* plants.](#)

(a) N₂ fixation rates of roots of individual plants with each symbol representing individual root pieces. (b) Relative abundance of 16S rRNA gene-based OTUs ('others' are classes with less than 1.5%) in individually analysed plants. Asterisks indicate plants with sequencing results below the threshold for the calculation of α-diversity. (c) Ratio of *Celerinatantimonas*-related reads to organellar reads. Lines, boxes, and error bars represent mean, 25th and 75th percentiles and standard deviation,

respectively. (d) Number of shared OTUs as a function of their presence in an increasing number of plants (as % of plants in both categories) with Venn diagram showing the number of unique and shared OTUs. (e) α -diversity indices where boxes indicate second and third quartiles, whiskers indicate first and fourth quartiles, lines indicate median values and black dots are the individual plants. Differences in α -diversity indices were not statistically significant (Kruskal-Wallis pairwise tests). (f) First two axes of a principal component analysis (PCA) using Aitchison distance based on OTU counts, highlighting statistically significant microbial community compositions in N₂-fixing (square symbols) and non-N₂-fixing (diamond symbols) plants (pairwise Permanova, pseudo-F = 7.9, q-value = 0.001). Colors indicate seasons (spring (blue), summer (yellow) and autumn (orange); the brown diamond represents a spring and a summer sample on top of each other). Arrow indicates the OTU contributing most to the clustering of samples, i.e. *Ca. C. neptuna*. (g) Ranking of differentially abundant OTUs (x-axis). Negative and positive log ratios indicate higher abundance in non-N₂-fixing and N₂-fixing plants, respectively. Highlighted in magenta is the *Ca. C. neptuna*-OTU, the highest ranked OTU positively associated with N₂ fixation. The number of plants (n) included in the analyses is indicated in parentheses in panels c and e.

Source Data

Extended Data Fig. 4 Distribution, abundance and activity of *Ca. C. neptuna*.

(a) Stitched epifluorescence images (black-and-white inverted) of root cross-sections from April 2016 (left), June 2014 (middle) and August 2016 (right) showing primary locations of *Ca. C. neptuna* cells (as visualized by FISH) in red (signal-amplified for easier visualization; [Methods](#)), (b) Epifluorescence image (from panel a) indicating the different root tissues (as in c), (c) absolute and relative abundance of *Ca. C. neptuna* cells (visualized by FISH) in the different root tissues of individual *P. oceanica* root sections (0.5 – 1.0 μm thickness), (d) Correlative epifluorescence (FISH and autofluorescence) and nanoSIMS (¹⁵N/¹⁴N) images (left) with the retrieved surface areas and correlative nanoSIMS data for *Ca. C.*

neptuna cells (middle) and root tissue (right) ([Methods](#)), x- and y-axes are in pixels with 256 pixels = 20 μm , and the color scale represents the $^{15}\text{N}/^{14}\text{N}$ ratio; (e) single-cell $^{15}\text{N}_2$ fixation activity (*Ca. C. neptuna* cells) or transfer of freshly fixed N (root tissue) as relative incorporation per day in *Ca. C. neptuna* cells, root tissue with *Ca. C. neptuna* cells in the same image (roots with cells) and root tissue without *Ca. C. neptuna* cells in the same image (roots without cells) obtained from a total of 167 measurements as shown in (d) from two cross-sections.

[Source Data](#)

[Extended Data Fig. 5 STEM and epifluorescence images of seagrass root sections.](#)

(a, b) STEM images with black arrowheads indicating bacteria residing in intercellular space (a) as well as inside plant cells (b). Based on the location as well as the relative and absolute abundances of *Ca. C. neptuna* in root cross-sections (Extended Data Fig. 4), the majority of bacteria visible are likely *Ca. C. neptuna* cells. (c) Epifluorescence image of a root cross-section showing *Ca. C. neptuna* cells (overlay image of DAPI (blue), autofluorescence (green/orange) and FISH-positive cells (pinkish color due to the overlap of DAPI and FISH probe (orange) signals) inside the stele.

[Extended Data Fig. 6 Genome and transcriptomes of *Ca. C. neptuna*.](#)

General genome characteristics for the novel species *Ca. C. neptuna* and its closest relative *C. diazotrophica* (DSM18577) (top). *Ca. C. neptuna* genome is represented circularized (bottom). Circles from the inside to the outside: (1) GC content, (2) GC skew, (3) Illumina metagenome (June 2014) coverage (coverage range 0-250), (4) PacBio metagenome (August 2016) coverage (coverage range 0-200), (5-9) gene transcription in Illumina-based transcriptomes plotted as TPM for protein coding locus tags (TPM range 0-1000; genes with TPM values > 1000 are highlighted in dark blue), (10) regions present in *C. diazotrophica* based on BLASTn searches, shown are regions with similarity between 50% (grey) – 100% (orange),

(11) selected features (genes and gene clusters) including highly transcribed ones (blue, mean expression across all transcriptomes >1000 TPM). A list of genes and features is included in the Supplementary Information.

Extended Data Fig. 7 Gene transcription of selected genes/pathways.

Normalized gene transcription (TPM/TPM_{house} = TPM relative to TPM of housekeeping genes) of selected *Ca. C. neptuna* pathways in transcriptomes of five N₂-fixing plants (1 plant = 1 column). The genes averaged for analysis are indicated in parentheses. For glycolysis and the TCA cycle, the following genes were pooled for analysis: *pgi*, *pfkA*, *fbaA*, *tpiA*, *gapA*, *pgk*, *gpmA*, *gpmI*, *eno*, *pykF* and *gltA*, *acnB*, *icd*, *lpdA*, *sucA*, *sucB*, *sucC*, *sucD*, *sdhC*, *sdhD*, *sdhA*, *sdhB*, *fumB*, *mdh*, *mqo*, respectively. Measured N₂ fixation rates of the respective plants are shown on top. The (non-linear) color-coding was chosen so that transcription similar to housekeeping genes or higher is presented as yellow while lower expression levels are gradient-colored between blue and yellow.

Source Data

Extended Data Fig. 8 Enrichment of ¹⁵N in amino acids.

Enrichment of ¹⁵N in individual amino acids relative to the enrichment of ¹⁵N in glutamate, the first amino acid synthesized upon N₂ fixation, in root-extracted proteins after incubations of plants with ¹⁵N₂ ([Methods](#); up to four plants that yielded sufficient amino acids for δ¹⁵N measurement are shown here). **a)** Amino acids synthesized from intermediates of the glycolysis or the pentose phosphate pathway and **b)** amino acids synthesized from intermediates of the TCA cycle or via transamination of another amino acid (or intermediates) with glutamate. Slopes of the linear regressions are given behind the three-letter code for each amino acid. A slope of 1.0 indicates ¹⁵N enrichments that are identical to those of glutamate with smaller slopes indicating lower ¹⁵N enrichments (see Supplementary Discussion). Phe, phenylalanine (R²=0.991); Leu, Leucine (R²=0.995); Val, Valine

($R^2=0.993$); Ser, Serine ($R^2=0.998$); Ala, Alanine ($R^2=0.819$); Lys, Lysine ($R^2=0.979$); Thr, Threonine ($R^2=0.986$); Pro, Proline ($R^2=0.988$); Gly, Glycine ($R^2=0.992$); Asp, Aspartate ($R^2=0.999$).

Source Data

Extended Data Fig. 9 Model of main metabolic interactions of *Ca. C. neptuna* and *P. oceanica*.

Blue and orange arrows indicate transfer of metabolites from *Ca. C. neptuna* to *P. oceanica* and vice versa, respectively. I, respiratory complex I (NADH dehydrogenase); II, respiratory complex II (succinate dehydrogenase); V, respiratory complex V (ATP synthase); CytbdI, *bdI*-type terminal quinol oxidase; OAA, oxaloacetate; TCA, tricarboxylic acid cycle; 2-OG, 2-oxoglutarate; SSA, succinic semialdehyde; PTS, phosphotransferase system (sugar transport); Aap/Liv, Amino acid permease / branched-chain amino acid transporter; GABA, 4-aminobutyrate; NQR, sodium-translocating NADH:ubiquinone reductase; Glu, glutamate; Leu, leucine; Phe, phenylalanine; AA, amino acids.

Extended Data Table 1 Nutrient concentrations associated with *P. oceanica* meadow and nearby sediments

Supplementary information

Supplementary Information

Supplementary Notes 1–11.

Reporting Summary

Supplementary Data 1

A list of sequence entries shown in the phylogenetic tree in Fig. 2c including accession numbers and references.

Supplementary Data 2

Tree-file belonging to the phylogenetic tree in Fig. 2c.

Supplementary Data 3

List of the presence or absence of genes involved in selected pathways retrieved from 34 available genomes (including the ones from this study).

Supplementary Data 4

PhyloFlash results from the metagenomes as presented in Extended Data Fig. 2.

Peer Review File

Source data

Source Data Fig. 1

Source Data Fig. 2

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. 7

Source Data Extended Data Fig. 8

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

Mohr, W., Lehn, N., Ahmerkamp, S. *et al.* Terrestrial-type nitrogen-fixing symbiosis between seagrass and a marine bacterium. *Nature* **600**, 105–109 (2021). <https://doi.org/10.1038/s41586-021-04063-4>

- Received: 29 March 2021
- Accepted: 22 September 2021
- Published: 03 November 2021
- Issue Date: 02 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04063-4>

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

- [**A seagrass harbours a nitrogen-fixing bacterial partner**](#)
 - Douglas G. Capone

Nature (2021)

[**A seagrass harbours a nitrogen-fixing bacterial partner**](#)

- Douglas G. Capone

News & Views 03 Nov 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04063-4>

- Article
- [Published: 24 November 2021](#)

The human microbiome encodes resistance to the antidiabetic drug acarbose

- [Jared Balaich](#)¹,
- [Michael Estrella](#) [ORCID: orcid.org/0000-0002-1721-9418](#)¹,
- [Guojun Wu](#) [ORCID: orcid.org/0000-0002-7910-2813](#)²,
- [Philip D. Jeffrey](#)¹,
- [Abhishek Biswas](#)^{1,3},
- [Liping Zhao](#)^{2,4},
- [Alexei Korennykh](#)¹ &
- [Mohamed S. Donia](#) [ORCID: orcid.org/0000-0002-9604-2912](#)^{1,5,6}

[Nature](#) volume 600, pages 110–115 (2021)

- 4247 Accesses
- 75 Altmetric
- [Metrics details](#)

Subjects

- [Enzyme mechanisms](#)
- [Microbiome](#)
- [X-ray crystallography](#)

Abstract

The human microbiome encodes a large repertoire of biochemical enzymes and pathways, most of which remain uncharacterized. Here, using a metagenomics-based search strategy, we discovered that bacterial members of the human gut and oral microbiome encode enzymes that selectively phosphorylate a clinically used antidiabetic drug, acarbose^{1,2}, resulting in its inactivation. Acarbose is an inhibitor of both human and bacterial α -glucosidases³, limiting the ability of the target organism to metabolize complex carbohydrates. Using biochemical assays, X-ray crystallography and metagenomic analyses, we show that microbiome-derived acarbose kinases are specific for acarbose, provide their harbouring organism with a protective advantage against the activity of acarbose, and are widespread in the microbiomes of western and non-western human populations. These results provide an example of widespread microbiome resistance to a non-antibiotic drug, and suggest that acarbose resistance has disseminated in the human microbiome as a defensive strategy against a potential endogenous producer of a closely related molecule.

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: Metagenomic discovery of Maks from the human microbiome.



Fig. 2: Mak1 and AcbK phosphorylate and inactivate acarbose with similar enzyme kinetics.

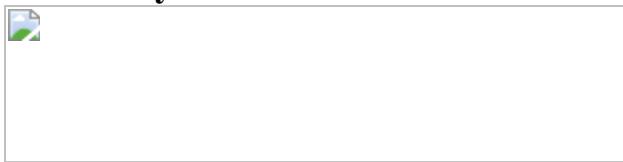
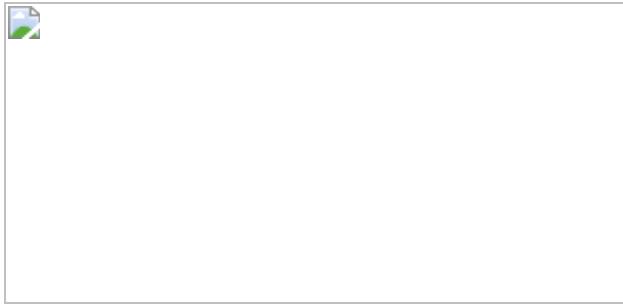


Fig. 3: Crystal structures of Mak1 and AcbK reveal their structural similarity and specific acarbose interacting residues.



Fig. 4: Biological relevance of Maks and a potential origin for Mak1 in the human oral microbiome.



Data availability

Associated data are provided in Extended Data Figs. 1–9 and Supplementary Tables 1–10. Structures of seleno-methionine Mak1, Mak1 bound to AMP-PNP and acarbose and AcbK bound to AMP-PMP and acarbose are available in the Protein Data Bank under accession codes [6WB4](#), [6WB5](#) and [6WB7](#), respectively. Publicly available datasets used in the metagenomic analyses performed here were obtained as follows: HMP-1-1 (43021), Chinese (PRJNA422434), MetaHIT (PRJEB1220), Fijicomp (PRJNA217052), HMP-1-2 ([PRJNA275349](#), [PRJNA48479](#)), human clinical trial with acarbose ([PRJEB14155](#)). The following datasets were used in metatranscriptomic analyses: PRJNA354235, PRJNA389280, PRJNA398089, PRJEB4673, PRJNA221620, <https://www.ncbi.nlm.nih.gov/bioproject/5148>, [PRJNA383868](#).

Code availability

MetaBGC (v.1.3.3, <https://github.com/donia-lab/MetaBGC>) was used to discover AcbK homologues in unassembled metagenomic sequencing data, following the details and parameters described in the [Methods](#).

References

1. 1.

- Chiasson, J. L. et al. Acarbose for prevention of type 2 diabetes mellitus: the STOP-NIDDM randomised trial. *Lancet* **359**, 2072–2077 (2002).

2. 2.

Wehmeier, U. F. & Piepersberg, W. Biotechnology and molecular biology of the alpha-glucosidase inhibitor acarbose. *Appl. Microbiol. Biotechnol.* **63**, 613–625 (2004).

3. 3.

Yoon, S.-H. & Robyt, J. F. Study of the inhibition of four alpha amylases by acarbose and its 4IV- α -maltohexaosyl and 4IV- α -maltododecaosyl analogues. *Carbohydr. Res.* **338**, 1969–1980 (2003).

4. 4.

Maurice, C. F., Haiser, H. J. & Turnbaugh, P. J. Xenobiotics shape the physiology and gene expression of the active human gut microbiome. *Cell* **152**, 39–50 (2013).

5. 5.

Wu, H. et al. Metformin alters the gut microbiome of individuals with treatment-naive type 2 diabetes, contributing to the therapeutic effects of the drug. *Nat. Med.* **23**, 850–858 (2017).

6. 6.

Maier, L. et al. Extensive impact of non-antibiotic drugs on human gut bacteria. *Nature* **555**, 623–628 (2018).

7. 7.

Whang, A., Nagpal, R. & Yadav, H. Bi-directional drug-microbiome interactions of anti-diabetics. *eBioMedicine* **39**, 591–602 (2019).

8. 8.

Le Bastard, Q. et al. Systematic review: human gut dysbiosis induced by non-antibiotic prescription medications. *Aliment. Pharmacol. Ther.* **47**, 332–345 (2018).

9. 9.

Maruhama, Y. et al. Effects of a glucoside-hydrolase inhibitor (Bay g 5421) on serum lipids, lipoproteins and bile acids, fecal fat and bacterial flora, and intestinal gas production in hyperlipidemic patients. *Tohoku J. Exp. Med.* **132**, 453–462 (1980).

10. 10.

Su, B. et al. Acarbose treatment affects the serum levels of inflammatory cytokines and the gut content of bifidobacteria in Chinese patients with type 2 diabetes mellitus. *J. Diabetes* **7**, 729–739 (2015).

11. 11.

Zhang, X. et al. Effects of acarbose on the gut microbiota of prediabetic patients: a randomized, double-blind, controlled crossover trial. *Diabetes Ther.* **8**, 293–307 (2017).

12. 12.

Santilli, A. D., Dawson, E. M., Whitehead, K. J. & Whitehead, D. C. Nonmicrobicidal small molecule inhibition of polysaccharide metabolism in human gut microbes: a potential therapeutic avenue. *ACS Chem. Biol.* **13**, 1165–1172 (2018).

13. 13.

Baxter, N. T., Lesniak, N. A., Sinani, H., Schloss, P. D. & Koropatkin, N. M. The glucoamylase inhibitor acarbose has a diet-dependent and reversible effect on the murine gut microbiome. *mSphere* **4**, <https://doi.org/10.1128/mSphere.00528-18> (2019).

14. 14.

Zhang, M. et al. Effects of metformin, acarbose, and sitagliptin monotherapy on gut microbiota in Zucker diabetic fatty rats. *BMJ Open Diabetes Res. Care* **7**, e000717 (2019).

15. 15.

Ahr, H. J. et al. Pharmacokinetics of acarbose. Part I: absorption, concentration in plasma, metabolism and excretion after single administration of [¹⁴C]acarbose to rats, dogs and man. *Arzneimittelforschung* **39**, 1254–1260 (1989).

16. 16.

Zhao, L. et al. Gut bacteria selectively promoted by dietary fibers alleviate type 2 diabetes. *Science* **359**, 1151–1156 (2018).

17. 17.

Wehmeier, U. F. The biosynthesis and metabolism of acarbose in *Actinoplanes* sp. SE 50/110: a progress report. *Biocatal. Biotransform.* **21**, 279–284 (2003).

18. 18.

Schmidt, D. D. et al. Alpha-glucosidase inhibitors. New complex oligosaccharides of microbial origin. *Naturwissenschaften* **64**, 535–536 (1977).

19. 19.

Drepper, A. & Pape, H. Acarbose 7-phosphotransferase from *Actinoplanes* sp.: purification, properties, and possible physiological function. *J. Antibiot.* **49**, 664–668 (1996).

20. 20.

Goeke, K., Drepper, A. & Pape, H. Formation of acarbose phosphate by a cell-free extract from the acarbose producer *Actinoplanes* sp. *J. Antibiot.* **49**, 661–663 (1996).

21. 21.

Human Microbiome Project, C. Structure, function and diversity of the healthy human microbiome. *Nature* **486**, 207–214 (2012).

22. 22.

Qin, J. et al. A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature* **490**, 55–60 (2012).

23. 23.

Nielsen, H. B. et al. Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. *Nat. Biotechnol.* **32**, 822–828 (2014).

24. 24.

Brito, I. L. et al. Mobile genes in the human microbiome are structured from global to individual scales. *Nature* **535**, 435–439 (2016).

25. 25.

Lloyd-Price, J. et al. Strains, functions and dynamics in the expanded Human Microbiome Project. *Nature* **550**, 61–66 (2017).

26. 26.

Sugimoto, Y. et al. A metagenomic strategy for harnessing the chemical repertoire of the human microbiome. *Science* **366**, eaax9176 (2019).

27. 27.

Rockser, Y. & Wehmeier, U. F. The gac-gene cluster for the production of acarbose from *Streptomyces glaucescens* GLA.O: identification, isolation and characterization. *J. Biotechnol.* **140**, 114–123 (2009).

28. 28.

Guo, X. et al. Draft genome sequence of *Streptomyces coelicoflavus* ZG0656 reveals the putative biosynthetic gene cluster of acarviostatin family alpha-amylase inhibitors. *Lett. Appl. Microbiol.* **55**, 162–169 (2012).

29. 29.

Parducci, R. E., Cabrera, R., Baez, M. & Guixe, V. Evidence for a catalytic Mg²⁺ ion and effect of phosphate on the activity of *Escherichia coli* phosphofructokinase-2: regulatory properties of a ribokinase family member. *Biochemistry* **45**, 9291–9299 (2006).

30. 30.

Miller, B. G. & Raines, R. T. Identifying latent enzyme activities: substrate ambiguity within modern bacterial sugar kinases. *Biochemistry* **43**, 6387–6392 (2004).

31. 31.

Fong, D. H. & Berghuis, A. M. Substrate promiscuity of an aminoglycoside antibiotic resistance enzyme via target mimicry. *EMBO J.* **21**, 2323–2331 (2002).

32. 32.

McAuley, M., Huang, M. & Timson, D. J. Dynamic origins of substrate promiscuity in bacterial galactokinases. *Carbohydr. Res.* **486**, 107839 (2019).

33. 33.

Sigrell, J. A., Cameron, A. D., Jones, T. A. & Mowbray, S. L. Structure of *Escherichia coli* ribokinase in complex with ribose and dinucleotide determined to 1.8 Å resolution: insights into a new family of kinase structures. *Structure* **6**, 183–193 (1998).

34. 34.

Yeung, M. K. & Kozelsky, C. S. Transformation of *Actinomyces* spp. by a gram-negative broad-host-range plasmid. *J. Bacteriol.* **176**, 4173–4176 (1994).

35. 35.

Flint, H. J., Scott, K. P., Duncan, S. H., Louis, P. & Forano, E. Microbial degradation of complex carbohydrates in the gut. *Gut Microbes* **3**, 289–306 (2012).

36. 36.

Patnode, M. L. et al. Interspecies competition impacts targeted manipulation of human gut bacteria by fiber-derived glycans. *Cell* **179**, 59–73 (2019).

37. 37.

Rakoff-Nahoum, S., Coyne, M. J. & Comstock, L. E. An ecological network of polysaccharide utilization among human intestinal symbionts. *Curr. Biol.* **24**, 40–49 (2014).

38. 38.

Leimena, M. M. et al. A comprehensive metatranscriptome analysis pipeline and its validation using human small intestine microbiota datasets. *BMC Genom.* **14**, 530 (2013).

39. 39.

Goodrich, J. K. et al. Genetic determinants of the gut microbiome in UK twins. *Cell Host Microbe* **19**, 731–743 (2016).

40. 40.

Villmones, H. C. et al. Species level description of the human ileal bacterial microbiota. *Sci. Rep.* **8**, 4736 (2018).

41. 41.

Fung, T. C. et al. Intestinal serotonin and fluoxetine exposure modulate bacterial colonization in the gut. *Nat. Microbiol.* **4**, 2064–2073 (2019).

42. 42.

Kumar, S., Stecher, G. & Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870–1874 (2016).

43. 43.

Abu-Ali, G. S. et al. Metatranscriptome of human faecal microbial communities in a cohort of adult men. *Nat. Microbiol.* **3**, 356–366 (2018).

44. 44.

Schirmer, M. et al. Dynamics of metatranscription in the inflammatory bowel disease gut microbiome. *Nat. Microbiol.* **3**, 337–346 (2018).

45. 45.

Lloyd-Price, J. et al. Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature* **569**, 655–662 (2019).

46. 46.

Peterson, S. N. et al. Functional expression of dental plaque microbiota. *Front. Cell Infect. Microbiol.* **4**, 108 (2014).

47. 47.

Benitez-Paez, A., Belda-Ferre, P., Simon-Soro, A. & Mira, A. Microbiota diversity and gene expression dynamics in human oral biofilms. *BMC Genom.* **15**, 311 (2014).

48. 48.

Jorth, P. et al. Metatranscriptomics of the human oral microbiome during health and disease. *mBio* **5**, e01012-14 (2014).

49. 49.

Szafranski, S. P. et al. Functional biomarkers for chronic periodontitis and insights into the roles of *Prevotella nigrescens* and *Fusobacterium nucleatum*; a metatranscriptome analysis. *NPJ Biofilms Microbiomes* **1**, 15017 (2015).

50. 50.

Schmieder, R. & Edwards, R. Quality control and preprocessing of metagenomic datasets. *Bioinformatics* **27**, 863–864 (2011).

51. 51.

Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).

52. 52.

Chen, I. A. et al. IMG/M v.5.0: an integrated data management and comparative analysis system for microbial genomes and microbiomes. *Nucleic Acids Res.* **47**, D666–D677 (2019).

53. 53.

Kabsch, W. Xds. *Acta Crystallogr. D* **66**, 125–132 (2010).

54. 54.

Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. D* **69**, 1204–1214 (2013).

55. 55.

Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D* **67**, 235–242 (2011).

56. 56.

Sheldrick, G. M. A short history of SHELX. *Acta Crystallogr. A* **64**, 112–122 (2008).

57. 57.

Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).

58. 58.

McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).

59. 59.

Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).

60. 60.

Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).

61. 61.

Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

62. 62.

Kearse, M. et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**, 1647–1649 (2012).

Acknowledgements

We thank the NSLS-2 AMX and FMX beamline staff for their assistance with data collection and the crystallography core facility at the Department of Molecular Biology, Princeton University. We thank M. Cahn and M. Elmassry for assistance with metagenomic data analysis; S. Chatterjee for general assistance; and the rest of the Donia laboratory for discussions. Funding for this project was provided by an NIH Director's New Innovator Award (1DP2AI124441) and the Pew Biomedical Scholars Program to M.S.D.; an NIH grant (1R01GM110161), a Burroughs Wellcome Foundation Grant (1013579) and an award from The Vallee Foundation to A.K. J.B. is funded by a training grant from the National Institute of General Medicine Sciences (NIGMS) (T32GM007388) and L.Z. is a CIFAR fellow. This research used the AMX and FMX beamlines of the National Synchrotron Light Source II, a United States Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Brookhaven National Laboratory under contract no. DE-SC0012704. The Life Science Biomedical Technology Research resource, which supports AMX and FMX, is primarily supported by the NIH (NIGMS) through a Biomedical Technology Research Resource P41 grant (P41GM111244), and by the DOE Office of Biological and Environmental Research (KP1605010).

Author information

Affiliations

1. Department of Molecular Biology, Princeton University, Princeton, NJ, USA
Jared Balaich, Michael Estrella, Philip D. Jeffrey, Abhishek Biswas, Alexei Korennyykh & Mohamed S. Donia
2. Center for Microbiome, Nutrition, and Health, New Jersey Institute for Food, Nutrition, and Health, Department of Biochemistry and Microbiology, Rutgers University, New Brunswick, NJ, USA

Guojun Wu & Liping Zhao

3. Research Computing, Office of Information Technology, Princeton University, Princeton, NJ, USA

Abhishek Biswas

4. State Key Laboratory of Microbial Metabolism, Ministry of Education Laboratory of Systems Biomedicine, Shanghai Jiao Tong University, Shanghai, China

Liping Zhao

5. Department of Chemical and Biological Engineering, Princeton University, Princeton, NJ, USA

Mohamed S. Donia

6. Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA

Mohamed S. Donia

Contributions

M.S.D. and J.B. conceived the project. J.B. and M.E. performed biochemical experiments and analysed the resulting data, with input from M.S.D. and A.K.; P.D.J. collected and analysed X-ray crystallography data, with input from J.B., M.E., M.S.D. and A.K.; and J.B. performed microbiological experiments and analysed the resulting data. A.B., J.B. and M.S.D. performed metagenomic and metatranscriptomic data analyses. G.W. and L.Z. performed human clinical data analysis. J.B., P.D.J. and M.S.D. wrote the paper, with input from all of the other authors.

Corresponding author

Correspondence to [Mohamed S. Donia](#).

Ethics declarations

Competing interests

M.S.D. is a member of the scientific advisory board of Deepbiome Therapeutics, and L.Z. is a co-founder of Notitia Biotechnologies.

Additional information

Peer review information *Nature* thanks Sofia Forslund 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 Identification of AcbK homologues using MetaBGC.](#)

a. Number of true positive (TP) reads identified in synthetic metagenomic dataset number 2 for all spHMMs created by MetaBGC-Build. **b.** Number of false positive (FP) reads identified in synthetic metagenomic dataset number 2 for all spHMMs created by MetaBGC-Build. **c.** F1 scores of all spHMMs created by MetaBGC-Build. Only spHMMs with F1 scores ≥ 0.5 were used in MetaBGC-Identify runs with real metagenomic data. **d.** spHMM logos for spHMMs included in MetaBGC-Identify. Please see Supplementary Table [2](#) for spHMM cutoffs used for all intervals.

[Extended Data Fig. 2 Phylogenetic analysis of Mcks/Maks and prevalence of selected *maks* and *pmaks* in metagenomic samples of various human cohorts.](#)

a. A maximum likelihood phylogenetic tree for all AcbK homologues discovered in this study (Mcks). The tree was constructed using MEGA7, with bootstrap values of more than 50% out of 1000 replicates displayed at the branch points (see [Methods](#)). The tree includes previously characterized acarbose kinases from soil-derived Actinobacteria (red label, AcbK, GacK, and ScatK), the canonical phosphofructokinase from *E. coli* (PfkB), Mcks that were experimentally verified to have (green label, designated as Maks) or lack (blue label) an acarbose-*O*6A-kinase activity, and ones that have not been experimentally tested (black label). pMaks indicate putative Maks in the AcbK clade. **b.** Prevalence of selected *maks* and *pmaks* expressed as the percentage of positive individuals out of the total, across different cohorts and sampling sites and subsites (brown: gut; teal: oral). The total number of individuals in each cohort is indicated above their corresponding bar chart. Prevalence of all *mcks* in all cohorts can be found in Supplementary Table [2](#).

Extended Data Fig. 3 Purification and biochemical characterization of a subset of Mcks/Maks.

a. Fast Protein Liquid Chromatography (FPLC) chromatograms of AcbK (**a**) and Mck1/Mak1 (**b**) purification by size exclusion chromatography, monitored at the indicated absorbance wavelengths. Coomassie stained SDS-PAGE of the collected fraction is shown on the right side of each chromatogram. **c.** Coomassie stained SDS-PAGE of the collected fractions for the eight additional Mcks that were tested. All Mcks were purified in the same manner as AcbK and Mck1/Mak1 (see [Methods](#)). The data shown in **a, b** are representative examples from among at least three different times the same experiment was performed and produced similar results. The data shown in **c** is a representative example from among at least two different times the same experiment was performed and produced similar results. **d.** Representative Extracted Ion Chromatograms (EICs) for acarbose ($m/z = 646.4, [M+H]^+$) from the products of a reaction of Mck1/Mak1 (red) or a no-protein control (blue) with acarbose. **e.** Representative EICs for acarbose-*O*6A-phosphate ($m/z = 726.4, [M+H]^+$) from the products of a reaction of Mck1/Mak1 (red) or a no-protein control (blue) with acarbose (the two peaks have identical MS, HRMS, and HRMS/MS and are likely isomers of the same molecule or charge variants that elute differently, as

seen with AcbK, see Extended Data Fig. 5). This product is often referred to as acarbose-7-phosphate in the literature, but for consistency, we will refer to all acarbose positions based on their displayed numbering in Fig. 1a, which matches previous PDB depositions (PDB Ligand Code: ACR). Note the complete conversion of acarbose to acarbose-*O*6A-phosphate. See [Methods](#) for the complete experimental details.

Extended Data Fig. 4 Abundances of all metagenomic read bins discovered by MetaBGC.

Heatmap showing the abundance of all identified bins in the five analysed cohorts as calculated by MetaBGC, and following the colour codes to the right. See Supplementary Table 2 for the complete results of this analysis and Supplementary Table 1 for corresponding Mck/Mak identifiers for each bin. Samples that had no reads mapped to any bin were excluded from the heatmap. Hierarchical clustering of the samples and bins was performed using UPGMA (unweighted pair group method with arithmetic mean) in pheatmap in R.

Extended Data Fig. 5 HPLC-HR-MS and HPLC-HR-MS/MS analyses of acarbose-*O*6A-phosphate produced by AcbK and Mak1.

Extracted Ion Chromatograms (EICs) of acarbose-*O*6A-phosphate produced by AcbK (a), Mak1 (d), and a co-injection of the two (f), indicating that their products are identical. EIC is displayed for $m/z = 726.2\text{-}726.3 [M+H]^+$ in all cases. b, c, e, g, HR-MS of the two acarbose-*O*6A-phosphate peaks produced by either AcbK (b, c) or Mak1 (e, g), following the numbering scheme on the individual EIC peaks. h, i, identical HR-MS/MS fragmentation patterns of the two EIC peaks of acarbose-*O*6A-phosphate from the co-injection analysis in f. The structure of acarbose-*O*6A-phosphate is shown on the right, along with its predicted fragments and their calculated mass to charge ratios that match observed ones.

Extended Data Fig. 6 Additional biochemical characterization of Mak1 and AcbK.

a. Phosphorylation rates (relative to the acarbose phosphorylation rate, which is set as 100%) at which Mak1 and AcbK phosphorylate a diverse panel of carbohydrates and aminoglycosides with structural similarities to acarbose. No phosphorylation was detected for all but one (validamycin) of the substrates under the same experimental conditions (see [Methods](#)). Experiments were done in duplicates with the average value used for rate comparisons; raw data is available in Supplementary Table [3](#). **b.** Michaelis-Menten saturation curves for AcbK (grey) and Mak1 (blue) performed at 1 μM enzyme concentration. K_m and k_{cat} values are indicated in their respective colours and individual k_{obs} measurement replicates are shown on the graph for both enzymes. Raw data is available in Supplementary Table [3](#). **c.** k_{obs} of both AcbK (grey) and Mak1 (blue) across different temperatures from 25-40 °C in 3 °C steps. The difference in k_{obs} between the two enzymes can be seen across different temperatures. **d, e.** Hill plot (logarithm of k_{obs} on the y axis and logarithm of protein concentration on the x axis) of Mak1 (**d**) and AcbK (**e**). The Hill coefficient (slope) is greater than 1 for Mak1 (**d**, $n=1.71 \pm 0.06$, $n=1$ is shown in grey dashed line for reference) but not for AcbK (**e**, $n=1.13 \pm 0.02$), suggesting that only Mak1 is a cooperative enzyme. **f.** Relative change in k_{obs} (y axis) for the single mutants D160A and D247A as well as the double mutant (D160A, H162A) of Mak1 as compared to wild type protein ($N=2$). **g.** Extracted Ion Chromatograms (EICs) for acarbose (left, $m/z = 646.4$, $[\text{M}+\text{H}]^+$) and acarbose-*O*6A-phosphate (right, $m/z = 726.4$, $[\text{M}+\text{H}]^+$), showing that the addition of EDTA (blue traces) abolishes the activity of Mak1, while the addition of excess MgCl_2 (red traces) restores it.

[Extended Data Fig. 7 Additional structural details of AcbK.](#)

a. AcbK forms homodimers with extensive interactions between the two monomers. These interactions include multiple β strands in the β -clasp domain, most notably the $\beta3$ of one monomer and the $\beta8$ of the other (magenta highlighted box, with β strands from each monomer shaded differently for clarity). A molecular surface view is shown on the right with one AcbK monomer coloured green and the other coloured blue, highlighting the extensive surface area ($1,298 \text{ \AA}^2$) involved in forming the

dimer via the β -clasp domains. **b.** Zoomed in view of the AcbK substrate binding pocket where an extensive network of hydrogen bonds (from residues Asp16, Asn99, Ser109, and Asp248) form with all the hydroxyl groups in the acarbose A ring and hold it in place. Distances are shown for each of the hydrogen bonds mentioned above. **c.** Zoomed in view of the AcbK active site with important residues shown (Asp162, His164, Asp248), all involved in priming the *O*6A hydroxyl of acarbose for nucleophilic attack and in facilitating the transfer of the phosphate from an ATP (AMP-PNP shown) to the *O*6A hydroxyl of acarbose. Distance shown is from the *O*6A hydroxyl to the γ -phosphate. **d.** Full-length amino acid sequence alignment of experimentally tested Mcks/Maks, as well as AcbK and PfkB. The bars on the top of the alignment denote the average pairwise percent identity at each residue. **e.** Selected segments of the amino acid sequence alignment of experimentally tested Mcks/Maks, as well as AcbK and PfkB. Blue colours highlight amino acid residues that are deemed important for hydrogen bonding with the A ring of acarbose, yellow colours highlight those involved in the transfer of the phosphate group to acarbose and green indicates a residue involved in both processes. Grey colours indicate other conserved residues in the alignment. The bars on the top of the alignment denote the average pairwise percent identity at each residue.

Extended Data Fig. 8 Additional functional and genetic analyses of *maks* and *end*.

a. Normalized bacterial growth at 48 h, measured as the optical density at 600 nm and presented as a percentage of the untreated control. Different bars indicate treatments with varying acarbose concentrations for each of the tested strains: *A. viscosus* expressing *mak1* (blue) and *A. viscosus* harbouring an empty-vector control (grey). *mak1* expression resulted in a statistically significant rescue of the carbohydrate-dependent growth inhibitory activity of acarbose at several concentrations. Error bars represent the standard error of the mean among four replicates and “*” denotes a *p*-value < 0.01 (using a two-sample t-test assuming unequal variances): 0.5 μ M, *p*= 3.33×10^{-5} ; 5 μ M, *p*= 2.17×10^{-5} ; 10 μ M, *p*= 2.37×10^{-8} ; 50 μ M, *p*= 2.96×10^{-5} ; and 100 μ M, *p*= 5.94×10^{-3} . See Supplementary Table 6 for the raw experimental data. **b, c.** HbA1c (**b**) and fasting blood glucose (**c**) levels in *mak*-negative (N=8, blue) and *mak*-positive (N=8, red) Type 2

Diabetes patients treated with acarbose. The bar graphs show mean \pm SD. Repeated measures analysis of variance (ANOVA) with Tukey's post hoc test was used for comparing different time points in each group. Compact letter displays were used to indicate the statistical differences ($P < 0.05$) between different time points. The difference between groups at each post-treatment time point was evaluated by an ANCOVA model controlling for the baseline measurements. * $P < 0.05$. We observed statistically significant reductions in both HbA1c and fasting blood glucose levels only in *mak*-negative patients. In addition, at day 84, *mak*-negative patients showed significantly larger reduction of HbA1c ($P_{\text{ANCOVA}} = 0.011$) and fasting blood glucose levels ($P_{\text{ANCOVA}} = 0.024$) from baseline than *mak*-positive patients. **d.** Genetic context of all *maks* and *pmaks*, as well as *acbK*, *gacK* and *scatK*, following the colour key to the right. **e.** A simplified biosynthetic scheme for acarbose by *gac*, with *gac* proteins involved at each step shown above the arrows. Homologous *end* proteins are shown below the arrow, accounting for most of the core biosynthetic steps needed for acarbose production.

Extended Data Fig. 9 Distribution of *mak1* and *end* variants in the human microbiome.

a. Percentage of samples across different cohorts that are positive for either *mak1* (blue) or *end* (green). *mak1* was considered “present” in a given sample if any reads mapped to it, and *end* was considered “present” in a given sample if it had a breadth coverage of $\geq 25\%$ of its length. SP stands for supragingival plaque. See Supplementary Table 8 for detailed results of this analysis. **b.** Genetic context of all five *end* variants discovered, following the colour code on the right. Note that most variants exist in close proximity to a transposase gene, suggesting a role in their mobility. **c.** Relative RPKM (y axis, defined as the RPKM of the gene of interest divided by the RPKM of the entire *end* cluster) of each *end* gene (x axis). Ten examples of human metagenomic samples (supragingival plaque, HMP) are shown in the top graph, where the depth of coverage is uniform across all *end* genes. Ten examples of human metagenomic samples (supragingival plaque, HMP) are shown in the bottom graph, where a clear spike in the depth of coverage can be observed around the *endM/mak1*

gene, indicating the presence of at least two different *end* genetic variants in these samples. Next to each graph is a representative example of metagenomic reads from supragingival plaque samples mapped to the *end* BGC. See Supplementary Table 10 for detailed results of this analysis. **d**. A heatmap showing the abundance (in RPKM) of all *end* genes in supragingival plaque samples of the HMP cohorts (HMP-1-1 and HMP-1-2). Different samples harbour different genetic contexts (or variants) of the *end* BGC: 14-gene, 8-gene, 4-gene, 2-gene, or stand-alone *makI/endM*. Samples are sorted according to their classification into one of these five genetic variants listed above. **e**. Metagenomic reads from five different supragingival plaque samples mapped to the *end* BGC (see [Methods](#)). Each of the five samples illustrates an example for one of the unique genetic variants described above. **f**. A Pie chart showing the distribution of the five genetic variants amongst HMP participants (supragingival plaque samples). If participants had multiple visits and different variants across visits, they were classified into *end* (multiple variants) (see Supplementary Table 10).

Supplementary information

[Supplementary Figure 1](#)

Raw unedited images for gels shown in Extended Data Figure 3. **a**, Raw unedited gel image that corresponds to the cropped gel image shown in Extended Data Figure 3a. The ladder bands are labelled on the y-axis and the x-axis labels indicate sample names. Samples labelled as NA belong to an experiment unrelated to this manuscript. **b**, Raw unedited gel image that corresponds to the cropped gel image shown in Extended Data Figure 3b. The ladder bands are labelled on the y-axis and the x-axis labels indicate sample names. **c**, Raw unedited gel image that corresponds to the cropped gel image shown in Extended Data Figure 3c. The ladder bands are labelled on the y-axis and the x-axis labels indicate sample.

[Reporting Summary](#)

[Supplementary Table 1](#)

Additional information about AcbK homologues identified from all computational searches.

Supplementary Table 2

Details and results of MetaBGC analyses.

Supplementary Table 3

Raw enzymatic assay data for Maks/Mcks characterized in this study.

Supplementary Table 4

Raw data for the amylase inhibition assay with acarbose and acarbose-*O*6A-phosphate.

Supplementary Table 5

AcbK and Mak1 crystal structure refinement statistics.

Supplementary Table 6

Raw data for the *A. viscosus* growth rescue experiment.

Supplementary Table 7

Quantification of *mak* and *pmak* genes in faecal metagenomic datasets from a recent acarbose human clinical trial¹⁶.

Supplementary Table 8

Quantification of *mak1* and *end* across all samples and cohorts.

Supplementary Table 9

Comparison between the *end* and *gac* BGCs.

Supplementary Table 10

Analysis of *end* genetic variants in supragingival plaque samples of HMP participants.

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Balaich, J., Estrella, M., Wu, G. *et al.* The human microbiome encodes resistance to the antidiabetic drug acarbose. *Nature* **600**, 110–115 (2021). <https://doi.org/10.1038/s41586-021-04091-0>

- Received: 29 July 2020
- Accepted: 01 October 2021
- Published: 24 November 2021
- Issue Date: 02 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04091-0>

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/s41586-021-04091-0>

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

- Article
- [Published: 01 December 2021](#)

Antiviral activity of bacterial TIR domains via immune signalling molecules

- [Gal Ofir](#) ORCID: [orcid.org/0000-0002-8307-9388¹](https://orcid.org/0000-0002-8307-9388),
- [Ehud Herbst¹](#),
- [Maya Baroz¹](#),
- [Daniel Cohen¹](#),
- [Adi Millman](#) ORCID: [orcid.org/0000-0002-6844-7613¹](https://orcid.org/0000-0002-6844-7613),
- [Shany Doron](#) ORCID: [orcid.org/0000-0002-8392-6110¹](https://orcid.org/0000-0002-8392-6110),
- [Nitzan Tal¹](#),
- [Daniel B. A. Malheiro²](#),
- [Sergey Malitsky](#) ORCID: [orcid.org/0000-0003-4619-7219³](https://orcid.org/0000-0003-4619-7219),
- [Gil Amitai](#) ORCID: [orcid.org/0000-0002-2954-2503¹](https://orcid.org/0000-0002-2954-2503) &
- [Rotem Sorek](#) ORCID: [orcid.org/0000-0002-3872-4982¹](https://orcid.org/0000-0002-3872-4982)

[Nature](#) volume 600, pages 116–120 (2021)

- 3726 Accesses
- 184 Altmetric
- [Metrics details](#)

Subjects

- [Bacterial genetics](#)
- [Bacteriophages](#)

Abstract

The Toll/interleukin-1 receptor (TIR) domain is a canonical component of animal and plant immune systems^{1,2}. In plants, intracellular pathogen sensing by immune receptors triggers their TIR domains to generate a molecule that is a variant of cyclic ADP-ribose^{3,4}. This molecule is hypothesized to mediate plant cell death through a pathway that has yet to be resolved⁵. TIR domains have also been shown to be involved in a bacterial anti-phage defence system called Thoeris⁶, but the mechanism of Thoeris defence remained unknown. Here we show that phage infection triggers Thoeris TIR-domain proteins to produce an isomer of cyclic ADP-ribose. This molecular signal activates a second protein, ThsA, which then depletes the cell of the essential molecule nicotinamide adenine dinucleotide (NAD) and leads to abortive infection and cell death. We also show that, similar to eukaryotic innate immune systems, bacterial TIR-domain proteins determine the immunological specificity to the invading pathogen. Our results describe an antiviral signalling pathway in bacteria, and suggest that the generation of intracellular signalling molecules is an ancient immunological function of TIR domains that is conserved in both plant and bacterial immunity.

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: Thoeris causes abortive infection and depletion of NAD⁺.

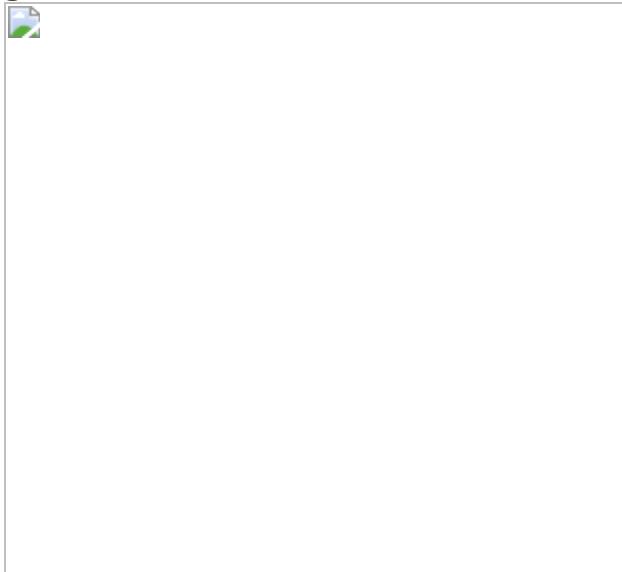


Fig. 2: The NADase activity of ThsA is triggered by small molecules generated in TIR-expressing infected cells.



Fig. 3: ThsB produces an isomer of cADPR after phage infection.



Fig. 4: Thoeris TIR proteins determine defence specificity.



Data availability

Data that support the findings of this study are available within the Article and its Extended Data. Gene accessions appear in the Methods section of the paper. Plasmid maps of the constructs used for the experiments are attached as Supplementary Files. Additional data are available from the corresponding authors upon reasonable request. [Source data](#) are provided with this paper.

References

1. 1.
Fitzgerald, K. A. & Kagan, J. C. Toll-like receptors and the control of immunity. *Cell* **180**, 1044–1066 (2020).
2. 2.
Burch-Smith, T. M. & Dinesh-Kumar, S. P. The functions of plant TIR domains. *Sci. STKE* **2007**, pe46 (2007).
3. 3.
Wan, L. et al. TIR domains of plant immune receptors are NAD⁺-cleaving enzymes that promote cell death. *Science* **365**, 799–803 (2019).
4. 4.
Horsefield, S. et al. NAD⁺ cleavage activity by animal and plant TIR domains in cell death pathways. *Science* **365**, 793–799 (2019).
5. 5.
Bayless, A. M. & Nishimura, M. T. Enzymatic functions for Toll/interleukin-1 receptor domain proteins in the plant immune system. *Front. Genet.* **11**, 539 (2020).
6. 6.

Doron, S. et al. Systematic discovery of antiphage defense systems in the microbial pangenome. *Science* **359**, eaar4120 (2018).

7. 7.

Balint-Kurti, P. The plant hypersensitive response: concepts, control and consequences. *Mol. Plant Pathol.* **20**, 1163–1178 (2019).

8. 8.

Duxbury, Z. et al. Induced proximity of a TIR signaling domain on a plant–mammalian NLR chimera activates defense in plants. *Proc. Natl Acad. Sci. USA* **117**, 18832–18839 (2020).

9. 9.

Ka, D., Oh, H., Park, E., Kim, J.-H. & Bae, E. Structural and functional evidence of bacterial antiphage protection by Thoeris defense system via NAD⁺ degradation. *Nat. Commun.* **11**, 2816 (2020).

10. 10.

Lopatina, A., Tal, N. & Sorek, R. Abortive infection: bacterial suicide as an antiviral immune strategy. *Annu. Rev. Virol.* **7**, 371–384 (2020).

11. 11.

Tzipilevich, E., Pollak-Fiyaksel, O. & Ben-Yehuda, S. Bacteria elicit a phage tolerance response subsequent to infection of their neighbors. Preprint at <https://doi.org/10.1101/2021.02.16.428622> (2021).

12. 12.

Morehouse, B. R. et al. STING cyclic dinucleotide sensing originated in bacteria. *Nature* **586**, 429–433 (2020).

13. 13.

Burroughs, A. M. & Aravind, L. Identification of uncharacterized components of prokaryotic immune systems and their diverse eukaryotic reformulations. *J. Bacteriol.* **202**, <https://doi.org/10.1128/JB.00365-20> (2020).

14. 14.

Burroughs, A. M., Zhang, D., Schäffer, D. E., Iyer, L. M. & Aravind, L. Comparative genomic analyses reveal a vast, novel network of nucleotide-centric systems in biological conflicts, immunity and signaling. *Nucleic Acids Res.* **43**, 10633–10654 (2015).

15. 15.

Huang, Y., Fliegert, R., Guse, A. H., Lü, W. & Du, J. A structural overview of the ion channels of the TRPM family. *Cell Calcium* **85**, 102111 (2020).

16. 16.

Huang, Y., Roth, B., Lü, W. & Du, J. Ligand recognition and gating mechanism through three ligand-binding sites of human TRPM2 channel. *eLife* **8**, e50175 (2019).

17. 17.

Cohen, D. et al. Cyclic GMP–AMP signalling protects bacteria against viral infection. *Nature* **574**, 691–695 (2019).

18. 18.

Ye, Q. et al. HORMA domain proteins and a Trip13-like ATPase regulate bacterial cGAS-like enzymes to mediate bacteriophage immunity. *Mol. Cell* **77**, 709–722 (2020).

19. 19.

Essuman, K. et al. The SARM1 Toll/interleukin-1 receptor domain possesses intrinsic NAD⁺ cleavage activity that promotes pathological

axonal degeneration. *Neuron* **93**, 1334–1343 (2017).

20. 20.

Essuman, K. et al. TIR domain proteins are an ancient family of NAD⁺-consuming enzymes. *Curr. Biol.* **28**, 421–430 (2018).

21. 21.

Coronas-Serna, J. M. et al. The TIR-domain containing effectors BtpA and BtpB from *Brucella abortus* impact NAD metabolism. *PLoS Pathog.* **16**, e1007979 (2020).

22. 22.

Watanabe, S., Shiwa, Y., Itaya, M. & Yoshikawa, H. Complete sequence of the first chimera genome constructed by cloning the whole genome of *Synechocystis* strain PCC6803 into the *Bacillus subtilis* 168 genome. *J. Bacteriol.* **194**, 7007 (2012).

23. 23.

Wilson, G. A. & Bott, K. F. Nutritional factors influencing the development of competence in the *Bacillus subtilis* transformation system. *J. Bacteriol.* **95**, 1439–1449 (1968).

24. 24.

Mazzocco, A., Waddell, T. E., Lingohr, E. & Johnson, R. P. Enumeration of bacteriophages using the small drop plaque assay system. *Methods Mol. Biol.* **501**, 81–85 (2009).

25. 25.

Zheng, L. et al. Fumarate induces redox-dependent senescence by modifying glutathione metabolism. *Nat. Commun.* **6**, 6001 (2015).

26. 26.

Pluskal, T., Castillo, S., Villar-Briones, A. & Orešič, M. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* **11**, 395 (2010).

27. 27.

Myers, O. D., Sumner, S. J., Li, S., Barnes, S. & Du, X. One step forward for reducing false positive and false negative compound identifications from mass spectrometry metabolomics data: new algorithms for constructing extracted ion chromatograms and detecting chromatographic peaks. *Anal. Chem.* **89**, 8696–8703 (2017).

28. 28.

Bernheim, A. et al. Prokaryotic viperins produce diverse antiviral molecules. *Nature* **589**, 120–124 (2020).

29. 29.

Berman, H. M. et al. The Protein Data Bank. *Nucleic Acids Res.* **28**, 235–242 (2000).

30. 30.

Sonn-Segev, A. et al. Quantifying the heterogeneity of macromolecular machines by mass photometry. *Nat. Commun.* **11**, 1772 (2020).

Acknowledgements

We thank the Sorek laboratory members, M. Voichek, A. Levy, D. Dar, V. Šikšnys and M. Zaremba for comments on earlier versions of this manuscript; Y. M. Bar-On for useful discussion during the project; C. Avraham and T. Fedorenko for their assistance with the experiments; A. Bernheim for her continuous advice and support throughout this project; A. Sonn-Segev for her assistance with mass photometry experiments; and Y. Peleg and S. Albeck for their help in purification of the ThsA protein. R.S. was supported, in part, by the European Research Council (grant ERC-CoG

681203), the Ernest and Bonnie Beutler Research Program of Excellence in Genomic Medicine, the Minerva Foundation with funding from the Federal German Ministry for Education and Research, the Knell Family Center for Microbiology, and the Yotam project and the Weizmann Institute Sustainability And Energy Research (SAERI) initiative. G.O. was supported by the SAERI doctoral fellowship. A.M. was supported by a fellowship from the Ariane de Rothschild Women Doctoral Program and, in part, by the Israeli Council for Higher Education via the Weizmann Data Science Research Center, and by a research grant from O. Klein-Astrachan.

Author information

Affiliations

1. Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel

Gal Ofir, Ehud Herbst, Maya Baroz, Daniel Cohen, Adi Millman, Shany Doron, Nitzan Tal, Gil Amitai & Rotem Sorek

2. MS-Omics, Vedbæk, Denmark

Daniel B. A. Malheiro

3. Life Science Core Facilities, Weizmann Institute of Science, Rehovot, Israel

Sergey Malitsky

Contributions

G.O. designed all the experiments, performed all of the experiments unless otherwise noted, analysed the data and wrote the manuscript. E.H. performed the experimental design and data analysis for the mass spectrometry experiments presented in Figs. [1d](#), [3a](#), [b](#), Extended Data Figs. [1h](#), [i](#), [3a](#). M.B. and D.C. performed the cloning and plaque assay experiments presented in Fig. [4b](#), [c](#), Extended Data Fig. [4](#). A.M. and S.D.

performed bioinformatic analysis leading to the identification of the Thoeris systems studied. N.T. participated in writing the manuscript. D.B.A.M. performed the mass spectrometry experiments and data analysis presented in Fig. [3a, b](#). S.M. performed the mass spectrometry experiments presented in Fig. [1d](#), Extended Data Fig. [1h, i](#). G.A. performed, together with G.O., the enzymatic assay experiments presented in Figs. [2b, c, 4d](#); the structural analysis presented in Fig. [3c](#); and the mass photometry experiments presented in Fig. [3d, e](#), Extended Data Fig. [3c–f](#). R.S. supervised the study and wrote the manuscript.

Corresponding authors

Correspondence to [Gil Amitai](#) or [Rotem Sorek](#).

Ethics declarations

Competing interests

R.S. is a scientific cofounder and advisor of BiomX, Pantheon Bioscience and Ecophage. The other 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 Effects of mutations in Thoeris genes on defence against phage SPO1.](#)

a, Plaques of phage SPO1 on control cells (black), cells expressing both WT Thoeris proteins (green), cells expressing mutant ThsA(N112A) and WT ThsB (magenta), or cells expressing WT ThsA and mutant ThsB(E85Q) (cyan). Ten-fold serial dilution of the phage lysate were dropped on the plates. **b**, Efficiency of plating (EOP) of phage SPO1 on control and Thoeris-containing strains, representing plaque-forming units per millilitre (PFU/ml). Asterisk marks statistically significant reduction in EOP (One-way ANOVA, followed by pairwise multiple comparison analysis according to Tukey's honest significant difference criterion $p=2 \times 10^{-8}$). **c**, Replication of phage SPO1 in the presence of Thoeris-containing and Thoeris-lacking (control) cells, or without cells (no bacteria). Lysates were collected 2.5 h following infection of liquid cultures at an initial MOI of 5, and phage titer was quantified by plating serial dilution of the lysates on the control strain. Asterisk marks statistically significant reduction in EOP compared to the control strain (One-way ANOVA, followed by pairwise multiple comparison analysis according to Tukey's honest significant difference criterion $p=0.016$). **d**, Remaining phage titre after culture recovery. Samples were collected 12 h following infection of liquid cultures at an initial MOI of 5, and phage titre was quantified by plating serial dilution of the samples on the control strain. Asterisk marks statistically significant reduction in EOP compared to the control strain (One-way ANOVA, followed by pairwise multiple comparison analysis according to Tukey's honest significant difference criterion $p = 0.0004$). **e**, Growth curves of infections with phage SPO1 at MOI of 5 of control cells (black), Thoeris-expressing cells (dark green), and 11 colonies isolated from recovered Thoeris-expressing cells 12 h post infection (light green). 3–4 Colonies were isolated from each of three independent infections. **f**, **g**, Growth curves of uninfected Thoeris mutants (**f**) and during infection by phage SPO1 at MOI of 5 or 0.05 (**g**). Three independent experiments are presented as individual curves. **h**, Adenosine diphosphate ribose (ADPR) levels in control culture (black) and cells expressing wild type Thoeris (green). Time 0 represents uninfected cells. Cells were infected by phage SPO1 at an MOI of 5. Each line represents the mean of three independent experiments, with individual data points shown. ADPR levels were measured by LC–MS and calculated from the area under the curve of the identified ADPR peak. **i**, NAD⁺ levels in cell expressing WT ThsB + ThsA(N112A) (magenta) and cells expressing WT ThsA +

ThsB(E85Q) (cyan). Time 0 represents uninfected cells. Each line represents the mean of three independent experiments, with individual data points shown. Cells were infected by phage SPO1 at an MOI of 5. NAD⁺ levels were measured by LC–MS and calculated from the area under the curve of the identified NAD⁺ peak.

[Source data](#)

Extended Data Fig. 2 NADase activity of ThsA after addition of standard cADPR.

NADase activity of ThsA, when incubated with filtered cell lysates derived from ThsB-expressing cells 70-min post infection by phage SPO1, or with buffer containing 0 μM – 1000 μM of synthetic cyclic ADP-ribose (cADPR). NADase activity was calculated as the rate of change in εNAD fluorescence during the linear phase of the reaction. Bars represent mean of three experiments, with individual data points overlaid. NADase activity between the cADPR-containing samples and the blank sample (0 μM cADPR) are not statistically significant (One-way ANOVA, followed by pairwise multiple comparison analysis according to Tukey’s honest significant difference).

[Source data](#)

Extended Data Fig. 3 MS/MS spectra of the cADPR isomer produced by ThsB following phage infection and the activity of the SLOG domain.

a, MS/MS fragmentation spectra of standard cADPR (top) and the Thoeris-derived cADPR isomer (bottom). Hypothesized structures of MS/MS fragments of cADPR are presented. **b**, EOP of phages on bacteria expressing WT ThsB and ThsA(R371A) compared to WT Thoeris and other mutants. Results for the control, WT Thoeris and ThsB(E85Q) and ThsA(N112A) are those presented in Extended Data Fig. 1. Asterisk marks statistically significant reduction in EOP (One-way ANOVA, followed by pairwise multiple comparison analysis according to Tukey’s honest

significant difference criterion $p = 2 \times 10^{-8}$). EOP of ThsA(R371A) is not statistically significant compared to the control strain ($p = 0.7$). **e–f**, Two additional replicates of the mass photometry measurement presented in Figs [3d, e](#). ThsA purified protein was incubated with lysates derived from infected cells expressing ThsB(E85Q) (**c, e**) or ThsB (**d, f**). Dashed lines represent masses of ThsA monomer, dimer and tetramer.

[Source data](#)

Extended Data Fig. 4 Efficiency of plating on cells expressing Thoeris systems of *B. cereus*, *B. dafuensis* and hybrid systems.

a, EOP of phages on control cells (black) and cells expressing combinations of ThsA from either *B. cereus* (green) or *B. dafuensis* (blue), TIR protein from *B. cereus* (ThsB), TIR proteins from *B. dafuensis* (TIR1 and TIR2) or both TIRs from *B. dafuensis* (TIR1+TIR2). Bars represent mean of 3 replicates, with individual data points overlaid. Asterisk marks statistically significant reduction in EOP (One-way ANOVA, followed by pairwise multiple comparison analysis according to Tukey's honest significant difference criterion $p < 0.0015$). **b**, EOP of phages on control cells (black) and cells expressing mutant ThsA(N112A) from *B. cereus* together with combinations of TIR proteins from *B. cereus* and *B. dafuensis*.

[Source data](#)

Supplementary information

Supplementary Information

This file contains Supplementary Tables 1 and 2.

Reporting Summary

Supplementary File 1

Plasmid map of pGO1_thrC_Pxyl_cereus_ThsA, shuttle vector for *B. cereus* MSX-D12 *thsA* gene.

Supplementary File 2

Plasmid map of pGO2_thrC_Pxyl_dafuensis_ThsA, shuttle vector for *B. dafuensis* FJAT-25496 *thsA*

Supplementary File 3

Plasmid map of pGO3_amyE_hspank_cereus_ThsB, shuttle vector for *B. cereus* MSX-D12 *thsB* gene.

Supplementary File 4

Plasmid map of pGO4_amyE_hspank_dafuensis_TIR1, shuttle vector for *B. dafuensis* FJAT-25496 *tir1* gene.

Supplementary File 5

Plasmid map of pGO5_amyE_hspank_dafuensis_TIR2, shuttle vector for *B. dafuensis* FJAT-25496 *tir2* gene.

Supplementary File 6

Plasmid map of pGO6_amyE_hspank_GFP, shuttle vector for GFP control construct.

Supplementary File 7

Plasmid map of pGO7_amyE_hspank_dafuensis_TIR1+TIR2, shuttle vector for *B. dafuensis* FJAT-25496 *tir1+tir2* genes.

Supplementary File 8

Plasmid map of pGO8_pET30a_strep_BdTIR_His, protein expression plasmid for *Brachypodium distachyon* BdTIR gene.

Supplementary File 9

Plasmid map of pGO9_pAB151_cereus_ThsA_strep, protein expression plasmid for *B. cereus* ThsA purification.

