

nature

ABOVE & BELOW

How fine-root traits fit into plant form and function



Trials and tribulations
The struggle of two scientists launching their own labs

Data-driven policy
Build predictive models to tackle growing crisis in mental health

The weather nowcast
Deep generative model offers near real-time predictions of rainfall

Nature.2021.10.02

[Sat, 02 Oct 2021]

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

This Week

- **[Concrete needs to lose its colossal carbon footprint](#)** [28 September 2021]
Editorial • Concrete will be crucial for much-needed climate-resilient construction. But the cement industry must set out its plan for decarbonization.
- **[What a personal saga reveals about scientists' lives — and about science itself](#)** [29 September 2021]
Editorial • Two scientists allowed Nature to chronicle their lives for three years. Their story speaks to the epic professional and personal struggles involved in establishing a career in research.
- **[The incoming Afghan government must allow immunizations](#)** [23 September 2021]
World View • The world must work with Kabul's new rulers to get polio and other diseases under control.
- **[An iron grip could help pee to produce electricity](#)** [24 September 2021]
Research Highlight • Catalyst based on nickel and iron allows electrical current to be harvested from the breakdown of urea.
- **[Life-sized camels in rock are a Stone Age masterpiece](#)** [21 September 2021]
Research Highlight • Saudi Arabia's Camel Site is a testimony to the dexterity of prehistoric sculptors.
- **[Tiny particles could make a powerful COVID vaccine](#)** [23 September 2021]
Research Highlight • Nanoparticles that bristle with dozens of copies of a SARS-CoV-2 protein segment could protect against a wide variety of coronaviruses.
- **[Bright lights at night scramble the sweet song of crickets](#)** [21 September 2021]
Research Highlight • Exposure to stronger artificial light led to greater disruption of the insects' night music.
- **[Thirty lasers act as one to shine a powerful light](#)** [24 September 2021]
Research Highlight • Cluster of tiny tube-shaped devices creates an intense, focused light source.

- **The oil-field ‘fugitive’ that can slip into the water supply**

[26 September 2021]

Research Highlight • Methane gas that steals away from leaky fossil-fuel wells can hide underground and taint groundwater.

- **A mega-penguin stood tall on prodigious limbs** [21

September 2021]

Research Highlight • Remains discovered by schoolchildren on a fossil-hunting trip belonged to an extinct giant bird.

- EDITORIAL
- 28 September 2021

Concrete needs to lose its colossal carbon footprint

Concrete will be crucial for much-needed climate-resilient construction. But the cement industry must set out its plan for decarbonization.



Cement manufacturing (such as that at this plant in Russia) accounts for 8% of the world's carbon dioxide emissions. Credit: Getty

Wet concrete has been poured into buildings, roads, bridges and more for centuries. Structures using concrete have survived wars and natural disasters, outlasting many of the civilizations that built them¹. Alongside its strength and resilience, concrete is also a staple of building because it is relatively cheap and simple to make. Worldwide, 30 billion tonnes of concrete is used each year. On a per capita basis, that is 3 times as much as

40 years ago — and the demand for concrete is growing more steeply than that for steel or wood².

Versatile and long-lasting, concrete buildings and structures are in many ways ideal for climate-resilient construction. But concrete has a colossal carbon footprint — at least 8% of global emissions caused by humans come from the cement industry alone³. We must decarbonize its production.

Concrete is made by adding sand and gravel to cement, whisking the mixture with water and pouring it into moulds before it dries. Making the cement is the most carbon-intensive part: it involves using fossil fuels to heat a mixture of limestone and clay to more than 1,400 °C in a kiln. Also, when limestone (calcium carbonate) is heated with clays, roughly 600 kilograms of carbon dioxide is released for every tonne of cement produced (see go.nature.com/3exhg82).



Lithium-ion batteries need to be greener and more ethical

There are alternatives to cement, but they're in the early stages of development, and cement itself remains highly profitable — two disincentives for companies to change.

Alternatives include the leftover compounds from iron and steel production, known as slag, and heaps of unused fly ash, a residual material from coal

plants. Researchers are also experimenting with reducing the temperatures needed in the cement-making process — thereby decreasing the energy used.

In carbon-accounting terms, such replacements and procedural changes will reduce cement's environmental impact, and that of concrete, too. But they still involve carbon emissions. Coal is being phased out, so fly ash isn't a long-term solution. And alternatives have yet to be certified for use in building; for this to happen, long-term studies on their performance are needed.

Cement will be around for the foreseeable future, so cement production itself needs to be decarbonized, which could happen in a number of ways. For example, low-carbon fuels — such as hydrogen or biomass — could be substituted for fossil fuels in heating the limestone and clay. And scientists are examining whether electricity — instead of combustion — could be used for the heating.

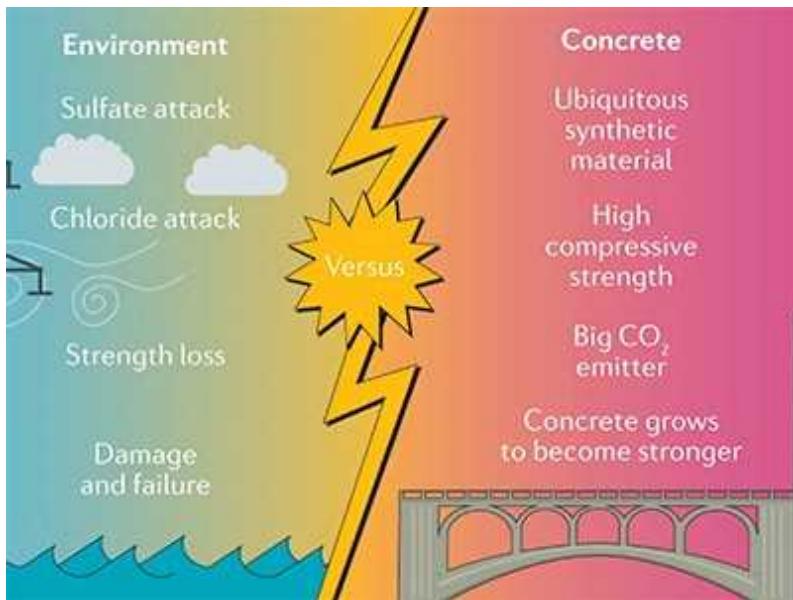
Carbon capture could be part of the cement industry's transition process⁴. In Sweden, for example, a company announced in July that it wants to capture 1.8 million tonnes of CO₂ from a cement plant and bury it in the North Sea. Another possibility is to pump the captured CO₂ into concrete itself, locking it up forever — which might also improve the properties of the resulting material. The injected CO₂ reacts with calcium ions in the cement, producing more calcium carbonate, and potentially making the concrete able to withstand larger loads.

Concrete options

Technological changes can be accelerated through regulation and legislation. A huge proportion of concrete is used in public building projects. In North America alone, public agencies buy as much as one-third of concrete manufactured annually. That means they have leverage in the low-carbon transition: they could work with researchers and manufacturers to reshape the concrete industry.

In New York and New Jersey, a bill is making its way through state legislatures that, if passed, will mandate that state agencies and departments

prioritize cement that has a lower carbon footprint.



An ancient battle between environment and concrete

Other states are introducing legislation that requires construction proposals to declare the environmental impact of cement mixes. Some regions, such as Honolulu, Hawaii, have added a requirement that city construction projects must consider using concrete that stores CO₂.

In Europe, the European Union's Waste Framework Directive requires 70% of construction waste to be reused. Another idea, known as materials passports, could also help. When buildings are demolished, the waste concrete is smashed up and discarded, or sold for low-grade use such as backfilling in road construction. But a passport would ensure that concrete is recorded 'at birth' and then tracked throughout its life cycle — making it accessible for more kinds of reuse.

Finally, the cement industry needs to publish better emissions data so that progress can be tracked. In a study published last month, researchers from Columbia University in New York City report that some of China's cement-making companies have an ambition for emissions to peak in 2023 (see go.nature.com/39z1sdd). But only one of the ten companies surveyed is reporting emissions data. China is the world's largest cement producer (55% of global capacity), and cement accounts for 15% of the country's carbon

emissions. Without data, it will be impossible to know whether national targets are being met.

Next month, a cement-industry campaign called Concrete Action for Climate will announce its road map for carbon neutrality by 2050. This is overdue, but the road map must also explain interim steps, how companies intend to achieve neutrality and how progress will be measured.

Parts of the ancient world were made with concrete, and the material was used to build much of the modern world, too. Researchers and governments must work with the cement industry to slash its carbon footprint, driving the climate-resilient construction the world now sorely needs.

Nature **597**, 593–594 (2021)

doi: <https://doi.org/10.1038/d41586-021-02612-5>

References

1. 1.

Jackson, M. D. *et al.* *Proc. Natl Acad. Sci. USA* **111**, 18484–18489 (2014).

2. 2.

Monteiro, P., Miller, S. & Horvath, A. *Nature Mater.* **16**, 698–699 (2017).

3. 3.

Ellis, L. D., Badel, A. F., Chiang, M. L., Park, R. J.-Y. & Chiang, Y.-M. *Proc. Natl Acad. Sci. USA* **117**, 12584–12591 (2020).

4. 4.

Fennell, P. S., Davis, S. J. & Mohammed, A. *Joule* **5**, 1305–1311 (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02612-5>

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

- EDITORIAL
- 29 September 2021

What a personal saga reveals about scientists' lives — and about science itself

Two scientists allowed *Nature* to chronicle their lives for three years. Their story speaks to the epic professional and personal struggles involved in establishing a career in research.



Molecular biologist Daniel Bose, pictured shortly before a grant interview at the Royal Society in London. Credit: Chris Maddaloni for *Nature*

In 2018, a team of *Nature* reporters and editors began documenting, in real time, the lives and experiences of two scientists at the University of Sheffield, UK. Alison Twelvetrees, a neuroscientist, and molecular biologist Daniel Bose are on a path to establishing their own research laboratories. They are also a married couple, and their stories — the highs and lows, the triumphs and tribulations — are told in a [three-part Feature](#) (and can be heard in a [four-part *Nature Podcast* series](#)).



[Starting up in science: two biologists struggle to launch their labs](#)

The intention, with Ali and Dan's agreement, was to present their lives in science over a year or more. Documenting such a process is not very common in science reporting, where the emphasis is more often on describing results. Our aim was to chronicle the journey involved in becoming a principal investigator (PI). But neither we at *Nature*, nor Ali and Dan, knew whether they would be able to build up their research groups, or that the story would run for more than three years. In addition to other crises that arose, the pandemic would shut down their experiments.

The United Kingdom's universities — like those of many countries — are powered by people such as Ali and Dan, who were employed as PIs on fixed-term contracts. In the United Kingdom, some 74,000 academic staff — out of a total of 223,000 — are on such contracts. In the smaller group of staff that do just research, 35,000 out of 50,000 are on fixed-term contracts,

according to data from the UK Higher Education Statistics Agency. For the aspiring academic researcher, such a precarious existence is, sadly, a rite of passage.

As if that wasn't hard enough, applying for a PI position — leading, eventually, to a professorship — is not like applying for a permanent job in many other professions. Researchers wanting to become PIs in universities are required to show evidence of a multiplicity of skills. They must be leaders and managers; entrepreneurs; mentors and teachers; accountants and administrators. And all the while, they must be doing world-leading research and building their publication list. Moreover, not all PI posts are permanent positions. In some cases, candidates must compete with their peers to secure income from external grants to pay their salary.



Neuroscientist Alison Twelvetrees, pictured in her lab at the University of Sheffield. Credit: Chris Maddaloni for *Nature*

Cycle of precariousness

Ali and Dan's attempts to win funding provide some of the most revealing aspects of the narrative — for example, having to face questions from up to 20 interviewers for a fellowship worth more than £1 million (US\$1.4 million). The researchers had a one-in-five chance of success. If they were successful, their funding would support new research talent in the form of PhDs and postdocs. But, as Ali and Dan's story indicates, these newer entrants would also be employed on temporary contracts — thus perpetuating the cycle of precariousness in future generations of researchers.



[Starting up in science: an agonizing search for cash confronts two labs](#)

The employers of doctors, teachers, architects and engineers do not expect candidates to raise funding to pay their own salaries. Scientists working at universities should not be expected to do that either.

In the United Kingdom, a previous generation of research planners anticipated a situation in which researchers might one day find themselves struggling to pay the rent. A funding principle called the dual support system followed. Its architects established two sources of public funding: one funding pot to pay salaries for staff, and a second for grants and fellowships. It meant that researchers had access to a secure income stream to support their families while applying for grants. Today, that principle is under strain. Ali and Dan's story emphasizes why something closer to the original plan, which provides greater job security, is needed.

But Dan and Ali's story isn't just one of funding. It is also a chronicle of the process of science. Media reporting of science typically covers major findings or policy decisions. But, as researchers know only too well, such outcomes are the final steps in a much longer and more-complex process that typically doesn't make it into news stories: the joy of receiving a new microscope, or seeing a student's experiment succeed; the stress of explaining complex science to a lecture theatre packed with students; or the disappointment of getting a funding rejection. These don't always get covered, and that can create an unbalanced view of what science is.

We're grateful to Ali and Dan for allowing *Nature* a glimpse into their lives, to witness the day-to-day struggles, the anxieties, the crises and the victories, large and small. None of us expected this project to last three years. In publishing Ali and Dan's experiences, we hope to redress some of the imbalance, and to provide a key missing piece of the picture of what it means to be an academic scientist today.

Nature **597**, 594 (2021)

doi: <https://doi.org/10.1038/d41586-021-02613-4>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02613-4>

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

- WORLD VIEW
- 23 September 2021

The incoming Afghan government must allow immunizations



The world must work with Kabul's new rulers to get polio and other diseases under control.

- [Zulfiqar A. Bhutta](#) ✉

Since the Taliban retook Afghanistan last month, I've been worrying about the resurgence of polio and other diseases there and beyond.

Children's and women's health was already precarious in Afghanistan, despite fragile gains in the past two decades. I fear that it will now get precipitously worse. That would be a local humanitarian disaster with global implications. Both polio and COVID-19 infections in Afghanistan could spread to neighbouring countries. Until transmission is interrupted, the entire world is at risk.

As of 25 August, the World Bank and the International Monetary Fund have suspended funding from development funds committed to Afghanistan, and the United States has frozen nearly US\$7 billion of Afghan government funds. There is now no money to pay non-governmental organizations (NGOs) and health-care workers who deliver basic health services. The country's stockpile of about three million doses of COVID-19 vaccines seems likely to expire. To avert the collapse of public-health programmes, leaders must engage with the Taliban running the country.



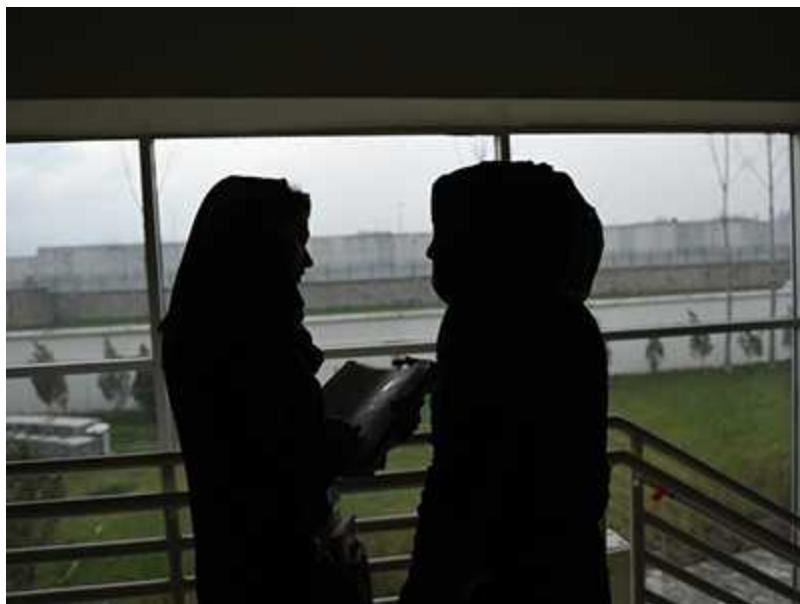
[Polio eradication hinges on child health in Pakistan](#)

I am the founding director of the Centre of Excellence in Women and Child Health at the Aga Khan University in Karachi, Pakistan, which includes a teaching hospital in Kabul; my group has worked across Afghanistan to monitor polio exposure and to improve health and nutrition for mothers and children.

At the turn of the century, fewer than one-third of Afghan children had received any immunizations. One of the earliest acts by United Nations agencies after the overthrow of the Taliban in 2001 to launch nationwide measles-immunization campaigns, which targeted 17 million children. In 2002 and 2003, the campaigns vaccinated 96% of the targeted population. That expanded into more preventive health services; by 2014, rates of

childhood immunizations doubled to about 60%, although this figure hid huge disparities between urban and rural populations.

As the Taliban gained more control across the country, they banned door-to-door visits by polio workers. Cases in Afghanistan tripled between 2018 and 2020. Approximately three million children, a third of those eligible, were thus left out of vaccination campaigns. At least eight vaccinators and polio workers were killed by unidentified gunmen in several attacks this year.



The global research community must not abandon Afghanistan

But the Taliban's objections to these health services were due to its perception that such home visits were being used as a front for intelligence gathering and potential air strikes. Although the US Central Intelligence Agency did collect information to target Osama Bin Laden using fake vaccination workers in 2011 (with catastrophic implications for trust in such a crucial arm of global public health), there is no clear evidence for more recent allegations of similar activities. Some restrictions on house-to-house vaccination campaigns were softened in 2019 (although COVID-19 lockdowns subsequently halted many efforts). I am optimistic that Taliban leaders could be persuaded to permit vaccinations again.

The Taliban now running the country has an opportunity to show a pragmatic, reformist face to the world and people of Afghanistan: it needs to run the health system, to care more about protecting women and children than obsessing about spies and political opponents. It should honour its stated commitment to allow female health-care workers and educators to continue working, for girls to get back to school and for the polio-immunization programme to resume.

The infrastructure for immunizations exists. In early September, the World Health Organization tallied its polio workers and estimated that 90% of the more than 2,000 health facilities were still (somewhat) operational.

Starting in 2018 with \$35 million from the World Bank's Global Financing Facility, as well as a \$140-million World Bank loan, Afghanistan created a \$600-million, 3-year programme to support family planning, health and nutrition services in 34 provinces of Afghanistan. Even with the deteriorating security situation, health-care visits for children under 5 increased by 25% and newborn vaccinations by 8%.



[Afghanistan's terrified scientists predict huge research losses](#)

Frozen assets are unlikely to pressure the Taliban into any major change in its policies. Global organizations should continue supporting the NGOs and existing primary-care workers and programmes on the ground. There has

been a promising confluence of neighbouring and friendly countries helping to stabilize Afghanistan's infrastructure (such as airports) and providing other support through a \$1.2-billion emergency fund, to help Afghans without food, health care or protection. Already, at least one-third of Afghans who are facing hunger. Funds are urgently needed to help secure food supplies through the winter, reopen financial services such as banks and let the estimated 500,000 people displaced this year return to their homes.

Countries should formalize their support, and include provisions for health care in the agreements. Critics protest that nothing should be done to legitimize the Taliban, and that their word (particularly regarding women and girls) cannot be trusted. Although international pressure and scrutiny are essential, the best strategy is to engage with the Taliban to scale up vaccination and health services.

I travelled through Afghanistan in the early 1970s, -tracing family links to Ghazni. I was struck by the resilience and -hospitality of a proud people, despite grinding poverty and inequity. They deserve unfettered support in this hour of need.

Nature **597**, 595 (2021)

doi: <https://doi.org/10.1038/d41586-021-02557-9>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02557-9>

- RESEARCH HIGHLIGHT
- 24 September 2021

An iron grip could help pee to produce electricity

Catalyst based on nickel and iron allows electrical current to be harvested from the breakdown of urea.



Reactions that break down urea, one of the primary components of urine, can produce an electrical current. Credit: Getty

A catalyst containing nickel and iron can degrade the main waste product from urine to generate electricity and harmless by-products¹.

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 **597**, 596 (2021)

doi: <https://doi.org/10.1038/d41586-021-02559-7>

References

1. 1.

Geng, S.-K. *et al.* *Nature Energ.* <https://doi.org/10.1038/s41560-021-00899-2> (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02559-7>

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

- RESEARCH HIGHLIGHT
- 21 September 2021

Life-sized camels in rock are a Stone Age masterpiece

Saudi Arabia's Camel Site is a testimony to the dexterity of prehistoric sculptors.



A camel's body, complete with hump, is depicted in the remnants of a prehistoric sculpture at the Camel Site in Saudi Arabia. Credit: Fayed Nureldine/AFP/Getty

Stunning reliefs of camels in a rock formation in Saudi Arabia are far older than was first thought: they were carved more than 7,000 years ago, when the climate of Arabian Peninsula was markedly cooler and wetter than it is today¹.

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 **597**, 596 (2021)

doi: <https://doi.org/10.1038/d41586-021-02548-w>

References

1. 1.

Guagnin, M. *et al.* *J. Arch. Sci. Rept.*
<https://doi.org/10.1016/j.jasrep.2021.103165> (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02548-w>

- RESEARCH HIGHLIGHT
- 23 September 2021

Tiny particles could make a powerful COVID vaccine

Nanoparticles that bristle with dozens of copies of a SARS-CoV-2 protein segment could protect against a wide variety of coronaviruses.

 Colorized scanning electron micrograph of a cell infected with SARS-CoV-2 virus particles

A human cell infected with SARS-CoV-2 particles (purple; artificially coloured), which can be fended off by a vaccine based on the virus's receptor binding domain. Credit: NIAID ([CC BY 2.0](#))

An experimental COVID-19 vaccine that uses an atypical ingredient to stimulate an immune response shows promise against several variants of the coronavirus — at least in monkeys.

The mRNA COVID-19 vaccines stimulate the immune system by exposing it to the SARS-CoV-2 protein, called spike, that enables the virus to latch on to cells. David Veesler and Neil King at the University of Washington in Seattle and their colleagues made a vaccine that relies instead on a spike fragment called the receptor binding domain (RBD). They injected monkeys with nanometre-scale particles studded with dozens of RBDs and found that the vaccine generated virus-blocking 'neutralizing' antibodies against the Alpha, Beta and Gamma variants, and related animal coronaviruses. This vaccine is currently in phase III clinical trials.

The team also tested nanoparticle vaccines containing RBDs from a mix of sarbecoviruses, a family of coronaviruses that includes SARS-CoV-2. In mice, these vaccines triggered production of neutralizing antibodies against multiple coronaviruses.

These findings pave the way for a pan-sarbecovirus vaccine that could protect against SARS-CoV-2 variants and sarbecoviruses that jump from animals to humans in the future.

Nature **597**, 596 (2021)

doi: <https://doi.org/10.1038/d41586-021-02561-z>

References

1. 1.

Walls, A. C. *et al.* *Cell* <https://doi.org/10.1016/j.cell.2021.09.015> (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02561-z>

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

- RESEARCH HIGHLIGHT
- 21 September 2021

Bright lights at night scramble the sweet song of crickets

Exposure to stronger artificial light led to greater disruption of the insects' night music.



Field crickets usually synchronize their chirps, but that coordination fails if the insects are exposed to artificial light after sunset. Credit: Sutthiwat Srikhrueadom/Getty

Around the world, night is marked by the male crickets' chirps, sounds they make by rubbing raspy modified forewings together in an effort to attract females. But artificial lights can disrupt these melodies, experiments show¹.

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 **597**, 596 (2021)

doi: <https://doi.org/10.1038/d41586-021-02560-0>

References

1. 1.

Levy, K. *et al.* *Proc. R. Soc. B.* <https://doi.org/10.1098/rspb.2021.1626> (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02560-0>

- RESEARCH HIGHLIGHT
- 24 September 2021

Thirty lasers act as one to shine a powerful light

Cluster of tiny tube-shaped devices creates an intense, focused light source.

 Artist's impression of a topological insulator vertical-cavity laser array

A single coherent light beam (pink) is emitted by an array of 30 individual lasers. Credit: SimplySci Animations

An array of 30 lasers that acts as a single light source represents a key step towards the development of large-scale, high-powered lasers on a microchip¹.

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 **597**, 597 (2021)

doi: <https://doi.org/10.1038/d41586-021-02584-6>

References

1. 1.

Dikopoltsev, A. *et al.* *Science* **373**, 1514–1517 (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02584-6>

- RESEARCH HIGHLIGHT
- 26 September 2021

The oil-field ‘fugitive’ that can slip into the water supply

Methane gas that steals away from leaky fossil-fuel wells can hide underground and taint groundwater.



A natural-gas wellhead in northern British Columbia, Canada. Such wells can leak methane into the environment. Credit: Aaron Black/Getty

Some of the natural gas that leaks from oil and gas wells can stay trapped underground, where it dissolves in groundwater and can affect water quality¹.

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 **597**, 597 (2021)

doi: <https://doi.org/10.1038/d41586-021-02550-2>

References

1. 1.

Chao, J. T.-H. *et al.* *Geophys. Res. Lett.*
<https://doi.org/10.1029/2021GL095200> (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02550-2>

- RESEARCH HIGHLIGHT
- 21 September 2021

A mega-penguin stood tall on prodigious limbs

Remains discovered by schoolchildren on a fossil-hunting trip belonged to an extinct giant bird.



The giant penguin *Kairuku waewaeroa* (artist's impression) had both a long bill and long legs. Credit: Simone Giovanardi

In 2006, school kids in New Zealand discovered the fossilized remains of a penguin the size of a ten-year-old child. The giant bird turns out to be a new species of penguin that, thanks to its long legs, was taller than any of its modern relatives¹.

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 **597**, 597 (2021)

doi: <https://doi.org/10.1038/d41586-021-02562-y>

References

1. 1.

Giovanardi, S., Ksepka, D. T. & Thomas, D. B. *J. Vertebr. Paleontol.*
<https://doi.org/10.1080/02724634.2021.1953047> (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02562-y>

News in Focus

- **[Moon ice, research imbalance and a new science minister](#)** [29 September 2021]
News Round-Up • The latest science news, in brief.
- **[Ancient footprints could be oldest traces of humans in the Americas](#)** [23 September 2021]
News • Children left tracks in New Mexico around 22,500 years ago — thousands of years before most scientists thought humans settled in North America.
- **[Closest known relatives of virus behind COVID-19 found in Laos](#)** [24 September 2021]
News • Studies of bats in China and Laos show southeast Asia is a hotspot for potentially dangerous viruses similar to SARS-CoV-2.
- **[Resistance to front-line malaria drugs confirmed in Africa](#)** [23 September 2021]
News • The artemisinin-based treatments are taking longer to clear infections. But they are still working — for now.
- **[Young people's climate anxiety revealed in landmark survey](#)** [22 September 2021]
News • Children worldwide worry about the future and feel let down by governments, a huge study on attitudes towards climate change has found.
- **[COVID vaccine immunity is waning — how much does that matter?](#)** [17 September 2021]
News Explainer • As debates about booster shots heat up, what's known about the duration of vaccine-based immunity is still evolving.
- **[Starting up in science: two biologists struggle to launch their labs](#)** [29 September 2021]
News Feature • Nature follows two scientists as they fight against the odds to build their careers.
- **[Starting up in science: an agonizing search for cash confronts two labs](#)** [29 September 2021]
News Feature • In part two of a series, a pair of researchers race to secure the funding that will keep their operations going.

- **Starting up in science: two labs face the pandemic and another shock** [29 September 2021]

News Feature • Just as two scientists manage to get their labs going, the COVID outbreak hits. Then comes a personal setback in part three of a series.

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

- NEWS ROUND-UP
- 29 September 2021

Moon ice, research imbalance and a new science minister

The latest science news, in brief.



Researchers at NASA's Glenn Research Center in Cleveland, Ohio, test a model of VIPER on simulated lunar terrain. Credit: NASA/GRC/Bridget Caswell

Will NASA's Moon rover find enough of the ice it seeks?

NASA plans to land its next lunar rover beside a crater, named Nobile, near the Moon's south pole. But some scientists question whether the mission, set to launch in 2023, will efficiently find the lunar ice that it's looking for.

The mission, known as the Volatiles Investigating Polar Exploration Rover (VIPER), will be the [first to visit the Moon's south pole](#), which holds great scientific promise: it receives little sunlight, so it has reserves of ice containing information about the origin and evolution of the Solar System.

The agency acknowledges that VIPER (pictured, a test version on Earth) does not yet have a detailed map of where the ice is around Nobile. "This is exactly why we're going," says Anthony Colaprete, VIPER's project scientist at NASA's Ames Research Center in Moffett Field, California.

But some scientists say VIPER would be better off if NASA first launched a different planned spacecraft: a small, US\$55-million satellite called Lunar Trailblazer. Lunar Trailblazer's mission is to map water on the Moon. Researchers say the VIPER team could use these maps to help the rover to prospect efficiently for ice during its fast-paced, 100-day mission. It's an "opportunity missed", says Clive Neal, a geoscientist at the University of Notre Dame in Indiana.

Development research led mainly by scholars in global north

Research on economic issues relating to developing countries is [led predominantly by people based in the global north](#), according to an analysis of nearly 25,000 papers. The findings show that although many studies focus on countries or regions in the global south, researchers based there have been vastly under-represented in the literature for decades.

Economists analysed data on journal articles, citations and conference presentations in development and development-policy research. They found that just 16% of 24,894 articles published in 20 high-profile development journals between 1990 and 2019 were authored by researchers based in the global south, compared with 73% authored by researchers in the global north

and 11% that were collaborations between researchers in the north and south (see ‘Under-represented in research’).

UNDER-REPRESENTED IN RESEARCH

A large proportion of research on economic development does not involve any researchers who are based in the global south.

■ Global north ■ Global south ■ Both

All articles*



Articles with a specific focus on country or region in global south†



*For 24,894 papers published in 20 high-profile development journals between 1990 and 2019.

†For a subset of 15,117 papers that focus on a country or region in the global south.

©nature

Source: [\(2021\)](https://doi.org/gxgq).

“Once you begin to quantify it in a systematic way, the gravity of the problem really just comes home,” says co-author Grieve Chelwa, an economist now at the New School in New York City.

In a subset of 15,117 articles explicitly focused on a country or region in the global south, the team found that 62% were authored by researchers based in the global north. ‘Southern’ researchers were also under-represented in article citations and among presenters at international development conferences ([V. Amarante et al. Appl. Econ. Lett. https://doi.org/gxgq; 2021](https://doi.org/gxgq)).

The authors defined ‘southern’ researchers as those working at a university or organization based in any country of Latin America, Asia or Africa, including the Middle East, whereas ‘northern’ researchers were defined as those based anywhere else.

Limited access to research funding and opportunities in some countries, and migration of researchers from the global south to the global north, might partially explain the dominance of the north in development studies, says Chelwa. But he and other development researchers suspect that exclusion and exploitation of global-south-based researchers could also contribute.

Chelwa adds that, in future, he would like to see more “equal partnerships” between researchers in the global north and south, and increased representation of academics from the global south on journal editorial boards.



George Freeman, pictured in 2019.Credit: Leon Neal/Getty

UK appoints new science minister

The United Kingdom has a [new science minister](#) — its ninth since 2010, following a reshuffle of Prime Minister Boris Johnson’s cabinet. George Freeman, a former investor in life-sciences companies, takes the role as the coronavirus pandemic puts renewed focus on research.

Freeman has been a Conservative Member of Parliament since 2010, and has previously served as a government life-sciences adviser and as transport

minister. He takes over his latest role — a junior ministerial job — from Amanda Solloway, who was appointed in early 2020.

His background, as a minister and a biotechnology venture capitalist, makes Freeman a relatively popular choice for some researchers.

The reshuffle comes less than two months before the government's spending review, in the run-up to which ministers must vie for funding for their departments. Scientists will be watching closely to see how Freeman plans to achieve the United Kingdom's bold target of increasing public and private research and development investment to 2.4% of gross domestic product by 2027.

The country's relationship with Europe is also likely to influence the science minister's agenda.

Nature **597**, 599 (2021)

doi: <https://doi.org/10.1038/d41586-021-02614-3>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02614-3>

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

- NEWS
- 23 September 2021

Ancient footprints could be oldest traces of humans in the Americas

Children left tracks in New Mexico around 22,500 years ago — thousands of years before most scientists thought humans settled in North America.

- [Ewen Callaway](#)



Ancient footprints found at the New Mexico site. The tracks were dated to between 21,000 and 23,000 years old, and probably belonged to children and teenagers. Credit: National Park Service, USGS and Bournemouth University

White Sands National Park, in southern New Mexico, is known for chalk-coloured dunes that stretch for hundreds of square kilometres. But at the height of the last Ice Age, the region was wetter and grassier. Mammoths,

giant sloths and other animals walked the muddy shores of shallow lakes that grew and shrank with the seasons. And they had company.

In a landmark study published on 23 September in *Science*¹, researchers suggest that human footprints from an ancient lakeshore in the park date to between 21,000 and 23,000 years old. If the dating is accurate — which specialists say is likely — the prints represent the earliest unequivocal evidence of human occupation anywhere in the Americas.



Ancient stone tools hint at settlers' epic trek to North America

“The evidence is very convincing and extremely exciting,” says Tom Higham, an archaeological scientist and radiocarbon-dating expert at the University of Vienna. “I am convinced that these footprints genuinely are of the age claimed.”

The dates raise questions about when and how humans from Siberia settled in the region, with evidence growing that they skirted the Pacific coast while inland routes were entrenched in ice. The authors of the study say the footprints give credence to contentious evidence of even earlier signs of settlement in the Americas.

“The paper makes a very compelling case that these footprints are not only human, but they’re older than 20,000 years,” says Spencer Lucas, a

palaeontologist at the New Mexico Museum of Natural History & Science in Albuquerque. “That’s a game-changer.”

Rocky evidence

For decades, archaeologists associated the earliest Americans with 11,000–13,000-year-old stone spear points and other vestiges of ‘Clovis’ culture (named after another New Mexico site, but found throughout North America). The dates coincide with the recession of a continent-size glacier, which created an ice-free corridor through central Canada.

The discovery of numerous 'pre-Clovis' archaeological sites, from Alaska to the tip of South America, dating to as old as 16,000 years, sowed doubts about the ‘Clovis-first’ hypothesis and argued for a coastal migration route from Siberia.

Research journals are dotted with claims of even earlier sites, including a controversial *Nature* paper that put humans in California 130,000 years ago². But many of these claims have been discounted because of the equivocality of the evidence: rocks potentially mistaken for tools, marks on animal bones that might have been made by natural processes — or diggers, in the case of the California claim — rather than butchery.



Excavations in White Sands National Park reveal human footprints at the base of a trench. Credit: National Park Service, USGS and Bournemouth University

White Sands is filled with human and animal fossil footprints — in 2018, the same team that found the tracks in the latest paper documented a giant sloth hunt on a dried-up lake bed known as a playa³. But these tracks are notoriously difficult to date, says study co-author Matthew Bennett, a geoscientist at Bournemouth University in Poole, UK, who specializes in the study of fossil footprints. “Every time you uncover something it’s potentially a different age. Dating is a nightmare.”

Ancient seeds

In 2019, study co-author David Bustos, an archaeologist and resource manager at White Sands, identified a site on the playa that had tracks that led right into layers of rock-hard sediment. The rock contained seeds of

spiral ditchgrass (*Ruppia cirrhosa*), an aquatic plant that could be carbon-dated to determine the age of the tracks. “That’s the holy grail of trying to date footprints,” says Bennett.

He and his colleagues weren’t surprised when radiocarbon dating by researchers at the US Geological Survey in Denver, Colorado, determined that the seeds were between 21,000 and 23,000 years old, because a previous small-scale excavation had dated the sediment to around the same time. But Bennett says the team knew that claims of human occupation at this age would draw extreme scrutiny.



Ancient voyage carried Native Americans' DNA to remote Pacific islands

So they attempted to address factors that could skew the seeds’ ages. The most likely was a phenomenon whereby organisms incorporate carbon that has leached into the water from nearby rocks, such as calcium carbonate in limestone. Such carbon sources tend to be much older than the carbon in Earth’s atmosphere.

The researchers say such ‘reservoir effects’ are unlikely. They dated hundreds of seeds in different sediment layers and their ages fell into line, with older seeds at the bottom, younger on top. If the seeds had incorporated old carbon, there would probably have been more variation, says co-author Daniel Odess, an archaeologist at the US National Park Service in

Washington DC. At a site in the region that didn't have any footprints, spiral ditchweed seeds date to the same age as charcoal in the same layer — which is not subject to reservoir effects.

"I really think those ages are okay," says Thomas Stafford, an experimental geochronologist at Stafford Research Laboratories in Lafayette, Colorado. Even a 1,000-year error wouldn't tarnish the importance of the footprints, he points out. "Whether people were here 20,000, or 22,000, or 19,000 years ago, does not change their incredible story," Stafford adds. "We have human footprints."

Teenage tracks

The team determined that the several dozen tracks probably belonged to numerous individuals, mostly children and teenagers. "To me this makes perfect sense," says Odess. "When I was young I was always heading to the water. Stream, river, pond, whatever it was. Given the chance, I would probably walk in mud more than dry ground."

Karen Moreno, a palaeoichnologist at Austral University of Chile in Valdivia, has no doubt that the tracks are human. She isn't yet convinced that they were mostly made by children, because these estimates are based on the statures of modern people. But she says the tracks could shine a light on the earliest humans in America. "This older community most probably had a different and complex way of life."



Is this cave painting humanity's oldest story?

Now that there is strong evidence that humans settled the Americas more than 20,000 years ago, researchers should grapple with the consequences, says Bennett. He hopes the White Sands footprints will force researchers to reconsider sites that have more equivocal evidence of early human occupation.

David Meltzer, an archaeologist at Southern Methodist University in Dallas, Texas, is convinced by the White Sands footprints, but disagrees that they give credence to the more controversial sites. However, if stone tools or other artefacts associated with the track-makers could be discovered, this could allow such connections to be drawn, Meltzer adds.

The footprints make it “extremely likely” that the ancestors of the White Sands humans and other early settlers travelled along the Pacific coast, says Higham. The next step will be to identify the people who arrived through these Ice Age voyages, he adds. “An urgent research priority is not just to find footprints such as these, but the remains of the people who made them.”

Nature **597**, 601-602 (2021)

doi: <https://doi.org/10.1038/d41586-021-02597-1>

References

1. 1.

Bennett, M. R. *et al. Science* <https://doi.org/10.1126/science.abg7586> (2021).

2. 2.

Holen, S. R. *et al. Nature* **544**, 479–483 (2017).

3. 3.

Bustos, D. *et al. Sci. Adv.* **4**, eaar7621 (2018).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02597-1>

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

- NEWS
- 24 September 2021
- Correction [27 September 2021](#)

Closest known relatives of virus behind COVID-19 found in Laos

Studies of bats in China and Laos show southeast Asia is a hotspot for potentially dangerous viruses similar to SARS-CoV-2.

- [Smriti Mallapaty](#)



A *Rhinolophus* bat exits a cave at dusk in Perak, Malaysia. Credit: Fletcher & Baylis/SPL

Scientists have found three viruses in bats in Laos that are more similar to SARS-CoV-2 than any known viruses. Researchers say that parts of their genetic code bolster claims that the virus behind COVID-19 has [a natural origin](#) — but their discovery also raises fears that there are numerous coronaviruses with the potential to infect people.

David Robertson, a virologist at the University of Glasgow, UK, calls the find “fascinating, and quite terrifying”.



Did the coronavirus jump from animals to people twice?

The results, which are not peer reviewed, have been posted on the preprint server Research Square¹. Particularly concerning is that the new viruses contain receptor binding domains that are almost identical to that of SARS-CoV-2, and can therefore infect human cells. The receptor binding domain allows SARS-CoV-2 to attach to a receptor called ACE2 on the surface of human cells to enter them.

To make the discovery, Marc Eloit, a virologist at the Pasteur Institute in Paris and his colleagues in France and Laos, took saliva, faeces and urine samples from 645 bats in caves in northern Laos. In three horseshoe (*Rhinolophus*) bat species, they found viruses that are each more than 95%

identical to SARS-CoV-2, which they named BANAL-52, BANAL-103 and BANAL-236.

Natural origin

“When SARS-CoV-2 was first sequenced, the receptor binding domain didn’t really look like anything we’d seen before,” says Edward Holmes, a virologist at the University of Sydney in Australia. This caused some people to speculate that the virus had been [created in a laboratory](#). But the Laos coronaviruses confirm these parts of SARS-CoV-2 exist in nature, he says.

“I am more convinced than ever that SARS-CoV-2 has a natural origin,” agrees Linfa Wang, a virologist at Duke–NUS Medical School in Singapore.

Together with relatives of SARS-CoV-2 discovered in Thailand², [Cambodia](#)³ and Yunnan in southern China⁴, the study demonstrates that southeast Asia is a “hotspot of diversity for SARS-CoV-2-related viruses”, says Alice Latinne, an evolutionary biologist at the Wildlife Conservation Society Vietnam in Hanoi.



[Coronaviruses closely related to the pandemic virus discovered in Japan and Cambodia](#)

In an extra step in their study, Eloit and his team showed in the laboratory that the receptor binding domains of these viruses could attach to the ACE2 receptor on human cells as efficiently as some early variants of SARS-CoV-2. The researchers also cultured BANAL-236 in cells, which Eloit says they will now use to study how pathogenic the virus is in animal models.

Last year, researchers described another close relative of SARS-CoV-2, called RaTG13, which was found in bats in Yunnan⁵. It is 96.1% identical to SARS-CoV-2 overall and the two viruses probably shared a common ancestor 40–70 years ago⁶. BANAL-52 is 96.8% identical to SARS-CoV-2, says Eloit — and all three newly discovered viruses have individual sections that are more similar to sections of SARS-CoV-2 than seen in any other viruses.

Viruses swap chunks of RNA with one another through a process called recombination, and one section in BANAL-103 and BANAL-52 could have shared an ancestor with sections of SARS-CoV-2 less than a decade ago, says Spyros Lytras, an evolutionary virologist at the University of Glasgow. “These viruses recombine so much that different bits of the genome have different evolutionary histories,” he says.

Missing links

The Laos study offers insight into the origins of the pandemic, but there are still missing links, say researchers. For example, the Laos viruses don’t contain the so-called furin cleavage site on the spike protein that further aids the entry of SARS-CoV-2 and other coronaviruses into human cells.

The study also doesn’t clarify how a progenitor of the virus could have travelled to Wuhan, in central China, where the first known cases of COVID-19 were identified — or whether the virus hitched a ride on an intermediate animal.



After the WHO report: what's next in the search for COVID's origins

Answers might come from sampling more bats and other wildlife in southeast Asia, which many groups are doing.

Another preprint, also posted on Research Square and not yet peer reviewed, sheds light on the work under way in China⁷. For that study, researchers sampled some 13,000 bats between 2016 and 2021 across China. But they did not find any close relatives of SARS-CoV-2, and conclude that these are “extremely rare in bats in China”.

But other researchers question this claim. “I strongly disagree with the suggestion that relatives of SARS-CoV-2 may not be circulating in Chinese bats, as such viruses have already been described in Yunnan,” says Holmes.

The corresponding author of the study declined to respond to *Nature*'s questions about the findings, because the paper is still under review.

Wang says that both studies highlight the importance of ramping up sampling in regions outside China to help uncover the origins of the pandemic.

Nature **597**, 603 (2021)

doi: <https://doi.org/10.1038/d41586-021-02596-2>

Updates & Corrections

- **Correction 27 September 2021:** An earlier version of this story was published with a different primary image, which was incorrectly captioned as featuring *Rhinolophus* bats. The story has been updated.

References

1. 1.
Temmam, S. *et al.* Preprint at Research Square
<https://doi.org/10.21203/rs.3.rs-871965/v1> (2021).
2. 2.
Wacharapluesadee, S. *et al.* *Nature Commun.* **12**, 972 (2021).
3. 3.
Hul, V. *et al.* Preprint at bioRxiv
<https://doi.org/10.1101/2021.01.26.428212> (2021).
4. 4.
Zhou, H. *et al.* *Cell* **184**, 4380–4391 (2021).
5. 5.
Zhou, P. *et al.* *Nature* **579**, 270–273 (2020).
6. 6.
Boni, M. F. *et al.* *Nature Microbiol.* **5**, 1408–1417 (2020).
7. 7.
Wu, Z. *et al.* Preprint at Research Square
<https://doi.org/10.21203/rs.3.rs-885194/v1> (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02596-2>

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

- NEWS
- 23 September 2021

Resistance to front-line malaria drugs confirmed in Africa

The artemisinin-based treatments are taking longer to clear infections. But they are still working — for now.

- [Max Kozlov](#)



A doctor shows packets of antimalarial drugs called artemisinin-combination therapies to a mother and son in Mali. Credit: Godong/BSIP/Science Photo Library

Scientists have confirmed that malaria parasites in Africa have developed resistance to a key family of drugs used to protect against them.

“We’ve all been expecting and dreading this for quite some time,” says Leann Tilley, a biochemist at the University of Melbourne in Australia, who researches the molecular basis of antimalarial resistance.

Signs of drug resistance have long been present in Africa: for instance, in Rwanda between 2012 and 2015, scientists detected¹ the existence of gene mutations associated with resistance in malaria parasites. A new study, published in the *New England Journal of Medicine* today², bolsters these findings by showing that such mutations are causing an observable drop in antimalarials’ ability to quickly treat people with the disease.



How to defuse malaria’s ticking time bomb

The ‘gold standard’ treatments for malaria — the drug family including artemisinin and its derivatives — are often administered alongside ‘partner’ drugs in what are called artemisinin-combination therapies (ACTs), because multiple drugs are more difficult for parasites to develop resistance against.

The first signs of resistance to artemisinin and its relatives appeared in Cambodia in the early 2000s. Within a few years, malaria parasites in southeast Asia began to evade some of the partner drugs in the ACTs, too,

rendering some of the most effective drug cocktails against malaria useless in the region and sending public-health officials scrambling to find combinations that still worked.

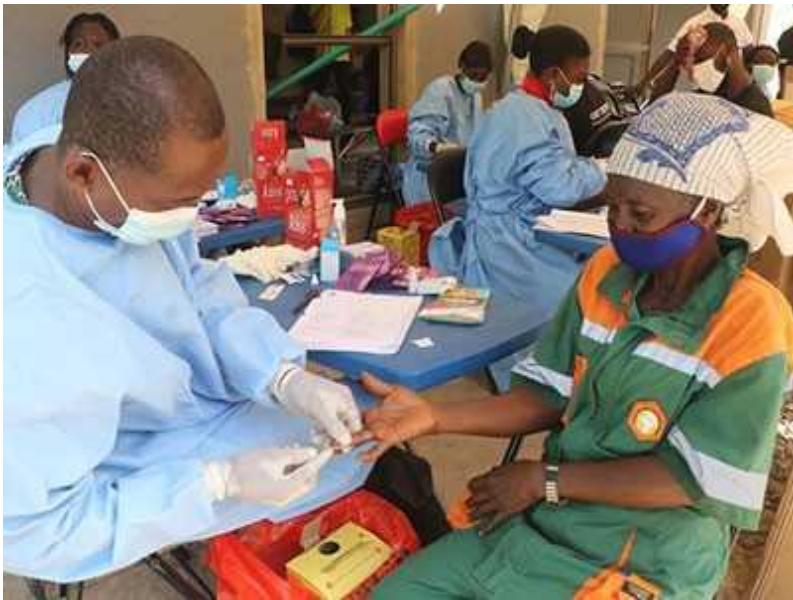
For resistance to now hit Africa is particularly dire, says Tilley. More than 90% of malaria cases and deaths worldwide occur on the continent. Also concerning is that the study found evidence that resistance in Africa arose independently of the resistant parasite strains in southeast Asia, meaning the strains now in Africa might continue evolving to culminate in a ‘super resistant’ parasite that becomes dominant, she says.

“It’s certainly a worrying finding, because we rely completely on these artemisinin-combination therapies,” says Arjen Dondorp, head of malaria research at the Mahidol Oxford Tropical Medicine Research Unit in Bangkok. Scientists fear that a similar scenario to the one in southeast Asia will unfold in Africa, he says — and in light of the lack of access to adequate health care in many parts of sub-Saharan Africa, it could exact a huge toll.