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

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Ofir, G., Herbst, E., Baroz, M. *et al.* Antiviral activity of bacterial TIR domains via immune signalling molecules. *Nature* **600**, 116–120 (2021). <https://doi.org/10.1038/s41586-021-04098-7>

- Received: 05 January 2021
- Accepted: 06 October 2021
- Published: 01 December 2021
- Issue Date: 02 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04098-7>

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/s41586-021-04098-7>

- Article
- [Published: 17 November 2021](#)

Mental health concerns during the COVID-19 pandemic as revealed by helpline calls

- [Marius Brülhart](#) ORCID: orcid.org/0000-0001-5483-0219^{1,2},
- [Valentin Klotzbücher](#) ORCID: orcid.org/0000-0001-9382-6757³,
- [Rafael Lalive](#)^{1,2} &
- [Stephanie K. Reich](#)³

Nature volume 600, pages 121–126 (2021)

- 12k Accesses
- 1 Citations
- 340 Altmetric
- [Metrics details](#)

Subjects

- [Human behaviour](#)
- [Society](#)

Abstract

Mental health is an important component of public health, especially in times of crisis. However, monitoring public mental health is difficult because data are often patchy and low-frequency^{1,2,3}. Here we complement established approaches by using data from helplines, which offer a real-time measure of ‘revealed’ distress and mental health concerns across a range of topics^{4,5,6,7,8,9}. We collected data on 8 million calls from 19 countries, focusing on the COVID-19 crisis. Call volumes peaked six weeks after the initial outbreak, at 35% above pre-pandemic levels. The increase was driven

mainly by fear (including fear of infection), loneliness and, later in the pandemic, concerns about physical health. Relationship issues, economic problems, violence and suicidal ideation, however, were less prevalent than before the pandemic. This pattern was apparent both during the first wave and during subsequent COVID-19 waves. Issues linked directly to the pandemic therefore seem to have replaced rather than exacerbated underlying anxieties. Conditional on infection rates, suicide-related calls increased when containment policies became more stringent and decreased when income support was extended. This implies that financial relief can allay the distress triggered by lockdown measures and illustrates the insights that can be gleaned from the statistical analysis of helpline data.

[Download PDF](#)

Main

The state of population mental health is difficult to measure. This could lead policymakers to neglect mental health issues relative to aspects that can be measured more easily—especially during fast-moving crisis situations^{1,2,3}. We propose using helpline data as a source of real-time information on the state of public mental health. Helpline data have two main advantages. First, helpline calls can be considered as a manifestation of revealed distress and mental health concerns. Callers incur the mental and time cost of reaching out without having been prompted to do so. Therefore, helpline calls resemble clinical data by offering a measure of mental health that is unaffected by researchers' study design and framing. Second, information about helpline calls is recorded digitally with daily frequency and covers a wide range of conversation topics.

Telephone helplines are well-established institutions for mental health protection and suicide prevention in many countries, and they offer support immediately, anonymously, cheaply and accessibly^{10,11,12}. Some helplines specialize in particular issues such as suicide, children or violence against women. Suicide helplines, for example, have been shown to reduce suicide rates¹³, and call volumes of suicide prevention helplines have been shown to relate to the incidence of actual suicides¹⁴.

Using this approach in relation to the COVID-19 crisis, we documented the growth and composition of helpline calls as well as their pandemic-related determinants. Helplines take on particular relevance in a pandemic, when face-to-face contacts carry infection risks and may even be impossible owing to stay-at-home orders^{4,5,6,7,8,9}. We collected data from 23 helplines in 14 European countries, the USA, China, Hong Kong, Israel and Lebanon. The total dataset covers 8 million individual calls made between 2019 and early 2021 (Extended Data Table 1). The panel structure of the data enables us to exploit differences in the timing of local infection waves and policy

measures to isolate their separate effects on helpline calls. This is a first-order issue for policymakers, as interventions designed to contain infections might also affect mental health by exacerbating unemployment, financial stress, loneliness, relationship problems and pre-existing mental vulnerabilities. These are, in turn, well-recognized risk factors for suicide^{[15,16,17,18,19,20,21,22](#)}.

We consider the analysis of helpline calls as a complement, and not a substitute, for established approaches. Mental health surveys^{[23,24,25,26](#)} and suicide statistics^{[27,28,29](#)} are highly informative, but they tend to be low frequency and available with a lag. Higher-frequency monitoring has been performed in the context of the COVID-19 pandemic on the basis of online searches as recorded by Google Trends^{[30,31,32,33](#)}, by tracking visits to emergency departments^{[34,35](#)}, and by monitoring calls to the police for help with domestic disputes^{[36,37,38](#)}. Helpline data contribute a measurement tool that is both broadly available and well targeted on the mental health concerns of a particularly vulnerable segment of the population.

Results

Increased call volumes across helplines

When we pool and size-weight the data for the 21 helplines for which we have daily data (Extended Data Table [1a,b](#)), we observe a peak call volume, reached six weeks after the outbreak of the pandemic, that exceeds the pre-pandemic level by 35% (95% confidence interval (CI): 22.6, 48.3; $P < 0.001$) (Fig. [1a](#)). With the country-specific outbreak defined as the date when more than 1 SARS-CoV-2 infection per 100,000 inhabitants was recorded^{[39](#)}, we see a significant increase of 13.5% (95% CI: 1.6, 25.5; $P = 0.027$) for the first time in week 3. After the peak in week 6, volumes gradually decreased again, to 6.2% (95% CI: -0.2, 12.6; $P = 0.058$) above pre-pandemic levels^{[39](#)} by around week 11. When we instead define the starting point of the pandemic as the entry into force of the first shelter-in-place (SIP) order^{[40](#)}, we observe an increase of 11.2% (95% CI: 3.1, 19.4; $P = 0.007$) (Fig. [1b](#)) by week 2, steadily elevated call volumes from week 3 (+27%; 95% CI: 19.1, 35.0; $P < 0.001$) until about week 8 (+22.6%; 95% CI: 15.2, 30.1; $P < 0.001$), and a decrease thereafter^{[40](#)}. The different time profiles are mainly explained by the fact that on average, SIP orders were issued 2 to 3 weeks after local outbreaks (Extended Data Fig. [1](#), Extended Data Table [1](#)).

Fig. 1: Evolution of total helpline call numbers during the first wave.

 **figure1**

a, b, Estimated coefficients for week indicators with 95% confidence intervals. The dependent variable is $\ln(\text{Calls})$. The sample includes daily data for 21 helplines during the period from 4 weeks before to 12 weeks after the event date in early 2020, and for 17 of the 21 helplines, the corresponding days in 2019. Average percentage change in call volumes relative to reference week 0. **a**, Week 0 is when the cumulative number of SARS-CoV-2 infections exceeded 1 per 100,000 population³⁹. **b**, Week 0 is when SIP orders were introduced⁴⁰. Results show data weighted by total number of calls recorded for each helpline during the sample period (black) and unweighted models (grey) ([Methods](#), equation (1)).

Source data.

The gradual nature of the increase in call volumes could, to some extent, be a result of capacity constraints⁴. Several helplines initially had to leave some of the additional calls unanswered and only gradually managed to adjust capacity to the new level of demand. Because of this issue, the evolution of recorded aggregate call numbers should be interpreted as a lower-bound estimate of the true increase in the number of people who sought to call a helpline in the first wave of the pandemic. However, unanswered calls are not pre-screened, and call answering is thus a random process unrelated to the motives of the caller. Thus our data provide representative information on the reasons for calling even if some calls were left unanswered because of capacity constraints.

Caller issues and conversation topics

We analysed the reasons for calling using data on the 12 helplines for which we have call-level information on conversation topics and caller characteristics. Our main results relating to call topics are presented in Fig. 2. Most pre-COVID-19 calls were made because of relationship issues (37%), loneliness (20%) or various fears and anxieties (13%) (Fig. 2a). Women placed 61% of total calls, and 63% of calls were placed by people between 30 and 60 years of age. The breakdown by topic was fairly similar across helplines, with relationship issues being the most prevalent topic in 8 of the 10 helplines for which this category is defined (34% overall) (Extended Data Fig. 2). More than 90% of ‘calls’ were voice calls, but for some helplines our data also includes text-based (online chat) conversations. Between 49% and 81% of calls were placed by first-time or sporadic callers, both before and after the onset of the pandemic (Extended Data Table 2).

Fig. 2: Conversation topics during the first COVID-19 wave.

figure2

- a, Pre-pandemic shares of main non-exclusive helpline conversation topics by sex and age group, before cumulative SARS-CoV-2 infections³⁹ reached 1 per 100,000

population. **b**, Estimated coefficients for the binary post-outbreak indicator variable with 95% confidence intervals. Separate linear probability regression models with the dependent variable set to ‘1’ for calls related to the indicated topic ([Methods](#), equation (2)).

[Source data.](#)

During the first wave of the pandemic, defined here as lasting until the end of June 2020, the composition of calls changed significantly. The biggest increase in calls was recorded in the category ‘fear’, with 2.4 percentage points (95% CI: 1.8, 2.9; $P < 0.001$) (Fig. 2b). This category includes calls made out of fear of infection with SARS-CoV-2. The other category of calls whose share increased during the first wave of the pandemic was ‘loneliness’ with 1.5 percentage points (95% CI: 1.1, 1.8; $P < 0.001$) (Fig. 2b). The share of all other conversation topics decreased during the first wave. Statistically significant relative decreases were observed for the topics ‘relationships’ (−2.5 percentage points; 95% CI: −3.2, −1.8; $P < 0.001$), ‘livelihood’ (that is, economic worries, −0.6 percentage points; 95% CI: −0.9, −0.3; $P < 0.001$), ‘violence’ (−0.3 percentage points; 95% CI: −0.5, −0.2; $P < 0.001$) and ‘addiction’ (−0.3 percentage points; 95% CI: −0.4, −0.1; $P = 0.002$) (Fig. 2b). We detected no statistically significant change in the share of calls related to suicidal ideation (−0.1 percentage points; 95% CI: −0.3, −0.1; $P = 0.476$ (two-sided t -test of difference) and $P < 0.006$ (two one-sided t -tests), against effect size <-0.35 and >0.35 , respectively)) (Extended Data Fig. 3). These results show that the first wave of the pandemic and the associated measures led to a less than proportional increase in calls about domestic violence, addiction and suicidal ideation relative to the overall increase in calls.

When we break down post-pandemic changes in topic shares by gender and age group, we observe that the increase in fear-related calls was driven entirely by the over-30s, both male and female (between 2.1 and 3.1 percentage points; 95% CI: 1.5, 2.7 to 2.2, 4.0; $P < 0.001$) (Extended Data Fig. 4). This is consistent with the fact that vulnerability to COVID-19 increases monotonically with age. The share of suicide-related calls placed by men under 30 fell particularly strongly (by 1.6 percentage points; 95% CI: −2.3, −0.9; $P < 0.001$) (Extended Data Fig. 4). Conversely, the category of women under 30 stands out, with a 0.9 percentage points increase in the share of calls related to violence (95% CI: 0.2, 1.6; $P = 0.010$) (Extended Data Fig. 4), despite the fact that it may well have been more difficult under stay-at-home orders to make helpline calls in situations of domestic violence.

For around one-third of the calls underlying our analysis of Fig. 2, operators recorded more than one conversation topic (Extended Data Fig. 5a). In particular, calls related to ‘violence’ and ‘livelihood’ also concerned ‘relationships’ (39% and 35%, respectively) (Extended Data Fig. 5b), but combinations of all eight topics

distinguished in our analyses were observed in the data. Dropping multiple-topic calls from the analysis left results almost unchanged (Extended Data Fig. 5c).

Overall, our results suggest that the observed increase in helpline calls during the first wave of the COVID-19 pandemic was driven to a large extent by fears of the virus itself and by loneliness in the context of SIP orders, rather than by domestic violence, addiction or suicidal ideation.

Call dynamics during subsequent waves

For two of the largest helplines in our sample, Telefonseelsorge (Germany) and SOS Amitié (France), we received data up to 31 March 2021, enabling us to analyse helpline calls beyond the first wave of the pandemic. Figure 3 shows that call volumes increased again in the second half of 2020, in parallel with an increase in infections and a tightening of non-pharmaceutical interventions (NPIs). Whereas in Germany the volume of calls increased continuously into early 2021 (Fig. 3a), in France it fell again after the peak in December 2020 (Fig. 3b). These diverging patterns correlate with stronger upswings and downswings in both infections and the stringency of government measures in the two countries.

Fig. 3: Helpline calls in Germany and France during the first and subsequent waves.

 figure3

a, c, Total number of daily helpline calls with seven-day moving average in black (right axis), government response stringency index in blue⁴⁰, and seven-day moving average of newly confirmed SARS-CoV-2 infections per 10,000 population and day in red (left axis)³⁹, for Germany (Telefonseelsorge) (**a**) and France (SOS Amitié) (**c**). Shaded areas indicate first wave (11 March 2020–30 June 2020) and subsequent waves (1 October 2020–31 March 2021). **b, d**, Estimated coefficients for binary variables denoting the two periods, and their associated 95% confidence intervals for

Germany (Telefonseelsorge) (**b**) and France (SOS Amitié) (**d**), based on separate linear probability regression models with the dependent variable set to ‘1’ for calls related to the indicated topic ([Methods](#), equation (4)).

[Source data.](#)

Conversation topic patterns resemble each other both between the two helplines and between the two distinct periods of the pandemic. In Germany, calls due to loneliness increased by 1.4 percentage points (95% CI: 0.9, 2.0; $P < 0.016$) in the first wave, and by 0.6 percentage points (95% CI: 0.1, 1.1; $P < 0.016$) in subsequent waves (Fig. [3b](#)). In France, those increases were 2.0 percentage points (95% CI: 1.4, 2.6; $P < 0.016$) and 0.8 percentage points (95% CI: 0.4, 1.2; $P < 0.001$), respectively (Fig. [3b](#)). During the first wave, the share of calls related to ‘fear’ (including the fear of infection) increased by 2.2 percentage points (95% CI: 1.4, 2.9; $P < 0.001$) in Germany, and by 2.7 percentage points (95% CI: 2.0, 3.5; $P < 0.001$) in France (Fig. [3b](#)). For France, we also observed a significant increase during subsequent waves, by 1.2 percentage points (95% CI: 0.8, 1.5; $P < 0.001$) (Fig. [3b](#)). The share of calls concerning relationship issues decreased in Germany by 3.5 percentage points (95% CI: -4.4, -2.5; $P < 0.001$) in the first wave and by 1.8 percentage points (95% CI: -2.3, -1.2; $P < 0.001$) in subsequent waves (Fig. [3b](#)). Decreases were also observed for France: -2.6 percentage points (95% CI: -3.9, -1.2; $P < 0.001$) in the first wave, and -1.1 percentage points (95% CI: -2.0, -0.3; $P < 0.001$) during the subsequent waves (Fig. [3b](#)). Conversations were less likely to relate to suicidality during subsequent waves (-0.6 percentage points in Germany and -0.9 percentage points in France; 95% CI: -0.9, -0.3 and -1.0, -0.7; $P < 0.001$) (Fig. [3b](#)).

Conversely, a larger proportion of calls in the second and third waves in France concerned physical health (+0.8 percentage points; 95% CI: 0.3, 1.2; $P = 0.001$) (Fig. [3b](#)). This could be related to a larger share of the population being infected with SARS-CoV-2 or to health worries because of restricted or postponed access to treatment facilities and fewer opportunities for physical activity. Similar to the first wave, additional calls focused predominantly on issues linked directly with the pandemic: fear of infection, loneliness, and—new to subsequent waves—physical health.

Infection rates and policy measures

Helpline call data enable us to use panel data regression to isolate partial correlations between policy measures and indicators of mental health. A particularly informative empirical laboratory for this analysis are calls to the National Suicide Prevention Lifeline (hereafter referred to as Lifeline) in the USA. We have data for 2019, 2020 and early 2021, which enables us to exploit the considerable intranational (state-level) variation of epidemiological situations and policy measures observed within the USA.

Thanks to coordination across the network of crisis centres that constitute the Lifeline through a common set of general guidelines, institutional and measurement issues that complicate comparisons across diverse sets of helplines and nations are less of a concern in this dataset. As a helpline focused on suicide, however, Lifeline does not enable us to track changes in the composition of mental health problems.

Our main findings are presented in Fig. 4. The aggregate time trend reveals that during the first wave, calls to Lifeline were no higher than in the corresponding period of 2019 (around 32,000 weekly calls) (Fig. 4a), but during subsequent waves they increased above pre-pandemic levels (more than 35,000 weekly calls in late 2020 and spring 2021). Figure 4b illustrates the heterogeneity in the time profiles of calls across states that we seek to ‘explain’ with state–week variation in our three explanatory variables: SARS-CoV-2 infection rates³⁹ (Fig. 4c), NPIs as measured by the components ‘containment and closure policies’ summarized in the stringency index (Fig. 4d), and the generosity of public compensation payments for labour costs (for example, furlough payments) as measured by the component ‘income support’ (Fig. 4e) in the Oxford COVID-19 Government Response Tracker⁴⁰.

Fig. 4: Lifeline calls, COVID-19 and policy measures in the USA.

 figure4

- a**, Total number of weekly calls routed to Lifeline centres by year, with three-week moving average (vertical axis is truncated). **b**, Deviation of logged calls from the time-averaged state-level mean (grey), with nationwide weekly average (black). **c–e**, Weekly average scores, with individual state values (grey). **c**, Newly confirmed SARS-CoV-2 infections³⁹ per 100,000 population (mean in red). **d**, Government response stringency index (mean in blue). **e**, Income support index⁴⁰ (mean in yellow). **f**, Estimated coefficients and associated 95% confidence intervals from sub-national

panel model including state and week fixed effects. The dependent variable is $\ln(\text{Lifeline calls} + 1)$ and natural log values of the independent variables are used ([Methods](#), equation (5)). **g**, Coefficient estimates for interaction terms with indicators for the two periods from January to August 2020 and September 2020 to March 2021, and associated 95% confidence intervals ([Methods](#), equation (6)).

[Source data](#).

In Fig. [4f](#), we summarize our regression results based on data up to March 2021. For given policy measures, increases in SARS-CoV-2 infections were associated with statistically significant decreases in the number of calls to the suicide helpline (elasticity = -0.012 , 95% CI: $-0.023, -0.001$; $P = 0.026$) (Fig. [4f](#)). The estimated coefficient implies that a 10% increase in SARS-CoV-2 infections is associated with a 0.1% reduction in calls to the suicide helpline.

One interpretation of this result is that the pandemic itself attenuates suicidal anxieties, perhaps by shifting people's focus towards the distress of others, or to their own fear of the pandemic. This interpretation is consistent with the evolution of calls to the US Disaster Distress Helpline, which was advertised for providing crisis counselling to people affected by COVID-19: calls to this helpline increased sharply during the initial phase of the pandemic (from around 500 to around 3,000 weekly calls) (Extended Data Fig. [6](#)), suggesting some displacement of pre-existing anxieties by more proximate COVID-19-related sources of distress.

Policy interventions in the shape of more stringent state-level NPIs or more generous state-level income support measures were not found to have statistically significant effects on Lifeline calls (effect of NPI stringency: 0.020 ; 95% CI: $-0.007, 0.047$; $P = 0.155$) (Fig. [4f](#)). Even though the data do not have the statistical power to reject the hypothesis of no effect, our estimates are consistent with stricter NPIs being followed by an increase in Lifeline calls and with income support policies having the opposite effect.

Figure [4g](#) shows the estimated effects of the three explanatory variables separately for the first and subsequent waves of the pandemic, with the cut-off date placed at 1 September 2020. We find that the dampening effect on Lifeline calls of the pandemic itself (measured as the number of SARS-CoV-2 infections) increased over time (-0.022 during the second sub-period; 95% CI: $-0.038, -0.006$; $P = 0.006$) (Fig. [4g](#)). The effects on Lifeline calls of more stringent NPIs or more generous income support, however, did not differ noticeably across waves of the pandemic. Together, these estimates confirm that the mental-health implications of the pandemic remained relatively stable across the first and subsequent waves. In Supplementary Tables [7, 8](#), we show that these qualitative results are robust across a range of panel regression specifications.

The pattern observed in US suicide helpline data is corroborated by a corresponding regression analysis based on the German and French helplines: all other things being equal, increasing SARS-CoV-2 infections and more generous income support policies were followed by falls in the number of helpline calls related to suicidality, with an elasticity of -0.024 (95% CI: $-0.035, -0.014$; $P < 0.001$) and -0.020 (95% CI: $-0.033, -0.006$; $P = 0.004$), respectively (Fig. 5). Conversely, more stringent NPIs were followed by more suicide-related calls ($+0.035$, 95% CI: $0.011, 0.060$; $P = 0.005$) (Fig. 5). These estimated effects are statistically significant and qualitatively consistent with those based on the Lifeline data.

Fig. 5: Daily calls by conversation topic, COVID-19, and policy in Germany and France.

 figure5

Coefficients from separate regression models by topic with 95% confidence intervals. The dependent variable is $\ln(\text{Calls} + 1)$, and natural log values of SARS-CoV-2 infections per 100,000 population³⁹ and policy indices⁴⁰ are used. The sample includes all calls to Telefonseelsorge (Germany) and SOS Amitié (France) for which at least

one conversation topic was recorded, aggregated to daily totals from 1 January 2019 to 31 March 2021 ([Methods](#), equation (2)).

[Source data](#).

Our findings suggest that public compensation payments for pandemic-induced losses not only reduce economic hardship but also have broader benefits: more generous income support leads to fewer calls due to fear (-0.042 ; 95% CI: -0.061 , -0.024 ; $P < 0.001$), loneliness (-0.024 ; 95% CI: -0.040 , -0.008 ; $P = 0.003$), physical health concerns (-0.026 ; 95% CI: -0.041 , -0.011 ; $P = 0.001$) and, as expected, economic anxiety ('livelihood'; -0.016 ; 95% CI: -0.030 , -0.002 ; $P = 0.023$) (Fig. 5).

Discussion

We drew on international helpline call data to shed light on a statistical blind spot of pandemic policy: mental health concerns and general distress of the population. Helpline calls increased during the pandemic, and this increase was driven primarily by concerns linked to the pandemic itself (such as fear of infection and loneliness). Conversely, on average, the share of calls due to other forms of distress, including suicidality, violence and addiction, decreased. The lack of an increase in the share of suicide-related calls is consistent with observed decreases in actual suicides during the early stages of the pandemic across several countries²⁸. Underlying these general patterns are helpline-specific evolutions that are documented in detail in Supplementary Figs. 1–37.

The panel structure of the data enables us to estimate multivariate models to disentangle the separate effects of the pandemic itself (SARS-CoV-2 infections), of the stringency of containment policies, and of the generosity of income support policies. We found that more stringent measures were associated with a higher number of calls due to fear, loneliness and suicidality, but that more generous income support had the opposite effect. This implies that compensation payments to workers and businesses affected economically by COVID-19, which were designed to preserve demand and productive capacity, have additional benefits in alleviating distress and mental health concerns.

Longitudinal helpline data offer an attractive complement to existing empirical approaches based on surveys, administrative and clinical data (such as suicide statistics, admissions to treatment centres and so on), and internet search data.

Helpline data also have their limitations. One such issue is that call counts may be influenced not only by demand but also by supply, as capacity constraints can force operators to leave some calls unanswered. This may cause a downward bias in measures of increases in demand. However, as calls are not pre-screened, capacity

constraints are unlikely to affect analyses of the composition of calls in terms of topics or caller characteristics.

Another limitation lies in our agnosticism about the representativeness of callers to helplines. We are aware of no rigorous evidence regarding the composition of the helpline caller population in terms of socio-demographic status, health, occupation, nationality and other factors. By focusing our analysis on changes in call volumes over time, we eliminate time-invariant specificities of the helpline caller population, which should remove a large proportion of any potential sample selection bias. Moreover, anecdotal evidence from helpline workers confirms that the caller population typically includes the most vulnerable members of society, which is the population of greatest interest in a study of distress and mental health concerns.

Methods

Helpline call data

Our sample of helplines includes large general crisis helplines and dedicated suicide prevention helplines, as well as some helplines that focus on specific groups such as children, parents or immigrants. Observations within helplines are self-selected, as they consist of callers to helplines. The selection of sample helplines was based on (1) an internet search of well-documented helplines, and (2) receiving data from those helplines. Out of 154 helplines that we contacted, we received data from 37 helplines. Where possible, we requested data from 1 January 2019 to the most recent available date, to enable a comparison of call patterns after the COVID-19 outbreak with call patterns at the same time of year before the pandemic. The information obtained from 23 helplines was of sufficiently detailed coverage and consistency to be included in our pooled analyses. Extended Data Table 1 lists the included helplines, grouped by the format in which the data were made available for this study.

The most detailed information was provided by the helplines in Extended Data Table 1a, where we received individual conversation-level data, including information on the callers' sex and approximate age, as well as on the issues discussed during the conversation. From the three additional helplines in Extended Data Table 1b, we received aggregate time series of daily call volumes, with separate series by gender, age category and topic. Moreover, for the two helplines in Extended Data Table 1c, we received sub-national weekly series of call volumes across US states. In contrast to the data from helplines in Extended Data Table 1a,b, the number of calls in Extended Data Table 1c does not refer to answered calls and actual conversations, but to the raw number of calls routed to local centres.

Data processing and analysis were conducted according to the guidelines of the Internal Review Board (IRB) of the Faculty of Business and Economics at the University of Lausanne. Two considerations were important. First, all helplines guarantee anonymity to their callers, both towards their operators and towards the outside world. Names and addresses are never asked for, and caller numbers are hidden by the system. It is therefore impossible to identify callers even from the call-level data provided by a subset of helplines, and the anonymous information is not covered by data protection considerations. Second, all of the helplines we analysed inform callers that anonymous call data are collected for reporting and statistical purposes, whether explicitly in the terms and conditions or statutes, and/or implicitly in annual reports and online publications. The analysis of those data conforms with the aim of the Ethics Charter of the International Federation of Telephone Emergency Services (IFOTES), which aims to “(c)ollect and disseminate data gathered by the Branches in connection with the challenges facing Mental Health and Prevention of Suicide” and to “(a)ssist and encourage research carried out in these fields”¹². The IRB exempted the study from a full review owing to the secondary nature of the data used.

Government response and epidemiological data

To measure the timing and intensity of government responses consistently across time, countries and sub-national regions, we rely on aggregate policy indices from the Oxford COVID-19 Government Response Tracker⁴⁰. In particular, we use two policy indices, the government response stringency index and the income support index. The stringency index shows the strictness of containment policies and restrictions of personal freedom, and is based on an unweighted average of eight component scores for shelter-in-place requirements, workplace and public transport closures, restrictions on public events, gatherings, domestic and international travel, and information campaigns. The income support index reflects the availability of financial support and is constructed the index score using the ordinal measure and the flag for sectoral targeting to arrive at a value between 0 and 100 (following the definition of the stringency index). For the sub-national information on policies in US states, it is important to note that we use the total index scores, where, whenever national policies were more restrictive than those of individual states, the higher score is imputed. Data on the daily number of newly confirmed SARS-CoV-2 infections are taken from the JHU CSSE COVID-19 Dataset³⁹.

Call volumes after the pandemic outbreak

For Fig. 1, we combine the time-series data (Extended Data Table 1b) with aggregates based on the call-level data (Extended Data Table 1a) in a panel of daily call volumes for 21 helplines, covering the time up to 30 June 2020 if available. For four of the

helplines (MIELI, SOS Détresse, Sahar and Muslimisches Seelsorgetelefon), no data were available for 2019. We then look exclusively on the period from 4 weeks before to 12 weeks after the country-specific event date in 2020, as well as, if available, the corresponding days in 2019. To summarize the overall dynamics, we estimate the following model:

$$\begin{aligned} \ln(n) = & \sum_{\tau=-4}^{-1} \gamma_\tau \ln(W) + \sum_{\tau=1}^{12} \gamma_\tau \ln(W) \\ & + \sum_{k=1}^K \zeta_k h_k + \xi_h + \theta_t + \epsilon_h \end{aligned} \quad (1)$$

The dependent variable is the natural logarithm of the number of calls to helpline h recorded on day t . We define the local outbreak as the date when (1) the cumulative number of SARS-CoV-2 infections in the population exceeded 1/100,000 or (2) when SIP orders were first introduced. For both versions, we define indicator variable Week τ , which is set ‘1’ for days in event week number τ in 2020. The model includes helpline fixed effects ζ_h interacted with year, week-of-year and day-of-week indicators, summarized in the vector Θ_t . The reference category is week 0 of the pandemic outbreak or SIP introduction, and the coefficient γ^τ allows us to track the percentage deviation in daily calls, controlling for seasonal effects and secular trends. See Supplementary Table 1 and Extended Data Fig. 1 for details on event dates and call volumes for each of the 21 included helplines. Supplementary Table 2 contains numerical estimation results.

Helpline data on individual calls

To investigate changes in conversation topics (Fig. 2), we focus on the call-level data and combine information from 12 helplines (Extended Data Table 1a) for which we have information on conversation topics and inferred caller characteristics. This yields a sample of up to 2.2 million calls. For each helpline, we categorize calls on the basis of the recorded information on the problems of callers and the topics discussed. Precise categorizations of call topics differ across helplines, but they are sufficiently similar to allow us to map them into the following common, non-exclusive categories: loneliness (social isolation, entrapment), fear (general fear, anxiety disorder, fear of infection with SARS-CoV-2), suicidality (suicidal ideation, suicidal thoughts or plans, suicide attempts, suicidality of others), addiction (drugs, alcohol, other addictions), violence (physical violence and abuse, sexual harassment, rape), physical health (disease, long-term illness, disability), and two broad categories for livelihood (work situation, unemployment, financial problems, housing), and relationships (family life, parenting, marriage and intimate relationships, separation).

Supplementary Tables 12–22 show the precise topic definitions for each helpline. As some topics are not recorded at all for some helplines, the sample size differs depending on which topic we look at: the largest sample includes data from 12 helplines, where we can distinguish calls related to suicide from calls concerning other issues (Extended Data Fig. 2). Recorded conversation topics can be non-exclusive. We document the joint distribution of topics in Extended Data Fig. 5a.

Additionally, we have coded the sex and age category of each caller, and (where possible) further characteristics such as marital status, living situation and occupational status. As helplines record age categories differently, our classification cannot be fully precise. Using the boundaries of available age groups, the group of callers below 30 includes only those that were recorded in an age group with an upper limit at or below 30. The same logic applies to the group of callers older than 60, and the middle category in some cases includes also individuals whose age is slightly below 30 years or above 60.

For Fig. 2, we restrict the sample to calls recorded for the time from 1 January 2019 through 30 June 2020, where information on sex and age group of callers is available. When estimating the relative importance of a topic, we define the dependent variable T as equal to ‘1’ for call i to helpline h on day t if the conversation was related to the respective topic (Fear, Loneliness, Suicide, Addiction, Violence, Physical health, Livelihood, or Relationships), and zero for unrelated calls, where another topic was recorded. Calls without information on caller issues or conversation topics are not included. Based on the date when the cumulative number of SARS-CoV-2 infections per population exceeded 1/100,000 in the country of operation³⁹, we define an indicator ‘Post outbreak’ and estimate a linear probability model as in equation (2):

$$\begin{aligned} \text{\$\$}\{ T \}_{i,h,t} = & \gamma_P \{ \text{\rm P} \} \{ \text{\rm o} \} \{ \text{\rm s} \} \{ \text{\rm t} \} \{ \text{\rm o} \} \{ \text{\rm u} \} \\ & \{ \text{\rm t} \} \{ \text{\rm b} \} \{ \text{\rm r} \} \{ \text{\rm e} \} \{ \text{\rm a} \} \{ \text{\rm k} \} \}_{h,t} + \xi_h \times \\ & \{ \boldsymbol{\Theta} \}_t + \epsilon_{i,h,t} \end{aligned} \quad (2)$$

The model includes the helpline indicator ξ_h to account for time-invariant differences among helplines. We further add year, week-of-year and day-of-week indicators, summarized in the vector $\boldsymbol{\Theta}_t$, interacted with the helpline fixed effects, to account for secular trends and for seasonal and day-of-week effects. Standard errors are clustered at the helpline–week level. Supplementary Table 3 contains numerical estimation results.

For the analysis of heterogeneous effects in Extended Data Fig. 4, we estimate an alternative specification including individual caller characteristics and interaction terms. To illustrate the change in topics for different groups, we classify callers into

six non-overlapping groups, denoted in the vectors **Sex** (male, female) and **Age group** (below 30, 30–60, above 60). In the model illustrated in equation (3), we interact the post-outbreak variable Post with all six group indicators, so that the coefficients represent the group-specific changes in topic shares:

$$\begin{aligned} \text{\$}\{ T \}_{i,h,t} = & \beta (\{ \bf{S} \} \{ \bf{e} \} \{ \{ \bf{x} \} \}_{i,t} \times \{ \bf{A} \} \{ \bf{g} \}) \\ & \{ \bf{e} \} \{ \bf{g} \} \{ \bf{r} \} \{ \bf{o} \} \{ \bf{u} \} \{ \bf{p} \}_{i,t} + \gamma \{ \rm{P} \} \\ & \{ \rm{o} \} \{ \rm{s} \} \{ \rm{t} \}_{h,t} (\{ \bf{S} \} \{ \bf{e} \} \{ \{ \bf{x} \} \}_{i,t} \times \{ \bf{A} \} \{ \bf{g} \} \{ \bf{e} \} \{ \bf{g} \} \{ \bf{r} \} \{ \bf{o} \} \{ \bf{u} \} \{ \bf{p} \}_{i,t}) + \xi \\ & \{ h \} \times \{ \Theta \}_{t} + \{ \epsilon \}_{i,h,t} \end{aligned} \quad (3)$$

For the main effects of caller sex and age groups, indicators for the reference group of male callers in the 30–60 age category are omitted. Supplementary Table 4 contains numerical estimation results.

For the analysis of the longer time horizon and subsequent waves in Fig. 3, we focus on call-level data from Germany and France, from 1 January 2019 to 31 March 2021. We estimate a specification similar to the previous approach, separately for the two helplines and each topic. To distinguish the changes around the outbreak from later adjustments during the subsequent wave, we define two indicator variables W1 and W2 denoting two periods. The first covers the time from 11 March 2020, when the World Health Organization declared the outbreak a pandemic, to 30 June 2020, when the number of infections decreased again and containment measures were relaxed both in Germany and in France. The second period indicator is equal to one for the time after 1 October 2020. Equation (4) illustrates the estimated model:

$$\begin{aligned} \text{\$}\{ T \}_{i,h,t} = & \gamma_1 \{ \rm{W} \}_1 \{ h,t \} + \gamma_2 \{ \rm{W} \}_2 \{ h,t \} + \{ \Theta \}_{t} + \{ \epsilon \}_{i,h,t} \end{aligned} \quad (4)$$

As we analyse the two helplines separately, we do not include helpline fixed effects here, but capture secular trends and seasonal patterns through the inclusion of year, week-of-year, and day-of-week indicators summarized in the vector Θ_t . Standard errors are clustered at the week level. See Supplementary Table 6 for numerical estimation results.

Call volumes across US states

The analysis of sub-national call volumes in Fig. 4 relies on data on weekly call volumes routed to Lifeline. The analysis is based on weekly call volumes for US states and territories over 116 weeks, starting in the week to 6 January 2019, and up to the week ending on Sunday, 21 March 2020 (Supplementary Fig. 37). Based on phone

numbers, the state from which calls were placed can be inferred, even though internal migration means that this classification is subject to measurement error. While the Lifeline will serve any calls regardless of country of origin, its mission is to serve calls originating from the US and US territories. We focus on calls from 50 US states and DC. Calls originating from Canadian provinces, US territories, as well as those of other international or unknown origin are not considered, to maximize consistency and because of the limited availability of data on policy responses. The panel structure allows us to exploit the idiosyncratic variation within states j over time (weeks w) while controlling for overall trends. We estimate a two-way error component model as illustrated in equation (5):

$$\begin{aligned}
 & \$\$ \{ \rm{l} \} \{ \rm{n} \} (\{ \rm{C} \} \{ \rm{a} \} \{ \rm{l} \} \{ \rm{l} \} \{ \rm{s} \}) _{j,w} + 1) = \\
 & \{ \rm{\pi} \}_{1} \{ \rm{l} \} \{ \rm{n} \} (\{ \rm{I} \} \{ \rm{n} \} \{ \rm{f} \} \{ \rm{e} \} \\
 & \{ \rm{c} \} \{ \rm{t} \} \{ \rm{i} \} \{ \rm{o} \} \{ \rm{n} \} \{ \rm{s} \}) _{j,w} + 1) + \{ \rm{\pi} \}_{2} \\
 & \{ \rm{l} \} \{ \rm{n} \} (\{ \rm{S} \} \{ \rm{t} \} \{ \rm{r} \} \{ \rm{i} \} \{ \rm{n} \} \{ \rm{g} \} \\
 & \{ \rm{e} \} \{ \rm{n} \} \{ \rm{c} \} \{ \rm{y} \}) _{j,w} + 1) + \{ \rm{\pi} \}_{3} \{ \rm{l} \} \\
 & \{ \rm{n} \} (\{ \rm{l} \} \{ \rm{n} \} \{ \rm{c} \} \{ \rm{o} \} \{ \rm{m} \} \{ \rm{e} \} \{ \rm{s} \} \\
 & \{ \rm{u} \} \{ \rm{p} \} \{ \rm{p} \} \{ \rm{o} \} \{ \rm{r} \} \{ \rm{t} \}) _{j,w} + 1) + \{ \rm{x}_i \}_{j} + \\
 & \{ \theta_w \} + \{ \epsilon \}_{j,w} \$\$ \\
 \end{aligned} \tag{5}$$

The dependent variable is the natural logarithm of the number of calls plus one, $\ln(\text{Calls}) + 1$. The independent variables include the natural logarithm of the population ($\ln(P)$), the natural logarithm of the number of infections per 100,000 population ($\ln(I)$), the natural logarithm of the stringency index ($\ln(S)$), the natural logarithm of the income support index ($\ln(O)$), and the natural logarithm of the month indicator ($\ln(M)$). State fixed effects ξ_j absorb all time-invariant factors, and our analysis is therefore based on the idiosyncratic within-state variation in call volumes over time. The inclusion of week indicators θ_w allows us to capture all nation-wide and global effects and to focus solely on the relative differences in pandemic exposure and policy response. Standard errors are clustered at the state-month level.

To investigate the extent to which the relationship changed over time, we re-estimate the model as in equation (6). Here, we include the three main explanatory variables, interacted with two indicator variables that are set to “1” for the time period from 1 January to 31 August 2020, and for 1 September 2020 to 21 March 2021, respectively. Supplementary Tables 7, 8 contain numerical estimation results from alternative specifications.

$$\begin{aligned}
 & \$\$ \begin{array}{ccc} \{ \rm{l} \} \{ \rm{n} \} (\{ \rm{C} \} \{ \rm{a} \} \{ \rm{l} \} \{ \rm{l} \} \{ \rm{s} \}) _{j,w} + 1) & = & \{ \rm{J} \} \{ \rm{a} \} \{ \rm{n} \} - \{ \rm{A} \} \{ \rm{u} \} \\
 & \{ \rm{g} \} \}_{w} \{ \varphi_1 \}_{1} \{ \rm{l} \} \{ \rm{n} \} (\{ \rm{I} \} \{ \rm{n} \} \{ \rm{f} \} \\
 & \{ \rm{e} \} \{ \rm{c} \} \{ \rm{t} \} \{ \rm{i} \} \{ \rm{o} \} \{ \rm{n} \} \{ \rm{s} \}) _{j,w} + 1) & + & \end{array} \\
 \end{aligned}$$

$$\begin{aligned}
& \{\backslash varphi\}_{2}\backslash,\{\backslash rm{l}\}\{\backslash rm{n}\}(\{\backslash rm{S}\}\{\backslash rm{t}\}\{\backslash rm{r}\}\{\backslash rm{i}\}\{\backslash rm{n}\}) \\
& \{\backslash rm{g}\}\{\backslash rm{e}\}\{\backslash rm{n}\}\{\backslash rm{c}\}\{\backslash rm{y}\}_{j,w}+1) \& +\{\backslash varphi\}_{3}\backslash, \\
& \{\backslash rm{l}\}\{\backslash rm{n}\}\{\backslash rm{l}\}\{\backslash rm{n}\}\{\backslash rm{c}\}\{\backslash rm{o}\}\{\backslash rm{m}\}\{\backslash rm{e}\}), \\
& \{\backslash rm{s}\}\{\backslash rm{u}\}\{\backslash rm{p}\}\{\backslash rm{p}\}\{\backslash rm{o}\}\{\backslash rm{r}\}\{\backslash rm{t}\}_{j,w}+1)] \& & \\
& +\{(\backslash rm{S}\}\{\backslash rm{e}\}\{\backslash rm{p}\}-\{\backslash rm{M}\}\{\backslash rm{a}\}\{\backslash rm{r}\})\}_{w}[\{\backslash psi\}_1], \\
& \{\backslash rm{l}\}\{\backslash rm{n}\}\{\backslash rm{l}\}\{\backslash rm{n}\}\{\backslash rm{f}\}\{\backslash rm{e}\}\{\backslash rm{c}\}\{\backslash rm{t}\}\{\backslash rm{i}\} \\
& \{\backslash rm{o}\}\{\backslash rm{n}\}\{\backslash rm{s}\}_{j,w}+1) \& & +\{\backslash psi\}_2\{\backslash rm{l}\}\{\backslash rm{n}\} \\
& (\{\backslash rm{S}\}\{\backslash rm{t}\}\{\backslash rm{r}\}\{\backslash rm{i}\}\{\backslash rm{n}\}\{\backslash rm{g}\}\{\backslash rm{e}\}\{\backslash rm{n}\}\{\backslash rm{c}\} \\
& \{\backslash rm{y}\}_{j,w}+1)+\{\backslash psi\}_3] \& & \& \& \{\backslash rm{l}\}\{\backslash rm{n}\}(\{\backslash rm{l}\}\{\backslash rm{n}\} \\
& \{\backslash rm{c}\}\{\backslash rm{o}\}\{\backslash rm{m}\}\{\backslash rm{e}\}\{\backslash rm{s}\}\{\backslash rm{u}\}\{\backslash rm{p}\}\{\backslash rm{p}\}\{\backslash rm{o}\} \\
& \{\backslash rm{r}\}\{\backslash rm{t}\}_{j,w}+1)] \& & +\{\backslash xi\}_j+\{\backslash theta\}_w+\{\backslash epsilon\} \\
& \}_{j,w}\backslash end{array}\$\$ \\
(6)
\end{aligned}$$

Call volumes in Germany and France

For the analysis in Fig. 5, we combine the previous approaches and estimate the relationship between call volumes and the three variables as illustrated in equation (7), based on topic-specific call volumes to Telefonseelsorge (Germany) and SOS Amitié (France) during the time from 1 January 2019 to 31 March 2021.

$$\begin{aligned}
& \$\$ \{\backslash rm{l}\}\{\backslash rm{n}\}(\{\backslash rm{C}\}\{\backslash rm{a}\}\{\backslash rm{l}\}\{\backslash rm{l}\}\{\backslash rm{s}\})_{h,t}+1)= \\
& \{\backslash rm{\pi}\}_1\{\backslash rm{l}\}\{\backslash rm{n}\}(\{\backslash rm{l}\}\{\backslash rm{n}\}\{\backslash rm{f}\}\{\backslash rm{e}\} \\
& \{\backslash rm{c}\}\{\backslash rm{t}\}\{\backslash rm{i}\}\{\backslash rm{o}\}\{\backslash rm{n}\}\{\backslash rm{s}\})_{h,t}+1)+\{\backslash rm{\pi}\}_2\{\backslash rm{l}\}\{\backslash rm{n}\} \\
& (\{\backslash rm{S}\}\{\backslash rm{t}\}\{\backslash rm{r}\}\{\backslash rm{i}\}\{\backslash rm{n}\}\{\backslash rm{g}\}) \\
& \{\backslash rm{e}\}\{\backslash rm{n}\}\{\backslash rm{c}\}\{\backslash rm{y}\})_{h,t}+1)+\{\backslash rm{\pi}\}_3\{\backslash rm{l}\}\{\backslash rm{n}\} \\
& \{\backslash rm{l}\}\{\backslash rm{n}\}\{\backslash rm{c}\}\{\backslash rm{o}\}\{\backslash rm{m}\}\{\backslash rm{e}\}\{\backslash rm{s}\} \\
& \{\backslash rm{u}\}\{\backslash rm{p}\}\{\backslash rm{p}\}\{\backslash rm{o}\}\{\backslash rm{r}\}\{\backslash rm{t}\})_{h,t}+1)+\{\backslash xi\}_h\times \\
& \{\boldsymbol{\Theta}_t\}+\{\backslash epsilon\}_h \$\$ \\
(7)
\end{aligned}$$

In contrast to the sub-national panel of US states, here we do not include week fixed effects but capture secular trends and seasonal patterns through helpline fixed effects ξ_h , interacted with year, week-of-year and day-of-week indicators summarized in the vector Θ_t . Standard errors are clustered at the helpline–week level. Numerical estimation results are shown in Supplementary Table 9.

Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

Data availability

Data were provided by helplines for the sole purpose of this research project, subject to confidentiality agreements. The full data underlying specific parts of the analysis are available from the authors upon reasonable request and conditional on permission of the respective helplines. To obtain (updated) helpline data, researchers have to sign agreements with individual helplines—for further information, contact: presse@telefonseelsorge.de (Telefonseelsorge, Germany), contact@sosamitieparisidf.fr (SOS Amitié), info@deluisterlijn.nl (De Luisterlijn), info@nummergegenkummer.de (Nummer gegen Kummer), federatie@tele-onthaal.be (Tele-Onthaal), telefonseelsorge@edw.or.at (Telefonseelsorge, Austria), liccenter@163.com (Hope Line), telefonoamico@telefonoamico.it (Telefono Amico), samarijan@gmail.com (Zaupni Telefon Samarijan), hana.regnerova@modralinka.cz (Modrá linka), support@sahar.org.il (Sahar), direccaolphm@gmail.com (SOS Voz Amiga), info@mutes.de (Muslimisches Seelsorgetelefon), info@embracelebanon.org (Embrace Lifeline), info@sosdetresse.lu (SOS Détresse), info@novageneracija.org (Plavi Telefon), sbhinfo@sbhk.org.hk (Samaritan Befrienders), verband@143.ch (Die Dargebotene Hand), sos-keskus@mieli.fi (MIELI), lesziroda@gmail.com (LESZ) and info@vibrant.org (Lifeline and Disaster Distress Helpline). Data on infection rates and policy measures are publicly available online from the JHU CSSE COVID-19 Dataset at <https://github.com/CSSEGISandData> and the Oxford COVID-19 Government Response Tracker at <https://github.com/OxCGRT>. Source data are provided with this paper.

Code availability

Files were collected in MS Excel 2016 and Notepad++ v7.9.5. Data preparation and analysis was carried out in Stata/SE 17.0, Do-files are available online at <https://doi.org/10.5281/zenodo.5495830>.

References

1. 1.

COVID-19 and the Need for Action on Mental Health UN Sustainable Development Group Policy Brief (United Nations, 2020).

2. 2.

Tackling the Mental Health Impact of the COVID-19 Crisis: An Integrated, Whole-of-Society Response (OECD, 2021).

3. 3.

Rehm, J. & Shield, K. D. Global burden of disease and the impact of mental and addictive disorders. *Curr. Psychiatry Rep.* **21**, 10 (2019).

4. 4.

Batchelor, S., Stoyanov, S., Pirkis, J. & Kõlves, K. Use of kids helpline by children and young people in Australia during the COVID-19 pandemic. *J. Adolesc. Health* **68**, 1067–1074 (2021).

5. 5.

Zalsman, G. et al. Suicide-related calls to a national crisis chat hotline service during the COVID-19 pandemic and lockdown. *J. Psychiatr. Res.* **139**, 193–196 (2021).

6. 6.

Turkington, R. et al. Behavior of callers to a crisis helpline before and during the COVID-19 pandemic: quantitative data analysis. *JMIR Ment. Health* **7**, e22984 (2020).

7. 7.

Halford, E. A., Lake, A. M. & Gould, M. S. Google searches for suicide and suicide risk factors in the early stages of the COVID-19 pandemic. *PLoS ONE* **15**, e0236777 (2020).

8. 8.

Armbruster, S. & Klotzbücher, V. Lost in lockdown? COVID-19, social distancing, and mental health in Germany. *Covid Econ.* **22**, 117–153 (2020).

9. 9.

Brülhart, M. & Lalive, R. Daily suffering: helpline calls during the COVID-19 crisis. *Covid Econ.* **19**, 143–158 (2020).

10. 10.

Leach, L. S. & Christensen, H. A systematic review of telephone-based interventions for mental disorders. *J. Telemed. Telecare* **12**, 122–129 (2006).

11. 11.
Coveney, C. M., Pollock, K., Armstrong, S. & Moore, J. Callers' experiences of contacting a national suicide prevention helpline. *Crisis* **33**, 313–324(2012).
12. 12.
IFOTES Ethical Charter (IFOTES, 1993).
13. 13.
De Leo, D., Buono, M. D. & Dwyer, J. Suicide among the elderly: the long-term impact of a telephone support and assessment intervention in northern Italy. *Br. J. Psychiatry* **181**, 226–229 (2002).
14. 14.
Choi, D. et al. Development of a machine learning model using multiple, heterogeneous data sources to estimate weekly US suicide fatalities. *JAMA Netw. Open* **3**, e2030932 (2020).
15. 15.
Ridley, M., Rao, G., Schilbach, F. & Patel, V. Poverty, depression, and anxiety: causal evidence and mechanisms. *Science* **370**, eaay0214 (2020).
16. 16.
McInerney, M., Mellor, J. M. & Nicholas, L. H. Recession depression: mental health effects of the 2008 stock market crash. *J. Health Econ.* **32**, 1090–1104 (2013).
17. 17.
Parmar, D., Stavropoulou, C. & Ioannidis, J. P. A. Health outcomes during the 2008 financial crisis in Europe: systematic literature review. *BMJ* **354**, i4588 (2016).
18. 18.
Chang, S.-S., Stuckler, D., Yip, P. & Gunnell, D. Impact of 2008 global economic crisis on suicide: time trend study in 54 countries. *BMJ* **347**, f5239 (2013).
19. 19.

Phillips, J. A. & Nugent, C. N. Suicide and the Great Recession of 2007–2009: the role of economic factors in the 50 US states. *Soc. Sci. Med.* **116**, 22–31 (2014).

20. 20.

Nordt, C., Warnke, I., Seifritz, E. & Kawohl, W. Modelling suicide and unemployment: a longitudinal analysis covering 63 countries, 2000–11. *Lancet Psychiatry* **2**, 239–245 (2015).

21. 21.

Gunnell, D. et al. Suicide risk and prevention during the COVID-19 pandemic. *Lancet Psychiatry* **7**, 468–471 (2020).

22. 22.

Zortea, T. C. et al. The impact of infectious disease-related public health emergencies on suicide, suicidal behavior, and suicidal thoughts. *Crisis* <https://doi.org/10.1027/0227-5910/a000753> (2020).

23. 23.

Richter, D., Riedel-Heller, S. & Zürcher, S. Mental health problems in the general population during and after the first lockdown phase due to the SARS-CoV-2 pandemic: rapid review of multi-wave studies. *Epidemiol. Psychiatr. Sci.* **30**, E27 (2021).

24. 24.

Banks, J., et al. The mental health effects of the first two months of lockdown during the COVID-19 pandemic in the UK. *Fisc. Stud.* **41**, 685–708 (2020).

25. 25.

Holman, E. A., Thompson, R. R., Garfin, D. R. & Silver, R. C. The unfolding COVID-19 pandemic: a probability-based, nationally representative study of mental health in the United States. *Sci. Adv.* **6**, eabd5390 (2020).

26. 26.

Holingue, C. et al. Mental distress during the COVID-19 pandemic among US adults without a preexisting mental health condition: Findings from American trend panel survey. *Prev. Med.* **139** (2020).

27. 27.

John, A., Pirkis, J., Gunnell, D., Appleby, L. & Morrissey, J. Trends in suicide during the COVID-19 pandemic. *BMJ* **371**, m4352 (2020).

28. 28.

Pirkis, J. et al. Suicide trends in the early months of the COVID-19 pandemic: Interrupted time series analysis of preliminary data from 21 countries. *Lancet Psychiatry* **8**, 579–588 (2021).

29. 29.

Tanaka, T. & Okamoto, S. Increase in suicide following an initial decline during the COVID-19 pandemic in Japan. *Nat. Hum. Behav.* **5**, 229–238 (2021).

30. 30.

Berger, L. M., Ferrari, G., Leturcq, M., Panico, L. & Solaz, A. COVID-19 lockdowns and demographically relevant Google Trends: a cross-national analysis. *PLoS ONE* **16**, e0248072 (2021).

31. 31.

Brodeur, A., Clark, A. E., Fleche, S. & Powdthavee, N. COVID-19, lockdowns and well-being: evidence from Google Trends. *J. Public Econ.* **193**, 104346 (2020).

32. 32.

Silverio-Murillo, A., Hoehn-Velasco, L., Tirado, A. R. & de la Miyar, J. R. B. COVID-19 blues: lockdowns and mental health-related Google searches in Latin America. *Soc. Sci. Med.* **281**, 114040 (2021).

33. 33.

Fetzer, T., Hensel, L., Hermle, J. & Roth, C. Coronavirus perceptions and economic anxiety. *Rev. Econ. Stat.* https://doi.org/10.1162/rest_a_00946 (2020).

34. 34.

Faust, J. S. et al. Suicide deaths during the COVID-19 stay-at-home advisory in Massachusetts, March to May 2020. *JAMA Netw. Open* **4**, e2034273 (2021).

35. 35.

Holland, K. M. et al. Trends in US emergency department visits for mental health, overdose, and violence outcomes before and during the COVID-19 pandemic. *JAMA Psychiatry* **78**, 372–379 (2021).

36. 36.

Baron, E. J., Goldstein, E. G. & Wallace, C. T. Suffering in silence: how COVID-19 school closures inhibit the reporting of child maltreatment. *J. Public Econ.* **190**, 1–13 (2020).

37. 37.

Bullinger, L. R., Carr, J. B. & Packham, A. COVID-19 and crime: effects of stay-at-home orders on domestic violence. *Am. J. Health Econ.* **7**, 249–280 (2021).

38. 38.

Leslie, E. & Wilson, R. Sheltering in place and domestic violence: evidence from calls for service during COVID-19. *J. Public Econ.* **189**, 104241 (2020).

39. 39.

Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* **20**, 533–534 (2020).

40. 40.

Hale, T. et al. A global panel database of pandemic policies (Oxford COVID-19 Government Response Tracker). *Nat. Hum. Behav.* **5**, 529–538 (2021).

Acknowledgements

We thank the following people and helplines for sharing their expertise and for granting us access to their data: C. Hochhauser and A. Kesselring (Telefonseelsorge, Austria), J. Pots (Tele-Onthaal, Belgium), M. Kovacevic (Plavi Telefon, Bosnia and Herzegovina), R. Ma and W. Ni (Hope Line, China), H. Regnerova (Modrá linka, Czech Republic), H. Dumont (SOS Amitié, France), S. Winter (MIELI, Finland), L. Storch and B. Blömeke (Telefonseelsorge, Germany), M. I. Sagir (Muslimisches Seelsorgetelefon, Germany), H. Schütz (Nummer gegen Kummer, Germany), H.-Chia (Samaritan Befrienders, Hong Kong), E. Brandisz (LESZ, Hungary), Y. Levy (Sahar, Israel), M. Petra (Telefono Amico, Italy), P. Zeinoun (Embrace Lifeline, Lebanon), S.

Hay (SOS Détresse, Luxembourg), J. Jakobs (De Luisterlijn, Netherlands), F. Paulino (SOS Voz Amiga, Portugal), K. Bogataj (Zaupni telefon Samarijan, Slovenia), S. Basler (Die Dargebotene Hand, Switzerland), A. Goldstein, J. Higgins, S. Murphy and J. Draper, Vibrant Emotional Health (National Suicide Prevention Lifeline and Disaster Distress Helpline, USA). We thank C. Efferson, E. Fehr, L. Keller, K. Kölves, J. Vornberger and S. Métille for comments and suggestions. We are grateful to the Swiss National Science Foundation (NCCR LIVES—‘Overcoming Vulnerability: Life Course Perspectives’) for financial support.

Author information

Affiliations

1. University of Lausanne, Lausanne, Switzerland

Marius Brülhart & Rafael Lalive

2. Centre for Economic Policy Research, London, UK

Marius Brülhart & Rafael Lalive

3. University of Freiburg, Freiburg im Breisgau, Germany

Valentin Klotzbücher & Stephanie K. Reich

Contributions

S.K.R., V.K., M.B. and R.L. collected data. R.L., M.B., V.K. and S.K.R. designed the empirical approach. V.K., M.B. and R.L. implemented the analysis. M.B., V.K., R.L. and S.K.R. wrote the paper.

Corresponding author

Correspondence to [Marius Brülhart](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Lola Kola, Kairi Kolves 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 Evolution of daily helpline call volumes during the first wave.](#)

Sum of daily helpline contacts with seven-day moving average, January–June 2020 (black) and 2019 (light grey, not available for all helplines). Note that the vertical axes are truncated and not equal across panels, and the magnitudes of changes are thus not directly comparable. The solid red line shows the date of the pandemic outbreak, when more than 100 SARS-CoV-2 infections per 100,000 population have been recorded³⁹, the dashed blue line shows the date when shelter-in-place requirements were first introduced in the country of operation⁴⁰, see Supplementary Figs. 1–34 for details on individual helplines.

[Source data](#)

[Extended Data Fig. 2 Conversation topic shares by helpline.](#)

Each cell shows the share of calls related to the conversation topic on the horizontal axis, in percent of all calls with the helpline indicated on the vertical axis. Full dataset, covering all calls for which at least one topic was recorded, from 1 January 2019 to the respective end of available data, see Extended Data Table 1a and Supplementary Figs. 1–34.

[Source data](#)

[Extended Data Fig. 3 Magnitude of post-outbreak changes and equivalence tests.](#)

a, Coefficient estimates from linear probability models as in Fig. 2b, with 95% confidence intervals and equivalence bounds, defined as 5% of the pre-pandemic share of the respective topic, indicated by light blue vertical bars. **b**, Results from a

normalized across conversation topics, with coefficient estimates and associated 95% confidence intervals, and equivalence bounds divided by the pre-pandemic share of calls related to the respective topic. **c**, Relevance tests, numerical coefficient estimate with corresponding equivalence bounds, with test statistics and p-values from two one-sided tests for equivalence.

[Source data](#)

Extended Data Fig. 4 Change in conversation topics by caller sex and age group.

Estimated coefficients for interaction terms of group indicators with binary post-outbreak variable, and associated 95% confidence intervals. Separate linear probability regression models with dependent variable set to one for calls related to the respective topic, see [Methods](#), equation (3) and Supplementary Table 4.

[Source data](#)

Extended Data Fig. 5 Non-exclusive conversation topics.

Relation among conversation topics for calls included in the estimation sample underlying Fig. 2, from 1 January 2019 to 30 June 2020, and where sex and age group of callers are observed. **a**, Distribution of recorded number of conversation topics per call, **b**, Overlap in conversation topics, where each row shows the distribution of second or further topics (horizontal axis), in percent of all calls that are related to one specific topic (vertical axis), **c**, Results from Fig. 2b, with alternative estimates based on a restricted sample of single-topic category calls, see [Methods](#), equation (2). Supplementary Table 5 contains the numerical estimates.

[Source data](#)

Extended Data Fig. 6 Disaster Distress Helpline.

a, Sum of weekly calls routed to centers by year with 3-week moving average, letters on the horizontal axis indicate calendar months. **b**, deviation of log calls from time-averaged state-level mean (gray), with overall weekly average (black). **c**, Estimated coefficients and associated 95% confidence intervals; sub-national panel model including state and week fixed effects. Dependent variable is $\ln(\text{Disaster Distress calls} + 1)$, and independent variables are measured in logs as well; see [Methods](#), equation (5). Supplementary Table 10 contains the numerical estimates. **d**, Coefficient estimates for interaction terms with indicators for the two periods from January–August 2020

and September 2020–March 2021, and associated 95% confidence intervals; see [Methods](#), equation (6). Supplementary Table [11](#) contains the numerical estimates.

[Source data](#)

Extended Data Table 1 Overview over helpline data

Extended Data Table 2 Caller characteristics before and after the pandemic outbreak

Supplementary information

[Supplementary Information](#)

The Supplementary Information contains A, numerical estimation results underlying the main Figures and Extended Data Figures in Supplementary Tables, and B, Supplementary Notes with background information on individual helplines, as well as Supplementary Tables and Figures, organized alphabetically by country of operation.

[Reporting Summary](#)

[Peer Review File](#)

Source data

[Source Data Fig. 1](#)

[Source Data Fig. 2](#)

[Source Data Fig. 3](#)

[Source Data Fig. 4](#)

[Source Data Fig. 5](#)

[Source Data Extended Data Fig. 1](#)

[Source Data Extended Data Fig. 2](#)

[Source Data Extended Data Fig. 3](#)

Source Data Extended Data Fig. 4

Source Data Extended Data Fig. 5

Source Data Extended Data Fig. 6

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Brülhart, M., Klotzbücher, V., Lalive, R. *et al.* Mental health concerns during the COVID-19 pandemic as revealed by helpline calls. *Nature* **600**, 121–126 (2021). <https://doi.org/10.1038/s41586-021-04099-6>

- Received: 19 May 2021
- Accepted: 06 October 2021
- Published: 17 November 2021
- Issue Date: 02 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04099-6>

Share this article

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

[Get shareable link](#)

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

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

Further reading

- **Helpline data used to monitor population distress in a pandemic**

- Cindy H. Liu
- Alexander C. Tsai

Nature (2021)

Millions of helpline calls reveal how COVID affected mental health

- Heidi Ledford

News 17 Nov 2021

Helpline data used to monitor population distress in a pandemic

- Cindy H. Liu
- Alexander C. Tsai

News & Views 17 Nov 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04099-6>

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

- Article
- Open Access
- [Published: 25 October 2021](#)

Cryptic transmission of SARS-CoV-2 and the first COVID-19 wave

- [Jessica T. Davis](#)¹ ✉nal,
- [Matteo Chinazzi](#) [ORCID: orcid.org/0000-0002-5955-1929](#)¹ ✉nal,
- [Nicola Perra](#) [ORCID: orcid.org/0000-0002-5559-3064](#)^{1,2} ✉nal,
- [Kunpeng Mu](#)¹,
- [Ana Pastore y Piontti](#) [ORCID: orcid.org/0000-0002-2894-0489](#)¹,
- [Marco Ajelli](#) [ORCID: orcid.org/0000-0003-1753-4749](#)³,
- [Natalie E. Dean](#) [ORCID: orcid.org/0000-0003-3884-0921](#)⁴,
- [Corrado Gioannini](#)⁵,
- [Maria Litvinova](#) [ORCID: orcid.org/0000-0001-6393-1943](#)³,
- [Stefano Merler](#) [ORCID: orcid.org/0000-0002-5117-0611](#)⁶,
- [Luca Rossi](#)⁵,
- [Kaiyuan Sun](#)⁷,
- [Xinyue Xiong](#)¹,
- [Ira M. Longini Jr](#)⁸,
- [M. Elizabeth Halloran](#) [ORCID: orcid.org/0000-0002-3127-1757](#)^{9,10},
- [Cécile Viboud](#) [ORCID: orcid.org/0000-0003-3243-4711](#)⁷ &
- [Alessandro Vespignani](#) [ORCID: orcid.org/0000-0003-3419-4205](#)¹

[Nature](#) volume 600, pages 127–132 (2021)

- 79k Accesses
- 1 Citations
- 870 Altmetric
- [Metrics details](#)

Subjects

- [Computational biology and bioinformatics](#)
- [Mathematics and computing](#)
- [SARS-CoV-2](#)

Abstract

Considerable uncertainty surrounds the timeline of introductions and onsets of local transmission of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) globally^{1,2,3,4,5,6,7}. Although a limited number of SARS-CoV-2 introductions were reported in January and February 2020 (refs.^{8,9}), the narrowness of the initial testing criteria, combined with a slow growth in testing capacity and porous travel screening¹⁰, left many countries vulnerable to unmitigated, cryptic transmission. Here we use a global metapopulation epidemic model to provide a mechanistic understanding of the early dispersal of infections and the temporal windows of the introduction of SARS-CoV-2 and onset of local transmission in Europe and the USA. We find that community transmission of SARS-CoV-2 was likely to have been present in several areas of Europe and the USA by January 2020, and estimate that by early March, only 1 to 4 in 100 SARS-CoV-2 infections were detected by surveillance systems. The modelling results highlight international travel as the key driver of the introduction of SARS-CoV-2, with possible introductions and transmission events as early as December 2019 to January 2020. We find a heterogeneous geographic distribution of cumulative infection attack rates by 4 July 2020, ranging from 0.78% to 15.2% across US states and 0.19% to 13.2% in European countries. Our approach complements phylogenetic analyses and other surveillance approaches and provides insights that can be used to design innovative, model-driven surveillance systems that guide enhanced testing and response strategies.

[Download PDF](#)

Main

A few weeks after the initial announcement of a cluster of atypical pneumonia cases in Wuhan, China, the first confirmed cases of coronavirus disease 2019 (COVID-19) in the USA and Europe were detected (on 21 January 2020 in WA, USA¹ and on 24 January 2020 in France²). Although many more states and countries began to report initial introductions in the following weeks, only a few cases were detected daily during this time period (Fig. 1a), and most countries adopted a testing policy that targeted symptomatic individuals with a travel history linked to China. Several reports suggest that the introduction of SARS-CoV-2 occurred earlier than initially recognized^{3,4,5,6,7,8}, raising questions about the effectiveness of the initial testing policies and travel-related restrictions, as well as the extent to which the SARS-CoV-2

virus spread through cryptic transmission in January and February 2020. To address these questions, we use the global epidemic and mobility (GLEAM) model, a data-driven, stochastic, spatial, age-structured metapopulation epidemic model^{11,12}, to study the global dynamic underlying the evolution of the COVID-19 pandemic in Europe and the USA. Our model maps the plausible pathways of the pandemic using information available at the early stages of the outbreak and provides a global picture of the cryptic phase as well as the ensuing first wave of the COVID-19 pandemic.

Fig. 1: Early picture of the COVID-19 outbreak in Europe and the USA.

 figure1

a, Timelines of the daily reported and confirmed cases of COVID-19 in Europe (left) and the USA (right). BEL, Belgium; ESP, Spain; EU, European Union; FIN, Finland; FRA, France; GER, Germany; ITA, Italy; SWE, Sweden. **b**, Model-based estimates for the daily number of new infections in Europe (left) and the USA (right). The model estimates reported are the median values with the IQR obtained with an ABC calibration method using $n = 200,000$ independent model realizations. The inset plots compare the weekly incidence of reported cases with the median, weekly incidence of infections estimated by the model for the week of 8–14 March 2020 for the contiguous US states and European countries that reported at least one case (Europe, $n = 30$; USA, $n = 48$). Circle size corresponds to the population size of each state and country. The correlations were calculated using the Pearson correlation coefficient with a two-sided P value (Europe: $\rho = 0.80$, $P < 0.001$; USA: $\rho = 0.79$, $P < 0.001$). **c**, The probability

that a city in Europe (left) and the USA (right) had generated at least 100 cumulative infections by 21 February 2020. Colour and circle size are proportional to the probability.

Source data

We consider data concerning the continental USA and 30 European countries (the full list is reported in Extended Data Table 1). The model integrates real-time human mobility and population data with a mechanistic epidemic model at a global scale, incorporating changes in contact patterns and mobility according to the non-pharmaceutical interventions (NPIs) implemented in each region. It is calibrated on international case introductions out of mainland China at the early stage of the pandemic using an approximate Bayesian computation (ABC) methodology¹³. The model returns an ensemble of stochastic realizations of the global epidemic spread including international and domestic infection importations, incidence of infections and deaths at a daily resolution (see [Methods](#)). In the following text, we provide a detailed discussion of the analyses and results concerning European countries and the US states; however, to further test and validate our approach, in the [Supplementary Information](#), we report the modelling results for 24 additional countries that are globally representative, including countries of world regions such as Latin America, the Middle East, Africa, East Asia and Oceania.

In Fig. 1b we show the model estimates of the median daily incidence of new infections up to 21 February 2020, for both the USA and Europe. These values are much larger than the number of officially reported cases (see Fig. 1a), highlighting the substantial number of potential transmission events that may have already occurred before many states and countries had implemented testing strategies independent of travel history. As validation we compare our model's estimates of the number of infections during the week of 8 March 2020 to the number of cases reported during that week within each US state and European country with at least one reported case (shown in Fig. 1b inset). While we see a strong correlation between the reported cases and our model's estimated number of infections (Pearson's correlation coefficient on log values, USA: 0.79, $P < 0.001$; Europe: 0.80, $P < 0.001$), far fewer cases had actually been reported by that time. If we assume that the number of reported cases and simulated infections are related through a simple binomial sampling process, we find that on average 9 in 1,000 infections (90% confidence interval (CI) 1–35 per 1,000) and 35 in 1,000 infections (90% CI 4–90 per 1,000) were detected by 8 March 2020 in the USA and Europe, respectively. As testing capacity increased, the ascertainment rate grows and our estimates increase to detecting 17 in 1,000 infections (90% CI 2–55 per 1,000) by 14 March 2020 in the USA and 77 in 1,000 infections (90% CI 5–166 per 1,000) in Europe. The estimated ascertainment rates are in agreement with independent results based on different statistical methodologies^{14,15,16}.

In Fig. 1c we show the probability that a city in the USA or Europe had generated at least 100 infections by 21 February 2020. We see that the progression of the virus through the USA and Europe was both temporally and spatially heterogeneous. While many cities had not yet experienced much community transmission by late February, a few areas such as New York City and London are very likely to have already had local outbreaks.

Onset of local transmission

The model's ensemble of realizations provides a statistical description of all the potential pandemic histories compatible with the initial evolution of the pandemic in China. Rather than describing a specific, causal chain of events, we can estimate possible time windows pertaining to the initial chains of transmission in different geographical regions. We define the onset of local transmission for a country or state as the earliest date when at least 10 new infections are generated per day. This number is chosen because at this threshold the likelihood of stochastic extinction is extremely small^{17,18}. As detailed in the [Supplementary Information](#), further calibration on the US states and European countries suggests posterior values of R_0 ranging from 2.4 to 2.8. These values are consistent with many other (country-dependent) estimates^{19,20,21,22,23,24}. At the same time, given the doubling time of the number of COVID-19 cases before the implementation of public health measures, any variation of a factor 2 around the 10 infections per day threshold corresponds to a small adjustment of 3–5 days to the presented timelines.

In Fig. 2, we show the posterior probability distribution, $p(t)$, of the week, t , of the onset of local transmission for 15 US states (Fig. 2a) and European countries (Fig. 2b) (see [Supplementary Information](#) for all states and countries). We also calculate, for each country or state, the median date, T , that identifies the first week in which the cumulative distribution function is larger than 50%. Among the US states, CA and NY have the earliest dates, T , by the week of 19 January (CA) and 2 February (NY) 2020. In Europe, Italy, the UK, Germany and France are the first countries with T close to the end of January 2020. However, it is worth noting that each distribution, $p(t)$, has a support spanning several weeks. In Italy, the 5th and 95th percentiles of the $p(t)$ distribution are the week of 6 January and the week of 30 January 2020, respectively. These dates also suggest that it is not possible to rule out introductions and transmission events as early as December 2019, although the probability of this is very small.

Fig. 2: Timing of the onset of local transmission.

 **figure2**

a, b, Posterior distributions of the week in which each US state (**a**) or European country (**b**) first reached 10 locally generated SARS-CoV-2 transmission events per day. Countries and states are ordered by the median date of their posterior distribution. The week of this date corresponds to the dates reported on the vertical axis.

Source data

For each state in the USA and each country in Europe, we compared the order in which they surpassed 100 cumulative infections in the model and in the surveillance data (gathered from the John Hopkins University Coronavirus Resource Center²⁵). In Extended Data Fig. 1a we plot the ordering for states and compute the Kendall rank correlation coefficient τ (see [Supplementary Information](#)). The correlation is positive ($(\{\tau\}_{\{\rm EU\}})=0.71, P<0.001$ and $(\{\tau\}_{\{\rm US\}})=0.68, P<0.001$) indicating that, despite the detection and testing issues, the expected patterns of epidemic diffusion are largely described by the model in both regions.

SARS-CoV-2 introductions

As the model allows the recording of the origin and destination of travellers carrying SARS-CoV-2 at the global scale, we can study the possible sources of SARS-CoV-2 introductions for each US state and European country. More specifically, we record the cumulative number of introductions in each stochastic realization of the model until 30 April 2020. In Fig. 3 we visualize the origin of the introductions considering some key geographical regions (for example, Europe and Asia) while keeping the USA and China separate and aggregating all of the other countries (Others). For both the USA and Europe, the contribution from mainland China is barely visible and the local share (that is, sources within Europe and the USA) becomes significantly higher across the

board. Hence, while introduction events in the early phases of the outbreak were key to start local spreading (see details in the [Supplementary Information](#)), the cryptic transmission phase was sustained largely by internal flows. Domestic SARS-CoV-2 introductions to 30 April 2020 account for 69% (interquartile range (IQR) 60%–81%) of the introductions in CA, 78% (IQR 71%–87%) in TX and 69% (IQR 60%–80%) in MA, which is supported by phylogenetic analysis²⁶. European origins account for 69% (IQR 60%–80%), 84% (IQR 79%–91%) and 58% (IQR 48%–68%) of the introductions in Italy, Spain and the UK, respectively. In the [Supplementary Information](#), we report the full breakdown for all states and countries.

Fig. 3: Importation sources from the beginning of the outbreak until the end of April 2020.

 [figure3](#)

a, b, Each US state (**a**) and European country (**b**) is displayed in a clockwise order with respect to the start of the local outbreak (as seen in Fig. 2). Importation flows are directed and weighted. We normalize links considering the total in-flow for each state so that the sum of importation flows, for each state, is 1. In the Supplementary Information, we report the complete list of countries contributing as importation sources in each geographical region.

[Source data](#)

It is also necessary to distinguish between the full volume of SARS-CoV-2 introductions and the introduction events that could be relevant to the early onset of local transmission in each stochastic realization of the model. To this point, it is worth stressing that seeding introductions are different from the actual number of times the

virus has been introduced to each location with subsequent onward transmission. Even after a local outbreak has started, future importation events may give rise to additional onward transmission forming independently introduced transmission lineages of the virus²⁷. In the model, we can investigate seeding events by recording introduction events before the local transmission chains were established. We report the results of this analysis in the [Supplementary Information](#), showing that importations from mainland China may be relevant in seeding the epidemic in January, but then play a comparatively small role in the expansion in the number of COVID-19 cases in the USA and Europe owing to the travel restrictions imposed to/from mainland China after 23 January 2020.

The early timing of the initial introductions and diffusion pattern of SARS-CoV-2 were driven by air travel. We find a positive correlation ($\langle\langle \tau_{\text{EU}} \rangle\rangle = 0.66, P < 0.001$ and $\langle\langle \tau_{\text{US}} \rangle\rangle = 0.66, P < 0.001$) comparing the ordering of states according to when they surpassed 100 cumulative, reported cases (referred to as the epidemic order) and their domestic and international air travel volume rank (Extended Data Fig. 1B). Similar observations have been reported in China, where the initial spreading of the virus outside Hubei was strongly correlated with the traffic to/from the province²⁸. Other factors such as population size are also correlated with both the travel flows ($\langle\langle \tau_{\text{EU}} \rangle\rangle = 0.59, P < 0.001$ and $\langle\langle \tau_{\text{US}} \rangle\rangle = 0.7, P < 0.001$) and the epidemic order ($\langle\langle \tau_{\text{EU}} \rangle\rangle = 0.46, P < 0.001$ and $\langle\langle \tau_{\text{US}} \rangle\rangle = 0.68, P < 0.001$), which are discussed in detail in the [Supplementary Information](#). In our model, it is not possible to exclude increased contacts in highly populated places before social distancing interventions and disentangle this effect from increased seeding due to the correlation between travel volume and population size.

COVID-19 burden

Starting in March 2020, the establishment and timing of NPIs as well as other epidemiological drivers (that is, population size and density, age structure and so on) determined the disease burden in the USA and Europe^{29,30,31,32}. We account for these features by calibrating the model results, individually, for each US state and European country. More precisely, we estimate the posterior distribution of the infection fatality ratio (IFR) and infection attack rate in each US state and European country. To this end, we adopt the ABC approach using as evidence the number of new deaths reported from 22 March 2020 to 27 June 2020. We consider a uniform prior for the average IFR in the range from 0.4% to 2% that is age stratified proportional to the IFR values reported in ref. ³³. We also consider a uniform prior for reporting delays between the date of death and reporting ranging from 2 to 22 days in both Europe and the USA³⁴. Details are provided in the [Supplementary Information](#).

In Fig. 4a–d, f–i, we report the model fit of the estimated weekly deaths of the first wave for selected states and countries. Additional model results for all investigated regions including a sensitivity analysis of different calibration methods can be found in the [Supplementary Information](#). We find a strong correlation between the weekly model-estimated deaths and the reported values with a Pearson correlation coefficient of 0.99 ($P < 0.001$) for both Europe and the USA (see Supplementary Fig. 6). As the data suggest, many European countries and US states saw peaks in April and May with various decreasing trajectories that depend on the mitigation strategies in place. Additionally, we report the estimated posteriors for the cumulative infection attack rates and IFRs as of 4 July 2020 in European countries experiencing more than 100 total deaths and the top 20 states ranked by infection attack rate in the USA.

Fig. 4: The burden of the first wave in Europe and the USA.



a–d, Model fit of the estimated weekly deaths for selected countries in Europe (France, **a**; Italy, **b**; Sweden, **c**; UK, **d**). **e**, Posterior distributions of the infection attack rates and IFRs by 4 July 2020, for European countries where there were at least 100 reported deaths. **f–i**, Model fit of the estimated weekly deaths for selected states in the USA (CA, **f**; IL, **g**; MA, **h**; NY, **i**). **j**, Posterior distributions of the estimated infection attack rates and IFRs by 4 July 2020 for the top 20 US states (ranked according to their infection attack rates). The curves in **a–d** and **f–i** show the median values and 90% CIs. For **e** and **j**, the outer, lighter boxes represent the 90% CI, the darker, inner boxes represent the IQR, and the vertical lines represent the median value. Posterior distributions in **e** and **j** are the result of the ABC analysis of 200,000 independent model realizations.

Source data

Within Europe, Belgium has the highest estimated infection attack rate of 13.2% (90% CI [8.5%–28.3%]) by 4 July 2020, in agreement with the results in ref. ¹⁴. Furthermore, by that time Belgium reported the highest COVID-19 mortality rate out of the European countries investigated with 8.5 deaths per 10,000 individuals. However, Italy is estimated to have the highest median IFR of 1.4% (90% CI [0.6%–1.8%]), which aligns with other ranges reported in the literature^{35,36}. The US states with the highest infection attack rates are located within the northeast and experienced a significant first wave during March–April 2020. NY and NJ are the top two states with infection attack rates of 13.4% (90% CI [9.1%–26.7%]) and 15.2% (90% CI [10.2%–31.3%]), respectively. These numbers are aligned with estimates from New York City reported in ref. ³⁷. In the [Supplementary Information](#), we report summary tables with estimated IFRs, infection attack rates and the reproductive number in the absence of mitigation measures for all calibrated US states and European countries. Additionally, we compare our attack rate estimates to the prevalence of individuals with SARS-CoV-2 antibodies from serological studies across the USA and Europe (Extended Data Fig. ^{1D}). The seroprevalence estimates are compared to the model estimates during the same time window the studies were performed (details on the seroprevalence data from this figure can be found in Supplementary Table ⁸ and Supplementary Section 9.3).

Discussion

The model presented here captures the spatial and temporal heterogeneity of the early stage of the pandemic, going beyond the single-country-level reconstruction. It provides a mechanistic understanding of the underlying dynamics of the pandemic's interconnected evolution. Furthermore, rather than showing specific evidence for early infection in a few locations, our study aims at providing a statistical characterization and quantification of the initial transmission pathways at a global scale. Our results can be compared to and complement analyses based on gene sequencing and travel volumes. We find that 72% of the early introductions to Italy, before the local outbreak, are linked to China, which is in agreement with ref. ³⁸ highlighting the key role of importations between these regions at the beginning of the pandemic.

Additionally, similar to our findings, ref. ²⁷ estimates that the majority of importation events to April 2020, associated with onward transmission in the UK, came from Europe. The contributions from China are quantified below 1% and limited to the very early phase. Furthermore, seeding events from the USA are estimated to be below 3%, which aligns with our estimate (8%; IQR 3%–9%). However, their results point to a larger share from Europe (~90%) compared to ours (58%; IQR 48%–68%), and conversely, we estimate a larger contribution from Asia (27%; IQR 19%–35%). As

our analysis is a statistical description of the possible introduction pathways, differences could arise due to our model design, and also from genomic sampling biases³⁹.

The sources of introduction of SARS-CoV-2 infections in Europe and the USA changed substantially and rapidly through time. This caused reactive response strategies, such as issuing travel restrictions targeting countries only after local transmission is confirmed, ineffective at preventing local outbreaks. Our results suggest that many regions in the USA and Europe experienced an onset of local transmission in January and February 2020, during the time when testing capacity was limited. If testing had been more widespread and not restricted to individuals with a travel history from China, there would have been more opportunities for earlier detection and interventions. In the Supplementary Information, we report a counterfactual scenario where we assume broader testing specifications not based on the individual travel history and find that the epidemic progression is considerably delayed (see Supplementary Section 8).

As testing capacity increased and more cases were detected, many governments began to issue social distancing guidelines to mitigate the spread of SARS-CoV-2. The first European country to implement a cordon sanitaire was Italy on 23 February 2020, for a few northern cities⁴⁰. Many other countries followed suit and implemented national lockdowns in March 2020 (refs. [30,41](#)); however, this was weeks after our model estimates that SARS-CoV-2 was introduced and locally spreading. We find a strong correlation between the number of cases reported by the date of a lockdown/social distancing measure and the cumulative infections projected by 4 July 2020 (Extended Data Fig. 1C), indicating that the earlier NPIs had been issued, the smaller the COVID-19 burden experienced during the first wave. This is in agreement with other analyses showing that the timing of NPIs is crucial in limiting the burden of COVID-19 (refs. [19,29,42,43,44,45,46,47,48](#)). Overall, our results strengthen the case for preparedness plans with broader indication for testing that are able to detect local transmission earlier.

As with all modelling analyses, results are subject to biases from the limitations and assumptions within the model as well as the data used in its calibration. The model's parameters, such as generation time, incubation period and the proportion of asymptomatic infections, are chosen according to the current knowledge of SARS-CoV-2. Although the model is robust to variations in these parameters (see the Supplementary Information for the sensitivity analysis), more information on the key characteristics of the disease would considerably reduce uncertainties. The model calibration does not consider correlations among importations (that is, family travel) and assumes that travel probabilities are age specific across all individuals in the catchment area of each transportation hub.

In light of the assumptions and limitations inherent to this modelling approach, the results are able to complement the SARS-CoV-2 genome sequencing analyses used to reconstruct the early epidemic history of the COVID-19 pandemic³⁸. The methods used in this analysis offer a blueprint to identify the most likely early spreading dynamics of emerging viruses, and they can be used as a real-time risk assessment tool. Anticipating the locations where a virus is most likely to spread to next could be instrumental in guiding enhanced testing and surveillance activities. The estimated SARS-CoV-2 importation patterns and the cryptic transmission phase dynamics are of potential use when planning and developing public health policies in relation to international travelling, and they could provide important insights into assessing the potential risk and impact of emerging SARS-CoV-2 variants in regions of the world with limited testing and genomic surveillance resources.

Methods

The GLEAM model

The GLEAM model is a stochastic, spatial, age-structured metapopulation model. Previously this model was used to characterize the early stage of the COVID-19 epidemic in mainland China to estimate the effectiveness of travel bans and restrictions⁴⁹. The GLEAM model divides the global population into more than 3,200 subpopulations in roughly 200 different countries and territories interconnected by realistic air-travel and commuting mobility networks. A subpopulation is defined as the catchment area around major transportation hubs. The airline transportation data encompass daily travel data in the origin–destination format from the Official Aviation Guide database⁵⁰ reflecting actual traffic changes that occurred during the pandemic. Ground mobility and commuting flows are derived from the analysis and modelling of data collected from the statistics offices of 30 countries on 5 continents^{11,12}. The international travel data account for travel restrictions and government-issued policies. Furthermore, the model accounts for the reduction of internal, country-wide mobility and changes in contact patterns in each country and state in 2020. Specific model details are reported in the Supplementary Information.

SARS-CoV-2 transmission dynamics

The transmission dynamics take place within each subpopulation and assume a classic compartmentalization scheme for disease progression similar to those used in several large-scale models of SARS-CoV-2 transmission^{15,51,52,53,54,55}. Each individual, at any given point in time, is assigned to a compartment corresponding to their particular disease-related state (specifically, one could be susceptible, latent, infectious or removed)⁴⁹. This state also controls the individual's ability to travel (details in the

Supplementary Information). Individuals transition between compartments through stochastic chain binomial processes. Susceptible individuals can acquire the virus through contact with individuals in the infectious category and can subsequently become latent (that is, infected but not yet able to transmit the infection). The process of infection is modelled using age-stratified contact patterns at the state and country level^{56,57}. Latent individuals progress to the infectious stage at a rate inversely proportional to the latent period, and infectious individuals progress to the removed stage at a rate inversely proportional to the infectious period. The sum of the mean latent and infectious periods defines the generation time. Removed individuals are those who can no longer infect others. To estimate the number of deaths, we consider a uniformly distributed prior of the IFRs (ranging from 0.4% to 2%) that is age stratified proportional to the values estimated by ref. ³³ and incorporates reporting delays. The transmission model does not assume heterogeneities due to age differences in susceptibility to the SARS-CoV-2 infection for younger children (1–10 years old). This is an intense area of discussion^{58,59,60}. The transmission dynamic and the offspring distribution of infectious individuals in the model will depend on the specific details of each population, local and global mobility, NPIs and so on. While overdispersion in transmission varies by location in our model, we find that overall, it is consistent with 25% of primary infections causing 75% of transmission in our simulations (Supplementary Fig. 9). Additional simulations considering a fixed level of dispersion, informed by past studies, result in differences of less than 3 days in onset times (Supplementary Fig. 10; see also the Supplementary Information for further discussion).

Model calibration

We assume a start date of the epidemic in Wuhan, China, that falls between 15 November 2019 and 1 December 2019, with 20 initial infections^{49,51,61,62,63}. This considers that our model has a posterior distribution for the emergence of the outbreak in China that includes the possibility of transmission starting in October, 2019 (refs. ^{64,65}). The model generates an ensemble of possible epidemic realizations and is initially calibrated using an approximate Bayesian computation (ABC) rejection approach¹³ based on the observed international importations from mainland China up to 21 January 2020 (ref. ⁴⁹). Only a fraction of imported cases is generally detected at the destination^{10,66}. According to the estimates proposed in ref. ⁶⁷, we stratify the detection capacity of countries into three groups: high, medium, and low surveillance capacity according to the Global Health Security Index⁶⁸, and assume that asymptomatic infections are never detected. The model calibration does not consider correlated importations (for example, family travel) and assumes that travel probabilities are homogeneous across all individuals in each subpopulation. We further calibrate our model using the temporal ordering of the onset of local transmission (as

defined in the section ‘Onset of local transmission’) of the countries investigated. If we consider the epidemiological evidence^{38,69,70}, Italy was the first European country to experience substantial community transmission. Therefore, throughout the paper, we constrain the ensemble of simulations focusing only on stochastic realizations in which Italy is the first country, in the group under examination, to experience sustained local transmission (see [Supplementary Information](#) for details and further analyses of unconstrained simulations). Furthermore, we perform for each state and country an additional ABC rejection analysis using as evidence the weekly reported deaths in the time window starting on 22 March 2020 and ending on 27 June 2020. A full description of the model calibration is provided in the [Supplementary Information](#).

Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

Data availability

Epidemic surveillance data were collected from the Johns Hopkins Coronavirus Resource Center (<https://coronavirus.jhu.edu/>). Proprietary airline data are commercially available from the Official Aviation Guide (<https://www.oag.com/>) and International Air Transport Association (<https://www.iata.org/>) databases. Other model intervention data include data from Google’s COVID-19 Community Mobility Reports available at <https://www.google.com/covid19/mobility/> and the Oxford COVID-19 Response Tracker available at <https://github.com/OxCGR/COVID-policy-tracker>. [Source data](#) are provided with this paper.

Code availability

The GLEAM model is publicly available at <http://www.bleamviz.org/>. All data analyses of model results were performed using Python v3.8.

References

1. 1.

CDC *First Travel-related Case of 2019 Novel Coronavirus Detected in United States*. <https://www.cdc.gov/media/releases/2020/p0121-novel-coronavirus-travel-case.html> (21 January 2020).

2. 2.

Spiteri, G. et al. First cases of coronavirus disease 2019 (COVID-19) in the WHO European Region, 24 January to 21 February 2020. *Euro Surveill.* **25**, 2000178 (2020).

3. 3.

Althoff, K. N. et al. Antibodies to SARS-CoV-2 in All of Us Research Program Participants, January 2–March 18, 2020. *Clin. Infect. Dis.* <https://doi.org/10.1093/cid/ciab519> (2021).

4. 4.

La Rosa, G. et al. SARS-CoV-2 has been circulating in northern Italy since December 2019: evidence from environmental monitoring. *Sci. Total Environ.* **750**, 141711 (2021).

5. 5.

Deslandes, A. et al. SARS-CoV-2 was already spreading in France in late December 2019. *Int. J. Antimicrob. Agents* **55**, 106006 (2020).

6. 6.

Bedford, T. et al. Cryptic transmission of SARS-CoV-2 in Washington state. *Science* **370**, 571–575 (2020).

7. 7.

Basavaraju, S. V. et al. Serologic testing of US blood donations to identify severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)–reactive antibodies: December 2019–January 2020. *Clin. Infect. Dis.* **72**, e1004–e1009 (2020).

8. 8.

CDC COVID-19 Response Team et al. Evidence for limited early spread of COVID-19 within the United States, January–February 2020. *MMWR Morb. Mortal. Wkly Rep.* **69**, 680–684 (2020).

9. 9.

Lescure, F.-X. et al. Clinical and virological data of the first cases of COVID-19 in Europe: a case series. *Lancet Infect. Dis.* **20**, 697–706 (2020).

10. 10.

Gostic, K., Gomez, A. C., Mummah, R. O., Kucharski, A. J. & Lloyd-Smith, J. O. Estimated effectiveness of symptom and risk screening to prevent the spread of COVID-19. *eLife* **9**, e55570 (2020).

11. 11.

Balcan, D. et al. Multiscale mobility networks and the spatial spreading of infectious diseases. *Proc. Natl Acad. Sci. USA* **106**, 21484–21489 (2009).

12. 12.

Balcan, D. et al. Modeling the spatial spread of infectious diseases: The GLoBal Epidemic and Mobility computational model. *J. Comput. Sci.* **1**, 132–145 (2010).

13. 13.

Sunnåker, M. et al. Approximate Bayesian computation. *PLoS Comput. Biol.* **9**, e1002803 (2013).

14. 14.

Russell, T. W. et al. Reconstructing the early global dynamics of under-ascertained COVID-19 cases and infections. *BMC Med.* **18**, 332 (2020).

15. 15.

Gatto, M. et al. Spread and dynamics of the COVID-19 epidemic in Italy: effects of emergency containment measures. *Proc. Natl Acad. Sci. USA* **117**, 10484–10491 (2020).

16. 16.

Havers, F. P. et al. Seroprevalence of antibodies to SARS-CoV-2 in 10 sites in the United States, March 23-May 12, 2020. *JAMA Intern. Med.* **180**, 1576–1586 (2020).

17. 17.

Bailey, N. T. et al. *The Mathematical Theory of Infectious Diseases and Its Applications* (Charles Griffin, 1975).

18. 18.

Colizza, V. & Vespignani, A. Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: theory and simulations. *J. Theor. Biol.* **251**, 450–467 (2008).

19. 19.

Perra, N. Non-pharmaceutical interventions during the COVID-19 pandemic: a review. *Phys. Rep.* <https://doi.org/10.1016/j.physrep.2021.02.001> (2021).

20. 20.

Salje, H. et al. Estimating the burden of SARS-CoV-2 in France. *Science* **369**, 208–211 (2020).

21. 21.

Domenico, L. D., Pullano, G., Sabbatini, C. E., Boëlle, P.-Y. & Colizza, V. Impact of lockdown on COVID-19 epidemic in Île-de-France and possible exit strategies. *BMC Med.* <https://doi.org/10.1186/s12916-020-01698-4> (2020).

22. 22.

Chang, S. et al. Mobility network models of COVID-19 explain inequities and inform reopening. *Nature* <https://doi.org/10.1038/s41586-020-2923-3> (2020).

23. 23.

Pei, S., Kandula, S. & Shaman, J. Differential effects of intervention timing on COVID-19 spread in the United States. *Sci. Adv.* **6**, eabd6370 (2020).

24. 24.

Lau, M. S. Y. et al. Characterizing superspreading events and age-specific infectiousness of SARS-CoV-2 transmission in Georgia, USA. *Proc. Natl Acad. Sci. USA* **117**, 22430–22435 (2020).

25. 25.

Dong, E., Du, H. & Gardner, L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* **20**, 533–534 (2020).

26. 26.

Fauver, J. R. et al. Coast-to-coast spread of SARS-CoV-2 during the early epidemic in the United States. *Cell* **181**, 990–996 (2020).

27. 27.

du Plessis, L. et al. Establishment and lineage dynamics of the SARS-CoV-2 epidemic in the UK. *Science* **371**, 708–712 (2021).

28. 28.

Kraemer, M. U. et al. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* **368**, 493–497 (2020).

29. 29.

White, E. R. & Hébert-Dufresne, L. State-level variation of initial COVID-19 dynamics in the United States. *PLoS ONE* **15**, e0240648 (2020).

30. 30.

Desvars-Larrive, A. et al. A structured open dataset of government interventions in response to COVID-19. *Sci. Data* **7**, 285 (2020).

31. 31.

Althouse, B. M. et al. The unintended consequences of inconsistent pandemic control policies. Preprint at *medRxiv* <https://doi.org/10.1101/2020.08.21.20179473> (2020).

32. 32.

Rader, B. et al. Crowding and the shape of COVID-19 epidemics. *Nat. Med.* **26**, 1829–1834 (2020).

33. 33.

Verity, R. et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. *Lancet Infect. Dis.* **20**, 669–677 (2020).

34. 34.

CDC & ASPR. COVID-19 pandemic planning scenarios. *CDC* <https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html> (2021).

35. 35.

O'Driscoll, M. et al. Age-specific mortality and immunity patterns of SARS-CoV-2. *Nature* **590**, 140–145 (2021).

36. 36.

Poletti, P. et al. Infection fatality ratio of SARS-CoV-2 in Italy. *Euro Surveill.* **25**, 2001381 (2020).

37. 37.

Kissler, S. et al. Reductions in commuting mobility correlate with geographic differences in SARS-CoV-2 prevalence in New York City. *Nat. Commun.* **11**, 4674 (2020).

38. 38.

Lemey, P. et al. Accommodating individual travel history and unsampled diversity in Bayesian phylogeographic inference of SARS-CoV-2. *Nat. Commun.* **11**, 51120 (2020).

39. 39.

Martin, M. A., Van Insberghe, D. & Koelle, K. Insights from SARS-CoV-2 sequences. *Science*. **371**, 466–467 (2021).

40. 40.

Lazzerini, M. & Putoto, G. COVID-19 in Italy: momentous decisions and many uncertainties. *Lancet Glob. Health* **8**, e641–e642 (2020).

41. 41.

Cheng, C., Barceló, J., Hartnett, A. S., Kubinec, R. & Messerschmidt, L. Covid-19 government response event dataset (coronanet v.1.0). *Nat. Hum. Behav.* **4**, 756–768 (2020).

42. 42.

Yang, W., Shaff, J. & Shaman, J. Effectiveness of non-pharmaceutical interventions to contain COVID-19: a case study of the 2020 spring pandemic wave in New York City. *J. R. Soc. Interface* **18**, 2020082.

43. 43.

Ali, S. T. et al. Serial interval of SARS-CoV-2 was shortened over time by nonpharmaceutical interventions. *Science* **369**, 1106–1109 (2020).

44. 44.

Pan, A. et al. Association of public health interventions with the epidemiology of the COVID-19 outbreak in Wuhan, China. *JAMA* **323**, 1915–1923 (2020).

45. 45.

Jefferies, S. et al. COVID-19 in New Zealand and the impact of the national response: a descriptive epidemiological study. *Lancet Public Health* **5**, e612–e623 (2020).

46. 46.

Auger, K. A. et al. Association between statewide school closure and COVID-19 incidence and mortality in the US. *JAMA* **324**, 859–870 (2020).

47. 47.

Islam, N. et al. Physical distancing interventions and incidence of coronavirus disease 2019: natural experiment in 149 countries. *BMJ* (2020).

48. 48.

Haug, N. et al. Ranking the effectiveness of worldwide covid-19 government interventions. *Nat. Hum. Behav.* **4**, 1303–1312 (2020).

49. 49.

Chinazzi, M. et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* **368**, 395–400 (2020).

50. 50.

Official Aviation Guide (OAG); <https://www.oag.com> (2020).

51. 51.

Imai, N. et al. *Report 3: Transmissibility of 2019-nCoV*.
<https://doi.org/10.25561/77148> (Imperial College London, 2020).

52. 52.
Kissler, S. M., Tedijanto, C., Goldstein, E., Grad, Y. H. & Lipsitch, M. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science* **368**, 860–868 (2020).
53. 53.
Li, R. et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* **368**, 489–493 (2020).
54. 54.
Wu, J. T., Leung, K. & Leung, G. M. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *Lancet.* **395**, 10225 (2020).
55. 55.
Lai, S. et al. Effect of non-pharmaceutical interventions to contain COVID-19 in China. *Nature*. **585**, 410–413 (2020).
56. 56.
Mistry, D. et al. Inferring high-resolution human mixing patterns for disease modeling. *Nature Commun.* **12**, 323 (2021).
57. 57.
Prem, K., Cook, A. R. & Jit, M. Projecting social contact matrices in 152 countries using contact surveys and demographic data. *PLoS Comput. Biol.* **13**, e1005697 (2017).
58. 58.
Zhang, J. et al. Changes in contact patterns shape the dynamics of the COVID-19 outbreak in China. *Science* **368**, 1481–1486 (2020).
59. 59.
Davies, N. G. et al. Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nat. Med.* **26**, 1205–1211 (2020).

60. 60.
Bi, Q. et al. Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study. *Lancet Infect. Dis.* **20**, 911–919 (2020).
61. 61.
Rambaut, A. Preliminary phylogenetic analysis of 11 nCoV2019 genomes, 2020-01-19. *Virological* <http://virological.org/t/preliminary-phylogenetic-analysis-of-11-ncov2019-genomes-2020-01-19/329> (2020).
62. 62.
Anderson, K. Clock and TMRCA based on 27 genomes. *Virological* <http://virological.org/t/clock-and-tmrca-based-on-27-genomes/347> (2020).
63. 63.
Bedford, T. et al. Genomic analysis of nCoV spread. Situation report 2020-01-23. *Nextstrain* <https://nextstrain.org/narratives/ncov/sit-rep/2020-01-23> (2020).
64. 64.
Pekar, J., Worobey, M., Moshiri, N., Scheffler, K. & Wertheim, J. O. Timing the SARS-CoV-2 index case in Hubei province. *Science*. **372**, 412–417 (2021).
65. 65.
van Dorp, L. et al. Emergence of genomic diversity and recurrent mutations in SARS-CoV-2. *Infect. Genet. Evol.* **83**, 104351 (2020).
66. 66.
De Salazar, P. M., Niehus, R., Taylor, A., Buckee, C. & Lipsitch, M. Identifying locations with possible undetected imported severe acute respiratory syndrome coronavirus 2 cases by using importation predictions. *Emerg. Infect. Dis.* **26**, 1465–1469 (2020).
67. 67.
Niehus, R., De Salazar, P. M., Taylor, A. R. & Lipsitch, M. Using observational data to quantify bias of traveller-derived COVID-19 prevalence estimates in Wuhan, China. *Lancet Infect. Dis.* **20**, 803–808 (2020).

68. 68.

Global Health Security Index (GHS); <https://www.ghsindex.org> (2020).

69. 69.

Nadeau, S. A., Vaughan, T. G., Scire, J., Huisman, J. S. & Stadler, T. The origin and early spread of SARS-CoV-2 in Europe. *Proc. Natl Acad. Sci. USA* **118**, e2012008118 (2021).

70. 70.

Worobey, M. et al. The emergence of SARS-CoV-2 in Europe and North America. *Science* **370**, 6516 (2020).

71. 71.

Kontis, V. et al. Magnitude, demographics, and dynamics of the effect of the first wave of the COVID-19 pandemic on all-cause mortality in 21 industrialized countries. *Nat. Med.* **26**, 1919–1928 (2020).

72. 72.

Managan, D. Trump issues ‘Coronavirus Guidelines’ for next 15 days to slow pandemic. *CNBC* <https://www.cnbc.com/2020/03/16/trumps-coronavirus-guidelines-for-next-15-days-to-slow-pandemic.html> (16 March 2020).

Acknowledgements

A.V., M.E.H., N.E.D. and I.M.L. acknowledge support from the award NIH-R56AI148284. S.M. acknowledges support from the EU H2020 MOOD project. C.G. and L.R. acknowledge support from the EU H2020 Icarus project. M.A., M.C. and A.V. acknowledge support from the COVID Supplement CDC-HHS-6U01IP001137-01. M.C. and A.V. acknowledge support from the Google Cloud Research Credits programme to fund this project. A.V. acknowledges support from the McGovern Foundation and the Chleck Foundation. The findings and conclusions in this study are those of the authors and do not necessarily represent the official position of the funding agencies, the National Institutes of Health or the US Department of Health and Human Services.

Author information

Author notes

1. These authors contributed equally: Jessica T. Davis, Matteo Chinazzi, Nicola Perra

Affiliations

1. Laboratory for the Modeling of Biological and Socio-technical Systems,
Northeastern University, Boston, MA, USA

Jessica T. Davis, Matteo Chinazzi, Nicola Perra, Kunpeng Mu, Ana Pastore y Piontti, Xinyue Xiong & Alessandro Vespignani

2. Networks and Urban Systems Centre, University of Greenwich, London, UK

Nicola Perra

3. Department of Epidemiology and Biostatistics, Indiana University School of Public Health, Bloomington, IN, USA

Marco Ajelli & Maria Litvinova

4. Department of Biostatistics and Bioinformatics, Emory University, Atlanta, GA, USA

Natalie E. Dean

5. ISI Foundation, Turin, Italy

Corrado Gioannini & Luca Rossi

6. Bruno Kessler Foundation, Trento, Italy

Stefano Merler

7. Division of International Epidemiology and Population Studies, Fogarty International Center, National Institutes of Health, Bethesda, MD, USA

Kaiyuan Sun & Cécile Viboud

8. Department of Biostatistics, College of Public Health and Health Professions, University of Florida, Gainesville, FL, USA

Ira M. Longini Jr

9. Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA

M. Elizabeth Halloran

10. Department of Biostatistics, University of Washington, Seattle, WA, USA

M. Elizabeth Halloran

Contributions

J.T.D., M.C., N.P. and A.V. designed research; M.C., J.T.D., N.P., M.A., C.G., M.L., S.M., A.P.P., K.M., L.R., K.S., C.V., X.X., M.E.H., I.M.L. and A.V. performed research; M.C., J.T.D., N.P., A.P.P., K.M. and A.V. analysed data; and M.C., J.T.D., N.P., M.A., C.G., M.L., S.M., A.P.P., K.M., N.E.D., L.R., K.S., C.V., X.X., M.E.H., I.M.L. and A.V. wrote and edited the paper.

Corresponding author

Correspondence to [Alessandro Vespignani](#).

Ethics declarations

Competing interests

M.A. reports research funding from Seqirus, not related to COVID-19. A.V., M.C. and A.P.P. report grants from Metabiota Inc., outside the submitted work. The authors declare no other relationships or activities that could appear to have influenced the submitted work.

Additional information

Peer review information *Nature* thanks Steven Riley 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 Correlation Analysis for European countries and US states.

(a) The correlation between the ordering of each country/state to reach 100 infections in the model-estimates and to reach 100 reported cases in the surveillance data (Europe: n = 23, US: n = 49). (b) The correlation between the ordering of each country/state considering the time needed to reach 100 reported cases in the surveillance data and the ranking of the combined international and domestic air traffic (Europe n = 23, US n = 49). Correlations in (a, b) are computed considering the Kendall rank correlation coefficient reported with a two-sided p-value, we consider European countries that reached at least 100 reported deaths by July 4, 2020 and countries in Scandinavia (c) Left: the correlation between the number of cases reported by the date of lockdown for European countries (from Table 4 in Ref. [71](#)) and the estimated total number of infections by July 4, 2020 (median values, n = 15). Right: the correlation between the number of cases reported by March 16, 2020 (the date the “15 days to slow the spread” guidelines were released in the US Ref. [72](#)) for each US state and the estimated total infections by July 4, 2020 (median values, n = 36). We consider states that reached at least 100 reported deaths by July 4, 2020. The circle sizes in (a–c) correspond to the population sizes of each country/state. (d) The correlation between the model-estimated infection attack rate and the serological prevalence collected from studies, n = 20. Estimated attack rates are the posterior distributions that are the result of the ABC analysis of 200,000 independent model realizations. Data points refer to different dates and the locations for which serological surveys were available (see table S8 in SI for study descriptions). The model-estimated attack rates use the median value, and the error bars represent the 90%CI. The correlations are calculated using the Pearson correlation coefficient in (c, d) reported with a two-sided p-value.

Extended Data Table 1 Regions under investigation

Supplementary information

Supplementary Information

This file contains Supplementary Sections 1–9, Figs. 1–20, Tables 1–8 and References.

Reporting Summary

Peer Review File

Source data

[Source Data Fig. 1](#)

[Source Data Fig. 2](#)

[Source Data Fig. 3](#)

[Source Data Fig. 4](#)

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

Davis, J.T., Chinazzi, M., Perra, N. *et al.* Cryptic transmission of SARS-CoV-2 and the first COVID-19 wave. *Nature* **600**, 127–132 (2021). <https://doi.org/10.1038/s41586-021-04130-w>

- Received: 24 March 2021
- Accepted: 13 October 2021
- Published: 25 October 2021
- Issue Date: 02 December 2021

- DOI: <https://doi.org/10.1038/s41586-021-04130-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

Further reading

- [**A reconstruction of early cryptic COVID spread**](#)

- Simon Cauchemez
- Paolo Bosetti

Nature (2021)

[**A reconstruction of early cryptic COVID spread**](#)

- Simon Cauchemez
- Paolo Bosetti

News & Views 01 Nov 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04130-w>

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

- Article
- Open Access
- [Published: 17 November 2021](#)

Independent infections of porcine deltacoronavirus among Haitian children

- [John A. Lednicky](#)^{1,2} na1,
- [Massimiliano S. Tagliamonte](#) ORCID: orcid.org/0000-0003-2699-7331^{1,3} na1,
- [Sarah K. White](#)^{1,2},
- [Maha A. Elbadry](#)^{1,2},
- [Md. Mahbubul Alam](#)^{1,2},
- [Caroline J. Stephenson](#)^{1,2},
- [Tania S. Bonny](#)^{1,2},
- [Julia C. Loeb](#)^{1,2},
- [Taina Telisma](#)⁴,
- [Sonese Chavannes](#)⁴,
- [David A. Ostrov](#) ORCID: orcid.org/0000-0002-4696-875X^{1,3},
- [Carla Mavian](#) ORCID: orcid.org/0000-0003-3819-5520^{1,3},
- [Valery Madsen Beau De Rochars](#)^{1,5},
- [Marco Salemi](#) ORCID: orcid.org/0000-0003-0136-2102^{1,3} &
- [J. Glenn Morris Jr](#) ORCID: orcid.org/0000-0001-9347-1984^{1,6}

Nature volume 600, pages 133–137 (2021)

- 10k Accesses
- 70 Altmetric
- [Metrics details](#)

Subjects

- [Computational biology and bioinformatics](#)
- [SARS-CoV-2](#)

Abstract

Coronaviruses have caused three major epidemics since 2003, including the ongoing SARS-CoV-2 pandemic. In each case, the emergence of coronavirus in our species has been associated with zoonotic transmissions from animal reservoirs^{1,2}, underscoring how prone such pathogens are to spill over and adapt to new species. Among the four recognized genera of the family *Coronaviridae*, human infections reported so far have been limited to alphacoronaviruses and betacoronaviruses^{3,4,5}. Here we identify porcine deltacoronavirus strains in plasma samples of three Haitian children with acute undifferentiated febrile illness. Genomic and evolutionary analyses reveal that human infections were the result of at least two independent zoonoses of distinct viral lineages that acquired the same mutational signature in the genes encoding Nsp15 and the spike glycoprotein. In particular, structural analysis predicts that one of the changes in the spike S1 subunit, which contains the receptor-binding domain, may affect the flexibility of the protein and its binding to the host cell receptor. Our findings highlight the potential for evolutionary change and adaptation leading to human infections by coronaviruses outside of the previously recognized human-associated coronavirus groups, particularly in settings where there may be close human–animal contact.

[Download PDF](#)

Main

Coronaviruses are enveloped, positive-sense single-stranded RNA viruses that belong to the family *Coronaviridae*. Porcine deltacoronavirus (PDCoV) is a member of the genus *Deltacoronavirus*, and was reported for the first time in Hong Kong, China, in 2012 (refs. [6,7](#)). PDCoV causes gastrointestinal symptoms in piglets, with dehydration and possibly death⁸. The jejunum and ileum are the primary sites of infection^{[8,9,10,11](#)}. There are reports of symptomatic infection, in experimental settings, in chickens, turkeys and calves^{[12,13](#)}. Human cells have also been reported to be permissive to PDCoV infection^{[14](#)}. Binding of the PDCoV spike glycoprotein to an interspecies conserved site, the host aminopeptidase N^{[14,15,16,17](#)}, may facilitate direct transmission to non-reservoir species, possibly including humans.

Haiti is part of the island of Hispaniola, and one of the poorest countries in the world^{[18](#)}. The local pig population was wiped out in the 1980s to eliminate African swine fever from the area^{[19](#)}, followed by subsequent and ongoing repopulation from North America, Europe and China^{[19,20](#)}. Pig farming in the country is at a subsistence level, and, to our knowledge, no cases of PDCoV infection have been reported to date in pigs. Between 2012 and 2020, our group monitored the occurrence of illness among

children seen at a free school clinic operated by the Christianville Foundation school system in the Gressier region of Haiti^{21,22,23,24}. Respiratory and diarrhoeal illnesses were most common among children presenting to the clinic for care, followed by acute undifferentiated febrile illnesses (that is, fever with no clear localizing symptoms), which accounted for approximately 16% of clinic cases²¹.

Detection of Hu-PDCoV

As part of ongoing studies at the Christianville school clinic, we collected plasma samples from 369 children with acute undifferentiated febrile illness seen at the clinic between May 2014 and December 2015. As previously reported²⁵, all children from whom samples were collected were screened for common pathogens associated with fever, including malaria and viral pathogens such as dengue, Zika and Chikungunya viruses. Plasma samples that were negative in these assays were then cultured in Vero E6 cells, as a ‘non-biased’ means of identifying potential new or emerging viruses. Cultures of three samples (0.8% of the 369 samples collected) were positive for coronavirus strains, which clustered with PDCoV. Information on the children from whom these three samples were obtained is provided in Table 1.

Table 1 Sample collection date and epidemiology of children with PDCoV infection

Cases 1 and 3 (samples 0081-4 and 0329-4, respectively) were in-patients from the main campus of the school (school A) that is attended by students from semi-urban areas. This school provides classes for grades K–13 (ages 5–20); the socioeconomic status of families of students ranges widely, with a subset of families coming from low-income families and receiving tuition support. Case 2 (sample 0256-1) was from a different campus (school B), which is an elementary school located in the mountains approximately 1-h drive from school A; the school is in a rural area, with students from very low socioeconomic backgrounds. All three children presented with a history of fever but recovered uneventfully: child 2 was febrile (40 °C) when seen in the clinic; child 1 and child 3 reported cough and abdominal pain. Although they reported a fever, child 3 did not have acute symptoms when seen in the clinic.

Nucleic acids purified as previously described^{26,27} from the plasma samples of the three children tested negative for alphavirus and flavivirus RNAs^{26,27} by PCR with reverse transcription (RT–PCR). Virus isolation was also attempted after inoculation of aliquots of the plasma onto Vero E6 cells^{26,27}. Nucleic acids, purified from the cell culture medium 7, 14, 21 and 30 days post-inoculation of the cells, again tested negative for alphavirus and flavivirus RNAs. Moreover, they tested negative for the DNA and RNA of common human respiratory viruses using a GenMark Respiratory Panel²⁸. However, subtle cytopathic effects (CPEs) were observed in Vero E6 cell

monolayers starting at about 11 days post-inoculation, suggesting that a virus had been isolated. The non-specific CPEs included granulation of the cells (Fig. 1).

Fig. 1: Non-specific CPEs formed by plasma from patient 0081-4 in Vero E6 cells.

 figure1

a, Mock-infected Vero E6 cells, 11 days post-inoculation with phosphate-buffered saline. **b**, Vero E6 cells 11 days post-inoculation with plasma from patient 0081-4. Original magnification, $\times 200$.

Because none of the tests produced evidence that could be used for a preliminary identification of a viral agent, an unbiased amplification and sequencing approach²⁹ was attempted for cells inoculated with plasma from sample 0081-4, which displayed more CPEs than cells inoculated with the other two plasma samples. PCR amplification yielded seven amplicons. Sequence analyses indicated that six amplicons were African green monkey sequences from the Vero E6 cells, whereas one 401-bp amplicon had 100% identity with the genome sequence of various PDCoV strains. Therefore, RNA purified from Vero E6 culture samples was retested using a pan-coronavirus RT-PCR test that amplifies a conserved 668-bp region within the RNA-dependent RNA polymerase gene of alphacoronaviruses, betacoronaviruses gammacoronaviruses and deltacoronaviruses³⁰, generating positive results. A 3-ml sample of cell culture medium collected from Vero E6 cells that displayed CPEs 11 days post-inoculation was subsequently screened by transmission electron microscopy. Rare spherical and pleiomorphic coronavirus particles that ranged from 90 to 120 nm in diameter were visualized with negative staining. A representative image of a 110-nm spherical coronavirus particle is shown in Extended Data Fig. 1.

A follow-up test using RNA directly purified from plasma generated the same 668-bp amplicons, providing further indications that PDCoV was present in the plasma and cell culture samples. Mock-inoculated Vero E6 cultures and 20 additional randomly selected plasma samples from the study tested negative for PDCoV RNAs. At the time, we had no PDCoV strains in our laboratory. Following this preliminary identification, whole-genome sequences for the three isolates were obtained by Sanger

sequencing. The GenBank accession numbers corresponding to the sequenced genomes are provided in Table 1.

Genomic analyses of Hu-PDCoV strains

In agreement with previous reports³¹, assessment of potential recombinants in the multiple sequence alignment including all currently available full-genome PDCoV sequences detected a signal for recombination. Recombination events involved 60 strains from pigs in China belonging to sequence clusters unrelated to the new human isolates. After all recombinant strains were removed, the pairwise homoplasy index (PHI)⁴⁷ test for recombination using the alignment of the remaining full genomes, as well as the recombinant fragments identified in the sequences from China, did not show any recombination signal ($P > 0.05$). Moreover, NeighborNets inferred from full genomes, as well as recombinant fragments, showed the human sequences from Haitian individuals consistently clustering with non-recombinant porcine strains of Chinese (child 1 and child 3) or US (child 2) origin (Extended Data Fig. 2). The human PDCoV (Hu-PDCoV) strains 0081-4 and 0329-4, identified 4 months apart in child 1 and child 3, respectively, while attending school A (Table 1), were highly similar (99.97%), and closely related (99.8%) to a pig strain detected 1 year later in Tianjin, China. Child 2, who attended school B, was infected with a variant, 0256-1, closely related to a pig strain detected in Arkansas, USA, in 2015.

Since recombination did not affect the new Hu-PDCoV strains nor their closest evolutionary relatives, we inferred a maximum likelihood tree from all full-genome PDCoV sequences available to investigate in more detail the origin of the human isolates. The tree clearly shows that strains 0081-4 and 0329-4, and strain 0256-1 belong, respectively, to two distinct and well-supported monophyletic clades: the first clade clustering strains from pigs in China and the second clade clustering strains from pigs in the USA (Extended Data Fig. 3). It is important to emphasize that while phylogenetic relationships in the deep branches of the maximum likelihood tree cannot be considered an accurate depiction of the evolutionary relationships among major PDCoV clades because of recombination, clustering within the two clades including the strains detected in Haiti is not affected by recombinant events (Extended Data Fig. 2). The three PDCoV sequences identified in the Haitian children were the result of at least two separate zoonotic transmissions from related non-recombinant porcine strains that probably occurred within a similar time frame. There are two possible scenarios that could explain why two Haitian children were infected with genetically similar PDCoV strains. There could have been two independent zoonoses from animals infected with highly genetically similar viruses. Alternatively, there could have been one initial zoonosis followed by human-to-human transmission. Since samples from pigs in the areas surrounding the two schools were not available, it is impossible at this time to discern which scenario is the most likely. Regardless, the

phylogeny demonstrates the occurrence of two distinct PDCoV lineages in school A and school B, highlighting the ability of deltacoronaviruses to spill over successfully in the human population.

Our next step was the calibration of a molecular clock to infer the time of the most recent ancestor (TMRCA) of Hu-PDCoV and their most closely related porcine strains. We tested for the presence of a temporal signal in the sequence dataset by calculating the linear regression between root-to-tip distances and sampling time in the maximum likelihood tree. After removal of sparrow outgroup sequences and the southeast Asian clade outliers, the tree inferred from the remaining ($n = 94$) sequences showed sufficient signal to calibrate a molecular clock (Extended Data Fig. 4). The topology of the Bayesian maximum clade credibility tree obtained using a strict molecular clock confirmed the findings of the maximum likelihood phylogeny (Fig. 2, Extended Data Fig. 5). Identical results were obtained with the relaxed clock model. The mean evolutionary rate estimated with the strict clock resulted in 7.3×10^{-4} nucleotide substitutions per site per year, with a 95% high posterior density (HPD) interval of $5\text{--}9 \times 10^{-4}$, which is slightly higher than previous estimates but with overlapping confidence intervals³². According to the clock calibration, 0081-4 and 0329-4 TMRCA dates to October 2014, with 95% HPD intervals essentially overlapping (October 2014 to January 2015) with the sampling dates (see Table 1) of the strains themselves. In turn, the strains detected in Haiti diverged from their MRCA with the pig strain detected in China in July 2014 (95% HPD: April to August 2014). Conversely, the isolate 0256-1 TMRCA discovered in Haiti and the isolate KR150443 detected in the USA date to 2011 (95% HPD: February 2011 to March 2012). It is possible that PDCoV strains had been circulating in pigs in Haiti for a few years, as also suggested by its relatively long terminal branch in the maximum likelihood tree (Extended Data Fig. 3), before infecting the human patient and that we are missing several intermediate links along the 0256-1 branch, either from pig or other human strains. Unfortunately, given the paucity of detailed information of recent livestock importations into Haiti as well as the lack of a surveillance system to monitor viral infections in farm animals, a detailed reconstruction of the events surrounding introduction of PDCoV into Haiti and its subsequent introduction into humans is currently not possible.

Fig. 2: Bayesian maximum clade credibility tree of PDCoV strains.

 **figure2**

The circles at the internal nodes indicate high posterior probability support of more than 0.9. The branch lengths were scaled in time by using a strict molecular clock. The information above the tree shows the inferred TMRCA between strains discovered in Haiti and their closest phylogenetic relative, with 95% HPD intervals.

Although the Hu-PDCoV strains belong to independent evolutionary lineages, introduced in humans through what would appear to be at least two separate zoonotic transmissions, a more in-depth analysis of the genomic changes shows that the three strains detected in Haiti share a signature of five conserved amino acid residues in the ORF1a/b polyprotein and two in the spike glycoprotein, unique among other currently known PDCoV sequences from pigs (Fig. 3a). The sole exception is the strain KY065120 discovered in China, which displays the same amino acid signature and is the one most closely related to the strains discovered in Haiti from school A (Table 1, Extended Data Fig. 3) and may represent a porcine strain pre-adapted for effective transmission to humans. Indeed, the convergent evolution of identical amino acid changes along distinct phylogenetic lineages is highly suggestive of an adaptive response. Mutations in the first five ORF1a/b amino acids that are part of the Hu-PDCoV-specific signature (Fig. 3a) are located at sites that do not correspond to solved crystal structures. The other ORF1a/b mutation maps in non-structural protein 15 (Nsp15): A30V (the amino acid position is numbered according to the reference sequence JQ065043). The carboxy-terminal domain of the Nsp15 protein possesses endoribonuclease with uridylate-specific activity³³. Although the protein is not necessary for RNA synthesis, it is necessary in coronaviruses to escape recognition of double-stranded RNA intermediates by the host³⁴. PDCoV NS15 inhibits the induction of interferon-β, the main intestinal antiviral cytokine, by preventing nuclear translocation of the interferon regulatory factor IRF1 (ref. 35). The last two mutations in the Hu-PDCoV-specific signature map in the spike glycoprotein. The first mutation, P8A, in the amino-terminal domain of the glycoprotein, is not resolved in the known crystallographic structure, possibly because the segment is too flexible to be seen by cryo-electron microscopy. The second mutation, V550A, is located in the S1 subunit (between the receptor-binding domain and the cleavage site between S1 and S2) on a short β-sheet forming intramolecular contact with a neighbouring loop (Fig. 3b). The V550A change observed in Hu-PDCoV (removal of two methyl groups) is present at relatively low frequency in other PDCoVs discovered in Asia, neither of which displays the additional amino acid changes observed in the strains detected in Haiti. This change, albeit minor, eliminates specific Van der Waals contact (with proline at position 535 and the backbone carbonyl at position 532), potentially enhancing protein flexibility and dynamic movement of S1. Since mutations that prevent intermolecular spike protein interactions between S1 and S2 of the SARS-CoV-2 variant B.1.1.7 have been observed (A570D, D614G and S982A)^{36,37}, V550A may represent a common mechanism that enhances dynamic movements, accelerating membrane fusion events and transmission of the virus.

Fig. 3: Analysis of conserved amino acids in human PDCoV strains.

 **figure3**

Residues are numbered based on the ORF1a/b absolute amino acid position in the reference sequence JQ065043 (pig/China). **a**, Amino acid signature pattern analysis on ORF1a/b. CN, China; SEA, southeast Asia. **b**, The trimer structure of the spike glycoprotein. Different colours are assigned to each monomer: red indicates residue 550, where Haitian sequences have a valine to alanine mutation.

Conclusions

To our knowledge, this is the first report of PDCoV infection in humans, consistent with viraemia and systemic dissemination. The recent divergence of human strains detected in Haiti from their closest pig strains detected in China and the USA in the phylogeny highlights how little we know about the spreading of PDCoV and its

introduction in Haiti. Recent data regarding the movements of live pigs and meat into the country are lacking³⁸, and movements of pigs and their pathogens across the globe can be unexpectedly complicated and difficult to track³⁹, stressing the need for further studies. Our findings, however, are consistent with a virus maintained in the swine population and is capable of successful spillover in humans. Children infected with PDCoV only had mild illness, with less than 1% acute undifferentiated febrile illness during the time period studied, suggesting that the strains identified do not represent a major human health threat. We would emphasize, however, that this study only identified symptomatic children who were acutely viraemic with PDCoV. Further serological studies will be needed to identify the frequency with which such infections occur in the general Haitian population, with the caution that serological studies may be difficult because of possible cross-reactivity with human endemic coronaviruses. Nonetheless, our data highlight the potential for PDCoV zoonoses into human populations, especially in rural or less-developed regions where contact with domestic animals is common.