Clinical confidence

In the study, conducted in Uganda from 2017 to 2019, researchers treated 240 people who had malaria by giving them intravenous artesunate, a potent derivative of artemisinin, three times over the course of a day, followed by a standard three-day course of ACT pills. Doctors typically administer artemisinins without partner drugs to people only when they have severe malaria.

The team found that, for 14 participants, it took longer than 5 hours to clear half of the malaria-causing parasites (*Plasmodium falciparum*), which meets the World Health Organization (WHO) definition for resistance. (People with malaria usually clear half the parasites within a couple hours of treatment with artesunate.) Parasites in 13 of these participants had one of two concerning mutations in their *kelch13* gene, which has been linked with antimalarial resistance in southeast Asia³.



How COVID is derailing the fight against HIV, TB and malaria

Although the mutations had already been detected in malaria parasites in Africa, “we didn’t know if these parasites were actually resistant to drugs in humans”, says Toshihiro Mita, a parasitologist at Juntendo University in Tokyo, and a co-author of the study. This investigation, and one published in April in *The Lancet Infectious Diseases*, confirmed scientists’ suspicions.

The April study reported⁴ the results of a three-day course of ACT pills in children with malaria in Rwanda. Some of the children still had parasites after completing the treatment, and more than 10% of the parasites had one of two *kelch13* mutations that are indicative of resistance, but are different from those observed in Uganda.

‘A major wake-up call’

For now, there appear to be few clinical consequences of the artemisinin resistance, explains Philip Rosenthal, an infectious-disease clinician at the University of California, San Francisco, who works with the WHO to study malaria in Uganda. The parasites take longer to clear in some severe cases, and they might return within a week or so, but the ACT of choice in much of sub-Saharan Africa — a combination of artemether, another derivative of

artemisinin, and a partner drug called lumefantrine — seems to still be effective.

Nonetheless, Rosenthal says the findings are a “major wake-up call”: if resistance continues to spread, as it is expected to, and the parasites become resistant to lumefantrine, the result could be disastrous. “In Africa, where huge numbers of young children are treated in rural clinics with very little infrastructure, losing your main drug could be really devastating,” he says.

Resistance might not spread as quickly as it did in southeast Asia, however. Some people in sub-Saharan Africa contract malaria at a young age and then live with parasites in their blood, which, researchers think, helps them to build an immune tolerance to those strains. Resistant parasites might not affect those people to the same degree that they would affect people in southeast Asia; the resistant strains would have to not only contend with this immune tolerance, but also out-compete the parasites that are already living in people, says Tilley.



Malaria vaccine shows promise — now come tougher trials

Still, scientists are nervous — they continue to reel from the failure of chloroquine, an antimalarial drug that *P. falciparum* quickly grew resistant to across the globe starting in the late 1950s. The loss of that drug probably led to millions of excess deaths, says Rosenthal.

Although the WHO [has mounted an aggressive campaign](#) to eliminate resistant parasites in areas of southeast Asia by rapidly deploying ACTs in sick individuals, the same approach is unrealistic in sub-Saharan Africa given how widespread the disease is there, says Pascal Ringwald, who leads the WHO Global Malaria Program's Drug Resistance and Containment Unit.

This study puts even more pressure on researchers and drug makers to find another viable treatment or a vaccine for malaria, in case there is more evidence that ACTs are failing in the future, says Ringwald. Several malaria vaccines are under development, one of which showed promising results⁵ in a small clinical trial earlier this year.

And new options can't arrive too soon. "I expect to see many more reports of artemisinin-resistant strains in the next few years," says Ringwald, "because they're already popping up like mushrooms."

Nature **597**, 604 (2021)

doi: <https://doi.org/10.1038/d41586-021-02592-6>

References

1. 1.

Uwimana, A. *et al.* *Nature Med.* **26**, 1602–1608 (2020).

2. 2.

Balikagala, B. *et al.* *N. Engl. J. Med.* **385**, 1163–1171 (2021).

3. 3.

Ariey, F. *et al.* *Nature* **505**, 50–55 (2014).

4. 4.

Uwimana, A. *et al.* *Lancet Infect. Dis.* **21**, 1120–1128(2021).

5. 5.

Mwakingwe-Omari, A. *et al.* *Nature* **595**, 289–294 (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02592-6>

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

- NEWS
- 22 September 2021

Young people's climate anxiety revealed in landmark survey

Children worldwide worry about the future and feel let down by governments, a huge study on attitudes towards climate change has found.

- [Tosin Thompson](#)



Children at a School Strike for Climate protest in Sydney, Australia.Credit: Richard Milnes/Shutterstock

Climate change is causing distress, anger and other negative emotions in children and young people worldwide, a survey of thousands of 16- to 25-year-olds has found. This ‘eco-anxiety’ has a negative impact on respondents’ daily lives, say the researchers who conducted the survey, and is partly caused by the feeling that governments aren’t doing enough to avoid a climate catastrophe.



The hard truths of climate change — by the numbers

“This study provides arguments for anyone who has any connection to youth mental health — climate change is a real dimension into their mental-health problems,” says Sarah Ray, who studies climate anxiety at Humboldt State University in Arcata, California.

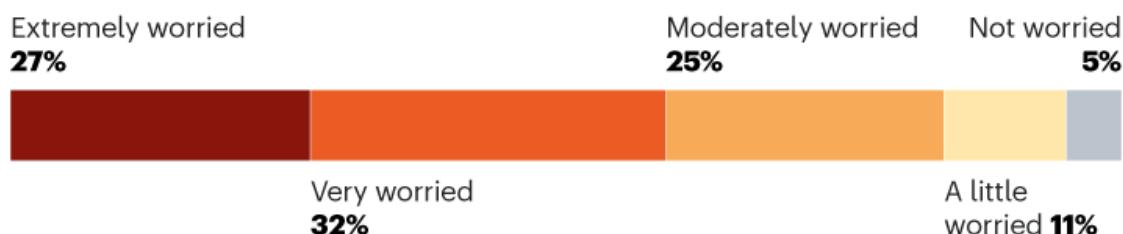
The survey — the largest of its kind — asked 10,000 young people in 10 countries how they felt about climate change and government responses to it.

The results, released in a preprint on 14 September¹, found that most respondents were concerned about climate change, with nearly 60% saying they felt ‘very worried’ or ‘extremely worried’. Many associated negative emotions with climate change — the most commonly chosen were ‘sad’, ‘afraid’, ‘anxious’, ‘angry’ and ‘powerless’ (see ‘Climate anxiety’). Overall, 45% of participants said their feelings about climate change impacted their daily lives.

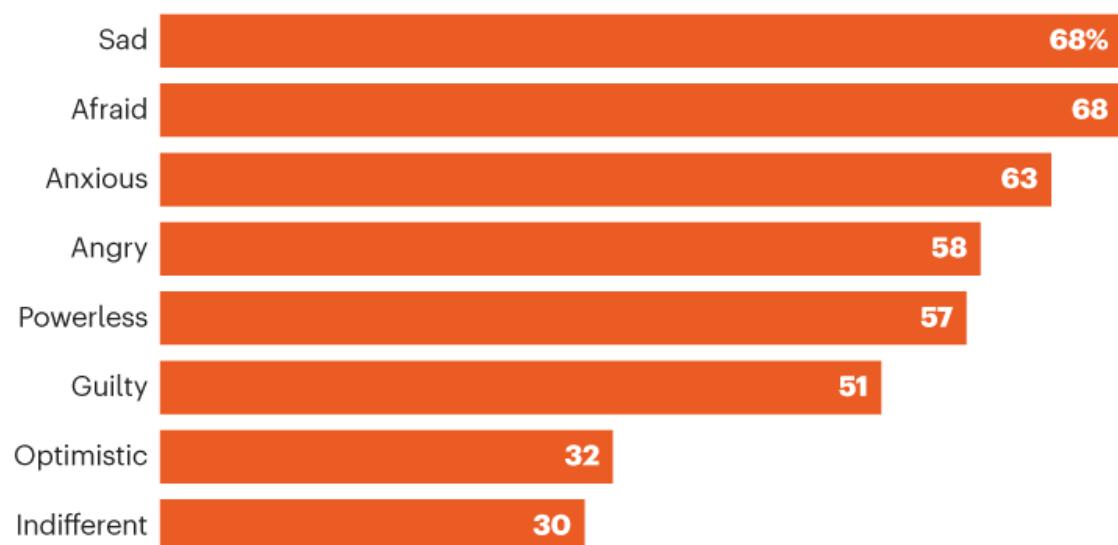
CLIMATE ANXIETY

A survey of 10,000 young people shows that negative feelings about climate change can cause psychological distress.

How worried are you about climate change?



Climate change makes me feel...



©nature

Source: Ref. 1

The countries with the highest proportion of respondents who felt 'very worried' or 'extremely worried' by climate change were the Philippines (84%), India (68%) and Brazil (67%), nations that have been hard-hit by climate change. Portugal — where wildfires are becoming increasingly severe — had the highest level of very worried or extremely worried respondents (65%) out of the high-income countries surveyed, which included France, Finland, Australia and the United States.

“As much as we try to make change, I also see government inaction — trees being cut up on the daily,” Jennifer Uchendu, a climate activist based in Lagos, told reporters at a press conference where the survey results were presented. She recalled feeling anger and grief about government-backed deforestation in her country. “Young people are having to suffer the brunt of these issues when crisis or disasters come in.” In the survey, nearly two-thirds of respondents from Nigeria said that their feelings about climate change had negatively affected their daily life.

A general ‘othering’

Among those who said they have talked to others about climate change (81%), nearly half reported being ignored or dismissed. “There is a general ‘othering’ of children in society, and children’s voices that threaten the predominant narrative of the most powerful group in society,” says study co-author Caroline Hickman, a climate-psychology researcher at the University of Bath, UK. Until now, there has been little research on the psychological impacts of climate change on children, she adds.



[Climate science is supporting lawsuits that could help save the world](#)

When asked about how governments are responding to climate change, 65% of respondents agreed with the statement that governments are failing young people, 64% agreed that they are lying about the impact of actions taken and 60% agreed they were dismissing people’s distress. Only 36% agreed that governments are acting according to science. However, young people do feel reassured when governments act.

Ray says that the results of surveys such as this could play a part in [climate-change lawsuits](#), because they might provide evidence of moral injury under human-rights law. “If the research can help hold government bodies accountable … that would be extraordinary. It would give a voice to children most vulnerable, and eco-anxiety can become less about ‘snowflakes’ in privileged settings worrying about polar bears — which is how detractors see it,” she says.

“I think this research will impact more audiences than other arguments about why we should do more on climate.”

Nature **597**, 605 (2021)

doi: <https://doi.org/10.1038/d41586-021-02582-8>

References

1. 1.

Hickman, C. *et al.* Preprint at <http://dx.doi.org/10.2139/ssrn.3918955> (2021).

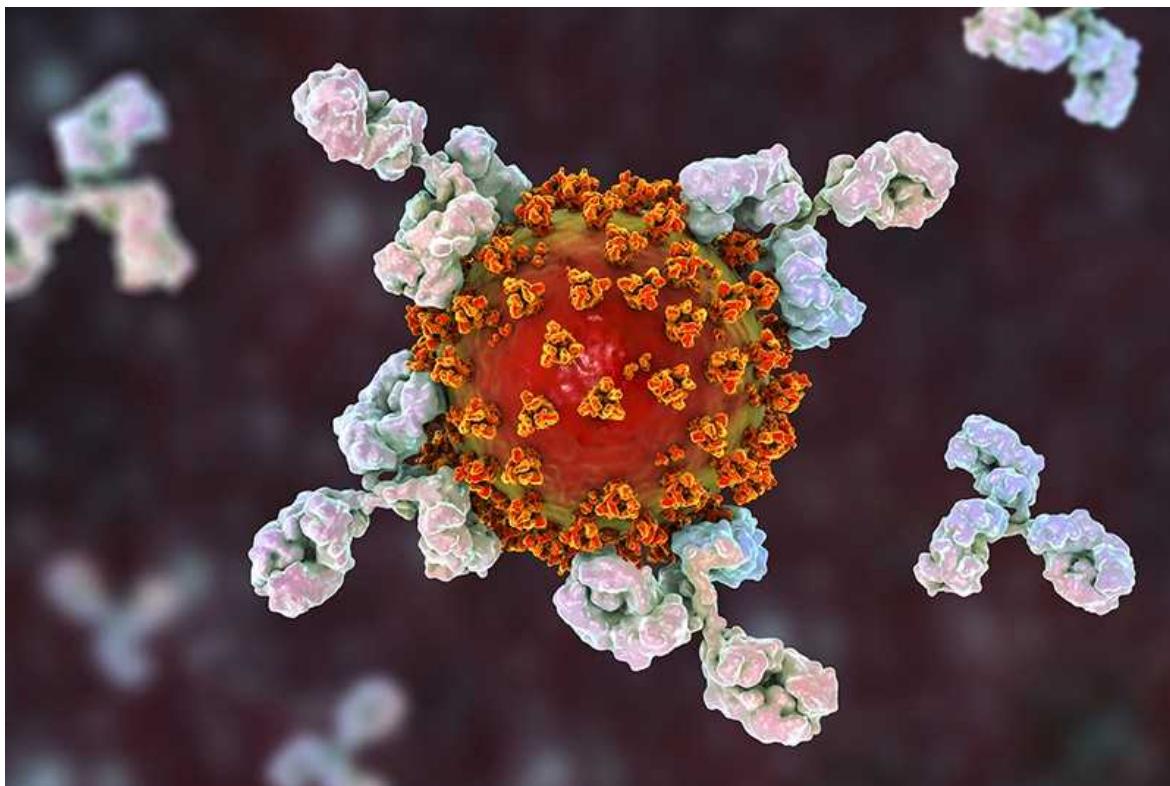
This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02582-8>

- NEWS EXPLAINER
- 17 September 2021

COVID vaccine immunity is waning — how much does that matter?

As debates about booster shots heat up, what's known about the duration of vaccine-based immunity is still evolving.

- [Elie Dolgin](#)



For those vaccinated against COVID-19, antibody levels eventually wane, but this is not the whole story.Credit: Kateryna Kon/Science Photo Library

Six months ago, Miles Davenport and his colleagues made a bold prediction. On the basis of published results from vaccine trials and other data sources, they estimated that people immunized against COVID-19 would lose approximately half of their defensive antibodies every 108 days or so. As a result, vaccines that initially offered, say, 90% protection against mild cases of disease might only be 70% effective after 6 or 7 months¹.

“It felt a little bit out on a limb at the time,” says Davenport, a computational immunologist at the University of New South Wales in Sydney, Australia. But on the whole, his group’s predictions have come true.

Immunological studies have documented a steady decline of antibody levels among vaccinated individuals². Long-term follow-up of vaccine trial participants has revealed a growing risk of breakthrough infection³. And health-care records from countries such as [Israel](#), the [United Kingdom](#) and [elsewhere](#) all show that COVID-19 vaccines are losing their strength, at least when it comes to keeping a lid on transmissible disease.



The fight to manufacture COVID vaccines in lower-income countries

That’s without accounting for the Delta threat either — and it’s clear that vaccine-induced antibodies do a worse job at recognizing SARS-CoV-2 variants compared with the ancestral strain of the virus⁴. What remains unclear, however, is to what degree the immune system’s safeguards that protect vaccinated people against severe disease, hospitalization and death

might be fading as well. “That,” says Davenport, “is the million-dollar question at the moment.”

As [discussions over booster programmes](#) heat up — with leading global-health authorities coming out publicly against the idea this week, UK officials endorsing boosters for the over-50s, and advisers to US regulators meeting on 17 September to discuss the issue — *Nature* takes stock of the data informing the debate.

How is vaccine-induced immunity holding up?

“Things wane,” says Nicole Doria-Rose, an immunologist at the US National Institute of Allergy and Infectious Diseases in Bethesda, Maryland. But not all things wane equally.

‘Neutralizing’ antibodies that can intercept viruses before they infiltrate cells might not have much staying power. Levels of these molecules typically shoot up after vaccination, then quickly taper off months later. “That’s how vaccines work,” Doria-Rose says.

But cellular immune responses are longer lasting — and as Jennifer Gommerman, an immunologist at the University of Toronto in Canada, explains: “Cellular immunity is what’s going to protect you from disease.” Memory B cells, which can rapidly deploy more antibodies in the event of re-exposure to the virus, tend to stick around, and so do T cells, which can attack already-infected cells. Both provide an added measure of protection should SARS-CoV-2 sneak past the body’s first line of defence.



COVID vaccine boosters: the most important questions

In one of the only long-term studies to consider these three planks of the immune system simultaneously — antibodies, B cells and T cells — researchers found that vaccination spurred durable cellular immunity⁵. Memory B cells continued to grow in numbers for at least six months, and got better at fighting the virus over time. T-cell counts remained relatively stable, dipping only slightly over the duration of the study period.

“So, you have this reserve,” says John Wherry, an immunologist at the University of Pennsylvania Perelman School of Medicine in Philadelphia, who led the study. “Circulating antibodies may be declining, but your immune system is capable of jumping into action once again.”

How can immune memory actually be getting stronger?

Research from immunologist Ali Ellebedy, at Washington University School of Medicine in St Louis, Missouri, helps to explain the vigour of the memory-B-cell response. His group took samples from the lymph nodes of vaccinated individuals and found tiny B-cell finishing schools — called germinal centres — that were churning out ever more potent immune cells as time went on⁶.

B cells in these structures randomly mutate their genes to create entire new sets of antibodies. Those cells that produce the best antibody repertoires eventually win out through an evolutionary process that augments the immune system's ability to fight off Delta and other SARS-CoV-2 variants of concern.

Ellebedy and his colleagues initially described the persistence of these germinal centres for 15 weeks post-immunization with an RNA-based jab — longer than anyone had ever seen before with older-technology vaccines for other ailments. Now, the researchers have unpublished data, following the germinal centres for up to six months. “The training camp is still going,” Ellebedy says. “It’s amazing.”

Should that immune memory give durable protection against severe disease?

For the most part, it should. But, says Theodora Hatzioannou, a virologist at the Rockefeller University in New York City, “if protection from disease relies at any level on circulating neutralizing antibodies” — and those molecules are clearly on the decline — “then, yes, the longer out you are from natural infection or from vaccination the worse you will be”.

Real-world data from diagnostic-testing records and hospital databases suggest that this might be the case. In Israel, for example, elderly people who got their shots at the beginning of the year seemed to have almost double the risk of severe illness during a July outbreak compared with similar individuals who were immunized more recently⁷. As researchers [reported this week](#), older individuals given a third dose of vaccine were less likely to become infected and much less likely to develop severe disease than those who had not received the boosters⁸.



Third-dose COVID-19 vaccine programmes have already begun in Israel, and data about their effectiveness are starting to come in. Credit: Kobi Wolf/Bloomberg via Getty

To Eran Segal, a computational biologist at the Weizmann Institute of Science in Rehovot, Israel, who advises the Israeli government on COVID-19 issues, the implications are clear-cut. “There’s compelling evidence that the third dose increases protection dramatically.”

But as University of Pennsylvania biostatistician Jeffrey Morris points out, inferences made from observational studies of this kind should be viewed with a critical eye. People going about their everyday lives are not clinical-trial participants. They are not randomized to account for behavioural and demographic differences. And although statistical modelling can help to correct for some of these variables, it’s impossible to account for all potential confounding factors.

“The bottom line,” Morris says, “is we need careful modelling and we need really thorough data to deconvolve all these things.”

What about other countries besides Israel?

Preliminary data from the United Kingdom and Qatar would seem to confirm the Israeli experience. Researchers at Public Health England [posted a preprint this week](#) detailing a modest but appreciable dip in vaccine effectiveness against hospital admission and death. This occurred about 20 weeks out from inoculation for recipients both of the mRNA vaccine from Pfizer–BioNTech and the viral vector vaccine from the AstraZeneca — although the effect was most pronounced for older individuals and those with underlying health conditions. Among the elderly, there was also some indication that spacing out the initial two doses of vaccine promoted more durable protective immunity.

Meanwhile, in Qatar, Laith Abu-Raddad and his colleagues described last month how the vaccine from Pfizer–BioNTech had provided consistently high protection against critical illness for up to six months post-immunization. Vaccine effectiveness against mild or symptom-free infections has declined gradually, as expected. But at the time that he posted a preprint online, on 27 August⁹, Abu-Raddad, an infectious-disease epidemiologist at Weill Cornell Medicine-Qatar in Doha, was unsure about the need for booster shots.



[What it will take to vaccinate the world against COVID-19](#)

Then, he says, he saw the data from seven months post-immunization. Although the results are preliminary, the vaccine's ability to ward off hospitalization and death seems to drop off. "The data now forced a change in thinking," says Abu-Raddad. (A comparable study from the United States has so far reported data only on waning protection against infection, not severe disease¹⁰.)

Still, globally, there is as yet no indication that the rates of severe illness among the vaccinated are spiking in any appreciable way. "The vaccines are really designed to prevent disease," says Julie McElrath, an infectious-disease specialist at the Fred Hutchinson Cancer Research Center in Seattle, Washington. "That is still holding up."

Even if most vaccinated individuals are not getting really sick, are there other benefits to boosters?

Reducing rates of infection should help break the cycle of viral transmission, which would ultimately result in fewer cases of severe COVID-19 and death. And according to Fyodor Kondrashov, an evolutionary geneticist at the Institute of Science and Technology Austria in Klosterneuburg, it should also help keep the emergence of vaccine-resistant variants at bay.

"Things that are good from the epidemiological perspective," he says, "are also good from the evolutionary perspective."

As Kondrashov's modelling work has shown, resistant viruses are most likely to emerge when transmission is not controlled¹¹. Getting more people vaccinated is the single most effective intervention to keep transmission rates low, but any bump in vaccine effectiveness can help as well.

So far, no human vaccine has been completely undermined by resistance in the way that many anti-infective drugs have, says Andrew Read, who studies the evolution of infectious diseases at Pennsylvania State University in University Park. "It's eroded their benefits," he says, but in ways that have been fixable with tweaks in vaccine design.

That's not to say it won't happen with COVID-19 vaccines. "We're on new territory here," says Read. The Delta variant took the world by surprise. Evolution — and our immunological responses — could have more surprises in store.



COVID boosters for wealthy nations spark outrage

But any discussion around the need for boosters cannot be had in a vacuum. In addition to considerations of immune kinetics among the immunized, there are also issues of vaccine equity and availability to factor in. And as long as vaccinated people are staying out of hospitals and morgues, then to Katrina Lythgoe, an evolutionary epidemiologist at the University of Oxford, UK, theoretical arguments around vaccine resistance are secondary. "In my view," she says, "apart from people who are particularly vulnerable, efforts should be directed to getting people, globally, vaccinated."

Nature **597**, 606-607 (2021)

doi: <https://doi.org/10.1038/d41586-021-02532-4>

References

1. 1.

Khoury, D. S. *et al.* *Nature Med.* **27**, 1205–1211 (2021).

2. 2.

Naaber, P. *et al.* *Lancet Reg. Health Eur.*
<https://doi.org/10.1016/j.lanepe.2021.100208> (2021).

3. 3.

Thomas, S. J. *et al.* *N. Engl. J. Med.*
<https://doi.org/10.1056/NEJMoa2110345> (2021).

4. 4.

Pegu, A. *et al.* *Science* **373**, 1372–1377 (2021).

5. 5.

Goel, R. R. *et al.* Preprint at bioRxiv
<https://doi.org/10.1101/2021.08.23.457229> (2021).

6. 6.

Turner, J. S. *et al.* *Nature* **595**, 421–425 (2021).

7. 7.

Goldberg, Y. *et al.* Preprint at medRxiv
<https://doi.org/10.1101/2021.08.24.21262423> (2021).

8. 8.

Bar-On, Y. M. *et al.* *N. Engl. J. Med.*
<https://doi.org/10.1056/NEJMoa2114255> (2021).

9. 9.

Chemaitelly, H. *et al.* Preprint at medRxiv
<https://doi.org/10.1101/2021.08.25.21262584> (2021).

10. 10.

Puranik, A. *et al.*, Preprint at medRxiv
<https://doi.org/10.1101/2021.09.04.21263115> (2021).

11. 11.

Rella, S. A., Kulikova, Y. A., Dermitzakis, E. T. & Kondrashov, F. A.
Sci. Rep. **11**, 15729 (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02532-4>

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

Starting up in science

Two researchers. Three years. One pandemic.

By Kerri Smith, Heidi Ledford and Richard Van Noorden
Photography by Chris Maddaloni

29 September 2021



After years of working in other scientists' laboratories, Alison Twelvetrees and Daniel Bose have finally landed positions as principal investigators and have grand plans to carve their own paths in science.

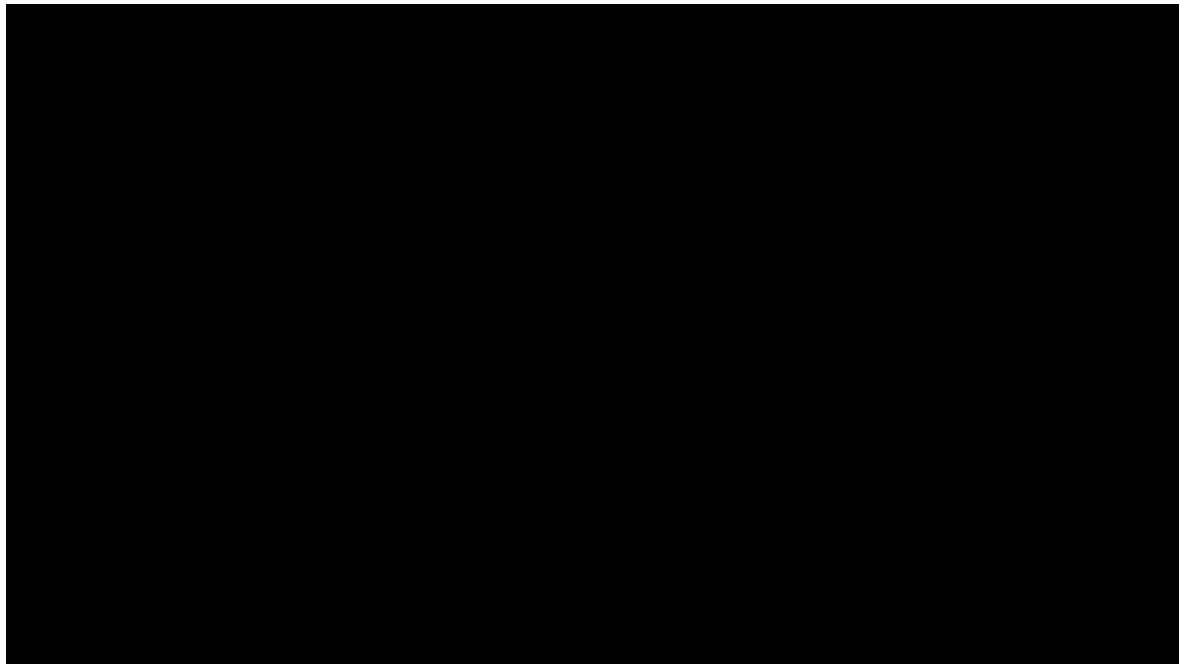


But Ali and Dan, who are married, won't keep their jobs if they can't win enough funding.



They didn't bank on a global pandemic – or another chance event that would turn their lives upside down.

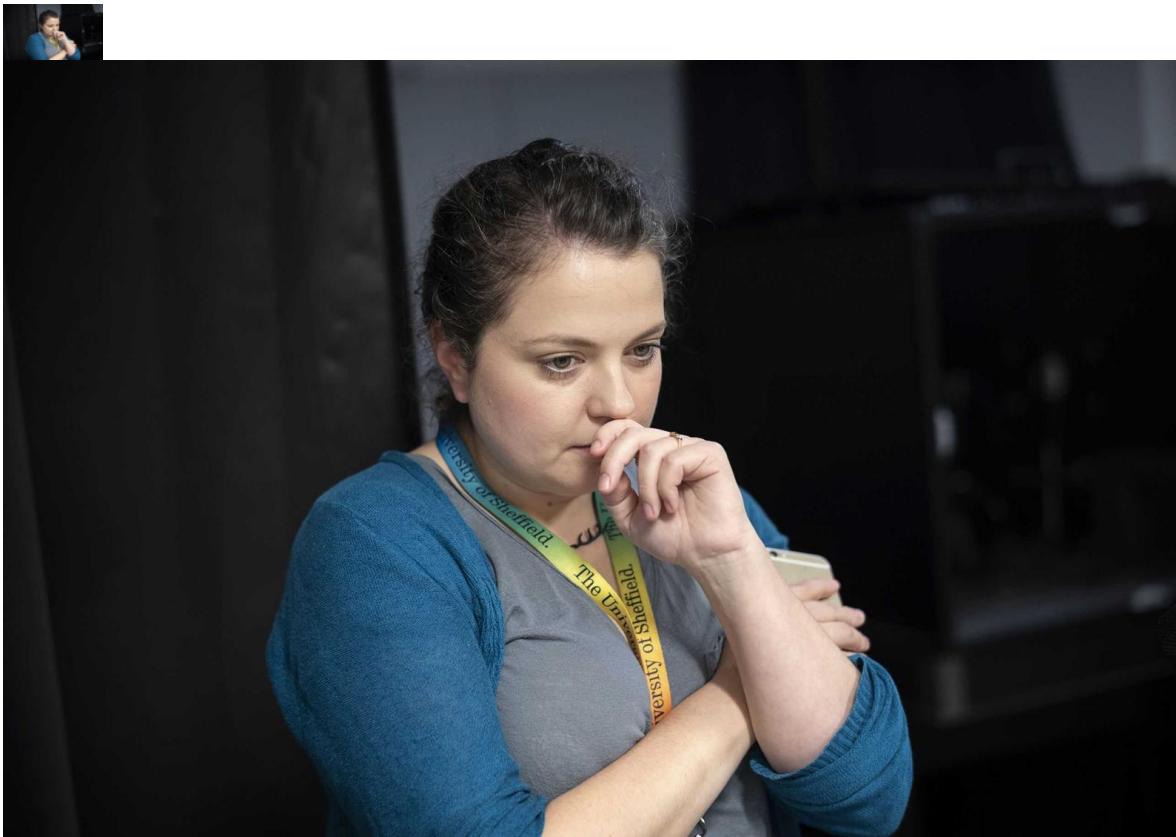




***Nature* followed Ali and Dan's journey over 3 years as they try to launch their careers.**



1. The pitch



On the morning of 7 June 2018, neuroscientist Alison Twelvetrees arrives at a slick five-storey building on the traffic-choked Euston Road in London. She has made sure she is wearing her lucky charm — a locket that belonged to her grandmother. Her bag is stuffed with crumpled neuroscience papers and notes that she's been carrying around like a comfort blanket for the past two weeks. She is here to make a pitch to win a £1.14-million (\$1.6-million) fellowship — money that she needs to sustain the fledgling research group she started the previous year in Sheffield, UK.

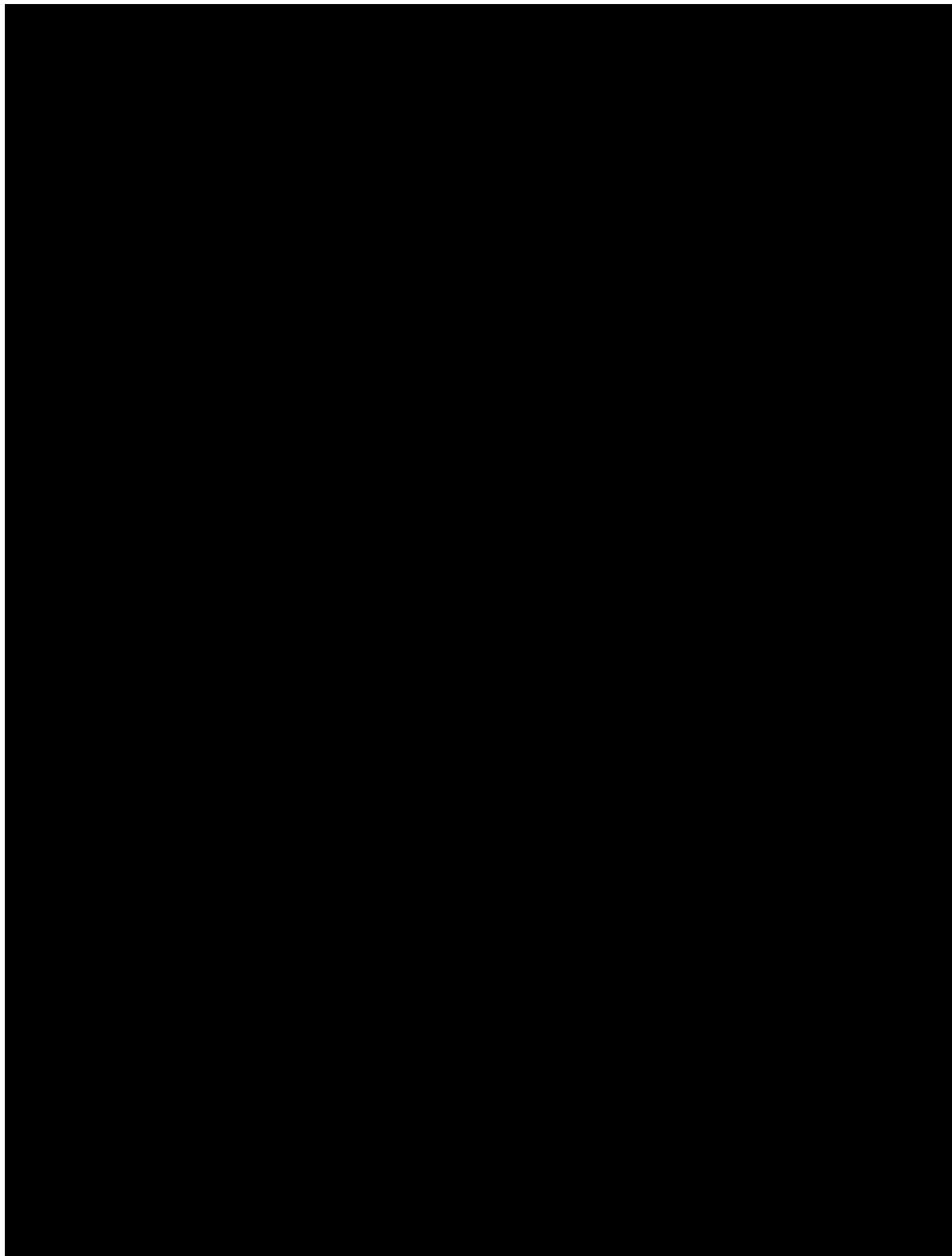
Outside the interview room, she watches the previous applicant emerge and tries to get a handle on her adrenaline. She has rehearsed for this moment intensively, fielding awkward questions in three practice interviews. The worst moment: when one panel member asked why she'd been so unproductive as a postdoc — a query so aggressive that it took Ali a while to gather herself sufficiently to reel off her scientific accomplishments and point out that she'd also had a baby.

Ali's now-three-year-old daughter, Ada, is back home in Sheffield with her husband, Daniel Bose. Dan knows what's at stake for his wife. Both of them

are new principal investigators (PIs) — independent scientists building up their own laboratories at the University of Sheffield. A year earlier, Dan had tried and failed to win a £1-million fellowship from the same scheme, funded by the UK charity Wellcome and the Royal Society. It is one of the most prestigious and lucrative grants available to PIs in the United Kingdom.

When Ali walks into the room at Wellcome headquarters, she finds a panel of 20 judges. The chairperson is so far away that she seems to exist in a different dimension. Ali has five minutes to take the panel through just two PowerPoint slides about her research into how cellular cargo is shuttled up and down neurons, and how these mechanisms break down in disease. She then has 25 minutes to answer the panel's questions. It will take the panel a few weeks to decide whether to award her the money she needs to get her lab off the ground.

The interview is simultaneously painfully slow and over too quickly. At a nearby pub afterwards, Ali sips a half-pint of beer, downcast. "I'm not crying, so there's that," she says. "The next three weeks is going to be stuff popping into my head: 'Why did I answer that? Why did he pull that face when I said that?'"





Firth Court at the University of Sheffield.

Firth Court at the University of Sheffield.

Every year, a few hundred scientists in the United Kingdom try to establish new labs from scratch; globally, thousands of researchers become heads of their own labs. From the outset, it's a chase for money and a time of intense pressure as scientists try to build research programmes while juggling teaching, fundraising, publishing and family life. Ali began her lab with just £15,000 in grants to cover equipment and experiments; Dan had £20,000. Both need to recruit PhD students, and Dan must also devise and deliver a programme of lectures.

They arrived at Sheffield in 2017 with three years in which to prove themselves, by attracting bigger, lab-sustaining grants and, they hope, publishing papers. If they can do that, they will pass their probation and their positions will become more stable, although the United Kingdom doesn't have a tenure system like that of the United States. If they fail, the university can terminate their contracts and they'll be out of a job. Dan and Ali agreed to allow a team of reporters from *Nature* to chronicle their lives as they

struggled to establish themselves in science (see endnote, 'How we found Ali and Dan').

2. "Okay, where do I start?"



When Dan first opened the door to his new lab in spring 2017, he was surprised to find that it was stuffed with someone else's junk. Half-full boxes of pipette tips and nitrile lab gloves littered the dusty shelves. In one corner sat a trolley with peeling, dirty white paint. On it stood an ancient computer monitor, a spaghetti tangle of defunct cables and several tired-looking machines.

Dan had one bench and one microcentrifuge. The rest of the space was assigned to people from another lab. He sprayed down his benchtop space, wiped off the dust and walked out again.

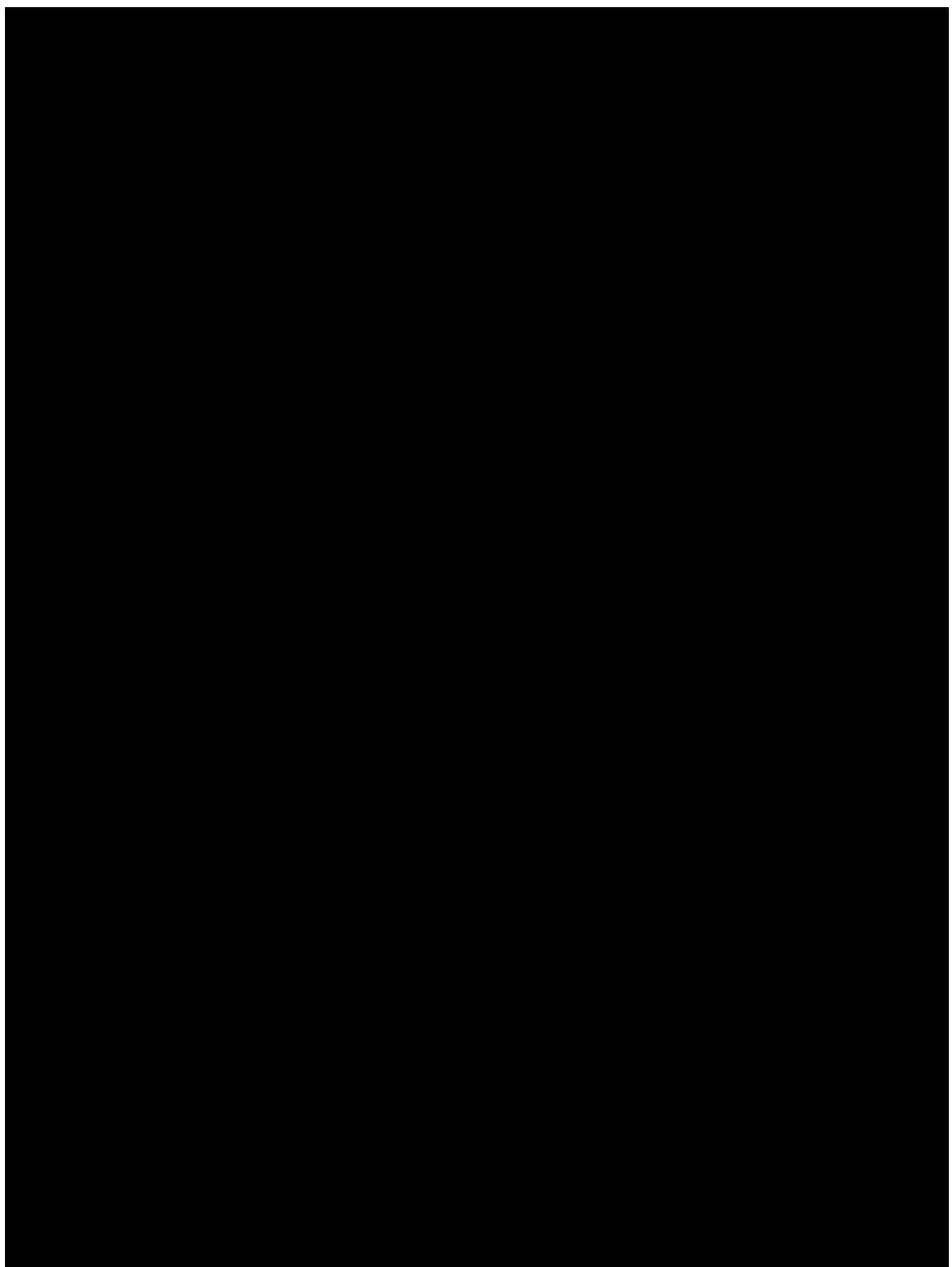
For the next two weeks, Dan's mind would race when he walked into the lab. "Okay, where do I start? What do I do? What do I need to get going?

How on earth do I afford what I need to get going?" he thought. Then he would retreat to the safety of his office down the hall to think about something else.

Hear Ali and Dan's story in our 4-part podcast miniseries. Via [Apple Podcasts](#) | Via [Google Podcasts](#) | Via [Spotify](#)

Gradually, things came together. Reagents he'd developed as a postdoc arrived from his former lab at the University of Pennsylvania in Philadelphia. And he learnt to scavenge. In corridors, meeting rooms and the backs of cupboards as well as on his new lab's shelves, he found an embarrassment of riches: barely used cartons of 96-well plates used to carry out reactions in small volumes, micropipette tips, forgotten tubes of enzymes.

Dan ordered bright orange tape to mark his territory in the shared space, a fridge freezer, a few test-tube racks and a partial set of micropipettes — his first purchases as a PI. But fiscal realities quickly set in. Dan wanted to study enhancer RNAs, which can control when and where genes are turned on. After years as a postdoc, dreaming of the experiments and projects he'd launch in his own lab, Dan realized he would have to rein in his ambitions. He didn't have the time or the money to race to the bench and get started. He needed to write grants, plan budgets, negotiate rates with vendors and recruit students.





Newly prepared buffers in Dan's lab at the University of Sheffield.

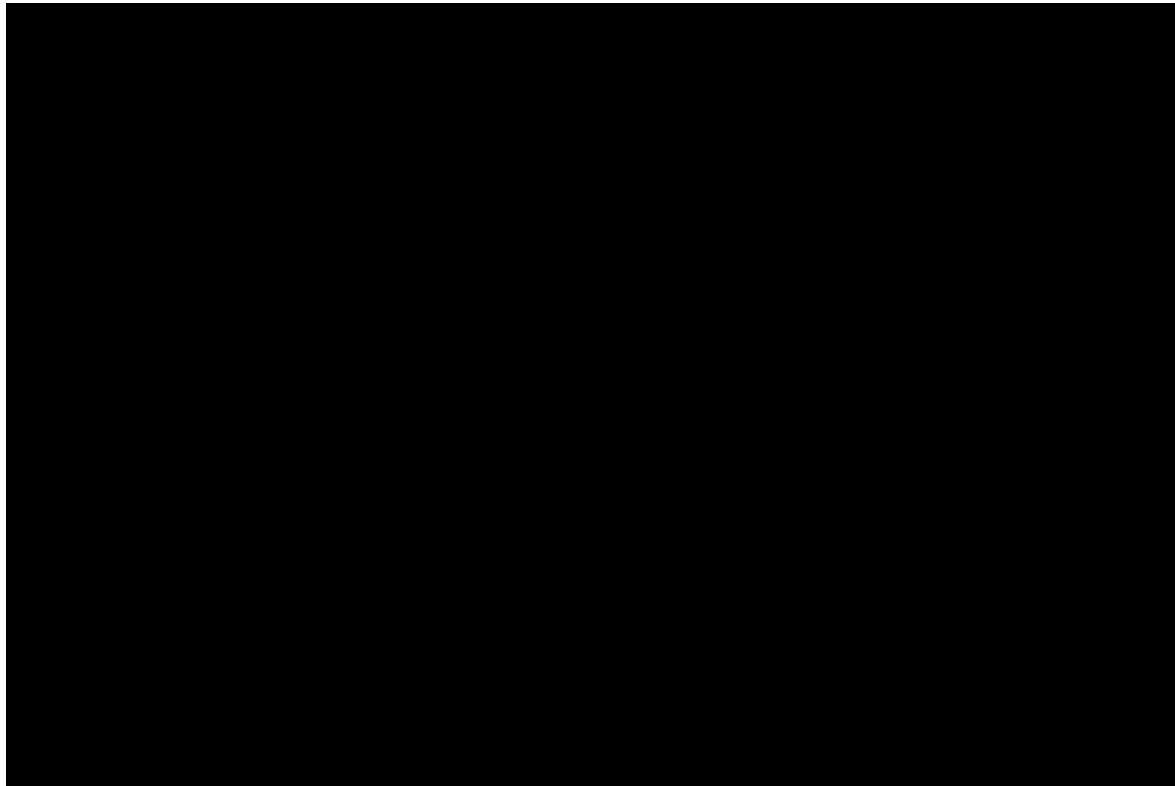
Newly prepared buffers in Dan's lab at the University of Sheffield.

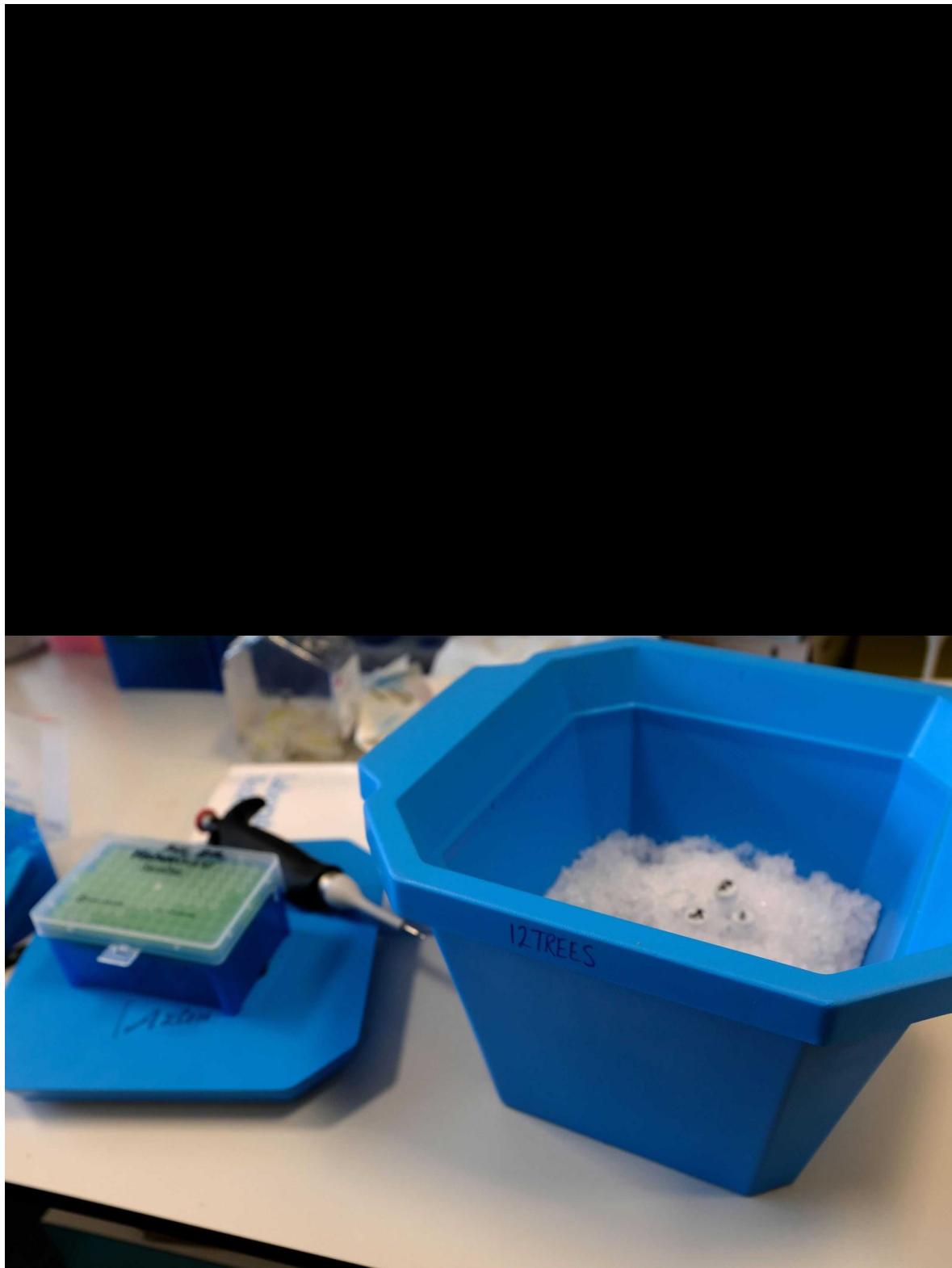
When he did make it to the bench, Dan would sometimes begin an experiment, only to realize that a key reagent was missing. Once a postdoc in an established and well-funded lab, Dan now found himself trying to stretch his scant resources. He borrowed basic molecular-biology staples such as agarose, a compound used to separate molecules on the basis of size, from people in neighbouring labs until he felt their patience wearing thin. When he did place an order, the price tag on some reagents made him shudder.