Methods

Clinical sample collection

From 2012 to 2020, our research group monitored a cohort of approximately 1,250 school children attending one of four schools in the Christianville Foundation school system in the Gressier region of Haiti²¹. Children attending schools in this school system had free access to medical care through a school-based clinic. The study was approved by the Institutional Review Board (IRB) at the University of Florida and the Haitian National IRB; written informed consent for sample collection was obtained from parents of participants, with assent from participants. For this study, children presenting to the school clinic between May 2014 and December 2015 with an acute undifferentiated febrile illness, defined as a history of fever and/or a measured temperature over 37.5 °C in the clinic with no localizing symptoms or signs (that is, no respiratory, skin, or urinary symptoms or signs) were invited to enrol²⁵. After enrolment, clinic healthcare providers recorded clinical data in a study questionnaire and a sample of venous blood (1–3 ml) was collected in an acid citrate dextrose blood collection tube. The blood samples were subsequently centrifuged to pellet the platelets, red blood cells and white blood cells, and the resulting plasma was aseptically transferred to cryovials and stored at –80 °C for subsequent analysis. Appropriate medical care on the basis of clinical presentation and laboratory studies was provided to study participants by clinic healthcare providers. Data on the identification of arboviruses and other virus species among children participating in the study have been previously reported^{22,23,25,26,27,40,41}. As the study was done in young children, the amount of plasma collected was limited, and samples have, in

most instances, been exhausted, owing to the range of studies initially conducted on the samples while screening for other pathogens. IRB restrictions limit our ability to share samples outside of our institution.

Cell culture lines

The African green monkey kidney cell line Vero E6 (Vero C1008) was obtained from the American Type Culture Collection (ATCC; catalogue no. CRL-1586), which authenticates the cells they sell. Upon culture, the cells displayed epithelial morphology (as expected). Before the preparation of seed stocks, the cells were treated for 3 weeks with plasmocin, then verified free of mycoplasma DNA by PCR using a Takara Bio USA mycoplasma detection kit.

Virus identification and sequencing

Attempts at next-generation sequencing using an Illumina MiSeq platform generated minimal coverage, so we sorted to Sanger sequencing using the primer system outlined by Liang et al.⁴², with one addition: to obtain the 5' ends of the viral genomes, a rapid amplification of cDNA ends (RACE) kit was used per the manufacturer's protocols (Life Technologies), and the resulting amplicons were TA-cloned into plasmids and sequenced. PCR amplicons for Sanger sequencing were amplified using AccuScript High-Fidelity reverse transcriptase (Agilent Technologies) in the presence of SUPERase-In RNase inhibitor (Ambion), followed by PCR with Q5 DNA polymerase (New England Biolabs). They were next purified using a QIAquick PCR purification kit (Qiagen) before TA cloning. The inserts in the plasmids were subsequently sequenced bidirectionally using a gene-walking approach, on the basis of obtaining at least 800 bp or non-ambiguous sequence. Briefly, pairs of non-overlapping primers and Q5 polymerase were used to produce 42 separate amplicons corresponding to the PDCoV genome, and each amplicon was Sanger sequenced bidirectionally.

Transmission electron microscopy

For visualization of negative-stained PDCoV by transmission electron microscopy (TEM), 3 ml of cell culture medium that had been harvested from Vero E6 cells that displayed subtle CPE 11 days post-inoculation with plasma was concentrated to approximately 200 µl since it was anticipated that the virus yield would be low and electron microscopy evaluation would be time-consuming. The cell medium was concentrated using an Amicon Ultra-15 centrifugal filter unit that has an Ultracel-100 membrane with a molecular mass cut-off of 100 kDa (Millipore). This was accomplished through centrifugation at 4,000g for 10 min at room temperature until the retentate had a volume of around 200 µl, after which it was recovered and

transferred to a sterile cryovial. Thereafter, 100 µl of the retentate was mixed with an equal volume of freshly prepared 2% paraformaldehyde in 0.1 M PBS (pH 7.20) in preparation for TEM. Thereafter, PCOV was visualized by TEM after an aliquot of the fixed sample was negatively stained at the UF Interdisciplinary Center for Biotechnology Research (ICBR) Electron Microscopy Laboratory (RRID: SCR_019146). For negative staining, a glow-discharged 400-mesh carbon-coated Formvar copper grid was floated on a 5 µl aliquot of virus suspension for 5 min, then washed twice with water. Excess solution was drawn off with filter paper, and the grid floated on 1% (w/v) aqueous uranyl acetate for 30 s. Excess stain was removed with filter paper, the grid was air dried and then examined using a FEI Tecnai G2 F20-TWIN transmission electron microscope (FEI Corporation) that was operated at 200 kV, with digital images acquired using a 4k × 4k CCD camera and Digital Micrograph software (Gatan). Other grids prepared in the same manner were also examined with a FEI Tecnai G2 Spirit Twin transmission electron microscope and digital images were acquired with a Gatan UltraScan 2k × 2k camera and Digital Micrograph software.

Sequence data assembly

The identity of the whole-genome sequences was confirmed via BLAST⁴³ of the nr/nt NCBI database. Following positive identification, available PDCoV sequences from pigs were downloaded from the NCBI (www.ncbi.nlm.nih.gov), together with closely related sparrow deltacoronavirus sequences⁴⁴ (Supplementary Table 1) to be used as outgroups in the phylogenetic analysis (see below).

The final full-genome dataset assembled included 104 PDCoV genomes from pigs, 4 from sparrows (Supplementary Table 1) and 3 newly sequenced Hu-PDCoV strains, which were aligned with MAFFT v.7.407 (ref. [45](#)). A potential recombinant origin of Haitian sequences was assessed using the NeighborNet algorithm⁴⁶, an algorithm based on the PHI test, as implemented in Splitstree^{47,48} v.4.14.8, and with the RDP4 (ref. [49](#)) software package v.4.97.

Phylogenetic and amino acid signature analysis

A phylogenetic signal was verified using likelihood mapping⁵⁰ (Extended Data Fig. 2), as implemented in IQTREE v.2.0.6 (ref. [51](#)). A maximum likelihood tree was calculated using the same version of IQTREE, with the best-fitting nucleotide substitution model according to the Bayesian information criterion and 1,000 bootstrap replicates. The correlation between root-to-tip genetic divergence and sampling dates to assess the clock signal of the alignment was performed with TempEst⁵² before Bayesian phylodynamic analyses. The time-scaled tree was calculated using the Bayesian phylodynamic inference framework in BEAST v1.10.4 (ref. [53](#)). Markov

chain Monte Carlo samplers were run for 200 million generations, with sampling every 20,000 generations, to ensure proper mixing, which was assessed by calculating the effective sampling size of each parameter estimate. The HKY nucleotide substitution model was used with empirical base frequencies and gamma distributions of site-specific rate heterogeneity⁵⁴. The molecular clock was calibrated with a strict clock choosing either a constant size or a Bayesian Skyline Plot demographic prior⁵⁵. A maximum clade credibility tree was inferred from the posterior distribution of trees using TreeAnnotator, specifying a burn-in of 104 million and median node heights. The maximum clade credibility tree was edited graphically using ggtree^{56,57,58}. Markov chain Monte Carlo runs with different demographic priors gave the same result. An analysis with the relaxed molecular clock resulted in a nearly identical mean rate estimate and did not show a rate significantly different from the mean along any of the branches of the tree (coefficient of variation 95% HPD including zero), including the branches leading to the Hu-PDCoV isolates.

Signature pattern analysis of the strains discovered in Haiti compared with the rest of the downloaded sequences was performed with the online version of VESPA⁵⁹; the PDCoV reference sequence JQ065043.2 was used as a guide for the codon coordinates. The spike glycoprotein three-dimensional structure (Protein Data Bank ID: 6B7N)⁶⁰ was used as a base, and figures were generated using PyMol⁶¹.

Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

Data availability

The GenBank accession numbers for sequence data are included in Table 1. Supplementary Table 1 includes a list of the accession numbers of deltacoronaviruses accessed from NCBI for the phylogenetic studies.

References

1. 1. Su, S. et al. Epidemiology, genetic recombination, and pathogenesis of coronaviruses. *Trends Microbiol.* **24**, 490–502 (2016).
2. 2.

Wu, F. et al. A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265–269 (2020).

3. 3.

Ye, Z. W. et al. Zoonotic origins of human coronaviruses. *Int. J. Biol. Sci.* **16**, 1686–1697 (2020).

4. 4.

Fehr, A. R. & Perlman, S. Coronaviruses: an overview of their replication and pathogenesis. *Methods Mol. Biol.* **1282**, 1–23 (2015).

5. 5.

Graham, R. L. & Baric, R. S. Recombination, reservoirs, and the modular spike: mechanisms of coronavirus cross-species transmission. *J. Virol.* **84**, 3134–3146 (2010).

6. 6.

Woo, P. C. et al. Discovery of seven novel mammalian and avian coronaviruses in the genus *Deltacoronavirus* supports bat coronaviruses as the gene source of *Alphacoronavirus* and *Betacoronavirus* and avian coronaviruses as the gene source of *Gammacoronavirus* and *Deltacoronavirus*. *J. Virol.* **86**, 3995–4008 (2012).

7. 7.

Wang, L., Byrum, B. & Zhang, Y. Detection and genetic characterization of deltacoronavirus in pigs, Ohio, USA, 2014. *Emerg. Infect. Dis.* **20**, 1227–1230 (2014).

8. 8.

Jung, K., Hu, H. & Saif, L. J. Porcine deltacoronavirus infection: etiology, cell culture for virus isolation and propagation, molecular epidemiology and pathogenesis. *Virus Res.* **226**, 50–59 (2016).

9. 9.

Ma, Y. et al. Origin, evolution, and virulence of porcine deltacoronaviruses in the United States. *mBio* **6**, e00064 (2015).

10. 10. Wang, Q., Vlasova, A. N., Kenney, S. P. & Saif, L. J. Emerging and re-emerging coronaviruses in pigs. *Curr. Opin. Virol.* **34**, 39–49 (2019).
11. 11. Wang, B. et al. Porcine deltacoronavirus engages the transmissible gastroenteritis virus functional receptor porcine aminopeptidase N for infectious cellular entry. *J. Virol.* **92**, e00318-18 (2018).
12. 12. Boley, P. A. et al. Porcine deltacoronavirus infection and transmission in poultry, United States. *Emerg. Infect. Dis.* **26**, 255–265 (2020).
13. 13. Jung, K., Hu, H. & Saif, L. J. Calves are susceptible to infection with the newly emerged porcine deltacoronavirus, but not with the swine enteric alphacoronavirus, porcine epidemic diarrhea virus. *Arch. Virol.* **162**, 2357–2362 (2017).
14. 14. Li, W. et al. Broad receptor engagement of an emerging global coronavirus may potentiate its diverse cross-species transmissibility. *Proc. Natl Acad. Sci. USA* **115**, E5135–E5143 (2018).
15. 15. He, W. T. et al. Genomic epidemiology, evolution, and transmission dynamics of porcine deltacoronavirus. *Mol. Biol. Evol.* **37**, 2641–2654 (2020).
16. 16. Tagliamonte, M. S. et al. Multiple recombination events and strong purifying selection at the origin of SARS-CoV-2 spike glycoprotein increased correlated dynamic movements. *Int. J. Mol. Sci.* **22**, 80 (2020).
17. 17. Xiao, Y. et al. RNA recombination enhances adaptability and is required for virus spread and virulence. *Cell Host Microbe* **19**, 493–503 (2016).

18. 18.

CIA. The world factbook—Haiti. *CIA* <https://www.cia.gov/the-world-factbook/countries/haiti/>

19. 19.

Alexander, F. C. Experiences with African swine fever in Haiti. *Ann. N. Y. Acad. Sci.* **653**, 251–256 (1992).

20. 20.

Gaertner, P. Whether pigs have wings: African swine fever eradication and pig repopulation in Haiti. *Webster* <http://faculty.webster.edu/corbetre/haiti/misctopic/pigs/gaertner.htm>

21. 21.

Beau De Rochars, V. E. M. et al. Spectrum of outpatient illness in a school-based cohort in Haiti, with a focus on diarrheal pathogens. *Am. J. Trop. Med. Hyg.* **92**, 752–757 (2015).

22. 22.

Beau De Rochars, V. M. et al. Isolation of coronavirus NL63 from blood from children in rural Haiti: phylogenetic similarities with recent isolates from Malaysia. *Am. J. Trop. Med. Hyg.* **96**, 144–147 (2017).

23. 23.

Bonny, T. S. et al. Complete genome sequence of human coronavirus strain 229E isolated from plasma collected from a Haitian child in 2016. *Genome Announc.* **5**, e01313–17 (2017).

24. 24.

Kim, Y. Y. et al. Acute respiratory illness in rural Haiti. *Int. J. Infect. Dis.* **81**, 176–183 (2019).

25. 25.

Ball, J. D. et al. Clinical and epidemiologic patterns of Chikungunya virus infection and coincident arboviral disease in a school cohort in Haiti, 2014–2015. *Clin. Infect. Dis.* **68**, 919–926 (2019).

26. 26.

Lednicky, J. et al. Zika virus outbreak in Haiti in 2014: molecular and clinical data. *PLoS Negl. Trop. Dis.* **10**, e0004687 (2016).

27. 27.

Lednicky, J. et al. Mayaro virus in child with acute febrile illness, Haiti, 2015. *Emerg. Infect. Dis.* **22**, 2000–2002 (2016).

28. 28.

Pan, M. et al. Collection of viable aerosolized influenza virus and other respiratory viruses in a student health care center through water-based condensation growth. *mSphere* **2**, e00251–17 (2017).

29. 29.

Lednicky, J. A. et al. Keystone virus isolated from a Florida teenager with rash and subjective fever: another endemic arbovirus in the southeastern United States? *Clin. Infect. Dis.* **68**, 143–145 (2019).

30. 30.

Hu, H., Jung, K., Wang, Q., Saif, L. J. & Vlasova, A. N. Development of a one-step RT-PCR assay for detection of pancoronaviruses (alpha-, beta-, gamma-, and delta-coronaviruses) using newly designed degenerate primers for porcine and avian fecal samples. *J. Virol. Methods* **256**, 116–122 (2018).

31. 31.

Zhao, Y. et al. Characterization and pathogenicity of the porcine deltacoronavirus isolated in Southwest China. *Viruses* **11**, 1074 (2019).

32. 32.

Homwong, N. et al. Characterization and evolution of porcine deltacoronavirus in the United States. *Prev. Vet. Med.* **123**, 168–174 (2016).

33. 33.

Ivanov, K. A. et al. Major genetic marker of nidoviruses encodes a replicative endoribonuclease. *Proc. Natl Acad. Sci. USA* **101**, 12694–12699 (2004).

34. 34.

Deng, X. et al. Coronavirus nonstructural protein 15 mediates evasion of dsRNA sensors and limits apoptosis in macrophages. *Proc. Natl Acad. Sci. USA* **114**, E4251–E4260 (2017).

35. 35.

Liu, S. et al. Porcine deltacoronavirus (PDCoV) infection antagonizes interferon- λ 1 production. *Vet. Microbiol.* **247**, 108785 (2020).

36. 36.

Public Health England. Investigation of novel SARS-CoV-2 variant: variant of concern 202012/01, technical briefing 3 (Public Health England, 2020).

37. 37.

Rambaut, A., et al Preliminary genomic characterisation of an emergent SARS-CoV-2 lineage in the UK defined by a novel set of spike mutations. COVID-19 Genomics Consortium UK (CoG-UK); <https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spike-mutations/563> (2020).

38. 38.

FAO. FAOSTAT—food and agriculture data. *FAO* <http://www.fao.org/faostat/> (2021).

39. 39.

Mena, I. et al. Origins of the 2009 H1N1 influenza pandemic in swine in Mexico. *eLife* **5**, e16777 (2016).

40. 40.

Blohm, G. et al. Mayaro as a Caribbean traveler: evidence for multiple introductions and transmission of the virus into Haiti. *Int. J. Infect. Dis.* **87**, 151–153 (2019).

41. 41.

ElBadry, M. et al. Isolation of an enterovirus D68 from blood from a child with pneumonia in rural Haiti: close phylogenetic linkage with New York strain.

Pediatr. Infect. Dis. J. **35**, 1048–1050 (2016).

42. 42.

Liang, Q., Li, B., Zhang, H. & Hu, H. Complete genome sequences of two porcine deltacoronavirus strains from Henan Province, China. *Microbiol. Resour. Announc.* **8**, e01517-18 (2019).

43. 43.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).

44. 44.

Chen, Q. et al. The emergence of novel sparrow deltacoronaviruses in the United States more closely related to porcine deltacoronaviruses than sparrow deltacoronavirus HKU17. *Emerg. Microbes Infect.* **7**, 105 (2018).

45. 45.

Nakamura, T., Yamada, K., Tomii, K. & Katoh, K. Parallelization of MAFFT for large-scale multiple sequence alignments. *Bioinformatics* **34**, 2490–2492 (2018).

46. 46.

Huson, D. H. & Bryant, D. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* **23**, 254–267 (2006).

47. 47.

Bruen, T. C., Philippe, H. & Bryant, D. A simple and robust statistical test for detecting the presence of recombination. *Genetics* **172**, 2665–2681 (2006).

48. 48.

Salemi, M., Gray, R. R. & Goodenow, M. M. An exploratory algorithm to identify intra-host recombinant viral sequences. *Mol. Phylogenet. Evol.* **49**, 618–628 (2008).

49. 49.

Martin, D. P., Murrell, B., Golden, M., Khoosal, A. & Muhire, B. RDP4: detection and analysis of recombination patterns in virus genomes. *Virus Evol.* **1**,

vev003 (2015).

50. 50.

Strimmer, K. & von Haeseler, A. Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. *Proc. Natl Acad. Sci. USA* **94**, 6815–6819 (1997).

51. 51.

Nguyen, L. T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).

52. 52.

Rambaut, A., Lam, T. T., Max Carvalho, L. & Pybus, O. G. Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol.* **2**, vew007 (2016).

53. 53.

Drummond, A. J. & Rambaut, A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* **7**, 214 (2007).

54. 54.

Hasegawa, M., Kishino, H. & Yano, T. Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22**, 160–174 (1985).

55. 55.

Hall, M. D., Woolhouse, M. E. & Rambaut, A. The effects of sampling strategy on the quality of reconstruction of viral population dynamics using Bayesian skyline family coalescent methods: a simulation study. *Virus Evol.* **2**, vew003 (2016).

56. 56.

Yu, G. Using ggtree to visualize data on tree-like structures. *Curr. Protoc. Bioinformatics* **69**, e96 (2020).

57. 57.

Yu, G., Lam, T. T., Zhu, H. & Guan, Y. Two methods for mapping and visualizing associated data on phylogeny using ggtree. *Mol. Biol. Evol.* **35**, 3041–3043 (2018).

58. 58.

Yu, G., Smith, D. K., Zhu, H., Guan, Y. & Lam, T. T.-Y. ggtree: an r package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods Ecol. Evol.* **8**, 28–36 (2017).

59. 59.

Korber, B. & Myers, G. Signature pattern analysis: a method for assessing viral sequence relatedness. *AIDS Res. Hum. Retroviruses* **8**, 1549–1560 (1992).

60. 60.

Shang, J. et al. Cryo-electron microscopy structure of porcine deltacoronavirus spike protein in the prefusion state. *J. Virol.* **92**, e01556-17 (2018).

61. 61.

The PyMOL Molecular Graphics System, version 2.0 (Schrodinger, L., 2015).

Acknowledgements

Work was supported in part by grants R01-AI123657S1 and R21AI164007 from the NIAID to J.G.M. Work was also funded in part by special funding provided by the University of Florida (UF) Office of Research and the UF Health Science Center. TEM work was performed at the UF ICBR Electron Microscopy, RRID:SCR_019146 by R. Alvarado and P. Chipman.

Author information

Author notes

1. These authors contributed equally: John A. Lednicky, Massimiliano S. Tagliamonte

Affiliations

1. Emerging Pathogens Institute, University of Florida, Gainesville, FL, USA

John A. Lednicky, Massimiliano S. Tagliamonte, Sarah K. White, Maha A. Elbadry, Md. Mahbubul Alam, Caroline J. Stephenson, Tania S. Bonny, Julia C. Loeb, David A. Ostrov, Carla Mavian, Valery Madsen Beau De Rochars, Marco Salemi & J. Glenn Morris Jr

2. Department of Environmental and Global Health, College of Public Health and Health Professions, University of Florida, Gainesville, FL, USA

John A. Lednicky, Sarah K. White, Maha A. Elbadry, Md. Mahbubul Alam, Caroline J. Stephenson, Tania S. Bonny & Julia C. Loeb

3. Department of Pathology, Immunology and Laboratory Medicine, College of Medicine, University of Florida, Gainesville, FL, USA

Massimiliano S. Tagliamonte, David A. Ostrov, Carla Mavian & Marco Salemi

4. Christianville Foundation, Gressier, Haiti

Taina Telisma & Sonese Chavannes

5. Department of Health Services Research, Management and Policy, College of Public Health and Health Professions, University of Florida, Gainesville, FL, USA

Valery Madsen Beau De Rochars

6. Department of Medicine, College of Medicine, University of Florida, Gainesville, FL, USA

J. Glenn Morris Jr

Contributions

J.A.L. supervised the virology work, identified the viruses subsequent to performing Sanger sequencing and assembly of the sequences, and contributed to manuscript writing. M.S.T. performed the genomic and evolutionary analyses and wrote the manuscript. S.K.W. performed the virology and molecular work, including GenMark analyses. M.A.E. performed the virology and molecular work and coordinated sample shipments and communications between stakeholders in Haiti and the University of Florida. M.M.A., C.J.S., T.S.B. and J.C.L. performed the virology and molecular work. T.T. and S.C. collected plasma samples, performed medical examinations and documented the findings thereof. C.M. contributed to the bioinformatic analyses and edited the manuscript. V.M.B.D.R. provided medical oversight over the project and

established the internal and external processes, including obtaining the necessary permissions for this project. D.A.O. carried out the structural analysis. J.A.L., M.S. and J.G.M. designed the study and contributed to manuscript writing.

Corresponding authors

Correspondence to [Marco Salemi](#) or [J. Glenn Morris Jr.](#)

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Fabian Leendertz, QiuHong Wang 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 Negative stained coronavirus particle viewed by transmission electron microscopy.](#)

Typical coronavirus spikes rim the spherical virus particle.

[Extended Data Fig. 2 Assessment of phylogenetic signal and Neighbor net plot of PDCoV fragments.](#)

a, Likelihood mapping of the final 109 sequences alignment (see [Methods](#)) showing extremely low phylogenetic noise (4.6% in the center of the triangle), as required for reliable phylogeny inference. **b**, Neighbor net inferred from pair-wise p-distances of 47 genomes for the major genome fragment identified by RDP4 (see [Methods](#)).

Recombination signal was assessed by the PHI test. No significant evidence of recombination was found in either fragment (PHI test $p > 0.05$). **c**, Neighbor net and PHI test for the minor genome fragment identified by RDP4 (see [Methods](#)). The only change in topology, in the split decomposition networks shown in panels **a** and **b**,

concerned two Chinese sequences (highlighted by the red circle) displaying mixed ancestry possibly due to homoplasy. These sequences were also removed before performing any further phylogenetic analyses. **d**, Neighbor net and PHI test for the full genome alignment only including Haitian Hu-PDCoV strains and closest Chinese and American strains (after removal recombinant strains detected by RDP4).

Extended Data Fig. 3 Maximum likelihood tree of PDCoV strains.

The tree was inferred from 109 PDCoV full genome sequences, including four sparrow CoV genomes for outgroup rooting (accession numbers MG812375, MG812376, MG812377, and MG812378, from Chen et al.⁴⁴) using the best fitting nucleotide substitution model (see [Methods](#)). For display purposes, outgroup sequences were removed from the figure. Vertical branches are scaled in number of nucleotide substitutions per site according to the bar on the left of the tree. Bootstrap values > 75% (1000 replicates) are indicated along supported branches. The table at the bottom of the tree shows nucleotide % dissimilarity (and total number of nucleotide differences in parenthesis) between Haitian strains and their closest non-Haitian relatives, which are labelled in tree by progressive letters according to the legend on the right.

Extended Data Fig. 4 Analysis of the temporal signal with TempEst.

a, b, Root-to-tip distance (y-axis) *vs.* sampling time linear regression in the ML likelihood phylogeny inferred from: all PDCoV sequences in the final data set ($n = 109$ sequences) including sparrow outgroup sequences (**a**); PDCoV sequences without outgroups ($n = 105$ sequences), with sequences in the red circle belonging to the South East Asia clade (see Fig. 1) showing a clear departure from the strict molecular clock model (**b**); and PDCoV sequences after removal of the South East Asia clade ($n = 94$ sequences), showing greatly improved clock signal ($R^2 = 0.68$) (**c**).

Extended Data Fig. 5 Bayesian maximum clade credibility (MCC) tree of PDCoV strains.

The MCC tree was evolutionary rate. Circles at internal nodes indicate high posterior probability (PP) support >0.9. The table on inferred from a subset of 94 full genome strains that displayed sufficient temporal signal for molecular clock calibration. Branch lengths were scaled in time, according to the bar at the bottom, by using a strict molecular clock and sampling dates to estimate PDCoV the left shows the inferred time of the most recent common ancestor (TMRCA) between Haitian strains and their closest phylogenetic relative, with 95% high posterior density intervals (95%HPD).

Supplementary information

Supplementary Table 1

Accession numbers of deltacoronaviruses accessed from NCBI.

Reporting Summary

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

Lednicky, J.A., Tagliamonte, M.S., White, S.K. *et al.* Independent infections of porcine deltacoronavirus among Haitian children. *Nature* **600**, 133–137 (2021). <https://doi.org/10.1038/s41586-021-04111-z>

- Received: 08 March 2021
- Accepted: 07 October 2021
- Published: 17 November 2021
- Issue Date: 02 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04111-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/s41586-021-04111-z>

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

- Article
- [Published: 10 November 2021](#)

Self-guarding of MORC3 enables virulence factor-triggered immunity

- [Moritz M. Gaidt](#) [ORCID: orcid.org/0000-0002-8009-294X¹](#),
- [Alyssa Morrow](#) [ORCID: orcid.org/0000-0002-5102-8156²](#),
- [Marian R. Fairgrieve¹](#),
- [Jonathan P. Karr³](#),
- [Nir Yosef](#) [ORCID: orcid.org/0000-0001-9004-1225^{2,4,5,6}](#) &
- [Russell E. Vance](#) [ORCID: orcid.org/0000-0002-6686-3912^{1,7,8}](#)

[Nature](#) volume 600, pages 138–142 (2021)

- 7331 Accesses
- 202 Altmetric
- [Metrics details](#)

Subjects

- [Gene regulation in immune cells](#)
- [Herpes virus](#)
- [Innate immunity](#)

Abstract

Pathogens use virulence factors to inhibit the immune system¹. The guard hypothesis^{2,3} postulates that hosts monitor (or ‘guard’) critical innate immune pathways such that their disruption by virulence factors provokes a secondary immune response¹. Here we describe a ‘self-guarded’ immune pathway in human monocytes, in which guarding and guarded functions are combined in one protein. We find that this pathway is triggered by ICP0, a key virulence factor of herpes simplex virus type 1, resulting in robust induction of anti-viral type I interferon (IFN). Notably, induction of IFN by ICP0 is independent of canonical immune pathways and the IRF3 and IRF7 transcription factors. A CRISPR screen identified the ICP0 target MORC3⁴ as an essential negative regulator of IFN. Loss of *MORC3* recapitulates the IRF3- and IRF7-independent IFN response induced by ICP0.

Mechanistically, ICP0 degrades MORC3, which leads to de-repression of a MORC3-regulated DNA element (MRE) adjacent to the *IFNB1* locus. The MRE is required in *cis* for *IFNB1* induction by the MORC3 pathway, but is not required for canonical IFN-inducing pathways. As well as repressing the MRE to regulate *IFNB1*, MORC3 is also a direct restriction factor of HSV-1⁵. Our results thus suggest a model in which the primary anti-viral function of MORC3 is self-guarded by its secondary IFN-repressing function—thus, a virus that degrades MORC3 to avoid its primary anti-viral function will unleash the secondary anti-viral IFN response.

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: Virulence factor-triggered immune sensing.

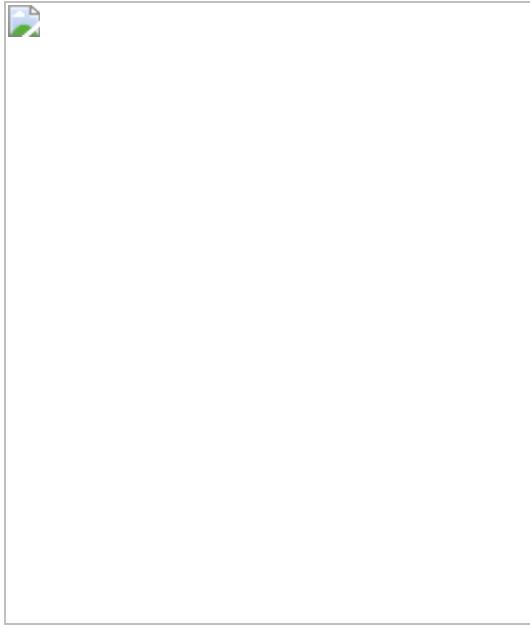


Fig. 2: MORC3 is a negative regulator of IFN.

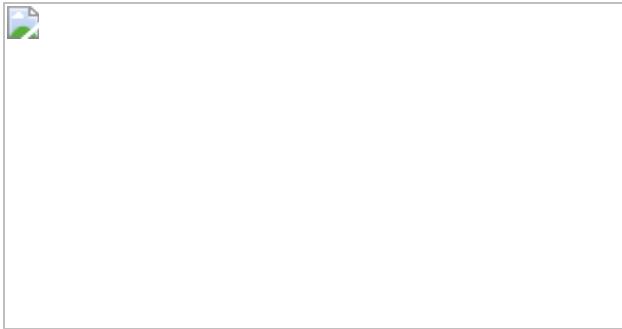


Fig. 3: Locus-specific repression of *IFNB1* by MORC3 enables self-guarding.

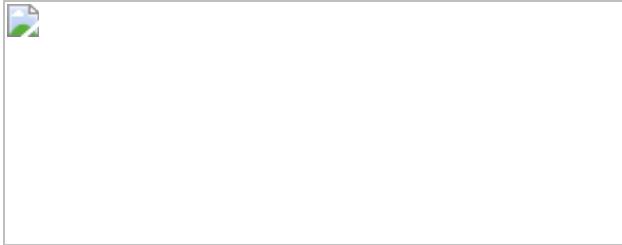


Fig. 4: A MRE explains positional *IFNB1* activation.



Data availability

RNA-seq and ATAC-seq data that support the findings of the study have been deposited in the Gene Expression Omnibus under accession [GSE183011](#). CRISPR–Cas9 screen data have been deposited in the Sequence Read Archive under accession [PRJNA759267](#). [Source data](#) are provided with this paper.

References

1. 1.

Lopes Fischer, N., Naseer, N., Shin, S. & Brodsky, I. E. Effector-triggered immunity and pathogen sensing in metazoans. *Nat. Microbiol.* **5**, 14–26 (2020).

2. 2.

Jones, J. D. & Dangl, J. L. The plant immune system. *Nature* **444**, 323–329 (2006).

3. 3.

Van der Biezen, E. A. & Jones, J. D. Plant disease-resistance proteins and the gene-for-gene concept. *Trends Biochem. Sci.* **23**, 454–456 (1998).

4. 4.

Sloan, E. et al. Analysis of the SUMO2 proteome during HSV-1 infection. *PLoS Pathog.* **11**, e1005059 (2015).

5. 5.

Sloan, E., Orr, A. & Everett, R. D. MORC3, a component of PML nuclear bodies, has a role in restricting herpes simplex virus 1 and human cytomegalovirus. *J. Virol.* **90**, 8621–8633, <https://doi.org/10.1128/JVI.00621-16> (2016).

6. 6.

Janeway, C. A., Jr Approaching the asymptote? Evolution and revolution in immunology. *Cold Spring Harb. Symp. Quant. Biol.* **54**, 1–13 (1989).

7. 7.

Hopfner, K. P. & Hornung, V. Molecular mechanisms and cellular functions of cGAS–STING signalling. *Nat. Rev. Mol. Cell Biol.* **21**, 501–521 (2020).

8. 8.

Ishikawa, H., Ma, Z. & Barber, G. N. STING regulates intracellular DNA-mediated, type I interferon-dependent innate immunity. *Nature* **461**, 788–792 (2009).

9. 9.

Scherer, M. & Stamminger, T. Emerging role of PML nuclear bodies in innate immune signaling. *J. Virol.* **90**, 5850–5854 (2016).

10. 10.

Everett, R. D., Boutell, C. & Hale, B. G. Interplay between viruses and host sumoylation pathways. *Nat. Rev. Microbiol.* **11**, 400–411 (2013).

11. 11.

Sacks, W. R. & Schaffer, P. A. Deletion mutants in the gene encoding the herpes simplex virus type 1 immediate-early protein ICP0 exhibit impaired growth in cell culture. *J. Virol.* **61**, 829–839 (1987).

12. 12.

Rapino, F. et al. C/EBP α induces highly efficient macrophage transdifferentiation of B lymphoma and leukemia cell lines and impairs their tumorigenicity. *Cell Rep.* **3**, 1153–1163 (2013).

13. 13.

Ullman, A. J. & Hearing, P. Cellular proteins PML and Daxx mediate an innate antiviral defense antagonized by the adenovirus E4 ORF3 protein. *J. Virol.* **82**, 7325–7335 (2008).

14. 14.

Everett, R. D. Construction and characterization of herpes simplex virus type 1 mutants with defined lesions in immediate early gene 1. *J. Gen. Virol.* **70**, 1185–1202 (1989).

15. 15.

Lopez, P., Van Sant, C. & Roizman, B. Requirements for the nuclear-cytoplasmic translocation of infected-cell protein 0 of herpes simplex virus 1. *J. Virol.* **75**, 3832–3840 (2001).

16. 16.

Crowl, J. T. & Stetson, D. B. SUMO2 and SUMO3 redundantly prevent a noncanonical type I interferon response. *Proc. Natl Acad. Sci. USA* **115**, 6798–6803 (2018).

17. 17.

Decque, A. et al. Sumoylation coordinates the repression of inflammatory and anti-viral gene-expression programs during innate sensing. *Nat. Immunol.* **17**, 140–149 (2016).

18. 18.

Ferri, F. et al. TRIM33 switches off Ifnb1 gene transcription during the late phase of macrophage activation. *Nat. Commun.* **6**, 8900 (2015).

19. 19.

Malakhova, O. A. et al. UBP43 is a novel regulator of interferon signaling independent of its ISG15 isopeptidase activity. *EMBO J.* **25**, 2358–2367 (2006).

20. 20.

Jadhav, G., Teguh, D., Kenny, J., Tickner, J. & Xu, J. Morc3 mutant mice exhibit reduced cortical area and thickness, accompanied by altered haematopoietic stem cells niche and bone cell differentiation. *Sci. Rep.* **6**, 25964 (2016).

21. 21.

Li, D. Q., Nair, S. S. & Kumar, R. The MORC family: new epigenetic regulators of transcription and DNA damage response. *Epigenetics* **8**, 685–693 (2013).

22. 22.

Rosendorff, A. et al. NXP-2 association with SUMO-2 depends on lysines required for transcriptional repression. *Proc. Natl Acad. Sci. USA* **103**, 5308–5313 (2006).

23. 23.

Mimura, Y., Takahashi, K., Kawata, K., Akazawa, T. & Inoue, N. Two-step colocalization of MORC3 with PML nuclear bodies. *J. Cell Sci.* **123**, 2014–2024 (2010).

24. 24.

Ver, L. S., Marcos-Villar, L., Landeras-Bueno, S., Nieto, A. & Ortín, J. The cellular factor NXP2/MORC3 is a positive regulator of influenza virus multiplication. *J. Virol.* **89**, 10023–10030 (2015).

25. 25.

Zhang, Y. et al. MORC3 is a target of the influenza A viral protein NS1. *Structure* **27**, 1029–1033.e1023 (2019).

26. 26.

Groh, S. et al. Morc3 silences endogenous retroviruses by enabling Daxx-mediated H3.3 incorporation. *Nat. Commun.* **12**, 5996 (2021).

27. 27.

Yu, P. et al. Nucleic acid-sensing Toll-like receptors are essential for the control of endogenous retrovirus viremia and ERV-induced tumors. *Immunity* **37**, 867–879 (2012).

28. 28.

Lopez-Flores, I. & Garrido-Ramos, M. A. The repetitive DNA content of eukaryotic genomes. *Genome Dyn.* **7**, 1–28 (2012).

29. 29.

Gaidt, M. M. et al. Human monocytes engage an alternative inflammasome pathway. *Immunity* **44**, 833–846 (2016).

30. 30.

Gaidt, M. M. et al. The DNA inflammasome in human myeloid cells is initiated by a STING-cell death program upstream of NLRP3. *Cell* **171**, 1110–1124.e1118 (2017).

31. 31.

Blaho, J. A., Morton, E. R. & Yedowitz, J. C. Herpes simplex virus: propagation, quantification, and storage. *Curr. Protoc. Microbiol.* Ch. 14, Unit 14E.11 (2005).

32. 32.

Schmidt, T., Schmid-Burgk, J. L., Ebert, T. S., Gaidt, M. M. & Hornung, V. Designer nuclease-mediated generation of knockout THP1 cells. *Methods Mol. Biol.* **1338**, 261–272 (2016).

33. 33.

Schmid-Burgk, J. L. et al. OutKnocker: a web tool for rapid and simple genotyping of designer nuclease edited cell lines. *Genome Res.* **24**, 1719–1723 (2014).

34. 34.

Labun, K. et al. CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Res.* **47**, W171–W174 (2019).

35. 35.

Schmidt, T., Schmid-Burgk, J. L. & Hornung, V. Synthesis of an arrayed sgRNA library targeting the human genome. *Sci. Rep.* **5**, 14987 (2015).

36. 36.

Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat. Methods* **11**, 783–784 (2014).

37. 37.

Bartok, E. et al. iGLuc: a luciferase-based inflammasome and protease activity reporter. *Nat. Methods* **10**, 147–154 (2013).

38. 38.

Spahn, P. N. et al. PinAPL-Py: a comprehensive web-application for the analysis of CRISPR/Cas9 screens. *Sci. Rep.* **7**, 15854 (2017).

39. 39.

Andrews, S. FastQC: a quality control tool for high throughput sequence data. (2010).

40. 40.

Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).

41. 41.

Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. *BMC Bioinf.* **12**, 323 (2011).

42. 42.

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).

43. 43.

Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287 (2012).

44. 44.

Liberzon, A. et al. Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).

45. 45.

Buenrostro, J. D., Wu, B., Chang, H. Y. & Greenleaf, W. J. ATAC-seq: a method for assaying chromatin accessibility genome-wide. *Curr. Protoc. Mol. Biol.* **109**, 21–29 (2015).

46. 46.

Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10–12 (2011).

47. 47.

Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

48. 48.

Amemiya, H. M., Kundaje, A. & Boyle, A. P. The ENCODE blacklist: identification of problematic regions of the genome. *Sci. Rep.* **9**, 9354 (2019).

49. 49.

Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods* **10**, 1213–1218, <https://doi.org/10.1038/nmeth.2688> (2013).

50. 50.

Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).

51. 51.

Cavalcante, R. G. & Sartor, M. A. annotatr: genomic regions in context. *Bioinformatics* **33**, 2381–2383 (2017).

52. 52.

Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

53. 53.

Criscione, S. W., Zhang, Y., Thompson, W., Sedivy, J. M. & Neretti, N. Transcriptional landscape of repetitive elements in normal and cancer human cells. *BMC Genomics* **15**, 583 (2014).

54. 54.

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).

55. 55.

Freund, E. C. et al. Efficient gene knockout in primary human and murine myeloid cells by non-viral delivery of CRISPR–Cas9. *J. Exp. Med.* **217**, e20191692 (2020).

Acknowledgements

We thank members of the Vance and Barton laboratories for advice and discussions; L. Coscoy, B. Glaunsinger and G. Barton for comments on the manuscript; D. Stetson, T. Graf, R. Tjian, X. Darzacq, D. Raulet and V. Hornung for cell lines; D. Kotov for the plasmid pMCP; B. Roizman for HSV-1 BACs; P. Dietzen and R. Chavez for technical assistance; A. Valeros and H. Nolla for assistance with automated cell sorting; M. West and P. He for assistance with robotics-based liquid handling; and F. Ramirez, C. Rose, C. Hann-Soden and S. McDevitt for deep sequencing. M.M.G. was supported by an EMBO Postdoctoral Fellowship (ALTF 1082-2018). R.E.V. is an Investigator of the Howard Hughes Medical Institute and is funded by NIH grants AI075039, AI063302 and AI155634.

Author information

Affiliations

1. Division of Immunology and Pathogenesis, Department of Molecular and Cell Biology, University of California, Berkeley, CA, USA

Moritz M. Gaidt, Marian R. Fairgrieve & Russell E. Vance

2. Electrical Engineering and Computer Science Department, University of California, Berkeley, CA, USA

Alyssa Morrow & Nir Yosef

3. Division of Genetics, Genomics and Development, Department of Molecular and Cell Biology, University of California, Berkeley, CA, USA

Jonathan P. Karr

4. Center for Computational Biology, University of California, Berkeley, CA, USA

Nir Yosef

5. Ragon Institute of Massachusetts General Hospital, Massachusetts Institute of Technology, and Harvard University, Boston, MA, USA

Nir Yosef

6. Chan Zuckerberg Biohub, San Francisco, CA, USA

Nir Yosef

7. Cancer Research Laboratory and the Immunotherapeutics and Vaccine Research Initiative, University of California, Berkeley, CA, USA

Russell E. Vance

8. Howard Hughes Medical Institute, University of California, Berkeley, CA, USA

Russell E. Vance

Contributions

Conceptualization: M.M.G. and R.E.V. Investigation: M.M.G. Data analysis: M.M.G. and A.M. Writing: M.M.G. and R.E.V., with input from all co-authors. Resources: M.F. and J.K. Funding acquisition: R.E.V. and N.Y. Supervision: N.Y. and R.E.V.

Corresponding authors

Correspondence to [Moritz M. Gaidt](#) or [Russell E. Vance](#).

Ethics declarations

Competing interests

R.E.V. consults for Ventus Therapeutics and Tempest Therapeutics. A provisional patent application covering these findings has been submitted.

Additional information

Peer review information *Nature* thanks John Schoggins 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 Redundancy between DNA- and ICP0-triggered sensing of HSV-1 in BLaER1 monocytes.](#)

(a-c) BLaER1 monocytes were infected with HSV-1 at MOI = 1 for 24h or transfected with DNA for 3h. Gene-expression analysis is depicted as mean + SEM of three independent experiments. **(d)** Immunoblot of BLaER1 monocytes of indicated genotypes. **(e)** BLaER1 monocytes were infected with HSV-1 at MOI = 1 for 24h and IFN β secretion was measured by ELISA. Data is depicted as mean + SEM of four independent experiments. **(f)** BLaER1 monocytes were infected with HSV-1 at MOI = 5. Gene-expression is depicted as mean + SEM of three independent experiments. * p < 0.05; ** p < 0.01; *** p < 0.001; ns = not significantly different than WT, tested by two-way ANOVA and Dunnett's or Bonferroni's post hoc test. See Source Data for exact p values

[Source data](#).

[Extended Data Fig. 2 Redundancy between DNA- and ICP0-triggered sensing of HSV-1 in THP1 and U937 cells.](#)

PMA-differentiated THP1 (**a, c**) or U937 (**b**) human myeloid-like cells were infected with HSV-1 at MOI = 1 or indicated MOI for 24h (**a, b**) or

indicated timepoints (**c**) or transfected with 2'-3'-cGAMP for 4h. Gene-expression analysis is depicted as mean + SEM of three independent experiments. ND = not detected. * p < 0.05; ** p < 0.01; *** p < 0.001; ns = not significantly different than WT, tested by two-way ANOVA or Mixed-effects model and Bonferroni's post hoc test. See Source Data for exact p values

[Source data.](#)

Extended Data Fig. 3 HSV-1 ΔICP0 is attenuated in BLaER1 monocytes.

(**a**) BLaER1 monocytes were infected with HSV-1 at indicated MOI and protein expression was analyzed by immunoblot at indicated timepoints. Data is depicted as one representative of two experiments. (**b**) BLaER1 monocytes were infected with HSV-1 at MOI = 1 and viral progeny was quantified in the supernatant. Mean of five independent experiments is shown. (**c**) BLaER1 monocytes were infected with HSV-1 at indicated MOI and analyzed by flow cytometry. Mean + SEM of three independent experiments is depicted. Data is duplicated in Fig. [3b](#). * p < 0.05; *** p < 0.001; ns = not significantly different than HSV-1 WT, tested by two-way ANOVA and Bonferroni's post hoc test. See Source Data for exact p values

[Source data.](#)

Extended Data Fig. 4 Virulence factors induce IFN in monocytes independently of PRR-signaling hubs.

BLaER1 monocytes (**a**), PMA-differentiated THP1 (**b**) or U937 (**c**) cells expressing doxycycline-inducible virulence factors were stimulated with doxycycline for 24h, left untreated or stimulated with DNA for 3h. Gene-expression as quantified by q-RT-PCR is depicted as mean + SEM of three independent experiments. *IFNB1* expression levels in (**a**) are partially duplicated from Fig. [1c](#). * p < 0.05; ** p < 0.01; *** p < 0.001; ns = not significantly different than the corresponding mCherry-expressing condition, tested by two-way ANOVA or Mixed-effects model and

Dunnett's post hoc test. ND = not detected. See Source Data for exact p values

[Source data](#).

Extended Data Fig. 5 Virulence factor activity is required for IFN induction.

(a, b) BLaER1 monocytes expressing doxycycline-inducible virulence factors were stimulated with doxycycline for 24h. Gene-expression as quantified by q-RT-PCR is depicted as mean ± SEM of three independent experiments and protein expression by immunoblot is depicted from one representative experiment of two. *** p < 0.001; ns = not significantly different than the corresponding mCherry-expressing condition, tested by two-way ANOVA or Mixed-effects model and Dunnett's post hoc test. See Source Data for exact p values. **(c-f)** BLaER1 monocytes expressing doxycycline-inducible virulence factors were stimulated with doxycycline for 24h. Transcriptomic changes in BLaER1 monocytes as detected by RNA-seq were analyzed by PCA of variance stabilizing transformed counts. Top 10 genes that contribute to individual PC directions are depicted (c). Heatmap of log normalized counts of the top 40 most variable genes (column normalized) from three independent RNA-seq experiments is shown (e). Modules from the Molecular Signatures Database that were found to be enriched in differentially expressed genes upon virulence factor expression in BLaER1 monocytes (f). Genes contributing to individual modules can be found in Supplementary Table 1

[Source data](#).

Extended Data Fig. 6 Validation of MORC3 as a repressor of IFN.

(a) Schematic of a genome wide CRISPR screen to identify negative regulators of IFN. **(b)** PinAPL.py analysis of the CRISPR screen to identify negative regulators of IFN. Significantly enriched genes are labeled. Raw data can be found in Supplementary Table 2. **(c)** U2OS cells were infected with HSV-1 for 24h. One representative immunoblot of two is shown. **(d, e)**

Gene expression of BLaER1 monocytes is shown as mean + SEM of 3 independent experiment from one representative clone or two (multiple KOs) or one clone (WT and *MORC3*^{-/-}) except CXCL10 quantification in (d) which is shown as mean of 2 independent experiments. *** p < 0.001; significantly different than the corresponding WT condition, tested by two-way ANOVA and Dunnett's post hoc test. See Source Data for exact p values. (f-i) Immunoblot of BLaER1 monocytes of indicated genotypes

[Source data.](#)

Extended Data Fig. 7 Characterization of *MORC3* deficiency in THP1, U937 and BLaER1 cells.

(a-d) Gene-expression in PMA-differentiated THP1-Cas9 or PMA-differentiated U937-Cas9 human myeloid-like cells expressing indicated sgRNAs is depicted as mean + SEM of three independent experiments. Protein expression in the same cells was analyzed by immunoblot. ND = not detected. *** p < 0.001; ns = not significantly different than the corresponding scramble sgRNA-expressing condition, tested by two-way ANOVA and Bonferroni's post hoc test. See Source Data for exact p values. (e) Transcriptional changes in BLaER1 monocytes as detected by RNA-seq in three independent experiments are depicted by PCA. These data are partially duplicated from Extended Data Fig. 5d. (f) Modules from the Molecular Signatures Database that were found to be enriched in differentially expressed genes in *MORC3*^{-/-} BLaER1 monocytes. Background gene set includes all genes that have base mean expression of at least 1 in *MORC3*^{-/-} BLaER1 monocytes. Genes contributing to individual modules can be found in Supplementary Table 1

[Source data.](#)

Extended Data Fig. 8 PML is not required for IFNB1 regulation by MORC3.

(a) Gene expression of BLaER1 monocytes of indicated genotype is depicted as mean + SEM of three independent experiments. (b) Protein

expression of BLaER1 monocytes was analyzed by immunoblot. (c) BLaER1 monocytes expressing doxycycline-inducible virulence factors were stimulated with doxycycline for 24h. Gene-expression as quantified by q-RT-PCR is depicted as mean + SEM of three independent experiments. (d, e) Gene-expression in HCT116-Cas9 cells expressing indicated sgRNAs is depicted as mean + SEM of three independent experiments. Protein expression in the same cells was analyzed by immunoblot. ND = not detected. ns = not significantly different than the corresponding scramble sgRNA-expressing condition, tested by Mixed-effects model and Bonferroni's post hoc test. See Source Data for exact p values. (f) BLaER1 monocytes of indicated genotype were infected with Δ ICP0 HSV-1 at indicated MOI for 6h and analyzed by flow cytometry. Mean + SEM of three independent experiment is depicted. * p < 0.05 ; ns = not significantly different than the $IFNAR1^{-/-}IFNAR2^{-/-}$ condition, tested by two-way ANOVA and Bonferroni's post hoc test. See Source Data for exact p values

[Source data.](#)

Extended Data Fig. 9 Quantification of expression of ERV families.

ERV family expression was quantified with RepEnrich2 in $IFNAR1^{-/-}IFNAR2^{-/-}$ mCherry vs $MORC3^{-/-}IFNAR1^{-/-}IFNAR2^{-/-}$ monocytes. ERVs with an FDR <0.05 are highlighted in red. The recommended RepEnrich2-EdgeR pipeline did not detect any regulation of ERVs upon *MORC3* deficiency. RepEnrich2-DeSeq2 analysis suggested minimal up- and down-regulation of ERV families. The only strongly de-repressed ERV family upon *MORC3* deficiency includes members within the *MORC3*-repressed region on chromosome 9, consistent with a positional rather than an ERV-specific de-repression

[Source data.](#)

Extended Data Fig. 10 MORC3 deficiency leads to de-repression of a gene cluster at chromosome 9.

(a) Column-normalized heatmap analysis of transcriptomic changes in BLaER1 monocytes as detected by RNA-seq in three independent experiments. **(b)** Transcriptional changes in BLaER1 monocytes as detected by RNA-seq in three independent experiments are depicted by PCA. These data are partially duplicated from Extended Data Fig. 7e. **(c)** Log transcripts per million (TPM) of genes in BLaER1 monocytes of genes clustered near *IFNB1* on chromosome 9 as detected by RNA-seq from three independent experiments. All protein coding genes within this region are depicted. **(d)** BLaER1 monocytes were stimulated with indicated PAMPs for 3h or 24h. Gene-expression analysis is depicted as mean + SEM of five independent experiments. **(e)** Log transcripts per million (TPM) of IFN genes in BLaER1 monocytes were detected by RNA-seq in three independent experiments. **(f, g)** BLaER1 monocytes were stimulated with DNA for 12h or left untreated. Cytokine secretion or gene expression from one representative clone of two per genotype except WT is shown as mean + SEM of four (**f**) or three (**g**) independent experiments. * p < 0.05 ; *** p < 0.001; ns = not significantly different than WT (unless otherwise indicated), tested by two-way ANOVA and Bonferroni's post hoc test. See Source Data for exact p values. **(h)** Gene expression from BLaER1 monocytes of indicated genotypes is depicted as mean + SEM of three independent experiments. *** p < 0.001; ns = not significantly different than WT, tested by two-way ANOVA and Dunnett's post hoc test. See Source Data for exact p values. **(i, j)** Transcriptomic changes in BLaER1 monocytes as detected by RNA-seq in three independent experiments are depicted by PCA. The data in **(i)** is partially duplicated from **(b)**. The distance on PC1-axis between samples and the mean of mCherry-expressing cells was calculated and is depicted as mean + SEM from three independent experiments. *** p < 0.001 significantly different than WT, tested by two-way ANOVA and Dunnett's post hoc test. See Source Data for exact p values

[Source data.](#)

Extended Data Fig. 11 Positional de-repression of *IFNB1* explains IFN induction by virulence factors.

(a) BLaER1 monocytes were infected with HSV-1 at MOI=1 for 24h. *MLT3* expression is depicted as mean + SEM of n = 3 independent

experiments. *** p < 0.001; ns = not significantly different than mock infection, tested by two-way ANOVA and Dunnett's post hoc test. See Source Data for exact p values. (b) Mean log transcripts per million (TPM) of a gene cluster at chromosome 9 in BLaER1 monocytes. Mean TPM is calculated for each condition across three independent experiments. All protein coding genes within this region are depicted. (c) Log transcripts per million (TPM) of IFN genes in BLaER1 monocytes upon virulence factor expression were detected by RNA-seq in three independent experiments. (d) BLaER1 monocytes expressing doxycycline-inducible virulence factors were stimulated with doxycycline for 24h. Gene expression is shown as mean + SEM of n = 3 independent experiments. ** p < 0.01; *** p < 0.001; ns = not significantly different than WT, tested by two-way ANOVA and Dunnett's post hoc test. See Source Data for exact p values. Note that *IFNB1*^{-/-} cells harbor small indels within *IFNB1* that allow detection of mRNA by q-RT-PCR. (e, f) BLaER1 monocytes expressing doxycycline-inducible virulence factors were stimulated with doxycycline for 24h. Transcriptomic changes as detected by RNA-seq in three independent experiments were analyzed PCA. Data is partially duplicated from Extended Data Fig. 10i. The distance on PC1-axis between samples and the mean of mCherry-expressing cells was calculated and is depicted as mean + SEM from three independent experiments. *** p < 0.001 significantly different than WT, tested by two-way ANOVA and Dunnett's post hoc test. See Source Data for exact p values

[Source data](#).

Extended Data Fig. 12 Validation of the MORC3-regulated DNA element.

(a, b) *STAT1*^{-/-}*STAT2*^{-/-}BLaER1-Cas9 expressing a randomly integrated *IFNB1*-promoter-Luciferase reporter were transduced with the indicated sgRNAs and stimulated with cytosolic DNA for 24h. Luciferase signal and IFN β secretion is depicted as mean + SEM of n = 3 independent experiments. * p < 0.05; ns = not significantly different, tested by paired, two-sided t-test. See Source Data for exact p values. (c) Protein expression of indicated BLaER1 monocytes was analyzed by immunoblot. This experiment was performed once. (d) Consensus sequences of amplicon

sequencing at the MRE locus from *STAT1*^{-/-}*STAT2*^{-/-} *MRE*^{Δ/Δ} BLaER1 Cas9 cells were aligned to the WT reference. 3166bp were omitted from the reference sequence. (e) Protein expression of indicated BLaER1 monocytes was analyzed by immunoblot. One representative experiment of two is depicted. (f) Gene expression of indicated BLaER1 monocytes is depicted as mean + SEM of three independent experiments from one representative clone of two (*MORC3*^{-/-} *MLLT3*^{-/-} *IFNARI*^{-/-} *IFNAR2*^{-/-} and *MORC3*^{-/-} *FOCAD*^{-/-} *IFNARI*^{-/-} *IFNAR2*^{-/-}) or one (*IFNARI*^{-/-}*IFNAR2*^{-/-} and *MORC3*^{-/-}*IFNARI*^{-/-}*IFNAR2*^{-/-}) per genotype. ns = not significantly different than the *IFNARI*^{-/-}*IFNAR2*^{-/-} condition, tested by two-way ANOVA and Dunnett's post hoc test. See Source Data for exact p values. (g) Protein expression of indicated BLaER1 monocytes was analyzed by immunoblot. This experiment was performed once. (h) Log transcripts per million (TPM) of high-confidence MORC3 targets (genes de-repressed in *MORC3*^{-/-} *IFNARI*^{-/-}*IFNAR2*^{-/-}, *MORC3*^{-/-} *IFNB1*^{-/-} and *MORC3*^{-/-} *STAT1*^{-/-}*STAT2*^{-/-} Cas9) in WT = *MORC3*^{-/-} *STAT1*^{-/-}*STAT2*^{-/-} Cas9 or indicated genotypes are depicted as mean of two independent RNA-seq experiments. (i) *STAT1*^{-/-}*STAT2*^{-/-} Cas9 ("WT") or *STAT1*^{-/-}*STAT2*^{-/-} *MRE*^{Δ/Δ} BLaER1 Cas9 were infected with HSV-1 at indicated MOI for 4-5h. One representative immunoblot of two is depicted. (j) Peripheral blood human monocytes were nucleofected with indicated sgRNA:Cas9 complexes and differentiated with M-CSF for 5 days. Gene expression from three independent donors is depicted as mean + SEM. * p < 0.05; ** p < 0.01; *** p < 0.001; significantly different than the scramble sgRNA condition, tested by two-way ANOVA and Bonferroni's post hoc test. ND = not detected. Source Data for exact p values

[Source data.](#)

Supplementary information

[Supplementary Information](#)

This file contains Supplementary Figs. 1–4 (showing FACS gating strategies and the full scans of the immunoblots), Supplementary Note 1 and Supplementary References.

Reporting Summary

Supplementary Table 1

Gene module enrichment in RNA-seq data.

Supplementary Table 2

Candidate negative regulators of IFN. PinAPL.py analysis of the CRISPR screen.

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**](#)

[**Source Data Extended Data Fig. 8**](#)

[**Source Data Extended Data Fig. 9**](#)

[**Source Data Extended Data Fig. 10**](#)

[**Source Data Extended Data Fig. 11**](#)

[**Source Data Extended Data Fig. 12**](#)

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Gaidt, M.M., Morrow, A., Fairgrieve, M.R. *et al.* Self-guarding of MORC3 enables virulence factor-triggered immunity. *Nature* **600**, 138–142 (2021).
<https://doi.org/10.1038/s41586-021-04054-5>

- Received: 22 December 2020
- Accepted: 23 September 2021
- Published: 10 November 2021
- Issue Date: 02 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04054-5>

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/s41586-021-04054-5>

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

- Article
- [Published: 13 October 2021](#)

Structural basis of cytokine-mediated activation of ALK family receptors

- [Steven De Munck^{1,2},](#)
- [Mathias Provost^{1,2},](#)
- [Michiko Kurikawa³,](#)
- [Ikuko Omori³,](#)
- [Junko Mukohyama³,](#)
- [Jan Felix ORCID: orcid.org/0000-0002-8436-9467^{1,2},](#)
- [Yehudi Bloch ORCID: orcid.org/0000-0001-7924-3539^{1,2},](#)
- [Omar Abdel-Wahab ORCID: orcid.org/0000-0002-3907-6171⁴,](#)
- [J. Fernando Bazan⁵,](#)
- [Akihide Yoshimi ORCID: orcid.org/0000-0002-0664-7281³ &](#)
- [Savvas N. Savvides ORCID: orcid.org/0000-0003-3420-5947^{1,2}](#)

[Nature](#) volume 600, pages 143–147 (2021)

- 4529 Accesses
- 1 Citations
- 71 Altmetric
- [Metrics details](#)

Subjects

- [Cancer](#)
- [Metabolism](#)
- [X-ray crystallography](#)

Abstract

Anaplastic lymphoma kinase (ALK)¹ and the related leukocyte tyrosine kinase (LTK)² are recently deorphanized receptor tyrosine kinases³. Together with their activating cytokines, ALKAL1 and ALKAL2^{4,5,6} (also called FAM150A and FAM150B or AUGβ and AUGα, respectively), they are involved in neural development⁷, cancer^{7,8,9} and autoimmune diseases¹⁰. Furthermore, mammalian ALK recently emerged as a key regulator of energy expenditure and weight gain¹¹, consistent with a metabolic role for *Drosophila* ALK¹². Despite such functional pleiotropy and growing therapeutic relevance^{13,14}, structural insights into ALK and LTK and their complexes with cognate cytokines have remained scarce. Here we show that the cytokine-binding segments of human ALK and LTK comprise a novel architectural chimera of a permuted TNF-like module that braces a glycine-rich subdomain featuring a hexagonal lattice of long polyglycine type II helices. The cognate cytokines ALKAL1 and ALKAL2 are monomeric three-helix bundles, yet their binding to ALK and LTK elicits similar dimeric assemblies with two-fold symmetry, that tent a single cytokine molecule proximal to the cell membrane. We show that the membrane-proximal EGF-like domain dictates the apparent cytokine preference of ALK. Assisted by these diverse structure–function findings, we propose a structural and mechanistic blueprint for complexes of ALK family receptors, and thereby extend the repertoire of ligand-mediated dimerization mechanisms adopted by receptor tyrosine kinases.

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 ALK family cytokine-binding domain.

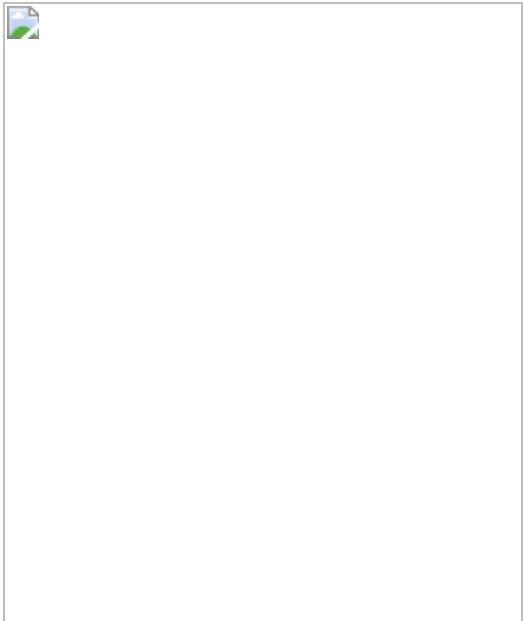


Fig. 2: Cytokine-mediated dimerization of ALK and LTK.