But at least he had space to play with. For a few weeks after her job started, Ali didn't have a bench — nowhere to actually do any lab work. "I turned up. And I sat in my office for a few weeks, and then at a meeting, I sort of raised... 'Can I have a bench, maybe?'"

Her new home, the Sheffield Institute for Translational Neuroscience (SITraN), had been established seven years earlier by drawing together

existing scientists from Sheffield. Ali was the first member of staff to join from outside the university, and most lab space was already allocated. Once she prodded, she got a small spot in a shared space.





An icebox in Ali's lab.

An icebox in Ali's lab.

Feeling that she needed more support from her peers, Ali set up a group on the messaging app Slack for new PIs in the United Kingdom, together with immunologist Sophie Acton, who had also just started her first lab at University College London. [Slack groups are a popular way for PIs to share experiences](#): one set up in 2016 by cell biologist Prachee Avasthi now has more than 3,000 members.

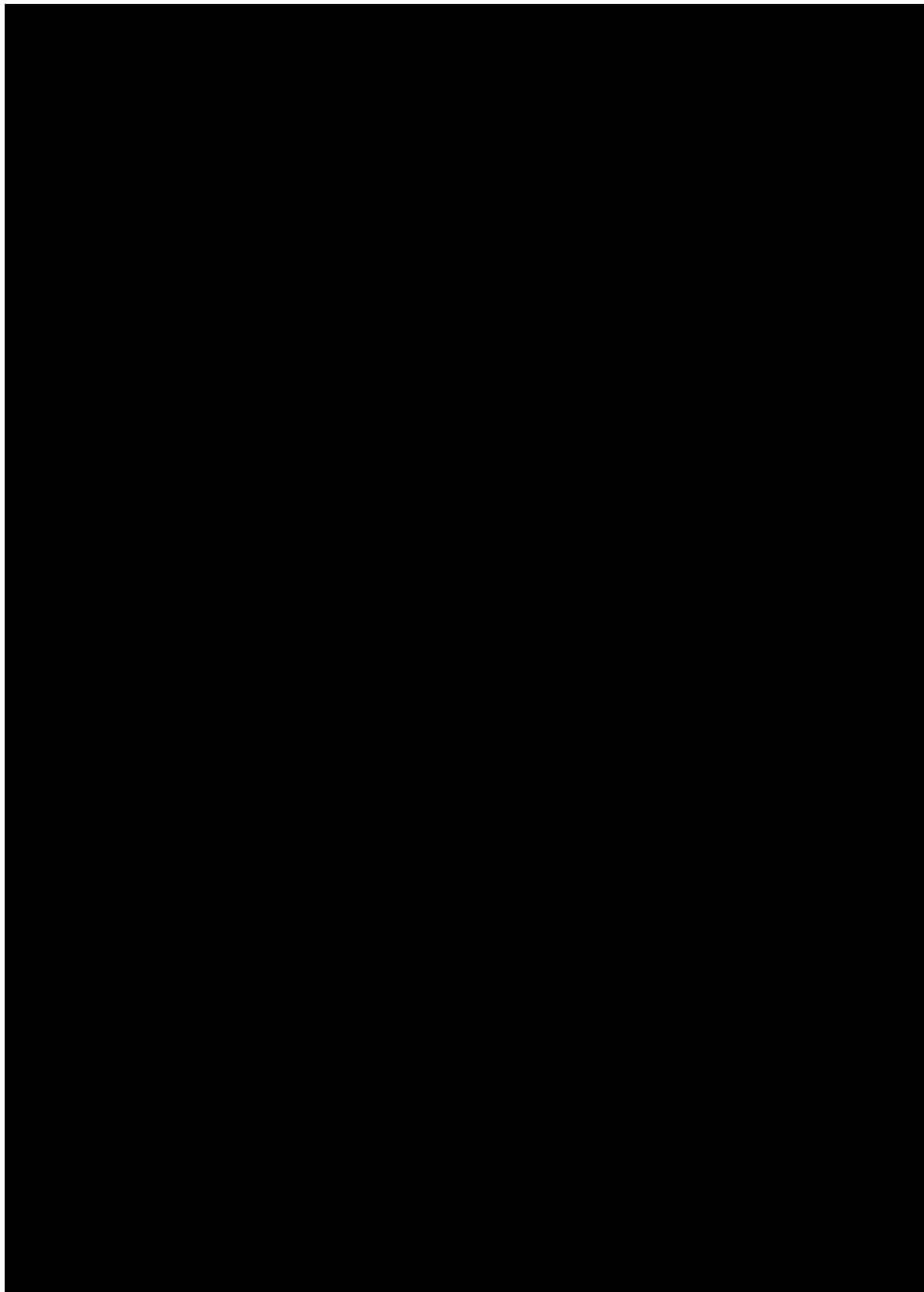
Despite the slow start, Ali began to make plans for her science and to collect some preliminary data. She started collaborating with a scientist from another department who had built a microscope she could use in her research. For a while, Dan and Ali had a game going to see how much they could stretch free samples from vendors. It was a far cry from his postdoc lab, Dan recalls, where all he had to do was wander over to tell the lab manager that he needed a reagent. “Suddenly, you’re hit with having a budget and learning purchasing,” he says. “I was spectacularly unprepared for that.”

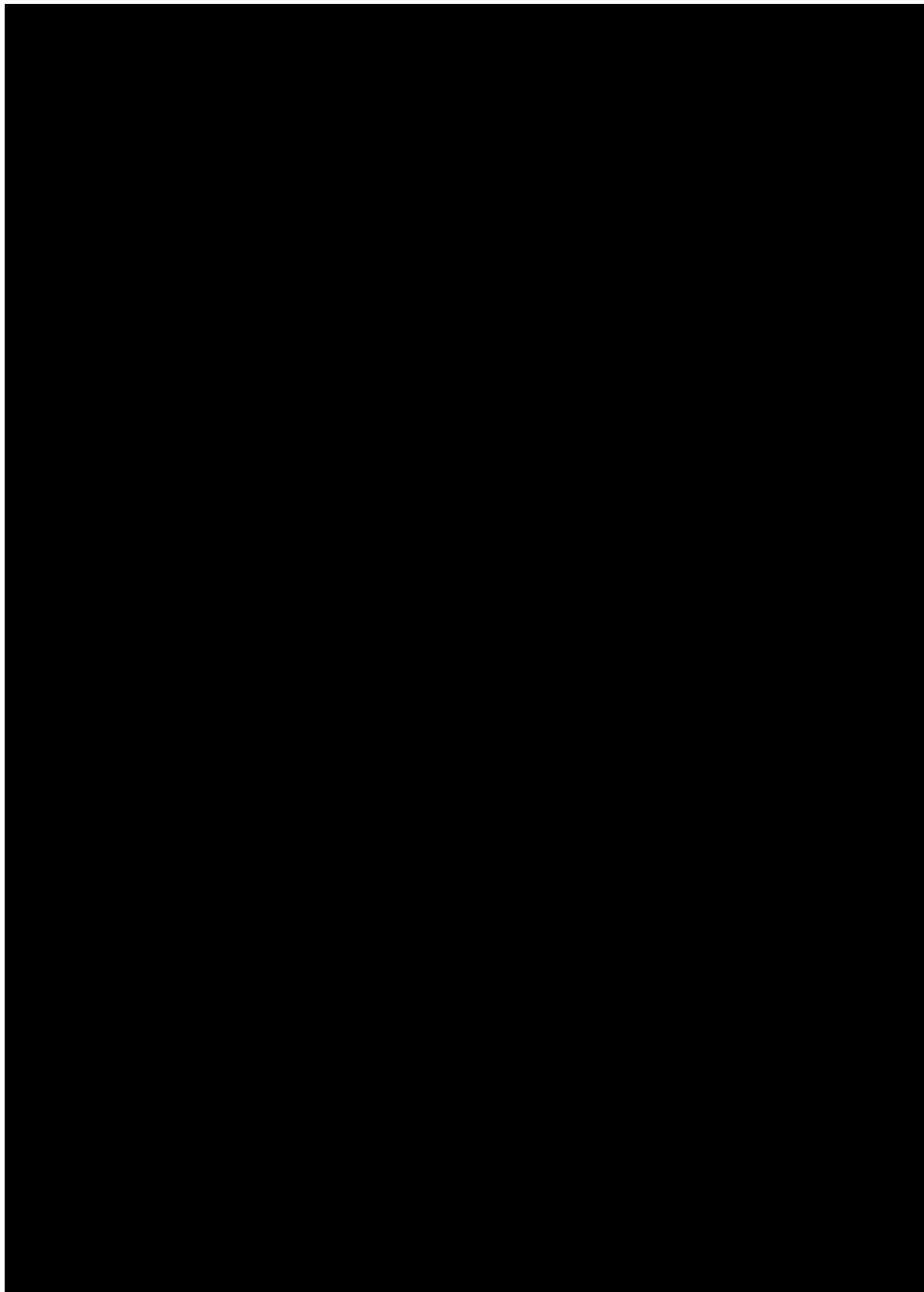
Dan was also unprepared for teaching. Unlike Ali, Dan’s position came with teaching responsibilities and, in the autumn of his first year at Sheffield, Dan walked into the first of four lectures he was slated to give for a class on RNA biology. He’d had almost no experience or training, and he was a little nervous, as he often is before public speaking events. The lecture started well as Dan ran through the basics of the CRISPR–Cas9 gene-editing technique. But when he delved into the enzyme mechanism underlying how this editing system works, 75 faces went blank. “You hear comedians talk about how you can tell when you lose an audience,” he says. “I didn’t expect to see this happening myself.”

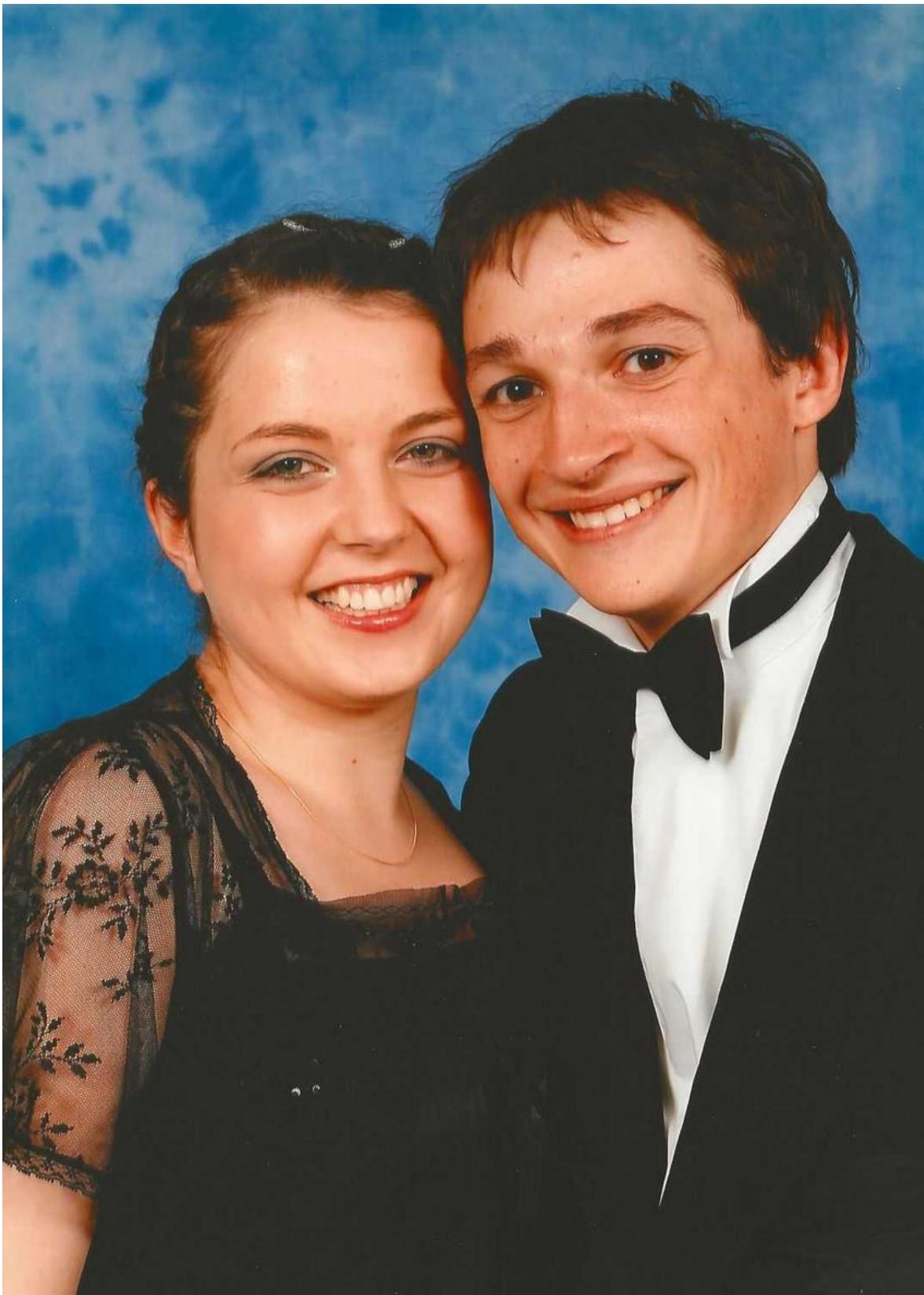
Dan felt a whiff of panic but ploughed on with his material, rather than going back over the topic — a choice he later regretted. Looking back on that day, he describes it as “a little bit soul destroying”, but he knows he’s not the only scientist who has had a rough first lecture: “It’s a rite of passage.”

3. “I finally decided I had nothing to lose”



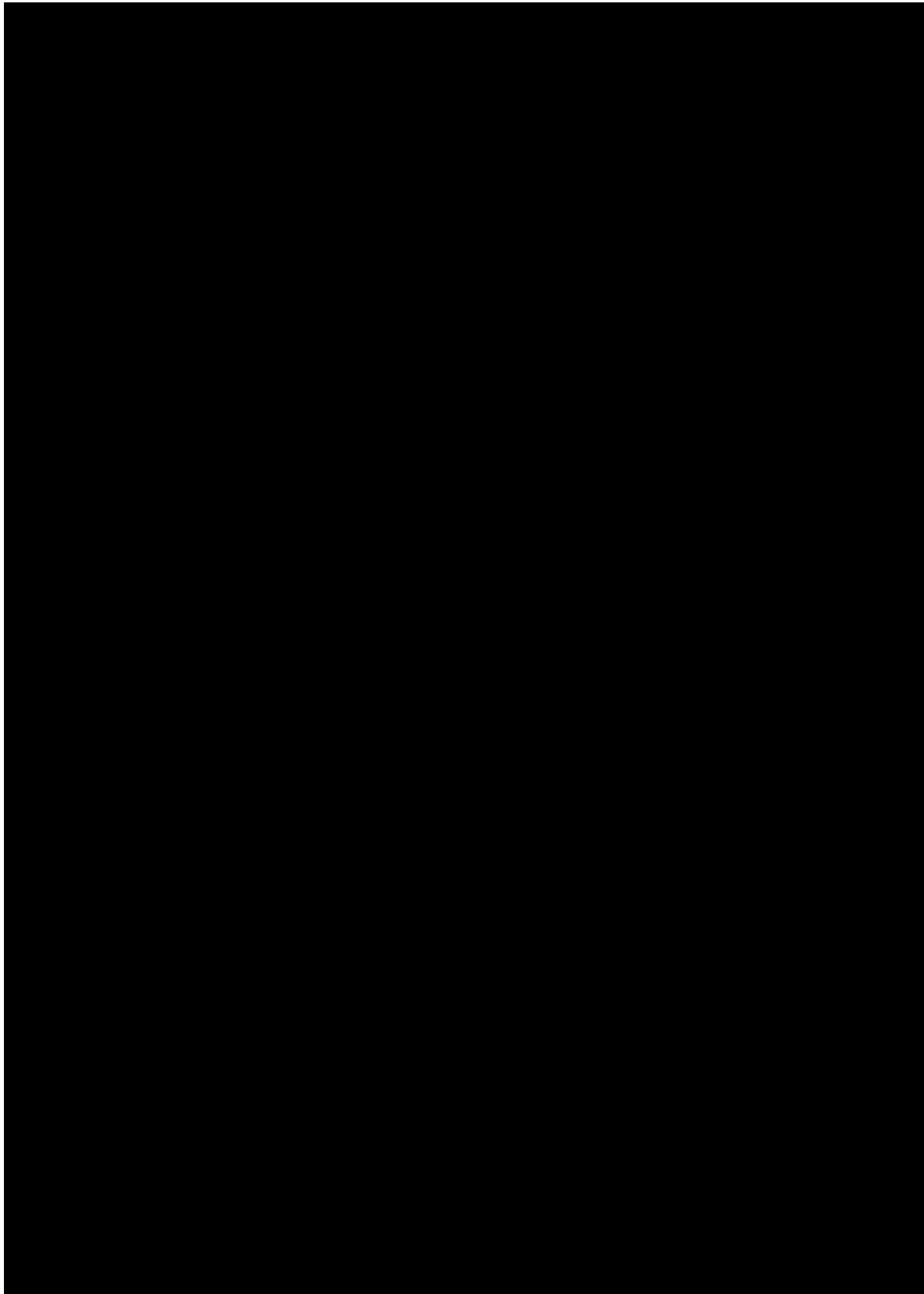


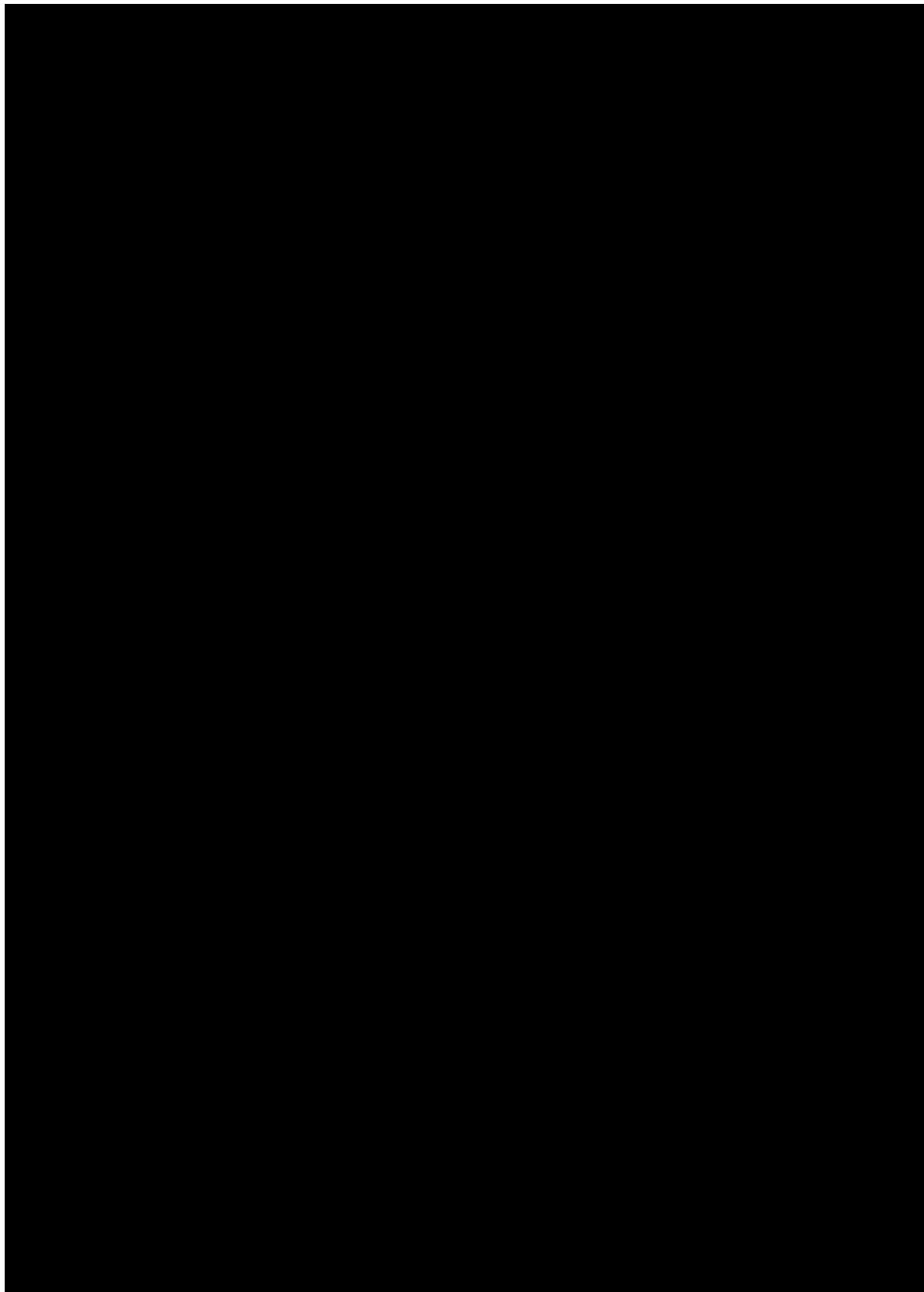




Ali and Dan in 2005.

Ali and Dan in 2005.

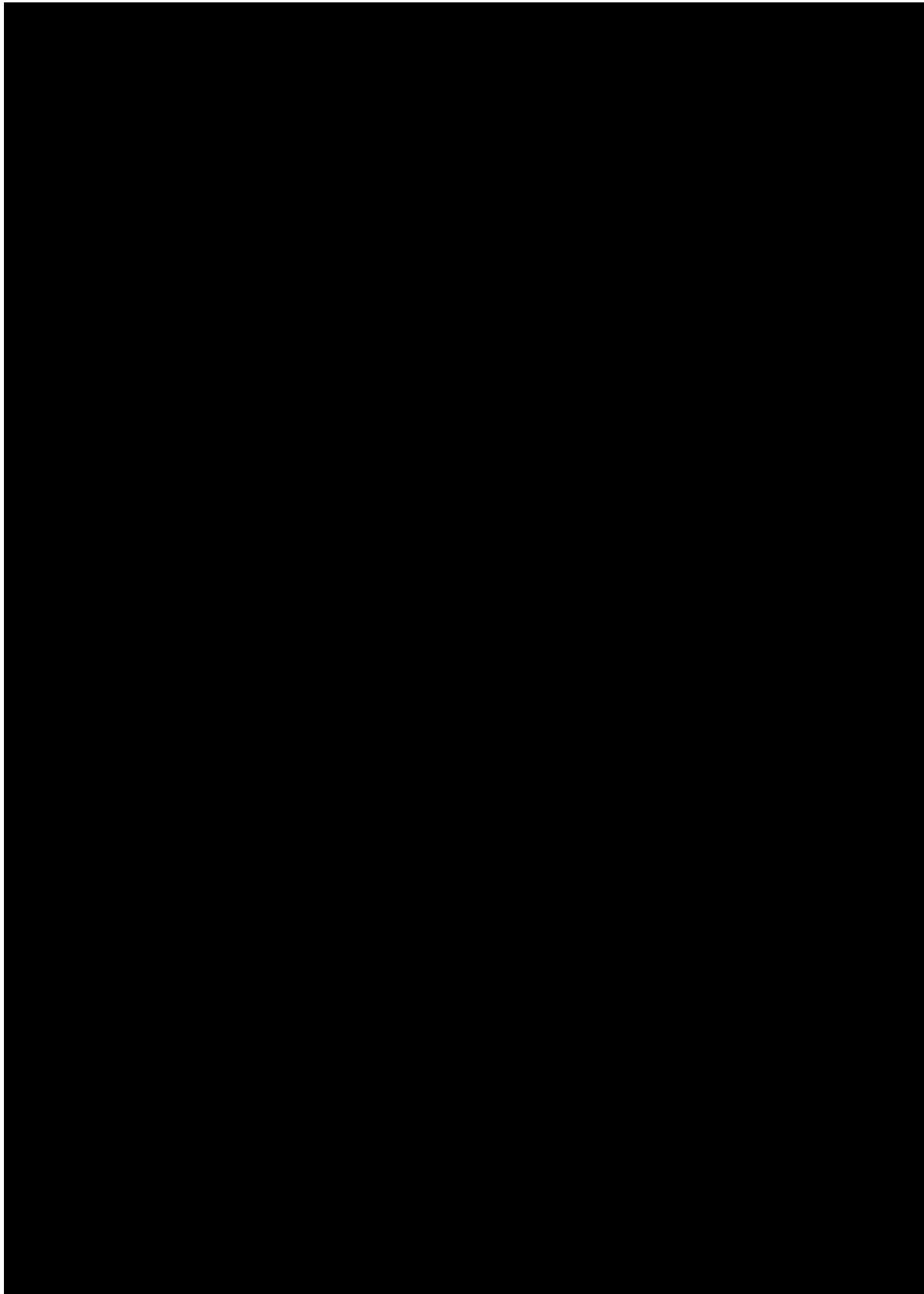






Dan in Philadelphia, Pennsylvania, in 2012.

Dan in Philadelphia, Pennsylvania, in 2012.

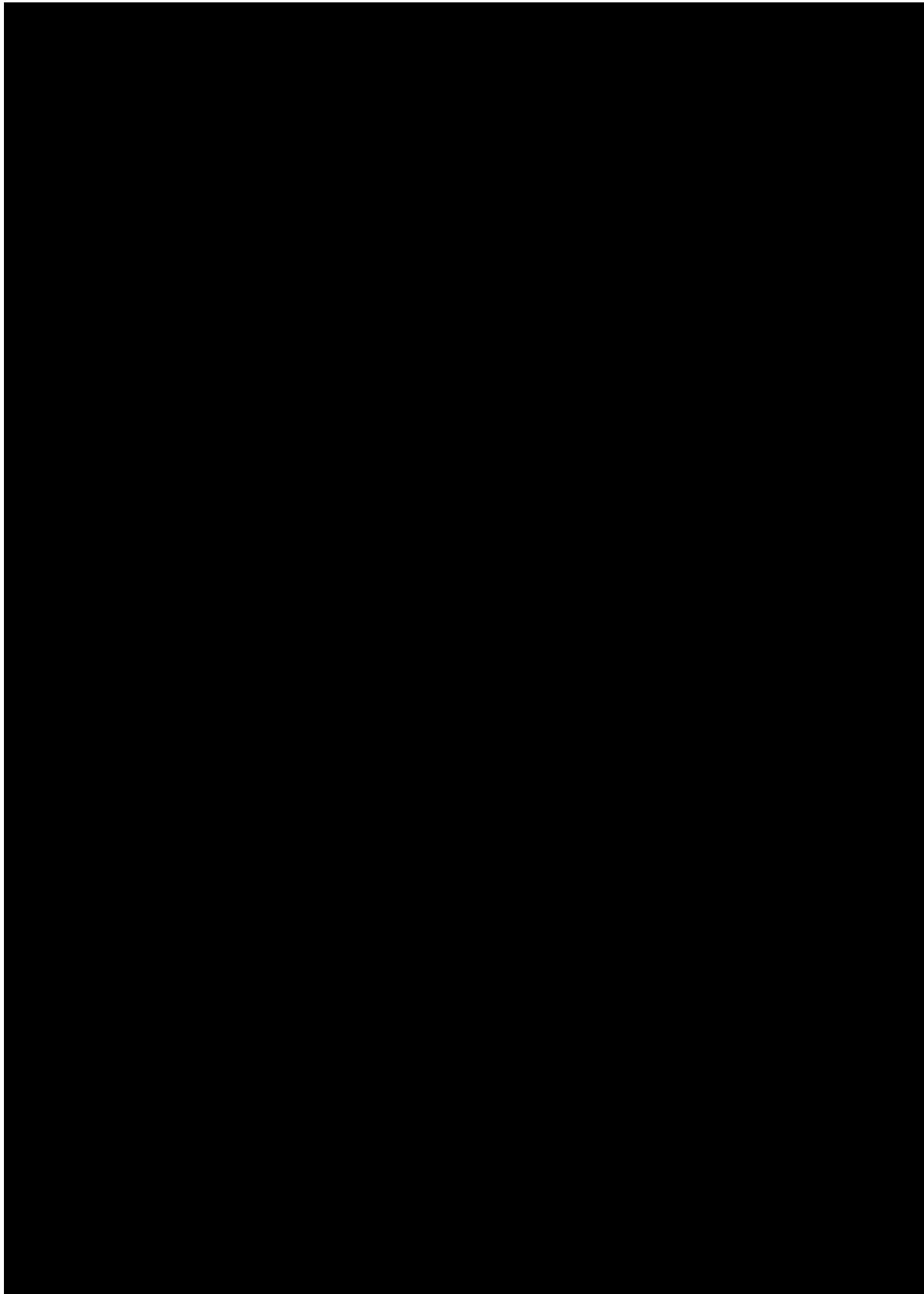


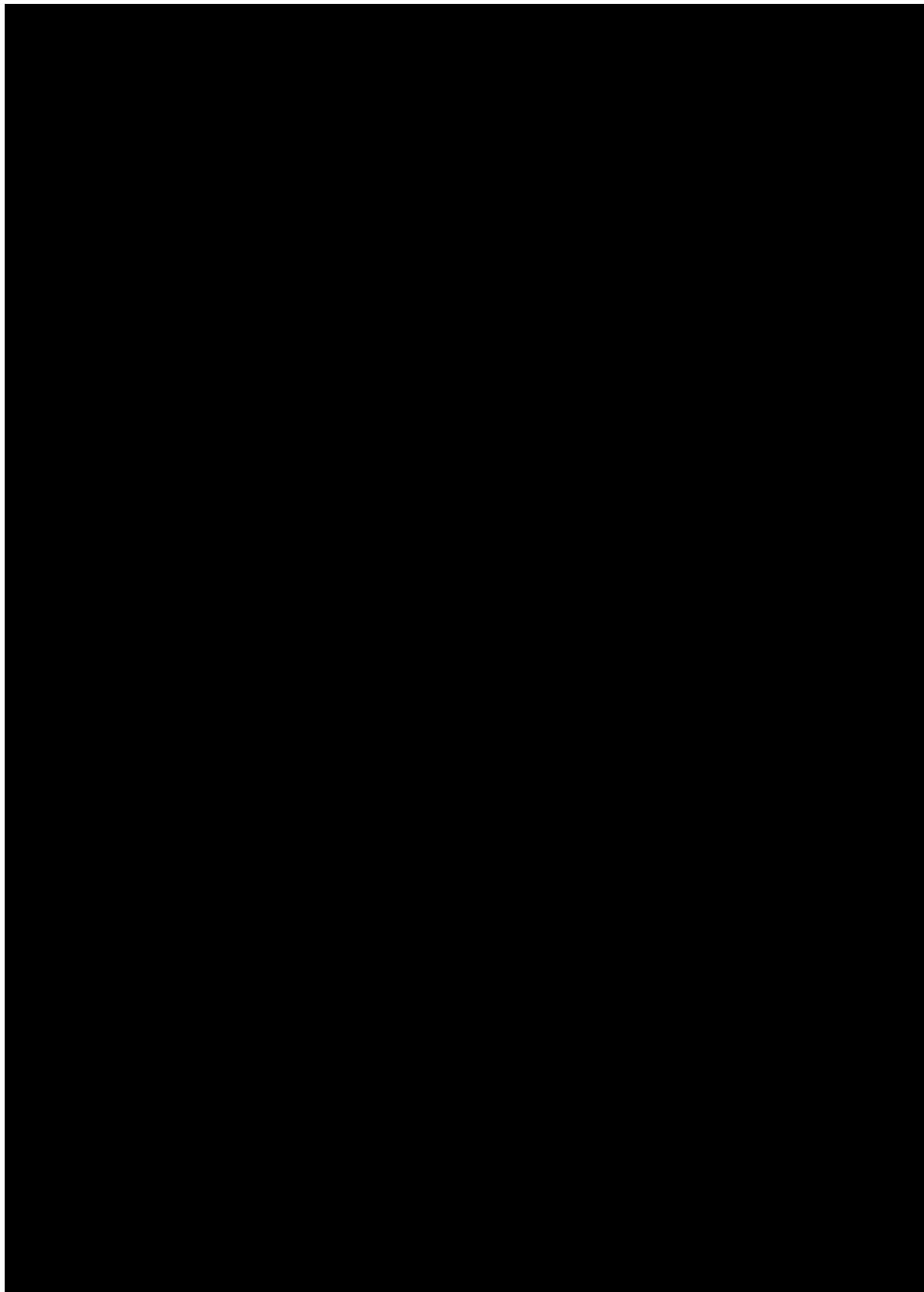




Ali and Ada in Philadelphia in 2015.

Ali and Ada in Philadelphia in 2015.

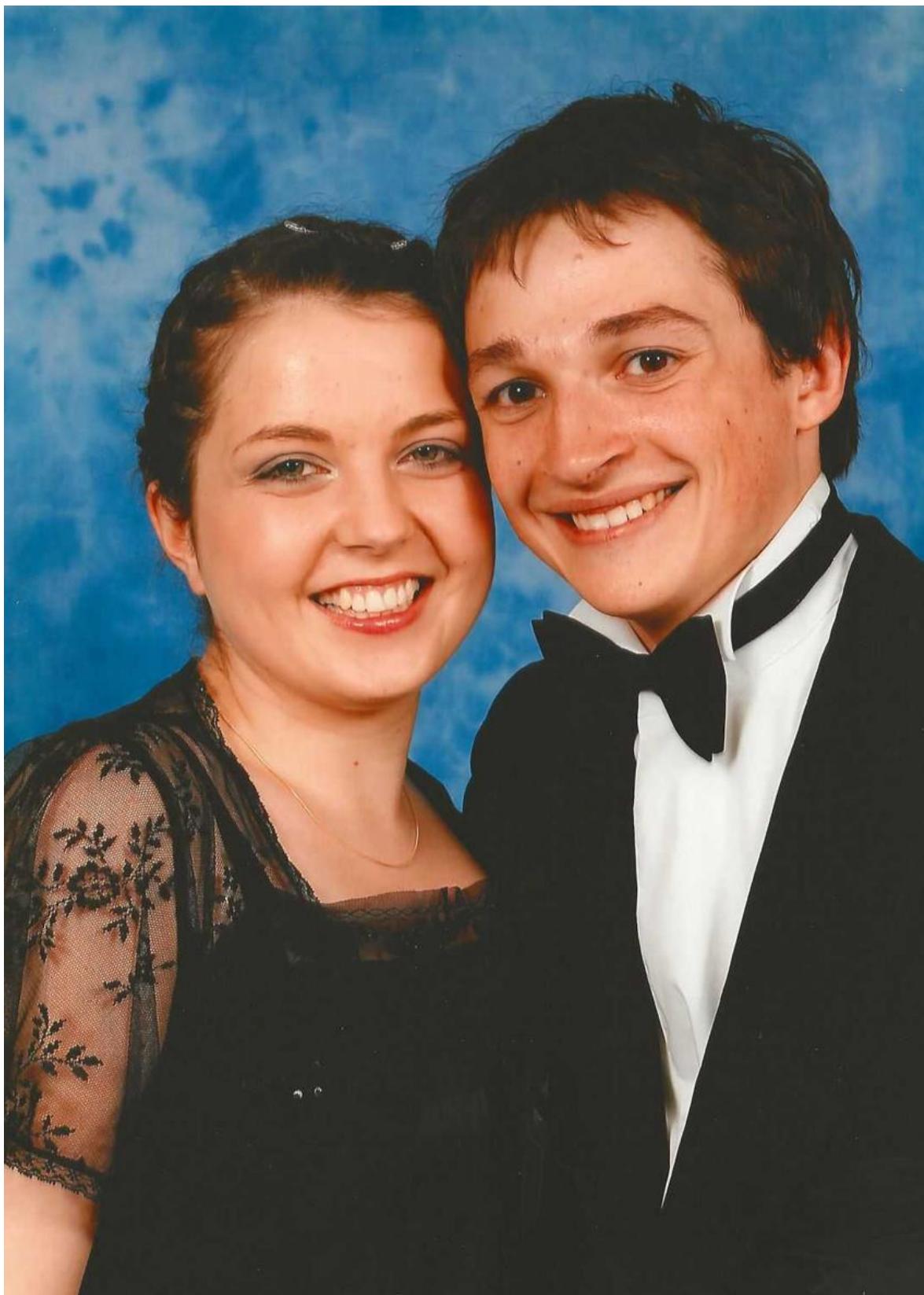






Dan with Ada on a family hike in the Peak District, near Sheffield.

Dan with Ada on a family hike in the Peak District, near Sheffield.



Ali and Dan in 2005.

Ali and Dan met as undergraduates at Imperial College London. They both studied biochemistry, but Dan sat at the back of the lecture theatres; Ali at the front. “I didn’t speak to Ali for three years,” says Dan. “I was kind of shy. She was kind of dating somebody.”

But at a summer ball towards the end of their third year in 2004, Ali persuaded a friend to ask Dan out for her. “I finally decided I had nothing to lose: ‘Well, I won’t see him again, so it doesn’t really matter,’” she says. “But I still wasn’t brave enough to ask him out in person.”

“You were braver than me,” says Dan.

A little over a year into the relationship, the two began what would become a long struggle with the ‘two-body problem’ — a difficulty encountered by many academic couples as they manoeuvre themselves through a tight job market in search of two jobs in the same location. They both got into graduate school in London, with Dan lingering there as a postdoc while Ali finished her PhD. Then Ali got a fellowship that allowed her to split her time between London and the University of Pennsylvania, where Dan found a second postdoc.



Dan in Philadelphia, Pennsylvania, in 2012.

They uprooted and moved to Pennsylvania in December 2011, just three weeks after getting married. A little more than three years later, their daughter Ada was born.

At the end of Ali's one-year maternity leave, the two-body problem loomed large. Ali had to return to London to finish her fellowship, but Dan was preparing to submit a paper — a crucial step in any young scientist's career — and he couldn't leave Pennsylvania just then.

In December 2015, the family packed up its belongings and the cat, and rushed to the airport to fly back to the United Kingdom for the holidays. At the start of the new year, Dan returned to the United States alone.

"It was a really tough time for everybody," he says.

Ultimately, Dan stayed in Philadelphia for another 11 months, while Ali juggled work and Ada alone in London. Dan did his best to parent by FaceTime. "Ada got very used to talking to an iPad," he says. "She would do things like carry the iPad around and tuck it up with her toys at bedtime."

Years later, Dan and Ali still carry the weight of that time apart. The separation seemed a necessary sacrifice to ensure that they both had a solid chance of long-term employment, says Ali. But it also felt like a risk to their family.



Ali and Ada in Philadelphia in 2015.

“It was a big gamble, or at least it felt that way,” she says. “It’s difficult. The ideal, perfect image of a mother... what kind of mother would separate their child from a loving father who wanted to be there for her?”

“Similarly, what kind of a father would leave his wife and daughter in another country and go and take a flight back to do science?” Dan says.

Dan has given talks to new postdocs about his academic career path, and likens the two-body problem to running a three-legged race. “You’re both trying to go in the same direction,” he says, “running with two legs tied together.”

After Dan’s offer from Sheffield had arrived, Ali faced a protracted wait to see whether her application was also successful. When it finally came through, Ali and Dan counted themselves lucky that they were able to find jobs at the same university. When they first moved to Sheffield in 2017, it was the first time that they’d all lived together as two working parents raising a child. “There was a lot of intense discussion about how it was all going to work because we hadn’t had to do it before,” says Ali. “It took a long time.”



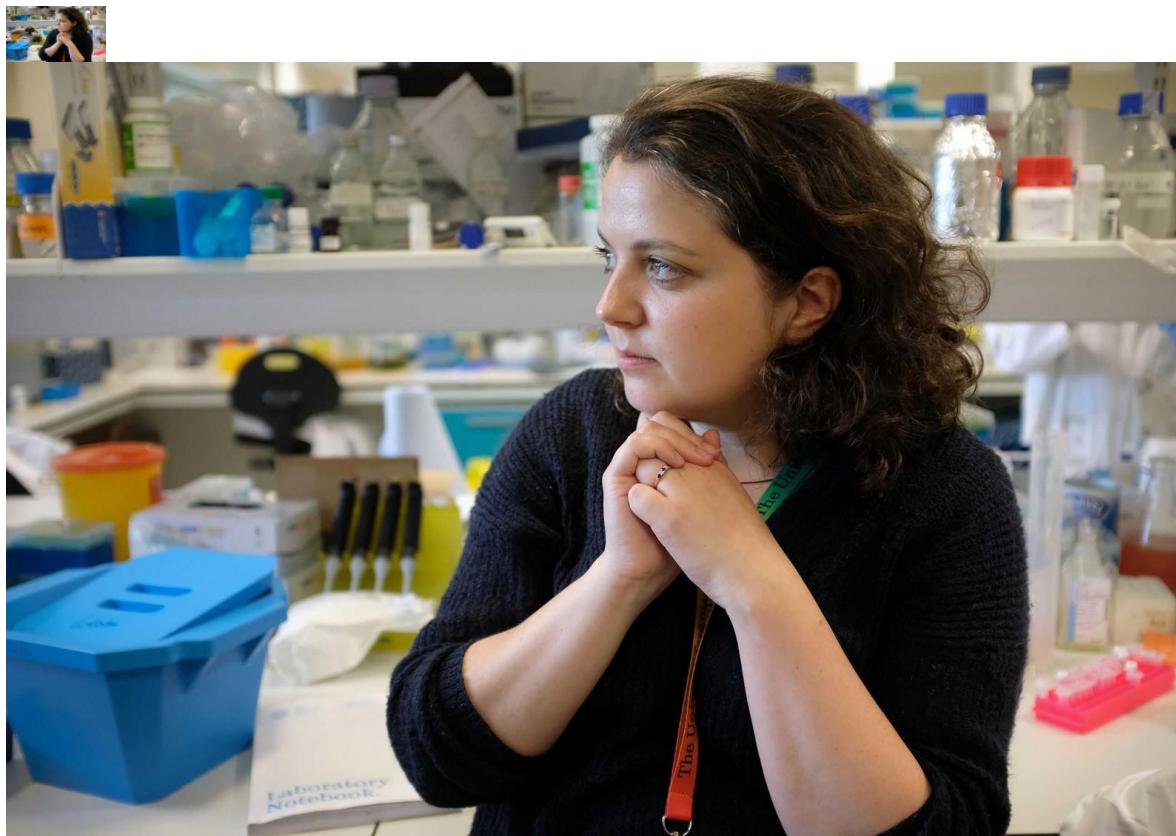
Dan with Ada on a family hike in the Peak District, near Sheffield.

But through trial and error, the family settled into a routine that worked. Ali and Dan alternated Ada's day-care pickup — one week on, one week off — to allow them to work late and tend to experiments.

"The fact that we managed to get here at all is somewhat surprising," says Dan. "There's definitely an element — when you're in the middle of paper reviews and neither of you have jobs sorted out — when you think maybe it's not going to happen for us."

Now they're in the same place, the challenge is to sustain their good fortune and use their stability to break new ground in science.

4. Nanometre steps

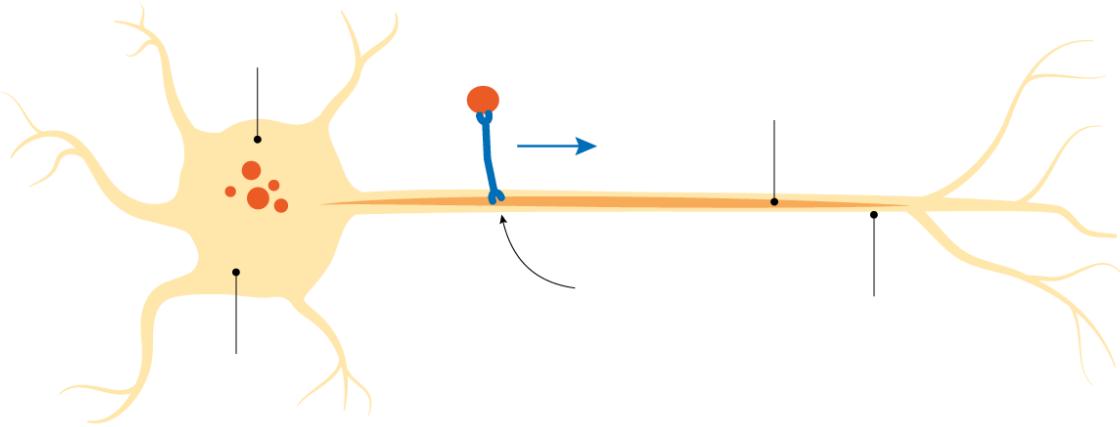


When Ali talks about the brain, her eyes widen and sparkle with child-like glee. "All cells are special, but I love neurons the most," she says. "And they have to function your whole life. It's crazy. It kind of gives you vertigo, thinking about all this stuff going on in your own cells."

In every one of these vertigo-inducing neurons, a fleet of motor proteins ensures the safe passage of cargo from the cell's main body down its axon, or nerve fibre, and back again. That cellular shuttle system is crucial because it supplies materials to the ends of the axon, where electrical signals jump from one neuron to another.

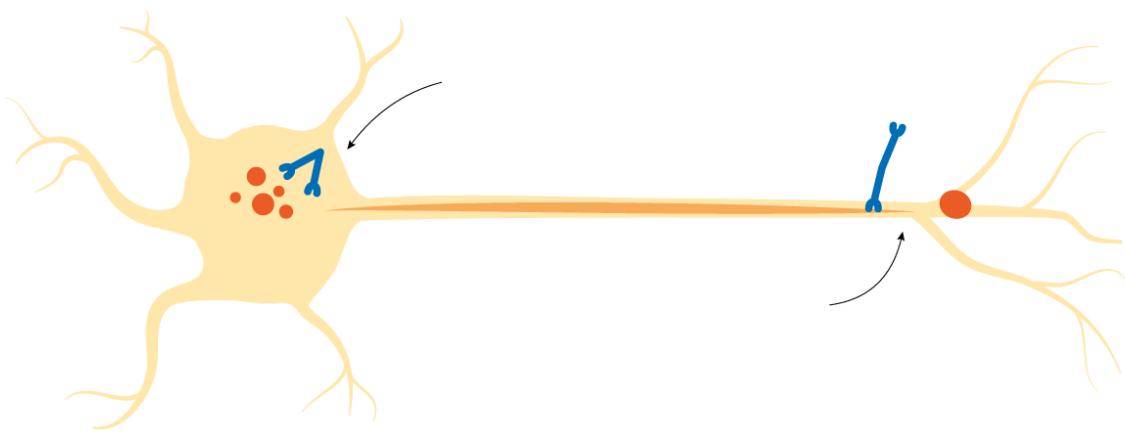
Inside neurons, the motor proteins travel along tracks called microtubules. For some long neurons in the spinal cord, the journey from end to end could take two weeks. When motor proteins fail to do their work, as they do in motor neuron disease, the effects on the network can be devastating.

Ali wants to know how these proteins load and lug their precious cargo. They can carry freight many times their size — large organelles such as mitochondria, or proteins contained in bubbles called vesicles. One class of motor protein called kinesins look like upside-down Ys, and their two 'legs' plod along the microtubules in eight-nanometre steps, shuttling newly made components to the ends of the axon.



ein
o
ubule
Kinesin
motor
protein
on
ll
ly

Ali knows that different cargos can travel to the ends of axons at very different rates. But the kinesin that does the hauling moves forwards at an almost constant speed. Ali wants to understand how kinesin activates — or loads its cargo — and how that sets the pace of the journey. If she can do that, she can take a stab at working out how the process breaks down in disease.



Activating

kinesin
Cargo

unloading

She can just about see these proteins at work in single neurons in a dish, but it's hard to observe them properly. Ali's dream microscope, which would allow her to track motor proteins as they carry cargo in live neurons, is a big part of her grant application and costs a quarter of a million pounds. "£250k

is a lot of money, but it's a lot of microscope," she says. "I always think of the money as science tokens — 'I can buy this much science'."

For now, she is having to settle for putting the proteins under a shared microscope and watching them float freely in fluid, spying on them as they open and close like the legs of synchronized swimmers.

Ali describes all of this to the judges in her pitch for the Sir Henry Dale Fellowship at Wellcome in June 2018, outlining how she plans to study kinesin and what she needs to make it happen. Afterwards, Ali recalls, there wasn't much time to take stock of how it had gone because Dan had a deadline a few days later. "I had to finish my interview, and then sort of drop everything to make sure that I was covering everything at home while Dan was finishing his grant application," she says.

A few weeks after her interview, Ali checks her phone after a lab meeting: "And literally the first thing I saw was this e-mail in my inbox. And I just knew, even though I hadn't seen the content of it, that it was negative."

She forces herself to read the e-mail, which is brief and confirms her fear: she hasn't got the fellowship. Then she calls Dan. "And then I ran away and had some lunch. And a big piece of chocolate brownie and a small cry and got back to work again."

Live cell imaging of microtubules (tagged with fluorescent molecules) in neuronal dendrites. This video has no sound.

Live cell imaging of microtubules (tagged with fluorescent molecules) in neuronal dendrites. This video has no sound.

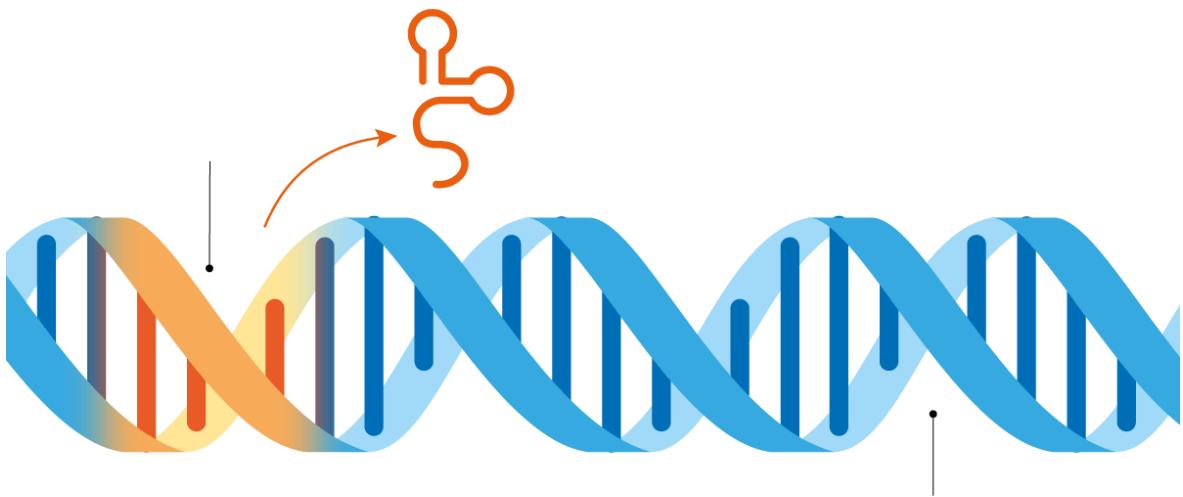
Ali is immediately pragmatic about what to do next. She knows she can apply to other funding bodies, and probably to the same fellowship programme, which she is most familiar with. She knows her weak spots. "Even though I can do all this stuff, I don't necessarily project impermeable confidence. When you're trying to essentially give a pitch and get people on board with your vision... you kind of have to be bullet-proof in the confidence department."

But it's hard to project that confidence when you know your job is at stake. Ali joined SITrN on a three-year, fixed-term contract and must pass her probation to convert it to an open-ended job. The main way to pass is to win a grant that can support the lab, and she fears she'll be let go if she doesn't get a cash infusion.

At lunch, Dan had suggested cocktails, but they'd decided to hold off until the evening. "He's probably more upset than I am," Ali says. "His phrase is that we win together and we lose together, and I think we get so tied up with wanting each other to succeed."

The couple had lost together before, in 2017, when Dan applied for a Sir Henry Dale Fellowship in a previous round and didn't get it. Dan is reworking his presentation for a second application this year, hoping that he can convince the judges to fund his work on a molecular mystery.

The enhancers that he studies reside in poorly understood regions of the genome that do not serve as templates for proteins, but instead encode snippets of RNA called enhancer RNAs.



ric

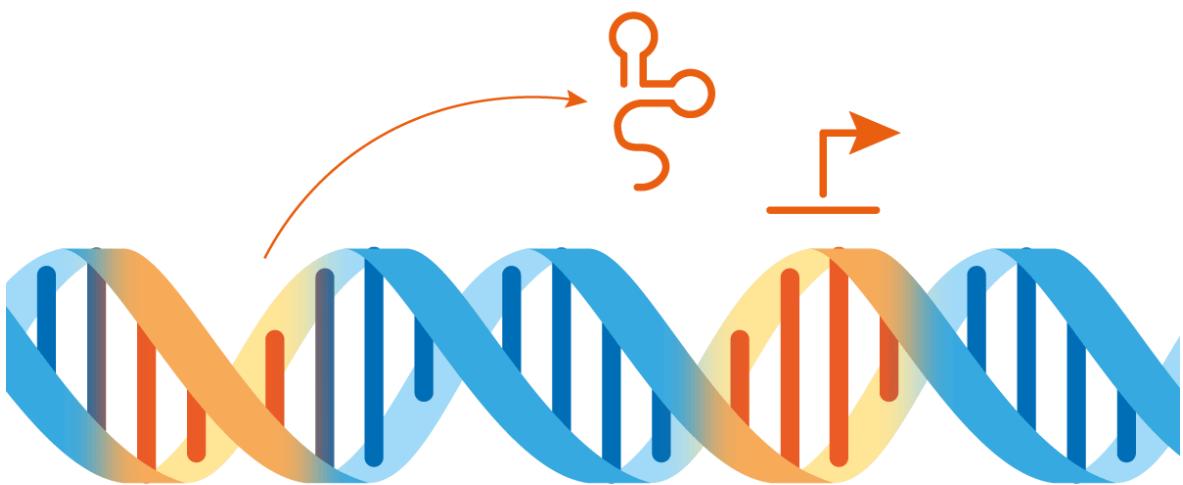
icer

n
Enhancer

RNA

ome

These are molecular switchboard operators, controlling whether genes are turned on or off at just the right time and place. When the switchboard goes awry, disease can follow. Errant enhancers have been linked to conditions such as cancer and neurodegenerative diseases. This means that these RNAs could provide an attractive drug target — but to design such drugs, researchers need to know how enhancer RNAs do their job.



l on

During his postdoc, Dan showed that some enhancer RNAs bind to a protein called CBP that turns genes on by loosening tightly packaged regions of DNA.



l on
RNA binds

to CBP

Now, he wants to use that as a jumping-off point to pick apart how those enhancer RNAs work. A key experiment will be to use Dan's favourite technique — cryo-electron microscopy (cryo-EM) — to see how one particular enhancer RNA changes the structure of CBP when it binds to it. He has been hooked on cryo-EM since graduate school. "I was incredibly taken with the fact that you could look down the microscope and see individual molecules of RNA polymerase in front of you," he says. "It's stuff like that that really stirs your excitement in science."

But to get to that point, there is a slew of molecular biology to be done. Dan and his team must first purify the CBP protein away from all other proteins and cellular gunk, then find the right conditions to coax it to bind to its enhancer RNAs outside the cell. All of that has to happen before Dan even touches the electron microscope and starts collecting data.

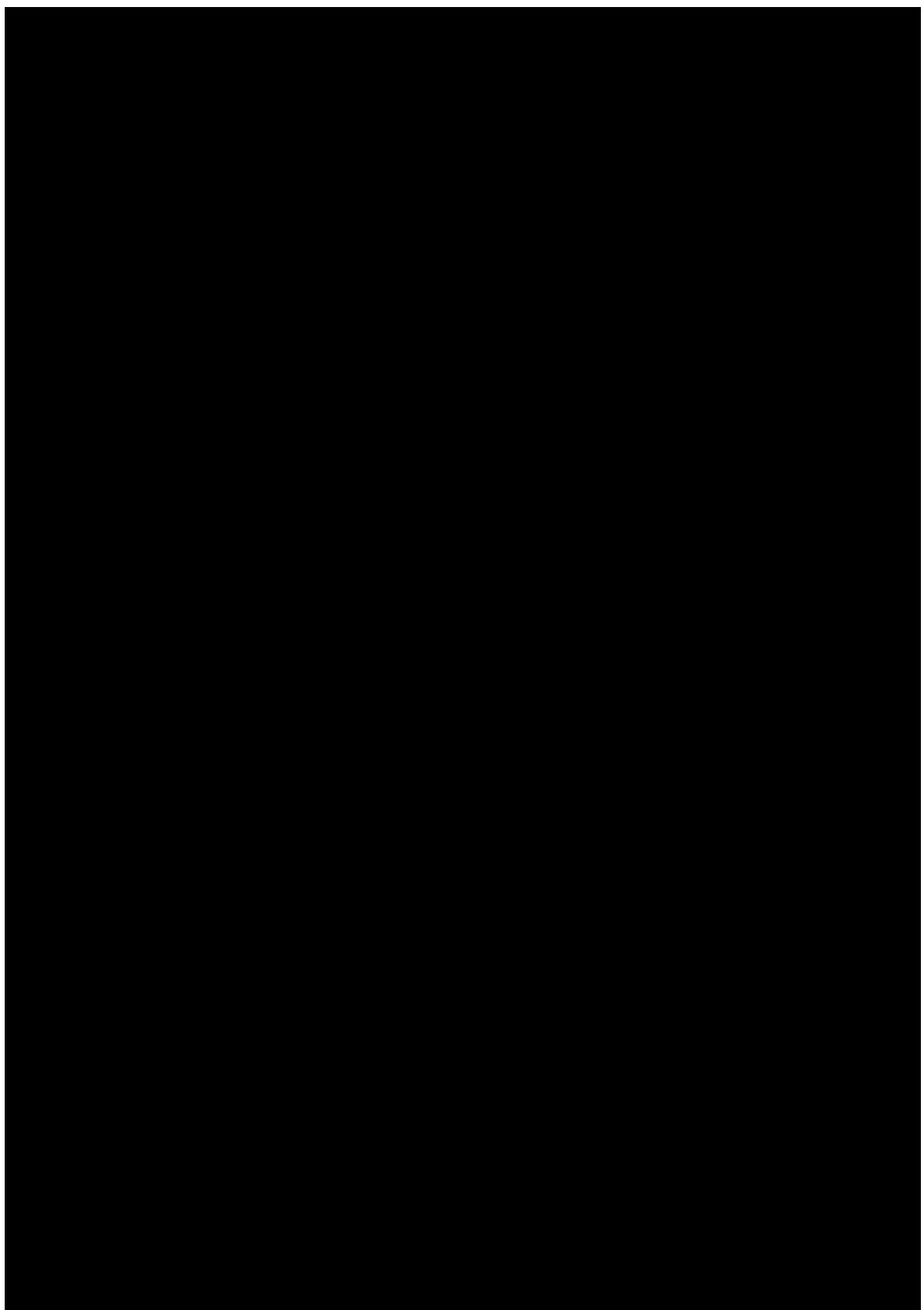
These are the crucial steps that molecular biologists sweat over — each represents an opportunity for a project to get hung up for months. The painstaking work to optimize lab conditions and troubleshoot failed experiments rarely makes it into a scientific publication, but it underlies every discovery.

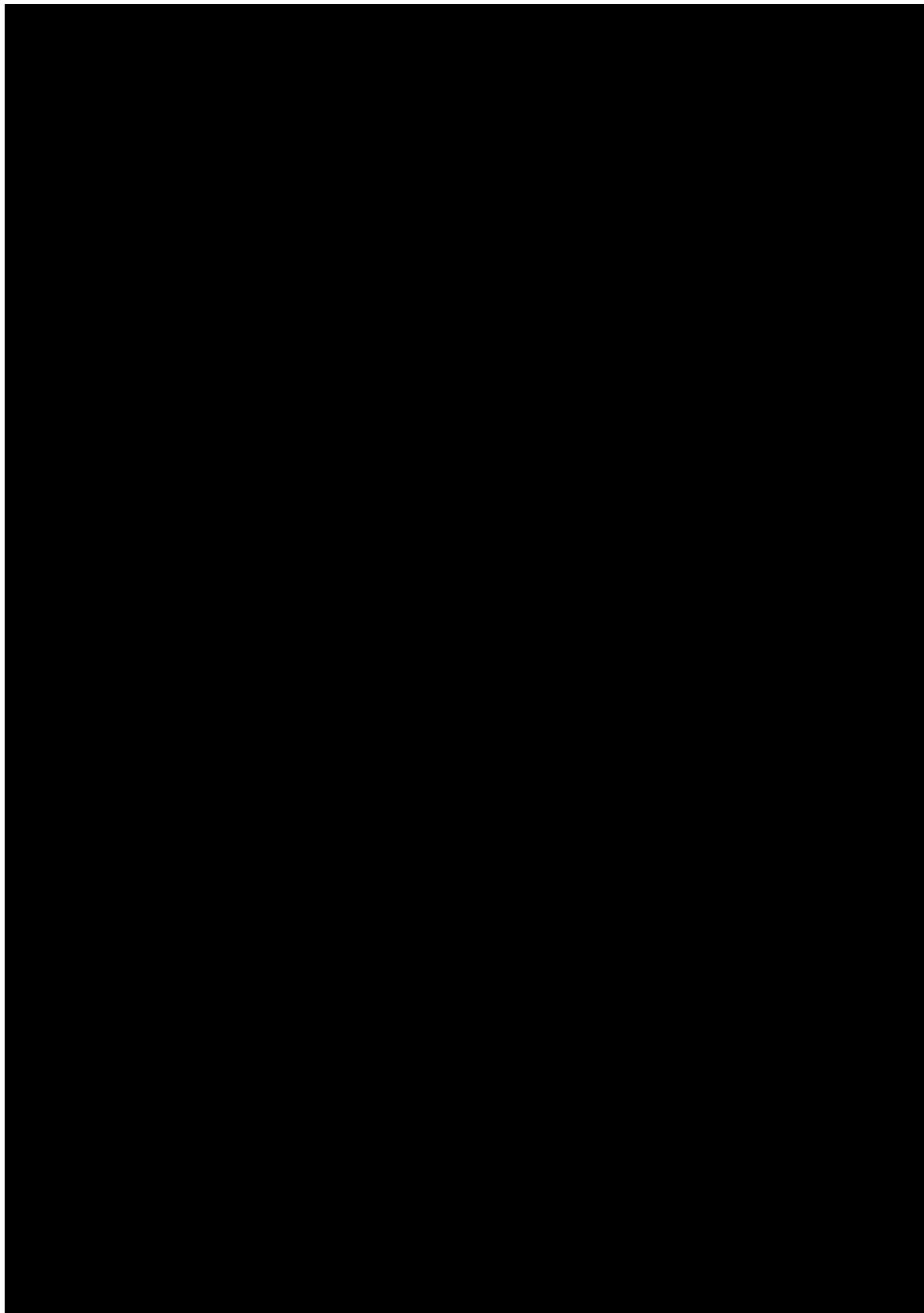
It is also not captivating conversation for a three-year-old.

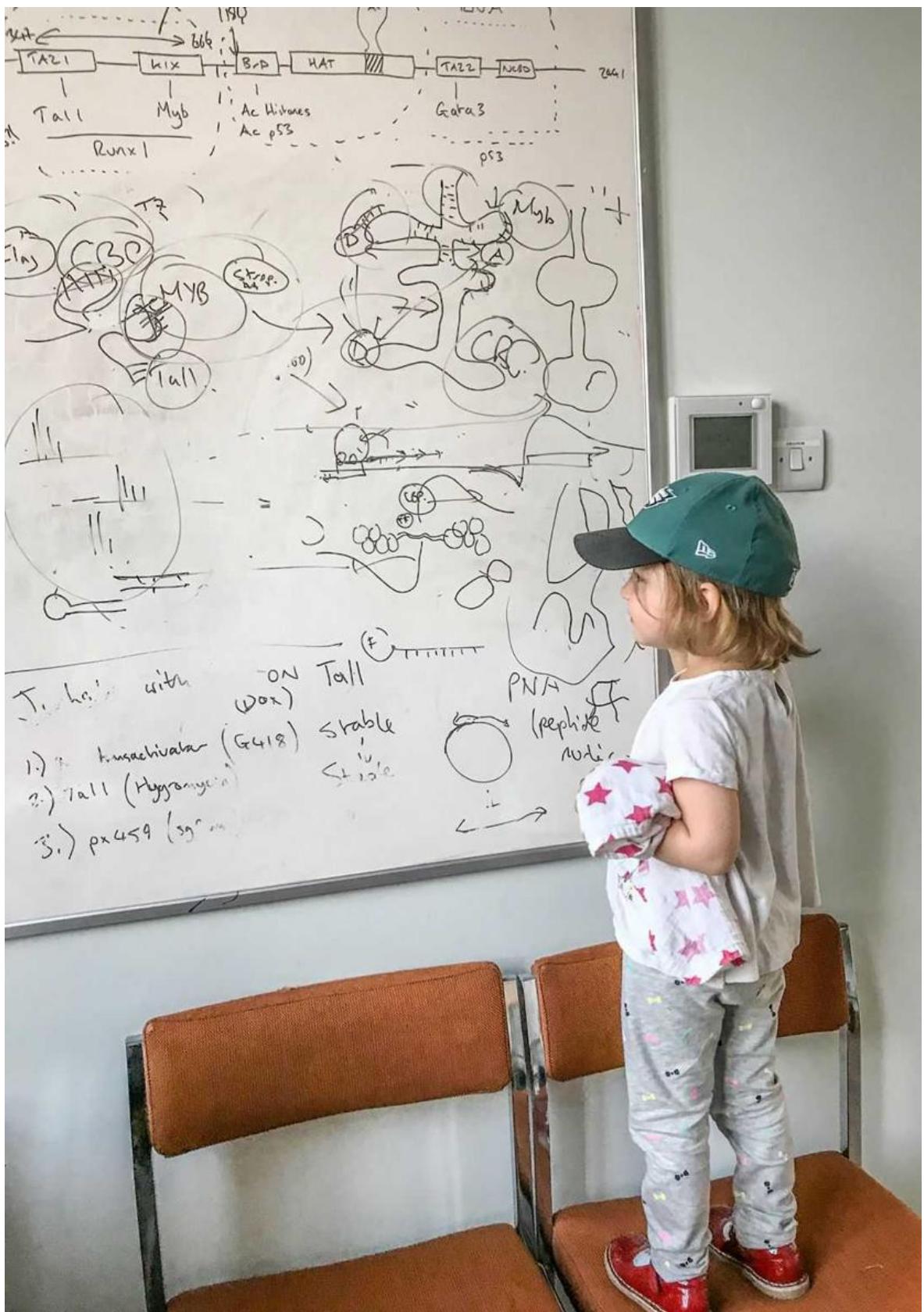
“Ada shouts at us now,” says Ali. “If she’s bored by the conversation, like if it’s too much work-based, she’s like, ‘What are you talking about? Stop talking.’”

“Which is a very important perspective,” says Dan.

“So we have an extra person to ensure that the work–life balance is enforced at the other end,” adds Ali.







Ada in front of Dan's office whiteboard.

Ada in front of Dan's office whiteboard.

It's not just for Ada's benefit that Ali and Dan avoid talking shop.

"Scientifically, we're pretty distant," says Dan. "We've tried quite hard to maintain that distance so we don't end up talking about science all the time at home."

Both rely on high-resolution microscopy to peer into the molecular world. But that is where the similarity of their work ends. They both claim to lack the deep background to fully grasp what the other is doing. "The funny thing is, although I'm very familiar with Dan's work, I have no context for it," says Ali. "I keep having to ask what enhancers are."

"Yeah, which probably means I'm not giving you a very good answer," Dan says.

"What are enhancers again?" Ali laughs, half-serious.

"Outside the realm of the basic techniques, there's a lot of stuff I really don't understand about what she does," says Dan.

They might not sell their science well to each other, but both must think hard about how best to market their work to the grant panel. Around 200 new PIs around the country submit a full application to the Sir Henry Dale scheme each year, and only about 40 will be offered a fellowship. For Dan and Ali, their best shot at keeping their fledgling labs open is to win two of those places.

Next: Good news and bad news

[Read part 2](#) of *Starting up in science*.

Nature **597**, 608-613 (2021)

doi: <https://doi.org/10.1038/d41586-021-02564-w>

How we found Ali and Dan: *Nature* journalists asked staff members across the journal for examples of scientists who were going through big changes, such as starting up new labs or institutions, moving a lab to another country or retiring. We wanted to know — in real time — what these experiences are like. We contacted several scientists to get a sense of their stories and, crucially, whether they would be open to multiple visits over years from our reporting team, based in London. This profile of Ali and Dan does not represent the journey of every new principal investigator, but many researchers will recognize aspects in their own path to independence — including moves to other countries to improve skills, having to attract funding to retain jobs, and the ‘two-body problem’, in which scientist couples seek jobs in the same location.

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

Starting up in science

Nature follows two researchers over three years as they struggle to launch their own labs.



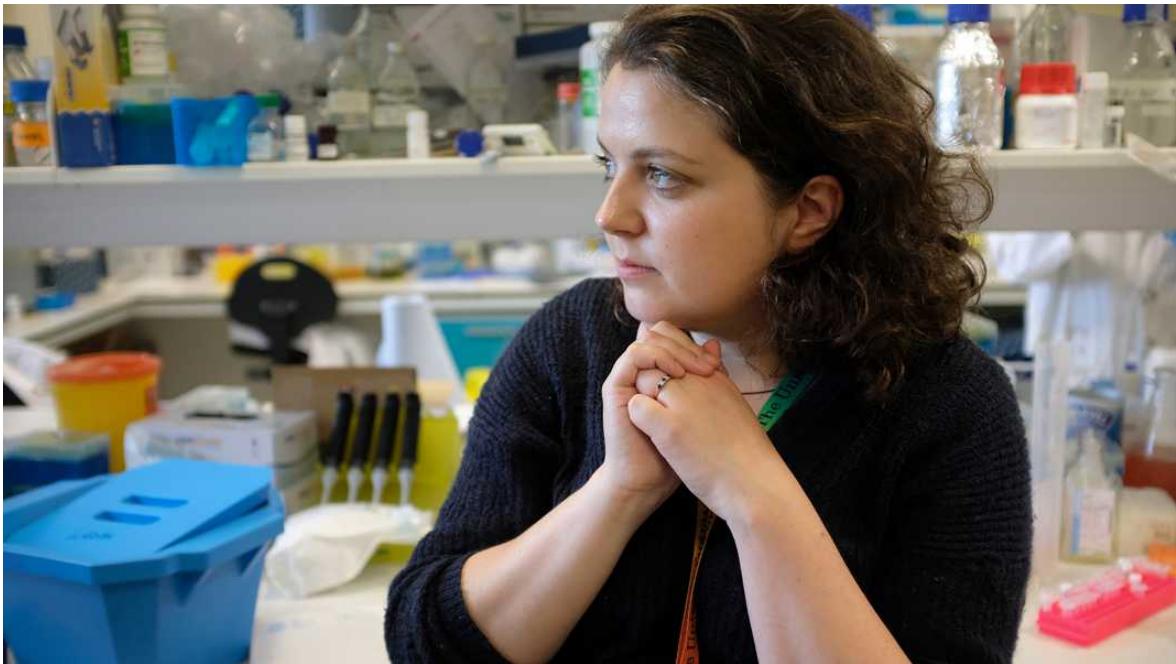
Part 2: The quest for cash

Ali and Dan race against time to secure stable funding to keep their labs going.



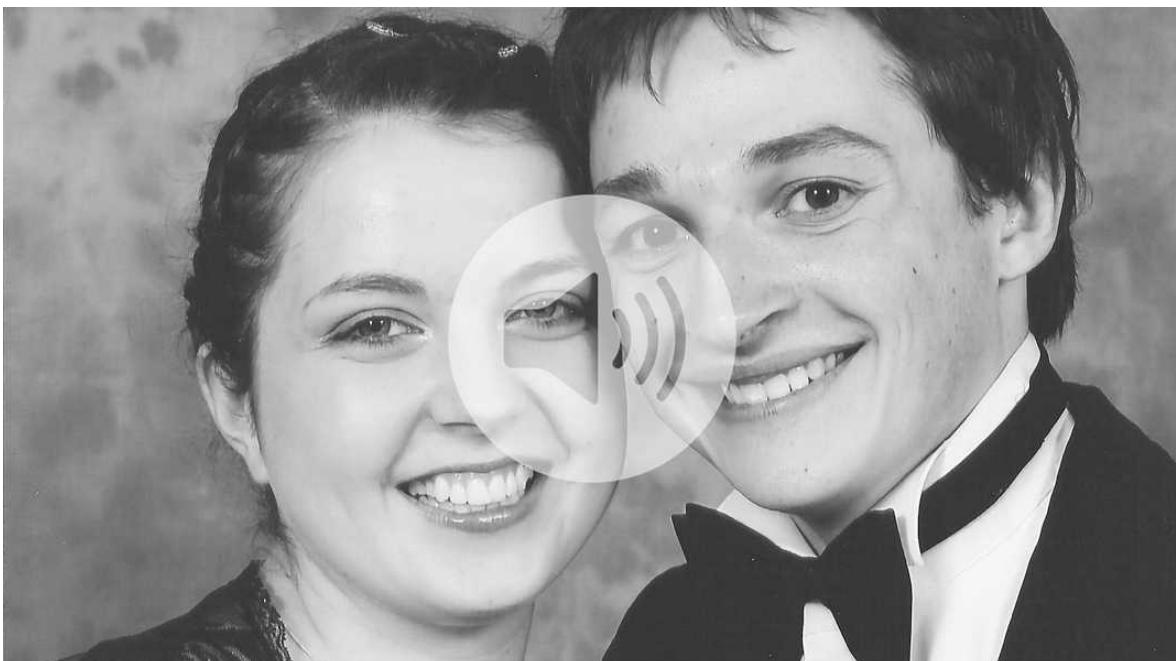
Part 3: When crisis hits

The pandemic shuts down the labs and sends both researchers reeling. Then comes personal setback.



[Editorial](#)

What a personal science saga reveals about funding — and science itself.



[Podcast](#)

Get to know Ali and Dan better in this 4-part series telling their story.

Authors: Kerri Smith, Heidi Ledford, Richard Van Noorden

Additional Reporting: Benjamin Thompson

Design: Lizzy Brown, Wesley Fernandes, Kelly Krause

Original photography: Chris Maddaloni for *Nature*

Subeditors: Anne Haggart, Anna Callender

Podcast editors: Benjamin Thompson, Kerri Smith

Editor: Richard Monastersky

Project leader: Kerri Smith

Personal photos provided by Alison Twelvetrees and Daniel Bose.

Related articles



[How junior scientists can land a seat at the leadership table](#)



How new principal investigators tackled a tumultuous year



These labs are remarkably diverse — here's why they're winning at science

SPRINGER NATURE

Springer Nature

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

© 2021 Springer Nature Limited. All rights reserved.

[Top](#)



[Built with Shorthand](#)

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02563-x>

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

Starting up in science

Part 2: The quest for cash

By Kerri Smith, Heidi Ledford and Richard Van Noorden
Photography by Chris Maddaloni

29 September 2021



Scientists Alison Twelvetrees and Daniel Bose opened the doors to their new labs at the University of Sheffield, UK, in 2017.

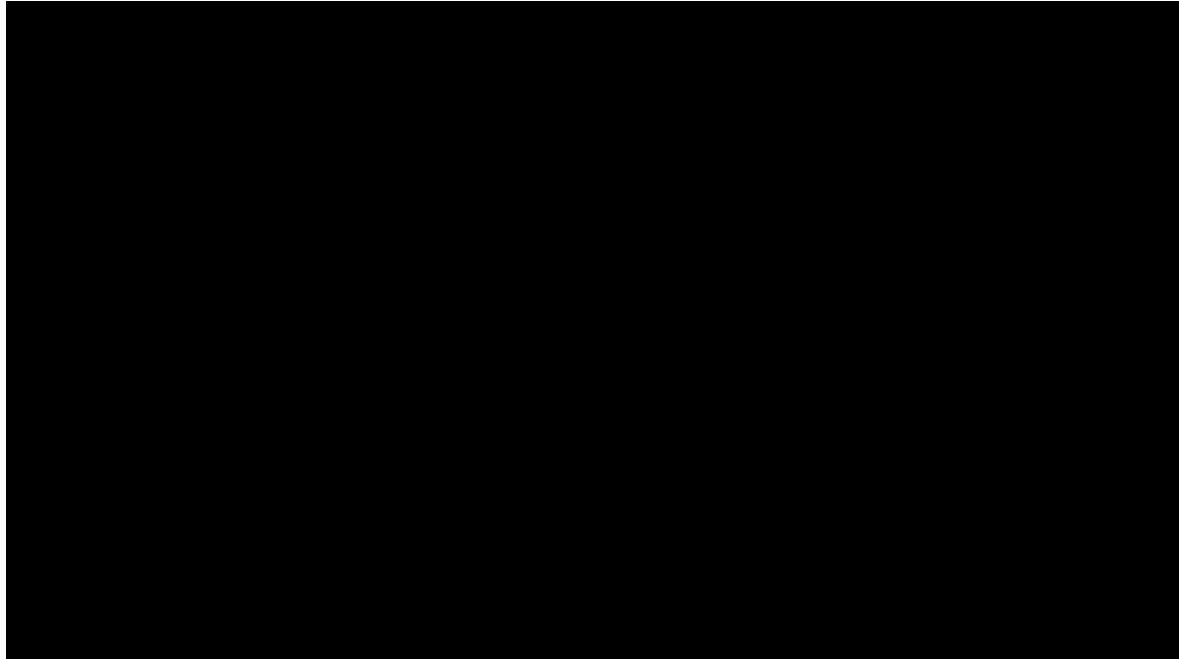


Their most urgent problem is funding — they can't keep the lights on without it.



In June 2018, Ali lost out on a £1-million grant. The year before, the same had happened to Dan.





Can they win funding while trying to design experiments, prepare lectures and supervise students?



This is part two of a three-part series ‘Starting up in science’. [Read part one.](#)



5. Team science



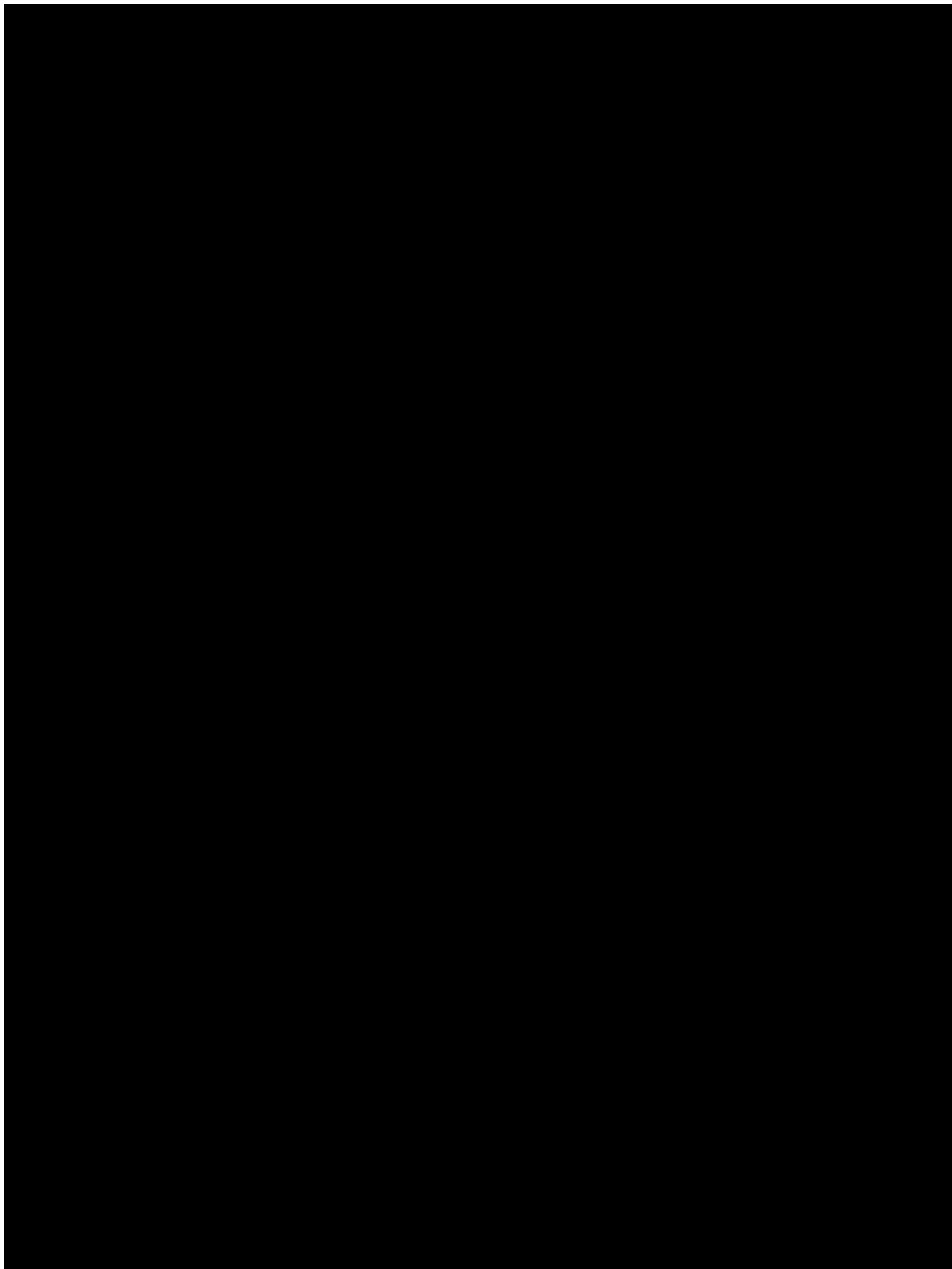


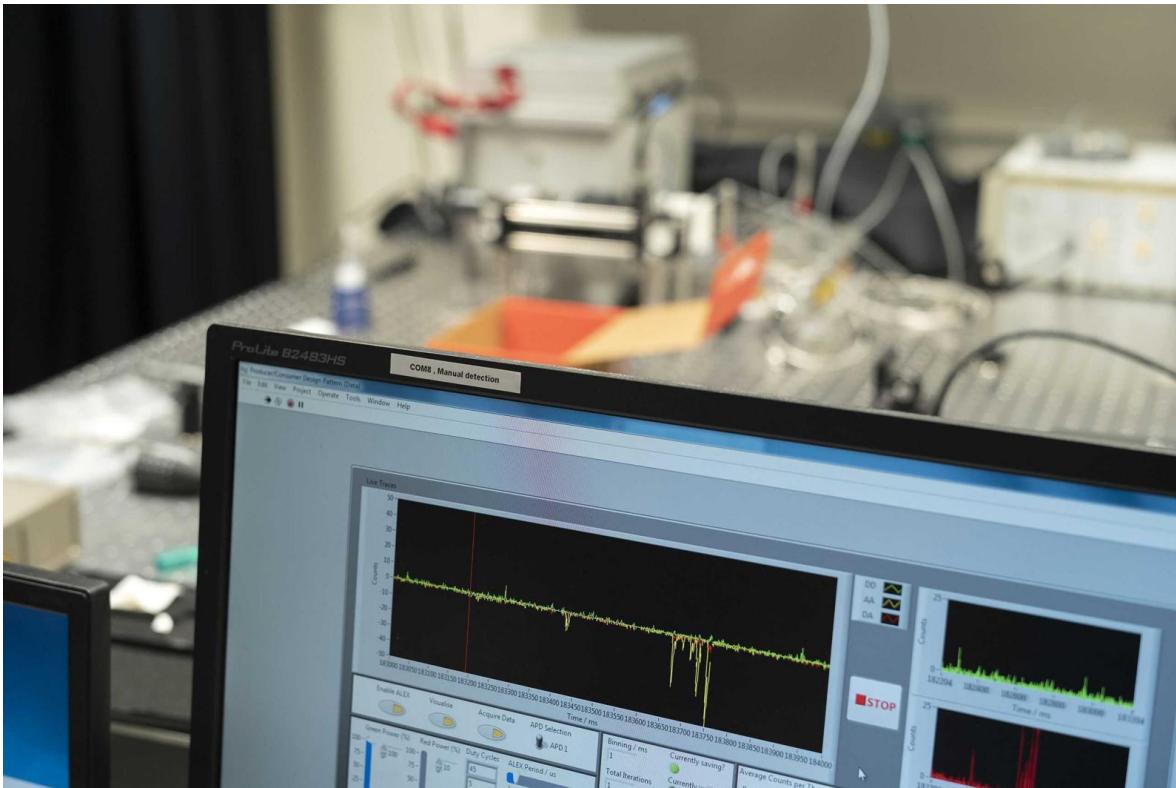
On an overcast morning in October 2018, Ali crosses a busy road clutching an ice box of test tubes. She's been carrying this precious cargo once a week to the lab of her chief collaborator, biophysicist Tim Craggs, in the University of Sheffield's physics department.

Their collaboration came about because of a chance encounter before Dan and Ali had even started at Sheffield. On a visit to campus, Dan met Tim and learnt that he had built a special kind of microscope for imaging single molecules. Ali immediately saw the potential for her own work.

"I'm like, 'Yes. That's what I need,'" she remembers.

So now she is making the ten-minute walk through the city at a careful pace to take her samples to Tim's microscope. She is trying to develop a way to study, one by one, the proteins that could form the foundation of her career as a principal investigator (PI). In the depths of the physics building, Ali enters a dimly lit room where a stainless-steel table is kitted out with lasers and lenses. "My son calls it an air-hockey table," says Tim, gesturing at the experimental set-up. It's essentially a way to spy on the shapes of the motor proteins called kinesins.





Data from Ali's FRET microscopy experiment.

Data from Ali's FRET microscopy experiment.

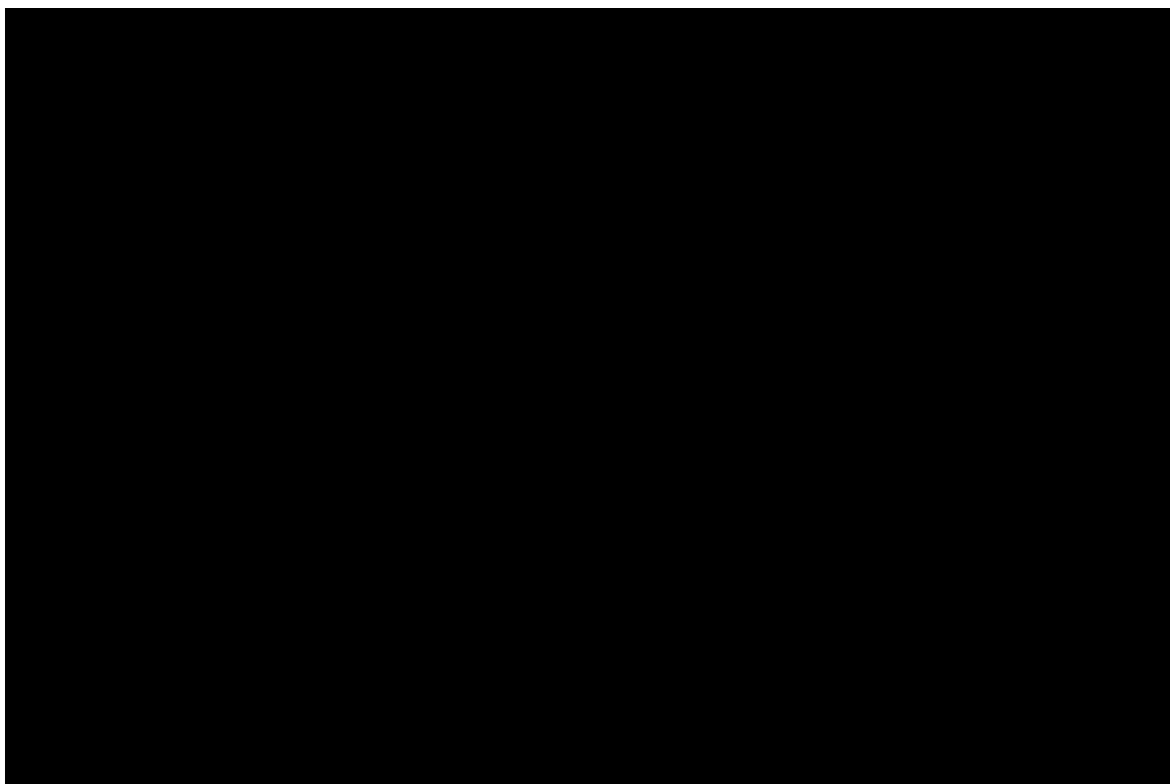
Kinesins are shaped like long sticks that have a hinge in the middle and two feet at one end. They are too small to see directly. But Ali has spent weeks creating kinesin proteins adorned with fluorescent tags at each end. The scientists shoot laser light through a diluted drop of cellular fluid from Ali's test tubes. When an individual kinesin molecule drifts under the laser, its fluorescent tags absorb energy and light up. The frequency of the fluorescence shows whether the motor protein is folded tightly at the hinge or stretched open into its active form. "The actual experiment — don't tell anyone — is incredibly easy," says Tim. The harder job, he says, is Ali's: preparing high-quality protein samples with the right kinds of tags inside cells.

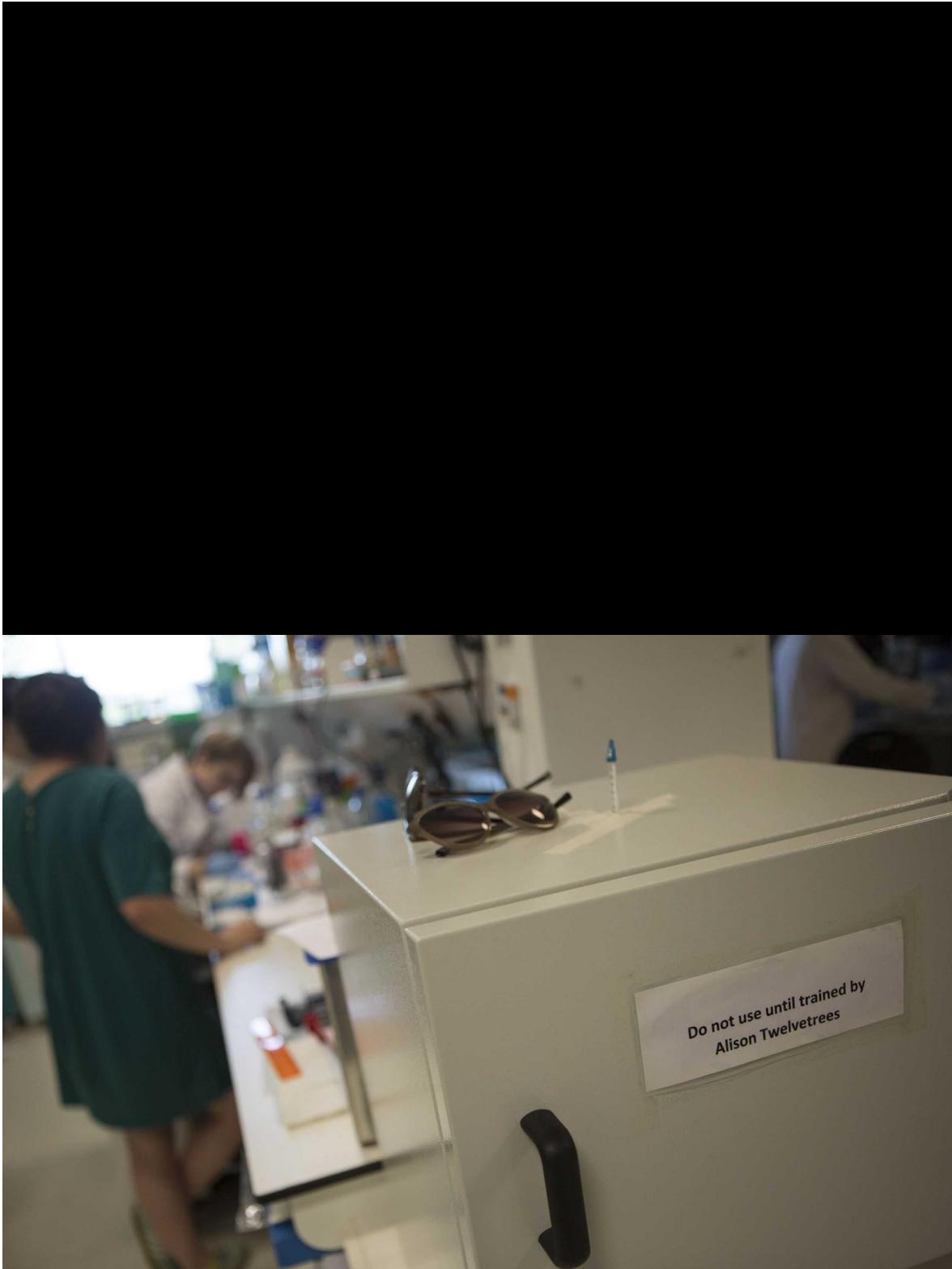
The technique is called fluorescence resonance energy transfer (FRET) microscopy and, much to Ali's delight, it seems to be working. She hopes to use it to see how single molecules of kinesin change shape when they grab cellular cargo, and then to ask what conditions inside a cell affect that

process. This could be the key advance that answers Ali's burning questions about how cargo travels down neurons at different speeds and what causes this process to go wrong. But it will take months, if not years, to progress from a "quick and dirty" experiment to a repeatable, high-quality protocol, Ali says.

For Ali, the collaboration with Tim and his team, which has been running for about a year, is crucial. It means she has access to a microscope despite having little cash of her own, and other people to bounce ideas off. "When you're starting out, you're a team of one. That's quite a lonely place to be," she says.

Now, 18 months on, Ali's lab is a team of three. She's taken on a PhD student, Evie Smith, who will work full time on the FRET experiment. A final-year undergraduate student, Ashleigh Davey, has also joined the group to work on a 'neuron-on-a-chip' project, which will enable the team to watch motor proteins move in cells. Ali has managed to negotiate a move to a larger lab space. But money is still incredibly tight. Evie is funded by a research grant for PhD students, but if Ashleigh is to stay for a PhD, Ali will need to get a grant to support her.





Ali's lab, with Ali and Evie in the background.

Ali's lab, with Ali and Evie in the background.