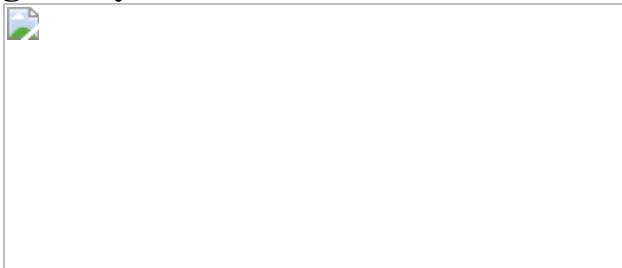


Fig. 3: ALK– and LTK–cytokine complexes harbour three distinct interaction sites.

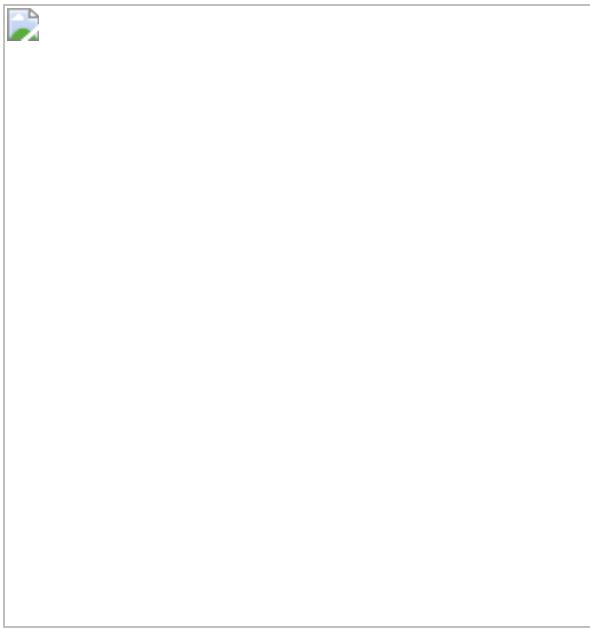
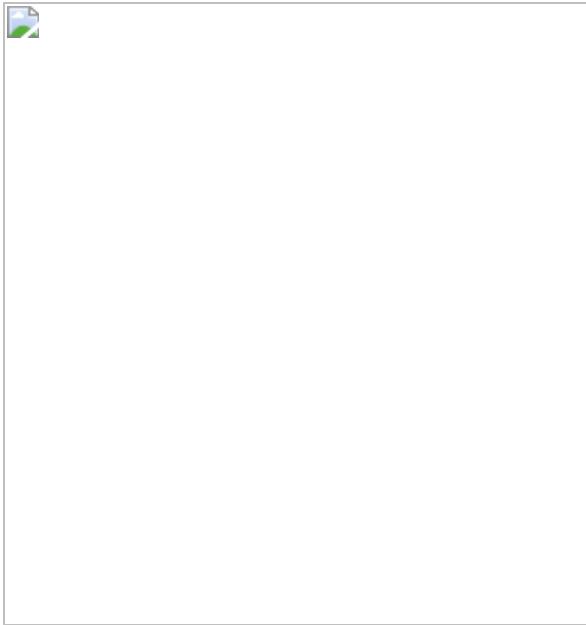


Fig. 4: The EGF-like domain dictates cytokine specificity in ALK.



Data availability

Coordinates and structure factors for the complexes have been deposited in the Protein Data Bank (PDB) under accessions [7NWZ](#) (ALK_{TG}–ALKAL2), [7NX0](#) (LTK_{TG}–ALKAL1–Nb3.16), [7NX1](#) (LTK_{TG}), [7NX2](#) (unbound Fab324), [7NX3](#) (ALK_{TG}–EGFL–Fab324), [7NX4](#) (ALK_{TG}–EGFL). Most common single-nucleotide polymorphisms were obtained from the COSMIC database (<https://cancer.sanger.ac.uk/cosmic>) and gnomAD database (<https://gnomad.broadinstitute.org>). The Ba/F3 cell lines expressing ALK(WT), ALK(R753Q) or ALK(F856S) can be obtained from the authors upon request. [Source data](#) are provided with this paper.

References

1. 1.

Morris, S. W. et al. Fusion of a kinase gene, ALK, to a nucleolar protein gene, NPM, in non-Hodgkin's lymphoma. *Science* **263**, 1281–1284 (1994).

2. 2.

Ben-Neriah, Y. & Bauskin, A. R. Leukocytes express a novel gene encoding a putative transmembrane protein-kinase devoid of an extracellular domain. *Nature* **333**, 672–676 (1988).

3. 3.

Zhang, H. et al. Deorphanization of the human leukocyte tyrosine kinase (LTK) receptor by a signaling screen of the extracellular proteome. *Proc. Natl Acad. Sci. USA* **111**, 15741–15745 (2014).

4. 4.

Reshetnyak, A. V. et al. Augmentor α and β (FAM150) are ligands of the receptor tyrosine kinases ALK and LTK: Hierarchy and specificity of ligand-receptor interactions. *Proc. Natl Acad. Sci. USA* **112**, 15862–15867 (2015).

5. 5.

Guan, J. et al. FAM150A and FAM150B are activating ligands for anaplastic lymphoma kinase. *eLife* **4**, e09811 (2015).

6. 6.

Reshetnyak, A. V. et al. Identification of a biologically active fragment of ALK and LTK-ligand 2 (augmentor-α). *Proc. Natl Acad. Sci. USA* **115**, 8340–8345 (2018).

7. 7.

Hallberg, B. & Palmer, R. H. Mechanistic insight into ALK receptor tyrosine kinase in human cancer biology. *Nat. Rev. Cancer* **13**, 685–700 (2013).

8. 8.

Janostiak, R., Malvi, P. & Wajapeyee, N. Anaplastic lymphoma kinase confers resistance to BRAF kinase inhibitors in melanoma. *iScience* **16**, 453–467 (2019).

9. 9.

Javanmardi, N. et al. Analysis of *ALK*, *MYCN*, and the ALK ligand *ALKAL2* (*FAM150B/AUGα*) in neuroblastoma patient samples with chromosome arm 2p rearrangements. *Genes Chromosomes Cancer* **59**, 50–57 (2020).

10. 10.

Li, N. et al. Gain-of-function polymorphism in mouse and human Ltk: implications for the pathogenesis of systemic lupus erythematosus. *Hum. Mol. Genet.* **13**, 171–179 (2004).

11. 11.

Orthofer, M. et al. Identification of ALK in thinness. *Cell* **181**, 1246–1262.e22 (2020).

12. 12.

Pospisilik, J. A. et al. *Drosophila* genome-wide obesity screen reveals Hedgehog as a determinant of brown versus white adipose cell fate. *Cell* **140**, 148–160 (2010).

13. 13.

Borenäs, M. et al. ALK ligand ALKAL2 potentiates MYCN-driven neuroblastoma in the absence of *ALK* mutation. *EMBO J.* **40**, e105784 (2021).

14. 14.

Sano, R. et al. An antibody-drug conjugate directed to the ALK receptor demonstrates efficacy in preclinical models of neuroblastoma. *Sci. Transl. Med.* **11**, eaau9732 (2019).

15. 15.

Dornburg, A. et al. Comparative genomics within and across bilaterians illuminates the evolutionary history of ALK and LTK proto-oncogene origination and diversification. *Genome Biol. Evol.* **13**, evaa228 (2021).

16. 16.

Murray, P. B. et al. Heparin is an activating ligand of the orphan receptor tyrosine kinase ALK. *Sci. Signal.* **8**, ra6 (2015).

17. 17.

Alvarado, D. et al. Anti-ALK antibodies and methods for use thereof. US patent 15/755421 (2021).

18. 18.

Crick, F. H. C. & Rich, A. Structure of polyglycine II. *Nature* **176**, 780–781 (1955).

19. 19.

Lorén, C. E. et al. A crucial role for the anaplastic lymphoma kinase receptor tyrosine kinase in gut development in *Drosophila melanogaster*. *EMBO Rep.* **4**, 781–786 (2003).

20. 20.

Kolodny, R. Searching protein space for ancient sub-domain segments. *Curr. Opin. Struct. Biol.* **68**, 105–112 (2021).

21. 21.

Fadeev, A. et al. ALKALs are in vivo ligands for ALK family receptor tyrosine kinases in the neural crest and derived cells. *Proc. Natl Acad. Sci. USA* **115**, E630–E638 (2018).

22. 22.

Wang, Y. W. et al. Identification of oncogenic point mutations and hyperphosphorylation of anaplastic lymphoma kinase in lung cancer. *Neoplasia* **13**, 704–715 (2011).

23. 23.

Maxson, J. E. et al. Therapeutically targetable ALK mutations in leukemia. *Cancer Res.* **75**, 2146–2150 (2015).

24. 24.

Durham, B. H. et al. Activating mutations in CSF1R and additional receptor tyrosine kinases in histiocytic neoplasms. *Nat. Med.* **25**, 1839–1842 (2019).

25. 25.

Malinauskas, T., Aricescu, A. R., Lu, W., Siebold, C. & Jones, E. Y. Modular mechanism of Wnt signaling inhibition by Wnt inhibitory factor 1. *Nat. Struct. Mol. Biol.* **18**, 886–893 (2011).

26. 26.

Wehrman, T. et al. Structural and mechanistic insights into nerve growth factor interactions with the TrkA and p75 receptors. *Neuron* **53**, 25–38 (2007).

27. 27.

Ogiso, H. et al. Crystal structure of the complex of human epidermal growth factor and receptor extracellular domains. *Cell* **110**, 775–787 (2002).

28. 28.

Elegherdt, J. et al. Extracellular complexes of the hematopoietic human and mouse CSF-1 receptor are driven by common assembly principles. *Structure* **19**, 1762–1772 (2011).

29. 29.

Schlessinger, J. et al. Crystal structure of a ternary FGF–FGFR–heparin complex reveals a dual role for heparin in FGFR binding and dimerization. *Mol. Cell* **6**, 743–750 (2000).

30. 30.

Uchikawa, E., Choi, E., Shang, G., Yu, H. & Xiao-Chen, B. Activation mechanism of the insulin receptor revealed by cryo-EM structure of the fully liganded receptor–ligand complex. *eLife* **8**, e48630 (2019).

31. 31.

Aricescu, A. R., Lu, W. & Jones, E. Y. A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallogr. D* **62**, 1243–1250 (2006).

32. 32.

Reeves, P. J., Callewaert, N., Contreras, R. & Khorana, H. G. Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous *N*-glycosylation by a tetracycline-inducible *N*-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. *Proc. Natl Acad. Sci. USA* **99**, 13419–13424 (2002).

33. 33.

Backliwal, G. et al. Valproic acid: A viable alternative to sodium butyrate for enhancing protein expression in mammalian cell cultures. *Biotechnol. Bioeng.* **101**, 182–189 (2008).

34. 34.

Gorrec, F. The MORPHEUS II protein crystallization screen. *Acta Crystallogr. F* **71**, 831–837 (2015).

35. 35.

Kabsch, W. XDS. *Acta Crystallogr. D* **66**, 125–132 (2010).

36. 36.

McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).

37. 37.

Bricogne G., et al. BUSTER 2.11.2 (United Kingdom Global Phasing Ltd, 2017).

38. 38.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).

39. 39.

Terwilliger, T. C. Maximum-likelihood density modification. *Acta Crystallogr. D* **56**, 965–972 (2000).

40. 40.

Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. Sect. D* **75**, 861–877 (2019).

41. 41.

Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat. Protoc.* **3**, 1171–1179 (2008).

42. 42.

Murshudov, G. N. et al. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. D* **67**, 355–367 (2011).

43. 43.

D'Arcy, A., Villard, F. & Marsh, M. An automated microseed matrix-screening method for protein crystallization. *Acta Crystallogr. D* **63**, 550–554 (2007).

44. 44.

Strong, M. et al. Toward the structural genomics of complexes: crystal structure of a PE/PPE protein complex from *Mycobacterium tuberculosis*. *Proc. Natl Acad. Sci. USA* **103**, 8060–8065 (2006).

45. 45.

Howarth, M. & Ting, A. Y. Imaging proteins in live mammalian cells with biotin ligase and monovalent streptavidin. *Nat. Protoc.* **3**, 534–545 (2008).

46. 46.

Hopkins, J. B., Gillilan, R. E. & Skou, S. BioXTAS RAW: Improvements to a free open-source program for small-angle X-ray scattering data reduction and analysis. *J. Appl. Crystallogr.* **50**, 1545–1553 (2017).

47. 47.

Biasini, M. et al. SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* **42**, W252 (2014).

48. 48.

Webb, B. & Sali, A. Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinforma.* **2016**, 5.6.1–5.6.37 (2016).

49. 49.

Yoshimi, A. et al. Coordinated alterations in RNA splicing and epigenetic regulation drive leukaemogenesis. *Nature* **574**, 273–277 (2019).

50. 50.

Ishihara, T. et al. HEN-1, a secretory protein with an LDL receptor motif, regulates sensory integration and learning in *Caenorhabditis elegans*. *Cell* **109**, 639–649 (2002).

51. 51.

Englund, C. et al. Jeb signals through the Alk receptor tyrosine kinase to drive visceral muscle fusion. *Nature* **425**, 512–516 (2003).

52. 52.

Lee, H. H., Norris, A., Weiss, J. B. & Frasch, M. Jelly belly protein activates the receptor tyrosine kinase Alk to specify visceral muscle pioneers. *Nature* **425**, 507–512 (2003).

53. 53.

Blum, M. et al. The InterPro protein families and domains database: 20 years on. *Nucleic Acids Res.* **49**, D344–D354 (2021).

54. 54.

Pentelute, B. L. et al. X-ray structure of snow flea antifreeze protein determined by racemic crystallization of synthetic protein enantiomers. *J. Am. Chem. Soc.* **130**, 9695–9701 (2008).

55. 55.

Buglino, J., Shen, V., Hakimian, P. & Lima, C. D. Structural and biochemical analysis of the Obg GTP binding protein. *Structure* **10**, 1581–1592 (2002).

56. 56.

Weidenweber, S. et al. Structure of the acetophenone carboxylase core complex: Prototype of a new class of ATP-dependent carboxylases/hydrolases. *Sci. Rep.* **7**, 1–10 (2017).

57. 57.

Dunne, M. et al. Salmonella phage S16 tail fiber adhesin features a rare polyglycine rich domain for host recognition. *Structure* **26**, 1573–1582.e4 (2018).

58. 58.

Krissinel, E. & Henrick, K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **60**, 2256–2268 (2004).

Acknowledgements

We thank the staff of beamlines P13 and P14 (Petra III, Deutsches Elektronen-Synchrotron), Proxima 2A (SOLEIL) and ID23-2 (ESRF) and SWING (SOLEIL) for their technical support and beamtime allocation. S.D.M. was supported by a predoctoral fellowship from the Flanders Agency for Innovation and Entrepreneurship (VLAIO-Flanders, Belgium). Y.B. is a post-doctoral research fellow of Research Foundation Flanders (FWO). A.Y. acknowledges support from the Japan Society for the Promotion of Science (JSPS) Home-Returning Researcher Development Research (grant number 19K24691), KAKENHI (grant number 21H04828) and National Cancer Center Research and Development Funds (2020-A-2). S.N.S. acknowledges research support from the FWO (grant number G0B4918N), Ghent University (grant number BOF17-GOA-028), the Hercules Foundation (no. AUGE-11-029) and the Flanders Institute for Biotechnology (VIB).

Author information

Affiliations

1. Unit for Structural Biology, Department of Biochemistry and Microbiology, Ghent University, Ghent, Belgium

Steven De Munck, Mathias Provost, Jan Felix, Yehudi Bloch & Savvas N. Savvides

2. Unit for Structural Biology, VIB-UGent Center for Inflammation Research, Ghent, Belgium

Steven De Munck, Mathias Provost, Jan Felix, Yehudi Bloch & Savvas N. Savvides

3. Cancer RNA Research Unit, National Cancer Center Research Institute, Tokyo, Japan

Michiko Kurikawa, Ikuko Omori, Junko Mukohyama & Akihide Yoshimi

4. Human Oncology and Pathogenesis Program, Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY, USA

Omar Abdel-Wahab

5. h Bioconsulting llc, Stillwater, MN, USA

J. Fernando Bazan

Contributions

S.D.M. performed protein expression and purification and performed crystallization experiments with contributions from M.P. S.D.M. determined and analysed crystal structures with contributions from Y.B. and S.N.S. J.F. performed analysis of SAXS data with contributions from S.D.M. S.D.M. performed ITC, SEC-MALLS and BLI experiments with contributions from M.P. O.A.-W. and A.Y. contributed materials and designed cell assays. M.K., I.O. and J.M. performed cellular assays under the supervision of A.Y. J.F.B. carried out evolutionary and structural

analyses with contributions from S.D.M. and S.N.S. S.D.M. and S.N.S. wrote the manuscript with contributions from all authors. S.N.S. conceived and supervised the project.

Corresponding author

Correspondence to [Savvas N. Savvides](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks the 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 Phylogenetic analysis and evolution of ALK and LTK.](#)

a, Phylogenetic analysis of ALK (**a**). Multiple sequence alignment was performed with MAFFT 7 using the E-INS-I algorithm for refinement. The phylogenetic tree was visualized using Phyl.io tool as implemented in MAFFT. ALK is an evolutionarily ancient RTK in *C. elegans* and *D. melanogaster*, where it is activated by HEN-1⁵⁰ and Jeb^{51,52}. **b**, Sequence alignment of ALK sequences focusing on the heparin binding motif. Positively charged residues are highlighted in blue. The multiple sequence

alignment was performed with MAFFT 7 using the E-INS-I algorithm for refinement. **c**, Phylogenetic analysis of LTK performed as for ALK. In mammals the LTK ectodomain consists of the TG and EGF-like domains while other vertebrates contain the additional MAM-LDLa-MAM domains at the N-terminus.

Extended Data Fig. 2 Purification of ALK_{TG-EGFL} and LTK_{TG} and structural details of the TNFL and GR subdomains of the novel TG supradomain fold.

a,b,c Representative chromatograms and SDS-PAGE gels for the purification of ALK_{TG-EGFL} (**a**), ALK_{TG-EGFL}—Fab324 (**b**) and LTK_{TG} (**c**). The arrow indicates the shift in elution volume after EndoH digest of ALK_{TG-EGFL}. Each protein was purified several times, chromatograms and SDS-PAGE analysis of each sample are representative for different protein batches. Uncropped gels are included in source data. **d**, ALK_{TG-EGFL} structure colored according to secondary structure elements. α -Helices (blue), β -strands (green), pGII-helices (orange), loops (grey). **e**, Structure of the ALK_{TG-EGFL}—Fab324 complex with ALK coloured according to its secondary structure elements. CDR loops of Fab324 are coloured yellow. The constant domains of Fab324 are omitted for clarity. **f**, LTK_{TG} structure colored according to secondary structure elements. **g**, Hexagonal pGII-helix arrangement surrounding pGII-helix d in LTK. Vermillion pGII-helices consist exclusively of glycine residues. pGII-helix d shown as sticks, hydrogen bonds to other residues in LTK are indicated as dotted lines. The glycine-rich segment has complicated detection of a globular fold but has led to its sequence-based classification as Glycine-rich PFAM domain PF12810⁵³. **h**, Schematic representations of pGII-helix arrangements in reported structures. Drastically less extensive pGII-helix arrangements than the one displayed in the GR subdomain of ALK and LTK have been observed in synthetic polyglycines²⁰ and four functionally diverse proteins^{54,55,56,57}. Full circles indicate pGII-helices coming out of the plane of the page while empty circles indicate helices going into the plane of the page. S16 adhesin (pdb: 6F45) Apc complex (pdb: 5L9W) obg (pdb: 5M04) Sf antifreeze protein (pdb: 3BOI). **i**, Sequence alignment performed with

Clustal Omega of human ALK, human LTK, *C. elegans* ALK (SCD-2) and *D. melanogaster* ALK covering pGII-helices j,k and l. Residues conserved across all four species are indicated with an orange background. Conserved hydrophobic residues involved in the hydrophobic groove between the TNF-like and glycine rich region are indicated by a green sphere. **j**, The β -sheet subregions of ALK and a trimmed view of TNF (pdb: 1TNF) are coloured in a N-(blue) to C-terminus (red) gradient and shown side by side after structural superposition. Topology diagram for the TNFL domain of ALK and the jelly-roll fold of TNF follow the same colour scheme. Jelly-roll fold nomenclature starts with strand B according to convention. For ALK_{TNFL} the nomenclature in black is according to the TG domain notation used in this study while the nomenclature according to the TNF convention (first β -strand labeled B) is shown in grey. Structural queries⁶³ using the TNFL subdomains retrieved TNF/C1q-class folds (r.m.s.d = 2.8 Å against C1q and TNF, 72 C α atoms). Topology-independent searches⁵⁸ covered an additional ~20 residues in the canonical TNF fold, and structure-based sequence alignments clarified the sequence homology between the A, D and E β -strands in ALK/LTK_{TNFL} and β -strands B, E and F in TNF or TRAIL. The distinctly connected β -strands in the ALK_{TNFL}/LTK_{TNFL} subdomain break up the alternating sheet-to-sheet register of the TNF/C1q β -jellyroll, and instead permit the spatially contiguous sprouting of the three glycine-rich loop inserts (between β -strands D and E, F and G, and H and H') towards the distinctive pGII-helix lattice of the ALK_{GR}/LTK_{GR} subdomain. The sequential B to I β -strands of the TNF/C1q β -jellyroll smoothly sew together the two β -sheets (that feature characteristic B'BIDG and FEHC faces) whereas the ALK/LTK_{TNFL} subdomain has AIDEH'' and H'HGFCB faces (primed small caps denote additional, edge β -strands). **k**, Annotated alignment of selected β -strands of human ALK, LTK, TNF and TRAIL. Conserved hydrophobic residues are indicated in red.

[Source data](#)

[Extended Data Fig. 3 ALKAL1- and ALKAL2-dependent proliferation of Ba/F3 cells expressing ALK.](#)

a, Molecular mass determination of ALKAL2, ALKAL1 and ALKAL1_{FL} by size-exclusion chromatography and in-line multi-angle laser light scattering (SEC-MALLS). The differential refractive index (left vertical axis) is plotted against the determined molecular weight (right vertical axis). ALKAL2 (blue trace), ALKAL1 (green trace) and ALKAL1_{FL} (vermillion trace). **b**, Cell proliferation of ALK^{WT} or EV (empty vector) expressing Ba/F3 cells treated with 10 nM ALKAL2 with and without addition of crizotinib ($n = 3$ biologically independent experiments; mean \pm s.d.). **c**, Cell proliferation of Ba/F3 cells expressing ALK^{WT} or EV upon stimulation by a concentration series of ALKAL2, ALKAL1 or ALKAL1_{FL} at indicated concentrations. ($n = 3$ biologically independent experiments; mean \pm s.d.; two-way ANOVA with Tukey's multiple comparison test compared with EV; * $P < 0.05$, ** $P < 0.01$ and *** $P < 0.001$; Exact P -values are provided in source data. The ratio of the observed ALKAL-induced cell growth with the IL-3 induced cell growth is shown in a heatmap representation for the measurements on day 4).

[Source data](#)

Extended Data Fig. 4 Biophysical characterization and purification of ALKAL-mediated complexes of ALK and LTK.

a, Calculated molecular masses for the ALK, LTK and ALKAL constructs under study. **b-e**, Experimental molecular mass determination of ALKAL-mediated complexes with LTK_{TG-EGFL} (**b**), ALK_{TG-EGFL} (deglycosylated) (**c**), LTK_{TG} (**d**), ALK_{TG} (deglycosylated) (**e**) by SEC-MALLS. The differential refractive index (left vertical axis) is plotted against the experimentally measured molecular mass (right vertical axis). Unbound LTK and ALK are in orange and pink respectively. Complexes with ALKAL1 or ALKAL2 are in green and blue respectively. The reported molecular mass represents the average molecular mass \pm s.d. across the elution peak. **f,g**, Representative SEC elution profiles and SDS-PAGE analysis for the purification of the ALK_{TG-EGFL}—ALKAL2 complex (**f**) and the LTK_{TG-EGFL}—ALKAL1—Nb3.16 complex (**g**). Each protein was purified several times, chromatograms and SDS-PAGE analysis of each sample are representative for different protein batches. Uncropped gels are

included in source data. **h**, Surface representation of the LTK_{TG}—EGFL—ALKAL1—Nb3.16 complex illustrating the location of the binding sites of the non-neutralizing Nb3.16 to LTK_{TG} far from any cytokine–receptor or receptor–receptor binding site.

Source data

Extended Data Fig. 5 Structure of ALKAL1 and ALKAL2.

a, Structural superposition of the conserved C-terminal segments of ALKAL1 and ALKAL2 (coloured in an N- (blue) to C-terminus (red) gradient) as observed in their complexes with cognate receptors. ALKAL1 and ALKAL2 share high sequence identity (66%) in their C-terminal domain, but have variable N-terminal regions. ALKAL1 residues in the interface between helix A and helices B and C are shown as sticks. α A connects via a conserved short loop to a helical hairpin constructed from α B and α C, which in turn are tethered by two conserved disulfides. A conserved stretch of ~10 residues preceding α A was not ordered in the reported structures, suggesting they might help to stabilize the soluble forms of these cytokines rather than contribute to direct receptor engagement or might help reduce the entropic cost of binding. **b**, Surface representation of ALKAL1 coloured according to the Eisenberg hydrophobicity scale. Hydroxyl groups surrounding the central cavity are shown as red spheres. **c**, Sequence alignment by ClustalOmega of human ALKAL1 with ALKAL2. Residues in the interface between helix A and helices B and C are coloured according to their position. Cysteines are coloured yellow and disulfide bonds are shown as yellow lines. **d**, Surface representation of the LTK_{TG}—ALKAL1 complex with LTK in ALKAL1 coloured according to the Eisenberg hydrophobicity scale (white is more hydrophobic). The black circle denotes the ALKAL1 hydrophobic cavity. **e**, Multiple sequence alignments of various vertebrate ALKAL1 and ALKAL2 sequences using the ESPript server (<http://escript.ibcp.fr/ESPript/ESPript/>) and structural annotation according to secondary structure elements. Symbols indicate residues participating in interaction sites 1 and 2 according to the graphical legend. *Hs* (*Homo sapiens*), *Mm* (*Mus musculus*), *Gg* (*Gallus gallus*), *Xl* (*Xenopus laevis*), *Xt* (*Xenopus tropicalis*), *Dr* (*Danio rerio*). **f**, ALKAL1 shown in green transparent surface and ribbon

representation. Residues differing with ALKAL2 are colored vermillion and labeled with the ALKAL1/ALKAL2 numbering.

Extended Data Fig. 6 Structural details of receptor–cytokine and receptor–receptor interactions in ALK/LTK–cytokine complexes.

a, 2Fo-Fc electron density maps contoured at +1 r.m.s.d. showing details of site 1, 2 and 3 of the LTK—ALKAL1 and ALK—ALKAL2 complexes. **b**, Transparent surface of ALKAL1 according to the Eisenberg hydrophobicity scale illustrating similarities and differences in site 1 of LTK/ALK–cytokine complexes. Shown is the central conserved hydrophobic patch formed by leucines (L97, L116 and L120) and the interacting residues of LTK (orange). The equivalent ALK residues (pink) are shown after alignment with LTK. **c**, View of Site 2 in the LTK—ALKAL1 and ALK—ALKAL2 complexes. ALKALs are coloured according to the Eisenberg hydrophobicity scale. Receptor residues surrounding the hydrophobic triad of helix A (L72, F76, F80 in ALKAL1 and M93, F97, L101 in ALKAL2) are shown as sticks for LTK (orange) and ALK (pink). **d**, Superposition of unbound ALK (dark gray), unbound LTK (light gray), bound ALK (pink, only helices shown) and bound LTK (orange, only helices shown). **e**, Superposition of the ALK—ALKAL2 and LTK—ALKAL1 complexes, zoomed in on the region around the e-f loop. **f**, View of the site 3 groove of ALK (top) and LTK (bottom). In LTK, site 3 centers on His153, which stacks against Gly74 and Arg241 on an LTK-specific loop with Asn369 residues hydrogen bonding across the twofold axis of the complex.

Extended Data Fig. 7 Functional interrogation of site 1, site 2, and site 3 interfaces in ALK/LTK–cytokine complexes.

a,b Representative response curves as measured by biolayer-interferometry (BLI) for the interaction of wild type ALKAL1 and ALKAL1 mutants (containing charge-reversal mutations of residues involved in site 1) (**a**) and WT ALKAL2 and ALKAL2 mutants (**b**) with ALK_{TG-EGFL} and LTK_{TG-EGFL}. For wild type ALKALs LTK curves were fitted with a 2:1 binding model (red) while for ALK a 1:1 model was used. Start and end

concentrations of the 2-fold dilution series used for the WT measurements is shown as an inset while for all mutants a 2-fold dilution series from 6.4 μM -400nM was used. **c**, BLI response curves for the interaction of the site 2 ALKAL1^{F76E} mutant with LTK_{TG-EGFL}. **d**, BLI response curves for the interaction of the site 2 ALKAL2^{F97E} and ALKAL2^{H100A} with LTK_{TG-EGFL}. **e**, BLI response curves for the interaction of the site 2 ALKAL2^{F97E} with ALK_{TG-EGFL}. **f**, SDS-PAGE analysis of purified ALKAL1 and ALKAL2 mutants used in Ba/F3 and SEC-MALLS assays. Each protein was purified several times, SDS-PAGE analysis of each sample are representative for different protein batches. Uncropped gels are included in source data. **g**, Western blot analysis of phosphorylated ALK (Y1278 and Y1604) after stimulation with ALKAL2^{WT}, ALKAL2^{R123E/R136E}, ALKAL2^{F97E} and ALKAL2^{H100A}. Uncropped western blot scans are provided in source data. **h**, Capacity of ALKAL1 and ALKAL2 mutants to form complexes with LTK_{TG-EGFL} and ALK_{TG-EGFL} respectively as characterized by SEC-MALLS. Differential refractive index (left axis) is plotted against the determined molecular weight (right axis). LTK_{TG-EGFL} (orange trace), LTK_{TG-EGFL}—ALKAL1^{R102E/R115E} (green trace) and LTK_{TG-EGFL}—ALKAL1^{F76E} (blue trace). ALK_{TG-EGFL} (pink trace), ALK_{TG-EGFL}—ALKAL2^{R123E/R136E} (green trace) and ALK_{TG-EGFL}—ALKAL2^{F97E} (blue trace). The ALKAL1 site 1 mutant is unable to form a complex with ALK while the site 2 mutant still forms a binary complex. The reported molecular mass represents the average molecular mass \pm s.d. across the elution peak. **i**, Capacity of the LTK_{TG-EGFL}^{R241A} LTK_{TG-EGFL}^{R241A/N369G} site 3 mutants to form complexes with ALKAL1 as characterized by SEC-MALLS. LTK_{TG-EGFL}^{R241A} (red trace), LTK_{TG-EGFL}^{R241A/N369G} (cyan trace), LTK_{TG-EGFL}^{R241A}—ALKAL1 (green trace), LTK_{TG-EGFL}^{R241A/N369G}—ALKAL1 (blue trace). LTK_{TG-EGFL} (orange trace) and LTK_{TG-EGFL}—ALKAL1 (pink trace) are shown for comparison. The reported molecular mass represents the average molecular mass \pm s.d. across the elution peak. **j**, Cell proliferation of Ba/F3 cells expressing ALK^{WT} or ALK^{M751T} upon stimulation with 50 nM ALKAL1 or 50 nM ALKAL2. Western blot analysis of ALK^{WT} or ALK^{M751T} expression is a representative of three

biologically independent experiments with similar results. Uncropped western blot scans are provided in source data.

[Source data](#)

Extended Data Fig. 8 Conservation of TG supradomains in ALK and LTK.

a, Structurally annotated multiple sequence alignments of the TG and EGFL domains of ALK and LTK using the ESPript server (<http://escript.ibcp.fr/ESPript/ESPript/>). *Hs* (*Homo sapiens*), *Mm* (*Mus musculus*), *Gg* (*Gallus gallus*) *Xl* (*Xenopus laevis*), *Xt* (*Xenopus tropicalis*) *Dr* (*Danio rerio*). Symbols indicate residues participating in the different interfaces according to the graphical legend. **b**, Bottom view of dimerized LTK (left) and ALK (right) in surface representation coloured according to residue conservation. The ALKAL binding sites are shown as an outline of ALKAL1 (green) and ALKAL2 (blue) are shown as dashed lines. Conservation analysis was performed using the Consurf server (<https://consurf.tau.ac.il>) based on an alignment of 248 vertebrate sequences for ALK and 225 for LTK by COBALT. **c**, Top view of dimerized LTK (left) and ALK (right) coloured according to sequence conservation levels.

Extended Data Fig. 9 Mapping of missense mutations on the structures of the ALK—ALKAL2 and LTK—ALKAL1 complexes.

a, Mapping of most frequent SNPs (GnomAD) to the ALKAL2—ALK_{TG} complex shown in top view. SNPs also found in COSMIC database (<https://cancer.sanger.ac.uk/cosmic>) are also indicated on the bottom view. Mutations further characterized in this study are coloured red. Inset shows the detailed position of R753 in the ALK_{TG}—ALKAL2 complex. **b**, Mapping of most frequent SNPs (GnomAD) to the ALKAL1—LTK_{TG} complex shown in top view. SNPs also found in COSMIC database (<https://cancer.sanger.ac.uk/cosmic>) are also indicated on the bottom view. **c**, Western blot analysis of the expression levels of ALK^{WT} and the ALK^{R753Q} and ALK^{F856S} mutants and their ERK phosphorylation (left). On

the right-side western blot analysis of phosphorylated ALK^{WT}, ALK^{R753Q} and ALK^{F856S} is shown. Representative results from three biologically independent experiments with similar results. Uncropped western blot scans are provided in source data. **d**, Sanger sequencing of cDNA showing WT or mutant ALK expression in isogenic Ba/F3 cells. **e**, Cell proliferation of ALK expressing Ba/F3 cells treated with 10 nM ALKAL2 with and without addition of crizotinib for ALK R753Q and F856S mutants. Data for EV and WT ALK are repeated from Extended Data Fig. 3 for direct comparison. Crizotinib is also able to inhibit ALKAL2 induced proliferation for mutant ALK, indicating ALK dependent signalling. ($n = 3$ biologically independent experiments; mean \pm s.d.; two-way ANOVA with Tukey's multiple comparison test compared with DMSO control). **f**, Proliferation of Ba/F3 cells expressing ALK carrying the R753Q or F856S mutation upon stimulation by a concentration series of ALKAL2, ALKAL1 or ALKAL1_{FL} at indicated concentration ($n=3$ biologically independent experiments; mean \pm s.d.; two-way ANOVA with Tukey's multiple comparison test compared with ALK^{WT}). ALKAL-induced cell growth relative to that of cells cultured with IL-3 is shown in a heatmap representation. EV and ALK^{WT} controls are included for comparison. * $P < 0.05$, ** $P < 0.01$ and *** $P < 0.001$. Exact P -values are provided in source data).

Source data

Extended Data Fig. 10 Mechanistic insights into the assembly of ALK/LTK–cytokine complexes derived from microcalorimetry, SAXS, and SEC-MALLS.

a,b ITC experiments for the titration of LTK_{TG-EGFL} (5 μ M) with ALKAL1 (56 μ M). **(a)** Titration of LTK_{TG-EGFL} (12 μ M) with ALKAL2 (56 μ M). **(b)**. ITC titration curves, the left panel shows the raw data with the differential electrical power (DP) plotted against time. The right panel represents the binding isotherm obtained from the integration of the raw data and fitted to a one-site model. mean \pm s.d were calculated based on 3 measurements. **c**, Small-angle X-ray scattering analysis and calculated FoXS fits of the binary ALK_{TG-EGFL}—ALKAL2 (pink), ternary ALK_{TG-EGFL}—ALKAL2 (purple), binary LTK_{TG-EGFL}—ALKAL1 (orange) and ternary LTK_{TG-EGFL}—

ALKAL1 (light orange) to experimental SAXS data (black curves). **d**, ITC experiments for the titration of LTK_{TG}-EGFL (7.2 μM) with ALKAL1_{FL} (55 μM) Data representative of 2 independent experiments. **e**, Ternary ALK_{TG}:ALKAL2 complex with regions differing with ALKAL1 coloured vermillion. C-termini of the TG domains leading towards the EGF-like domains are coloured red. **f**, ITC experiments for the titration of LTK_{TG} (10 μM) with ALKAL1 (70 μM) Data representative of 2 independent experiments. **g**, ITC experiments for the titration of LTK_{TG} (10 μM) with ALKAL2 (100 μM) Data representative of 2 independent experiments. **h**, ITC experiments for the titration of LTK_{TG} (10 μM) with ALKAL1_{FL} (40 μM) Data representative of 2 independent experiments. **i**, ITC experiments for the titration of ALK_{FL} (8 μM) with ALKAL2 (82 μM) Data representative of 2 independent experiments. **j**, Characterization of heparin-induced ALK_{FL} dimerization by SEC-MALLS. Differential refractive index (left vertical axis) is plotted against the determined molecular weight (right vertical axis). ALK_{FL} (pink), ALK_{FL} complexes with heparin dp20 (green) and ALK_{FL} complexed with heparin dp20 and ALKAL2 (blue trace). The reported molecular mass represents the average molecular mass ± s.d. across the elution peak. **k**, Overview of the different ligand-mediated extracellular assemblies across RTKs. Trka—NGF (PDB: 2IFG), EGFR—EGF (PDB: 1IVO), INSR—INS (PDB: 6PXW), CSF1R—CSF1 (PDB: 4WRM), FGFR—heparin—FGF (PDB: 1FQ9), ALK_{TG}—ALKAL2 (PDB: 7NWZ).

[Source data](#)

Extended Data Table 1 Crystallographic data and refinement statistics
Extended Data Table 2 Interaction interface analysis of ALK—ALKAL2 and LTK—ALKAL1 complexes

Supplementary information

[Supplementary Figure 1](#)

Uncropped blots shown in Extended Data Figs. 2, 4, 7, 9.

Reporting Summary

Peer Review File

Source data

[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. 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

De Munck, S., Provost, M., Kurikawa, M. *et al.* Structural basis of cytokine-mediated activation of ALK family receptors. *Nature* **600**, 143–147 (2021). <https://doi.org/10.1038/s41586-021-03959-5>

- Received: 24 March 2021
- Accepted: 25 August 2021
- Published: 13 October 2021
- Issue Date: 02 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-03959-5>

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/s41586-021-03959-5>

- Article
- [Published: 24 November 2021](#)

Structural basis for ligand reception by anaplastic lymphoma kinase

- [Tongqing Li^{1,2},](#)
- [Steven E. Stayrook ORCID: orcid.org/0000-0002-1677-8293^{1,2},](#)
- [Yuko Tsutsui^{1,2},](#)
- [Jianan Zhang^{1,2},](#)
- [Yueyue Wang²,](#)
- [Hengyi Li ORCID: orcid.org/0000-0001-6807-7432^{1,2},](#)
- [Andrew Proffitt³,](#)
- [Stefan G. Krimmer¹,](#)
- [Mansoor Ahmed¹,](#)
- [Olivia Belliveau²,](#)
- [Ian X. Walker^{1,2},](#)
- [Krishna C. Mudumbi ORCID: orcid.org/0000-0002-5991-9375^{1,2},](#)
- [Yoshihisa Suzuki¹,](#)
- [Irit Lax^{1,2},](#)
- [Diego Alvarado³,](#)
- [Mark A. Lemmon ORCID: orcid.org/0000-0002-3379-5319^{1,2},](#)
- [Joseph Schlessinger ORCID: orcid.org/0000-0002-5085-5969^{1,2} &](#)
- [Daryl E. Klein ORCID: orcid.org/0000-0002-7188-0450^{1,2}](#)

Nature volume 600, pages 148–152 (2021)

- 2663 Accesses

- 1 Citations
- 60 Altmetric
- [Metrics details](#)

Subjects

- [Antibody therapy](#)
- [CNS cancer](#)
- [Kinases](#)
- [X-ray crystallography](#)

Abstract

The proto-oncogene *ALK* encodes anaplastic lymphoma kinase, a receptor tyrosine kinase that is expressed primarily in the developing nervous system. After development, ALK activity is associated with learning and memory¹ and controls energy expenditure, and inhibition of ALK can prevent diet-induced obesity². Aberrant ALK signalling causes numerous cancers³. In particular, full-length ALK is an important driver in paediatric neuroblastoma^{4,5}, in which it is either mutated⁶ or activated by ligand⁷. Here we report crystal structures of the extracellular glycine-rich domain (GRD) of ALK, which regulates receptor activity by binding to activating peptides^{8,9}. Fusing the ALK GRD to its ligand enabled us to capture a dimeric receptor complex that reveals how ALK responds to its regulatory ligands. We show that repetitive glycines in the GRD form rigid helices that separate the major ligand-binding site from a distal polyglycine extension loop (PXL) that mediates ALK dimerization. The PXL of one receptor acts as a sensor for the complex by interacting with a ligand-bound second receptor. ALK activation can be abolished through PXL mutation or with PXL-targeting antibodies. Together, these results explain how ALK uses its atypical architecture for its regulation, and suggest new therapeutic opportunities for ALK-expressing cancers such as paediatric neuroblastoma.

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 ALK–ALKAL complex and the ALK GRD.

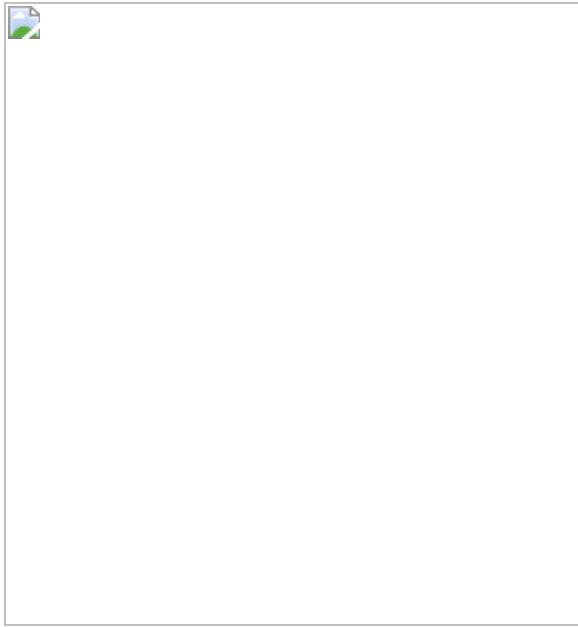


Fig. 2: Ligand binding is directed by the C-terminal loop of ALK.

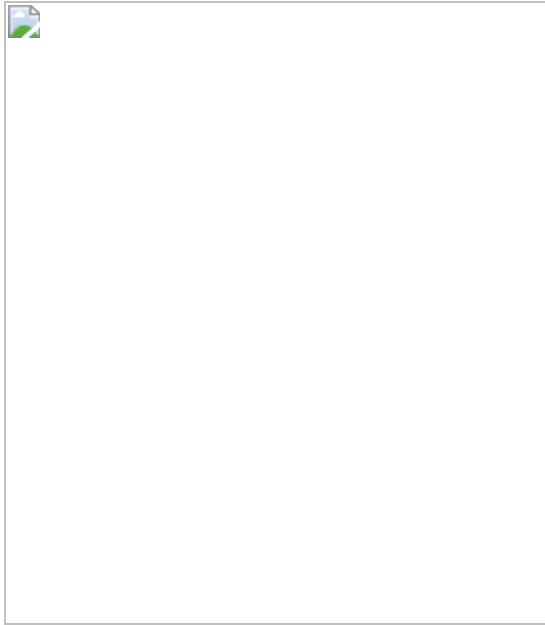


Fig. 3: Dimerization of ALK is directed by the PXL.

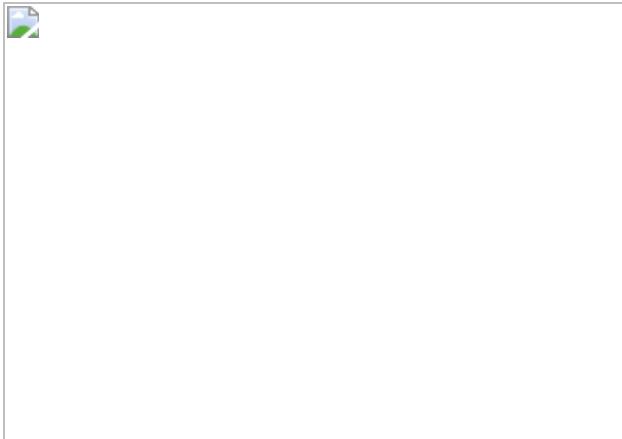
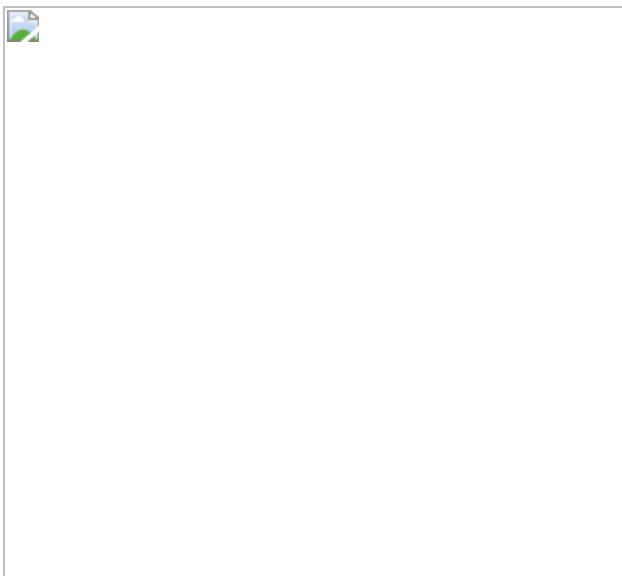


Fig. 4: Antibodies to the C-terminal loop or PXL prevent ALK activation.



Data availability

The refined structural protein models and corresponding structure–factor amplitudes have been deposited under PDB accession codes [7LS0](#) (human ALK GRD–ALKAL fusion complex), [7LRZ](#) (human ALK GRD), [7LIR](#) (worm ALK GRD) and [7MK7](#) (ALKAL2-AD).

References

1. 1.

Weiss, J. B. et al. Anaplastic lymphoma kinase and leukocyte tyrosine kinase: functions and genetic interactions in learning, memory and adult neurogenesis. *Pharmacol. Biochem. Behav.* **100**, 566–574 (2012).

2. 2.

Orthofer, M. et al. Identification of ALK in thinness. *Cell* **181**, 1246–1262.e1222, (2020).

3. 3.

Hallberg, B. & Palmer, R. H. Mechanistic insight into ALK receptor tyrosine kinase in human cancer biology. *Nat. Rev. Cancer* **13**, 685–700 (2013).

4. 4.

Carpenter, E. L. et al. Antibody targeting of anaplastic lymphoma kinase induces cytotoxicity of human neuroblastoma. *Oncogene* **31**, 4859–4867 (2012).

5. 5.

Mosse, Y. P. et al. Identification of ALK as a major familial neuroblastoma predisposition gene. *Nature* **455**, 930–935 (2008).

6. 6.

Trigg, R. M. & Turner, S. D. ALK in neuroblastoma: biological and therapeutic implications. *Cancers* **10**, 113 (2018).

7. 7.

Borenas, M. et al. ALK ligand ALKAL2 potentiates MYCN-driven neuroblastoma in the absence of ALK mutation. *EMBO J.* **40**, e105784 (2021).

8. 8.

Reshetnyak, A. V. et al. Augmentor α and β (FAM150) are ligands of the receptor tyrosine kinases ALK and LTK: hierarchy and specificity of ligand–receptor interactions. *Proc. Natl Acad. Sci. USA* **112**, 15862–15867 (2015).

9. 9.

Guan, J. et al. FAM150A and FAM150B are activating ligands for anaplastic lymphoma kinase. *eLife* **4**, e09811 (2015).

10. 10.

Lemmon, M. A. & Schlessinger, J. Cell signaling by receptor tyrosine kinases. *Cell* **141**, 1117–1134 (2010).

11. 11.

Loren, C. E. et al. A crucial role for the Anaplastic lymphoma kinase receptor tyrosine kinase in gut development in *Drosophila melanogaster*. *EMBO Rep.* **4**, 781–786 (2003).

12. 12.

Zhang, H. et al. Deorphanization of the human leukocyte tyrosine kinase (LTK) receptor by a signaling screen of the extracellular proteome. *Proc. Natl Acad. Sci. USA* **111**, 15741–15745 (2014).

13. 13.

Reshetnyak, A. V. et al. Identification of a biologically active fragment of ALK and LTK-ligand 2 (augmentor- α). *Proc. Natl Acad. Sci. USA* **115**, 8340–8345 (2018).

14. 14.

Qin, L. Y. et al. Discovery of 7-(3-(piperazin-1-yl)phenyl)pyrrolo[2,1-f][1,2,4]triazin-4-amine derivatives as highly potent and selective PI3K δ inhibitors. *Bioorg. Med. Chem. Lett.* **27**, 855–861 (2017).

15. 15.

Youn, S. J. et al. Construction of novel repeat proteins with rigid and predictable structures using a shared helix method. *Sci. Rep.* **7**, 2595 (2017).

16. 16.

Holm, L. DALI and the persistence of protein shape. *Protein Sci.* **29**, 128–140 (2020).

17. 17.

Eck, M. J. & Sprang, S. R. The structure of tumor necrosis factor- α at 2.6 Å resolution. Implications for receptor binding. *J. Biol. Chem.* **264**, 17595–17605 (1989).

18. 18.

Warkentin, E. et al. A rare polyglycine type II-like helix motif in naturally occurring proteins. *Proteins* **85**, 2017–2023 (2017).

19. 19.

Crick, F. H. & Rich, A. Structure of polyglycine II. *Nature* **176**, 780–781 (1955).

20. 20.

Dunne, M. et al. Salmonella phage S16 tail fiber adhesin features a rare polyglycine rich domain for host recognition. *Structure* **26**, 1573–1582.e1574 (2018).

21. 21.

Vadas, O., Jenkins, M. L., Dornan, G. L. & Burke, J. E. Using hydrogen–deuterium exchange mass spectrometry to examine protein–membrane interactions. *Methods Enzymol.* **583**, 143–172 (2017).

22. 22.

Sano, R. et al. An antibody–drug conjugate directed to the ALK receptor demonstrates efficacy in preclinical models of neuroblastoma. *Sci. Transl. Med.* **11**, eaau9732 (2019).

23. 23.

Tate, J. G. et al. COSMIC: the catalogue of somatic mutations in cancer. *Nucleic Acids Res.* **47**, D941–D947 (2019).

24. 24.

Ishihara, T. et al. HEN-1, a secretory protein with an LDL receptor motif, regulates sensory integration and learning in *Caenorhabditis elegans*. *Cell* **109**, 639–649 (2002).

25. 25.

Englund, C. et al. Jeb signals through the Alk receptor tyrosine kinase to drive visceral muscle fusion. *Nature* **425**, 512–516 (2003).

26. 26.

Lee, H. H., Norris, A., Weiss, J. B. & Frasch, M. Jelly belly protein activates the receptor tyrosine kinase Alk to specify visceral muscle pioneers. *Nature* **425**, 507–512 (2003).

27. 27.

Murray, P. B. et al. Heparin is an activating ligand of the orphan receptor tyrosine kinase ALK. *Sci. Signal.* **8**, ra6 (2015).

28. 28.

Reshetnyak, A. V. et al. Mechanism for the activation of the anaplastic lymphoma kinase receptor. *Nature* <https://doi.org/10.1038/s41586-021-04140-8> (2021).

29. 29.

Jenni, S., Goyal, Y., von Grotthuss, M., Shvartsman, S. Y. & Klein, D. E. Structural basis of neurohormone perception by the receptor tyrosine kinase torso. *Mol. Cell.* **60**, 941–952 (2015).

30. 30.

Klein, D. E., Stayrook, S. E., Shi, F., Narayan, K. & Lemmon, M. A. Structural basis for EGFR ligand sequestration by Argos. *Nature* **453**, 1271–1275 (2008).

31. 31.

Vonrhein, C., Blanc, E., Roversi, P. & Bricogne, G. Automated structure solution with autoSHARP. *Methods Mol. Biol.* **364**, 215–230 (2007).

32. 32.

Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).

33. 33.

McCoy, A. J. et al. Phaser crystallographic software. *J. Appl. Crystallogr.* **40**, 658–674 (2007).

34. 34.

Langer, G., Cohen, S. X., Lamzin, V. S. & Perrakis, A. Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. *Nat. Protoc.* **3**, 1171–1179 (2008).

35. 35.

Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D* **75**, 861–877 (2019).

36. 36.

Morin, A. et al. Collaboration gets the most out of software. *eLife* **2**, e01456 (2013).

37. 37.

Patil, K., et al Computational studies of anaplastic lymphoma kinase mutations reveal common mechanisms of oncogenic activation. *Proc. Natl Acad. Sci. USA* (in the press).

38. 38.

Grimm, J. B. et al. A general method to improve fluorophores for live-cell and single-molecule microscopy. *Nat. Methods* **12**, 244–250 (2015).

39. 39.

Kabsch, W. Xds. *Acta Crystallogr. D* **66**, 125–132 (2010).

40. 40.

Winn, M. D. et al. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D* **67**, 235–242 (2011).

41. 41.

Adams, P. D. et al. The Phenix software for automated determination of macromolecular structures. *Methods* **55**, 94–106 (2011).

Acknowledgements

We thank C. Alarcon, Y. Liu, J. Abraham and their laboratories for valuable discussions, as well as members of the Klein, Lemmon and Schlessinger laboratories; A. Reshetnyak, D. Puleo and J. Mohanty for their contributions. This work was supported by the NIH, NIGMS grant R35 GM122485 (to M.A.L.), and NCI grant R01 CA248532 (to D.E.K.). This

work is based in part on research conducted at the Northeastern Collaborative Access Team beamlines, which are funded by the National Institute of General Medical Sciences from the National Institutes of Health (P30 GM124165). The Pilatus 6M detector on 24-ID-C beam line is funded by a NIH-ORIP HEI grant (S10 RR029205). 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. GM/CA@APS has been funded in whole or in part with Federal funds from the National Cancer Institute (ACB-12002) and the National Institute of General Medical Sciences (AGM-12006). 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.

Author information

Affiliations

1. Department of Pharmacology, Yale University School of Medicine, New Haven, CT, USA

Tongqing Li, Steven E. Stayrook, Yuko Tsutsui, Jianan Zhang, Hengyi Li, Stefan G. Krimmer, Mansoor Ahmed, Ian X. Walker, Krishna C. Mudumbi, Yoshihisa Suzuki, Irit Lax, Mark A. Lemmon, Joseph Schlessinger & Daryl E. Klein

2. Yale Cancer Biology Institute, Yale University, West Haven, CT, USA

Tongqing Li, Steven E. Stayrook, Yuko Tsutsui, Jianan Zhang, Yueyue Wang, Hengyi Li, Olivia Belliveau, Ian X. Walker, Krishna C. Mudumbi, Irit Lax, Mark A. Lemmon, Joseph Schlessinger & Daryl E. Klein

3. Celldex Therapeutics, New Haven, CT, USA

Andrew Proffitt & Diego Alvarado

Contributions

D.E.K. designed the overall project, with input from M.A.L. and J.S. D.E.K. wrote the manuscript assisted by T.L., with input from all authors. T.L. generated all materials (assisted by J.Z., O.B. and I.X.W.) and performed all solution biophysical studies. T.L., S.E.S. and D.E.K. analysed ALK structures. T.L. performed cell assays (assisted by Y.W. and H.L.). T.L. and Y.T. carried out HDX-MS studies, supervised by M.A.L. T.L. and K.C.M. performed fluorescence studies in membranes, supervised by M.A.L. I.L. carried out full length ALK studies with help from M.A. and Y.S. ALKAL1 AD-MBP structural studies and analysis was performed by S.G.K., supervised by J.S. Purified CDX antibodies were provided by A.P. and D.A.

Corresponding author

Correspondence to [Daryl E. Klein](#).

Ethics declarations

Competing interests

J.S. is a member of the scientific advisory board of Celldex. D.A. is an employee of Celldex Therapeutics.

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 Structural comparison of human and invertebrate ALK.

a, Size exclusion chromatography of each protein was carried out using a Superdex 75 Increase 10/300 GL column. The non-glycosylated mass was calculated from the protein sequence. The mass corresponding to each SEC peak was determined based on the log (MW) versus elution volume plot of standards for the column. The SEC peak mass of each protein is consistent with a glycosylated monomer. **b**, ALK's glycine helices project downward from the TNF- α like region. Unique to ALK, strand 10 crosses over strands 4 and 5 to terminate the fold, producing a surface terminal loop, "C-term loop". The GRD has a distinct topology, and does not form the jelly-roll characteristic of TNF- α domains **c**, The GRD of human ALK. **d**, The hexagonal array of the Pole, and **e**, the order and topology of the Pole (red circles into the page, black out of the page). **f**, *C. elegans* ALK adopts a similar overall architecture to the human GRD. However, the invertebrate PXL is structurally different, forming a β -hairpin rather than helices (**orange**). **g**, The Pole of invertebrate ALK is also smaller, with 11 glycine helices that form 2 complete hexagons. **h**, The helical strands are in the same order and topology as in human ALK. Interestingly, the missing glycine helices of the invertebrate Pole are partially matched by non PG-II loops that interact with the hexagonal array (dashed circles). The invertebrate GRD structure additionally includes a C-terminal cysteine-rich region that leads up to the transmembrane domain. This region has 10 cysteines and forms 2 EGF-like domains (**f, blue**).

Extended Data Fig. 2 Structure of invertebrate ALK's EGF-like domains.

a, The first EGF-like domain has canonical disulfide pairing. **b**, The second EGF-like domain is atypical in that it lacks the first 1-3 disulfide bond, the stabilizing role of which is replaced by hydrophobic interactions involving Y871 and F859. **c**, The EGF-like domains pack tightly to the TNF- α like domain with hydrophobic interactions. Y836 and F837 of the first EGF

domain bind to the proximal C-terminal loop. F874 of the second EGF like domain is buried in a hydrophobic cavity.

Extended Data Fig. 3 HDX analysis of ALKAL binding to ALK.

HDX-MS percent exchange butterfly plot for ALKAL2-AD binding to human ALK GRD. Each peptide is assigned a peptide ID number (Extended Data Table 2) from the N- (left) to the C-terminus (right). Each grey bar shows the sum of $\Delta\%$ Exchange at all labeling timepoints ($\Delta\%sum$) for each peptide. The dotted purple line corresponds to statistically significant $\Delta\%sum$ (16%, or ± 0.77 Da difference in deuterium uptake between unliganded and ALKAL-bound GRD) with 98 % confidence limit calculated based on the measured standard deviation of deuterium uptake for each peptide. Regions with positive or negative $\Delta\%$ Exchange become more stable or flexible, respectively, upon ALKAL binding. Statistics were derived from two independent biological repeats, each with three technical repeats. Data represent mean \pm SD.

Extended Data Fig. 4 ALK binding and signaling by ALKAL2-AD.

a, Representative BLI sensograms traces (black) and kinetic fittings (red). Fitting was carried out using ForteBio Data Analysis 10.0 software using 1:1 model with Rmax linked global fitting. Where appropriate, kinetic fit parameters are included. **b, c**, NIH 3T3 cells stably expressing wild type (WT) or mutated full length ALK, stimulated with high concentrations (10 nM) of ALKAL2 to assess residual signaling ability. **b**, Single point mutations of conserved binding-site residues. E978R has the greatest impact on ALKAL induced ALK signaling in agreement with binding data. **c**, C-terminal conserved glutamates mutated to the residues found at the same position in invertebrate ALK. The double (E974L/E978Y) mutant fails to signal, consistent with it being the only mutation shown here that completely abolished ligand binding in (a) (Extended Data Table 3). For gel source data, see Supplementary Fig. 2. The stimulation experiment was repeated three times with similar results.

Extended Data Fig. 5 Ligand binding and receptor dimerization are coupled to Pole rotation and PXL changes.

a, View looking down the long axis of the Pole with apo ALK (gray) aligned to the complex structure (color). Compared to unliganded ALK, the ligand-bound ALK dimer undergoes a clockwise rotation of the Pole about the central glycine helix (number 4). **b**, Unliganded ALK (gray) aligned to both protomers of the complex dimer (color). The PXL residues surrounding Q788 adopt a helical structure upon ligand binding and dimerization. **c**, Upon ligand binding, two additional peptides – not directly involved in ligand binding and discussed in Fig. 2 – are significantly protected. Statistics were derived from two independent biological repeats, each with three technical repeats. Data represent mean ± SD. **d**, these peptides (cyan) (peptide #1 and #21, Extended Data Fig. 3) correspond to the dimer interface including the disulfide linked helix of the PXL (#21) and the helix on the TNF- α like domain it forms a bond with (#1). **e**, The core helix-turn-helix of the AD is largely unaltered upon binding. ALKAL2 from the ALK-ALKAL fusion complex structure (green) is aligned to an un-complexed ALKAL1-AD MBP (Maltose Binding Protein) fusion (cyan).

Extended Data Fig. 6 The PXL region is necessary for ALK signaling.

a, Binding of ALKAL2-AD to GRD PXL mutations detected by BLI. The sensors were loaded with ALKAL2-AD. REQ is a mutant that alters the interface residues 795-“IGE”-797 (shown in Fig. 3b,c) to REQ. Δ PXL is a mutant that removes the entire “disulfided helix” 783-797 (shown in Fig. 3b,c). The relative binding is the binding normalized to the maximum responses. Data represent mean ± SD of four measurements. **b**, Expression of Halo-ALK on NIH/3T3 cell membranes. NIH/3T3 cells were untransfected, transfected with wildtype Halo-ALK or ALK mutants. Cells were stained with membrane impermeable dye, JF549i, before imaging. For each construct fluorescence (top row), brightfield (middle row), and a merge (bottom row) were shown. Scale bar 5 μ m. The experiment was repeated two times with similar results. **c**, Binding of CDX123 or CDX125

to GRD or GRD Δ PXL detected by BLI. Data represent mean \pm SD from three independent experiments.