Dan's team has also grown. Several short-term students have come and gone, but PhD student Petra Celadova is a central figure. Petra joined in November 2017 with years of lab experience already under her belt, and quickly took ownership of her own project even as she helped Dan to set up the lab.

One of her first steps will be to use the gene-editing technology CRISPR–Cas9 to mutate an enhancer in leukaemia cells, looking for those mutations that impair the enhancer RNA's function. Less than a year into her project, Dan says, Petra is already independent. "She will tell you what to do, which I think is a great sign in a student," he says. "I take a lot of pleasure when she is like 'no, this is how we should be doing this'."

Petra is also wise to the harsh realities of lab work: she hopes to have her cell lines ready by the end of the year, but she knows that research will not necessarily bend to her timetable. "In science, you always think you will have done more than you actually can do in a year," she says.

Hear Ali and Dan's story in our 4-part podcast miniseries. Via [Apple Podcasts](#) | Via [Google Podcasts](#) | Via [Spotify](#)

6. Second shot



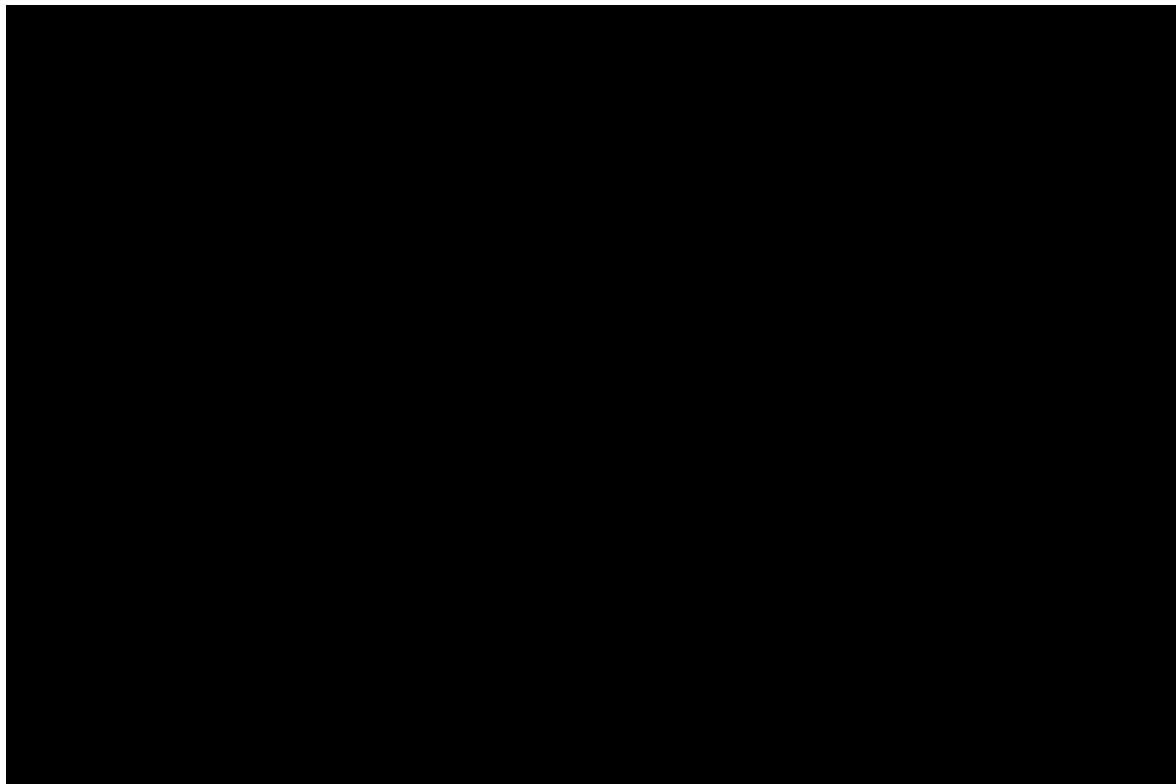


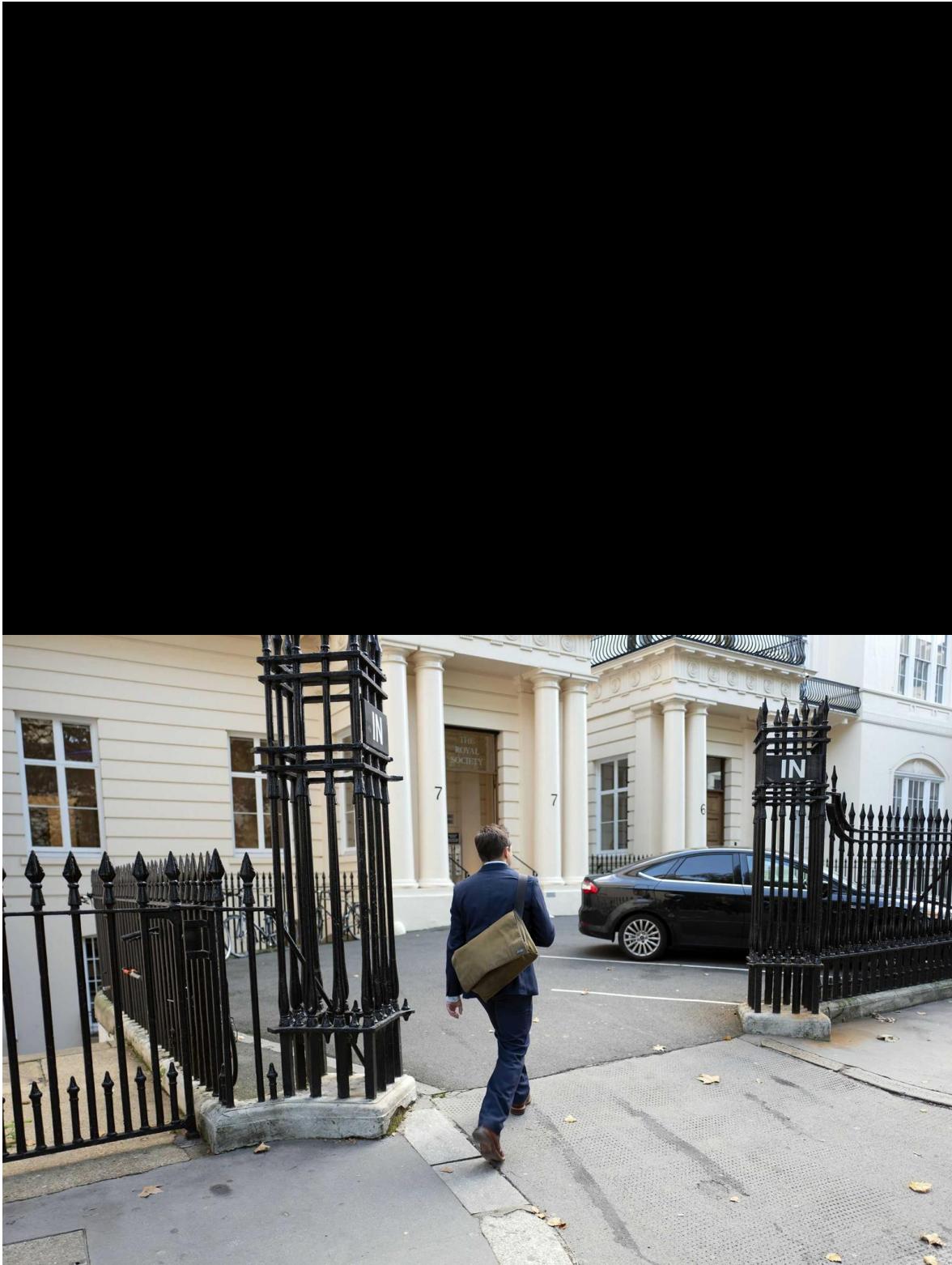
It's the eve of Dan's interview for the prestigious Sir Henry Dale Fellowship awarded by Wellcome and the Royal Society, which he first applied for in 2017 but didn't get. It's worth five years of lab funding that he desperately needs. He arrives at a hotel near King's Cross Station in London just in time to hold a goodnight video call with his three-year-old daughter, Ada.

The next morning, 19 October, Dan wakes up and has an extra-large breakfast to keep his stomach full and his mood positive. As Ada sends him a string of incoherent emoji texts, Dan runs through his practice talk. The interview will take place at the Royal Society's headquarters near Trafalgar Square, and will consist of [a five-minute presentation](#) followed by questions. "It's quite a lot to get in five minutes," Dan says. "So I've drunk lots of coffee."

With a bag full of snacks and his nerves jangling, Dan arrives too early at the Royal Society. He strolls through nearby St James's Park, pausing to watch a group of schoolchildren chase the geese. Then it's back to the Royal Society, where he sits in the lobby, has some water and tries to gather his thoughts.

There is cause for optimism: the reviews of his proposal were generally positive. He feels confident in the experiments he has proposed. And, in a way, he welcomes the interview experience, however agonizing it might be. The ability to stand up and defend your work is a great skill to have, he thinks. “Do I like it? No. I’m scared,” he says that morning before he leaves for the interview. “But I can see the value of it.”





Dan walks to his interview at the Royal Society in London.

Dan walks to his interview at the Royal Society in London.

The panel members call him in after their lunch break, and he walks into a room with about 20 people seated in a horseshoe shape. Only two or three seem to be leading the discussion; Dan focuses his eye contact on them. Some of the questions are surprisingly technical, but nothing catches him flat-footed. The 30-minute interview seems like only 10. “I felt like I gave a good impression of myself, and I think that’s all I can do here,” he says. “Other than that, I feel kind of tired.”

Once out of the Royal Society building, his mind returns immediately to his omnipresent to-do list: exam questions to write and all the other things he has pushed to the side while preparing. But first he is going to meet university friends for drinks in a few hours.

He calls Ali, who has spent the day distracting herself at the microscope, and Dan remembers his own anxiety during her interview for the same fellowship four months earlier. He wonders whether the family might go for a walk in the Peak District, a scenic national park near Sheffield, over the weekend to celebrate his new-found liberation from interview prep. “Ada has been bouncing off the walls,” he says. “I wasn’t a very diligent parent for a couple of weeks. I think we might unleash her on some sheep.”

7. Personal best





On 1 November 2018, Dan is on a train, sweating over the final details of a presentation he will give that afternoon at University College London. He used to think he would never be the sort of person who would still be preparing their talk on the way to give it. But now he's alternating between presentation prep and procrastination as he tries to avoid eye contact with his inbox, where a notification from Wellcome could arrive at any moment. "Waiting for e-mails to come in is very anxiety-inducing," he says.

Temptation eventually gets the better of him and, when he checks, the e-mail is there. "Dear Dr. Bose, I am pleased to inform you . . ." he reads. Dan's eyes jump forward to the word "successful" and then he looks away. He glances back and revisits the same words to be sure he hasn't misread them. "And then I sent it to Ali to double-check and probably read it again," he recalled later. "But I never made it beyond the first two lines."

Dan forwards the e-mail to collaborators, his department and people who'd helped with the application. He calls Ali. "We do all of this together. You ride the ups and the downs and the pain and the rejections and all of the stuff you do at the same time," he says. "So it's just lovely to be able to share that

kind of happiness and release with her.” After a few minutes, the train goes through a tunnel and cuts the call short.

The fellowship will give his lab £1.34 million over 5 years, half of which will be soaked up by salaries for the people who work there. But there will be some money for new equipment. “I’ve been promising Petra here for the last year, all the things we’re going to buy when we’re rich,” says Dan. He thinks about completing her set of micropipettes, buying extra time on the electron microscope and purchasing a machine for fast protein liquid chromatography (FPLC) that purifies proteins. Instead of spending time trying to find ways to stretch his funds to get particular experiments done, the lab could just do them, he thinks. “It just enables everything,” he says. “It makes such a big difference to everything we want to do.”

Dan’s lab in July 2019

And the lab in September 2021

The grant also buys Dan time. It funds him for research, guaranteeing some set-aside periods free from the demands of university teaching and committees. Many grants run for only three years; the five-year length of his would allow Dan and Ali to feel more settled. After years of juggling careers, coping with uncertainty and separation, they could look for a school for Ada with the knowledge that she would probably spend much of her primary-school education there. “It’s the first time we’ve been able to say, ‘This is where we’re going to be in five years,’” says Dan. “And that is just making a big difference to everything.”

Would the fellowship affect Dan’s salary? He isn’t sure. “I haven’t really looked at it,” he says. “It’s really very much secondary.”

But for the lab, he has lofty ambitions ready to hand. “Yeah, it’s just going to let us take off as a group,” he says. “That’s what we hope for.”

The next day, Dan kneels down in the middle of the corridor leading to his lab, cradling a bottle of champagne.

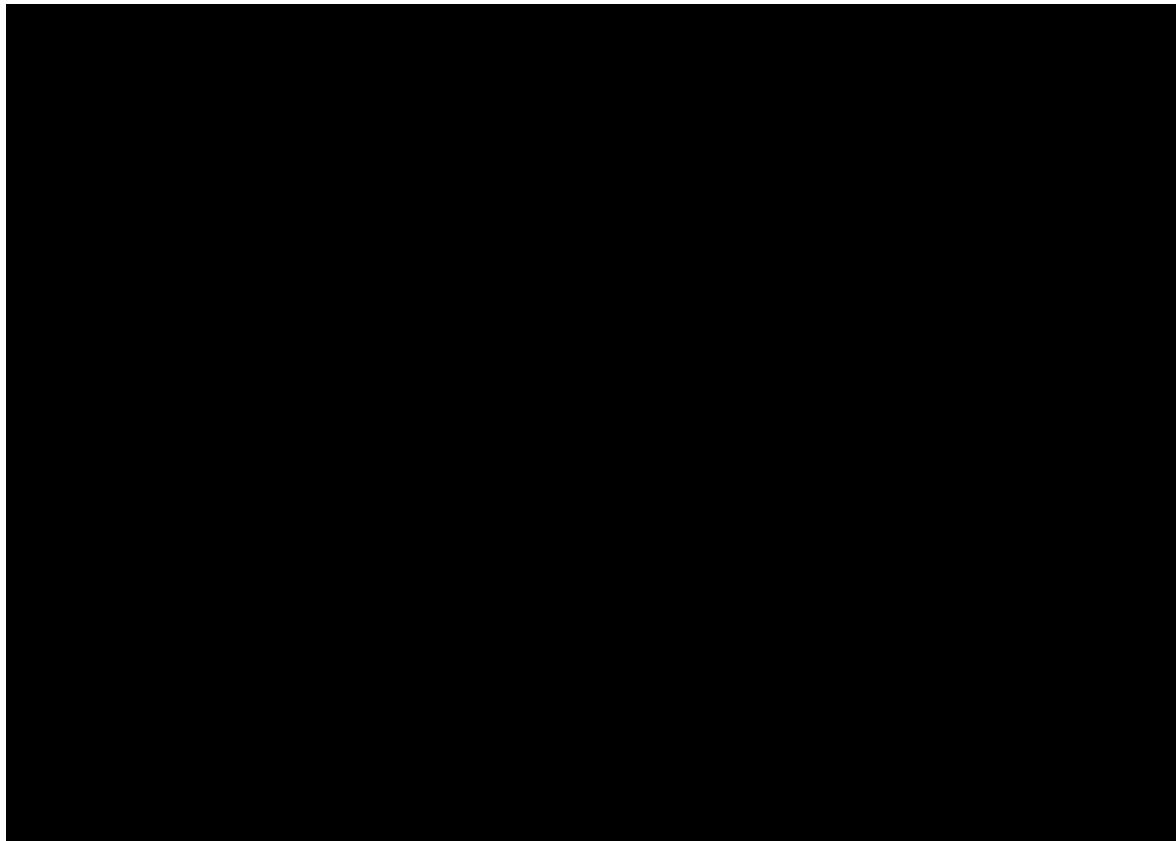
“Are we ready?” He shuffles slightly on his knee.

Behind him a gaggle of his colleagues monitor his position on the floor. Their chatter turns to whoops as Dan ejects the bottle's cork and watches it ping down the corridor and hit the wall just above a poster about a cancer biomarker.

It's the lab's tradition to fire off a champagne cork for each new grant; Dan's last effort, a small award from the Royal Society that began at Easter, was worth £20,000.

Where the cork hits the wall, his colleague affixes a piece of orange tape with Dan's achievement and the date.

"That's a PB!" someone shouts.



Dan B SHD 2/11/18
fellowship .

Dan's personal best — on tape.

Dan's personal best — on tape.

Dan having the security of a fellowship “takes the pressure off a little bit”, says Ali. Emotions have been running high as they both digested the news. “Knowing that I didn’t have to worry about Dan any more just kind of came over me and I just felt very overwhelmed.”

The feeling of relief is fleeting. As 2018 draws to a close, Ali is more than halfway through her three-year probation, and time for her to win a grant is fast running out.

In late December, they put those concerns aside to celebrate Christmas with their labs. Last year, Ali and Dan organized a dinner with another new PI, molecular biologist Emma Thomson, who shares lab space with Dan. There were just five diners, including the three lab heads. This year, Ali says, she and Dan had to move a table, chairs and plates to Emma’s house so a dozen people would fit, and “suddenly it was, ‘Ah’, I have built something.””

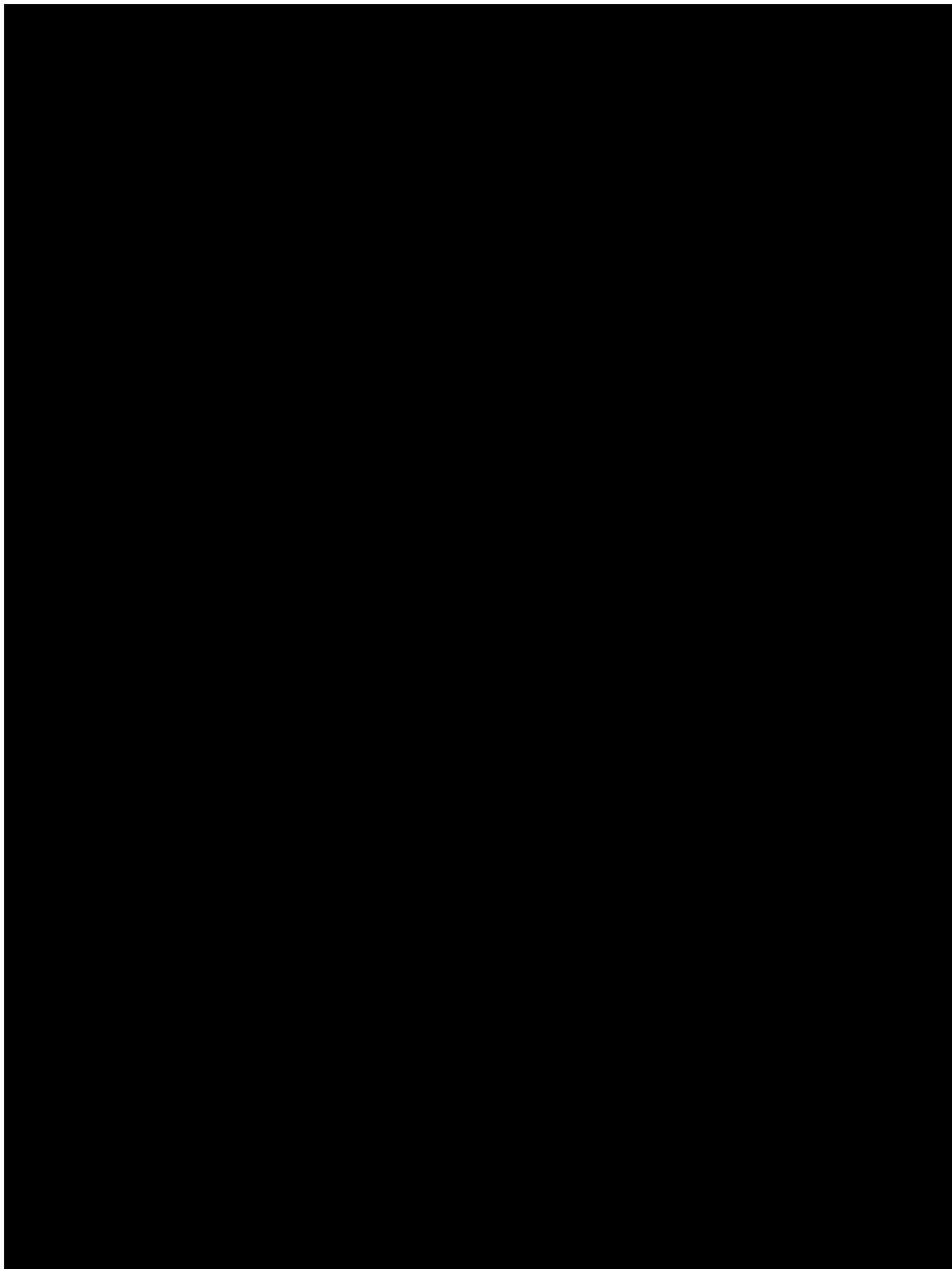
8. “I can see myself within that data”





In the new year, the family rents a larger, more comfortable house and instantly feels happier, after two years in what Ali calls an “adequate” rental home. Their kitchen now has a dishwasher, the garden has a swing and room for plants — Ali and Dan start growing Tuscan kale — and the walk to work now takes them through Sheffield’s floral botanic gardens. Ada’s school, which she’ll start this year, is only ten minutes away.

After so many years of moving around, the family feels more settled and able to bring out more of the books and other possessions they’d kept in storage. They make resolutions to throw away what they don’t need. “We’ve accumulated three continents’ worth of possessions,” Dan says.





Ali with the couple's daughter, Ada, in their garden in Sheffield, UK.

Ali with the couple's daughter, Ada, in their garden in Sheffield, UK.

At work, Ali has spent much of the past few months setting up protocols and tweaking them over and over again until they work. "It seems so glacially slow sometimes," says Ali. "But I feel like that's inevitable, because you're pulled in so many directions at once. It takes real effort to find the time to do the stuff that you need to do in lab."

One notable success is that Ali and Emma Thomson have joined together to offer a PhD position split between their labs. Ali's undergraduate student Ashleigh applied for and won that position, and will start in October 2019 using the neuron-on-a-chip that she had developed to grow single neurons in channels and study ribosomes in the axon. "It's super, super exciting," says Ali — but it also puts pressure on her to bring in more research funding.

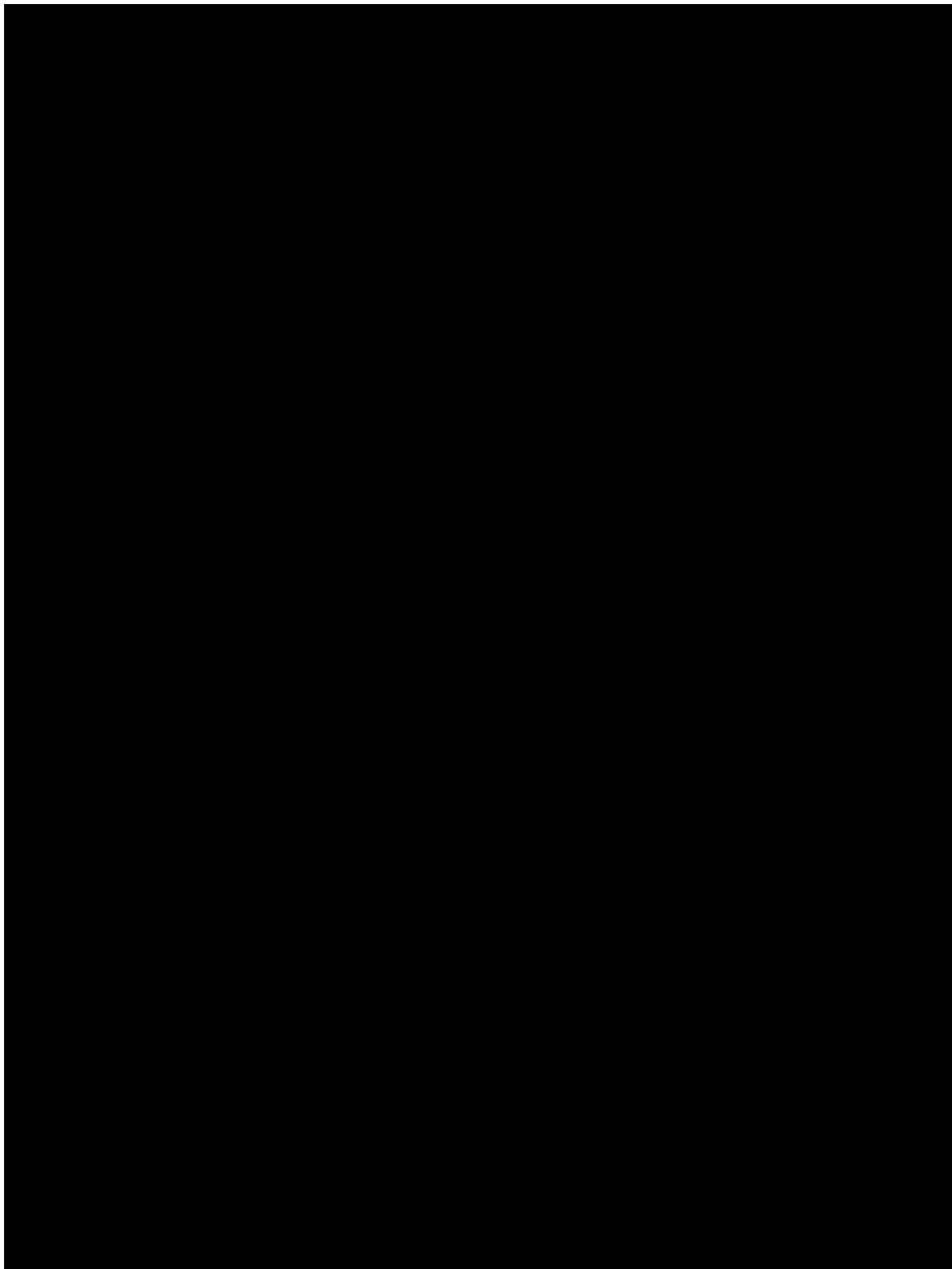
The Slack messaging group she helped to set up with immunologist Sophie Acton at University College London in 2018 has generated its own workload. Ali and Sophie kicked off a survey of new PIs and their

experiences, and by March 2019 they had crunched the data and posted their analysis to the preprint server bioRxiv — since published in the journal *eLife* ([S. E. Acton et al. *eLife* 8, e46827; 2019](#)).

Ali has been diverted by [some stark results from the survey](#), which attracted responses from more than 350 PIs who had launched their first UK labs between 2012 and 2018. Women starting their first research labs tended to have a lower salary and receive less university start-up funding than did their male peers, the results show. Female PIs were also less successful at securing further funding in the first five years.

“I can see myself within that data,” Ali says. Of the two, Dan got a higher salary and more start-up funding — although, Ali adds, they were hired into different positions in different departments, so aren’t directly comparable. “It’s difficult to not compare myself to Dan a lot of the time, because he’s right there. And that’s not particularly healthy,” she says. For Ali, the survey made it clear that the gender pay gap applies to new female lab heads as soon as they start work. “When you’re appointed, you believe that you’re equal to everybody who’s appointed at that stage,” she says. “Actually, it’s not true. We’re being appointed at a deficit immediately.”

Ali takes immediate action on one of the survey findings: that women without mentors reported the lowest level of optimism about their careers. “I could do with more mentorship,” she says. She signs up for a mentorship programme at Sheffield.





Dan pipettes reagents in his lab.

Dan pipettes reagents in his lab.

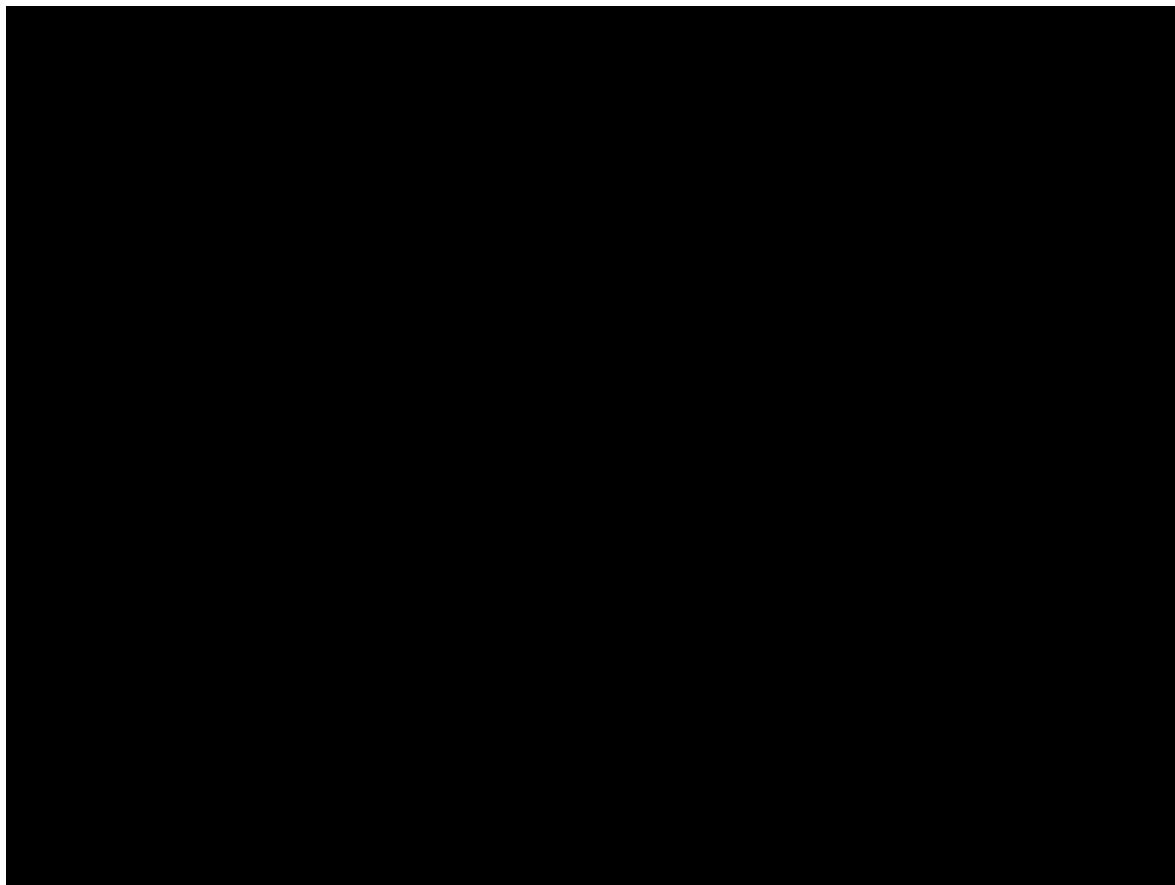
Dan is finding ways to progress some key experiments while dealing with a burst of teaching, admin and recruitment. He taught a group of students in the lab to clone and purify some of the proteins he'll need for his cryo-EM experiments. Those students have moved on, but the samples are nearly ready.

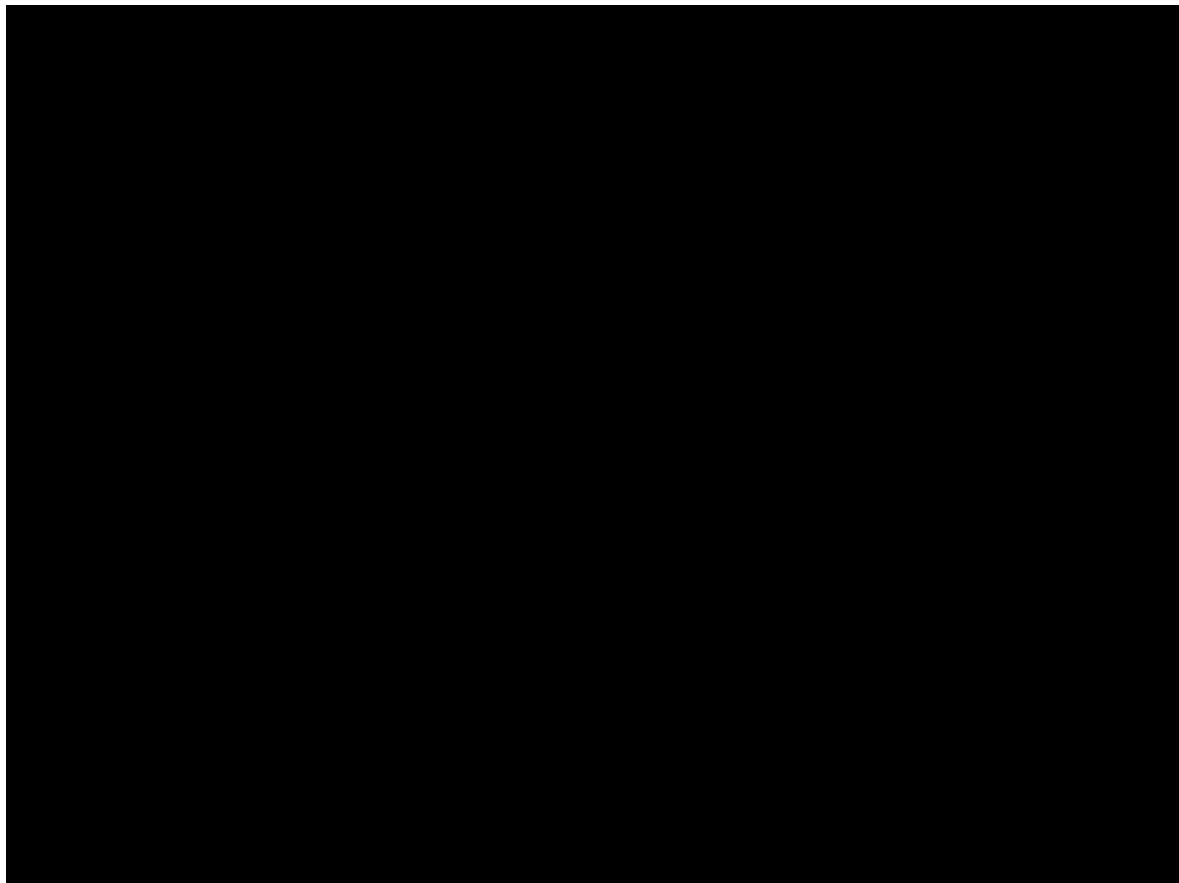
He's also interviewing for new PhD students. With the United Kingdom's formal exit from the European Union approaching, he's noticing a drop in applications from Europe. "It's quite sad. You try to build a culturally diverse laboratory and suddenly no one wants to come because of the uncertainty," he says.

In late March, Ali decides not to apply for the Sir Henry Dale Fellowship by the rapidly approaching deadline. She is juggling too many priorities to make a strong showing, and is gambling that if she waits until a new round for applications in August, there will be stronger preliminary data from her

lab's work to bolster her bid. But it is a big internal struggle to postpone. If Ali doesn't win a grant in the next year, she could be out of a job after her university reviews her three-year record. "That will be a disaster for everybody."

The stress is starting to shake her. "I'm quite up and down. I think I am struggling generally at the moment," she says when we catch up in April. Starting at Sheffield has not been easy, and before that she felt she was still playing catch-up after maternity leave. "It's just constantly felt like an uphill battle to try and have any sort of confidence in what I'm doing, and it's really difficult to kind of be upbeat. Sometimes I just do still feel like I want to quit."







The family in Patagonia in 2019.

The family in Patagonia in 2019.

Still, there are jubilant moments. Evie and Ali make their FRET microscope results public for the first time at a conference in Chile in April 2019.

“We’re really, really excited about what this method can show us in the future about kinesin activation, although we haven’t discovered any new biology just quite yet,” Ali tells scientists at a small session titled ‘The next generation of cytoskeleton researchers’. Dan flies out with Ada to Chile, so that the family can take advantage of the conference to hike in Patagonia — one of Dan’s bucket-list destinations.

9. Rainbow tape



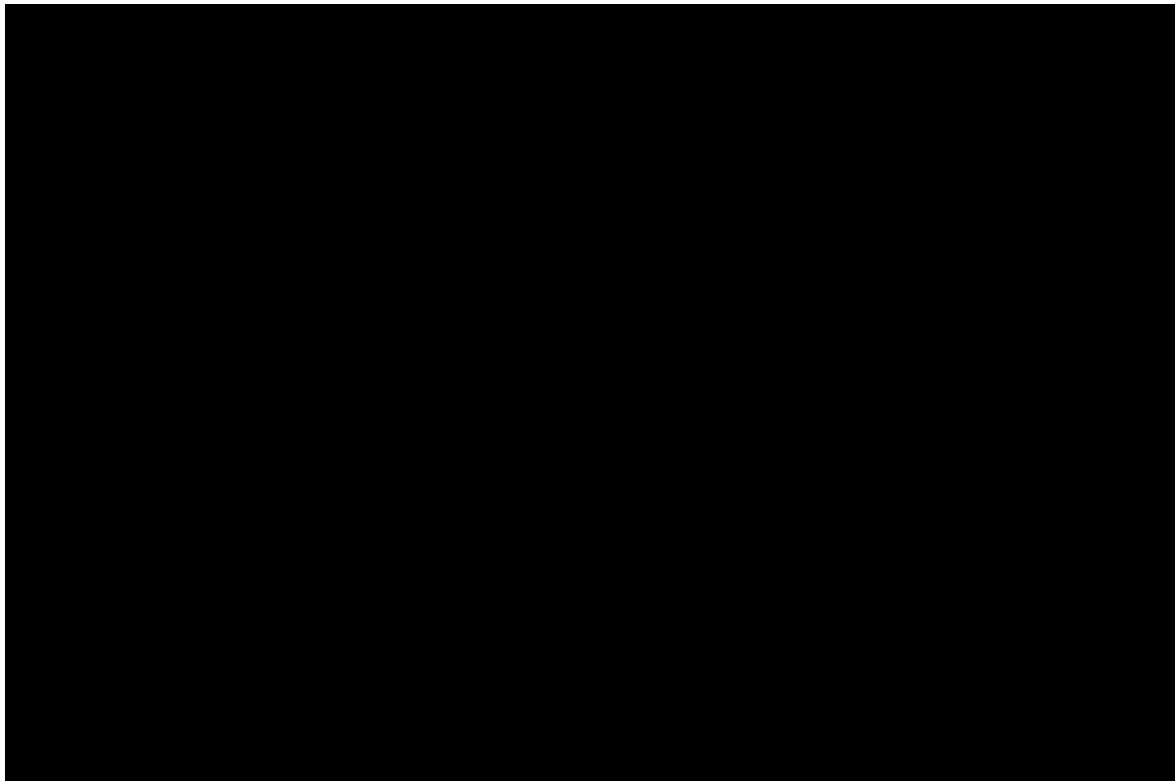


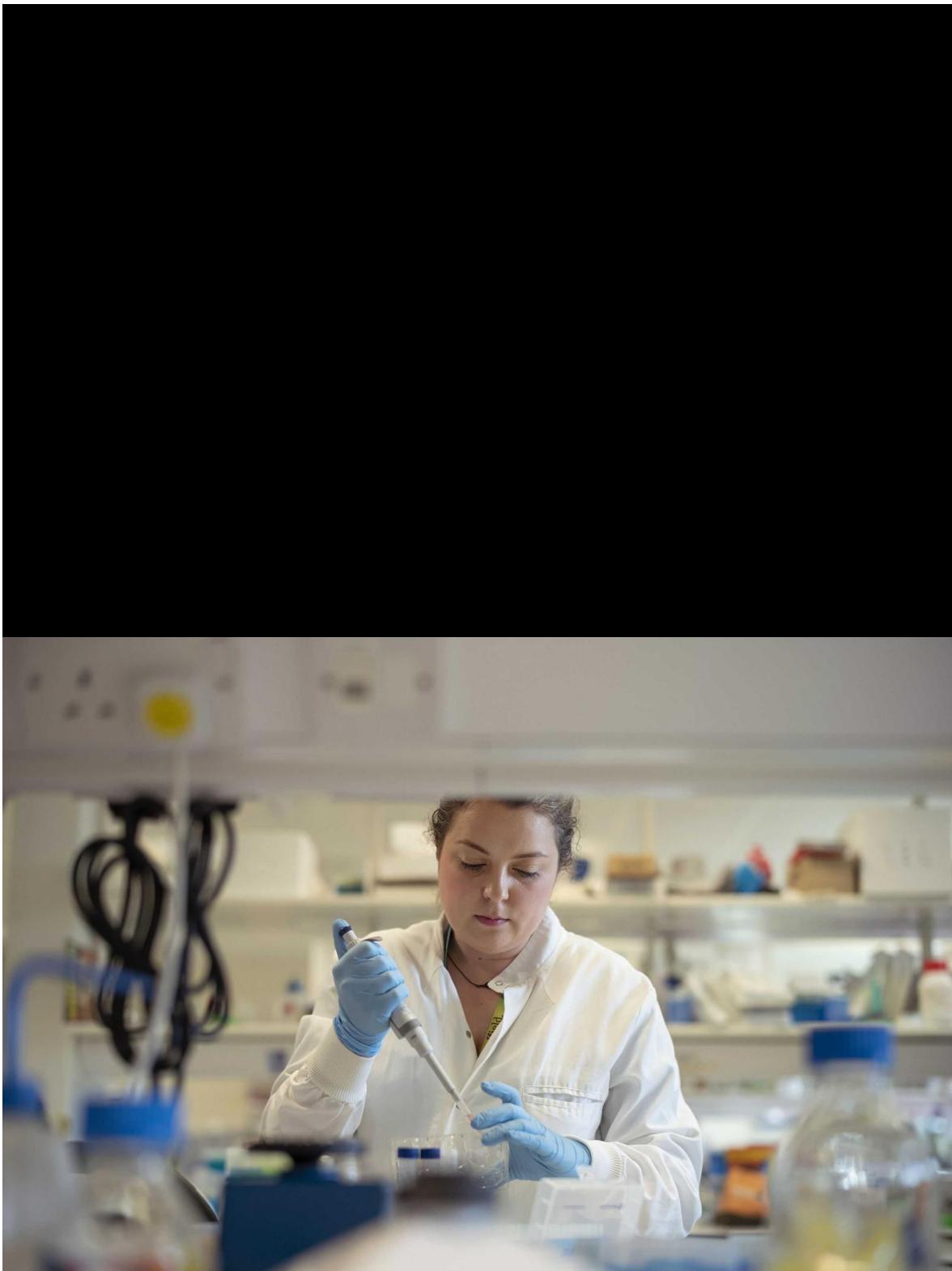
Dan's fellowship begins on 1 June 2019, and he quickly sets about laying purchasing plans. He agonizes over what he should buy first, consulting Ali at length. "I had this quandary: you get this grant, what's the first thing you buy with it?" he says. "It has to be semi-useful but also frivolous. So I got a seven-roll tape dispenser and rainbow tape to put on it." Then he needs to move on to more-practical purchases, such as centrifuges and shaking incubators. The biggest purchase is the FPLC machine for purifying proteins, with a price tag hovering around £65,000 — more than triple Dan's budget when he first opened his lab.

Is he at risk of forgetting his humble beginnings? "Maybe ask me in six months' time, when the gold-plated centrifuge turns up."

But the clock keeps ticking and, a little more than two weeks after his fellowship began, Dan still hasn't been able to bring himself to buy the FPLC machine. "Every time I think about it, it brings me out in some kind of horrible sweat," he says, with a tone that conveys a blend of guilty shopaholic and anxious spendthrift. "It's kind of terrifying, especially when I've been getting by on scavenging the backs of cupboards for a couple of years."

Petra's project seems to be going well: she has engineered the different leukaemia cell lines that she will need, and is on the verge of collecting data. Dan's goal is to get some samples on the cryo-EM over the summer to see how RNAs bind to a protein called CBP that switches genes on and off. Thus far, he has dealt with these molecular complexes only in a test tube. "To actually physically see an interacting RNA and CBP for the first time is going to be super exciting for me," he says.





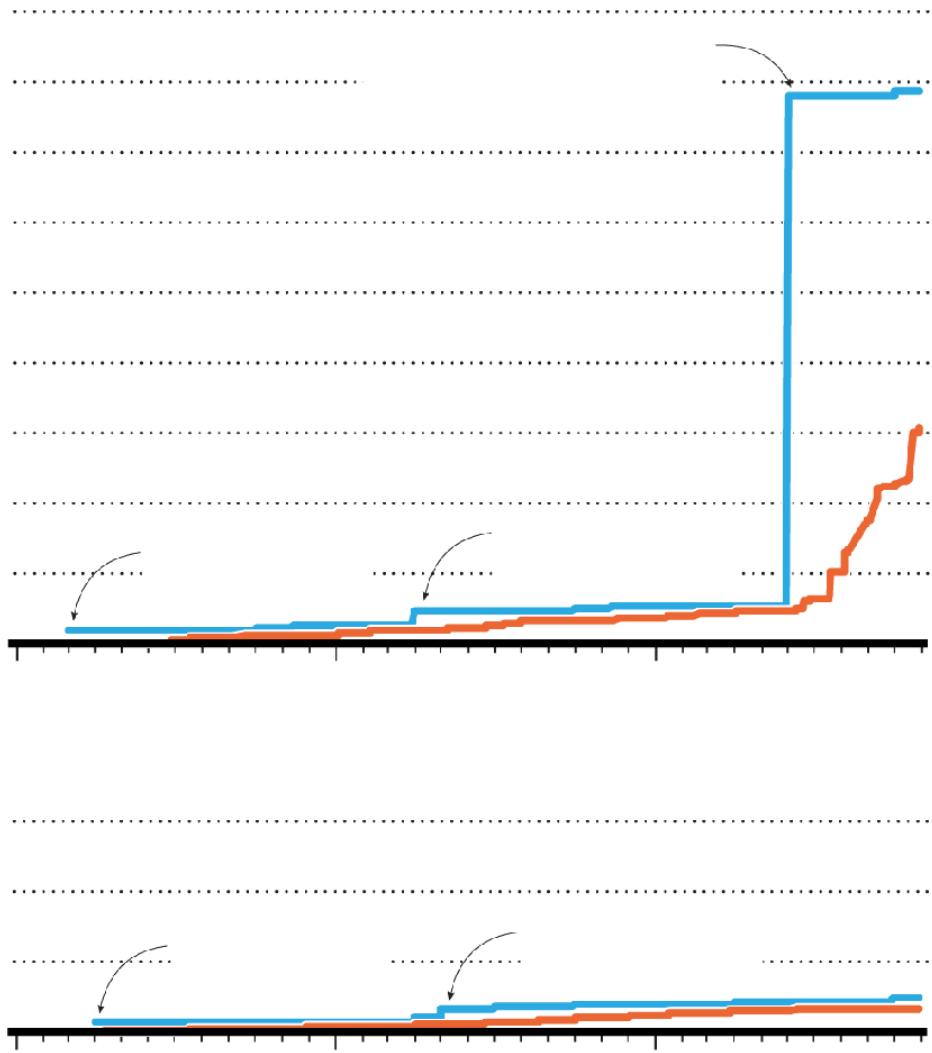
Ali prepares cell samples.

Ali prepares cell samples.

Ali, meanwhile, is starting to seriously worry about her finances (see ‘*Contrasting fortunes*’). She’s managed to convert a master’s student at Sheffield, Emma Turner, into a PhD student in her lab, starting in the autumn. For her project, Emma will be engineering mutations into kinesin — the same suite of mutations seen in people with motor neuron disease — to see what effects they have on kinesin walking and unfolding. Her presence will mean that Ali will have doubled the size of her lab team this year, but not increased her income for instruments, reagents and other consumables. “This is my biggest anxiety and stress,” she says.

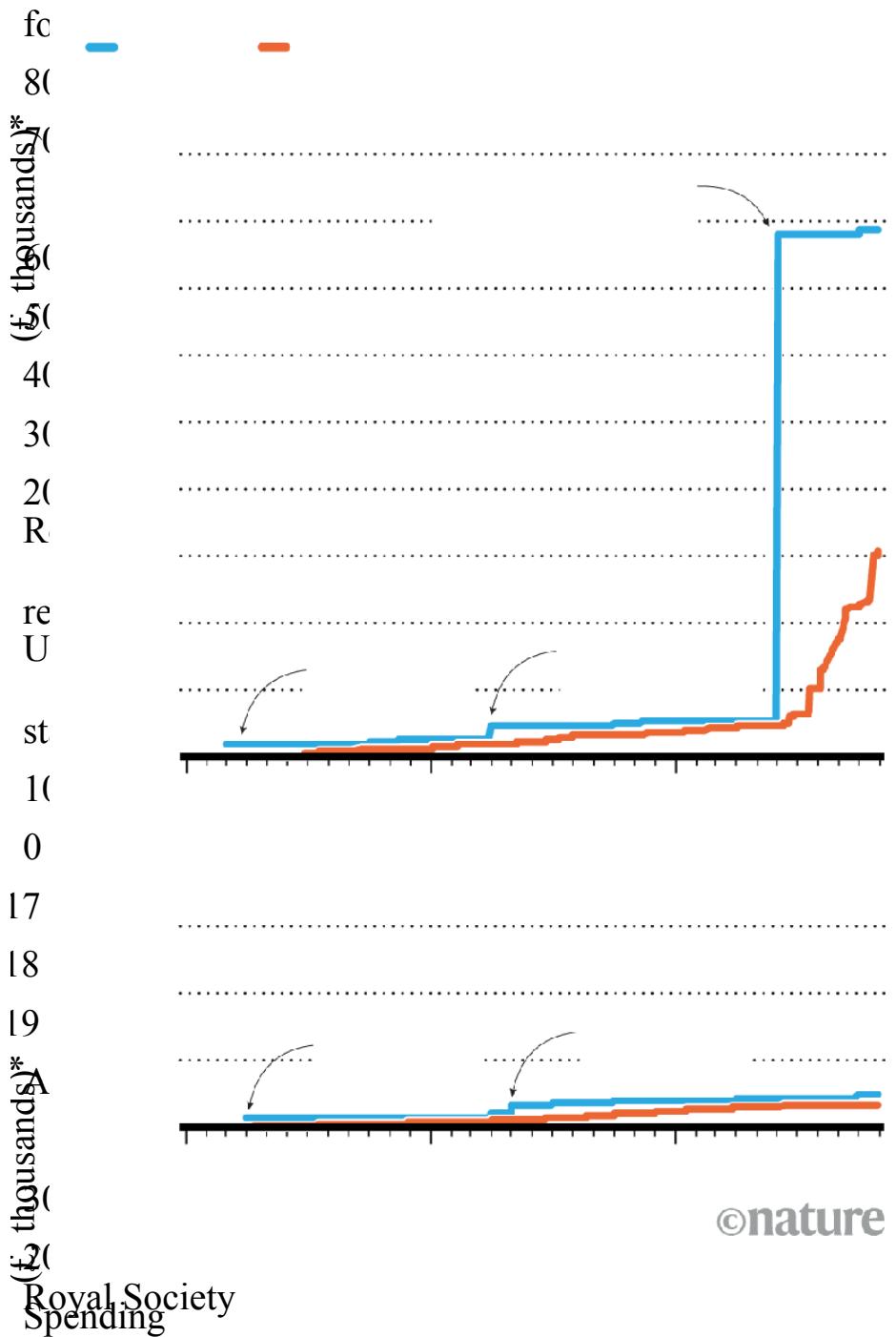
Contrasting fortunes

Whereas Dan’s large grant allows him to buy new equipment and reagents, Ali has to watch her spending carefully. (Chart excludes staff salaries.)



©nature

Spending
Money in
Daniel Bose's lab
\$900 Sir Henry Dale award
(amount available



©nature

Royal Society
Spending

Money fund
University

Daniel Rose's
start up fund
lab

The students get money to cover their tuition fees and living expenses, but they don't always get cash for research materials. Evie does have a support

Grant of about £5,000 a year, which is largely spent on essential equipment such as pipettes and biochemical reagents. Her project can go ahead only because they are using Tim Craggs's microscope. "I have to be upfront about what we can and can't afford," Ali says. She is eyeing a couple of small grants to kickstart and maintain projects for her students.

To top it all, Evie's microscope experiments, which seemed to be going well, have mysteriously stopped working. "It's super frustrating," says Evie. Ali calls it a "summer of anguish".

By the end of October 2019, as Dan is slowly and anxiously spending his grant money, Ali is midway through her second application for the Sir Henry Dale Fellowship. "Feeling unexpectedly excited/proud of it!" she writes in an e-mail after submitting the full pitch. But this is only step two of four; if they like her application, the judges will interview her in March next year. Then it will be another anxious wait to see if she, too, has secured the money she needs for her lab to survive.

Next: The world changes

Read part 3 of *Starting up in science*.

Nature 597, 613–618 (2021)

doi: <https://doi.org/10.1038/d41586-021-02564-w>

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

©nature

Starting up in science

Money

Nature follows two researchers over three years as they struggle to launch their own labs.

Daniel Boses

Alison Twelvetrees'

200



400

Part 1: The pitch

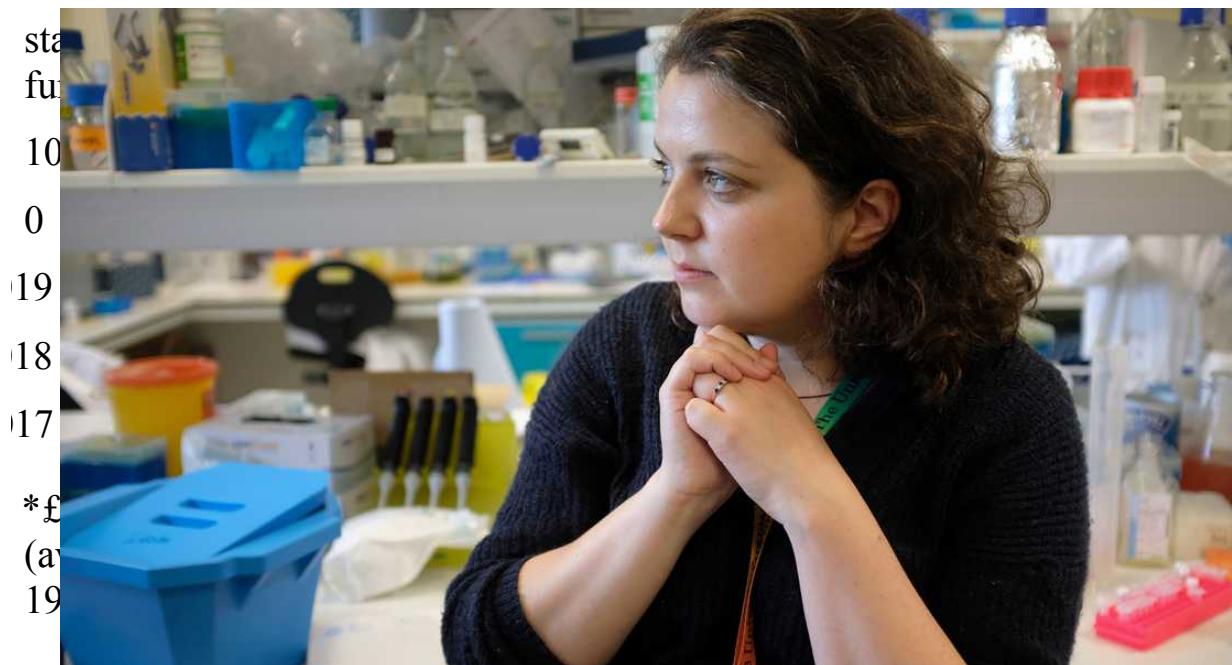
200 and Dan face a series of challenges as they begin their journeys as lab
leaders at University



300
holands)*

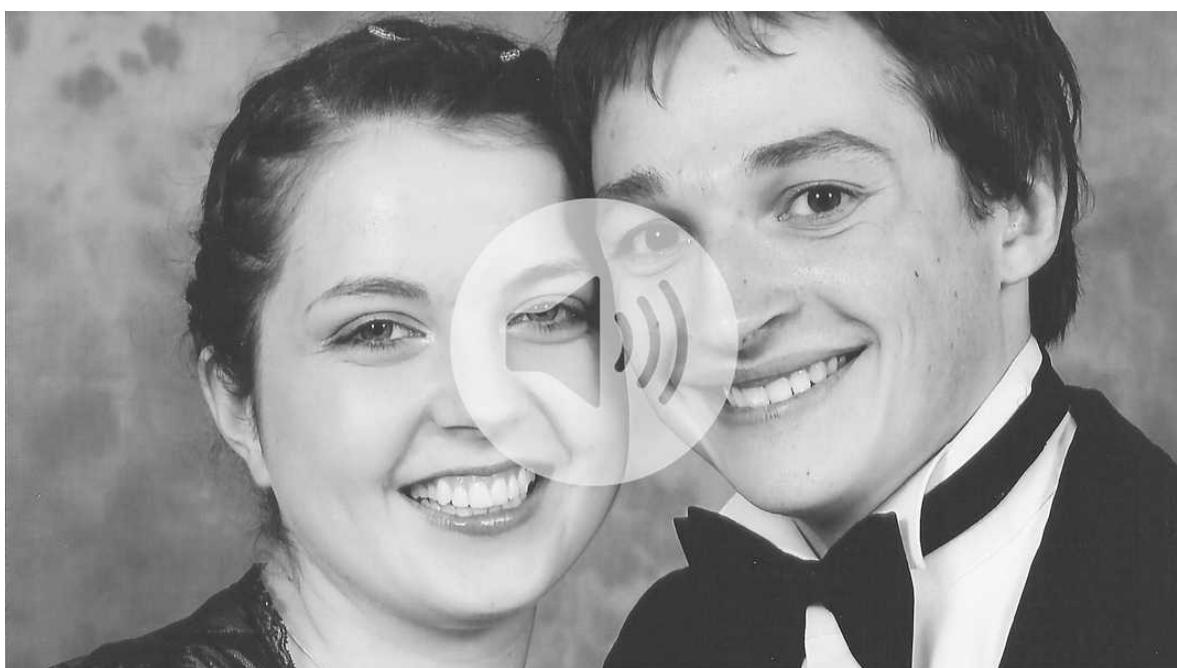
Part 3: When crisis hits

The pandemic shuts down the labs and sends both researchers reeling. Then comes personal setback.



Editorial

What a personal science saga reveals about funding — and science itself.



Podcast

Get to know Ali and Dan better in this 4-part series telling their story.

Authors: Kerri Smith, Heidi Ledford, Richard Van Noorden

Additional Reporting: Benjamin Thompson

Design: Lizzy Brown, Wesley Fernandes, Kelly Krause

Original photography: Chris Maddaloni for *Nature*

Subeditors: Anne Haggart, Anna Callender

Podcast editors: Benjamin Thompson, Kerri Smith

Editor: Richard Monastersky

Project leader: Kerri Smith

Personal photos provided by Alison Twelvetrees and Daniel Bose.

Related articles



How junior scientists can land a seat at the leadership table



How new principal investigators tackled a tumultuous year



These labs are remarkably diverse — here's why they're winning at science

SPRINGER NATURE

Springer Nature

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

© 2021 Springer Nature Limited. All rights reserved.

[Top](#)



[Built with Shorthand](#)

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02564-w>

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

Starting up in science

Part 3: Crisis point

By Kerri Smith, Heidi Ledford and Richard Van Noorden
Photography by Chris Maddaloni

29 September 2021



Alison Twelvetrees and Daniel Bose are newly minted lab heads at the University of Sheffield, UK, and they need to win grants to keep their operations going.

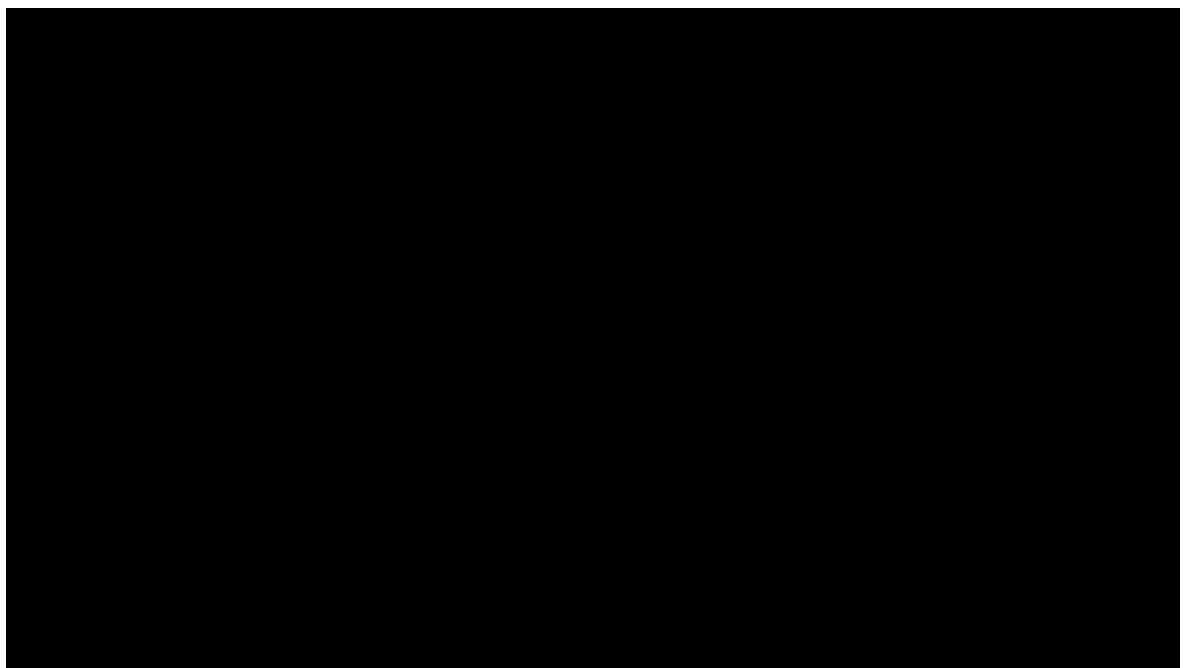


The financial pressure is off for Dan, who won a large fellowship at the end of 2018. But, by 2019, Ali still doesn't have the money to support her lab's future.



In 2020, just as both are finding their feet, there's a global pandemic. And then, something even more world-shattering happens to the family.



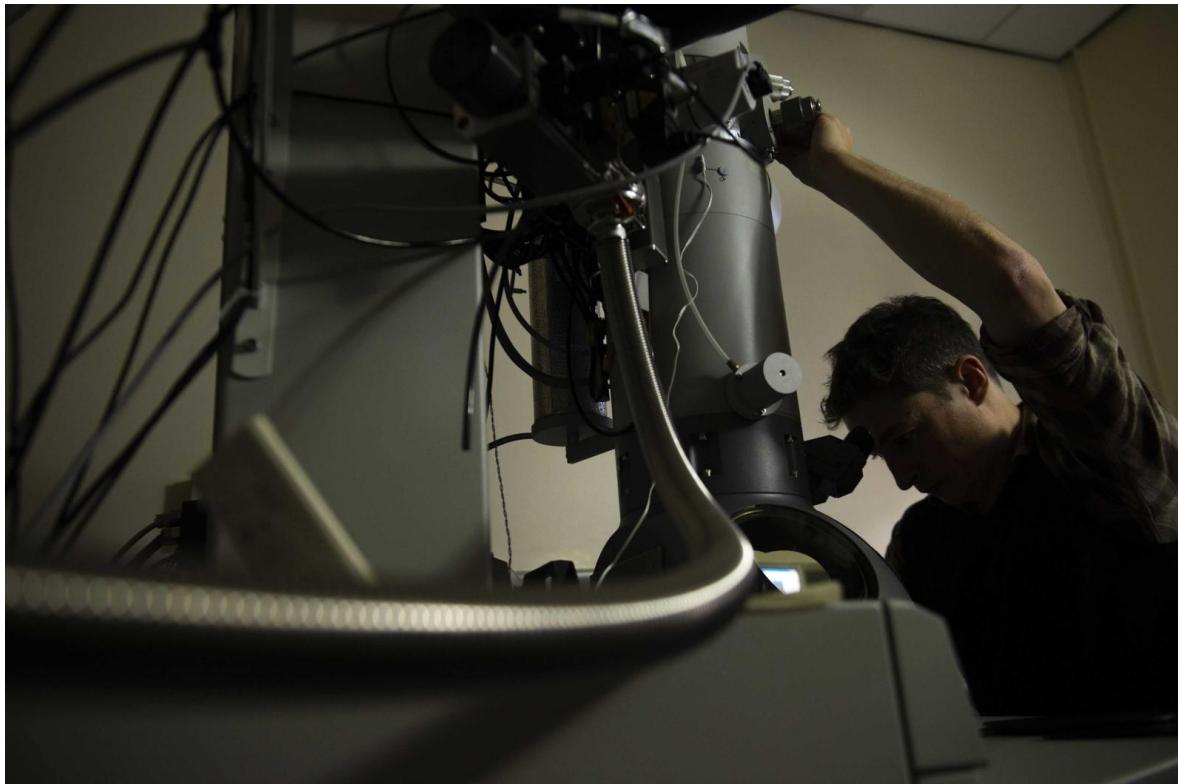


This is the third and final part of ‘Starting up in science’. Read [parts one](#) and [two](#).



10. Protein universe





In a small, dark room inside the university's biology building, Dan carefully transfers a sample onto a metal grid that will go into a large transmission electron microscope. The microscope takes up half the space in the basement room.

"I always find this the most exciting step," he says. "When you've spent a long time with a protein, you feel how it behaves in a test tube and how it interacts. To stick it under a microscope and see it for the first time is always quite a nice experience."

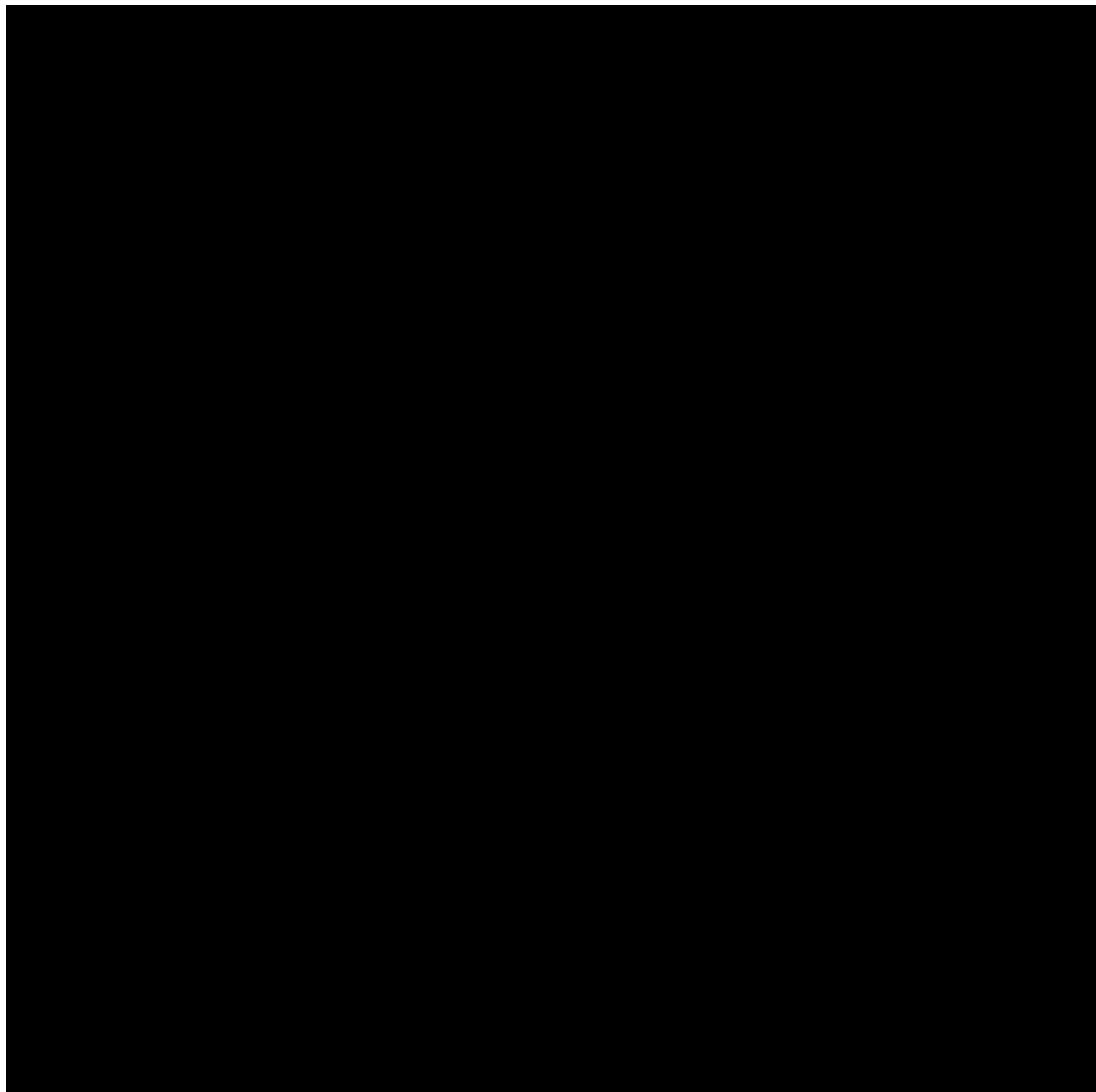
It's September 2019, and Dan is finally starting the experiment he's long dreamt of doing. For most of the past decade, he has been studying a protein known as CBP, which controls how easily a gene can be switched on. Researchers have glimpsed some of CBP's structure, but they haven't worked out all of it — and they don't know how CBP interacts with an 'enhancer RNA' that Dan's lab is studying.

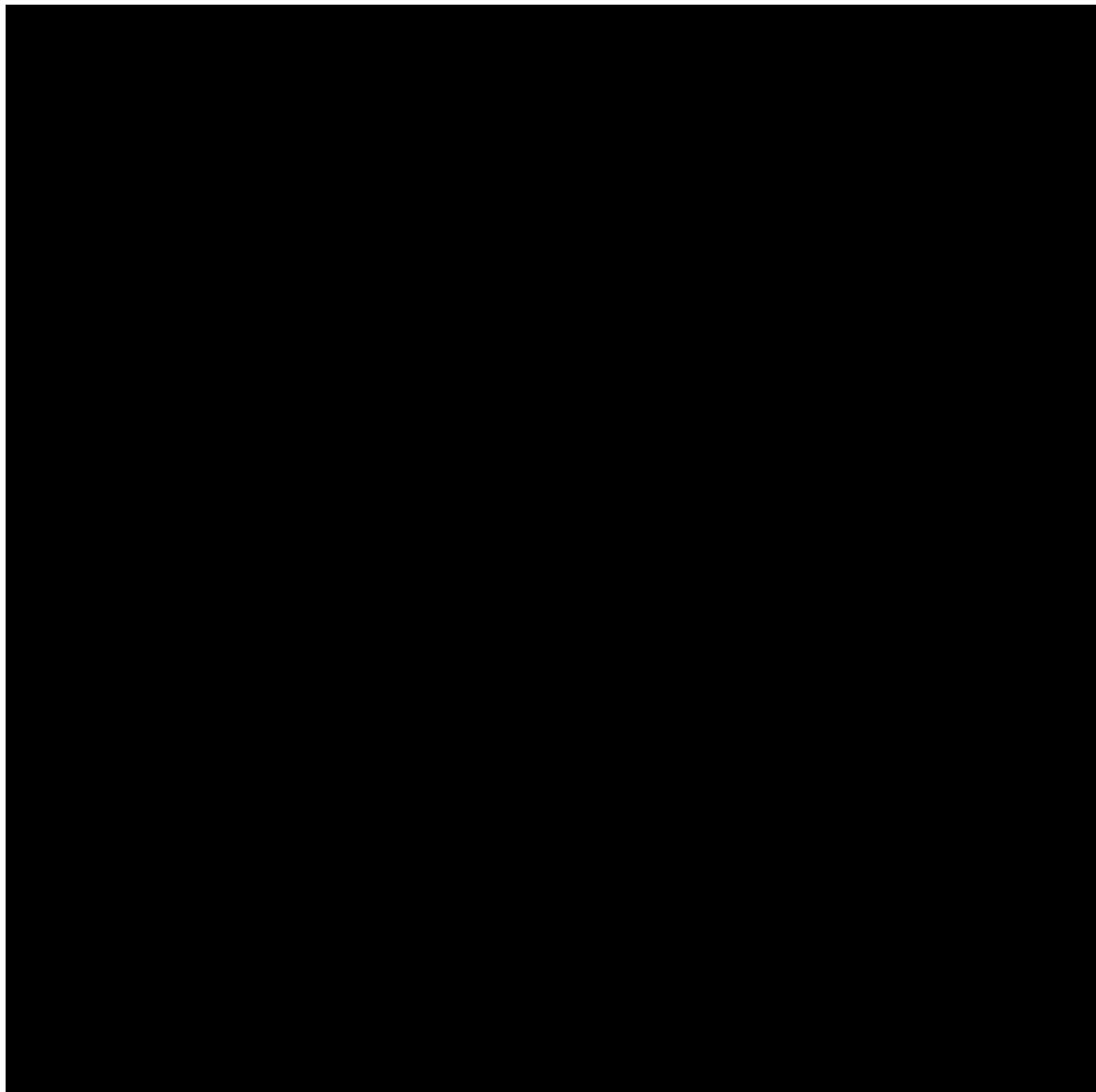
Now, with a break in his teaching schedule, Dan will try to decipher the RNA–CBP structure.

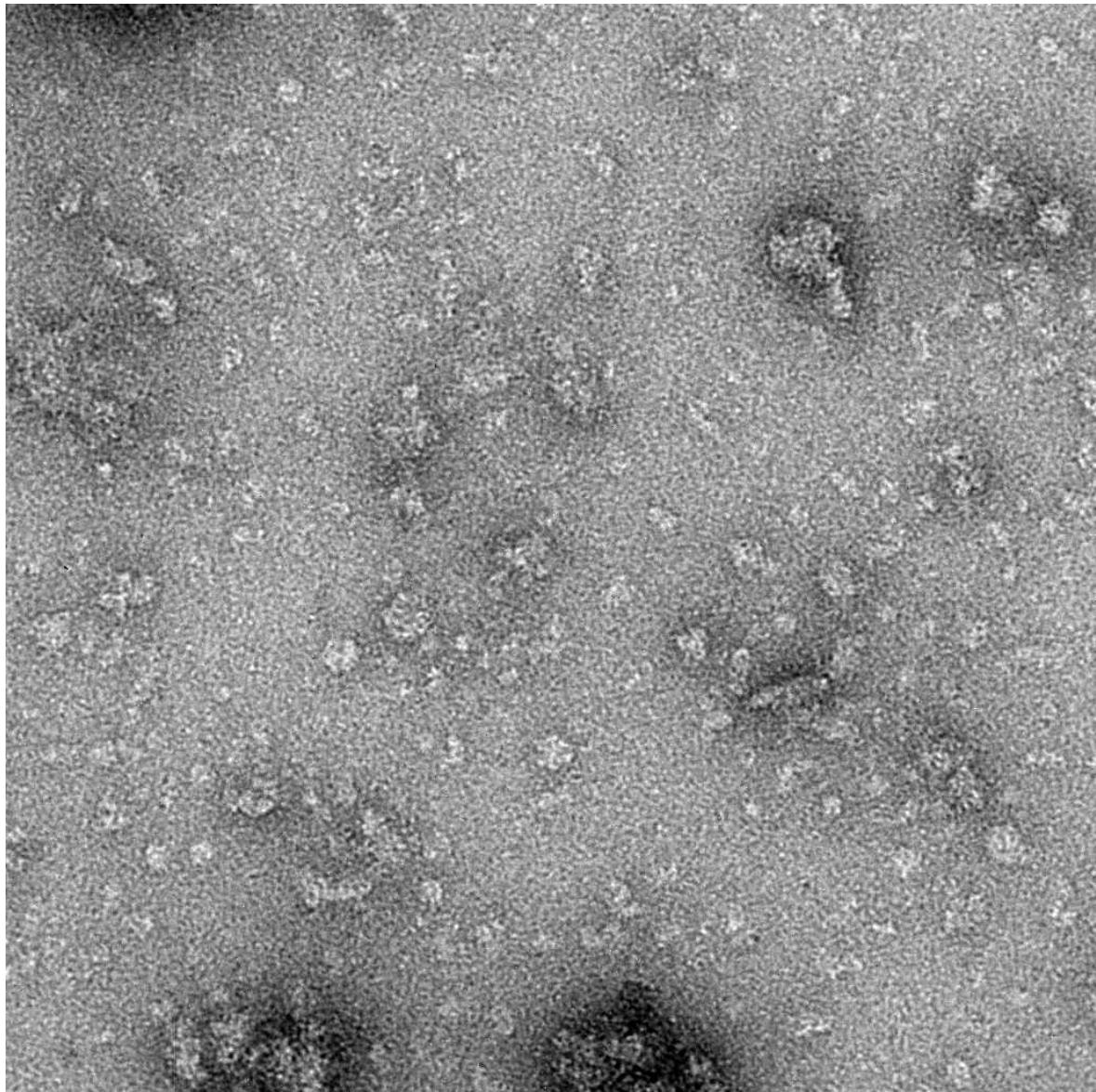
The microscope hums faintly as, inside it, electrons rain down on the sample. For Dan, it's a nostalgic sound. "I did this for years, before my first postdoc. I used to be regularly falling asleep in a dark room with a hum in the background."

The electron beam shoots through the protein sample, and hits a fluorescent screen below, forming a shadow-like image. Dan patiently moves the beam around the grid, hoping to spot a glimpse of a RNA–CBP complex; the molecules are just 30 nanometres across. He's turned up the zoom to 50,000 \times magnification. "You can spend hours flying around in this different protein universe," he says.

Eventually, he spots some promising-looking blobs. They seem to be bigger than CBP on its own, suggesting that he has got CBP and RNA binding together. "This is the first time I've seen it. That's pretty cool," he says.







Blobs in this electron microscope view show a complex of the protein CBP bound to RNA.

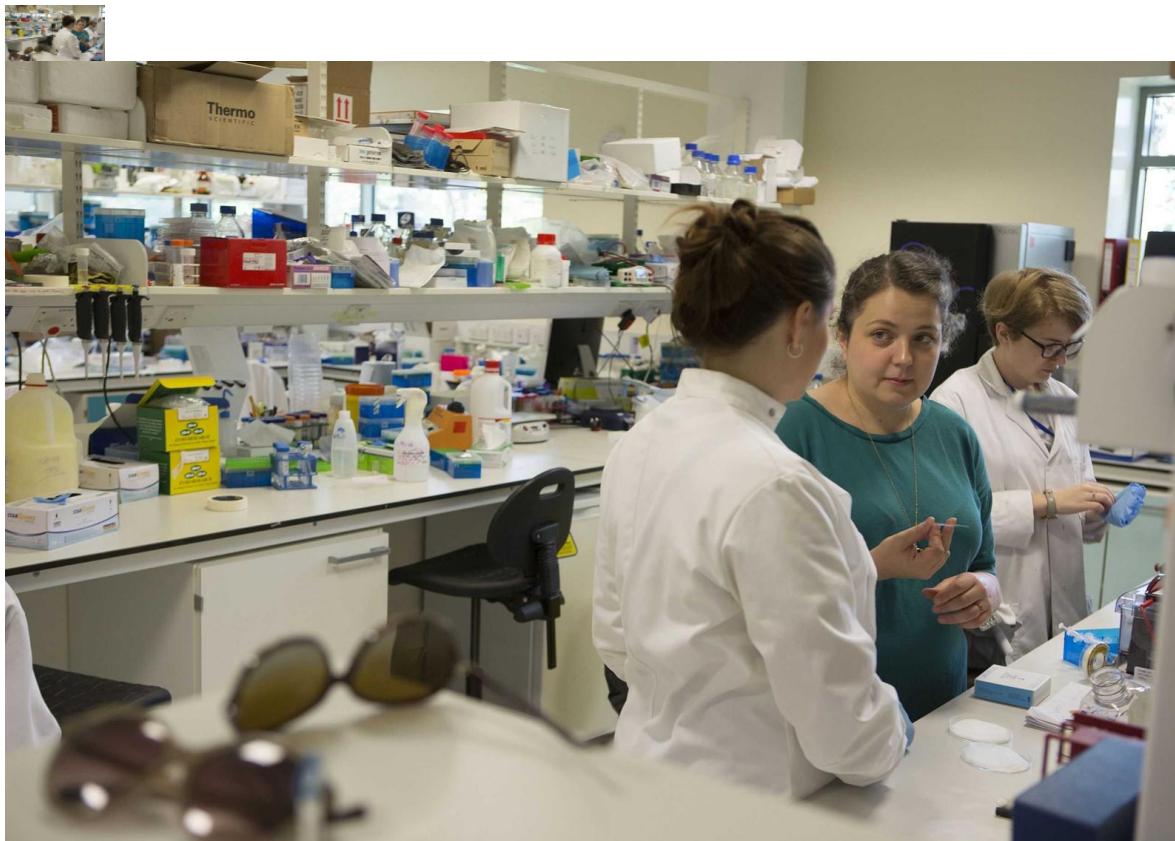
Blobs in this electron microscope view show a complex of the protein CBP bound to RNA.

It's just the start of what will be a long series of imaging experiments to construct the 3D details of the complex that Dan is using his £65,000 FPLC (fast protein liquid chromatography) machine to purify; Dan aims to recruit a postdoctoral student to do most of the work next year. That will involve preparing more samples for use on this machine and the larger, more

sophisticated cryo-electron microscope (cryo-EM) next door. Not one to show his feelings readily, Dan allows himself a smile. It's hard to hear his soft voice over the hum of the machine. "It feels like quite a big deal. This is a sample we can look at, and to build up the structure side of the lab is pretty exciting. It's a nice victory."

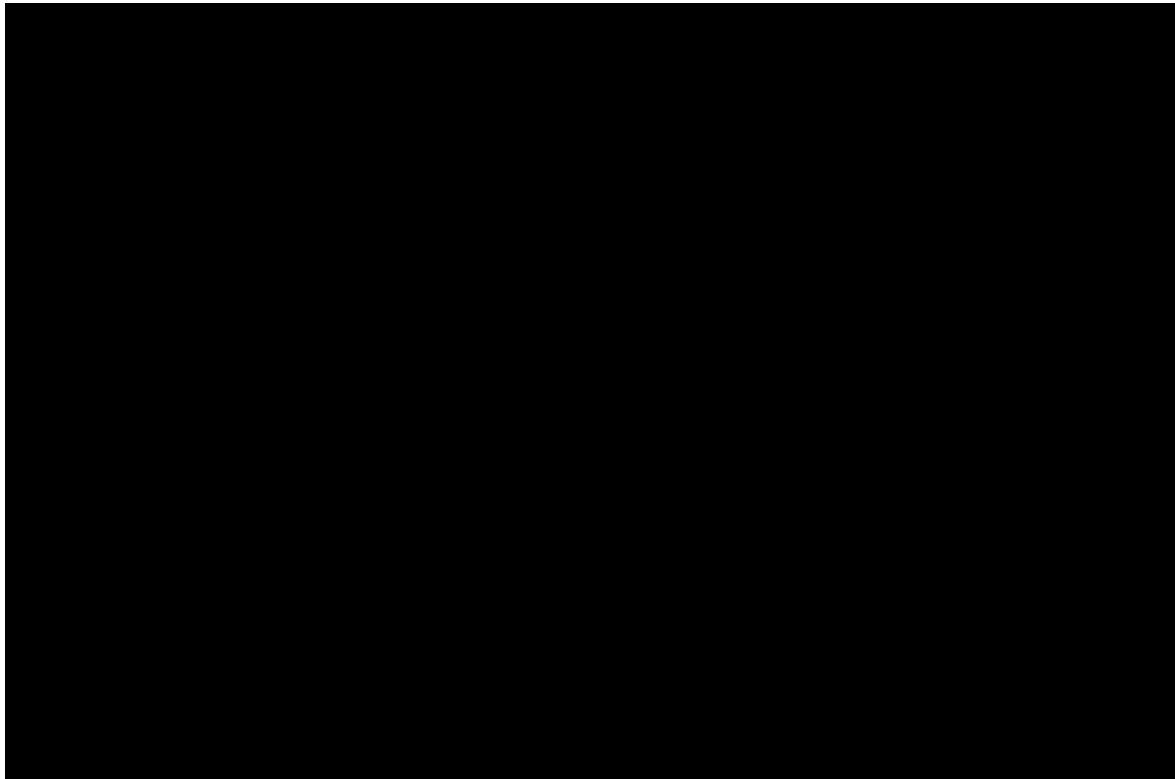
Hear Ali and Dan's story in our 4-part podcast miniseries. Via [Apple Podcasts](#) | Via [Google Podcasts](#) | Via [Spotify](#)

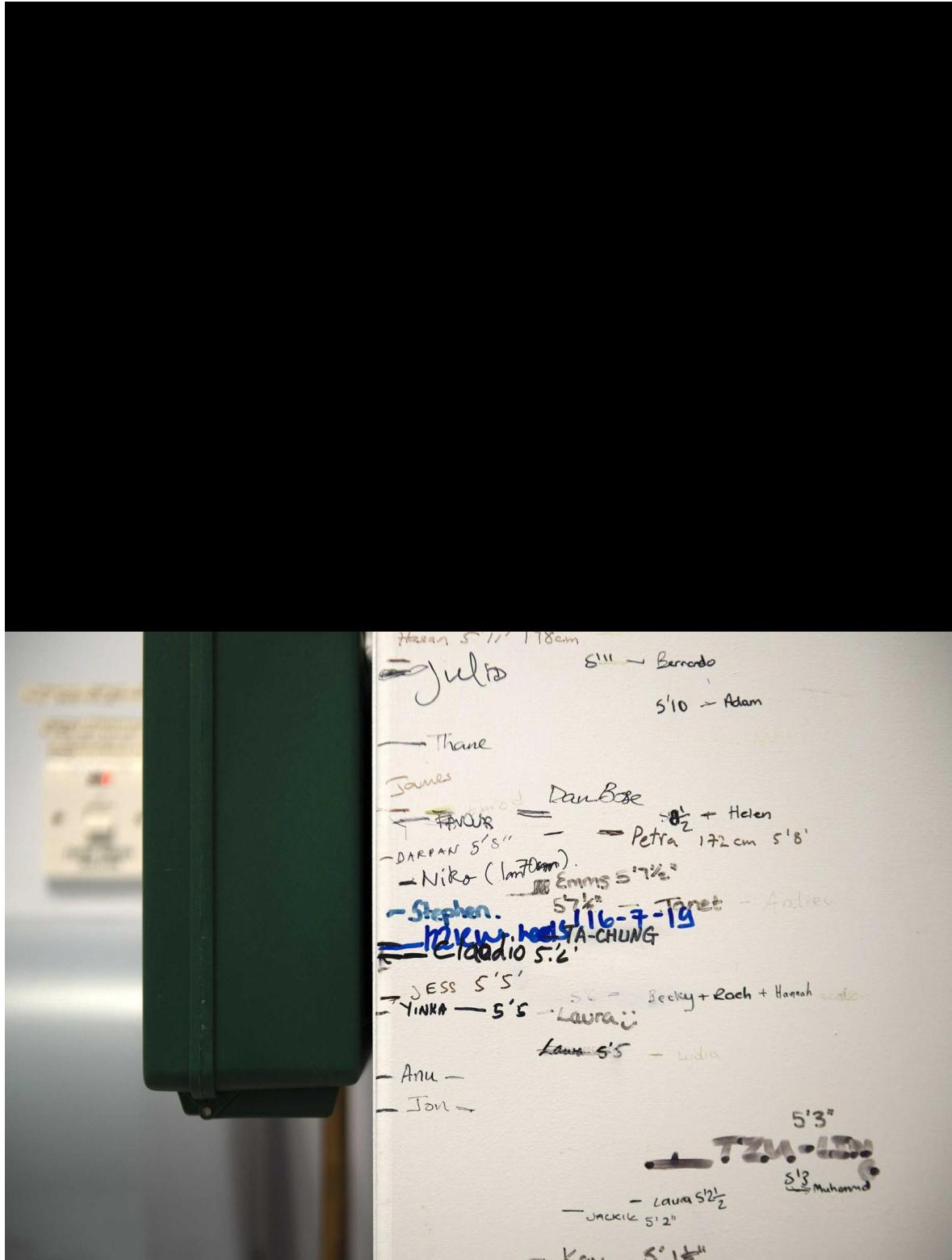
11. Pack mentality



January 2020 ushers in a wave of productivity and optimism. Both labs are teeming with workers; at the work Christmas dinner a couple of weeks earlier, when they collected their own groups together with that of fellow biologist, Emma Thomson, they found themselves catering for more than 20 people, where 2 years ago there were 5.

Dan's students go together to the pub or to the cinema, bring in snacks for their colleagues and share memes on the messaging app Slack. Dan has been concentrating on kickstarting their projects, and they are becoming increasingly autonomous. "Things move without me in the lab," he says. The result is an atmosphere in which the students feel relaxed, and free to bounce ideas off one another, says master's student Luc Costello-Heaven. "I enjoy coming in and seeing everyone," he says.





People who pass through Dan Bose's lab measure their heights on a whiteboard.

People who pass through Dan Bose's lab measure their heights on a whiteboard.

The camaraderie helps them endure the pain of inevitable failed experiments. "If you're having a bad day, if your experiment is not working, everyone's always like, 'It's fine. What's the worst that can happen? We just do it again,'" says Katie Gelder, a PhD student who started in the lab in October 2019. PhD student Petra Celadova is still working towards identifying key regions of the enhancer RNAs in leukaemia cells, but her cell lines are misbehaving and she's spent months trying to sort out what is wrong. She's still hoping to salvage the lines after investing so much time in them.

Petra has enjoyed the work, and her role as mentor to the crew of students now crowding the lab. But as she heads into the last year or so of her PhD, she has already decided against setting up a lab of her own some day. Dan's position seems like several jobs rolled into one. "You're supposed to be a lecturer, you're supposed to be leading a lab, you're supposed to be doing research to collect data for your grants," she says. "And you should be doing all of this full time, but obviously you can't."

Timelapse video of Dan Bose, Katie Gelder and Petra Celadova working in the lab. This video has no sound.

Timelapse video of Dan Bose, Katie Gelder and Petra Celadova working in the lab. This video has no sound.

In the last months of 2019, Ali managed to tick a few big goals off her list. She resubmitted her application, and she finally published a paper she refers to as the 'albatross'. It presented the last tranche of data from her postdoctoral stint, which describe two proteins that team up to help neuronal motors bind to their cargo ([A. E. Twelvetrees et al. J. Cell Sci. 132, jcs215822; 2019](#)). And in her lab, work on the kinesin motor proteins and their different shapes is proceeding apace after a difficult few months in which the samples weren't behaving.

Just as in Dan's lab, Ali's group is close-knit.

“My lot are crazy busy today,” Ali says, when we visit her and Dan in Sheffield at the end of January. “There’s a definite pack mentality.”

Dan chips in. “It’s very pronounced, you’ve obviously built a cohesive team.”

“In a really nice way,” says Ali.

“Just mildly aggressive,” Dan replies. His own group is not so pack-like. They’re collegial, he says, but “they don’t have the same air of intimidation. They might go for lunch together.”

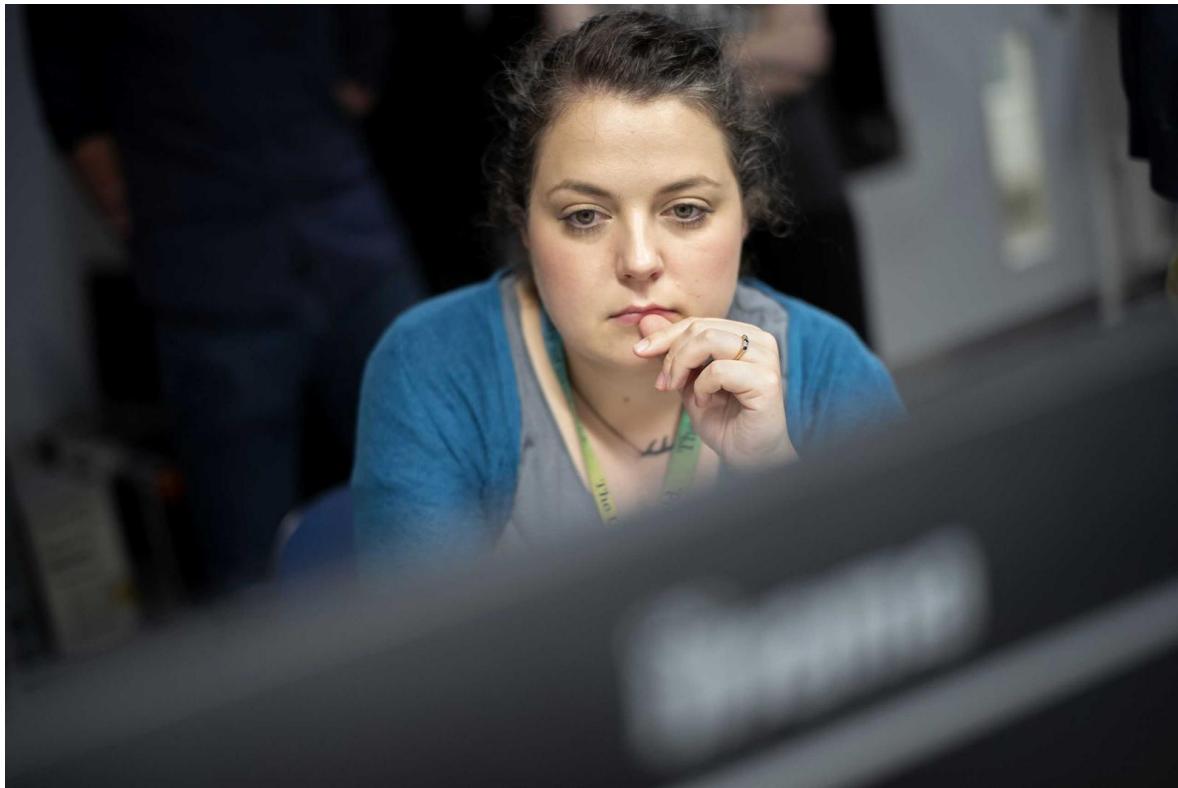
“I’m kind of encouraging it, to a certain extent,” Ali says. “Their confidence is going to feed off each other.”

Ali’s own confidence got a boost in mid-January when she heard that she’d secured a second interview for the Sir Henry Dale Fellowship. But her self-assurance was already quietly growing, and she says that was reflected in her latest application. “Your science comes from your personal creativity. And if you’re feeling not very good about yourself, it’s difficult to stop that leaching into how you perceive what you’re capable of,” she says.

The first time around, “I kept that grant within what I was capable of, and they wanted more from me”. She says she was afraid to ask such fundamental, unaddressed questions about neuronal cell biology as she has in her latest plan, and to prove that she was the scientist who could answer them. “And now I’m like, yeah, fine. You either think ‘I can do it now’ or you don’t, basically, but I’m pretty confident I can do it.”

12. “Such a stereotypical scientist’s nightmare”





The interview is in London on a torrential day in early March 2020. That morning, Ali had woken at 5.30 a.m. from a nightmare in which she had left the lab's -80 °C freezer open. She is disappointed that her sleeping brain wasn't more inventive. "Such a stereotypical scientist's nightmare to have. Such a cliché."