Extended Data Fig. 7 Implications for GRD mutation and evolution.

a, Invertebrate ALK GRD structure with known glycine-to-acidic residue (D/E) mutations highlighted (purple spheres)⁹. **b**, Human ALK GRD structure with all of glycine-to-acidic residue mutations found in the COSMIC database (<https://cancer.sanger.ac.uk>) highlighted (purple spheres). **c**, Invertebrate ALK GRD does not bind to human ALKAL2-AD. The relative binding is the binding normalized to the maximum responses. Data represent mean \pm SD of three measurements. **d**, The EGF-like domain of invertebrate ALK occupies a region of the C-terminal loop required for ALKAL binding. This would prevent binding of any similar helix-loop-helix like ligand.

Extended Data Table 1 Data collection and refinement statistics

Extended Data Table 2 HDX peptides

Extended Data Table 3 Binding analysis of GRD and ALKAL mutants

Supplementary information

Supplementary Information

See Supplementary Information for contents

Reporting Summary

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Li, T., Stayrook, S.E., Tsutsui, Y. *et al.* Structural basis for ligand reception by anaplastic lymphoma kinase. *Nature* **600**, 148–152 (2021).
<https://doi.org/10.1038/s41586-021-04141-7>

- Received: 25 May 2021
- Accepted: 14 October 2021
- Published: 24 November 2021
- Issue Date: 02 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04141-7>

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

- [**Mechanism for the activation of the anaplastic lymphoma kinase receptor**](#)
 - Andrey V. Reshetnyak
 - Paolo Rossi
 - Charalampos G. Kalodimos

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04141-7>

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

- Article
- [Published: 24 November 2021](#)

Mechanism for the activation of the anaplastic lymphoma kinase receptor

- [Andrey V. Reshetnyak](#) ORCID: [orcid.org/0000-0003-1751-3964^{1 na1}](https://orcid.org/0000-0003-1751-3964),
- [Paolo Rossi^{1 na1}](#),
- [Alexander G. Myasnikov](#) ORCID: [orcid.org/0000-0003-2607-7121^{1 na1}](https://orcid.org/0000-0003-2607-7121),
- [Munia Sowaileh¹](#),
- [Jyotidarsini Mohanty²](#),
- [Amanda Nourse](#) ORCID: [orcid.org/0000-0003-4595-5321¹](https://orcid.org/0000-0003-4595-5321),
- [Darcie J. Miller¹](#),
- [Irit Lax²](#),
- [Joseph Schlessinger](#) ORCID: [orcid.org/0000-0002-5085-5969²](https://orcid.org/0000-0002-5085-5969) &
- [Charalampos G. Kalodimos](#) ORCID: [orcid.org/0000-0001-6354-2796¹](https://orcid.org/0000-0001-6354-2796)

[Nature](#) volume 600, pages 153–157 (2021)

- 3059 Accesses
- 59 Altmetric
- [Metrics details](#)

Subjects

- [Cryoelectron microscopy](#)

- [Kinases](#)
- [Oncogene proteins](#)

Abstract

Anaplastic lymphoma kinase (ALK) is a receptor tyrosine kinase (RTK) that regulates important functions in the central nervous system^{1,2}. The *ALK* gene is a hotspot for chromosomal translocation events that result in several fusion proteins that cause a variety of human malignancies³. Somatic and germline gain-of-function mutations in ALK were identified in paediatric neuroblastoma^{4,5,6,7}. ALK is composed of an extracellular region (ECR), a single transmembrane helix and an intracellular tyrosine kinase domain^{8,9}. ALK is activated by the binding of ALKAL1 and ALKAL2 ligands^{10,11,12,13,14} to its ECR, but the lack of structural information for the ALK-ECR or for ALKAL ligands has limited our understanding of ALK activation. Here we used cryo-electron microscopy, nuclear magnetic resonance and X-ray crystallography to determine the atomic details of human ALK dimerization and activation by ALKAL1 and ALKAL2. Our data reveal a mechanism of RTK activation that allows dimerization by either dimeric (ALKAL2) or monomeric (ALKAL1) ligands. This mechanism is underpinned by an unusual architecture of the receptor–ligand complex. The ALK-ECR undergoes a pronounced ligand-induced rearrangement and adopts an orientation parallel to the membrane surface. This orientation is further stabilized by an interaction between the ligand and the membrane. Our findings highlight the diversity in RTK oligomerization and activation mechanisms.

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: Structural features of ALK-ECR^{ABR} and ALKAL2.

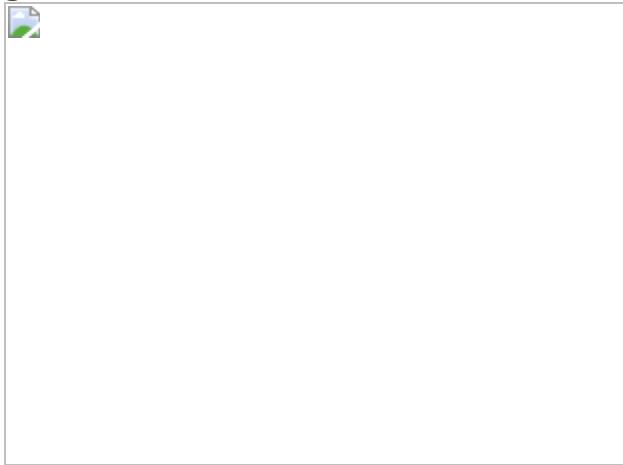


Fig. 2: Cryo-EM structure of the heterotetrameric ALK-ECR^{ABR}-ALKAL2 complex.

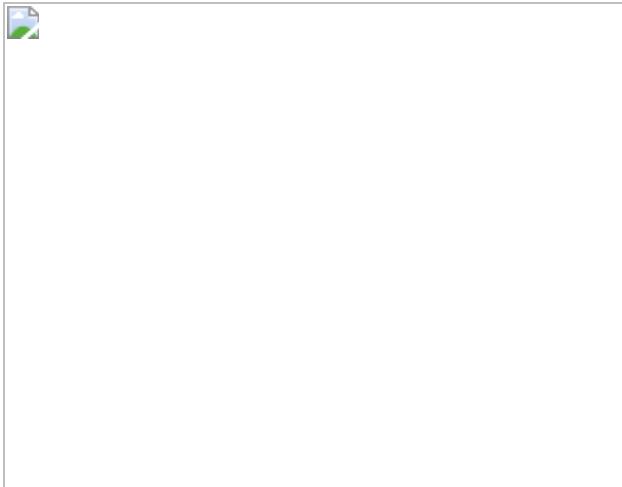


Fig. 3: Structural features of the ALK-ECR^{ABR}-ALKAL2 tetramerization interface.

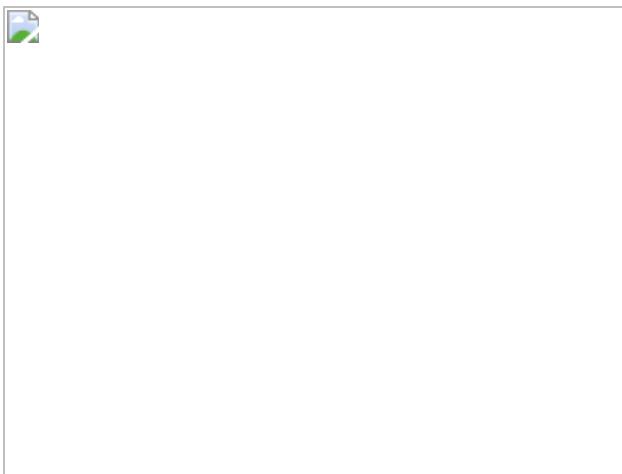


Fig. 4: ALKAL2-mediated receptor dimerization.



Data availability

Atomic coordinates have been deposited in the Protein Data Bank (PDB) with accession codes [7N00](#) (cryoEM structure of ALK-ECR^{ABR}-ALKAL2), [7MZY](#) (X-ray structure of ALK-ECR^{ABR}(ΔEGF)), [7MZW](#) (NMR structure of ALK-ECR^{ABR}), [7MZX](#) (NMR structure of ALKAL2)

and [7MZZ](#) (NMR structure of ALKAL1). The cryo-EM density map of ALK-ECR-ALKAL2 has been deposited in the Electron Microscopy Data Bank (EMDB) under the accession code [EMD-24095](#). NMR resonance assignments have been deposited in the Biological Magnetic Resonance Bank (BMRB) under the accession codes [30910](#) (ALK-ECR^{ABR}), [30911](#) (ALKAL2) and [30912](#) (ALKAL1).

References

1. 1.
Morris, S. W. et al. Fusion of a kinase gene, *ALK*, to a nucleolar protein gene, *NPM*, in non-Hodgkin's lymphoma. *Science* **263**, 1281–1284 (1994).
2. 2.
Orthofer, M. et al. Identification of ALK in thinness. *Cell* **181**, 1246–1262 (2020).
3. 3.
Hallberg, B. & Palmer, R. H. Mechanistic insight into ALK receptor tyrosine kinase in human cancer biology. *Nat. Rev. Cancer* **13**, 685–700 (2013).
4. 4.
Chen, Y. et al. Oncogenic mutations of ALK kinase in neuroblastoma. *Nature* **455**, 971–974 (2008).
5. 5.
George, R. E. et al. Activating mutations in ALK provide a therapeutic target in neuroblastoma. *Nature* **455**, 975–978 (2008).
6. 6.

Janoueix-Lerosey, I. et al. Somatic and germline activating mutations of the ALK kinase receptor in neuroblastoma. *Nature* **455**, 967–970 (2008).

7. 7.

Mosse, Y. P. et al. Identification of *ALK* as a major familial neuroblastoma predisposition gene. *Nature* **455**, 930–935 (2008).

8. 8.

Morris, S. W. et al. *ALK*, the chromosome 2 gene locus altered by the t(2;5) in non-Hodgkin's lymphoma, encodes a novel neural receptor tyrosine kinase that is highly related to leukocyte tyrosine kinase (LTK). *Oncogene* **14**, 2175–2188 (1997).

9. 9.

Lemmon, M. A. & Schlessinger, J. Cell signaling by receptor tyrosine kinases. *Cell* **141**, 1117–1134 (2010).

10. 10.

Zhang, H. et al. Deorphanization of the human leukocyte tyrosine kinase (LTK) receptor by a signaling screen of the extracellular proteome. *Proc. Natl Acad. Sci. USA* **111**, 15741–15745 (2014).

11. 11.

Guan, J. et al. FAM150A and FAM150B are activating ligands for anaplastic lymphoma kinase. *eLife* **4**, e09811 (2015).

12. 12.

Reshetnyak, A. V. et al. Augmentor α and β (FAM150) are ligands of the receptor tyrosine kinases ALK and LTK: hierarchy and specificity of ligand–receptor interactions. *Proc. Natl Acad. Sci. USA* **112**, 15862–15867 (2015).

13. 13.

Mo, E. S., Cheng, Q., Reshetnyak, A. V., Schlessinger, J. & Nicoli, S. Alk and Ltk ligands are essential for iridophore development in zebrafish mediated by the receptor tyrosine kinase Ltk. *Proc. Natl Acad. Sci. USA* **114**, 12027–12032 (2017).

14. 14.

Fadeev, A. et al. ALKALs are in vivo ligands for ALK family receptor tyrosine kinases in the neural crest and derived cells. *Proc. Natl Acad. Sci. USA* **115**, E630–E638 (2018).

15. 15.

Reshetnyak, A. V. et al. Identification of a biologically active fragment of ALK and LTK-ligand 2 (augmentor- α). *Proc. Natl Acad. Sci. USA* **115**, 8340–8345 (2018).

16. 16.

Crick, F. H. & Rich, A. Structure of polyglycine II. *Nature* **176**, 780–781 (1955).

17. 17.

Warkentin, E. et al. A rare polyglycine type II-like helix motif in naturally occurring proteins. *Proteins* **85**, 2017–2023 (2017).

18. 18.

Loren, C. E. et al. A crucial role for the Anaplastic lymphoma kinase receptor tyrosine kinase in gut development in *Drosophila melanogaster*. *EMBO Rep.* **4**, 781–786 (2003).

19. 19.

Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

20. 20.

Endres, N. F. et al. Conformational coupling across the plasma membrane in activation of the EGF receptor. *Cell* **152**, 543–556 (2013).

21. 21.

Myers, K. V., Amend, S. R. & Pienta, K. J. Targeting Tyro3, Axl and MerTK (TAM receptors): implications for macrophages in the tumor microenvironment. *Mol. Cancer* **18**, 94 (2019).

22. 22.

Grasberger, B., Minton, A. P., DeLisi, C. & Metzger, H. Interaction between proteins localized in membranes. *Proc. Natl Acad. Sci. USA* **83**, 6258–6262 (1986).

23. 23.

Klein, P., Mattoon, D., Lemmon, M. A. & Schlessinger, J. A structure-based model for ligand binding and dimerization of EGF receptors. *Proc. Natl Acad. Sci. USA* **101**, 929–934 (2004).

24. 24.

Kuriyan, J. & Eisenberg, D. The origin of protein interactions and allostery in colocalization. *Nature* **450**, 983–990 (2007).

25. 25.

Diwanji, D., Thaker, T. & Jura, N. More than the sum of the parts: toward full-length receptor tyrosine kinase structures. *IUBMB Life* **71**, 706–720 (2019).

26. 26.

Murray, P. B. et al. Heparin is an activating ligand of the orphan receptor tyrosine kinase ALK. *Sci. Signal.* **8**, ra6 (2015).

27. 27.

Li, T. et al. Structural basis for ligand reception by anaplastic lymphoma kinase. *Nature* <https://doi.org/10.1038/s41586-021-04141-7> (2021).

28. 28.

Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L. & Clardy, J. Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J. Mol. Biol.* **229**, 105–124 (1993).

29. 29.

Monneau, Y. R. et al. Exploiting *E. coli* auxotrophs for leucine, valine, and threonine specific methyl labeling of large proteins for NMR applications. *J. Biomol. NMR* **65**, 99–108 (2016).

30. 30.

Rossi, P., Monneau, Y. R., Xia, Y., Ishida, Y. & Kalodimos, C. G. Toolkit for NMR studies of methyl-labeled proteins. *Methods Enzymol.* **614**, 107–142 (2019).

31. 31.

Xie, T., Saleh, T., Rossi, P. & Kalodimos, C. G. Conformational states dynamically populated by a kinase determine its function. *Science* **370**, eabc2754 (2020).

32. 32.

Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).

33. 33.

Hendrickson, W. A. & Ogata, C. M. Phase determination from multiwavelength anomalous diffraction measurements. *Methods Enzymol.* **276**, 494–523 (1997).

34. 34.

Terwilliger, T. C. et al. Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. *Acta Crystallogr. D* **65**, 582–601 (2009).

35. 35.

Liebschner, D. et al. Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallogr. D* **75**, 861–877 (2019).

36. 36.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).

37. 37.

Delaglio, F. et al. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).

38. 38.

Lee, W., Tonelli, M. & Markley, J. L. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, 1325–1327 (2015).

39. 39.

Lange, O. F. et al. Determination of solution structures of proteins up to 40 kDa using CS-Rosetta with sparse NMR data from deuterated samples. *Proc. Natl Acad. Sci. USA* **109**, 10873–10878 (2012).

40. 40.

Shen, Y. & Bax, A. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56**, 227–241 (2013).

41. 41.

Rossi, P., Xia, Y., Khanra, N., Veglia, G. & Kalodimos, C. G. ^{15}N and ^{13}C - SOFAST-HMQC editing enhances 3D-NOESY sensitivity in highly deuterated, selectively [$^1\text{H}, ^{13}\text{C}$]-labeled proteins. *J. Biomol. NMR* **66**, 259–271 (2016).

42. 42.

Monneau, Y. R. et al. Automatic methyl assignment in large proteins by the MAGIC algorithm. *J. Biomol. NMR* **69**, 215–227 (2017).

43. 43.

Guntert, P. Automated NMR structure calculation with CYANA. *Methods Mol. Biol.* **278**, 353–378 (2004).

44. 44.

Linge, J. P., Williams, M. A., Spronk, C. A., Bonvin, A. M. & Nilges, M. Refinement of protein structures in explicit solvent. *Proteins* **50**, 496–506 (2003).

45. 45.

Brunger, A. T. Version 1.2 of the Crystallography and NMR system. *Nat. Protoc.* **2**, 2728–2733 (2007).

46. 46.

Bhattacharya, A., Tejero, R. & Montelione, G. T. Evaluating protein structures determined by structural genomics consortia. *Proteins* **66**, 778–795 (2007).

47. 47.

Tejero, R., Snyder, D., Mao, B., Aramini, J. M. & Montelione, G. T. PDBStat: a universal restraint converter and restraint analysis software package for protein NMR. *J. Biomol. NMR* **56**, 337–351 (2013).

48. 48.

Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).

49. 49.

Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).

50. 50.

Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).

51. 51.

Grant, T., Rohou, A. & Grigorieff, N. *cisTEM*, user-friendly software for single-particle image processing. *eLife* **7**, e35383 (2018).

52. 52.

Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).

53. 53.

Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

54. 54.

Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D* **74**, 531–544 (2018).

55. 55.

Williams, C. J. et al. MolProbity: more and better reference data for improved all-atom structure validation. *Protein Sci.* **27**, 293–315 (2018).

56. 56.

Kortt, A. A., Nice, E. & Gruen, L. C. Analysis of the binding of the Fab fragment of monoclonal antibody NC10 to influenza virus N9 neuraminidase from tern and whale using the BIACore biosensor: effect of immobilization level and flow rate on kinetic analysis. *Anal. Biochem.* **273**, 133–141 (1999).

57. 57.

Zhao, H., Brautigam, C. A., Ghirlando, R. & Schuck, P. Overview of current methods in sedimentation velocity and sedimentation equilibrium analytical ultracentrifugation. *Curr. Protoc. Protein Sci.* **71**, 20.12.1–20.12.49 (2013).

58. 58.

Schuck, P. Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys. J.* **78**, 1606–1619 (2000).

59. 59.

Zhao, H. et al. A multilaboratory comparison of calibration accuracy and the performance of external references in analytical ultracentrifugation. *PLoS ONE* **10**, e0126420 (2015).

60. 60.

Brautigam, C. A. Calculations and publication-quality illustrations for analytical ultracentrifugation data. *Methods Enzymol.* **562**, 109–133 (2015).

61. 61.

Folta-Stogniew, E. & Williams, K. R. Determination of molecular masses of proteins in solution: implementation of an HPLC size exclusion chromatography and laser light scattering service in a core laboratory. *J. Biomol. Tech.* **10**, 51–63 (1999).

Acknowledgements

This work was supported by the American Lebanese Syrian Associated Charities (ALSAC) and NIH grant R35 GM122462. Cryo-EM and NMR data were acquired at the St. Jude Cryo-EM Center and Biomolecular NMR Spectroscopy Center, respectively. X-ray data were collected at SERCAT ID and BM beamlines. We thank M. Rana and R. Kalathur for their help with protein expression and purification; M. Clay for assisting with NMR graphics; I. Chen for critical reading of the manuscript; E. Folta-Stogniew for discussions of SEC–MALS data; and Z. Luo for assisting with the preparation of figures and the animation.

Author information

Author notes

1. These authors contributed equally: Andrey V. Reshetnyak, Paolo Rossi, Alexander G. Myasnikov

Affiliations

1. Department of Structural Biology, St. Jude Children's Research Hospital, Memphis, TN, USA

Andrey V. Reshetnyak, Paolo Rossi, Alexander G. Myasnikov, Munia Sowaileh, Amanda Nourse, Darcie J. Miller & Charalampos G. Kalodimos

2. Department of Pharmacology, Yale School of Medicine, New Haven, CT, USA

Jyotidarsini Mohanty, Irit Lax & Joseph Schlessinger

Contributions

A.V.R. and C.G.K. supervised the project and wrote the manuscript. P.R., J.S. and A.G.M. contributed to the final version of the manuscript. A.V.R. and M.S. purified and prepared protein samples for NMR, X-ray, cryo-EM and biophysical studies. P.R. and M.S. performed NMR data collection, processing and structure calculations. A.V.R. and D.J.M. performed data collection and processing and solved the X-ray structures. A.V.R. and A.G.M. screened cryo-EM grids and performed cryo-EM data collection and processing. A.V.R. built atomic models for X-ray and cryo-EM studies and performed SEC–MALS experiments. A.V.R. and A.N. performed SV–AUC experiments and analysed data. I.L. and J.M. performed ALK cell-based phosphorylation assays. C.G.K. acquired funding and resources for the study.

Corresponding authors

Correspondence to [Joseph Schlessinger](#) or [Charalampos G. Kalodimos](#).

Ethics declarations

Competing interests

C.G.K. is a consultant for BridgeBio.

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 Structural features of ALK-ECR^{ABR}.

a, ^1H - ^{15}N -correlated (left panel) and ^1H - ^{13}C -correlated (right panel) spectra of [U- ^2H , ^{15}N ; Ala- $^{13}\text{CH}_3$; Met- $^{13}\text{CH}_3$; Ile- $\delta 1$ - $^{13}\text{CH}_3$; Leu, Val- $^{13}\text{CH}_3$ / $^{13}\text{CH}_3$; Thr- $^{13}\text{CH}_3$]-labelled ALK-ECR^{ABR}. **b**, ^1H - ^{13}C -correlated spectra of [U- ^2H ; Phe- δ - ^{13}CH ; Tyr- ϵ - ^{13}CH]-labelled ALK-ECR^{ABR}. **c**, Select strips from ^{13}C -edited NOESY experiments highlighting intra-domain NOEs between TNF-like and EGF-like. **d**, Close-up view of the TNF-like–EGF-like interface and close-range inter-proton contacts (within ~7 Å) observed from the NOESY NMR analysis of spectra shown in (c). **e**, Asymmetric unit content of ALK-ECR^{ABR}(ΔEGF) crystals. **f**, 2Fo-Fc map of ALK-ECR^{ABR}(ΔEGF) depicted at 0.951 contour level for chain A. **g**, B-factors of ALK-ECR^{ABR}(ΔEGF) (chain A) are mapped on its structure. The tube radius is proportional to the B-factor. Low-to-high B-factors are also denoted in a blue-to-red colour gradient. **h**, Topology diagram of ALK-ECR^{ABR}(ΔEGF). PG_{II} helices are shown in green tubes, β-strands in arrows and α helices in cylinders. **i**, Superposition of the following structures: ALK-ECR^{ABR} GlyR (grey, this work), glycine rich domain of GTP-binding protein Obg (red, PDB ID 1LNZ), acetophenone carboxylase (blue, PDB ID 5L9W), antifreeze protein (yellow, PDB ID 3BOG), and gp38 bacteriophage adhesin tip (green, PDB ID 6F45). **j**, Schematic representation of the GlyR PG_{II} array shown in a top view. Solid hexagons denote PGII helices with an N-to-C direction towards the reader whereas open hexagons denote PGII helices in the opposite direction. **k**, Side view of the GlyR PG_{II} array in ball-and-stick representation. **l**, Top view of GlyR

PG_{II} array showing inter-chain hydrogen bond network (grey dashes). **m**, Close-up view of the TNF-like–EGF-like interface.

Extended Data Fig. 2 ALK and LTK sequence comparison and structural characterization of ALKAL2.

a, Schematic representation of domain organization for human ALK and LTK receptors (left panel). Sequence alignment of ALK and LTK ECRs (right panel). The secondary structure diagram is shown based on the ALK-ECR^{ABR} structure determined in this work. Cys residues are coloured red and disulfide bridges are shown with red lines. Residues participating in ALKAL2 binding are coloured magenta, and residues participating in inter-protomer dimerization are underlined and coloured blue (contacts with TNF-like) and grey (contacts with THB). The key residues involved in ALKAL2 binding are conserved between ALK and LTK (highlighted with magenta in LTK sequence), with the exception of F143, S260, L361, Q362, A365, T367, E374, R376, D388, Q390 and L401 (LTK numbering, underlined with black lines). The difference in ALKAL2/1 specificity might be explained by the H120/Y99 and D124/E103 substitutions (ALKAL2/1 respectively) and/or difference in the receptor-receptor dimerization interface. **b**, SEC–MALS (upper panel) analysis of ALKAL2-AD at eluted concentrations of ~ 67 μM (red), 11 μM (blue) and 1 μM (magenta). Molecular masses in kDa determined by in-line MALS (left axis) are included. SV–AUC profile of ALKAL2-AD (lower panel). Concentrations used are: 233.53 (purple), 111.25 (blue) and 55.62 μM (cyan). **c**, ¹H-¹⁵N–correlated (left panel) and ¹H-¹³C–correlated (right panel) NMR spectra of ALKAL2-AD. **d**, NMR ensemble of the 20 lowest-energy conformers of ALKAL2-AD. **e**, Electrostatic surface representation of ALKAL2-AD. The electrostatic potential is measured in eV, with range as shown in the corresponding colour bar (from -5.000 to +5.000 eV). **f**, NMR ensemble of the 20 lowest-energy conformers of ALKAL1-AD. **g**, SV–AUC profile and sedimentation coefficient distribution model c(s) of MBP–ALKAL2(C66Y) (left panel). Concentrations used are: 96.9 μM (purple), 48.5 μM (blue), 29.4 μM (cyan), 12.1 μM (green), 6.9 μM (yellow), 3.5 μM (orange) and 1.7 μM (red). Isotherm of the signal-weight-average s-values (sw) for MBP–ALKAL2(C66Y) obtained by integration of c(s) distributions over

the s-range of 2.5 and 5 S for each loading concentration in a dilution series (right panel). The confidence intervals of the fits are presented in the lower panel. **h**, Superposition of NMR-solved and AlphaFold-predicted structures of ALKAL2 (left panel) and ALKAL1 (right panel). AD and variable region (VR) regions are labelled. **i**, SV–AUC profile of ALK-ECR^{ABR}–ALKAL2-AD. Concentrations used are: 177.8 (purple), 87.87 (blue), 43.93 (cyan), 20.92 (green), 10.46 (yellow), 5.44 (orange) and 2.72 μ M (red). **j**, SV–AUC profile and sedimentation coefficient distribution model c(s) of MBP–ALKAL1. The highest (96.6 μ M - purple) and lowest (4.8 μ M - orange) concentrations are shown. **k**, SEC–MALS profile for the ALK-ECR^{ABR}–MBP–ALKAL1 complex (blue, theoretical mass of 91 kDa for 1:1 complex). The profile for the ALK-ECR^{ABR}–MBP–ALKAL2(C66Y) complex (red, theoretical mass of 189 kDa for 2:2 complex) is included for direct comparison. **l**, Sedimentation velocity analytical ultracentrifugation profile of ALK-ECR^{ABR}–MBP–ALKAL1 (loading concentration 74 μ M).

Extended Data Fig. 3 Characterization of the binding mode and oligomeric state of the ALK-ECR^{ABR}–ALKAL2 complex.

a, SEC–MALS profiles of the ALK-ECR^{ABR}–ALKAL2(C66Y) complex. The corresponding isotherm of the signal-weight-average MW as a function of concentration is shown in the right panel. Standard errors were determined according to a previous study⁶¹. **b**, Sedimentation velocity analytical ultracentrifugation isotherm of the signal-weight-average s-values for ALK-ECR^{ABR}–ALKAL2(C66Y) complex. **c**, SEC–MALS profile (red) of ALK-ECR^{ABR}–MBP–ALKAL2(C66Y) complex mixed at 2:1 ratio. SDS/PAGE for corresponding fractions are shown in the right panel. Bands corresponding to ALK-ECR^{ABR} or MBP–ALKAL2(C66Y) are labelled, position of molecular weight markers are indicated. **d, e**, ALK autophosphorylation assays. **d**, ALK variants (as indicated in the labels) stably expressed in NIH/3T3 cells were stimulated with 10 nM of purified WT ALKAL2-AD. **e**, Wild type ALK stably expressed in NIH/3T3 cells was stimulated with 10 nM of purified ALKAL2-AD variants. Cell lysates were subjected to immunoprecipitation using anti-ALK antibodies followed by SDS/PAGE and immunoblotting with anti-pTyr (pY) and anti-ALK (ALK) antibodies. Relative position of the band for 180-kDa Mw marker is

shown. **f**, SEC–MALS profiles of ALK-ECR^{ABR}(4M):ALKAL2(2M) complex. ALK-ECR^{ABR}(4M) stands for T686A/N787A/Q788A/I795A mutations in ALK-ECR^{ABR}, and ALKAL2(2M) for full-size ALKAL2(C66Y) containing I127A/Y130A mutations. The corresponding isotherm of the signal-weight-average MW as a function of concentration is shown in the right panel. Standard errors were determined according to⁶¹. **g**, Sedimentation velocity analytical ultracentrifugation isotherm of the signal-weight-average s-values for ALK-ECR^{ABR}(4M)–ALKAL2(2M) complex.

Extended Data Fig. 4 Cryo-EM data processing workflow of ALK-ECR^{ABR}–ALKAL2(C66Y) and evaluation of the reconstruction.

Stages of cryo-EM data preprocessing, image classification and refinement for ALK-ECR^{ABR}–ALKAL2(C66Y) complex. Local resolution maps and fourier shell correlation are presented for final cryo-EM structures.

Extended Data Fig. 5 ALK-ECR^{ABR} rearrangements after ALKAL2 binding.

a, d, e, Superposition of ALKAL2 (**a**), GlyR-TNF-like (**d**) and EGF-like (**e**) structures in unliganded (pink) and liganded (blue) states. **b**, ALKAL2-induced repositioning of ALK-ECR^{ABR}. **c**, Cartoon representation of the heterotetrameric ALK-ECR^{ABR}–ALKAL2 complex wherein the 13-residue-long linker tethering EGF-like to TMH has been modelled in an extended conformation. The modelling shows that if EGF-like did not change its position upon ALKAL2 binding (EGF-like unliganded position shown in dark red) TMH dimerization would not be possible because the linker is too short. The model for the linker was manually built in Coot and follows the shortest possible path to reach the TMH. **f**, Superposition of EGF-like structures in the unliganded (orange) and ALKAL2-bound (grey) states demonstrates the conformational changes in EGF-like between the two states. **g**, Residues at the interface between EGF-like in unliganded state and TNF-like (left panel) and between EGF-like in ALKAL2-bound state and ALKAL2 (right panel). **i**, Wild type trx–ALKAL2-AD (grey) and trx–

ALKAL2-AD variants (magenta) were tested for their ability to bind wild type ALK-ECR^{ABR} using BLI. Steady-state dissociation constants and standard errors were determined according to a previous study⁵⁶. **j**, Comparison of tyrosine autophosphorylation of WT ALK stimulated by 10 nM of purified ALKAL2 variants as indicated. ALKAL2-AD^{RC} stands for mutation of four charged residues - K94E/K96E/K99E/H100E; Trx-ALKAL2-AD is N-terminal fusion of ALKAL2 with thioredoxin, and ALKAL2-AD^{Δɑ1} is deletion of ɑ1 helix (residue boundaries 103-152). NIH3T3 cells stably expressing WT ALK were lysed after ALKAL2 stimulation and were subjected to immunoprecipitation using anti-ALK antibodies followed by SDS/PAGE and immunoblotting with anti-pTyr (pY) and anti-ALK (ALK) antibodies. Relative position of the band for 180-kDa Mw marker is shown. **h**, Close-up views of residues (CA atoms shown) D732, H996 and T733, M997 is shown. These residues were mutated to Cys for the cross-linking experiments.

Extended Data Fig. 6 NMR analysis of ALK-ECR^{ABR}-ALKAL2-AD and ALK-ECR^{ABR}-ALKAL1-AD complexes.

a, b, Superimposed ¹H-¹³C-correlated spectra of ALK-ECR^{ABR}-ALKAL2-AD (**a**) and ALK-ECR^{ABR}-ALKAL1-AD (**b**) complexes. ALK-ECR^{ABR} and ALKAL proteins are ¹H-¹³C labelled in the indicated methyl groups. **c, d**, Chemical shift perturbation induced by ALKAL1-AD binding to ALK-ECR^{ABR} to combined ¹H and ¹⁵N amide atoms (**c**) and ¹H and ¹³C methyl atoms (**d**). **e, f**, Chemical shift perturbation induced by ALKAL1-AD mapped onto the ALK-ECR^{ABR} structure. **g**, NMR characterization of ALKAL1-AD binding to ALK-ECR^{ABR}. Select strips from ¹³C-edited NOESY experiments showing intermolecular NOEs between ALK-ECR^{ABR} and ALKAL1-AD. Similar results were obtained when ALKAL2-AD was used, confirming that the structure observed in the frozen sample used in cryo-EM is the same in solution. **h**, NMR characterization of the EGF-like domain repositioning upon ALKAL1 binding to ECR^{ABR}. Select strips from ¹³C-edited NOESY experiments for ALK-ECR^{ABR} showing interdomain NOEs in the unbound form. Characteristic NOE patterns between Met997 of the EGF-like domain and the indicated residues of the

TNF-like domain (right panel) changed dramatically upon ligand binding and demonstrate pronounced re-orientation of the EGF-like domain as shown schematically on the right panel.

Extended Data Table 1 NMR and structure statistics

Extended Data Table 2 Cryo-EM data collection and refinement statistics

Extended Data Table 3 ALK-ECR^{ABR}–ALKAL2 interface contacts

Extended Data Table 4 X-ray data collection and refinement statistics for ALK-ECR^{ABR}(ΔEGF) (aa 673–986)

Supplementary information

Supplementary Information

This file contains Supplementary Figures 1–3 and their accompanying legends.

Reporting Summary

Peer Review File

ALKAL-induced activation of ALK receptor

Supplementary Video 1 . The video demonstrates how ALK is activated by monomeric ALKAL ligands (ALKAL1 or ALKAL1/2-AD).

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Reshetnyak, A.V., Rossi, P., Myasnikov, A.G. *et al.* Mechanism for the activation of the anaplastic lymphoma kinase receptor. *Nature* **600**, 153–157 (2021). <https://doi.org/10.1038/s41586-021-04140-8>

- Received: 23 May 2021
- Accepted: 14 October 2021
- Published: 24 November 2021
- Issue Date: 02 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04140-8>

Share this article

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

[Get shareable link](#)

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

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04140-8>

- Article
- [Published: 24 November 2021](#)

Aldehyde-driven transcriptional stress triggers an anorexic DNA damage response

- [Lee Mulderrig](#)^{1,2} na1,
- [Juan I. Garaycoechea](#)³ na1,
- [Zewen K. Tuong](#) [ORCID: orcid.org/0000-0002-6735-6808](#)^{4,5},
- [Christopher L. Millington](#)¹,
- [Felix A. Dingler](#) [ORCID: orcid.org/0000-0002-6297-2986](#)¹,
- [John R. Ferdinand](#) [ORCID: orcid.org/0000-0003-0936-0128](#)⁴,
- [Liam Gaul](#) [ORCID: orcid.org/0000-0002-3159-8127](#)⁶,
- [John A. Tadross](#) [ORCID: orcid.org/0000-0002-8424-1252](#)^{7,8},
- [Mark J. Arends](#) [ORCID: orcid.org/0000-0002-6826-8770](#)⁹,
- [Stephen O'Rahilly](#)⁷,
- [Gerry P. Crossan](#) [ORCID: orcid.org/0000-0002-7974-009X](#)²,
- [Menna R. Clatworthy](#)^{4,5,10} &
- [Ketan J. Patel](#) ^{1,2}

Nature volume **600**, pages 158–163 (2021)

- 5019 Accesses
- 48 Altmetric
- [Metrics details](#)

Subjects

- [DNA damage response](#)
- [DNA metabolism](#)
- [Kidney](#)
- [Transcription](#)

Abstract

Endogenous DNA damage can perturb transcription, triggering a multifaceted cellular response that repairs the damage, degrades RNA polymerase II and shuts down global transcription^{1,2,3,4}. This response is absent in the human disease Cockayne syndrome, which is caused by loss of the Cockayne syndrome A (CSA) or CSB proteins^{5,6,7}. However, the source of endogenous DNA damage and how this leads to the prominent degenerative features of this disease remain unknown. Here we find that endogenous formaldehyde impedes transcription, with marked physiological consequences. Mice deficient in formaldehyde clearance (*Adh5*^{-/-}) and CSB (*Csb*^{m/m}; *Csb* is also known as *Ercc6*) develop cachexia and neurodegeneration, and succumb to kidney failure, features that resemble human Cockayne syndrome. Using single-cell RNA sequencing, we find that formaldehyde-driven transcriptional stress stimulates the expression of the anorexiogenic peptide GDF15 by a subset of kidney proximal tubule cells. Blocking this response with an anti-GDF15 antibody alleviates cachexia in *Adh5*^{-/-} *Csb*^{m/m} mice. Therefore, CSB provides protection to the kidney and brain against DNA damage caused by endogenous formaldehyde, while also suppressing an anorexic endocrine signal. The activation of this signal might contribute to the cachexia observed in Cockayne syndrome as well as chemotherapy-induced anorectic weight loss. A plausible evolutionary purpose for such a response is to ensure aversion to genotoxins in food.

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: Formaldehyde causes transcriptional stress.

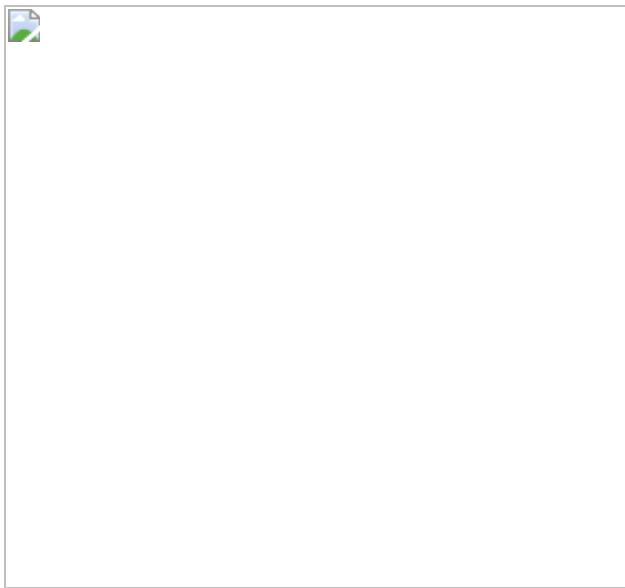


Fig. 2: Endogenous formaldehyde accumulation reveals Cockayne syndrome.

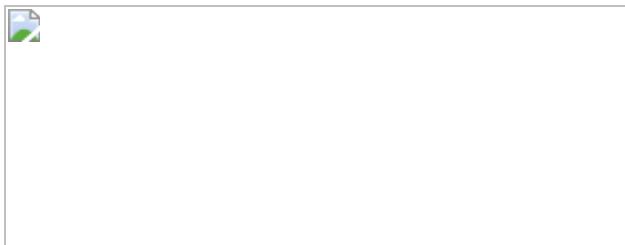


Fig. 3: Single cell RNA-seq identifies regions in the nephron that are susceptible to transcriptional stress.

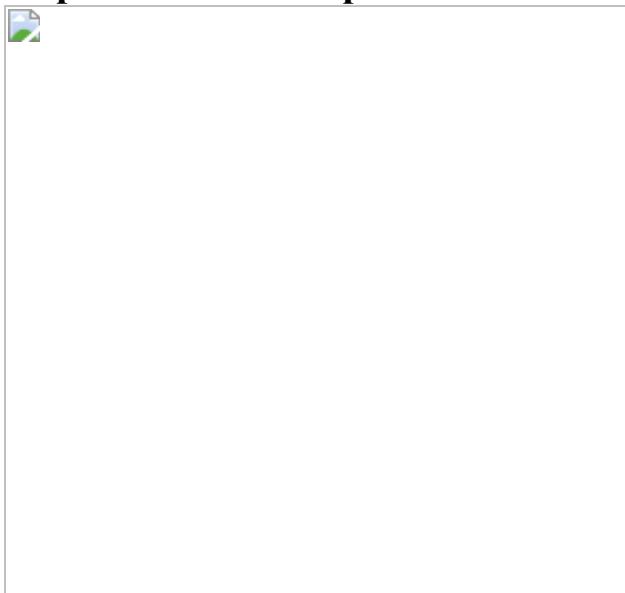


Fig. 4: Transcriptional stress damages a subset of proximal tubule cells that express GDF15.

Fig. 5: GDF15 mediates DNA damage induced cachexia.

Data availability

All scRNA-seq data have been deposited in the Gene Expression Omnibus under accession [GSE175792](#).

References

1. 1.

Brueckner, F., Hennecke, U., Carell, T. & Cramer, P. CPD damage recognition by transcribing RNA polymerase II. *Science* **315**, 859–862 (2007).

2. 2.

Xu, J. et al. Structural basis for the initiation of eukaryotic transcription-coupled DNA repair. *Nature* **551**, 653–657 (2017).

3. 3.

Gregersen, L. H. & Svejstrup, J. Q. The cellular response to transcription-blocking DNA damage. *Trends Biochem. Sci.* **43**, 327–341 (2018).

4. 4.

Tufegdžić Vidaković, A. et al. Regulation of the RNAPII pool is integral to the DNA damage response. *Cell* **180**, 1245–1261.e21 (2020).

5. 5.

Laugel, V. Cockayne syndrome: the expanding clinical and mutational spectrum. *Mech. Ageing Dev.* **134**, 161–170 (2013).

6. 6.

Stern-Delfils, A. et al. Renal disease in Cockayne syndrome. *Eur. J. Med. Genet.* **63**, 103612 (2020).

7. 7.

Kubota, M. et al. Nationwide survey of Cockayne syndrome in Japan: incidence, clinical course and prognosis. *Pediatr. Int.* **57**, 339–347 (2015).

8. 8.

Lans, H., Hoeijmakers, J. H. J., Vermeulen, W. & Marteijn, J. A. The DNA damage response to transcription stress. *Nat. Rev. Mol. Cell Biol.* **20**, (2019).

9. 9.

Wilson, M. D., Harreman, M. & Svejstrup, J. Q. Ubiquitylation and degradation of elongating RNA polymerase II: the last resort. *Biochim. Biophys. Acta* **1829**, 151–157 (2013).

10. 10.

van der Horst, G. T. et al. Defective transcription-coupled repair in Cockayne syndrome B mice is associated with skin cancer predisposition. *Cell* **89**, 425–435 (1997).

11. 11.

van der Horst, G. T. J. et al. UVB radiation-induced cancer predisposition in Cockayne syndrome group A (*Csa*) mutant mice. *DNA Repair* **1**, 143–157 (2002).

12. 12.

Pontel, L. B. et al. Endogenous formaldehyde is a hematopoietic stem cell genotoxin and metabolic carcinogen. *Mol. Cell* **60**, 177–188 (2015).

13. 13.

Cheng, G. et al. Reactions of formaldehyde plus acetaldehyde with deoxyguanosine and DNA: formation of cyclic deoxyguanosine adducts and formaldehyde cross-links. *Chem. Res. Toxicol.* **16**, 145–152 (2003).

14. 14.

Marteijn, J. A., Lans, H., Vermeulen, W. & Hoeijmakers, J. H. J. Understanding nucleotide excision repair and its roles in cancer and ageing. *Nat. Rev. Mol. Cell Biol.* **15**, 465–481 (2014).

15. 15.

Ranes, M. et al. A ubiquitylation site in Cockayne syndrome B required for repair of oxidative DNA damage, but not for transcription-

coupled nucleotide excision repair. *Nucleic Acids Res.* **44**, 5246–5255 (2016).

16. 16.

Wei, W. et al. S-nitrosylation from GSNOR deficiency impairs DNA repair and promotes hepatocarcinogenesis. *Sci. Transl. Med.* **2**, 19ra13 (2010).

17. 17.

Niedernhofer, L. J., Bohr, V. A., Sander, M. & Kraemer, K. H. Xeroderma pigmentosum and other diseases of human premature aging and DNA repair: Molecules to patients. *Mech. Ageing Dev.* **132**, 340–347 (2011).

18. 18.

Langevin, F., Crossan, G. P., Rosado, I. V., Arends, M. J. & Patel, K. J. Fancd2 counteracts the toxic effects of naturally produced aldehydes in mice. *Nature* **475**, 53–59 (2011).

19. 19.

Garaycoechea, J. I. et al. Genotoxic consequences of endogenous aldehydes on mouse haematopoietic stem cell function. *Nature* **489**, 571–575 (2012).

20. 20.

Young, M. D. et al. Single-cell transcriptomes from human kidneys reveal the cellular identity of renal tumors. *Science* **361**, 594–599 (2018).

21. 21.

Feng, D., Ngov, C., Henley, N., Boufaied, N. & Gerarduzzi, C. Characterization of matricellular protein expression signatures in

mechanistically diverse mouse models of kidney injury. *Sci. Rep.* **9**, 16736 (2019).

22. 22.

Vaidya, V. S., Ramirez, V., Ichimura, T., Bobadilla, N. A. & Bonventre, J. V. Urinary kidney injury molecule-1: a sensitive quantitative biomarker for early detection of kidney tubular injury. *Am. J. Physiol.* **290**, 517–529 (2006).

23. 23.

Dieterle, F. et al. Urinary clusterin, cystatin C, B2-microglobulin and total protein as markers to detect drug-induced kidney injury. *Nat. Biotechnol.* **28**, 463–469 (2010).

24. 24.

Brooks, C. R. et al. KIM-1/TIM-1-mediated phagocytosis links ATG5/ULK1-dependent clearance of apoptotic cells to antigen presentation. *EMBO J.* **34**, 2441–2464 (2015).

25. 25.

Freitas, M. C. S. et al. Type I interferon pathway mediates renal ischemia/reperfusion injury. *Transplantation* **92**, 131–138 (2011).

26. 26.

Liu, J. et al. Molecular characterization of the transition from acute to chronic kidney injury following ischemia/reperfusion. *JCI Insight* **2**, (2017).

27. 27.

Espinosa, J. M., Verdun, R. E. & Emerson, B. M. p53 functions through stress- and promoter-specific recruitment of transcription initiation components before and after DNA damage. *Mol. Cell* **12**, 1015–1027 (2003).

28. 28.

Campisi, J. & D'Adda Di Fagagna, F. Cellular senescence: when bad things happen to good cells. *Nat. Rev. Mol. Cell Biol.* **8**, 729–740 (2007).

29. 29.

Mullican, S. E. et al. GFRAL is the receptor for GDF15 and the ligand promotes weight loss in mice and nonhuman primates. *Nat. Med.* **23**, 1150–1157 (2017).

30. 30.

Tufegdzic Vidakovic, A. et al. Analysis of RNA polymerase II ubiquitylation and proteasomal degradation. *Methods* **159–160**, 146–156 (2019).

31. 31.

Wilson, B. T. et al. The Cockayne Syndrome Natural History (CoSyNH) study: clinical findings in 102 individuals and recommendations for care. *Genet. Med.* **18**, 483–493 (2016).

32. 32.

Hsu, J. Y. et al. Non-homeostatic body weight regulation through a brainstem-restricted receptor for GDF15. *Nature* **550**, 255–259 (2017).

33. 33.

Breen, D. M. et al. GDF-15 neutralization alleviates platinum-based chemotherapy-induced emesis, anorexia, and weight loss in mice and nonhuman primates. *Cell Metab.* **32**, 938–950.e6 (2020).

34. 34.

Enoiu, M., Jiricny, J. & Schärer, O. D. Repair of cisplatin-induced DNA interstrand crosslinks by a replication-independent pathway

involving transcription-coupled repair and translesion synthesis.
Nucleic Acids Res. **40**, 8953–8964 (2012).

35. 35.

Knipscheer, P. et al. The Fanconi anemia pathway promotes replication-dependent DNA interstrand cross-link repair. *Science* **326**, 1698–1701 (2009).

36. 36.

Mulderigg, L. & Garaycoechea, J. I. XPF-ERCC1 protects liver, kidney and blood homeostasis outside the canonical excision repair pathways. *PLoS Genet.* **16**, e1008555 (2020).

37. 37.

Vermeij, W. P. et al. Restricted diet delays accelerated ageing and genomic stress in DNA-repair-deficient mice. *Nature* **537**, 427–431 (2016).

38. 38.

Watson, H. J. et al. Genome-wide association study identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa. *Nat. Genet.* **51**, 1207–1214 (2019).

39. 39.

Crossan, G. P. et al. Disruption of mouse Slx4, a regulator of structure-specific nucleases, phenocopies Fanconi anemia. *Nat. Genet.* **43**, 147–152 (2011).

40. 40.

de Vries, A. et al. Increased susceptibility to ultraviolet-B and carcinogens of mice lacking the DNA excision repair gene XPA. *Nature* **377**, 169–173 (1995).

41. 41.

Cheo, D. L. et al. Characterization of defective nucleotide excision repair in XPC mutant mice. *Mutat. Res.* **374**, 1–9 (1997).

42. 42.

Liu, L. et al. Essential roles of S-nitrosothiols in vascular homeostasis and endotoxic shock. *Cell* **116**, 617–628 (2004).

43. 43.

Donehower, L. A. et al. Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. *Nature* **356**, 215–221 (1992).

44. 44.

Guyenet, S. J. et al. A simple composite phenotype scoring system for evaluating mouse models of cerebellar ataxia. *J. Vis. Exp.* **39**, 1787 (2010).

45. 45.

Dingler, F. A. et al. Two aldehyde clearance systems are essential to prevent lethal formaldehyde accumulation in mice and humans. *Mol. Cell* **80**, 996–1012.e9 (2020).

46. 46.

Young, M. D. & Behjati, S. SoupX removes ambient RNA contamination from droplet based single cell RNA sequencing data. *Gigascience* **9**,giaa151 (2020).

47. 47.

Wolock, S. L., Lopez, R. & Klein, A. M. Scrublet: computational identification of cell doublets in single-cell transcriptomic data. *Cell Syst.* **8**, 281–291.e9 (2019).

48. 48.

Popescu, D. M. et al. Decoding human fetal liver haematopoiesis. *Nature* **574**, 365–371 (2019).

49. 49.

Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902.e21 (2019).

50. 50.

Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).

51. 51.

Benjamini, Y. & Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289–300 (1995).

52. 52.

Park, J. et al. Single-cell transcriptomics of the mouse kidney reveals potential cellular targets of kidney disease. *Science* **360**, 758–763 (2018).

53. 53.

Traag, V. A., Waltman, L. & van Eck, N. J. From Louvain to Leiden: guaranteeing well-connected communities. *Sci. Rep.* **9**, 5233 (2019).

54. 54.

McInnes, L., Healy, J., Saul, N. & Großberger, L. UMAP: uniform manifold approximation and projection. *J. Open Source Softw.* **3**, 861 (2018).

Acknowledgements

We thank members of the K.J.P., G.P.C. and J.I.G. laboratories for critical reading of the manuscript. We thank J. Svejstrup and A. Tufegdzic Vidakovic for the pGEX-Dsk2 plasmid, E. Friedberg, G. T. van der Horst and J. Hoeijmakers for sharing the *Xpc*^{-/-}, *Xpa*^{-/-} and *Csb*^{m/m} mice and S. Wells, M. Stewart and H. Cater for work done at MRC Harwell. We thank Pfizer for the GDF15 monoclonal antibody. K.J.P. is supported by MRC, CRUK (C42693/A23273), Wellcome Trust (106202/Z/14/Z) and European Union Research and Innovation programme Horizon 2020 (Grant Agreement Number 730879). L.M., F.A.D. and G.P.C. were supported by CRUK (C42693/A23273). J.I.G. is supported by the Hubrecht Institute. C.L.M. is supported by the Wellcome Trust (106202/Z/14/Z). M.R.C. is funded by an NIHR Research Professorship (RP-2017-08-ST2-002). J.R.F. is funded by NIHR Cambridge Blood and Transplant Research Unit Organ Donation. Z.K.T. and M.R.C. are funded by MRC HCA grant MR/S035842/1. L.G. is funded from CRUK (FC0001166), the UK Medical Research Council (FC0001166) and the Wellcome Trust (FC0001166). S.O. is supported by the Medical Research Council MRC.MC.UU.12012.1, a Wellcome Senior Investigator Award 214274/Z/18/Z and the NIHR Cambridge Biomedical Research Centre. J.A.T. is supported by an NIHR Clinical Lectureship (CL-2019-14-504).

Author information

Author notes

1. These authors contributed equally: Lee Mulderrig, Juan I. Garaycoechea

Affiliations

1. MRC Weatherall Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Oxford, UK

Lee Mulderrig, Christopher L. Millington, Felix A. Dingler & Ketan J. Patel

2. MRC Laboratory of Molecular Biology, Cambridge, UK

Lee Mulderrig, Gerry P. Crossan & Ketan J. Patel

3. Hubrecht Institute, Utrecht, The Netherlands

Juan I. Garaycoechea

4. Molecular Immunity Unit, Department of Medicine, University of Cambridge, Cambridge, UK

Zewen K. Tuong, John R. Ferdinand & Menna R. Clatworthy

5. Cellular Genetics, Wellcome Sanger Institute, Wellcome Genome Campus, Hinxton, UK

Zewen K. Tuong & Menna R. Clatworthy

6. The Francis Crick Institute, London, UK

Liam Gaul

7. MRC Metabolic Diseases Unit, Wellcome Trust-Medical Research Council Institute of Metabolic Science, University of Cambridge, Cambridge, UK

John A. Tadross & Stephen O'Rahilly

8. Department of Pathology, University of Cambridge, Cambridge, UK

John A. Tadross

9. Division of Pathology, University of Edinburgh, Cancer Research UK Edinburgh Centre, IGMM, Edinburgh, UK

Mark J. Arends

10. Cambridge Institute of Therapeutic Immunology and Infectious Diseases, University of Cambridge and NIHR Cambridge Biomedical Research Centre, Cambridge, UK

Menna R. Clatworthy

Contributions

K.J.P., J.I.G. and G.P.C. conceived the study. K.J.P. wrote the manuscript. L.M. and J.I.G. designed and performed the majority of the experiments. F.A.D. performed GDF15 measurements. M.R.C. and J.R.F. designed the scRNA-seq experiment, J.R.F. prepared tissues for sequencing and processed the scRNA-seq data. Z.K.T. analysed the scRNA-seq data. C.L.M. measured formaldehyde adducts. L.G. generated CSB KO HEK 293 cells. J.A.T. performed *in situ* hybridization experiments. M.J.A. characterized the kidney pathology. S.O. provided crucial insight on the study.

Corresponding author

Correspondence to [Ketan J. Patel](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Stephan Herzig, Ruben van Boxtel 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 TC-NER and formaldehyde detoxification cooperate to protect cells from formaldehyde toxicity.

a-c, Cytotoxicity of UV (a) and formaldehyde (b) in *Csb^{m/m}* (a,b) or *Adh5^{-/-}* *Csb^{m/m}* (c) tMEF cell lines complemented with either WT or K991R CSB. Data plotted as mean and s.e.m; experiments performed in triplicate. **d-e**, Cytotoxicity of UV and formaldehyde in HEK293 cell lines, data shown as mean and s.e.m., experiments performed in triplicate.

Extended Data Fig. 2 *Adh5^{-/-}* *Csb^{m/m}* mice are born at sub-Mendelian ratios and have reduced body weight.

a, Ratios of pups from *Csb^{+/m}*, *Xpc^{+/ -}* and *Xpa^{+/ -}* crosses showing that homozygous mice are born at Mendelian ratios (*P* calculated by two-sided Chi-squared test) **b**, Ratios of pups born from *Adh5^{-/-}* *Csb^{+/m}*, *Adh5^{-/-}* *Xpc^{+/ -}* and *Adh5^{-/-}* *Xpa^{+/ -}* crosses showing that both *Adh5^{-/-}* *Csb^{m/m}* and *Adh5^{-/-}* *Xpa^{+/ -}* pups are born at sub-Mendelian ratios (*P* calculated by two-sided Chi-squared test). **c**, Ratios of pups born from *Adh5^{+/ -}* *Csb^{+/m}* and *Adh5^{-/-}* *Csb^{+/m}* crosses showing the ratio of *Adh5^{-/-}* *Csb^{m/m}* pups is partially rescued when the mother is aldehyde-detoxification proficient (*Adh5* *+/ -* instead of *Adh5* *-/-* in **b**, *P* calculated by two-sided Chi-squared test). For **a**, **b**, and **c**, mice were genotyped between 2-3 wk of age. **d**, Weights of adult male and female mice at 8 wk of age (Data shown as mean and s.e.m.; *P* calculated by two-sided Mann-Whitney test; *n* = 14, 19, 13, 12, 4, 9, 3, 7 for males left to right and 24, 30, 22, 21, 12, 9, 4, 5 for females left to right). **e**, Image of *Csb^{m/m}* and *Adh5^{-/-}* *Csb^{m/m}* littermates at 12 months of age.

Extended Data Fig. 3 *Adh5^{-/-}* *Csb^{m/m}* mice exhibit features of human Cockayne Syndrome.

a, Growth curves of female mice based on weekly weights. Data shown as mean and s.e.m. $n = 14, 14, 18$ and 14 . Along with fat mass from EchoMRI performed at 11 wk (young) and 53 wk (old). P calculated by two-sided Mann-Whitney test; data shown as mean and s.e.m.; $n = 13, 14, 18, 14, 11, 13, 6$ and 6 left to right. **b**, Bar graph of grip strength for young (3 months) and old (1 yr) mice, determined by placing all four limbs on a grid attached to a force gauge. P calculated by two-sided Mann-Whitney test, data shown as mean and s.e.m.; $n = 13, 15, 18, 20, 5, 12, 12$ and 8 for males left to right and $13, 14, 18, 14, 5, 3, 14$ and 8 for females left to right. **c**, Left, age of onset plot for kyphosis. P calculated by two-sided Mantel-Cox logrank test, $n = 21, 19, 22$ and 9 . Right, representative x-rays of $Csb^{m/m}$ and $Adh5^{-/-}Csb^{m/m}$ mice at 1 yr of age showing kyphosis in the $Adh5^{-/-}Csb^{m/m}$ mouse. **d**, Brain weights of mice taken at 3 months (young) and 18 months (old). P calculated by two-tailed Student's t -test; data shown as mean and s.e.m., $n = 3$ mice.

Extended Data Fig. 4 $Adh5^{-/-}Csb^{m/m}$ mice succumb to chronic kidney failure and have liver abnormalities.

a, Survival and cancer-free survival curve of $Adh5^{-/-}Csb^{m/m}$ and control mice. **b**, Bar chart indicating the cause of death for wild type, $Adh5^{-/-}$, $Csb^{m/m}$ and $Adh5^{-/-}Csb^{m/m}$ mice ($n = 18, 31, 16$ and 19). **c**, Weights of $Adh5^{-/-}Csb^{m/m}$ male and female kidneys at 12 months relative to tibia length, data shown as mean and s.e.m. ($n=8, 22, 20, 16, 10, 4, 22, 17$ from left to right). **d**, Representative H&E stained sections of kidney from $Csb^{m/m}$ and $Adh5^{-/-}Csb^{m/m}$ mice at sequential timepoints ($n=4, 3, 4, 5$, from left to right). **e**, Bar chart of the percentage of fibrosis in the cortex of H&E stained kidney sections at sequential time points in $Adh5^{-/-}Csb^{m/m}$ and terminal controls, data shown as mean and s.e.m. ($n=4, 4, 4, 3, 4, 5$, from left to right). **f**, Representative PAS stained sections of kidney showing intratubular casts from $Csb^{m/m}$ and $Adh5^{-/-}Csb^{m/m}$ terminal mice ($n=3$). **g**, Urine obtained from indicated mice were tested for the presence of proteinuria by multistix 10SG. **h**, Blood counts from terminal blood samples of $Adh5^{-/-}Csb^{m/m}$ and controls P calculated by two-sided Mann-Whitney test; data show as mean and s.e.m.; $n = 9, 8, 6, 5$ from left to right. **i**, Terminal serum measurements of albumin, alkaline phosphatase and

alanine transaminase. P calculated by two-sided Mann-Whitney test; data show as mean and s.e.m.; $n = 33, 10, 13$ and 16 . **j**, Representative H&E stained sections of liver from age-matched $Csb^{m/m}$ and $Adh5^{-/-}Csb^{m/m}$ mice, arrows indicate cells with enlarged nuclei ($n=3$). **k**, Quantification of hepatocyte nuclear DNA content in young (3 month) and old (18 month) mice. P calculated by two-sided Student's t -test for the content of 8n nuclei; data shown as mean and s.e.m.; $n = 3$ mice.

Extended Data Fig. 5 Methanol exposure exacerbates the Cockayne Syndrome phenotype in $Adh5^{-/-}Csb^{m/m}$ mice.

a, Scheme outlining weekly intra-peritoneal (I.P.) injection of 1.5 g/kg methanol (or saline) and analysis of treated mice. **b-c**, Monthly serum levels of urea (b) and creatinine (c) from mice exposed to methanol and saline controls, data plotted as mean and s.e.m.; $n = 4$. **d**, Kidney failure-free survival curve of $Adh5^{-/-}Csb^{m/m}$ mice with and without methanol exposure (P calculated by two-sided Mantel-Cox logrank test; $n = 7$ and 19). **e**, Representative image of kidneys from age-matched mice exposed to methanol or saline, taken after 24 hr of fixation. **f**, Representative H&E stained sections of kidney from $Adh5^{-/-}Csb^{m/m}$ mice exposed to methanol or saline, G indicates glomeruli and arrows indicate atrophic tubules ($n=4$). **g**, Brain weights of age-matched mice exposed to methanol or saline. P calculated by two tailed Student's t -test; data shown as mean and s.e.m., $n = 3$ mice. **h**, Quantification of the number of MAC2+ cells per field in mice exposed to methanol or saline. P calculated by two tailed Student's t -test; data shown as mean and s.e.m., $n = 3$ mice. **i**, Representative immunofluorescence images of the cerebellum of $Adh5^{-/-}Csb^{m/m}$ mice exposed to methanol or saline stained with MAC2 and DAPI at 40x ($n=3$).