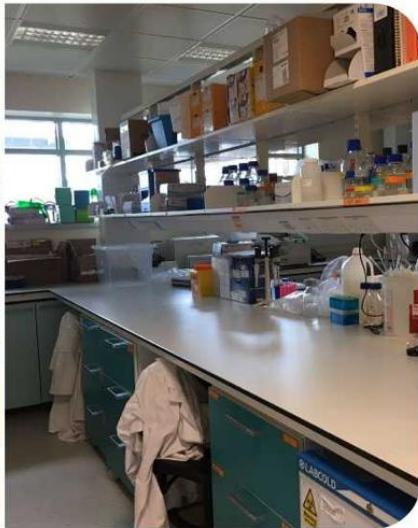
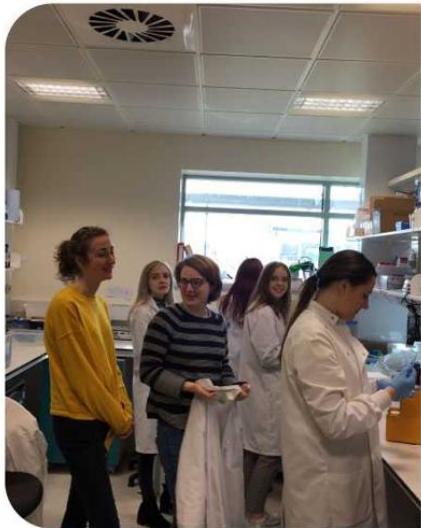
She has memorized [her five-minute talk](#), practised it over and over again and had three mock interviews with colleagues. Her lab members gave her a good-luck card and she is wearing a locket in which there are some strands of her daughter Ada's hair and some of Dan's.

She has left her lab in a state of flux. The coronavirus SARS-CoV-2 has established itself in the United Kingdom; COVID-19 cases are ticking up steadily. The university hasn't yet decided on a plan of action, Ali says, but she and Dan are thinking ahead. "Potentially, they're going to shut the university down, which is what they've done in Italy."

For now, she focuses on the task at hand. What's the first line of her presentation? "My lab studies slow axonal transport," she says without hesitation. "Hopefully, once I go in, it'll all just come out."

And mostly it does. After the interview, she is giddy with relief and pretty satisfied with how it went. “I didn’t say anything that I regret,” she says, her sentence dissolving into high-pitched laughter.

Ali barely has time to reflect on how her interview went before global events intrude. COVID-19 cases continue to rise, and Ali starts talking to her students about the possibility of shutting the lab. “You could see what was happening in Europe,” she says. Finally, on 17 March, the day after UK Prime Minister Boris Johnson made a televised address to advise people to stop ‘non-essential contact’, [she shut the lab down](#). That day, her department e-mailed the same advice. (It would be another week before lockdown began officially.)



T

@



Twelvetrees Lab today. Stay safe everyone x

Stow Met Lab, 2020

Before and after. This is us just over a week ago and our bench today. Stay safe everyone x

1:56 PM · Mar 17, 2020

Before and after: Ali's lab group tweets about the shutdown on 17 March 2020.

"I don't think anybody has begun to grasp how massive an effect this is going to have," she says on a video call the next day.

It's frustrating for Ali to see scientists from large lab groups saying on Twitter that they'll make the best of the enforced break by writing up papers. "If you are sitting on papers like that, you have too much research money," she says, viewing it from the perspective of her own group, in which two PhD students are only half a year into their projects. "It's going to have such a damaging effect."

Ali starts video calls with her team. “I want to make sure they’re coping OK.” And she’s angry that, it seems, universities still expect researchers to be as productive as ever. “This concept that we can continue to work from home while caring for sick people or looking after children. What fantasy land do people live in who can imagine we can do that?”

That day, Dan goes into work to give his lecture to an empty theatre — filmed for the students. “My jokes usually go flat anyway, so no difference there,” he says. The lockdown is necessary, but to both Dan and Ali, it means unknown territory for their science, their lab group and their home lives. Ada has had only six months at school — and she knows it’s going to close soon. “She’s got her red boots on today to keep cheerful,” Dan says.

In the afternoon of 25 March, Ali is making a sandwich for Ada when she checks her e-mail on her phone — and sees the grant decision has arrived. The subject line is the same as in her rejection e-mail. But when she opens up the message, the content is different: “Dear Dr. Twelvetrees. I am pleased to confirm that your application was successful.” She’s won £1.26 million — a grant that saves the lab

Show transcript

Ali Twelvetrees sits next to her daughter behind a desk in their home. They are taking part in a virtual meeting with Ali’s lab group on a laptop.

Twelvetrees Lab

Ali: So, er, I have something important that’s just happened. Er, we’ve got a
[@TwelvetreesLab](#)

Ba **re** **and** **a** **list** **of** **gr** **au** **s**
Ali **laugh** **ago** **and** **our** **bench**
today. Stay safe everyone x
Voice from laptop 2020 Well done!

“This is everything I’ve worked for,” she tells us after a raucous call with her lab group. “We can do all of the things that we’re excited about for five years. That’s the most certainty you ever get in a scientific career. It’s like winning the science lottery.”

“It’s not just the science. It’s the security that it brings as well — it’s so precious. To have me and Dan both have that now. We’ve never had that before.”

13. How to run a lab during a pandemic



Two months into England’s first lockdown and with schools closed, Ali and Dan have forged a routine: one spends the morning with Ada, the other the afternoon. With Ada they work on phonics, draw birds and grow crystals. At work, they are coping with an endless stream of online meetings, while trying to ensure their students continue to progress — and stay in good mental health. Dan is also teaching, grading papers and holding virtual office hours for undergraduates.

It’s impossible to do it all well, says Ali. She struggles to focus with Ada underfoot, even when Dan is on teaching duty. “Her voice, it just shuts my brain down completely,” she says.

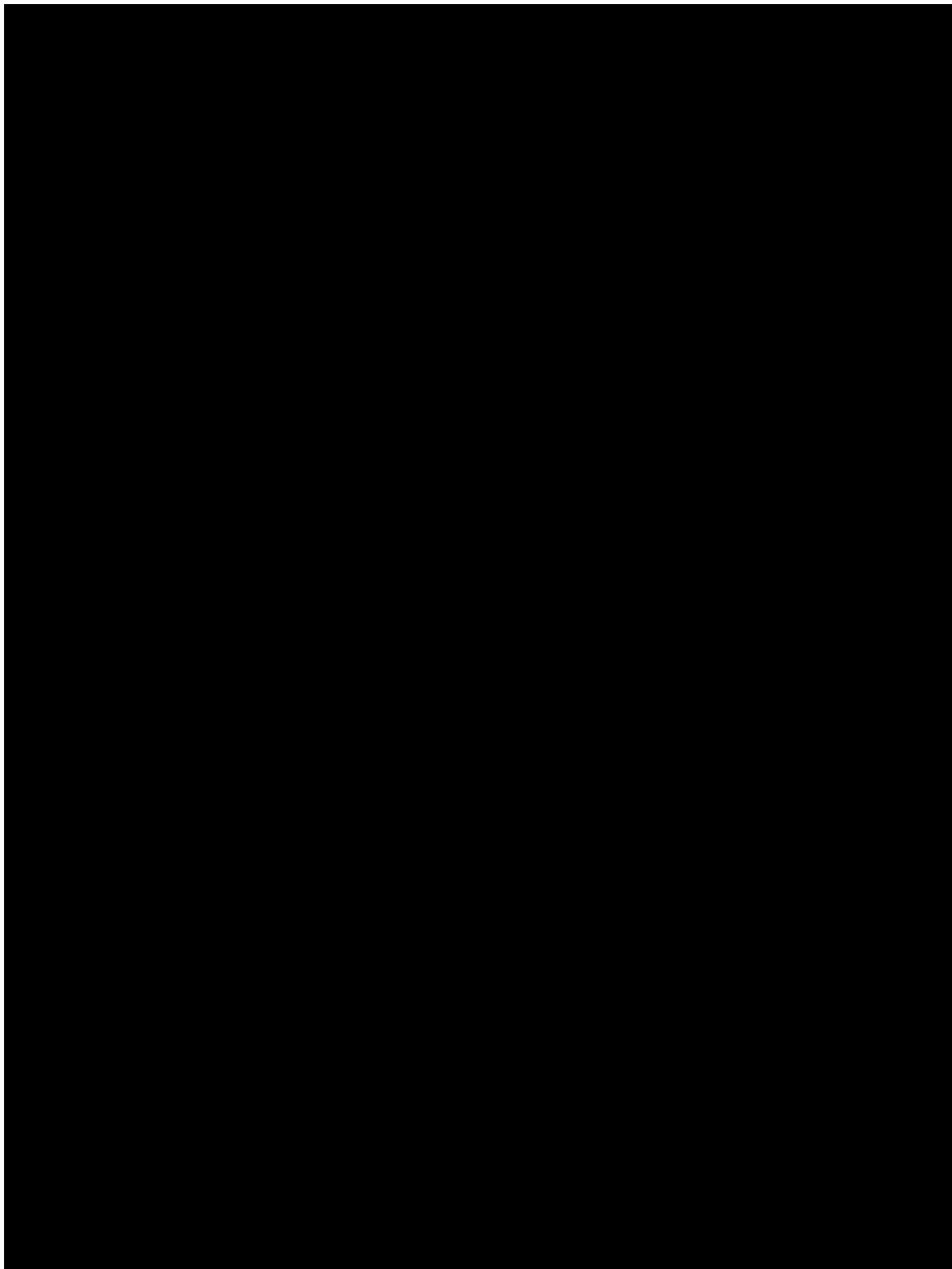
Both have tried to shift their lab members on to work that can be done at home. Ali ticks through her students' projects on a call: Evie Smith has been developing an automated data-analysis pipeline; Ashleigh Davey and Emma Turner are doing database searches ... Ali trails off. "Ada, could you just be quiet for a minute?" she says calmly. Ashleigh and Emma are searching databases to gather data to back up their hypotheses, so that they can return to the lab focused and ready to push their projects forward. Nicola Carruthers, one of Dan's students, has taken a leave of absence to volunteer in a coronavirus testing lab.

"They're so resilient," Ali says of her team. But each student has a strict window of time to complete their degree, and Ali knows that the uncertainty about how the pandemic will affect their education is adding to their lockdown stress. Students in both their labs are worried about being unproductive. For Dan and Ali, seeing other scientists brag on social media about attending online conferences and writing grants is frustrating when their major goal is to look after Ada and keep their students afloat remotely.

But there has been good news. Dan has passed his probation and switched from a fixed-term to a more permanent contract with the university. Although the criteria for promotion technically included publications, which Dan has not yet had, he was able to demonstrate that he had made progress towards publishing, and had met the other two criteria of teaching and bringing in funding.

Ali still bristles at mention of the publishing requirement. "Three years is not enough," she says. "I think if you're, you know, a theoretical physicist, right, you don't have to set up a lab or work at the pace at which things grow."

In Ali's case, the criteria are not as clear and the process is not as formal, but she has learnt that her department has shifted her to an open-ended contract and included her in its financial projections, suggesting that her lab will be able to stay.





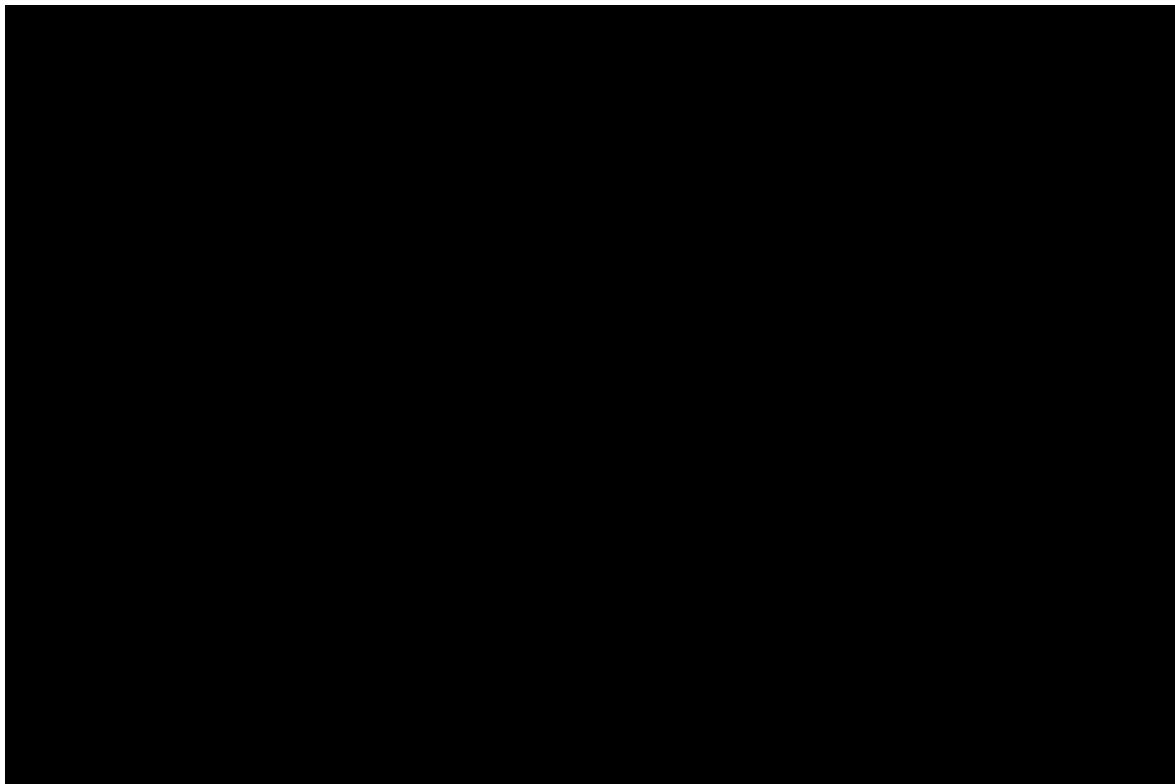
Dan with PhD student Nicola Carruthers.

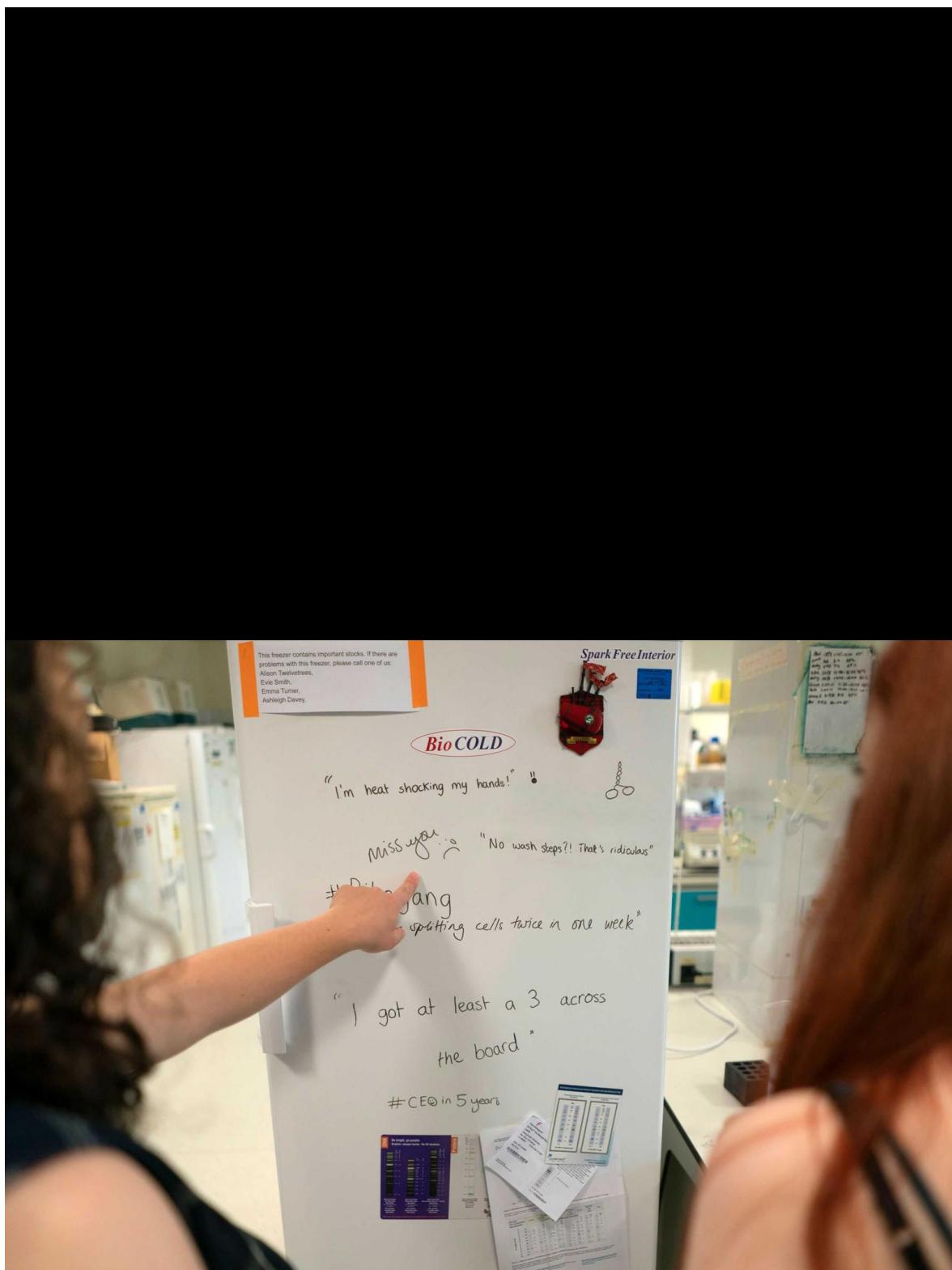
Dan with PhD student Nicola Carruthers.

When the labs reopen in the summer, Dan and Ali's tidy schedule is demolished. Their household routine is continually shifting as their work conditions change. Dan's lab is in a classical university building, with long corridors and closed doors; Ali's is in an open-plan building with shared spaces and a large atrium — their return to the lab is markedly different. Dan is able to bring back his students, and they work almost normally, minding the new social distancing and safety protocols.

For Ali, however, lab work is a logistical nightmare. The students are assigned to shifts, some working in the morning, some in the afternoon. This means that an experiment started in the morning must now be handed over to someone on the afternoon shift to finish. It is a level of collaboration, planning and restricted movement that leaves no margin for error, and progress has slowed as a result.

Some experiments remain off-limits: Ali's lab can no longer access the shared microscopes in other buildings on campus. And perhaps hardest of all is the physical separation, which makes it difficult to recapture the collaborative atmosphere that Ali had so cherished.





Ali's students wrote jokes and messages on a freezer door as they worked shifts in the lab.

Ali's students wrote jokes and messages on a freezer door as they worked shifts in the lab.

"At the start of the year, I felt like we'd finally turned a corner and started to build momentum," she says. "People were motivated, experiments were working. There was an energy."

Ali thinks back to her pre-pandemic lab and struggles to hold back tears. "Mostly I'm crying because I just miss them. For that brief window it was really nice to work with them and support them and see everything coming together."

But she is still moving forwards. At the end of October, Ali sends e-mails to start the release of her grant money, allowing her and the team to begin planning purchases. They buy a couple of small pieces of bench-top equipment to make lab life easier and save them having to move between rooms.

Lockdown has disrupted Dan's plan to recruit a postdoc to run the cryo-EM prong of his work, but in October, he finally puts out a job advert. It's a "big deal" to start this work, to be able to look at the 3D structure of the RNA-CBP complex and see how it is organized and how it behaves. The rest of the lab has weathered the storm relatively well — Dan keeps on top of projects through Slack and a morning video meet-up, and the students support each other when he isn't in the lab. He is delighted that they decided on their own to rearrange and organize the lab fridge, even if some of his old stuff got thrown out.

The lab seems to have changed gear. In March and December, Petra and Dan's names appear on papers for work led by other researchers to which they contributed some data and experiments. A new PhD student has just joined. And Petra — his first-ever PhD student — is starting to think about writing up her project and could soon be moving on.

14. "It'll be fine. Eventually."





On 3 December 2020, Dan is wrapping up work when Ali texts him: “Ugh miserable weather. Are you late home again?” Dan replies that he plans to leave around 6 p.m.. But he never shows up.

What happened that evening is still a bit fuzzy, even now, so many months later. Dan was cycling home along his usual route. But he either fell off his bike or dismounted as he passed a children’s hospital. A team of doctors walking out of the hospital found him having a seizure and diagnosed him on the spot: a brain haemorrhage.

An ambulance arrived quickly and took Dan to the university hospital, where doctors placed him in a coma. “They thought he wasn’t going to make it,” says Ali. “His condition going into the hospital was so bad, and that’s correlated with how well you recover.”

Because of pandemic restrictions on hospital visits, Ali was allowed to see Dan for only a short visit before he was transferred to the neurology department. “My department,” Ali says with a sharp laugh. “But I wasn’t allowed to cross the threshold into the ward.”

Dan was in a coma for 24 hours. After that, he gradually spent more and more time awake each day. Ada sent him lots of messages and pictures, practising her handwriting. Dan chatted with Ali by phone, but only after she had ensured he couldn't use most of the apps.

"She disabled everything except FaceTime," says Dan. "So that I couldn't check work e-mails." While Dan made optimistic plans for his return to work, Ali was haunted by the statistics: only about half of people with Dan's condition survive long enough to leave hospital, and even then they face varying degrees of disability.

But Dan was one of the lucky ones: he went home on 22 December. His haemorrhage had been caused by a weak blood vessel in his brain, called an aneurysm, which had burst. Dan had had a headache for weeks before the rupture, but had dismissed it as being the result of too much coffee, or stress, or screen time.

His medical team scanned his brain to find out if he was at risk of a repeat haemorrhage, and found nothing.

"In the doctor's words: my brain is 'pristine,'" he says with a mischievous smile.

"Dan brings that one out on a regular basis," says Ali.

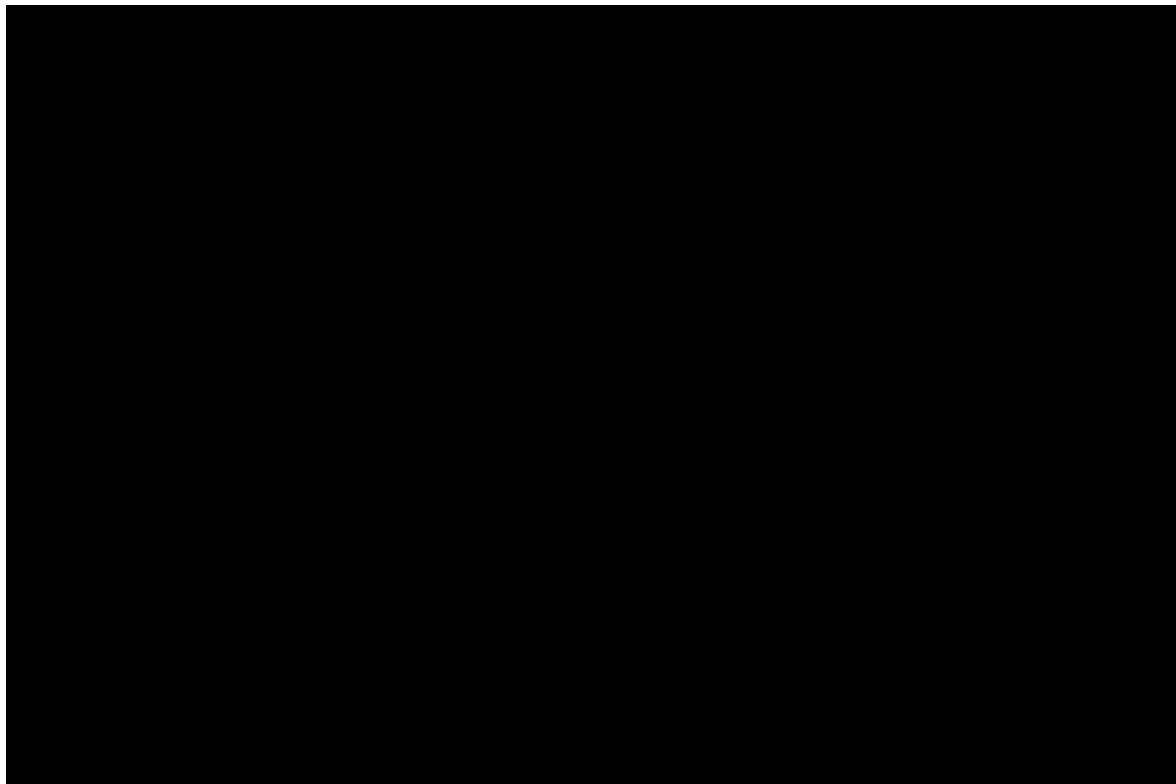
After the holidays, another COVID-19 lockdown closed the schools, so Ada was at home, too, wanting to check in on Dan when he was trying to rest. "She would tiptoe up," says Ali. "But she wasn't very quiet."

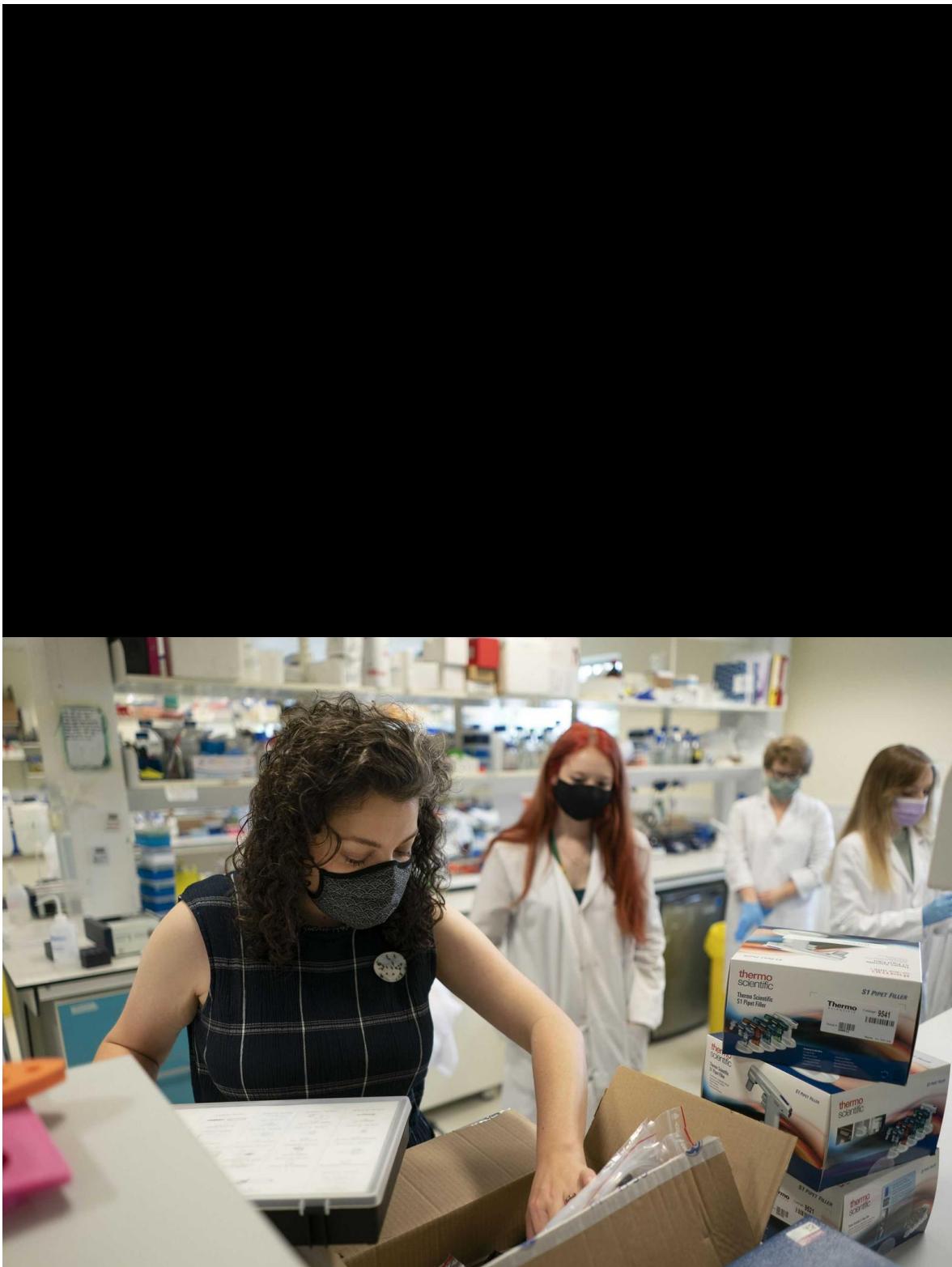
While Dan recovered, Ali and Emma worked to keep his job responsibilities at bay. Colleagues in the department took over his teaching, and Emma helped to supervise his team's day-to-day work. Ali kept a close eye on Dan to make sure he gave himself the time he needed to recover.

"He was planning all of these things, and I was like, 'you know you can't walk?'," says Ali. "He would say, 'It'll be fine. Eventually.'"

"Medication is a powerful thing," says Dan.

Gradually, Dan was able to do more and more. He checked in on his students over video calls, and selected a postdoc from applications he had received just before his accident. He interviewed candidates in February, and selected a new recruit to start in May.





Ali unloads new lab equipment.

Ali unloads new lab equipment.

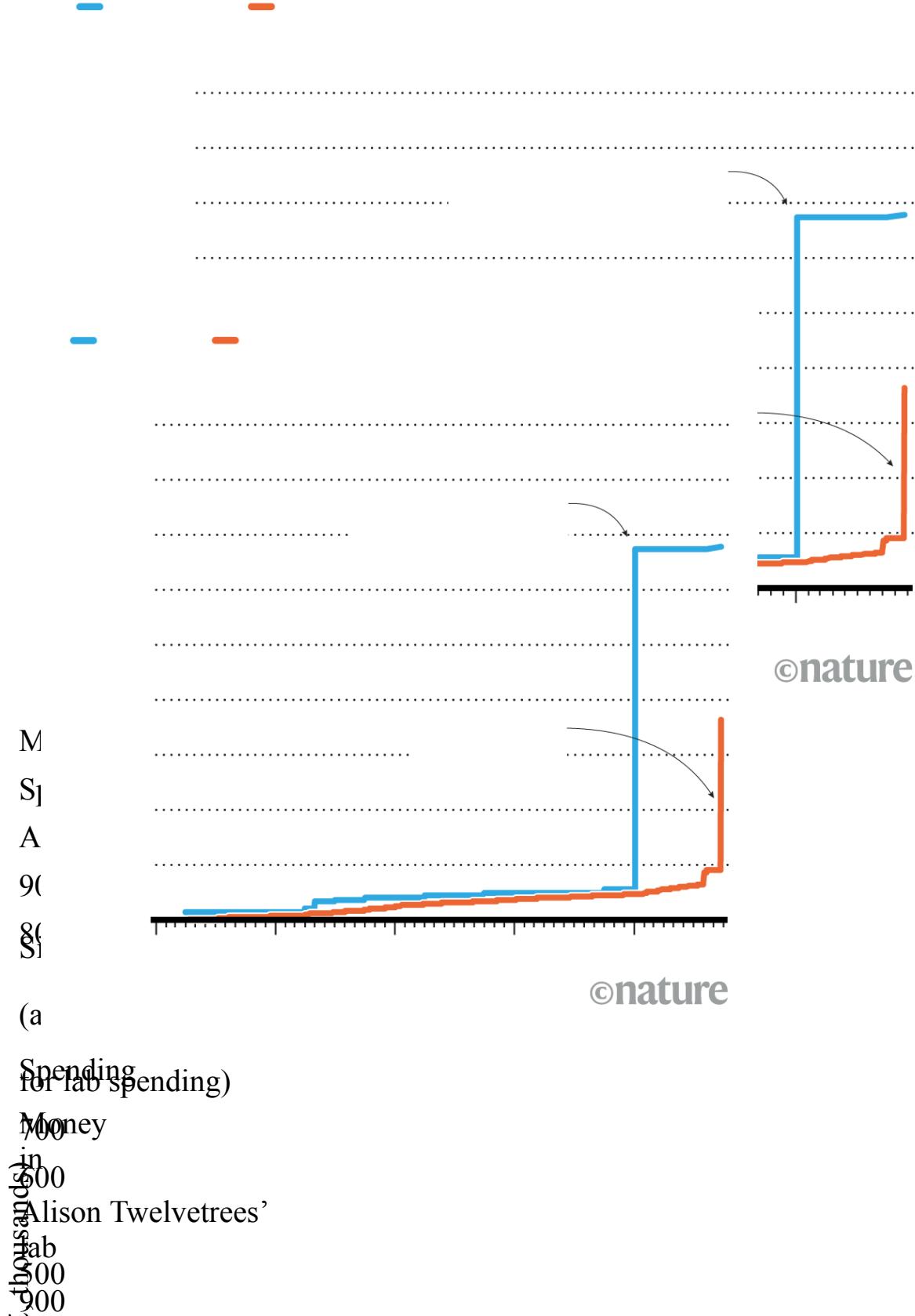
As Dan recovered, Ali was able to turn back to her own work, pushing her own students' projects forwards and making some changes to their shift patterns to enable a bit more overlap and make everyone happier.

By the summer, she had begun making plans to spend her grant money: she had finally found a place in her institute for the expensive fluorescence microscope — an act of diplomacy more than design — and begun the “miles of paperwork” its purchase would require (see ‘Finally funded’). The total cost would ultimately come to £272,906. “Sort of weird to think it costs the same as a house,” Ali says.

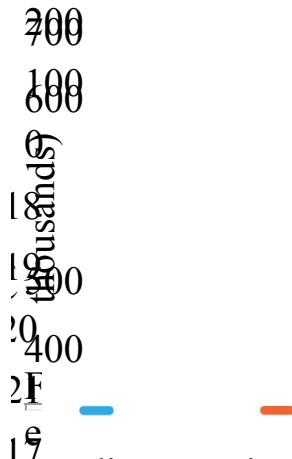
Meanwhile, using the fluorescence resonance energy transfer (FRET) microscope, her team had made some intriguing discoveries about their motor proteins: they seem to exist in more than just ‘on’ and ‘off’ states, which could have implications for how they do their jobs carrying cargo in neurons. “Still planning that slightly elusive, but hopefully closer, publication,” she says.

Finally funded

With her grant secured, Ali can at last order the microscope of her dreams. (Chart excludes staff salaries).



£100
Sir Henry Dale
award
microscope
(amount available
£272,906)
£100 lab spending)



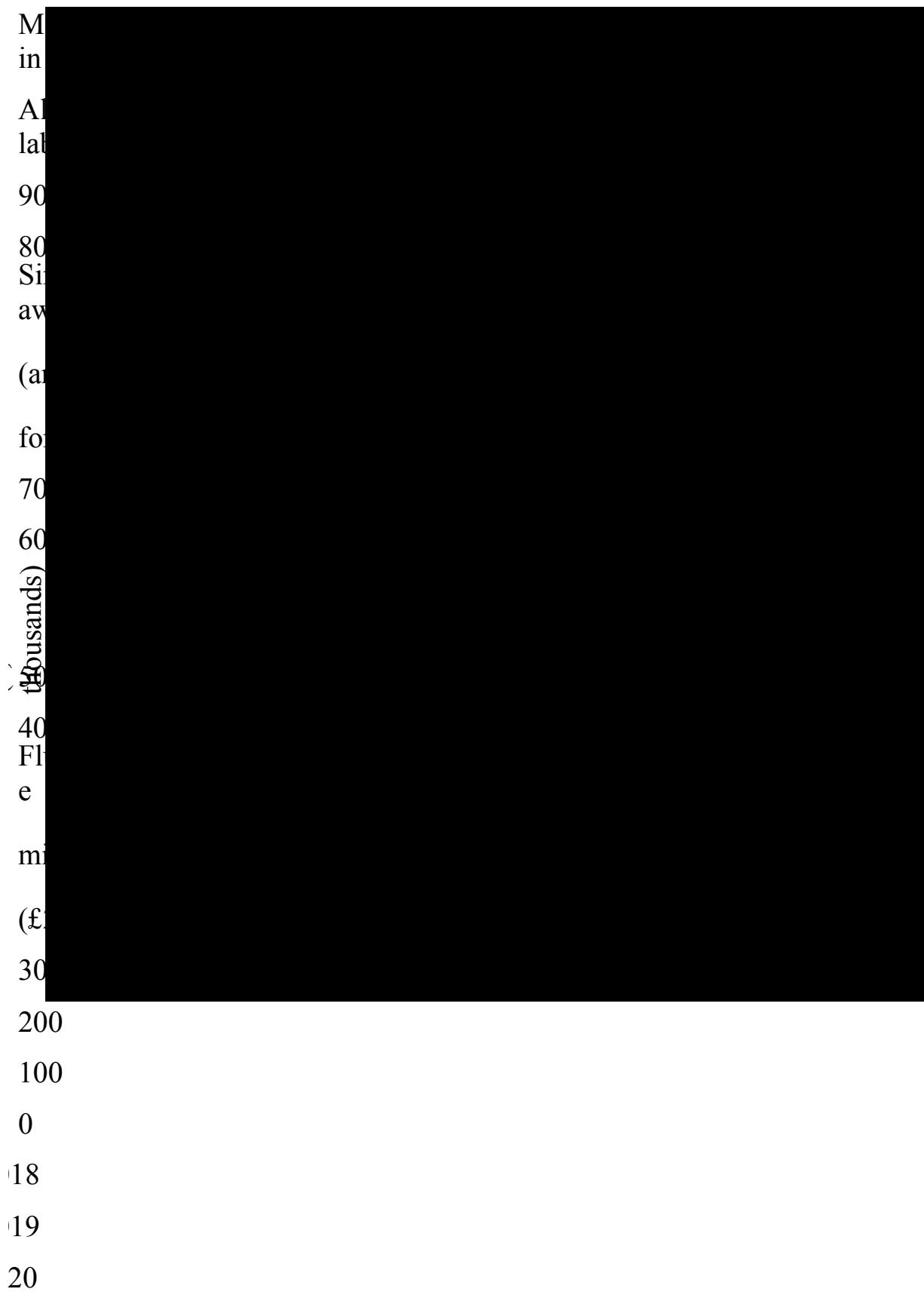
¹Funding remained a nightmare — literally. One weekend in June, Ali woke up in the middle of the night after an oppressive dream about losing her job. “Persistent career insecurity over many many years leaves anxiety scars,” [she tweeted](#).

³A few months earlier, she had asked fellow new PIs on Twitter how many times they’d thought about quitting academic science in the past year. More than one-third said they’d done so at least every week.

¹On June, about six months after his haemorrhage, Dan still got tired doing tasks that were routine before, and one of his feet would go numb if he forgot to do his physiotherapy. He worried about how he would deal with large crowds once the lecture theatres were open again.

²⁶But he was improving quickly, and both he and Ali were thinking about the future and the never-ending quest for research funding. “It feels like there’s this sort of balance between doing work that’s related to the fellowship and then doing work with half an eye down the road to where the next fellowship or the next bit of funding is going to come from,” Dan said.

¹es)
Spending



21

17

*£

(a

20



Dan and Ada on a holiday in Durness, Scotland, where Dan grew up.

Dan and Ada on a holiday in Durness, Scotland, where Dan grew up.

It has now been more than four years since Ali and Dan arrived in Sheffield. If she could give advice to her younger self, Ali says, she would have suggested nurturing more faith in her own abilities. “I can see how little confidence I had in myself and my ideas when I started, and how paralysing that was in terms of being able to write anything that someone was going to fund,” she says.

Dan says that, at the beginning, he beat himself up about all the things he wasn’t getting done. But, in retrospect, he can see that work in his lab was progressing — just more slowly than he’d hoped. “First off — know that slow is OK,” says Dan. “I had expectations of just carrying on at postdoc speeds, and, actually, that’s not healthy. And it wasn’t possible.”

In September, the UK school year starts. Ali and Dan are searching for their routine: they alternate week by week who picks Ada up after school so that they can each count on some evenings to work late.

Dan is once again the first one in the household to wake up each morning, which Ali finds comforting. “The weirdest thing for me after the aneurysm

was how much he slept,” she says.

Although he will be returning to teaching in a few weeks, Dan keeps his early mornings sacrosanct. He gets up, feeds the cats, brings a sleeping Ali some tea and sits reading a book until Ali or Ada wakes up. He tries to avoid checking e-mails.

He now tries to make sure that he has time set aside each day to think about his science, and he guards his free time much better. “It’s not normal to wake up in hospital after that and your first instinct is ‘I’ve got to get back to work because I’ve got teaching in a few weeks’,” says Dan.

Ever since his stint in the hospital, he can separate from work more easily. “I will often turn off the phone over the weekend and play with Ada for longer,” he says. “I feel much lighter now.”

Nature **597**, 618-623 (2021)

doi: <https://doi.org/10.1038/d41586-021-02565-9>

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

Starting up in science

Nature follows two researchers over three years as they struggle to launch their own labs.



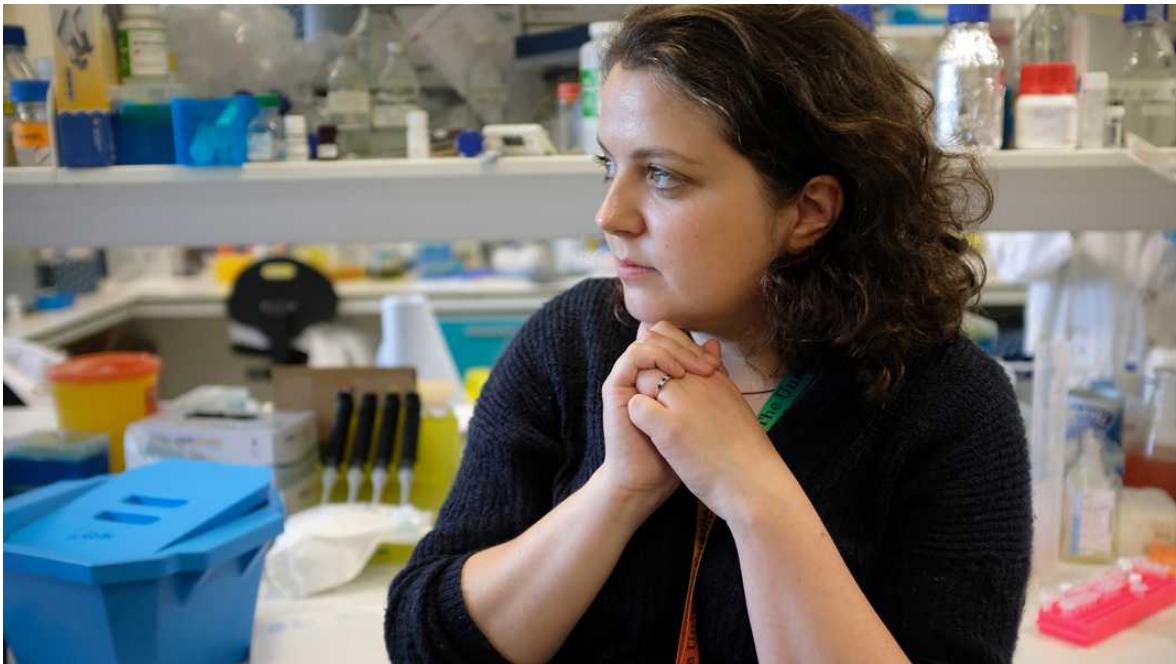
Part 1: The pitch

Ali and Dan face a series of challenges as they begin their journeys as lab leaders.



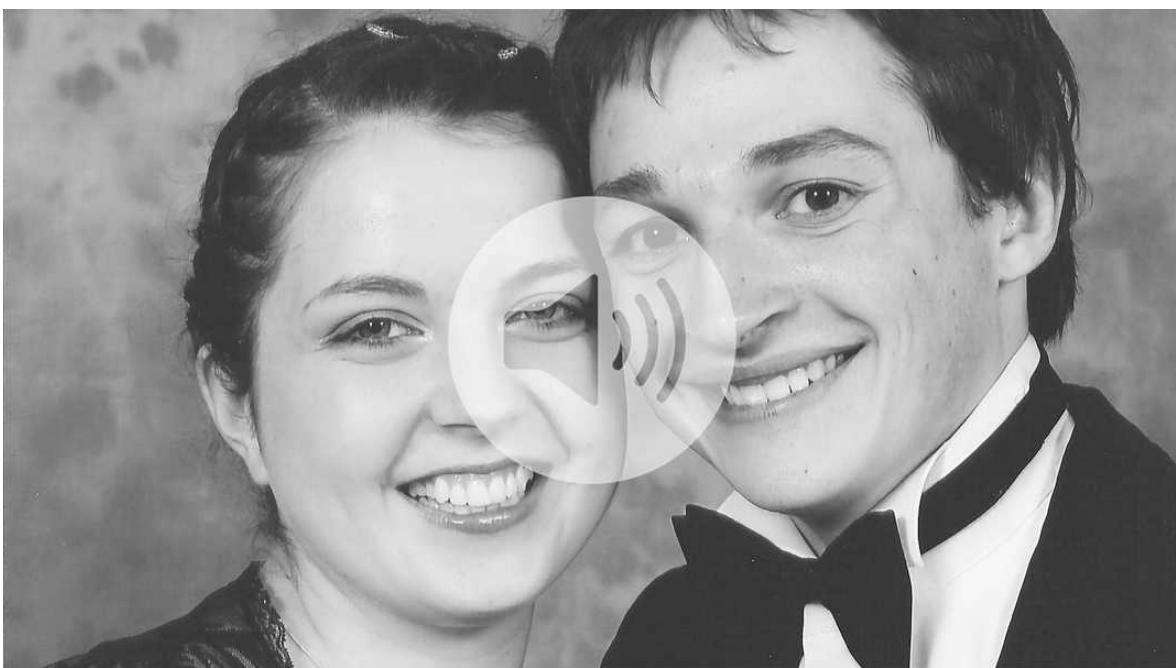
Part 2: The quest for cash

Ali and Dan race against time to secure stable funding to keep their labs going.



Editorial

What a personal science saga reveals about funding — and science itself.



Podcast

Get to know Ali and Dan better in this 4-part series telling their story.

Authors: Kerri Smith, Heidi Ledford, Richard Van Noorden

Additional Reporting: Benjamin Thompson

Design: Lizzy Brown, Wesley Fernandes, Kelly Krause

Original photography: Chris Maddaloni for *Nature*

Subeditors: Anne Haggart, Anna Callender

Podcast editors: Benjamin Thompson, Kerri Smith

Editor: Richard Monastersky

Project leader: Kerri Smith

Personal photos provided by Alison Twelvetrees and Daniel Bose.

Related articles



How junior scientists can land a seat at the leadership table



How new principal investigators tackled a tumultuous year



These labs are remarkably diverse — here's why they're winning at science

SPRINGER NATURE

Springer Nature

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

© 2021 Springer Nature Limited. All rights reserved.

[Top](#)



[Built with Shorthand](#)

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02565-9>

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

Books & Arts

- **Transformations in medical care were never easy or safe**

[27 September 2021]

Book Review • How innovation in drugs and surgery can take a toll on patients and public confidence.

- **From gold to Bitcoin and beyond** [28 September 2021]

Book Review • An economist envisions a mostly cashless future, which could make finance more inclusive.

- BOOK REVIEW
- 27 September 2021

Transformations in medical care were never easy or safe

How innovation in drugs and surgery can take a toll on patients and public confidence.

- [Hannah Wunsch](#) 0



Some of the first US children to receive the eagerly awaited Salk polio vaccine in 1955 were in San Diego, California. Credit: Bettmann/Getty

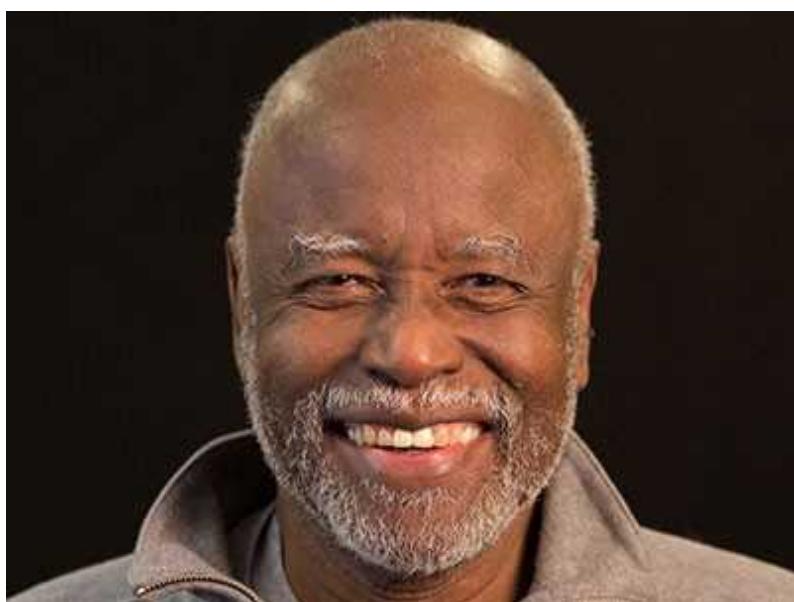
You Bet Your Life: From Blood Transfusions to Mass Vaccination, the Long and Risky History of Medical Innovation Paul A. Offit Basic (2021)

Author Paul Offit sits on the US Food and Drug Administration's COVID-19 vaccine advisory panel. An international expert on vaccine development, he co-invented the rotavirus vaccine. He has written many books, some prophetic — such as his 2010 treatise *Deadly Choices: How the Anti-Vaccine Movement Threatens Us All* — and is an attending physician in the Division of Infectious Diseases at the Children's Hospital of Philadelphia in Pennsylvania.