Extended Data Fig. 6 Methanol reveals kidney failure in $Adh5^{-/-}Xpa^{-/-}$ mice but ethanol does not in $Aldh2^{-/-}Csb^{m/m}$ mice.

a, Serum measurements of albumin, alkaline phosphatase and alanine transaminase after 6 months of methanol treatment along with saline-treated

controls. P calculated by two-sided Mann-Whitney test; data show as mean and s.e.m.; $n = 4$ mice. **b**, Kaplan-Meier survival curve of wild type, $Adh5^{-/-}$, $Xpa^{-/-}$ and $Adh5^{-/-}Xpa^{-/-}$ mice ($n = 18, 31, 10$ and 13). **c**, Kidney failure-free survival curve of $Adh5^{-/-}Xpa^{-/-}$ mice with and without methanol exposure (P calculated by two-sided Mantel-Cox logrank test; $n = 4$ and 13). **d**, Monthly serum levels of urea and creatinine from mice exposed to methanol or saline, data plotted as mean and s.e.m.; $n = 4$. **e**, Kaplan-Meier survival curve of wild-type, $Aldh2^{-/-}$, $Csb^{m/m}$ and $Aldh2^{-/-}Csb^{m/m}$ mice ($n = 18, 5, 16$ and 5). **f**, Kaplan-Meier survival curve of wild-type, $Aldh2^{-/-}$, $Csb^{m/m}$ and $Aldh2^{-/-}Csb^{m/m}$ mice treated with 20% ethanol continuously in the drinking water ($n = 6$). **g**, Monthly serum levels of urea and creatinine from mice exposed to 20% ethanol continuously in the drinking water, data plotted as mean and s.e.m.; $n = 6$.

Extended Data Fig. 7 scRNA-seq reveals cells susceptible to formaldehyde transcriptional stress.

a, Scheme outlining the scRNA-seq experiment. **b**, UMAP plots of murine kidney scRNA-seq data for wild type, $Adh5^{-/-}$, $Csb^{m/m}$ and $Adh5^{-/-}Csb^{m/m}$ mice, $n = 3$ and $n = 31,624, 16,023, 29,082, 29,802$ cells (pDC, plasmacytoid dendritic cell; MNP, mononuclear phagocyte; LOH, loop of henle; CT, connecting tubule; DCT, distal convoluted tubules; CD-PC, collecting duct – principle cell; CD-IC, collecting duct – intercalated cell; T_NK, T cells/NK cells; B, B cells). **c**, Bar chart showing the composition of cell types in the scRNA-seq data. **d**, UMAP plot of all scRNA-seq data with each genotype labelled a different colour, arrow indicates PT cells that are distinct to $Adh5^{-/-}Csb^{m/m}$ kidneys. **e**, UMAP plot of PT sub-clusters split by genotype. **f**, Heatmap indicating the top two marker genes for each PT sub-cluster (marker genes are calculated by comparing the expression of each PT sub-cluster against all remaining PTs). **g, top**, Feature plots of *Kim-1* expression from PT sub-clusters of $Adh5^{-/-}Csb^{m/m}$ cells and controls, **bottom**, immunofluorescence images of $Adh5^{-/-}Csb^{m/m}$ and control kidney sections stained with Kim-1 taken from mice aged 12 months. **h**, Log2 fold change plot of DEGs from PT-4 made relative to wild-type PT-4 cells. **i, top**, Feature plots of *Gdf15* expression from PT sub-clusters of $Adh5^{-/-}Csb^{m/m}$ cells and controls, **bottom**, representative images of *In situ*

hybridisation for *Gdf15* mRNA (red spots) performed on *Adh5^{-/-}Csb^{m/m}* and control kidney sections, n=3.

Extended Data Fig. 8 scRNA-seq gene expression analysis of PT-4 sub-cluster cells.

a, Feature plots of *Cyr61*, *Spp1* and *B2m* expression from PT sub-clusters of *Csb^{m/m}* and *Adh5^{-/-}Csb^{m/m}* cells. **b**, Network of top 100 marker genes from PT sub-cluster PT-4 (marker genes were calculated by comparing expression of genes in PT-4 to all other PT cells). Data visualised in cytoscape v3.7.1 using the STRING app. Nodes are coloured based on top GO term enrichment pathway. **c**, Feature plots of *p21*, *Phlda3* and *Btg2* expression from PT sub-clusters of *Csb^{m/m}* and *Adh5^{-/-}Csb^{m/m}* cells. **d**, Heatmap of expression of p53 target genes in PT cells of *Adh5^{-/-}Csb^{m/m}* and control mice.

Extended Data Fig. 9 *Gdf15* expression in mouse tissues.

a and b, Representative images of *in situ* hybridisation for *Gdf15* mRNA (red spots) performed on *Adh5^{-/-}Csb^{m/m}* and control kidney sections, n=3 **c**, Feature plots of *Gdf15* expression from LOH and CD cells for *Adh5^{-/-}Csb^{m/m}* and control mice. **d**, Dot plot of mean expression and fraction of cells expressing *Gdf15* in the loop of henle (LOH) and collecting duct (CD) for *Adh5^{-/-}Csb^{m/m}* and control mice. **e**, Feature plots of *Gdf15* expression from LOH and CD cells for *Adh5^{-/-}Csb^{m/m}* mice treated with methanol or untreated. **f**, Dot plot of mean expression and fraction of cells expressing *Gdf15* in *Adh5^{-/-}Csb^{m/m}* mice treated with methanol or untreated.

Extended Data Fig. 10 Metabolic analysis of *Adh5^{-/-}Csb^{m/m}* mice.

a-b, Male and female mice were placed in metabolic chambers and respiratory exchange rate (R.E.R) and energy expenditure were measured over a 20-hr period, data shown as mean and s.e.m. n = 3, 5, 13 and 2 for males and n = 7, 4, 11 and 10 for females. **c**, male and female mice were

placed in open field chambers and activity was measured over 20 min, data shown as mean and s.e.m ($n=5, 10, 16$ and 14 for males and $n= 4, 4, 14$ and 12 for females). **d**, Bar chart of weekly food intake from $Adh5^{-/-}Csb^{m/m}$ and control mice singly housed and averaged out over 5 wk, data shown as mean and s.e.m, $n=3$.

Extended Data Fig. 11 GDF15 release in response to cisplatin is increased in $Csb^{m/m}$ mice but absent in $p53^{-/-}$ mice.

a, Daily weights of wild-type and $Csb^{m/m}$ mice exposed to weekly 0.5 mg/kg cisplatin intra-peritoneal injections alongside serum GDF15 measurements taken before and 4 wk after weekly injections. P calculated using two-sided Mann-Whitney test; data shown as mean and s.e.m.; $n = 6$ mice. **b**, Daily weights of wild-type and $p53^{-/-}$ mice exposed to weekly 4 mg/kg cisplatin intra-peritoneal injections alongside serum GDF15 measurements taken before and 6 wk after weekly injections. P calculated using two-sided Mann-Whitney test; data shown as mean and s.e.m.; $n = 6$ mice. **c**, *In situ* hybridisation for $Gdf15$ mRNA (red spots) performed on wild-type and $p53^{-/-}$ kidney sections 24 h after 4 mg/kg cisplatin treatment ($n=3$).

Supplementary information

Supplementary Fig. 1

| **Uncropped western blots.** Uncropped western blots for Fig. 1e–g.

Reporting Summary

Supplementary Table 1

| **Top 100 marker genes for each PT sub-cluster.** Table of the top 100 marker genes from each PT sub-cluster from scRNA-seq analysis of murine kidneys.

Peer Review File

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Mulderrig, L., Garaycoechea, J.I., Tuong, Z.K. *et al.* Aldehyde-driven transcriptional stress triggers an anorexic DNA damage response. *Nature* **600**, 158–163 (2021). <https://doi.org/10.1038/s41586-021-04133-7>

- Received: 16 June 2020
- Accepted: 13 October 2021
- Published: 24 November 2021
- Issue Date: 02 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04133-7>

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/s41586-021-04133-7>

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

- Article
- [Published: 17 November 2021](#)

Structure, function and pharmacology of human itch receptor complexes

- [Fan Yang^{1,2,3,4,5 na1}](#),
- [Lulu Guo^{3 na1}](#),
- [Yu Li^{1,2 na1}](#),
- [Guopeng Wang^{2,6 na1}](#),
- [Jia Wang^{3 na1}](#),
- [Chao Zhang^{3 na1}](#),
- [Guo-Xing Fang^{3,7}](#),
- [Xu Chen⁴](#),
- [Lei Liu³](#),
- [Xu Yan⁴](#),
- [Qun Liu³](#),
- [Changxiu Qu^{1,3}](#),
- [Yunfei Xu ORCID: orcid.org/0000-0002-0580-7116⁸](#),
- [Peng Xiao³](#),
- [Zhongliang Zhu⁹](#),
- [Zijian Li¹⁰](#),
- [Jiuyao Zhou⁷](#),
- [Xiao Yu⁴](#),
- [Ning Gao ORCID: orcid.org/0000-0003-3067-9993²](#) &
- [Jin-Peng Sun ORCID: orcid.org/0000-0003-3572-1580^{1,3,5}](#)

Nature volume 600, pages 164–169 (2021)

- 5649 Accesses
- 44 Altmetric
- [Metrics details](#)

Subjects

- [Cryoelectron microscopy](#)
- [Receptor pharmacology](#)

Abstract

In the clades of animals that diverged from the bony fish, a group of Mas-related G-protein-coupled receptors (MRGPRs) evolved that have an active role in itch and allergic signals^{1,2}. As an MRGPR, MRGPRX2 is known to sense basic secretagogues (agents that promote secretion) and is involved in itch signals and eliciting pseudoallergic reactions^{3,4,5,6}. MRGPRX2 has been targeted by drug development efforts to prevent the side effects induced by certain drugs or to treat allergic diseases. Here we report a set of cryo-electron microscopy structures of the MRGPRX2–G_{i1} trimer in complex with polycationic compound 48/80 or with inflammatory peptides. The structures of the MRGPRX2–G_{i1} complex exhibited shallow, solvent-exposed ligand-binding pockets. We identified key common structural features of MRGPRX2 and describe a consensus motif for peptidic allergens. Beneath the ligand-binding pocket, the unusual kink formation at transmembrane domain 6 (TM6) and the replacement of the general toggle switch from Trp^{6,48} to Gly^{6,48} (superscript annotations as per Ballesteros–Weinstein nomenclature) suggest a distinct activation process. We characterized the interfaces of MRGPRX2 and the G_i trimer, and mapped the residues associated with key single-nucleotide polymorphisms on both the ligand and G-protein interfaces of MRGPRX2. Collectively, our results provide a structural basis for the sensing of cationic allergens by MRGPRX2, potentially facilitating the rational design of therapies to prevent unwanted pseudoallergic reactions.

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: Cryo-EM structure of MRGPRX2–G_i complexes and the C48/80-binding pocket.

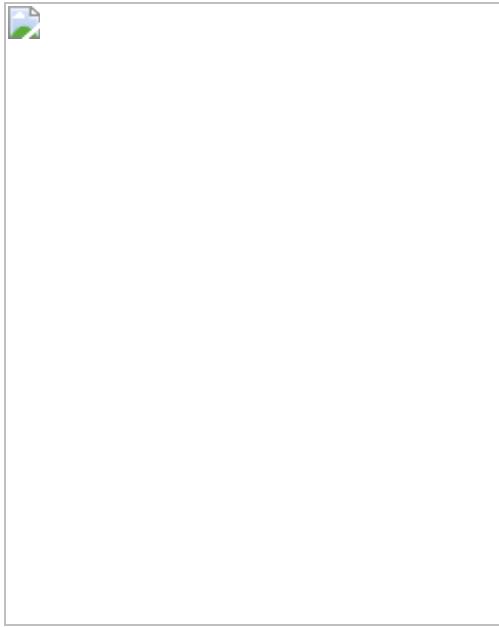


Fig. 2: Consensus peptidic motif recognized by MRGPRX2.

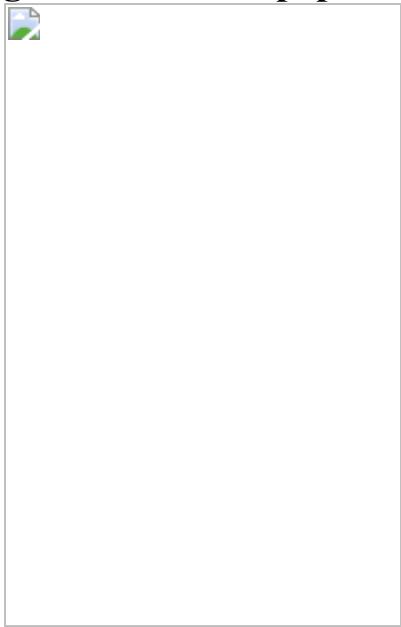


Fig. 3: Activation mechanism of MRGPRX2.

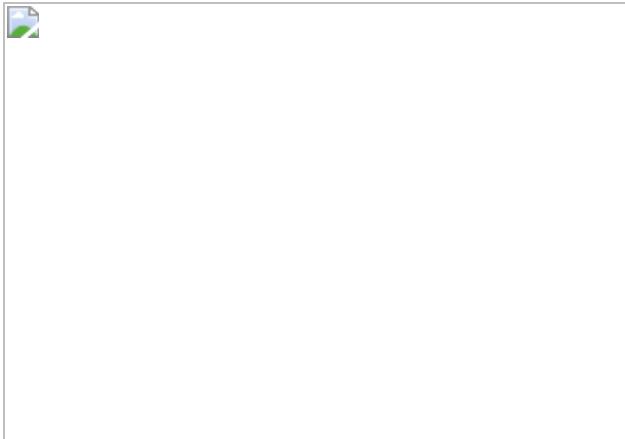


Fig. 4: The coupling of MRGPRX2 to G_{i1} .



Data availability

The cryo-EM density maps for the PAMP-12–MRGPRX2– G_{i1} state 1 complex, PAMP-12–MRGPRX2– G_{i1} state 2 complex, SP–MRGPRX2– G_{i1} complex, C14^{circular}–MRGPRX2– G_{i1} complex, C14^{linear}–MRGPRX2– G_{i1} complex, C48/80–MRGPRX2– G_{i1} state 1 complex and C48/80–MRGPRX2– G_{i1} state 2 complex have been deposited at the EMDB (Electron Microscopy Data Bank) under accession codes [EMD-32131](#), [EMD-32132](#), [EMD-31923](#), [EMD-31922](#), [EMD-32136](#), [EMD-32138](#) and [EMD-31918](#), respectively. The local maps of the receptor and ligand regions of the PAMP-12–MRGPRX2– G_{i1} complex, C14^{linear}–MRGPRX2– G_{i1} complex and C48/80–MRGPRX2– G_{i1} complex have been deposited at the EMDB (Electron Microscopy Data Bank) under accession codes [EMD-32133](#), [EMD-32137](#) and [EMD-32139](#), respectively. The coordinates for the model of the PAMP-12–MRGPRX2– G_{i1} state 1 complex (only receptor and G_{i1}), PAMP-12–MRGPRX2– G_{i1} state 2 complex (only receptor and G_{i1}), PAMP-12–MRGPRX2complex (only ligand and

receptor), SP–MRGPRX2–G_{i1} complex, C14^{circular}–MRGPRX2–G_{i1} complex, C14^{linear}–MRGPRX2–G_{i1} complex (only receptor and G_{i1}), C14^{linear}–MRGPRX2 complex (only ligand and receptor), C48/80–MRGPRX2–G_{i1} state 1 complex, C48/80–MRGPRX2–G_{i1} state 2 complex and C48/80–MRGPRX2 complex (only ligand and receptor) have been deposited at the PDB under accession codes [7VUY](#), [7VUZ](#), [7VV0](#), [7VDM](#), [7VDL](#), [7VV3](#), [7VV4](#), [7VV5](#), [7VDH](#) and [7VV6](#), respectively. All other data are available from the corresponding authors on request.

References

1. 1.
Bader, M., Alenina, N., Andrade-Navarro, M. A. & Santos, R. A. MAS and its related G protein-coupled receptors, Mrgprs. *Pharmacol. Rev.* **66**, 1080–1105 (2014).
2. 2.
Meixiong, J. & Dong, X. Mas-related G protein-coupled receptors and the biology of itch sensation. *Ann. Rev. Genet.* **51**, 103–121 (2017).
3. 3.
Solinski, H. J., Gudermann, T. & Breit, A. Pharmacology and signaling of MAS-related G protein-coupled receptors. *Pharmacol. Rev.* **66**, 570–597 (2014).
4. 4.
McNeil, B. D. et al. Identification of a mast-cell-specific receptor crucial for pseudo-allergic drug reactions. *Nature* **519**, 237–241 (2015).
5. 5.

Wedi, B., Gehring, M. & Kapp, A. The pseudoallergen receptor MRGPRX2 on peripheral blood basophils and eosinophils: expression and function. *Allergy* **75**, 2229–2242 (2020).

6. 6.

Yang, S. et al. Adaptive evolution of MRGX2, a human sensory neuron specific gene involved in nociception. *Gene* **352**, 30–35 (2005).

7. 7.

Meixiong, J. et al. Activation of mast-cell-expressed Mas-related G-protein-coupled receptors drives non-histaminergic itch. *Immunity* **50**, 1163–1171 (2019).

8. 8.

Green, D. P., Limjunyawong, N., Gour, N., Pundir, P. & Dong, X. A mast-cell-specific receptor mediates neurogenic inflammation and pain. *Neuron* **101**, 412–420 (2019).

9. 9.

Porebski, G., Kwiecien, K., Pawica, M. & Kwitniewski, M. Mas-related G protein-coupled receptor-X2 (MRGPRX2) in drug hypersensitivity reactions. *Front. Immunol.* **9**, 3027 (2018).

10. 10.

Azimi, E., Reddy, V. B. & Lerner, E. A. MRGPRX2, atopic dermatitis and red man syndrome. *Itch* **2**, e5 (2017).

11. 11.

Subramanian, H., Gupta, K. & Ali, H. Roles of Mas-related G protein-coupled receptor X2 on mast cell-mediated host defense, pseudoallergic drug reactions, and chronic inflammatory diseases. *J. Allergy Clin. Immunol.* **138**, 700–710 (2016).

12. 12.

Grimes, J. et al. MrgX2 is a promiscuous receptor for basic peptides causing mast cell pseudo-allergic and anaphylactoid reactions. *Pharmacol. Res. Perspect.* **7**, e00547 (2019).

13. 13.

Liang, Y. L. et al. Dominant negative G proteins enhance formation and purification of agonist-GPCR-G protein complexes for structure determination. *ACS Pharmacol. Transl. Sci.* **1**, 12–20 (2018).

14. 14.

Lansu, K. et al. In silico design of novel probes for the atypical opioid receptor MRGPRX2. *Nat. Chem. Biol.* **13**, 529–536 (2017).

15. 15.

Kamohara, M. et al. Identification of MrgX2 as a human G-protein-coupled receptor for proadrenomedullin N-terminal peptides. *Biochem. Biophys. Res. Commun.* **330**, 1146–1152 (2005).

16. 16.

Krishna Kumar, K. et al. Structure of a signaling cannabinoid receptor 1-G protein complex. *Cell* **176**, 448–458 (2019).

17. 17.

Liu, X. et al. Structural insights into the process of GPCR-G protein complex formation. *Cell* **177**, 1243–1251 (2019).

18. 18.

Ping, Y. Q. et al. Structures of the glucocorticoid-bound adhesion receptor GPR97-G_o complex. *Nature* **589**, 620–626 (2021).

19. 19.

Bokoch, M. P. et al. Ligand-specific regulation of the extracellular surface of a G-protein-coupled receptor. *Nature* **463**, 108–112 (2010).

20. 20.

Alkanfari, I., Gupta, K., Jahan, T. & Ali, H. Naturally occurring missense MRGPRX2 variants display loss of function phenotype for mast cell degranulation in response to substance P, hemokinin-1, human beta-defensin-3, and icatibant. *J. Immunol.* **201**, 343–349 (2018).

21. 21.

Gonzalez-Rey, E., Chorny, A., Robledo, G. & Delgado, M. Cortistatin, a new antiinflammatory peptide with therapeutic effect on lethal endotoxemia. *J. Exp. Med.* **203**, 563–571 (2006).

22. 22.

Li, R. et al. Molecular mechanism of ERK dephosphorylation by striatal-enriched protein tyrosine phosphatase. *J. Neurochem.* **128**, 315–329 (2014).

23. 23.

Ogasawara, H., Furuno, M., Edamura, K. & Noguchi, M. Novel MRGPRX2 antagonists inhibit IgE-independent activation of human umbilical cord blood-derived mast cells. *J. Leukocyte Biol.* **106**, 1069–1077 (2019).

24. 24.

Maeda, S. et al. Development of an antibody fragment that stabilizes GPCR/G-protein complexes. *Nat. Commun.* **9**, 3712 (2018).

25. 25.

Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51

(2005).

26. 26.

Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).

27. 27.

Rohou, A. & Grigorieff, N. CTFFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).

28. 28.

Zivanov, J., Nakane, T. & Scheres, S. H. W. Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in RELION-3.1. *IUCrJ* **7**, 253–267 (2020).

29. 29.

Kucukelbir, A., Sigworth, F. J. & Tagare, H. D. Quantifying the local resolution of cryo-EM density maps. *Nat. Methods* **11**, 63–65 (2014).

30. 30.

Biasini, M. et al. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* **42**, W252–W258 (2014).

31. 31.

Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

32. 32.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).

33. 33.

Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D Struct. Biol.* **74**, 531–544 (2018).

34. 34.

Morris, G. M. et al. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785–2791 (2009).

35. 35.

Bianco, G., Forli, S., Goodsell, D. S. & Olson, A. J. Covalent docking using autodock: two-point attractor and flexible side chain methods. *Protein Sci.* **25**, 295–301 (2016).

36. 36.

Lee, J. et al. CHARMM-GUI input generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM simulations using the CHARMM36 additive force field. *J. Chem. Theory Comput.* **12**, 405–413 (2016).

37. 37.

Huang, J. et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **14**, 71–73 (2017).

38. 38.

Metropolis, N. & Ulam, S. The Monte Carlo method. *J. Am. Stat. Assoc.* **44**, 335–341 (1949).

39. 39.

Van Der Spoel, D. et al. GROMACS: fast, flexible, and free. *J. Comput. Chem.* **26**, 1701–1718 (2005).

40. 40.

Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graphics* **14**, 33–38 (1996).

41. 41.

Kumari, R., Kumar, R., Open Source Drug Discovery, C. & Lynn, A. g_mmmpbsa—a GROMACS tool for high-throughput MM-PBSA calculations. *J. Chem. Inf. Model.* **54**, 1951–1962 (2014).

42. 42.

Srinivasan, J., Cheatham, T. E., Cieplak, P., Kollman, P. A. & Case, D. A. Continuum solvent studies of the stability of DNA, RNA, and phosphoramidate—DNA helices. *J. Am. Chem. Soc.* **120**, 9401–9409 (1998).

43. 43.

Yang, F. et al. Structural basis of GPBAR activation and bile acid recognition. *Nature* **587**, 499–504 (2020).

44. 44.

Li, T. et al. Homocysteine directly interacts and activates the angiotensin II type I receptor to aggravate vascular injury. *Nat. Commun.* **9**, 11 (2018).

45. 45.

Olsen, R. H. J. et al. TRUPATH, an open-source biosensor platform for interrogating the GPCR transducerome. *Nat. Chem. Biol.* **16**, 841–849 (2020).

Acknowledgements

We thank the staff at the Core Facilities at the School of Life Sciences, Peking University for help with negative staining EM; the staff at the Cryo-EM Platform and the Electron Microscopy Laboratory of Peking University for help with data collection; the staff at the High-performance Computing Platform of Peking University for help with computation; the staff at the National Centre for Protein Sciences at Peking University for technical assistance; and the staff at the cryo-EM facilities of Southern University of Science and Technology. This work was supported by the National Key R&D Program of China (2018YFC1003600 to X. Yu and J.-P.S.; 2019YFA0904200 to J.-P.S. and P.X.; and 2019YFA0508904 to N.G.), the National Natural Science Foundation of China (91939301 to J.-P.S.; 92057121 to X. Yu; 31900936 to F.Y.; 31971195 to P.X.), the National Science Fund for Distinguished Young Scholars Grant (81825022 to J.-P.S.; and 31725007 to N.G.), the National Science Fund for Excellent Young Scholars (81822008 to X. Yu and 82122070 to F.Y.), the Major Basic Research Project of Shandong Natural Science Foundation (ZR2020ZD39 to J.-P.S.), the Beijing Natural Science Foundation (Z200019 to J.-P.S.). The scientific calculations in this paper were performed on the HPC Cloud Platform of Shandong University.

Author information

Author notes

1. These authors contributed equally: Fan Yang, Lulu Guo, Yu Li, Guo-Peng Wang, Jia Wang, Chao Zhang

Affiliations

1. Key Laboratory of Molecular Cardiovascular Science, Department of Physiology and Pathophysiology, School of Basic Medical Sciences, Peking University, Ministry of Education, Beijing, China

Fan Yang, Yu Li, Changxiu Qu & Jin-Peng Sun

2. State Key Laboratory of Membrane Biology, Peking-Tsinghua Joint Center for Life Sciences, School of Life Sciences, Peking University, Beijing, China

Fan Yang, Yu Li, Guopeng Wang & Ning Gao

3. Key Laboratory of Experimental Teratology of the Ministry of Education, Department of Biochemistry and Molecular Biology, Shandong University School of Medicine, Jinan, China

Fan Yang, Lulu Guo, Jia Wang, Chao Zhang, Guo-Xing Fang, Lei Liu, Qun Liu, Changxiu Qu, Peng Xiao & Jin-Peng Sun

4. Key Laboratory of Experimental Teratology of the Ministry of Education, Department of Physiology, Shandong University School of Medicine, Jinan, China

Fan Yang, Xu Chen, Xu Yan & Xiao Yu

5. Advanced Medical Research Institute, Shandong University, Jinan, China

Fan Yang & Jin-Peng Sun

6. Cryo-EM platform, School of Life Sciences, Peking University, Beijing, China

Guopeng Wang

7. Department of Pharmacology, School of Pharmaceutical Sciences, Guangzhou University of Chinese Medicine, Guangzhou, China

Guo-Xing Fang & Jiuyao Zhou

8. Department of General Surgery, Qilu Hospital of Shandong University, Jinan, China

Yunfei Xu

9. School of Life Sciences, University of Science and Technology of China, Hefei, China

Zhongliang Zhu

10. Department of Cardiology and Institute of Vascular Medicine, Peking University Third Hospital, Research, Beijing, China

Zijian Li

Contributions

J.-P.S. and N.G. designed and supervised the overall project. J.-P.S., F.Y. and N.G. participated in all data analysis and interpretation. J.-P.S. and N.G. guided all structural analysis. F.Y. and L.G. generated the MRGPRX2 insect cell expression construct, established the C48/80–MRGPRX2–G_{i1}–scFv16, PAMP-12–MRGPRX2–G_{i1}–scFv16 and C14^{linear}–MRGPRX2–G_{i1}–scFV16 complex purification protocol, and prepared samples for the cryo-EM. Y.L. and G.W. prepared the cryo-EM grids, collected the cryo-EM data, and performed the cryo-EM map calculation, model building and refinement. X. Yu, J.W. and L.G. performed the G_i–G_q dissociation assay. C.Z. performed MD simulations. Z.Z. analysed the structure the C14^{linear}–MRGPRX2–G_{i1}–scFV16 complex. J.W., G.-X.F., J.Z., L.L. and X. Yan generated all of the MRGPRX2 constructs and mutants for the cell-based G-protein activity assays. C.Z. and Y.F.X. performed MD simulations. J.-P.S. and X. Yu designed the cellular functional assays for MRGPRX2. J.-P.S. and X. Yu designed the FlAsH–BRET assays and provided insightful ideas and experimental designs. L.G., J.W., Q.L., Z.L., X.C. and F.Y. designed all of the mutations for determining ligand-binding sites. F.Y., L.G., P.X. and J.W. prepared the figures. J.-P.S. wrote the manuscript. N.G. and X. Yu revised the manuscript. All of the authors have seen and commented on the manuscript.

Corresponding authors

Correspondence to [Ning Gao](#) or [Jin-Peng Sun](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Hugo Gutierrez-de-Teran 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 Purification of MRGPRX2–G_{i1}–scFv complex.

a-d: To increase receptor expression, thermostabilized cytochrome b₅₆₂RIL (BRIL) was incorporated at the N-terminus of full-length MRGPRX2, and the chimaeric protein was found to have G protein coupling activity similar to that of wild-type MRGPRX2. Dose response curves of the C48/80 (a), PAMP-12 (b), C14^{linear} (c) and SP (d) induced G_{αi}–G_γ dissociation in MRGPRX2-WT or Bril-MRGPRX2 overexpressing cells. Data from three independent experiments are presented as the mean ± SEM (n=3). **e-h:** Representative elution profile of in vitro reconstituted C48/80–MRGPRX2–G_{i1} trimer complex (e), PAMP-12–MRGPRX2–G_{i1} trimer complex (f), C14^{linear}–MRGPRX2–G_{i1} trimer complex (g) and SP–MRGPRX2–G_{i1} trimer complex (h) on Superose 6 Increase 10/30 column and SDS-PAGE of the size-exclusion chromatography peak. Representative Figures from three independent experiments were shown.

Extended Data Fig. 2 Cryo-EM images and single particle reconstruction of the C48/80–MRGPRX2–G_{i1} complex.

a–b, Cryo-EM micrograph (a) and reference-free two-dimensional class averages of the C48/80–MRGPRX2–G_{i1} trimer complex (b). Representative Cryo-EM micrograph from 6562 movies (shown in c) and representative two-dimensional class averages determined using approximately 0.26 million (state 1) or 0.32 million (state 2) particles after 3D classification were shown. **c**, Flow chart of cryo-EM data processing of C48/80–MRGPRX2–G_{i1} complex. **d**, Fourier shell correlation curves for the final 3D density maps of C48/80- bound MRGPRX2–G_{i1} state 1 (Blue) and state 2 (Orange). At the FSC 0.143 cut-off, the overall resolution for state 1 and state 2 are 2.76 Å and 2.90 Å, respectively (left panel). **e**, 3D density map colored according to local resolution (Å) of the C48/80–MRGPRX2–G_{i1} trimer complex.

Extended Data Fig. 3 Electron microscopy density map of C48/80–MRGPRX2–G_{i1} and PAMP-12–MRGPRX2–G_{i1} complex.

a–d, EM density of the TM helices of MRGPRX2 and the α5 helix of G_{αi1} of C 48/80–MRGPRX2–G_{i1} (a), PAMP-12–MRGPRX2–G_{i1} (b), C14^{linear}–MRGPRX2–G_{i1} complex (c) and SP–MRGPRX2–G_{i1} complex (d) respectively. All seven-TM bundles were unambiguously traceable in the cryo-EM density map, and the densities of large hydrophobic residues were utilized to assign the primary sequence of MRGPRX2. **e**, Position of ligand-binding pockets in the C14^{linear}–MRGPRX2–G_{i1}, SP–MRGPRX2–G_{i1}, β2AR–G_s (PDB: 3SN6), CB1–G_i (PDB: 6N4B), μ-opioid–G_i (PDB: 6DDF) and GPR97–mG_o (PDB: 7D76) complex. The distance between ligand and “toggle switch” were shown. **f**, Dose response curves of the C48/80, PAMP-12, C14^{linear} and SP induced G_{αq}–G_γ dissociation in MRGPRX2 overexpressing cells. Data from three independent experiments are presented as the mean ± SEM (n=3). All data were analysed by two-sided one-way ANOVA with Turkey test.

Extended Data Fig. 4 Effects of mutations in the ligand-binding pocket of MRGPRX2 on C48/80 induced dissociation of G_{αi}–G_γ or G_{αq}–G_γ.

a, Key interactions between the ethylamine group of the C48/80 and W243^{6.55}, F170^{ECL2}, C168^{4.64}, C180^{5.34}, W248^{6.60} of the MRGPRX2. The disulfide bond between the C168 and C180 was highlight in cyan. **b**, Key interactions between the amide of the ethylamine group of C48/80 and E164, D184 of MRGPRX2. Hydrogen bonds were showed in red dash. **c**, Important residues in MRGPRX2 that recognize the upper “Y” arm of C48/80. **d**, Elisa experiments to determine the expression levels of the MRGPRX2 wild type and corresponding mutants. Values are mean ± SEM from three independent experiments (n=3). n.s., no significance. All data were analysed by two-sided one-way ANOVA with Turkey test. **e**, Effects of mutations in the ligand-binding pocket residues of MRGPRX2 on C48/80–induced G_{αi}–G_γ dissociation in MRGPRX2 overexpressing HEK293 cells. Natural occurred SNPs were highlight in blue. ND, not detectable due to low signal. Values are the mean ± SEM of three independent experiments for the wild type (WT) and mutants (n=3). All data were determined by two-sided one-way ANOVA with Tukey test. *, P < 0.05; **, P < 0.01; ***, P < 0.001, n.s., no significant difference. (P = 0.0028, 0.0001, 0.0171, < 0.0001, 0.0045, ND, < 0.0001, < 0.0001, ND, 0.0224, 0.0178, < 0.0001, 0.0050, < 0.0001, 0.0008, 0.0007 from top to bottom.) **f**, Effects of mutations of key residues of MRGPRX2 on C48/80–induced G_{αq}–G_γ dissociation in MRGPRX2 overexpressing HEK293 cells. Values are the mean ± SEM of three independent experiments for the wild type (WT) and mutants (n=3) except E164A and D254A which are the mean ± SEM of five independent experiments (n=5). All data were determined by two-sided one-way ANOVA with Tukey test. ***, P < 0.001. (P < 0.0001, < 0.0001, < 0.0001, < 0.0001, < 0.0001, < 0.0001 from top to bottom).

Extended Data Fig. 5 FlAsH insertion site screening and FlAsH BRET experiments.

a, Schematic representation of the FlAsH-BRET assay. The Nluc was inserted at the N-terminal of wild-type MRGPRX2 and the FlAsH motifs were inserted at different positions of extracellular loops of the MRGPRX2. **b**, Detailed description of the FlAsH motif insertion sites at the extracellular loops of MRGPRX2. FlAsH motifs are labeled in red. **c**, Elisa experiments to determine the expression levels of the wild-type MRGPRX2 and FlAsH motif incorporated MRGRX2 mutants. Values are the mean ± SEM of three independent experiments for the wild type (WT) and mutants (n=3). n.s., no significance. All data were analysed by two-sided one-way ANOVA with Turkey test. **d**, Representative dose response curves of the conformational changes in response to binding of C48/80 reported by different MRGPRX2 FlAsH-BRET sensors. Data from three independent experiments are presented as the mean ± SEM (n=3). All data were analysed by two-sided one-way ANOVA with Turkey test. **e**, Elisa experiments to determine the expression levels of the wild type and corresponding mutants of MRGPRX2 using FlAsH-BRET sensor-3 construct. Data from three independent experiments are presented as the mean ± SEM (n=3). n.s., no significance. All data were analysed by two-sided one-way ANOVA with Turkey test. **f**, Representative dose response curve of the C48/80 induced BRET ratio in HEK 293 cell over-expressing FlAsH-BRET sensor-3 of MRGPRX2 (including wild type or corresponding mutants). Data from three independent experiments are presented as the mean ± SEM (n=3). **g**, Effects of mutations of the ligand-binding pocket residues of MRGPRX2 on changes in ΔpEC50 in response to stimulation of C48/80 or different polycationic compounds, evaluated using a FlAsH-BRET assay. Values are the mean ± SEM of three independent experiments for the wild type (WT) and mutants (n=3). All data were determined by two-sided one-way ANOVA with Tukey's test. *, P<0.05; **, P<0.01; ***, P<0.001, n.s., no significant difference. (From top to bottom, C48/80 panel: P<0.0001, <0.0001, 0.7568, <0.0001, <0.0001, <0.0001, <0.0001, 0.4929, <0.0001, <0.0001, 0.0003, <0.0001. Atracurium panel: P<0.0001, 0.0003, 0.0008, <0.0001, <0.0001, 0.001. Tubocurarine panel: P=0.0206, 0.011, 0.0005, 0.0064, 0.0011, 0.0017. Ciprofloxacin panel: P<0.0001, 0.0008, <0.0001, ND, ND, <0.0001. Rocuronium panel: P =0.0423, 0.0024, <0.0001, 0.0003, <0.0001, 0.0008). **h**, Structures of the representative polycationic compounds used in this study. **i**, Sequence alignment of the MRGPRX2 residues responsible for specific interaction with C48/80 with

the corresponding sequences of other MRGPR members. The key residues involved in the binding pocket are colored in red.

Extended Data Fig. 6 Effects of mutations of key residues of MRGPRX2 on polycationic compound drugs induced MRGPRX2 activation.

a, Effects of mutations of key residues of MRGPRX2 on polycationic compound drugs induced MRGPRX2 activation via $G_{\alpha i}-G_{\gamma}$ dissociation assay. The SNP mutations were highlighted in blue. ND, not detectable due to low signal. Data from three independent experiments are presented as the mean \pm SEM ($n=3$). (Atracurium: $P = 0.0018, 0.0013, \text{ND}, \text{ND}, < 0.0001, \text{ND}, 0.0064$, ND from top to bottom. Tubocurarine: $P = 0.0003, 0.0013, 0.0037, \text{ND}, 0.0064, < 0.0001, < 0.0001, \text{ND}$ from top to bottom. Ciprofloxacin: $P < 0.0001, 0.0003, \text{ND}, 0.0006, \text{ND}, \text{ND}, 0.0267, 0.0012$ from top to bottom. Rocuronium: $P = 0.0410, 0.0132, \text{ND}, \text{ND}, \text{ND}, 0.0078, 0.9383, \text{ND}$ from top to bottom). **b**, Effects of mutations of key residues of MRGPRX2 on polycationic compound drugs induced MRGPRX2 activation via $G_{\alpha q}-G_{\gamma}$ dissociation assay. The SNPs were highlight in blue. Values are the mean \pm SEM of three independent experiments for the wild type (WT) and mutants ($n=3$) except values of E164A and W243R in the ligand Ciprofloxacin, values of F170A, D184A, D184H, W243A and W243R in the ligand Rocuronium are the mean \pm SEM of five independent experiments ($n=5$). (Atracurium: $P < 0.0001, < 0.0001, 0.0181, 0.0010, 0.0013, < 0.0001, < 0.0001, < 0.0001$ from top to bottom. Tubocurarine: $P < 0.0001, < 0.0001, 0.0002, 0.0021, < 0.0001, 0.0056, 0.0004, < 0.0001$ from top to bottom. Ciprofloxacin: $P < 0.0001, 0.0001, < 0.0001, < 0.0001, < 0.0001, < 0.0001, 0.0007, < 0.0001$ from top to bottom. Rocuronium: $P < 0.0001, < 0.0001, < 0.0001, < 0.0001, 0.0020, < 0.0001, < 0.0001, < 0.0001$). **a-b**, Upper panel, bar graph representation of pEC50. Lower panel, dose response curves. Statistical differences between WT and mutations were determined by two-sided one-way ANOVA with Tukey test. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$, n.s., no significant difference.

Extended Data Fig. 7 The mode of interaction between peptidic allergens and MRGPRX2.

a: The EM density corresponding to PAMP-12 and potential interaction residues in MRGPRX2. In the EM density map of the PAMP-12–MRGPRX2–G_i complex, a potential density corresponding to PAMP-12 was identified. PAMP-12 lay flat and parallel to the plasma membrane, flanked by TM2-TM3 and ECL1 on one side and TM4-TM5 and ECL3 on the other side. **b:** The EM density corresponding to C14^{linear} and potential interaction residues in MRGPRX2. **c:** Alanine scanning mutagenesis and subsequent ligand-binding assays identified key residues in MRGPRX2 that recognize PAMP-12. Values are the mean ± SEM of three independent experiments for the wild type (WT) and mutants (n=3). Statistical differences between WT and mutant strains were determined by two-sided one-way ANOVA with Tukey's test. *, P<0.05; **, P<0.01; ***, P<0.001, n.s., no significant difference. (P=0.0003, 0.0805, 0.0757, 0.0053, 0.4707, 0.4720, 0.0009, 0.0208, 0.0049, 0.0002, 0.0001 from top to bottom). **d,** Effects of different PAMP-12 mutations on PAMP-12 binding. Values are the mean ± SEM of three independent experiments for the wild type (WT) and mutants (n=3) except W13A and R20 A which are the mean ± SEM of five independent experiments (n=5). **e,** Effects of different PAMP-12 mutations on PAMP-12 induced G_{αi}–G_γ dissociation in MRGPRX2 overexpressing HEK293 cells. Bar graph for EC50 was presented. The EC50 with five-fold change vs. WT was indicated by a dotted blue line. Statistical differences between PAMP-12 WT and mutations were determined by two-sided one-way ANOVA with Tukey test. ***, P<0.001. The curve data from three independent measurements are measured as mean ± SEM (n=3). (P=0.0004, <0.0001, <0.0001, 0.0057, <0.0001, 0.0004, <0.0001, <0.0001, 0.0022, 0.0015, <0.0001 from left to right). **f,** Alanine scanning mutagenesis and subsequent ligand-binding assays of the key residues in MRGPRX2 that recognize PAMP-12. Values are the mean ± SEM of three independent experiments for the wild type (WT) and mutants (n=3) except the T106A and F257A which are the mean ± SEM of five independent experiments (n=5). Statistical differences between WT and mutant strains were determined by two-sided one-way ANOVA with Tukey's test. *, P<0.05; **, P<0.01; ***, P<0.001, n.s., no significant difference. (P<0.0001, 0.0004, <0.0001, 0.0016, 0.0024, <0.0001, <0.0001, <0.0001, <0.0001, 0.0005, 0.0002, 0.0757, 0.0010, 0.2988 from top to bottom) **g,** Effects of different mutations within the

ligand-binding pocket of MRGPRX2 on PAMP-12 induced $G_{\alpha i}$ - G_{γ} dissociation in MRGPRX2 overexpressing cells. Statistical differences between MRGPRX2 WT and mutations were determined by two-sided one-way ANOVA with Tukey test. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$, n.s., no significant difference. The curve data from at least three independent measurements are measured as mean \pm SEM ($n=3$). ($P=0.0014, 0.0004, 0.0004, 0.0012, <0.0001, <0.0001, <0.0001, <0.0001, 0.0004, 0.0004, 0.4361, 0.0401, 0.2415$ from top to bottom).

Extended Data Fig. 8 Binding of PAMP-12 to the MRGPRX2.

a, Pairing of PAMP-12 mutants with MRGPRX2 mutants to identify hot spot interactions. We compared the binding capacity of PAMP-12 mutants with that of the PAMP-12 WT for MRGPRX2 mutations with significant deficiency in $G_{\alpha i}$ - G_{γ} dissociation. For a receptor mutation that had impaired response to PAMP-12 WT but show little deteriorated response to an PAMP-12 with a selective alanine substitution, the receptor mutated site and ligand mutated site could be paired with each other to identify potential hot spot interactions (highlight in red). Data are from three independent experiments. The original data of Figure 3a. **b**, The EM density corresponding to five residues from F9 to W13 of PAMP-12 and their interaction residues in ligand-binding pocket of MRGPRX2. EM density corresponding to PAMP-12 were highlight in orange. **c**, Structural representation of seven representative clusters derived by covalent docking of seven residues of PAMP-12 using Autodock4. **d**: Docking clusters were ranked by numbers in each cluster. The representative conformations in the classification with either the largest number (cluster 1) and the lowest binding energy (cluster 2) were used for MD simulations, respectively. **e**, The average RMSD value of the seven residues N14-R20 (N14, K15, W16, A17, L18, S19, R20) in PAMP-12 (upper panel) and RMSD of key residues in MRGPRX2 which directly interact with the seven residues (lower panel) during triplicate 200 ns MD simulations. **f-h**, Structural representations of the interactions surrounding the R10 (g), W13 (h), F9 and W16 (i) of PAMP-12 of MRGPRX2. Hydrogen bonds were showed in red dash. **i**, Effects of different mutations within the ligand-binding pocket of MRGPRX2 on PAMP-12 induced $G_{\alpha q}$ - G_{γ} dissociation in MRGPRX2

overexpressing cells. Statistical differences between MRGPRX2 WT and mutations were determined by two-sided one-way ANOVA with Tukey test. ***, P < 0.001. The curve data from at least three independent measurements are measured as mean ± SEM (n=3). (P < 0.0001, < 0.0001, < 0.0001, < 0.0001, 0.0007, < 0.0001, < 0.0001 from top to bottom).

Extended Data Fig. 9 Binding of C14^{linear} to the MRGPRX2.

a, Heatmap of pairing of C14^{linear} mutants with MRGPRX2 WT and MRGPRX2 alanine scanning mutants. The receptor mutants that did not show significantly decreased EC50 values compared to those of the WT receptor when binding to a specific C14^{linear} mutant are highlighted by red color. Biochemical characterization indicated that P1 and K3 of C14^{linear} paired with the hydrophobic residue F170^{ECL2} and the negatively charged residues E164^{4.60} and D184^{5.38} of MRGPRX2, respectively. Whereas F5 of C14^{linear} interacted with F170^{ECL2} and F257^{7.31} of MRGPRX2, the F6 mutation of C14^{linear} functionally coupled with the F239^{6.51} and W243^{6.55} mutations. **b**, Effects of different C14^{linear} mutations on C14^{linear} induced G_{αi}–G_γ dissociation in MRGPRX2 overexpressing HEK293 cells.

Statistical differences between C14^{linear} WT and mutations were determined by two-sided one-way ANOVA with Tukey test. P < 0.01; ***, P < 0.001, n.s., no significant difference. Data are from three independent experiments. (n=3). (P < 0.0001, < 0.0001, 0.0557, 0.0011, < 0.0001, < 0.0001, 0.0199, 0.0036, 0.0006, 0.0074, 0.0071, < 0.0001 from left to right). **c**, Effects of different mutations within the ligand-binding pocket of MRGPRX2 on C14^{linear} induced G_{αi}–G_γ dissociation in MRGPRX2 overexpressing cells.

Statistical differences between MRGPRX2 WT and mutations were determined by two-sided one-way ANOVA with Tukey test. **, P < 0.01; ***, P < 0.001, n.s., no significant difference. Data are from three independent experiments (n=3). (P=0.0023, 0.0005, 0.0002, < 0.0001, 0.0002, < 0.0001, 0.0034, 0.0029, 0.0450, 0.0168 from left to right). **d**,

Pairing of C14^{linear} mutants with MRGPRX2 mutants to identify hot spot interactions. We compared the binding capacity of C14^{linear} mutants with that of the C14^{linear} WT for MRGPRX2 mutations with significant deficiency in G_{αi}–G_γ dissociation. For a receptor mutation that had impaired

response to C14^{linear} WT but show little deteriorated response to an C14^{linear} with a selective alanine substitution, the receptor mutated site and ligand mutated site could be paired with each other to identify potential hot spot interactions (highlight in red). Data are from three independent experiments. **e-f**, Structural representation and EM density of C14^{linear} peptide (e) and C14^{circular} peptide (f). Because the C14^{linear}-MRGPRX2-G_i complex had higher resolution, unbiased alanine scanning mutagenesis of each residue in C14^{linear} and the pocket residues of MRGPRX2 was exploited to assign the primary sequence of C14^{linear} in the EM density. **g**, Structural model of the interactions between the P1-W7 segment of the C14^{linear} peptide and MRGPRX2, which were derived by computational simulation. **h**, RMSD of the seven residues P1-W7 in C14^{linear} (upper panel) and of key residues in MRGPRX2 which directly interact with the seven residues (lower panel) during triplicate 200 ns MD simulations. **i-j**, The average RMSD and +/- s.e.m. values of the seven residues P1-W7 in C14^{linear} (upper panel) and of key residues in MRGPRX2 which directly interact with the seven residues (lower panel) from triplicate 200 ns MD simulations. **k**, Structural representation and EM density corresponding to S11, S12, C13 and K14 of C14^{linear}.

[Extended Data Fig. 10 Binding of SP and Consensus peptidic motif recognized by MRGPRX2.](#)

a, Structural representation and EM density corresponding to four residues from R1 to P4 of SP. The EM density of the SP-MRGPRX2-G_i complex structure enabled us to position the main chain of the first four residues of SP, which assumed different configurations compared to PAMP-12 and C14^{linear}, likely due to the presence of the two prolines in the SP sequence. There was continuous EM density following the first four residues. **b**, The average RMSD value of the seven residues R1-F7 in SP (upper panel) and of key residues in MRGPRX2 that directly interact with the seven residues (lower panel) during triplicate 200 ns of MD simulation. **c**, Effects of different mutations within the ligand-binding pocket of MRGPRX2 on SP-induced G_{αi}-G_γ dissociation in MRGPRX2-overexpressing cells. Data from three independent experiments are presented as the mean ± SEM (n=3).

(P=0.0007, 0.0012, 0.0003, 0.0041, <0.0001, ND, 0.0002, ND, 0.0003 from top to bottom). **d**, Interaction differences in the pocket of the MRGPRX2 bound by C48/80, PAMP-12 and C14^{linear} and SP. Residues of MRGPRX2 that interact with ligands are indicated by cyan circle. Residues of MRGPRX2 showing no interaction with C48/80 or PAMP-12 are indicated by blank circle. Residues of MRGPRX2 that interact with simulated parts of ligands and verified by unbiased alanine scanning are indicated by green. **e**, The binding energy contribution of residues in PAMP-12 as calculated by the molecular mechanics Poisson–Boltzmann surface area (MM-PBSA) method. **f**, Mutations of PAMP-12, C14^{linear}, Sermorelin, Exenatide and Enfuvirtide used in Figure 2e. **g**, Effects of the $\varphi^{p9}(X_{0-1}) R/K^{p10}(X_2) \varphi^{p13}(X_{2-3}) \varphi^{p16}(X_3) R/K^{p20}$ motif mutations on PAMP-12-, C14^{linear} induced MRGPRX2 activation evaluated via G_{αi}–G_γ dissociation assay. Data from three independent experiments are presented as the mean ± SEM (n=3). All data were analysed by two-sided one-way ANOVA with Turkey test. **h**, Effects of mutations of key residues recognizing the $\varphi p9(X0-1) R/Kp10(X2) \varphi p13(X2-3) \varphi p16(X3) R/Kp20$ motif on MRGPRX2 activation in response to PAMP-12, cortistatin-14, sermorelin, enfuvirtide or exenatide evaluated via G_{αi}–G_γ dissociation assay. (From top to bottom, for PAMP-12 panel: P=0.0004, 0.0004, 0.0012, <0.0001, <0.0001, <0.0001, <0.0001; C14^{linear} panel: 0.0023, 0.0013, 0.0020, 0.0005, 0.0002, <0.0001, <0.0001; sermorelin panel: 0.0007, ND, <0.0001, <0.0001, <0.0001, <0.0001, 0.0004; enfuvirtide panel: ND, ND, 0.0001, <0.0001, <0.0001, 0.1193, <0.0001; exenatide panel: 0.0075, ND, ND, <0.0001, ND, 0.0010, 0.0010).

Extended Data Fig. 11 Active state and G protein coupling of MRGPRX2.

a, Structural comparison of TM3 and TM6 in the MRGPRX2–G_{i1} complex (green), CB1–G_i complex (cyan, PDB: 6N4B) and μ-opioid–G_i complex (blue, PDB: 6N4B). Compared to the corresponding structures in the CB1 and μ-opioid receptors, the upper TM6 of MRGPRX2 is distorted inwardly at the position of G236^{6,48} in MRGPRX2. **b**, Structural comparison of residues involved in kink formation in MRGPRX2 (green), CB1 (cyan,

PDB: 6N4B) and μ -opioid receptor (blue, PDB: 6N4B). The engagement of the phenolic oxygen of Y113^{3,36} and the main chain of G236^{6,48} in MRGPRX2 at the kink position was further stabilized by packing between F239^{6,51} and Y113^{3,36}. Toggle switch W^{6,48} of the CB1, ADGRG3 and μ -opioid receptors was replaced by G236^{6,48} in MRGPRX2. H-bonds are depicted as dashed lines. **c**, Structural comparison of TM6 in MRGPRX2 (green), CB1 (cyan, PDB: 6N4B), μ -opioid receptor (light blue, PDB: 6N4B), ADGRG3 (yellow orange, PDB: 7D76), GLPR1 (blue, PDB: 6B3J) and GABA1 (grey, PDB: 7C7Q). **d**, Cutaway view of G^{6,48}XXF^{6,52} motif in C14^{linear}–MRGPRX2–G_{i1} and C48/80–MRGPRX2–G_{i1} complex. Hydrogen bonds are depicted as red dashed lines. **e**, Schematic representation of the FlAsH-BRET assay. The Nluc was inserted at the P102 position close to extracellular end of TM3 and the FlAsH motif were inserted at the K251 position of the extracellular end of TM6 (upper panel). Representative dose response curves of the conformational changes in response to binding of C 48/80 or Compound 2 reported by FlAsH-BRET assay. A future solution of the inactive structure of MRGPRX2 and comparison with our active structure will provide further evidence of whether kink formation of TM6 is required for the activation of MRGPRX2. Data from three independent experiments are presented as the mean \pm SEM (n=3) (lower panel). **f**, Comparison of the D/ERY motif and NPXXY motif in MRGPRX2–G_i (green), CB1–G_i (gray) and μ -opioid (yellow) complex, which showed similar conformations. **g**, Structural representation of the MRGPRX2–G_{i1} complex (green). TM3-TM5 and ICL1-ICL3 of MRGPRX2 and the α N-helix, α 5-helix, i3 loop, and β 2- β 3 loop of G _{α i} are highlighted. **h**, Comparison of the G_i coupling interfaces in cryo-EM structures of MRGPRX2–G_{i1} complexes, the μ -opioid–G_i complex (PDB: 6N4B), the CB1–G_i complex (PDB: 6N4B), and the model of MRGPRX2–G _{α q} generated by computational simulation. Residues of MRGPRX2, μ -opioid receptor or CB1 in contact with G_i are illustrated as green or blue dots. The basic charged R that contributes to the interface is highlighted in blue. Residues that contact G _{α i} but not G _{α q} are illustrated as brown dots. **i**, Detailed interactions between ICL1 and ICL3 of MRGPRX2 and the G_{i1} trimer. R60^{ICL1} inserted in a hydrophobic pocket created by L55 and F335

and formed a hydrogen bond with S334 of the G_{β1} subunit, R214^{ICL3} formed a cation-π interaction with Y320 and potential charge interactions with E318 of G_{αi1}.

Extended Data Fig. 12 MRGPRX2 interactions with G_i and G_q:

a, Effects of mutations in the MRGPRX2 along the G_i trimer interface on PAMP-12 induced G_{αi}–G_γ dissociation. Values are the mean ± SEM of three independent experiments for the wild type (WT) and mutants (n=3). Statistical differences between WT and mutations were determined by two-sided one-way ANOVA with Tukey test. *, P < 0.05; **, P < 0.01; ***, P < 0.001, n.s., no significant difference. (P < 0.0001, < 0.0001, 0.0022, 0.0018, < 0.0001, 0.0008, < 0.0001, 0.0395, 0.0009, 0.0009, 0.0003, 0.0006, 0.0010, 0.0178, 0.0018, 0.0003, 0.0003 from top to bottom).

b, RMSD of PAMP-12–MRGPRX2 (upper panel) and G_{αi} (lower panel) during 200 ns MD simulations. **c**, RMSD of PAMP-12–MRGPRX2 (upper panel) and G_{aq} (lower panel) during 200 ns MD simulations. **d**, PAMP-12–MRGPRX2–G_{aq} complex model after 200 ns MD simulation using M1R-G_q/G₁₁ complex as the G_q template (PDB: 6OIL). **e**, Structural comparison of PAMP-12–MRGPRX2–G_{aq} stimulated model (orange) and PAMP-12–MRGPRX2–G_i complex (green). **f**, Effects of mutations in the MRGPRX2 along the G_{aq} trimer interface on PAMP-12 induced G_{aq}–G_γ protein dissociation. Values are the mean ± SEM of three independent experiments for the wild type (WT) and mutants (n=3). Statistical differences between WT and mutations were determined by two-sided one-way ANOVA with Tukey test. *, P < 0.05; **, P < 0.01; ***, P < 0.001, n.s., no significant difference. (P < 0.0001, 0.0016, < 0.0001, < 0.0001, 0.9226, 0.8226, 0.8815, < 0.0001, < 0.0001, < 0.0001, < 0.0001, < 0.0001, < 0.0001, < 0.0001, < 0.0001, < 0.0001, 0.0319 from top to bottom). **g-i**, Because the OD1 and OD2 in the carboxylate group of aspartic acid (Asp), or OE1 and OE2 in the carboxylate group of glutamic acid (Glu), or the NH1 and NH2 in guanidinium group of arginine (Arg) are equal atoms respectively, we named the ODc:NH_C pair as closer pair for the contacting Asp: Arg,

OE_C : NH_C pair as closer pair for the contacting Glu: Arg; and the remote OD_R atom, OE_R atom, or NH_R atom was named in Asp, Glu or Arg respectively. **g**, The average distances between NH_C in R140 of MRGPRX2 and oxygen atom of N198 of $\text{G}_{\alpha q}$ in MRGPRX2– $\text{G}_{\alpha q}$ complex or nitrogen atom of the guanidium group of R140 (N_E) of MRGPRX2 and OD_C of D193 of $\text{G}_{\alpha i 1}$ in MRGPRX2– $\text{G}_{i 1}$ complex calculated from triplicate 200 ns MD simulations. **h**, The average distances between NH_C in R143 of MRGPRX2 and NH_C in R24 of $\text{G}_{\alpha q}$ in MRGPRX2– $\text{G}_{\alpha q}$ complex or NH_C in R143 of MRGPRX2 and OE_C in E28 of $\text{G}_{\alpha i 1}$ in MRGPRX2– $\text{G}_{\alpha i 1}$ complex calculated from triplicate 200 ns MD simulations. **i**, The average distances between NH_C in R214 of MRGPRX2 and carbon atom of the methyl group I323(C_D) in MRGPRX2– G_q complex or NH_C in R214 of MRGPRX2 and OE_C in E318 of $\text{G}_{\alpha i 1}$ in MRGPRX2– $\text{G}_{\alpha i 1}$ complex calculated from triplicate 200 ns MD simulations. **j**: The effects of selected SNPs on ligand-induced MRGPRX2 activation. SNPs located in the ligand or G_i protein interface of MRGPRX2 are summarized. The $\text{G}_{\alpha i}/\text{G}_{\alpha q}$ – G_γ dissociation assays were measured in the current study, whereas calcium mobilization and mast cell degranulation were obtained from previous studies²⁰.

Extended Data Table 1 Cryo-EM data collection, model refinement and validation statistics

Supplementary information

Supplementary Information

Supplementary Figs. 1–7 and Supplementary Tables 1–6.

Reporting Summary

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Yang, F., Guo, L., Li, Y. *et al.* Structure, function and pharmacology of human itch receptor complexes. *Nature* **600**, 164–169 (2021).
<https://doi.org/10.1038/s41586-021-04077-y>

- Received: 15 February 2021
- Accepted: 30 September 2021
- Published: 17 November 2021
- Issue Date: 02 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04077-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

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04077-y>.

- Article
- [Published: 17 November 2021](#)

Structure, function and pharmacology of human itch GPCRs

- [Can Cao¹](#),
na1
- [Hye Jin Kang¹](#),
na1
- [Isha Singh²](#),
- [He Chen³](#),
- [Chengwei Zhang³](#),
- [Wenlei Ye](#) [ORCID: orcid.org/0000-0002-4694-1493⁴](#),
- [Byron W. Hayes⁵](#),
- [Jing Liu](#) [ORCID: orcid.org/0000-0003-4740-243X³](#),
- [Ryan H. Gumpfer¹](#),
- [Brian J. Bender²](#),
- [Samuel T. Slocum¹](#),
- [Brian E. Krumm](#) [ORCID: orcid.org/0000-0001-6039-4767¹](#),
- [Katherine Lansu¹](#),
- [John D. McCory](#) [ORCID: orcid.org/0000-0001-7555-9413¹](#) [nAff9](#),
- [Wesley K. Kroese¹](#),
- [Justin G. English¹](#) [nAff10](#),
- [Jeffrey F. DiBerto¹](#),
- [Reid H. J. Olsen¹](#),
- [Xi-Ping Huang¹](#),
- [Shicheng Zhang](#) [ORCID: orcid.org/0000-0002-5245-443X¹](#),
- [Yongfeng Liu](#) [ORCID: orcid.org/0000-0002-8649-7892¹](#),
- [Kuglae Kim¹](#),
- [Joel Karpiak²](#),

- [Lily Y. Jan](#) [ORCID: orcid.org/0000-0003-3938-8498^{4,6}](#),
- [Soman N. Abraham^{5,7}](#),
- [Jian Jin](#) [ORCID: orcid.org/0000-0002-2387-3862³](#),
- [Brian K. Shoichet](#) [ORCID: orcid.org/0000-0002-6098-7367²](#),
- [Jonathan F. Fay](#) [ORCID: orcid.org/0000-0003-1822-2384⁸](#) &
- [Bryan L. Roth](#) [ORCID: orcid.org/0000-0002-0561-6520¹](#)

[*Nature*](#) volume 600, pages 170–175 (2021)

- 8263 Accesses
- 122 Altmetric
- [Metrics details](#)

Subjects

- [Chemical biology](#)
- [Structural biology](#)

Abstract

The MRGPRX family of receptors (MRGPRX1–4) is a family of mast-related G-protein-coupled receptors that have evolved relatively recently¹. Of these, MRGPRX2 and MRGPRX4 are key physiological and pathological mediators of itch and related mast cell-mediated hypersensitivity reactions^{2,3,4,5}. MRGPRX2 couples to both G_i and G_q in mast cells⁶. Here we describe agonist-stabilized structures of MRGPRX2 coupled to G_{i1} and G_q in ternary complexes with the endogenous peptide cortistatin-14 and with a synthetic agonist probe, respectively, and the development of potent antagonist probes for MRGPRX2. We also describe a specific MRGPRX4 agonist and the structure of this agonist in a complex with MRGPRX4 and G_q . Together, these findings should accelerate the structure-guided discovery of therapeutic agents for pain, itch and mast cell-mediated hypersensitivity.

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: Cryo-EM structures of MRGPRX2 complexes.

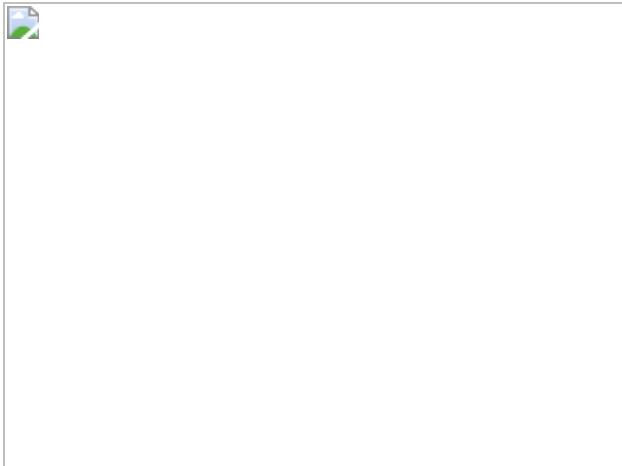


Fig. 2: G-protein coupling of MRGPRX2.

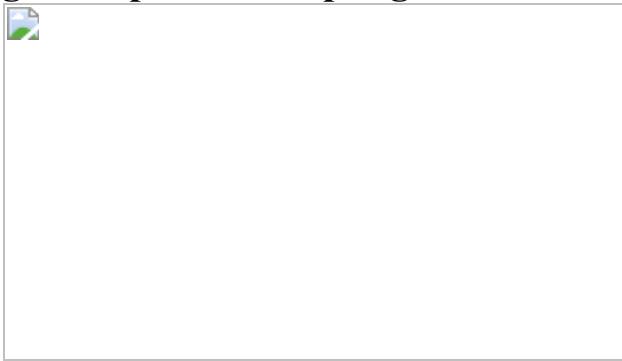


Fig. 3: Discovery of MRGPRX2-selective inverse agonists.

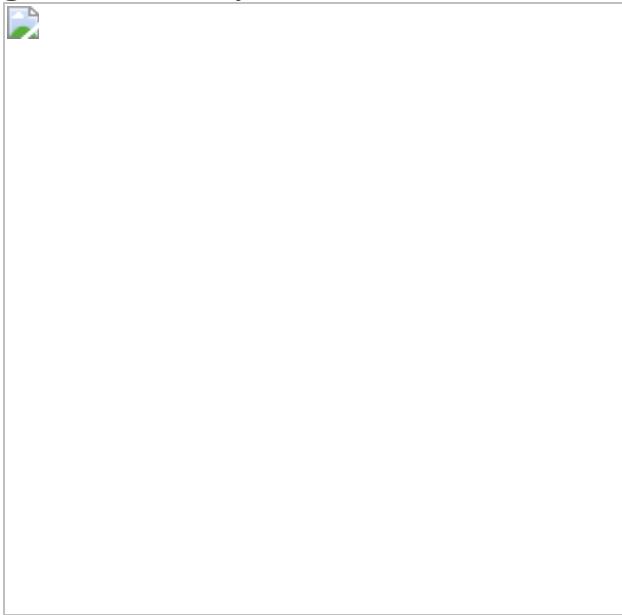
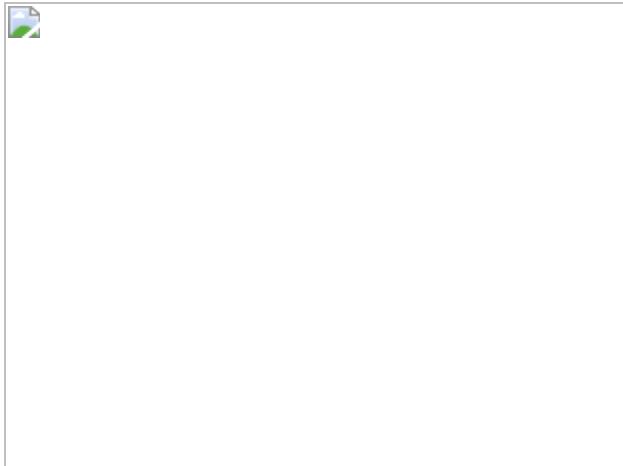


Fig. 4: Agonist discovery and the cryo-EM structure of MRGPRX4.



Data availability

The coordinate and cryo-EM map of MRGPRX2–G_q–cortistatin-14, MRGPRX2–G_{i1}–cortistatin-14, MRGPRX2–G_q–(R)-ZINC-3573, MRGPRX2–G_{i1}–(R)-ZINC-3573 and MRGPRX4–G_q–MS47134 have been deposited to PDB (EMDB) with accession codes [7S8L \(EMD-24896\)](#), [7S8M \(EMD-24897\)](#), [7S8N \(EMD-24898\)](#), [7S8O \(EMD-24899\)](#) and [7S8P \(EMD-24900\)](#), respectively. The cryoEM micrographs of MRGPRX4–G_q–MS47134, MRGPRX2–G_q–cortistatin-14, MRGPRX2–G_q–(R)-ZINC-3573, MRGPRX2–G_{i1}–cortistatin-14 and MRGPRX2–G_{i1}–(R)-ZINC-3573 have been deposited in the EMPIAR database (<https://www.ebi.ac.uk/empiar/>) with accession numbers EMPIAR-10852, EMPIAR-10853, EMPIAR-10854, EMPIAR-10855 and EMPIAR-10856, respectively. The MRGPRX2 antagonist C9 and C9-6 and negative control C7, and the MRGPRX4 agonist MS47134 and negative control X2-2 will be made available via Sigma-Millipore.

References

1. 1.

Zylka, M. J., Dong, X., Southwell, A. L. & Anderson, D. J. Atypical expansion in mice of the sensory neuron-specific Mrg G protein-

coupled receptor family. *Proc. Natl Acad. Sci. USA* **100**, 10043–10048 (2003).

2. 2.

McNeil, B. D. et al. Identification of a mast-cell-specific receptor crucial for pseudo-allergic drug reactions. *Nature* **519**, 237–241 (2015).

3. 3.

Green, D. P., Limjunyawong, N., Gour, N., Pundir, P. & Dong, X. A mast-cell-specific receptor mediates neurogenic inflammation and pain. *Neuron* **101**, 412–420.e413 (2019).

4. 4.

Yu, H. et al. MRGPRX4 is a bile acid receptor for human cholestatic itch. *eLife* **8**, e48431 (2019).

5. 5.

Meixiong, J., Vasavda, C., Snyder, S. H. & Dong, X. MRGPRX4 is a G protein-coupled receptor activated by bile acids that may contribute to cholestatic pruritus. *Proc. Natl Acad. Sci. USA* **116**, 10525–10530 (2019).

6. 6.

Chompunud Na Ayudhya, C., Roy, S., Alkanfari, I., Ganguly, A. & Ali, H. Identification of gain and loss of function missense variants in MRGPRX2's transmembrane and intracellular domains for mast cell activation by substance P. *Int. J. Mol. Sci.* **20**, 5247 (2019).

7. 7.

Ikoma, A., Steinhoff, M., Stander, S., Yosipovitch, G. & Schmelz, M. The neurobiology of itch. *Nat. Rev. Neurosci.* **7**, 535–547 (2006).

8. 8.

Greaves, M. W. & Wall, P. D. Pathophysiology of itching. *Lancet* **348**, 938–940 (1996).

9. 9.

Liu, Q. et al. Sensory neuron-specific GPCR Mrgprs are itch receptors mediating chloroquine-induced pruritus. *Cell* **139**, 1353–1365 (2009).

10. 10.

Lembo, P. M. et al. Proenkephalin A gene products activate a new family of sensory neuron–specific GPCRs. *Nat. Neurosci.* **5**, 201–209 (2002).

11. 11.

Azimi, E. et al. Dual action of neurokinin-1 antagonists on Mas-related GPCRs. *JCI Insight* **1**, e89362 (2016).

12. 12.

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).

13. 13.

Lansu, K. et al. In silico design of novel probes for the atypical opioid receptor MRGPRX2. *Nat. Chem. Biol.* **13**, 529–536 (2017).

14. 14.

Subramanian, H. et al. β -Defensins activate human mast cells via Mas-related gene X2. *J. Immunol.* **191**, 345–352 (2013).

15. 15.

Olsen, R. H. J. et al. TRUPATH, an open-source biosensor platform for interrogating the GPCR transducerome. *Nat. Chem. Biol.* **16**, 841–849 (2020).

16. 16.

English, J. G. et al. VEGAS as a platform for facile directed evolution in mammalian cells. *Cell* **178**, 748–761.e717 (2019).

17. 17.

Wacker, D. et al. Structural features for functional selectivity at serotonin receptors. *Science* **340**, 615–619 (2013).

18. 18.

Lu, L., Kulka, M. & Unsworth, L. D. Peptide-mediated mast cell activation: ligand similarities for receptor recognition and protease-induced regulation. *J. Leukoc. Biol.* **102**, 237–251 (2017).

19. 19.

Che, T. et al. Structure of the nanobody-stabilized active state of the kappa opioid receptor. *Cell* **172**, 55–67.e15 (2018).

20. 20.

Che, T. et al. Nanobody-enabled monitoring of kappa opioid receptor states. *Nat. Commun.* **11**, 1145 (2020).

21. 21.

Rasmussen, S. G. F. et al. Crystal structure of the β 2 adrenergic receptor–G_s protein complex. *Nature* **477**, 549–555 (2011).

22. 22.

Rosenbaum, D. M. et al. GPCR engineering yields high-resolution structural insights into β 2-adrenergic receptor function. *Science* **318**,

1266–1273 (2007).

23. 23.

Ogasawara, H., Furuno, M., Edamura, K. & Noguchi, M. Novel MRGPRX2 antagonists inhibit IgE-independent activation of human umbilical cord blood-derived mast cells. *J. Leukoc. Biol.* **106**, 1069–1077 (2019).

24. 24.

Shinkai, H. et al. *N*-(cyclohexylcarbonyl)-d-phenylalanines and related compounds. A new class of oral hypoglycemic agents. 2. *J. Med. Chem.* **32**, 1436–1441 (1989).

25. 25.

Irwin, J. J. & Shoichet, B. K. ZINC—a free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **45**, 177–182 (2005).

26. 26.

Meixiong, J. et al. Identification of a bilirubin receptor that may mediate a component of cholestatic itch. *eLife* **8**, e44116 (2019).

27. 27.

Azimi, E., Reddy, V. B. & Lerner, E. A. Brief communication: MRGPRX2, atopic dermatitis and red man syndrome. *Itch (Phila)* **2**, e5 (2017).

28. 28.

Chen, E. et al. Inflamed ulcerative colitis regions associated to MRGPRX2-mediated mast cell degranulation and cell activation modules, defining a new therapeutic target. *Gastroenterology* **160**, 1709–1724 (2021).

29. 29.

Kozlitina, J. et al. An African-specific haplotype in MRGPRX4 is associated with menthol cigarette smoking. *PLoS Genet.* **15**, e1007916 (2019).

30. 30.

Li, Z. et al. Targeting human Mas-related G protein-coupled receptor X1 to inhibit persistent pain. *Proc. Natl Acad. Sci. U.S.A.* **114**, E1996–E2005 (2017).

31. 31.

Thapaliya, M., Chompunud Na Ayudhya, C., Amponnawarat, A., Roy, S. & Ali, H. Mast cell-specific MRGPRX2: a key modulator of neuro-immune interaction in allergic diseases. *Curr. Allergy Asthma Rep.* **21**, 3 (2021).

32. 32.

Kim, K. et al. Structure of a hallucinogen activated G_q-coupled 5-HT2a serotonin receptor. *Cell* **182**, 1574–1588.e1519 (2020).

33. 33.

Draper-Joyce, C. J. et al. Structure of the adenosine-bound human adenosine A₁ receptor-G_i complex. *Nature* **558**, 559–563 (2018).

34. 34.

Koehl, A. et al. Structure of the μ-opioid receptor–G_i protein complex. *Nature* **558**, 547–552 (2018).

35. 35.

Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51

(2005).

36. 36.

Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).

37. 37.

Punjani, A., Zhang, H. & Fleet, D. J. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nat. Methods* **17**, 1214–1221 (2020).

38. 38.

Bepler, T., Kelley, K., Noble, A. J. & Berger, B. Topaz-Denoise: general deep denoising models for cryoEM and cryoET. *Nat. Commun.* **11**, 5208 (2020).

39. 39.

Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003).

40. 40.

Heymann, J. B. & Belnap, D. M. Bsoft: image processing and molecular modelling for electron microscopy. *J. Struct. Biol.* **157**, 3–18 (2007).

41. 41.

Sanchez-Garcia, R. et al. DeepEMhacer: a deep learning solution for cryo-EM volume post-processing. *Commun. Biol.* **4**, 874 (2021).

42. 42.

Grant, T., Rohou, A. & Grigorieff, N. cisTEM, user-friendly software for single-particle image processing. *eLife* **7**, e35383 (2018).

43. 43.

Xing, C. et al. Cryo-EM structure of the human cannabinoid receptor CB₂–G_i signaling complex. *Cell* **180**, 645–654.e613 (2020).

44. 44.

Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

45. 45.

Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).

46. 46.

Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).

47. 47.

Robertson, M. J., van Zundert, G. C. P., Borrelli, K. & Skiniotis, G. GemSpot: a pipeline for robust modeling of ligands into cryo-EM maps. *Structure* **28**, 707–716.e703 (2020).

48. 48.

Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D* **66**, 12–21 (2010).

49. 49.

Besnard, J. et al. Automated design of ligands to polypharmacological profiles. *Nature* **492**, 215–220 (2012).

50. 50.

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).

51. 51.

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).

Acknowledgements

This work was supported by NIH grants U24DA116195 (to B.L.R., B.K.S. and J.J.) and R35GM122481 (to B.K.S.), and by the Michael Hooker Distinguished Professorship to B.L.R. and NIH grant R01-DK121969, R01-DK121032 and R56-AI139620 to S.N.A. We thank J. Peck and J. Strauss of the UNC Cryo-EM Core Facility for technical assistance in this project. W.Y. is partly supported by Program for Breakthrough Biomedical Research funded by the Sandler Foundation, University of California, San Francisco. L.Y.J. is a Howard Hughes Medical Institute investigator. We thank S.-L. Shyng for the plasmids encoding human Kir6.2 and SUR1. The Titan X Pascal used for this research was kindly donated to J.F.F. by the NVIDIA Corporation.

Author information

Author notes

1. John D. McCory

Present address: Department of Cell Biology, Neurobiology and Anatomy, Medical College of Wisconsin, Milwaukee, WI, USA

2. Justin G. English

Present address: Department of Biochemistry, University of Utah
School of Medicine, Salt Lake City, UT, USA

3. These authors contributed equally: Can Cao, Hye Jin Kang

Affiliations

1. Department of Pharmacology, University of North Carolina at Chapel Hill School of Medicine, Chapel Hill, NC, USA

Can Cao, Hye Jin Kang, Ryan H. Gumpfer, Samuel T. Slocum, Brian E. Krumm, Katherine Lansu, John D. McCorry, Wesley K. Kroeze, Justin G. English, Jeffrey F. DiBerto, Reid H. J. Olsen, Xiping Huang, Shicheng Zhang, Yongfeng Liu, Kuglae Kim & Bryan L. Roth

2. Department of Pharmaceutical Sciences, University of California San Francisco, School of Medicine, San Francisco, CA, USA

Isha Singh, Brian J. Bender, Joel Karpiak & Brian K. Shoichet

3. Mount Sinai Center for Therapeutics Discovery, Departments of Pharmacological Sciences, Oncological Sciences and Neuroscience, Tisch Cancer Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA

He Chen, Chengwei Zhang, Jing Liu & Jian Jin

4. Department of Physiology, University of California, San Francisco, San Francisco, CA, USA

Wenlei Ye & Lily Y. Jan

5. Department of Pathology, Duke University Medical Center, Durham, NC, USA

Byron W. Hayes & Soman N. Abraham

6. Howard Hughes Medical Institute, San Francisco, CA, USA

Lily Y. Jan

7. Program in Emerging Infectious Diseases, Duke–National University of Singapore, Singapore, Singapore

Soman N. Abraham

8. Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill School of Medicine, Chapel Hill, NC, USA

Jonathan F. Fay

Contributions

C.C. designed the experiments, performed the cloning, expression and purification of all the signalling complexes for cryo-EM study, built the models, refined the structures, performed BRET assays, participated in MRGPRX4 drug screening, analysed the data, and prepared the figures, tables and manuscript. H.J.K. designed experiments, performed drug screening and functional assays, analysed the data and assisted in preparing the manuscript, and reviewed and edited the manuscript. J.F.F. made the grids, and collected and processed the cryo-EM data. I.S. performed the SAR study for MRGPRX2. H.C., C.Z. and J.L. performed the SAR study for MRGPRX4, and designed, synthesized and characterized the MRGPRX4 agonists. W.Y. performed electrophysiology for Kir6.2/SUR1. R.H.G. assisted in the modelling and validation of structures. B.W.H. performed the LAD2 mast cell degranulation assay. B.J.B. assisted in the structure modelling. J.K. supported agonist discovery and optimization for MRGPRX4. S.T.S. performed the GPCRome assay. B.E.K. designed the G_q protein construct and assisted in the reviewing and editing of the manuscript. K.L., J.D.M., W.K.K., J.F.D. and R.H.J.O. performed the initial screening for MRGPRX4 compounds. J.G.E., X.-P.H. and Y.L. helped with the initial functional assays. S.Z. and K.K. assisted in the protein expression. S.N.A. guided the LAD2 mast cell degranulation assay. J.J.

supervised the medicinal chemistry experiments. B.K.S. supervised the docking and compound design and edited the manuscript. L.Y.J. contributed to funding application. B.L.R. supervised the entire project, guided the structural and functional work and prepared the manuscript.

Corresponding authors

Correspondence to [Brian K. Shoichet](#), [Jonathan F. Fay](#) or [Bryan L. Roth](#).

Ethics declarations

Competing interests

A patent describing the MRGPRX2 antagonists has been filed by UCSF listing B.L.R., B.K.S., C.C., I.S. and H.J.K. 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 MRGPRX2 transducerome screening using TRUPATH.](#)

a, MRGPRX2 effectively couples to 14 distinct G proteins upon stimulation of agonists (*R*)-ZINC-3573 and cortistatin-14 (C-14) in HEK 293T cells. Net BRET values of MRGPRX2 together with positive controls of either neurotensin receptor 1 (NTSR1, agonist NT1-13) or β_2 AR (agonist isoproterenol) are shown in each panel. Data represent mean \pm s.e.m. of $n = 3$ biological replicates. **b**, Heatmap of the relative log potency (logEC₅₀) of

(*R*)-ZINC-3573 and cortistatin-14 for 14 distinct G proteins. **c**, Heatmap of the relative efficacy (E_{\max}) of (*R*)-ZINC-3573 and cortistatin-14 for 14 distinct G proteins.

Extended Data Fig. 2 CryoEM images and data-processing of MRGPRX2-G_q-(*R*)-ZINC-3573, MRGPRX2-G_{i1}-(*R*)-ZINC-3573 and MRGPRX4-G_q-MS47134 complex.

a–c, Representative motion corrected cryo-EM micrographs (scale bar, 100 nm) of respective ligand bound GPCR heterotrimeric complex particles imaged at a nominal 45k x magnification and representative two-dimensional class averages. The experiment was repeated three times with similar result. The exact number of movies and particles used for each complex are shown in the flow chart. **d–f**, Flow chart of cryo-EM data processing, GSFSC plot of auto-masked final map (black) and map-to-model real-space cross correlation (red) as calculated from phenix.mtriangle. **g–i**, Respective polar plots of particle angular distributions and local resolution estimations heat maps. **j–l**, Local cryo-EM density maps of TM1-7, respective ligands, and α 5 and α N helix of respective G-protein.

Extended Data Fig. 3 CryoEM images and data-processing of MRGPRX2-G_q-Cortistatin-14 and MRGPRX2-G_{i1}-Cortistatin-14 complex.

a, b, Representative motion corrected cryo-EM micrograph (scale bar, 100 nm) of MRGPRX2 G-protein cortistatin-14 (C14) particles imaged at a nominal 45k x magnification and representative two-dimensional class averages. The experiment was repeated three times with similar result. The exact number of movies and particles used for each complex are shown in the flow chart. **c, d**, Flow chart of cryo-EM data processing. GSFSC plot of auto-masked final map (black) and map-to-model real-space cross correlation (red) as calculated from phenix.mtriangle. **e, f**, Viewing direction distribution and local resolution estimation heat maps. **g, h**, Local cryo-EM density maps of TM1-7, Cortistatin-14 ligand, α 5 and α N helix

of respective G-protein. Also shown inset are residues W151 and F82 of the b-subunit (blue).

Extended Data Fig. 4 Structural comparison of G_q- and G_{i1}-coupled MRGPRX2 complex.

a, b, Structural comparison of the MRGPRX2-G_{i1}-cortistatin-14 complex (blue) with MRGPRX2-G_q-cortistatin-14 complex (cyan). Top view for the key interactions in sub-pocket 1 (**a**). Side view to show the overall conformational of cortistatin-14 (**b**). **c–e**, structural comparison of MRGPRX2-G_{i1}-(R)-ZINC-3573 complex with MRGPRX2-G_q-(R)-ZINC-3573 complex. G_{i1} and G_q are shown in green and salmon, respectively. G_{i1}-coupled MRGPRX2 and G_q-coupled MRGPRX2 are shown in blue and cyan, respectively. Side view of the whole complex (**c**), top view (**d**) and bottom view (**e**) of MRGPRX2. **f**, ICL3 of G_q is not clearly resolved in the G_q-coupled MRGPRX2 complex. **g**, Close-up view of the ICL3 in the G_{i1}-coupled MRGPRX2 structure with surrounding EM map at a threshold of 0.14. **h, i**, MRGPRX2 ICL3 mutations R214^{ICL3}A and L216^{ICL3}A impair cortistatin-14 (**h**) and (R)-ZINC-3573 (**i**) stimulated G_{i1} activation. Data represent mean ± s.e.m. of *n* = 3 biological replicates. **j, k**, BRET2 G_{i1} assays reveal that I135^{ICL2}A mutation of MRGPRX2 attenuates cortistatin-14 (**j**) and (R)-ZINC-3573 (**k**) stimulated G_{i1} activation. Data represent mean ± s.e.m. of *n* = 3 biological replicates. **l–m**, BRET2 G_q assays reveal that I135^{ICL2}A mutation of MRGPRX2 greatly reduced cortistatin-14 (**l**) and (R)-ZINC-3573 (**m**) stimulated G_q activation. Data represent mean ± s.e.m. of *n* = 3 biological replicates.

Extended Data Fig. 5 Non-conserved motifs in Mas-related GPCRs and the critical role of acidic residues E164^{4.60} and D184^{5.38} in MRGPRX2 activation.

a, Sequence alignment of the key residues in sodium site, DRY motif, PIF motif and CWxP motif, as well as residues involved in disulfide bond

formation in Mas-related GPCRs. Class A conserved residues are highlighted in green. **b**, cryoEM map of the TM4-TM5 disulfide bond in MRGPRX2-G_q-(R)-ZINC-3573 complex. **c, d**, Break of the TM4-TM5 disulfide bond by C168^{4.64}A and C180^{5.34}A mutations abolishes the cortistatin-14 stimulated G_q activation (**c**) and reduces the E_{max} of (R)-ZINC-3573 stimulated G_q activation by 60% (**d**). Data represent mean ± s.e.m. of $n = 3$ biological replicates. **e–g**, Compared with WT (**e**), E164^{4.60}A (**f**) and D184^{5.38}A (**g**) totally abolish the peptide stimulated G_q activation of MRGPRX2. Data represent mean ± s.e.m. of $n = 3$ biological replicates.

Extended Data Fig. 6 Unique structural features of MRGPRX2 and MRGPRX4.

a, MRGPRX2 and MRGPRX4 have a unique structural arrangement at the PIF motif compared with the G protein coupled active structures of 5-HT_{2A}R (PDB ID 6WHA), A_{2A}R (PDB ID 5G53) and β₂AR (PDB ID 3SN6). Residue 5.50 shifts away from the TM3-TM6 interface and does not engage L^{3.40} and F^{6.44} in MRGPRX2 and MRGPRX4. **b**, With G^{6.48}, TM6 of both MRGPRX2 and MRGPRX4 packs closer to TM3 compared with the G protein coupled active structures of 5-HT_{2A}R (PDB ID 6WHA), A_{2A}R (PDB ID 5G53) and β₂AR (PDB ID 3SN6), leading to an occluded canonical agonist binding pocket. **c**, (R)-ZINC-3573, cortistatin-14 and MS47134 bind to MRGPRX2 and MRGPRX4 at a position that is far away from residue 6.48, respectively. Cortistatin-14 is shown as cartoon. Small molecule compounds of receptors are shown as spheres.

Extended Data Fig. 7 Analog screening and functional characterization of MRGPRX2 antagonists.

a, b, Dose-response curves of initial 14 analogs of ‘1592 (**a**) and 8 analogs of C9 (**b**) in the presence of EC₈₀ concentration of (R)-ZINC-3573 using MRGPRX2 FLIPR Ca²⁺ assay. Data represent mean ± s.e.m. of $n = 3$ biological replicates. **c**, Dose-response curves of two potent MRGPRX2 antagonists C9 and C9-6 and an inactive compound C7 in the presence of

EC_{80} of each MRGPRX2 peptide using MRGPRX2 FLIPR Ca^{2+} assay. Data represent mean \pm s.e.m. of $n = 3$ biological replicates.

Extended Data Fig. 8 Functional characterization of optimized MRGPRX4 agonists.

a, Dose-response curves of Kir6.2/SUR1 current inhibition by indicated chemicals. Data represent mean \pm s.e.m. from $n=4$ biological replicates. **b**, **d**, **f**, **h**. Current-voltage relationships of whole-cell traces recorded in 150 mM KCl with the supplements of indicated chemicals of the labeled concentrations. **c**, **e**, **g**, Time courses showing the whole-cell-current responses to the indicated chemicals of the labeled concentrations. **i**, MRGPRX4 agonists X4-4 and MS47134 have a higher selectivity over Kir6.2/SUR1 channel compared with nateglinide. EC_{50} (nM) of each tested compound is shown. **j**, Screening of MS47134 across the GPCRome (at 320 receptors) using the PRESTO-Tango platform with 3 μ M MS47134. Red dashed line indicated threefold of basal levels. Data represent mean \pm s.e.m. of fold over basal for each receptor ($n=4$ technical replicates).

Extended Data Fig. 9 Structural comparison of G_q -coupled MRGPRX2 and MRGPRX4.

a–d, Structural comparison of the MRGPRX4- G_q -MS47134 complex with the MRGPRX2- G_q -(R)-ZINC-3573 complex. The receptor and G_q protein of MRGPRX4- G_q complex are colored by green and blue, respectively. The receptor and G_q protein of MRGPRX2- G_q complex are colored by cyan and salmon, respectively. Side view (**a**), Close-up view of α N-ICL2 interaction region (**b**), α 5 helix region (**c**), and the cytoplasmic side of receptors (**d**). **e**, The acidic residues E157^{4.60} and D177^{5.38} of MRGPRX4 are shielded by the inserted ECL2. Side chain of D177 is not resolved but modeled here for a better visual interpretation. **f**, Residues E164^{4.60} and D184^{5.38} of MRGPRX2 extend to the cationic agonists accessible pocket. **g**, Due to the variance in residue 2.39, Y243^{H5.23} of G_q adopts different side-chain conformations to interact with Y130^{ICL2} of MRGPRX4 and Y137^{ICL2} of

MRGPRX2. **h**, **i**, BRET2 G_q assays for Y130^{ICL2}A of MRGPRX4 (**h**) and Y137^{ICL2}A of MRGPRX2 (**i**). Data represent mean ± s.e.m. of $n = 3$ biological replicates.

Extended Data Table 1 Cryo-EM data collection, refinement and validation statistics

Supplementary information

Supplementary Information

See Supplementary Information page 1 for contents

Reporting Summary

Supplementary Tables

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Cao, C., Kang, H.J., Singh, I. *et al.* Structure, function and pharmacology of human itch GPCRs. *Nature* **600**, 170–175 (2021).
<https://doi.org/10.1038/s41586-021-04126-6>

- Received: 14 May 2021
- Accepted: 08 October 2021
- Published: 17 November 2021

- Issue Date: 02 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04126-6>

Share this article

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

[Get shareable link](#)

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

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04126-6>

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

- Matters Arising
- [Published: 01 December 2021](#)

Non-trivial role of internal climate feedback on interglacial temperature evolution

- [Xu Zhang](#) [ORCID: orcid.org/0000-0003-1833-9689^{1,2}](#) &
- [Fahu Chen](#) [ORCID: orcid.org/0000-0002-8874-1035^{1,2}](#)

[Nature](#) volume **600**, pages E1–E3 (2021)

- 819 Accesses
- 2 Altmetric
- [Metrics details](#)

Subjects

- [Cryospheric science](#)
- [Palaeoceanography](#)
- [Palaeoclimate](#)

[Matters Arising](#) to this article was published on 01 December 2021

The [Original Article](#) was published on 27 January 2021

arising from S. Bova et al. *Nature* <https://doi.org/10.1038/s41586-020-03155-x> (2021)

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: Evolution and drivers of Holocene and LIG SST.

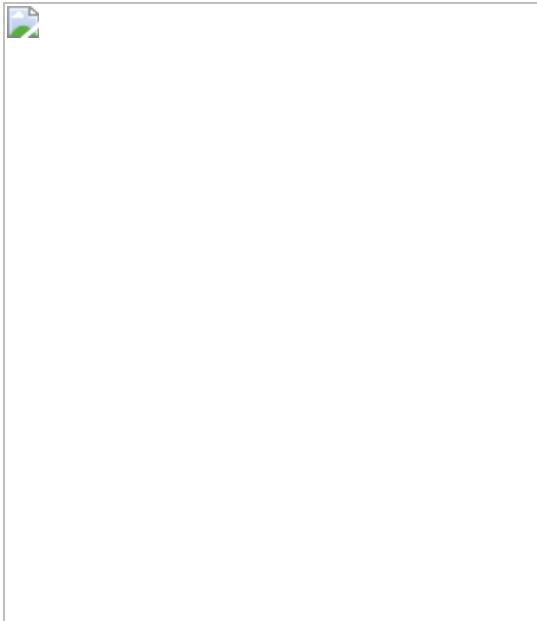
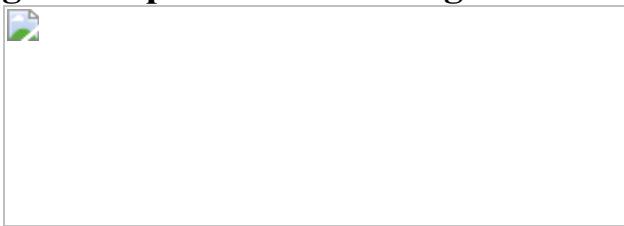


Fig. 2: Tropical mean cooling induced by sea-ice expansion.



Data availability

The datasets used for this study are available in the original papers.

References

1. 1.

Marcott, S. A., Shakun, J. D., Clark, P. U. & Mix, A. C. A reconstruction of regional and global temperature for the past 11,300 years. *Science* **339**, 1198–1201 (2013).

2. 2.

Liu, Z. et al. The Holocene temperature conundrum. *Proc. Natl Acad. Sci. USA* **111**, E3501–E3505 (2014).

3. 3.

Bova, S., Rosenthal, Y., Liu, Z., Godad, S. P. & Yan, M. Seasonal origin of the thermal maxima at the Holocene and the last interglacial. *Nature* **589**, 548–553 (2021).

4. 4.

Wolff, E. W. et al. Southern Ocean sea-ice extent, productivity and iron flux over the past eight glacial cycles. *Nature* **440**, 491–496 (2006).

5. 5.

Stein, R., Fahl, K., Gierz, P., Niessen, F. & Lohmann, G. Arctic Ocean sea ice cover during the penultimate glacial and the last interglacial. *Nat. Commun.* **8**, 373 (2017).

6. 6.

Holloway, M. D. et al. Antarctic last interglacial isotope peak in response to sea ice retreat not ice-sheet collapse. *Nat. Commun.* **7**, 12293 (2016).

7. 7.

Guarino, M. V. et al. Sea-ice-free Arctic during the Last Interglacial supports fast future loss. *Nat. Clim. Change* **10**, 928–932 (2020).

8. 8.

England, M. R., Polvani, L. M., Sun, L. & Deser, C. Tropical climate responses to projected Arctic and Antarctic sea-ice loss. *Nat. Geosci.* **13**, 275–281 (2020).

9. 9.

Bakker, P. et al. Temperature trends during the present and last interglacial periods - a multi-model-data comparison. *Quat. Sci. Rev.* **99**, 224–243 (2014).

10. 10.

Wu, Z., Yin, Q., Guo, Z. & Berger, A. Hemisphere differences in response of sea surface temperature and sea ice to precession and obliquity. *Glob. Planet. Change* **192**, 103223 (2020).

11. 11.

Lisiecki, L. E. & Raymo, M. E. A Pliocene-Pleistocene stack of 57 globally distributed benthic $\delta^{18}\text{O}$ records. *Paleoceanography* **20**, PA1003 (2005).

12. 12.

Clark, P. U. et al. Oceanic forcing of penultimate deglacial and last interglacial sea-level rise. *Nature* **577**, 660–664 (2020).

13. 13.

Varma, V., Prange, M. & Schulz, M. Transient simulations of the present and the last interglacial climate using the Community Climate System Model version 3: effects of orbital acceleration. *Geosci. Model Dev.* **9**, 3859–3873 (2016).

14. 14.

Barker, S. et al. Early interglacial legacy of deglacial climate instability. *Paleoceanogr. Paleoclimatol.* **34**, 1455–1475 (2019).

15. 15.

Capron, E. et al. Temporal and spatial structure of multi-millennial temperature changes at high latitudes during the Last Interglacial. *Quat. Sci. Rev.* **103**, 116–133 (2014).

16. 16.

The NCAR Command Language (NCL) v.6.62,
<https://doi.org/10.5065/D6WD3XH5> (UCAR/NCAR/CISL/TDD,

2019).

Acknowledgements

We acknowledge instructive comments by P. Bakker to improve this study. We also thank Y. Sun and Z. Fu for preparing the figures, and E. J. Gowan, X. Xiao and W. Xiao for helpful discussions. This study is supported by the Basic Science Center for Tibetan Plateau Earth System (BSCTPES, NSFC project No. 41988101) and the National Science Foundation of China (no. 42075047).

Author information

Affiliations

1. Alpine Paleoecology and Human Adaption Group (ALPHA), State Key Laboratory of Tibetan Plateau Earth System, Resources and Environment (TPESRE), Institute of Tibetan Plateau Research, Chinese Academy of Sciences, Beijing, China

Xu Zhang & Fahu Chen

2. Key Laboratory of Western China's Environmental Systems (Ministry of Education), College of Earth and Environmental Science, Lanzhou University, Lanzhou, China

Xu Zhang & Fahu Chen

Contributions

X.Z. initiated and developed this study by pointing out roles of sea ice change in global mean annual temperature. X.Z. and F.C. contributed equally to the writing of the Comment.

Corresponding author

Correspondence to [Xu Zhang](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

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

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Zhang, X., Chen, F. Non-trivial role of internal climate feedback on interglacial temperature evolution. *Nature* **600**, E1–E3 (2021).
<https://doi.org/10.1038/s41586-021-03930-4>

- Received: 02 March 2021
- Accepted: 18 August 2021
- Published: 01 December 2021
- Issue Date: 02 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-03930-4>

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/s41586-021-03930-4>

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

- Matters Arising
- [Published: 01 December 2021](#)

Reply to: Non-trivial role of internal climate feedback on interglacial temperature evolution

- [Samantha Bova](#) ORCID: [orcid.org/0000-0002-5064-8775^{1 nAff7}](https://orcid.org/0000-0002-5064-8775),
- [Yair Rosenthal](#) ORCID: [orcid.org/0000-0002-7546-6011^{1,2}](https://orcid.org/0000-0002-7546-6011),
- [Zhengyu Liu](#) ORCID: [orcid.org/0000-0003-4554-2666³](https://orcid.org/0000-0003-4554-2666),
- [Mi Yan](#) ORCID: [orcid.org/0000-0003-4730-3781^{4,5}](https://orcid.org/0000-0003-4730-3781),
- [Anthony J. Broccoli](#) ORCID: [orcid.org/0000-0003-2619-1434⁶](https://orcid.org/0000-0003-2619-1434),
- [Shital P. Godad](#) ORCID: [orcid.org/0000-0002-4182-6234^{1 nAff8}](https://orcid.org/0000-0002-4182-6234) &
- [Cheng Zeng](#) ORCID: [orcid.org/0000-0002-0771-970X^{4,5}](https://orcid.org/0000-0002-0771-970X)

[Nature](#) volume 600, pages E4–E6 (2021)

- 340 Accesses
- 1 Altmetric
- [Metrics details](#)

Subjects

- [Climate and Earth system modelling](#)
- [Palaeoceanography](#)
- [Palaeoclimate](#)

The [Original Article](#) was published on 01 December 2021

replying to X. Zhang and F. Chen *Nature* <https://doi.org/10.1038/s41586-021-03930-4> (2021)

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](#)

References

1. 1.

Zhang, X. & Chen, F. Non-trivial role of internal climate feedback on interglacial temperature evolution. *Nature* <https://doi.org/10.1038/s41586-021-03930-4> (2021).

2. 2.

Bova, S., Rosenthal, Y., Liu, Z., Godad, S. P. & Yan, M. Seasonal origin of the thermal maxima at the Holocene and the last interglacial. *Nature* **589**, 548–553 (2021).

3. 3.

Kaufman, D. et al. Holocene global mean surface temperature, a multi-method reconstruction approach. *Sci. Data* **7**, 201 (2020).

4. 4.

Marcott, S. A., Shakun, J. D., Clark, P. U. & Mix, A. C. A reconstruction of regional and global temperature for the past 11,300 years. *Science* **339**, 1198–1201 (2013).

5. 5.

Liu, Z. et al. The Holocene temperature conundrum. *Proc. Natl Acad. Sci. USA* **111**, E3501–E3505 (2014).

6. 6.

Brierley, C. M. et al. Large-scale features and evaluation of the PMIP4-CMIP6 *midHolocene* simulations. *Clim. Past* <https://doi.org/10.5194/cp-2019-168> (2020).

7. 7.

Leduc, G., Schneider, R., Kim, J.-H. & Lohmann, G. Holocene and Eemian sea surface temperature trends as revealed by alkenone and Mg/Ca paleothermometry. *Quat. Sci. Rev.* **29**, 989–1004 (2010).

8. 8.

Timmermann, A., Sachs, J. & Timm, O. E. Assessing divergent SST behavior during the last 21 ka derived from alkenones and *G. ruber*-Mg/Ca in the equatorial Pacific. *Paleoceanogr. Paleoclimatol.* **29**, 680–696 (2014).

9. 9.

Mohtadi, M. et al. Low-latitude control on seasonal and interannual changes in planktonic foraminiferal flux and shell geochemistry off south Java: a sediment trap study. *Paleoceanogr. Paleoclimatol.* <https://doi.org/10.1029/2008PA001636> (2009).

10. 10.

Lin, H.-L., Wang, W.-C. & Hung, G.-W. Seasonal variation of planktonic foraminiferal isotopic composition from sediment traps in the South China Sea. *Mar. Micropaleontol.* **53**, 447–460 (2004).

11. 11.

Sautter, L. R. & Thunell, R. C. Seasonal variability in the $\delta^{18}\text{O}$ and $\delta^{13}\text{C}$ of planktonic foraminifera from an upwelling environment: sediment trap results from the San Pedro Basin, Southern California Bight. *Paleoceanogr. Paleoclimatol.* **6**, 307–334 (1991).

12. 12.

Thunell, R. C. & Reynolds, L. A. Sedimentation of planktonic foraminifera: seasonal changes in species flux in the Panama Basin. *Micropaleontology* **30**, 243–262 (1984).

13. 13.

Sawada, K., Handa, N. & Nakatsuka, T. Production and transport of long-chain alkenones and alkyl alkenoates in a sea water column in the

northwestern Pacific off central Japan. *Mar. Chem.* **59**, 219–234 (1998).

14. 14.

Sikes, E. L., O’Leary, T., Nodder, S. D. & Volkman, J. K. Alkenone temperature records and biomarker flux at the subtropical front on the chatham rise, SW Pacific Ocean. *Deep Sea Res. I* **52**, 721–748 (2005).

15. 15.

Kienast, M. et al. Alkenone unsaturation in surface sediments from the eastern equatorial Pacific: implications for SST reconstructions. *Paleoceanogr. Paleoclimatol.* <https://doi.org/10.1029/2011PA002254> (2012).

16. 16.

Hertzberg, J. E. & Schmidt, M. W. Refining *Globigerinoides ruber* Mg/Ca paleothermometry in the Atlantic Ocean. *Earth Planet. Sci. Lett.* **383**, 123–133 (2013).

17. 17.

Wolff, E. W. et al. Southern Ocean sea-ice extent, productivity and iron flux over the past eight glacial cycles. *Nature* **440**, 491–496 (2006).

18. 18.

Malmierca-Vallet, I. et al. Simulating the Last Interglacial Greenland stable water isotope peak: the role of Arctic sea ice changes. *Quat. Sci. Rev.* **198**, 1–14 (2018).

19. 19.

Bader, J. et al. Global temperature modes shed light on the Holocene temperature conundrum. *Nat. Commun.* **11**, 4726 (2020).

20. 20.

Lisiecki, L. E. & Raymo, M. E. A Pliocene-Pleistocene stack of 57 globally distributed benthic $\delta^{18}\text{O}$ records. *Paleoceanogr. Paleoclimatol.* <https://doi.org/10.1029/2004pa001071> (2005).

21. 21.

Shackleton, S. et al. Global ocean heat content in the Last Interglacial. *Nat. Geosci.* **13**, 77–81 (2020).

22. 22.

Bereiter, B., Shackleton, S., Baggenstos, D., Kawamura, K. & Severinghaus, J. Mean global ocean temperatures during the last glacial transition. *Nature* **553**, 39–44 (2018).

23. 23.

Rosenthal, Y., Linsley, B. K. & Oppo, D. W. Pacific Ocean heat content during the past 10,000 years. *Science* **342**, 617–621 (2013).

Acknowledgements

Funding for this research was provided by the NSF grants OCE-1834208 and OCE-1810681, the NSF-sponsored US Science Support Program for IODP, the Institute of Earth, Ocean, and Atmospheric Sciences at Rutgers University, the Chinese NSF (41630527), the School of Geography, Nanjing Normal University, and the USIEF-Fulbright Program.

Author information

Author notes

1. Samantha Bova

Present address: Department of Geological Sciences, San Diego State University, San Diego, USA

2. Shital P. Godad

Present address: Department of Geosciences, National Taiwan University, Taipei, Taiwan

Affiliations

1. Department of Marine and Coastal Sciences, Rutgers, State University of New Jersey, New Brunswick, NJ, USA

Samantha Bova, Yair Rosenthal & Shital P. Godad

2. Department of Earth and Planetary Sciences, Rutgers, State University of New Jersey, New Brunswick, NJ, USA

Yair Rosenthal

3. Atmospheric Science Program, Department of Geography, The Ohio State University, Columbus, OH, USA

Zhengyu Liu

4. School of Geography, Nanjing Normal University, Nanjing, China

Mi Yan & Cheng Zeng

5. Open Studio for Ocean-Climate-Isotope Modeling, Pilot National Laboratory for Marine Science and Technology, Qingdao, China

Mi Yan & Cheng Zeng

6. Department of Environmental Sciences, Rutgers, State University of New Jersey, New Brunswick, NJ, USA

Anthony J. Broccoli

Contributions

All authors contributed to conception of the presented ideas. S.B. wrote the first manuscript draft. All authors provided review and editing. Two authors not on the original paper were added to the author list. C.Z. provided additional analysis of model results. A.J.B. provided critical feedback and discussion.

Corresponding author

Correspondence to [Samantha Bova](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

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

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Bova, S., Rosenthal, Y., Liu, Z. *et al.* Reply to: Non-trivial role of internal climate feedback on interglacial temperature evolution. *Nature* **600**, E4–E6 (2021). <https://doi.org/10.1038/s41586-021-03931-3>

- Published: 01 December 2021

- Issue Date: 02 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-03931-3>

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

[Seasonal origin of the thermal maxima at the Holocene and the last interglacial](#)

- Samantha Bova
- Yair Rosenthal
- Mi Yan

Article 27 Jan 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03931-3>

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

Amendments & Corrections

- [**Author Correction: Hepatic stellate cells suppress NK cell-sustained breast cancer dormancy**](#) [11 November 2021]
Author Correction •
- [**Author Correction: Isolation of SARS-CoV-2-related coronavirus from Malayan pangolins**](#) [11 November 2021]
Author Correction •
- [**Author Correction: RecA finds homologous DNA by reduced dimensionality search**](#) [12 November 2021]
Author Correction •

- Author Correction
- [Published: 11 November 2021](#)

Author Correction: Hepatic stellate cells suppress NK cell-sustained breast cancer dormancy

- [Ana Luísa Correia](#) ORCID: [orcid.org/0000-0003-2414-0131](#)^{1,2,3},
- [Joao C. Guimaraes](#) ORCID: [orcid.org/0000-0002-1664-472X](#)⁴ nAff11,
- [Priska Auf der Maur](#) ORCID: [orcid.org/0000-0002-9059-8979](#)^{1,3},
- [Duvini De Silva](#)^{1,2} nAff11,
- [Marcel P. Trefny](#) ORCID: [orcid.org/0000-0001-6755-7899](#)¹,
- [Ryoko Okamoto](#)^{1,2,3} nAff12,
- [Sandro Bruno](#)² nAff13,
- [Alexander Schmidt](#) ORCID: [orcid.org/0000-0002-3149-2381](#)⁵,
- [Kirsten Mertz](#) ORCID: [orcid.org/0000-0002-3074-6925](#)⁶,
- [Katrin Volkmann](#)^{1,3},
- [Luigi Terracciano](#)⁷,
- [Alfred Zippelius](#) ORCID: [orcid.org/0000-0003-1933-8178](#)^{1,8},
- [Marcus Vetter](#)^{8,9,10},
- [Christian Kurzeder](#)^{9,10},
- [Walter Paul Weber](#)^{3,10} &
- [Mohamed Bentires-Alj](#) ORCID: [orcid.org/0000-0001-6344-1127](#)^{1,2,3}

[Nature](#) volume 600, page E7 (2021)

- 722 Accesses
- 4 Altmetric

- [Metrics details](#)

Subjects

- [Breast cancer](#)
- [Cancer microenvironment](#)

The [Original Article](#) was published on 02 June 2021

[Download PDF](#)

Correction to: *Nature* <https://doi.org/10.1038/s41586-021-03614-z>

Published online 2 June 2021

In the version of this Article initially published, the Acknowledgements statement for author Joao C. Guimaraes was incomplete. The statement has been expanded to now read: “Part of J.C.G.’s work was performed while he was in the laboratory of Mihaela Zavolan at the University of Basel supported by a SystemsX.ch Transitional Postdoctoral Fellowship (grant 51FSP0_157344) to J.C.G., a Novartis University of Basel Excellence Scholarship for Life Sciences to J.C.G., and the Swiss National Science Foundation grant 51NF40_141735 (National Center for Competence in Research ‘RNA & Disease’; co-PI Mihaela Zavolan).” The changes have been made to the online version of the Article.

Author information

Author notes

1. Joao C. Guimaraes & Duvini De Silva

Present address: F. Hoffmann-La Roche AG, Basel, Switzerland

2. Ryoko Okamoto

Present address: Department of Biosystems Science and Engineering,
ETH Zürich, Basel, Switzerland

3. Sandro Bruno

Present address: Novartis Institutes for BioMedical Research, Basel,
Switzerland

Affiliations

1. Department of Biomedicine, University of Basel, Basel, Switzerland

Ana Luísa Correia, Priska Auf der Maur, Duvini De Silva, Marcel P. Trefny, Ryoko Okamoto, Katrin Volkmann, Alfred Zippelius & Mohamed Bentires-Alj

2. Friedrich Miescher Institute for Biomedical Research, Basel,
Switzerland

Ana Luísa Correia, Duvini De Silva, Ryoko Okamoto, Sandro Bruno & Mohamed Bentires-Alj

3. Department of Surgery, University Hospital Basel, Basel, Switzerland

Ana Luísa Correia, Priska Auf der Maur, Ryoko Okamoto, Katrin Volkmann, Walter Paul Weber & Mohamed Bentires-Alj

4. Computational and Systems Biology, Biozentrum, University of Basel,
Basel, Switzerland

Joao C. Guimaraes

5. Proteomics Core Facility, Biozentrum, University of Basel, Basel,
Switzerland

Alexander Schmidt

6. Institute of Pathology Liestal, Cantonal Hospital Basel-Land, Liestal,
Switzerland

Kirsten Mertz

7. Institute of Pathology, University Hospital Basel, Basel, Switzerland

Luigi Terracciano

8. Department of Medical Oncology, University Hospital Basel, Basel, Switzerland

Alfred Zippelius & Marcus Vetter

9. Gynecologic Cancer Center, University Hospital Basel, Basel, Switzerland

Marcus Vetter & Christian Kurzeder

10. Breast Center, University of Basel and University Hospital Basel, Basel, Switzerland

Marcus Vetter, Christian Kurzeder & Walter Paul Weber

Corresponding authors

Correspondence to [Ana Luísa Correia](#) or [Mohamed Bentires-Alj](#).

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Correia, A.L., Guimaraes, J.C., Auf der Maur, P. *et al.* Author Correction: Hepatic stellate cells suppress NK cell-sustained breast cancer dormancy. *Nature* **600**, E7 (2021). <https://doi.org/10.1038/s41586-021-04104-y>

- Published: 11 November 2021
- Issue Date: 02 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04104-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

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-04104-y>.

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

- Author Correction
- [Published: 11 November 2021](#)

Author Correction: Isolation of SARS-CoV-2-related coronavirus from Malayan pangolins

- [Kangpeng Xiao^{1,2}](#), [na1](#),
- [Junqiong Zhai³](#), [na1](#),
- [Yaoyu Feng^{1,2}](#),
- [Niu Zhou](#) [ORCID: orcid.org/0000-0002-0884-5943³](#),
- [Xu Zhang^{1,2}](#),
- [Jie-Jian Zou⁴](#),
- [Na Li^{1,2}](#),
- [Yaqiong Guo^{1,2}](#),
- [Xiaobing Li¹](#),
- [Xuejuan Shen¹](#),
- [Zhipeng Zhang¹](#),
- [Fanfan Shu^{1,2}](#),
- [Wanyi Huang^{1,2}](#),
- [Yu Li⁵](#),
- [Ziding Zhang⁵](#),
- [Rui-Ai Chen^{1,6}](#),
- [Ya-Jiang Wu³](#),
- [Shi-Ming Peng³](#),
- [Mian Huang³](#),
- [Wei-Jun Xie³](#),
- [Qin-Hui Cai³](#),
- [Fang-Hui Hou⁴](#),
- [Wu Chen](#) [ORCID: orcid.org/0000-0002-4033-254X³](#),
- [Lihua Xiao](#) [ORCID: orcid.org/0000-0001-8532-2727^{1,2}](#) &
- [Yongyi Shen](#) [ORCID: orcid.org/0000-0001-7660-5522^{1,2}](#)

- 3615 Accesses
- 112 Altmetric
- [Metrics details](#)

Subjects

- [SARS-CoV-2](#)
- [SARS virus](#)
- [Viral infection](#)

The [Original Article](#) was published on 07 May 2020

[Download PDF](#)

Correction to: *Nature* <https://doi.org/10.1038/s41586-020-2313-x> Published online 7 May 2020

In this Article, data in Extended Data Table 3 and Extended Data Fig. 4 were mislabelled and attributed incorrectly. The Pangolin-CoV genome reported was built using the metagenomic dataset described previously by Liu et al. in *Viruses*¹ (ref. 15 in our *Nature* paper) and targeted PCR. The *Viruses* article¹ and its metagenomic data ([PRJNA573298](#)) were cited in our Article. All 21 animals from the March 2019 seizure of smuggled pangolins were used for our Article, and 11 of these were used in the *Viruses* study.

Wu Chen, one of the corresponding authors of the *Nature* Article as well as a coauthor of the study by Liu et al., provided all the samples and the associated data for both studies. Dr Chen, who archives data from clinical specimens by animal ID, provided all samples and data for the *Nature* Article with matched IDs, including some metagenomic data. Because the numbering system for the metagenomic data was different from that used for the *Viruses* paper, the informaticians of our team believed that metagenomic data from M2, M3, M4 and M8 provided by Dr Chen were new. Instead, the samples labelled M2, M3, M4 and M8 in the Extended Data Table 3 of our Article correspond to samples labelled lung07, lung02, lung08 and lung11 in the *Viruses* paper. The lack of face-to-face meetings imposed by various restrictions due to COVID-19 among the four research groups involved with the study led to a delay in finding out the problem. Extended Data Table 3 of the original Article has been corrected to clarify the relationship between these samples.

The original Extended Data Table 3 listed nine animals, including the four samples (M2, M3, M4 and M8) that overlap with data from the *Viruses* paper. The corrected table lists 12 pangolins. Data from two samples (P59 and P60) submitted to NCBI during the initial submission to *Nature* were inadvertently omitted from Extended Data Table 3, and data from A22 were submitted to NCBI later. In revising our *Nature* Article, we became aware of both the overlap of metagenomic data with the *Viruses* paper and the fact that Jin-Ping Chen (corresponding author of the *Viruses* paper) was using the *Viruses* dataset in the preparation of another manuscript². We therefore added new metagenomic data from pangolin A22 that were generated in March 2020 and have almost full coverage of the Pangolin-CoV genome. We neglected to upload these data to NCBI. During the preparation of another manuscript on the pangolin coronavirus³, which used some of the metagenomic data in the *Nature* Article, we realized that data from A22 had not been submitted to NCBI. We therefore updated our dataset on 19 June 2020 and added the A22 data to BioProject [PRJNA607174](#) (released on 22 June 2020). The labels in BioProject have been updated using the animal IDs to avoid confusion.

In addition, in Extended Data Table 3, the Chinese pangolin M10 was incorrectly labelled as a Malayan pangolin and the numbers of total reads for samples M1 and M6 were incorrect. As described in the Methods, these samples were sequenced with the paired-end approach (two reads in each pair). The original table showed values for the number of paired reads (107,267,359 and 232,433,120, respectively, for M1 and M6) instead of the number of total reads: 214,534,718 and 464,866,240. The original published Extended Data Table 3 is shown as [Supplementary Information](#) to this Amendment, for transparency to readers.

The first two sentences in the fourth paragraph misstated the numbers of samples: ‘Illumina RNA sequencing was used to identify viruses in the lung from nine pangolins. Mapping sequence data to the reference SARS-CoV-2 WHCV genome identified coronavirus sequence reads in seven samples (Extended Data Table 3).’ Instead these sentences should read ‘Illumina RNA sequencing was used to identify viruses in the lung from **12 pangolins (including four that were reported previously¹)**. Mapping sequence data to the reference SARS-CoV-2 WHCV genome identified coronavirus sequence reads in nine samples (Extended Data Table 3)’.

The following two sentences in that paragraph said ‘For one sample, higher genome coverage was obtained by remapping the total reads to the reference genome (Extended Data Fig. 4). We obtained the completed coronavirus genome (29,825bp)—which we designated pangolin-CoV—using the assembled contigs, short sequence reads and targeted PCR analysis’. This completed genome used the metagenomic dataset of lung08 (labelled M4 in the *Nature* Article) published by Liu et al.¹ in the initial metagenome assembly. We used PCR in filling the numerous sequence gaps and

ambiguities in the metagenome assembly. Altogether, we obtained sequences through PCR to cover ~90% of the genome. The primers are listed in Supplementary Table [1](#) of this Amendment.

Extended Data Fig. 4 of the original Article was based on a composite of data from lung08 and M1. The sequence reads from metagenomic sequencing were mapped to the SARS-CoV-2 WHCV genome initially in the identification of viral sequences in the samples and in primer design. They were subsequently mapped to the assembled Pangolin-CoV genome for confirmation of the detection of Pangolin-CoV in these samples. In both analyses, mapped reads were further extracted from the two best metagenomic datasets (lung08 and M1) at the time for additional mapping of pooled reads to guide the primer design and to check the accuracy of the genome sequence. The intention was to visually show the presence of Pangolin-CoV sequence reads in these samples, using the data from the best sample lung08 (M4 in the *Nature* Article) as an example in Extended Data Fig. 4. However, the mapping plot from pooled reads of lung08 and M1 was mistakenly used for the figure. The corrected Extended Data Fig. 4 uses data from lung08 only, and the legend clarifies that lung08 was M4 in the *Nature* Article. The original, published Extended Data Fig. 4 is shown as [Supplementary Information](#) of this Amendment, for transparency to readers.

The raw sequence data (including the trace files) generated by PCR for the assembly of the Pangolin-CoV genome have been deposited to the SRA database of NCBI (accession no. SRX9503273), the six full sequences of the S gene generated by PCR from pangolin samples in the study to GenBank (accession nos. MT799521–MT799526), and we have added a new table on primers used in PCR ([Supplementary Table 1](#) of this Amendment).

Fig. 1

 **figure1**

This figure shows the incorrect, as-published version and the corrected version of Extended Data Fig. 4 of the original Article.

Fig. 2

 **figure2**

This figure shows the incorrect, as-published version and the corrected version of Extended Data Table 3 of the original Article.

We thank Yujia Alina Chan and Shing Hei Zhan for bringing the errors to our attention⁴. The original Article has been corrected online.

References

1. 1.

Liu, P., Chen, W. & Chen, J. P. Viral metagenomics revealed sendai virus and coronavirus infection of Malayan pangolins (*Manis javanica*). *Viruses* **11**, 979 (2019).

2. 2.

Liu, P. et al. Are pangolins the intermediate host of the 2019 novel coronavirus (SARS-CoV-2)? *PLoS Pathog.* **16**, e1008421 (2020).

3. 3.

Li, X. et al. Pathogenicity, tissue tropism and potential vertical transmission of SARSr-CoV-2 in Malayan pangolins. Preprint at *bioRxiv* <https://doi.org/10.1101/2020.06.22.164442> (2020).

4. 4.

Chan, Y. A. & Zhan, S. H. Single source of pangolin CoVs with a near identical Spike RBD to SARS-CoV-2. Preprint at *bioRxiv* <https://doi.org/10.1101/2020.07.07.184374> (2020).

Author information

Author notes

1. These authors contributed equally: Kangpeng Xiao, Junqiong Zhai

Affiliations

1. Center for Emerging and Zoonotic Diseases, College of Veterinary Medicine, South China Agricultural University, Guangzhou, China

Kangpeng Xiao, Yaoyu Feng, Xu Zhang, Na Li, Yaqiong Guo, Xiaobing Li, Xuejuan Shen, Zhipeng Zhang, Fanfan Shu, Wanyi Huang, Rui-Ai Chen, Lihua Xiao & Yongyi Shen

2. Guangdong Laboratory for Lingnan Modern Agriculture, Guangzhou, China

Kangpeng Xiao, Yaoyu Feng, Xu Zhang, Na Li, Yaqiong Guo, Fanfan Shu, Wanyi Huang, Lihua Xiao & Yongyi Shen

3. Guangzhou Zoo & Guangzhou Wildlife Research Center, Guangzhou, China

Junqiong Zhai, Niu Zhou, Ya-Jiang Wu, Shi-Ming Peng, Mian Huang, Wei-Jun Xie, Qin-Hui Cai & Wu Chen

4. Guangdong Provincial Wildlife Rescue Center, Guangzhou, China

Jie-Jian Zou & Fang-Hui Hou

5. State Key Laboratory of Agrobiotechnology, College of Biological Sciences, China Agricultural University, Beijing, China

Yu Li & Ziding Zhang

6. Zhaoqing Branch Center of Guangdong Laboratory for Lingnan Modern Agricultural Science and Technology, Zhaoqing, China

Rui-Ai Chen

Corresponding authors

Correspondence to [Wu Chen](#), [Lihua Xiao](#) or [Yongyi Shen](#).

Additional information

Supplementary information is available in the online version of this Amendment.

Supplementary information

[Supplementary Table 1](#)

PCR primers used to amplify fragments of the Pangolin-CoV genome.

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Xiao, K., Zhai, J., Feng, Y. *et al.* Author Correction: Isolation of SARS-CoV-2-related coronavirus from Malayan pangolins. *Nature* **600**, E8–E10 (2021).
<https://doi.org/10.1038/s41586-021-03838-z>

- Published: 11 November 2021
- Issue Date: 02 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-03838-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/s41586-021-03838-z>

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

- Author Correction
- Open Access
- [Published: 12 November 2021](#)

Author Correction: RecA finds homologous DNA by reduced dimensionality search

- [Jakub Wiktor¹](#) na1,
- [Arvid H. Gynnå](#) [ORCID: orcid.org/0000-0001-8087-7715](#)¹ na1,
- [Prune Leroy¹](#),
- [Jimmy Larsson¹](#),
- [Giovanna Coceano²](#),
- [Ilaria Testa](#) [ORCID: orcid.org/0000-0003-4005-4997](#)² &
- [Johan Elf](#) [ORCID: orcid.org/0000-0001-5522-1810](#)¹

[Nature](#) volume 600, page E11 (2021)

- 778 Accesses
- 2 Altmetric
- [Metrics details](#)

Subjects

- [DNA recombination](#)
- [Homologous recombination](#)
- [Super-resolution microscopy](#)

The [Original Article](#) was published on 01 September 2021

[Download PDF](#)

Correction to: *Nature* <https://doi.org/10.1038/s41586-021-03877-6>

Published online 1 September 2021

In the version of this Article initially published, a typo appeared in the Acknowledgements section for researcher name Zikrin. The sentence reading, in part, “We thank … S. Zirkin and P. Karempudi for help with the image analysis,” has now been corrected to read “We thank … S. Zikrin and P. Karempudi for help with the image analysis.” The original Article has been corrected online.

Author information

Author notes

1. These authors contributed equally: Jakub Wiktor, Arvid H. Gynnå

Affiliations

1. Department of Cell and Molecular Biology, Science for Life Laboratory, Uppsala University, Uppsala, Sweden

Jakub Wiktor, Arvid H. Gynnå, Prune Leroy, Jimmy Larsson & Johan Elf

2. Department of Applied Physics, Science for Life Laboratory, KTH Royal Institute of Technology, Stockholm, Sweden

Giovanna Coceano & Ilaria Testa

Corresponding author

Correspondence to [Johan Elf](#).

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

Wiktor, J., Gynnå, A.H., Leroy, P. *et al.* Author Correction: RecA finds homologous DNA by reduced dimensionality search. *Nature* **600**, E11 (2021). <https://doi.org/10.1038/s41586-021-04154-2>

- Published: 12 November 2021
- Issue Date: 02 December 2021
- DOI: <https://doi.org/10.1038/s41586-021-04154-2>

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/s41586-021-04154-2>

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

Nature Outline