In short, he thinks a lot about the risks and benefits of medical care.

His latest book – *You Bet Your Life* — couldn't be more timely. Using the stories of key patients, he weaves a narrative of incredible transformations in medical care that were not easy or safe linear progressions. Charting some of the “greatest advances”, Offit focuses on the long history of mistakes, negligence, paternalism, lack of oversight and the many casualties along the way, including individuals and public confidence.

The book breaks down key developments into three parts, traversing a wide range of medical breakthroughs — from heart transplants and anaesthesia ('Risk') to antibiotics and X-rays ('Oversight') and, finally, chemotherapy and gene therapy ('Serendipity').



Bill Carter Jenkins (1945–2019), Tuskegee whistle-blower

The stories are riveting and filled with fascinating details, many from the days before ethics review boards. Offit writes of slightly crazed doctors and dentists playing with anaesthetics; surgeons putting hearts into people with little ability to stop the organs being rejected; and potentially life-saving treatments gone wrong.

A poignant example is the story of folic acid antagonists in chemotherapy. Paediatric pathologist Sidney Farber (founder of the Dana-Farber Cancer Institute in Boston, Massachusetts) read about work in mice published in 1945 that demonstrated a miraculous reduction in cancer through the use of “folic acid”. The words were carefully kept in quotation marks in the early research publications because no one was sure what the extracts being used really contained.

In Farber’s rush to help 11 paediatric patients who were terminally ill with acute lymphocytic leukaemia, he treated them with a purified form of folic acid that he thought would be safer. This was before the bench scientists discovered that the extract they had used to treat mice was, in fact, not pure folic acid, but an antagonist that blocks its effects. Farber had inadvertently hastened the deaths of all 11 children.



Vaccines — lessons from three centuries of protest

In ‘Risk’, Offit describes the desperation of individuals such as Louis Washkansky. Dying of congestive heart failure, Washkansky’s only hope was a new heart. But no human had ever received such a transplant. Out of other options, he chose to go ahead with the surgery by Christiaan Barnard in 1967. Washkansky lived for just 18 days. He was at least aware of the experimental nature of the treatment. Others were not, such as Hannah Greener, a girl who died in the 1840s after being given chloroform, an anaesthetic that was only just coming into use.

It is in the section titled ‘Oversight’ that Offit feels on surest footing, discussing antibiotics and vaccines. In the context of the COVID-19 pandemic, the story of how other vaccines were developed is compelling. Offit centres his narrative around Anne Gottsanker. In 1955, aged 4, she received the poliovirus vaccine that Jonas Salk had just developed for a world desperate for reprieve from the devastating disease.

In a heartbreakng description, she remembers “lying in the hospital and not being able to move anything”. The Salk polio vaccine required the virus to be killed as part of its production. Through a lack of government oversight and sloppy manufacturing as vaccine production was scaled up, Anne was one of 70,000 people inadvertently given live polio virus. The tragedy, called the Cutter incident after the company that made the faulty doses, left 164 people severely paralysed and 10 dead. (Offit wrote a 2005 book on the episode and on how vaccine safety subsequently improved.)



The poisonous history of chemotherapy

The final section, ‘Serendipity’, traces work that has led to chemotherapy and gene therapy. Offit is a master at braiding multiple threads of a story, describing the fits and starts of research and the unexpected findings that lead to new treatments. He describes the first person whose tumour was successfully treated with a folic acid antagonist in 1946–47: Babe Ruth, the US baseball star. Such a footnote is a reminder that new drugs often go to the well-connected (recall the cocktail of drugs Donald Trump received as US president when hospitalized with COVID-19).

In the pursuit of exuberant storytelling, some points get exaggerated, and a few facts and dates are off. For example, Offit calls heart transplants today “as common as bypass surgery”, which is far from the case. The choice to provide a bibliography for each chapter without specific referencing makes it challenging to track down and read more about intriguing asides, such as the statement that couples in an unspecified country were barred from marrying if they had different rhesus blood types.

Offit’s most important messages are contained in a short epilogue. Here he reflects on the downstream consequences of unanticipated tragedies that can follow new discoveries. Harm to a few can result in harm to many, if the erosion of trust means that people opt out of preventions or treatments. For example, in 1956 in the United States alone, polio paralysed 15,000 unvaccinated people because of vaccine hesitancy in the wake of the Cutter incident.

Sadly, medical history is repeating itself. As Offit concludes, people are much more fearful of doing something that has a minuscule chance of causing them harm, such as having a COVID-19 vaccine, than they are of not doing something that is likely to result in harm — such as remaining unprotected from a virus that has killed millions of people in less than two years.

Nature **597**, 625–626 (2021)

doi: <https://doi.org/10.1038/d41586-021-02616-1>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02616-1>

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

- BOOK REVIEW
- 28 September 2021

From gold to Bitcoin and beyond

An economist envisions a mostly cashless future, which could make finance more inclusive.

- [Paola Subacchi](#) 0



A demonstrator in San Salvador protests against El Salvador's adoption of Bitcoin as legal tender. Credit: Camilo Freedman/Bloomberg/Getty

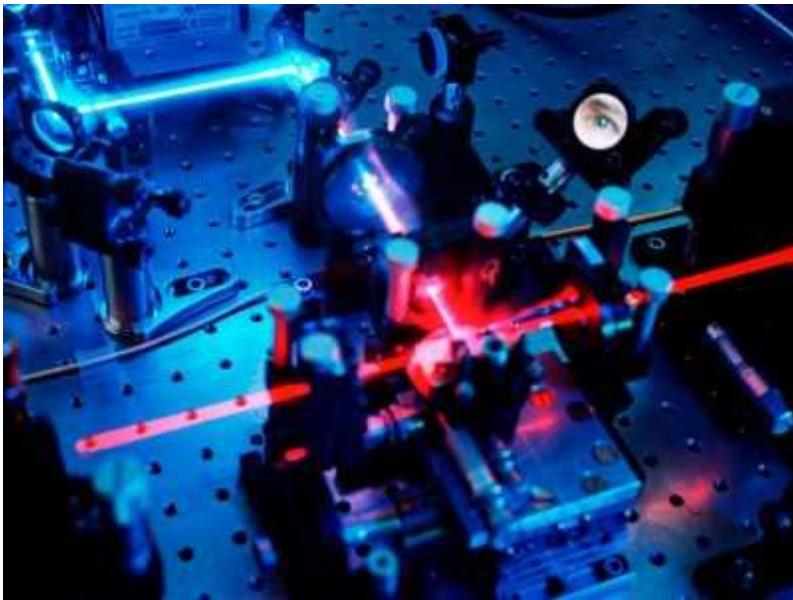
The Future of Money: How the Digital Revolution Is Transforming Currencies and Finance *Eswar S. Prasad* Belknap (2021)

Looking to understand why citizens have taken to the streets over the adoption of Bitcoin as legal tender in El Salvador? Enter economist Eswar Prasad, with his thoroughly researched book *The Future of Money*.

Innovations in payment systems and currencies, he explains, reflect both the significant changes in the world economy as it integrates large developing countries, such as China, and the all-encompassing digital transformation. As money has become free to move around the world, Prasad probes how digital innovation is reshaping it as both a tool and a concept.

Two things epitomize the transformation. The first concerns money as a physical entity. The digitalization of payments challenges the idea of cash as the most definitive form of money. Cashless transactions, however, entail a loss of privacy because they leave electronic traces. This does not seem to be a problem for the majority of people and businesses that make such transactions legally. And requiring digital payments for transactions above a certain amount would help to shrink the off-ledger economy. In Italy, for instance, where this shadow economy amounts to about 15% of gross domestic product, payments of €2,000 (US\$2,350) or more can no longer be settled in cash.

The governance of data collected through digital transactions remains an open question, one that cannot be dismissed as the product of libertarian paranoia over “harmful or misguided government intrusions”, as one commentator puts it. Safeguarding individual liberties and privacy is a legitimate issue in an open, democratic society in which concerns about the use of personal data are addressed in a transparent and accountable way.



Quantum computers put blockchain security at risk

Prasad scratches the surface of this complex topic. He seems satisfied with the prospect of broadening existing financial regulation. For example, he discusses regulating how and when financial-technology (fintech) companies should hand over personal data, rather than rethinking the whole regulatory framework.

The second aspect of the digital challenge to money is what counts as legal tender, and who is the issuer and guarantor. Conventionally, currencies issued and guaranteed by central banks are legal tender — for example, they must be accepted for settling any debt obligation to a private party. But cryptocurrencies issued by private entities, especially Bitcoin, are now part of an increasing number of transactions. Bitcoin was even briefly accepted by the electric carmaker Tesla, until the firm's chief executive, Elon Musk, rowed back on that idea over the gargantuan carbon footprint of the currency.

So who or what is at threat here? Private cryptocurrencies are fiat money, meaning that their value is not backed by a tangible asset such as gold. They share features with central-bank currencies, such as being able to settle payments. But cryptocurrencies lack transparency, accountability and governance. They are less safe (if you hold Bitcoin, for instance, and lose

your unique password, you lose access to your account forever), more volatile and cannot assure the same level of anonymity as cash.



[COVID-19 economics — first book hits shelves](#)

Prasad convincingly concludes that private cryptocurrencies — such as Facebook's proposed Diem (formerly known as Libra) — are more of a fad than a serious threat. They are likely to become niche assets held by speculators, or just as curios, while central banks prepare to launch their own digital currencies.

Despite its subtitle, *The Future of Money* is not just concerned with changes to banking and financial services. Prasad is deeply interested in the social impact of the digitalization of money, such as making finance more inclusive. Approximately 1.7 billion adults worldwide lack a bank account, according to the World Bank. Among other things, this means that immigrants who want to send money home often rely on expensive remittance services or personal connections.

Fintech can provide cheap financial services to all parts of society, including rural and low-income households. In Kenya, for example, the mobile banking service M-Pesa, introduced in 2007, allows small businesses in remote areas to safely and easily store and transfer money through mobile phones.

As with all attempts to define an uncertain future, *The Future of Money* leaves many questions open. On one point, however, Prasad seems confident — that although cash will become more marginal, it will never be phased out. This comes as a bit of a surprise after reading his account of the possibilities offered by digital money, as well as its flexibility and convenience. Prasad seems to accept that people wish to protect their privacy. He concludes that as long as people value convenience and anonymity, a cashless society is not on the cards. It is hard to disagree.

To completely phase out cash would require fully digital societies, with extensive and reliable infrastructure, universal access to devices and people who are comfortable with leaving electronic traces. Even though the pandemic has fast-tracked digital payments, perhaps we'll always carry a bit of cash, just in case.

Nature **597**, 626-627 (2021)

doi: <https://doi.org/10.1038/d41586-021-02615-2>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02615-2>

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

Opinion

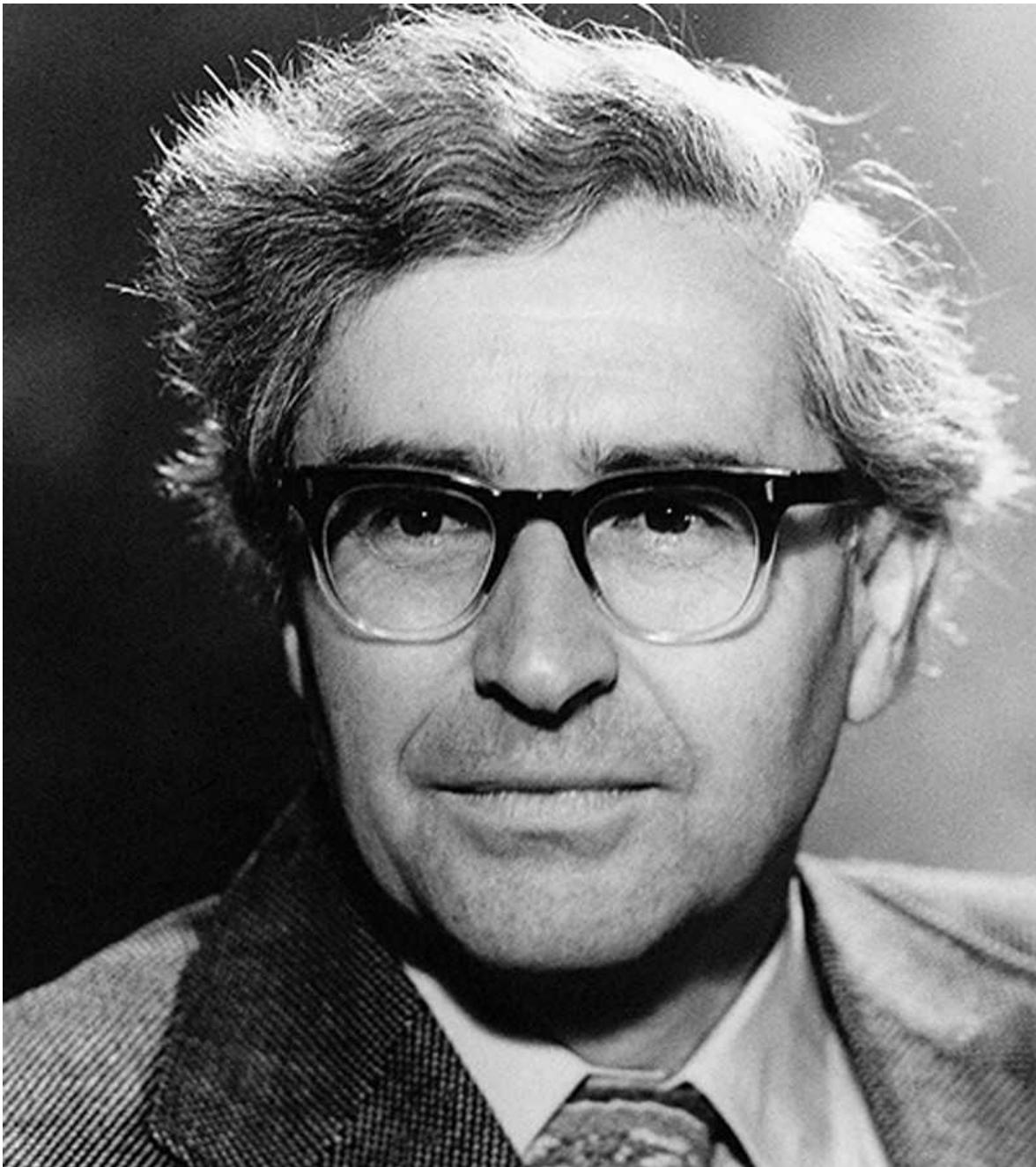
- [**Antony Hewish \(1924–2021\)**](#) [24 September 2021]
Obituary • Radioastronomer who won share of Nobel for role in discovering pulsars.
- [**Microsoft's million-tonne CO2-removal purchase — lessons for net zero**](#) [29 September 2021]
Comment • Strengthen markets, measures and definitions for removing carbon dioxide from the atmosphere to fight climate change.
- [**Mental health: build predictive models to steer policy**](#) [26 September 2021]
Comment • Combine economic, social and medical data to forecast need and design services to address the growing crisis.

- OBITUARY
- 24 September 2021

Antony Hewish (1924–2021)

Radioastronomer who won share of Nobel for role in discovering pulsars.

- [Malcolm Longair](#) 0



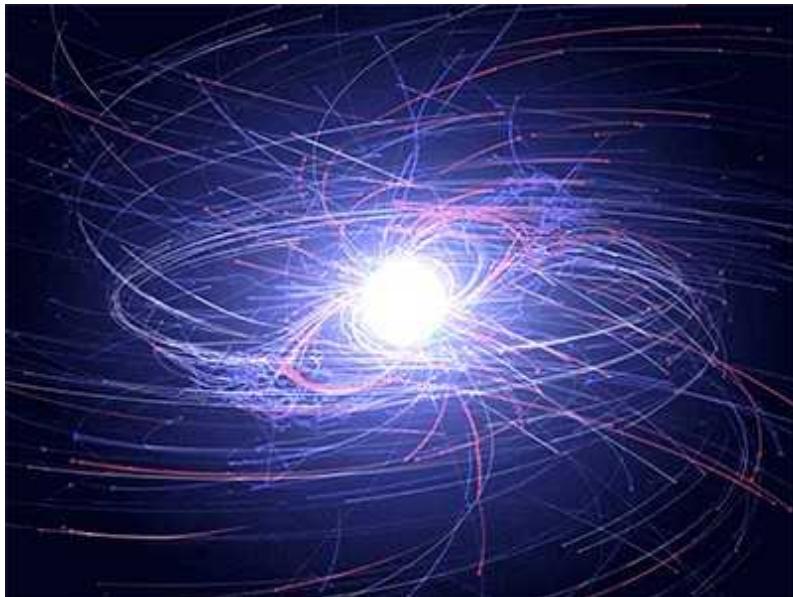
Credit: Photo 12/Universal Images Group/Getty

Antony (Tony) Hewish was a pioneering radioastronomer. His research student Jocelyn Bell (later Bell Burnell) made the first detection of a strange scintillating radio source that they subsequently showed was the first identified pulsar. These sources emit intense bursts of radio emission at precise time intervals, like the beam of a lighthouse. At first jokingly called LGM for ‘little green men’, these enigmatic sources were almost

immediately identified as magnetized, rotating neutron stars, one of the end points of stellar evolution. Hewish was awarded the 1974 Nobel Prize in Physics for his “decisive role” in the discovery. He has died aged 97.

This finding opened up new areas of high-energy and relativistic astrophysics. It confirmed the existence of neutron stars, predicted in the 1930s, and showed that Albert Einstein’s theory of general relativity could not be ignored in studies of their structures and stability.

As a member of the Cavendish Laboratory’s radioastronomy group at the University of Cambridge, UK, Hewish worked initially under the dynamic leadership of physicist Martin Ryle. Ryle and Hewish co-authored papers on aperture synthesis, the technique by which the signals from many small telescopes could be combined to reconstruct the imaging capability of a single, very large one. Ryle shared the Nobel prize with Hewish; Bell’s role led many to argue she should have shared the Nobel.



The golden age of neutron-star physics has arrived

Hewish grew up in Newquay on England’s Atlantic coast, where he developed a love of the sea and boats. In 1944, two years into his undergraduate studies in natural sciences at Cambridge, he was sent for war service to the Telecommunications Research Establishment in Malvern, UK, where Ryle was head of radar counter-measures. Hewish worked on a device

to jam the interception radar of hostile night-fighter aircraft. He wrote how checking the radiation pattern of the antennas on the ground required crouching in the slipstream of a B-17 Flying Fortress aeroplane, “while bombarded with gravel, dead rabbits, etc.”

Hewish returned to Cambridge to complete his undergraduate degree in 1946. He did not hesitate to join Ryle’s fledgling radioastronomy group there, following up the recent discovery that astronomical objects, including the Sun, emitted radio waves. Ryle showed Hewish to a large pile of brass tubing, and asked him to cut it into dipole antennas for his arrays. For the next ten years, Hewish worked on a succession of innovative radio telescopes.

Hewish was the group’s expert on radio scintillation, the flickering or twinkling of radio sources. He worked out the theory of radio source scintillation in detail in 1951–52. In 1954, he noted that strong scintillations would be observed from sources of small angular diameter because of plasma irregularities in the interplanetary medium. In 1964, such scintillations were observed in compact radio sources. Many of these were the newly discovered radio quasars. Associated with hyperluminous active galactic nuclei, they were soon to be linked with supermassive black holes.

Hewish realized that a large array designed to detect low-frequency scintillation would address three important astronomical problems: the discovery of many more radio quasars; the measurement of their angular sizes; and the determination of the structure and velocity of the solar wind, the stream of charged particles flowing out of the Sun. He won a grant of £17,286 (about US\$48,000 at the time) to design and build a 1.82-hectare array at the Mullard Radio Astronomy Observatory near Cambridge to detect fluctuating intensities on a timescale of one-tenth of a second.



Pulsar discoverer Jocelyn Bell Burnell wins \$3-million Breakthrough Prize

In October 1965, Jocelyn Bell joined the team as a graduate student. She became responsible for the network of cables connecting the dipoles. The telescope was commissioned in July 1967. Bell meticulously analysed the huge amount of data arriving each day, entirely by hand. On 6 August, she noticed a strange scintillating source that had an unfamiliar radio signature. On 28 November, the apparent scintillation proved to be a train of stable pulses arriving every 1.33 seconds. Nothing like this had been observed before. Bell discovered three similar sources, including one with a period of only 0.25 seconds.

Within a few months of the discovery's publication in *Nature* ([A. Hewish et al. *Nature* 217, 709–713; 1968](#)), the former Cambridge physicist Thomas Gold, then at Cornell University in Ithaca, New York, identified the pulsars as magnetized, rotating neutron stars. The discovery of neutron stars by this radio technique — they were much too faint to be detected by optical telescopes — came as a complete surprise. In the years after, other researchers detected many radio pulsars, including in binary neutron-star systems that provided precision tests of general relativity. Astronomers also inferred that the acceleration of pulsars' orbits is caused by the emission of gravitational radiation.

Hewish also used the scintillation technique to study interplanetary weather. In particular, he identified major plasma outbursts from the Sun that could

affect GPS navigation. He led the Cambridge radioastronomy group from 1977–89, and was head of the Mullard Radio Astronomy Observatory from 1982–88.

Tony — a friendly and hands-on colleague and supervisor — had the ingenuity to realize what could be achieved by constructing a low-cost radio telescope with a small team of dedicated research students and assistants. Within this vibrant research environment, discovery thrived.

Nature **597**, 628 (2021)

doi: <https://doi.org/10.1038/d41586-021-02617-0>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02617-0>

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

- COMMENT
- 29 September 2021

Microsoft's million-tonne CO₂-removal purchase — lessons for net zero

Strengthen markets, measures and definitions for removing carbon dioxide from the atmosphere to fight climate change.

- [Lucas Joppa](#)⁰,
- [Amy Luers](#)¹,
- [Elizabeth Willmott](#)²,
- [S. Julio Friedmann](#)³,
- [Steven P. Hamburg](#)⁴ &
- [Rafael Broze](#)⁵



Climeworks' plant for direct air capture in Iceland removes carbon dioxide permanently from the atmosphere. Credit: Julia Dunlop/Climeworks

In January this year, Microsoft made a major announcement: it had paid for the removal of 1.3 million tonnes of carbon dioxide from the atmosphere. Among its purchases were projects to expand forests in Peru, Nicaragua and the United States, as well as initiatives to regenerate soil across US farms. Microsoft will pay the Swiss firm Climeworks to operate a machine in Iceland that pulls CO₂ from the air and injects it into the ground, where it mineralizes and turns to stone. The amount of CO₂ to be removed is equivalent to about 11% of the annual emissions from Microsoft's value chain; of this, the company will count less than half as being certified to officially compensate for its emissions. It is the largest corporate procurement of carbon removal so far.

Microsoft did this as part of its 2020 commitment to slash its greenhouse-gas emissions to 'net zero' — as one of more than 120 nations and 1,500

companies to set such goals¹. By 2030, the company will reduce its emissions by half or more, and will have 100% of its electricity consumption matched by zero-carbon energy purchases. It will electrify its vehicle fleet, stop using diesel for backup energy and reduce emissions across its value chain. Emissions that are harder to abate, including historical emissions, will be compensated for by withdrawing carbon from the atmosphere. The firm is levying an internal carbon tax across all types of greenhouse-gas emission. It has set up a US\$1-billion fund to invest in carbon reduction and removal technologies, and partnerships to provide social and environmental benefits. The aim is that, by 2030, the company will be carbon negative. By 2050, it will have removed all of its emissions since it was founded in 1975.

Here we summarize the lessons learnt from Microsoft's carbon-removal efforts, along with those from another early corporate procurement — the \$9-million purchases of carbon removal in 2020 and 2021 by the US-Irish financial-infrastructure company Stripe. Although these are just two companies' efforts, they are the first significant open solicitations focused exclusively on carbon removal. We write as a team composed of Microsoft staff working on the company's carbon-negative programme and research scientists who analyse carbon reduction and removal strategies.

We highlight three ‘bugs’ in the current system: inconsistent definitions of net zero, poor measurement and accounting of carbon, and an immature market in CO₂ removal and offsets. These challenges need to be overcome if the world is to reach net zero by mid-century.

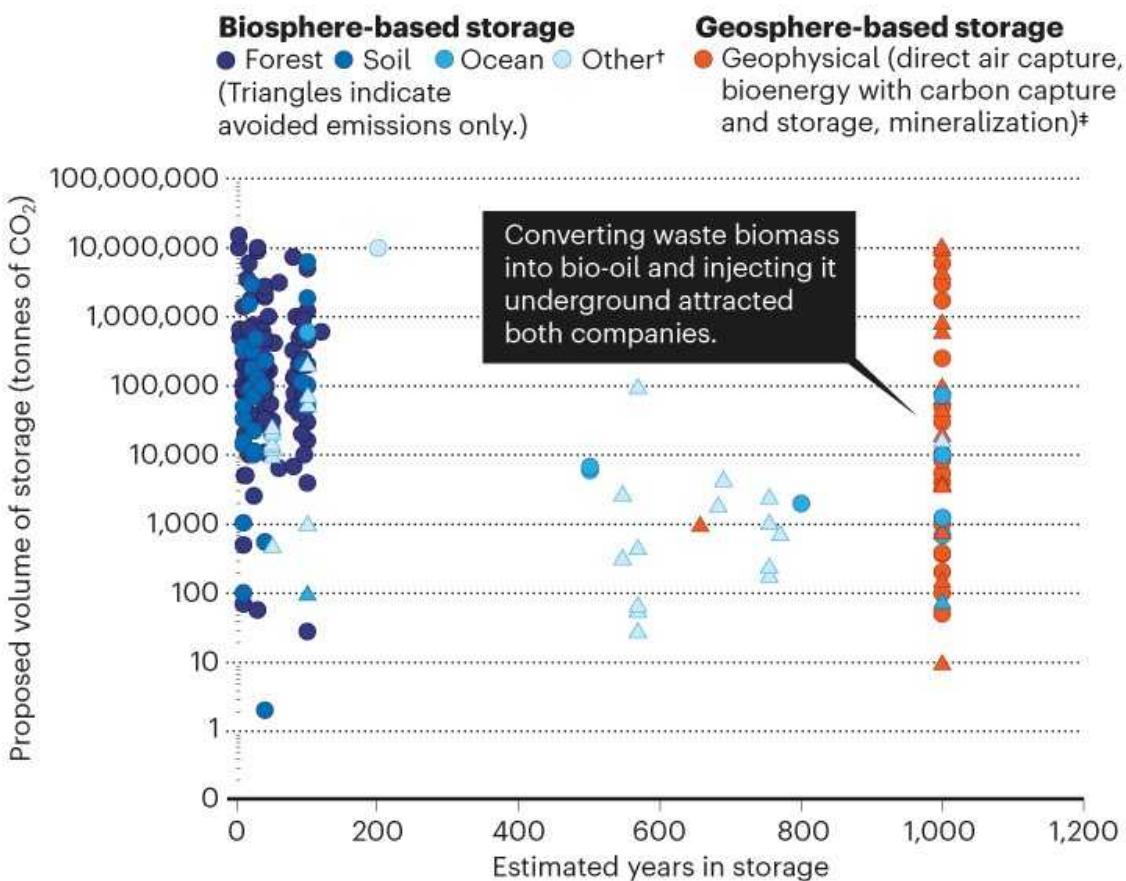
Three lessons

First, the supply of solutions capable of removing and storing carbon viably is a tiny proportion of that needed to reach global net-zero emissions by 2050 (which is an anticipated 2–10 gigatonnes of CO₂ per year)². Although Microsoft received 189 proposals offering 154 megatonnes of CO₂ (MtCO₂) over the coming years, only 55 MtCO₂ were available immediately, and a mere 2 MtCO₂ met [Microsoft's criteria for high-quality CO₂ removal](#). Stripe's 47 carbon-removal proposals amounted to 16 MtCO₂, but only

0.024 MtCO₂ met the company's requirement that [carbon remain sequestered for at least 1,000 years](#) (see 'Carbon-market snapshot').

CARBON-MARKET SNAPSHOT

In 2020, Microsoft and financial-services firm Stripe received 189* and 47 proposals from companies, respectively, for locking away carbon dioxide. Of these, 95% used nature-based storage, which is less durable than geosphere-based. Few options were available for permanent removal. Only about 2 million tonnes' worth was judged reliable enough to purchase, of the around 170 million tonnes offered.



*Data on 161 proposals compiled by CarbonPlan (<https://carbonplan.org>); these exclude 28 further proposals to Microsoft that lacked sufficient information.

†Biomass, wood products and biochar. *Many geosphere-based solutions were classified as >1,000 years duration, but are shown here as 1,000 years for simplicity.

©nature

Sources: CarbonPlan/Authors' analysis

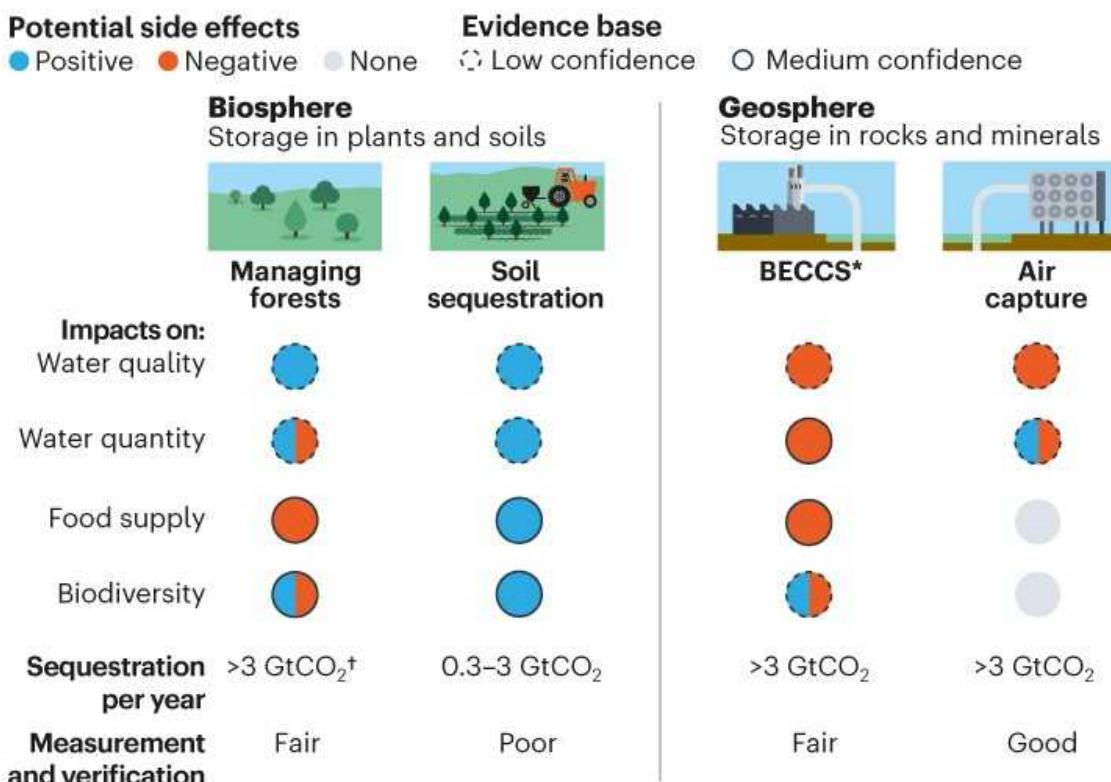
Second, the scarcity of proposals that met the companies' criteria reflects a lack of standards and clear definitions. Roughly one-fifth of proposals to Microsoft focused on avoiding new emissions, not on withdrawing CO₂ from the atmosphere; these were rejected. Others lacked the technical

information needed to ensure reliability. Indeed, there's no standard way to measure, report and verify carbon removed. Such ambiguity is a barrier to investment.

Third, systems for accounting for carbon removal do not distinguish between short- and long-term forms of CO₂ storage (see 'Some carbon-removal strategies'). This distorts the market and discourages investments in more-durable solutions. Nature-based storage projects sequestering carbon for less than 100 years accounted for most proposals that Microsoft received (in total, more than 95% of CO₂ volume). It is cheaper and easier to establish trees and enrich soils than to deploy nascent technologies that capture carbon and store it geologically.

SOME CARBON-REMOVAL STRATEGIES

Nature-based methods for storing carbon dioxide are relatively cheap and currently available. But carbon stored in terrestrial ecosystems is at risk of release by fires and pests, for example. Geological storage could be permanent, but today's technologies are pricey and immature.



Impact ratings are from the Intergovernmental Panel on Climate Change (IPCC) Sixth Assessment Report, apart from 'Measurement and verification', which are based on the authors' judgement.

*BECCS, bioenergy with carbon capture and storage; [†]GtCO₂, gigatonnes of CO₂.

©nature

Sources: IPCC Sixth Assessment report (2021)/Authors' analysis

On average, in the pitches that Stripe received, biosphere-based storage projects cost only \$16 per tonne of CO₂ (tCO₂), whereas geosphere-based storage costs, on average, \$141 per tCO₂ (\$20–10,000 per tCO₂) — similar to the costs in Microsoft's proposals. Geosphere-based costs are higher, close to the \$30–200 per tCO₂ social cost of carbon emissions³. However, many factors are not priced in. Nature-based solutions face risks of reversal by fires, pests, storms and changes in land use. These risks can be reduced with insurance and by accounting for carbon across larger areas⁴. But reliable tools for tracking carbon at scale are lacking. Co-benefits such as water conservation, hazard protection and biodiversity are also unaccounted for.

Three priorities

The following three aspects of the world's carbon-removal efforts need urgent attention.

Meaning. The Intergovernmental Panel on Climate Change's definition of net zero is simple enough at a global scale: when “anthropogenic emissions of greenhouse gases to the atmosphere are balanced by anthropogenic removals”. But it is too broad to tell individual companies how they can reach net zero⁵.

Businesses have lots of options. For example, offsetting emissions — by paying someone else not to emit as a way to compensate for ongoing emissions — can slow the rate at which CO₂ builds up in the atmosphere, but it does not remove any. That's why, in 2020, Microsoft pivoted to purchasing only carbon removal. It also expanded the scope of its programme to include its whole value chain and historical emissions, more than quadrupling the tonnage of carbon that the company needs to compensate for.

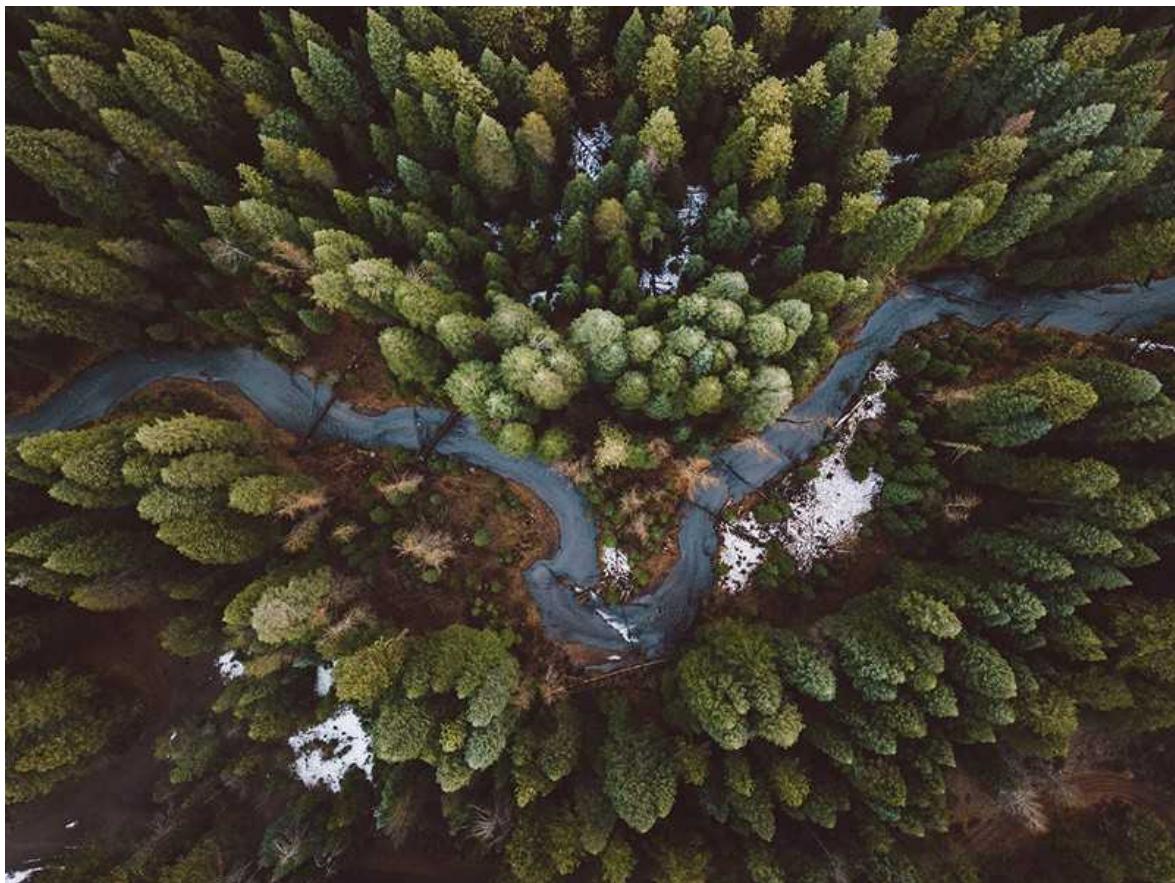


[Climate policy models need to get real about people — here's how](#)

Companies need standards to gauge whether their carbon commitments are consistent with global net zero. Efforts to develop them include the international non-profit [Science-based Targets Initiative](#), the [Oxford offsetting principles](#) from researchers at the University of Oxford, UK, and the cross-sector business initiative [Transform to Net Zero](#). These emphasize reducing all greenhouse-gas emissions as much as possible along the value chain; setting interim targets; purchasing carbon removal; and shifting towards long-term carbon storage. Such actions must be in addition to those designed to protect and enhance stocks and sinks of carbon in the biosphere.

Many organizations assume there is no limit to carbon-removal possibilities, but there is. Nature-based removal is constrained by area and competing uses of land⁶. Engineering solutions can scale up, but are currently scarce, expensive and resource intensive. Competition for supply will grow as more companies act. The most effective measures might become oversubscribed, making many net-zero commitments impossible to fulfil. Without rapid growth in supply, the world might run out of high-quality options to compensate for remaining emissions, even after drastic reductions. Too little is being invested in durable technological approaches and geological storage systems.

What can be done? Companies should start by reducing to zero those emissions they have most control over, such as from energy use and land management. In the meantime, they should invest in expanding the supply and lowering the cost of the most effective carbon-removal technologies, as Microsoft and Stripe are doing. Firms should consider purchasing removal for emissions beyond their control that are hardest to abate, such as transport of goods and materials by air and sea.



Planting forests and improving their management are nature-based solutions that companies can use to remove and store carbon. Credit: Getty

Researchers need to define a global budget for carbon removal, including evolving scenarios for the supply of nature-based and technology-based removal and storage. And they should assess the future demand for carbon removal driven by net-zero commitments of diverse organizations around the globe.

Social equity is crucial. To reach net zero, developing economies and under-served communities must benefit. For this reason, Microsoft is partnering with Sol Systems, a solar-energy finance and development firm in Washington DC, to create a fund that links the purchase of renewable energy to career training, habitat restoration and clean-energy grants. Similar efforts should be undertaken for carbon removal.

Measurement. Corporations need more-accurate, automated and consistent ways of measuring and accounting for carbon. The non-profit organization [Greenhouse Gas Protocol](#) provides guidelines for assessing emissions from internal operations, such as vehicle use and manufacturing, and from purchases of energy sourced off-site. Estimating emissions from supply and value chains is more difficult⁷. It requires calculations from all suppliers and users of a company's products and services. Three-quarters of Microsoft's emissions come from these, including building materials, business travel, product life cycles and the electricity that customers consume when using Microsoft's products. The company has been using expenditure data and industry-average emissions for reporting purposes. But these have large uncertainties and are of limited use in reducing emissions in practice. They do not factor in the impacts of making different choices in the value chain for greenhouse-gas emissions.

Microsoft is making a start by requiring suppliers to make annual disclosures of their greenhouse-gas emissions and to adopt plans to reduce emissions. However, suppliers face a plethora of carbon-reporting requests.



Protect global supply chains for low-carbon technologies

Digital tools are emerging that can automate and increase the accuracy of emissions measurement. Systems that combine remote-sensing images, sensors and machine learning are being developed — for example, the European Space Agency's Copernicus CO₂ monitoring mission and the methane-tracking satellite MethaneSAT, backed by the non-profit Environmental Defense Fund in New York City.

Microsoft's FarmBeats team is developing low-cost, scalable methods for measuring soil carbon in agricultural fields. Microsoft is also collaborating with start-ups, such as NCX (formerly SilviaTerra) in San Francisco, California, to process terabytes of satellite imagery to count trees in the United States, estimate their potential for carbon sequestration and create a marketplace for private landowners to reduce deforestation. For projects such as these to succeed, rural broadband will be needed to collect and transmit data from networks of Internet-connected devices — the Internet of Things.

A growing number of enterprise-software companies, including SAP in Walldorf, Germany, Salesforce in San Francisco and Microsoft, are developing platforms for automating carbon accounting. For example, Microsoft's Cloud for Sustainability connects organizations to real-time

sources of data to track carbon and show performance against net-zero goals. Such platforms are still in their early stages, however.

Automated systems will become more important as greenhouse-gas reporting and emissions reductions become mandatory — as the leaders of the G7 group of countries announced in June that they would implement. Investors and customers increasingly demand that companies demonstrate progress against environmental, social and corporate-governance goals. National governments are drafting regulations for corporate climate-related disclosure. These rules could help to create common standards for carbon accounting and climate-change data.

Markets. Companies need better economic incentives to promote the most effective forms of CO₂ removal. Nature-based removal and storage, and technology-enabled removal and geosphere-based storage are not equivalent commodities and should not be valued as such.

Today's pricing on a per-tonne basis encourages companies to buy the lowest-quality carbon offsets. It does not monetize the duration of carbon storage, the risk of premature release, or the social equity or environmental benefits of removal. At current prices, credits for avoided emissions are the cheapest (as low as \$3 per tCO₂). Nature-based carbon-removal costs more (\$5–50 per tCO₂), although it is much less expensive than geo-based removal.



Nature-based solutions can help cool the planet — if we act now

Price and supply will shift over time. The cost of nature-based removal is likely to increase as the requirement increases and supply declines, as available forests and soils become saturated. Meanwhile, geo-based technologies will develop and scale up, becoming more accessible and cheaper. Companies making commitments to become net zero by 2050 have to make decisions now about operations in 30 years' time, yet there is little economic modelling to project how CO₂-removal markets might change.

Governments, researchers and companies need to develop a robust and effective carbon-removal market that can meet the demand for global net zero. A key advance would be to set consistent standards for measuring, verifying and accounting for carbon removal that internalize differences in the quality and durability of carbon stored in the biosphere and the geosphere. Multiple approaches have been proposed^{8,9}. However, these comparisons hinge on choices of a few parameters, such as economic rates. Factors such as competition for land use, limits to ecosystem-carrying capacity and social and environmental impacts need to be accounted for.

Microsoft is currently executing a strategy for its investments over the next decade, while setting conditions for the next three decades. The company is prioritizing funding across three pillars: reducing greenhouse-gas emissions; removing carbon from the atmosphere and storing it in the biosphere; and

removing and storing it in the geosphere. It is implementing internal targets, grants and other incentives to encourage innovations in emissions reduction. Nature-based solutions will be a major portion of its carbon-removal strategy in the near term. The company will include more geo-based storage as this becomes more widely available. To hasten that day, Microsoft, through its \$1-billion Climate Innovation Fund, is investing in projects such as the Orca direct-air-capture facility in Iceland developed by Climeworks. Anyone who can do more should do more. It is time to step up to develop the science, technology and markets for successful carbon removal.

Nature **597**, 629-632 (2021)

doi: <https://doi.org/10.1038/d41586-021-02606-3>

References

1. 1.

Black, R. *et al.* *Taking Stock: A Global Assessment of Net Zero Targets* (Energy & Climate Intelligence Unit/Oxford Net Zero, 2021).

2. 2.

Intergovernmental Panel on Climate Change. *Global Warming of 1.5°C* (eds Masson-Delmotte, V. *et al.*) (World Meteorological Organization, 2018).

3. 3.

Moore, F. C. & Diaz, D. B. *Nature Clim. Change* **5**, 127–131 (2015).

4. 4.

Schwartzman, S. *et al.* *Environ. Res. Lett.* **16**, 091001 (2021).

5. 5.

Rogelj, J., Geden, O., Cowie, A. & Reisinger, A. *Nature* **591**, 365–368 (2021).

6. 6.

Nolan, C. J., Field, C. B. & Mach, K. J. *Nature Rev. Earth Environ.* **2**, 436–446 (2021).

7. 7.

Patchell, J. J. *Cleaner Production* **185**, 941–958 (2018).

8. 8.

Herzog, H., Caldeira, K. & Reilly, J. *Clim. Change* **59**, 293–310 (2003).

9. 9.

Friedmann, S. J. *et al.* *Net-Zero and Geospheric Return: Actions Today for 2030 and Beyond* (Columbia Univ./Global CCS Institute, 2020).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02606-3>

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

- COMMENT
- 26 September 2021

Mental health: build predictive models to steer policy

Combine economic, social and medical data to forecast need and design services to address the growing crisis.

- [Jo-An Occhipinti](#) ⁰,
- [Adam Skinner](#) ¹,
- [P. Murali Doraiswamy](#) ²,
- [Cameron Fox](#) ³,
- [Helen Herrman](#) ⁴,
- [Shekhar Saxena](#) ⁵,
- [Elisha London](#) ⁶,
- [Yun Ju Christine Song](#) ⁷ &
- [Ian B. Hickie](#) ⁸



A student during a COVID-19 quarantine in St Petersburg, Russia.Credit: Alexander Demianchuk/TASS/Getty

A crisis is growing in mental health as the widespread impacts of the COVID-19 pandemic and the economic hardship it has brought bites deeper. In Japan, suicides rose by 16% during the second wave of the pandemic, from July to October 2020, compared with the rate in previous years¹. In the United States, 25% of people aged 18–24 surveyed in June 2020 reported increased substance use to cope with pandemic-related stress². This year, the flagship report of the United Nations children's charity UNICEF, *The State of the World's Children*, focused on child and adolescent mental health and well-being for the first time. If urgent and effective action is not taken, the protracted and global scale of the pandemic disruption will cast a long shadow on mental health, particularly that of young people.

Decades of research suggests that the response must be all-encompassing and long term. The fact that this is neither feasible nor affordable in many

contexts gives rise to two types of response. Some governments or agencies allocate available resources over too broad a range of evidence-based programmes and services; without the scale and intensity needed, these cannot achieve real and sustained impact. An example is Australia's struggle, over three decades, to shift the needle on many mental-health conditions. Other governments and agencies take a reactive and ad hoc approach — as exemplified by the US response to the synthetic-opioid overdose epidemic. Neither approach will be adequate to tackle today's mental-health crisis.

Instead, policymakers must account for how the pandemic has fundamentally changed the state of mental health across society. It is time to examine preconceptions about what interventions are effective.



[COVID-19 mental-health responses neglect social realities](#)

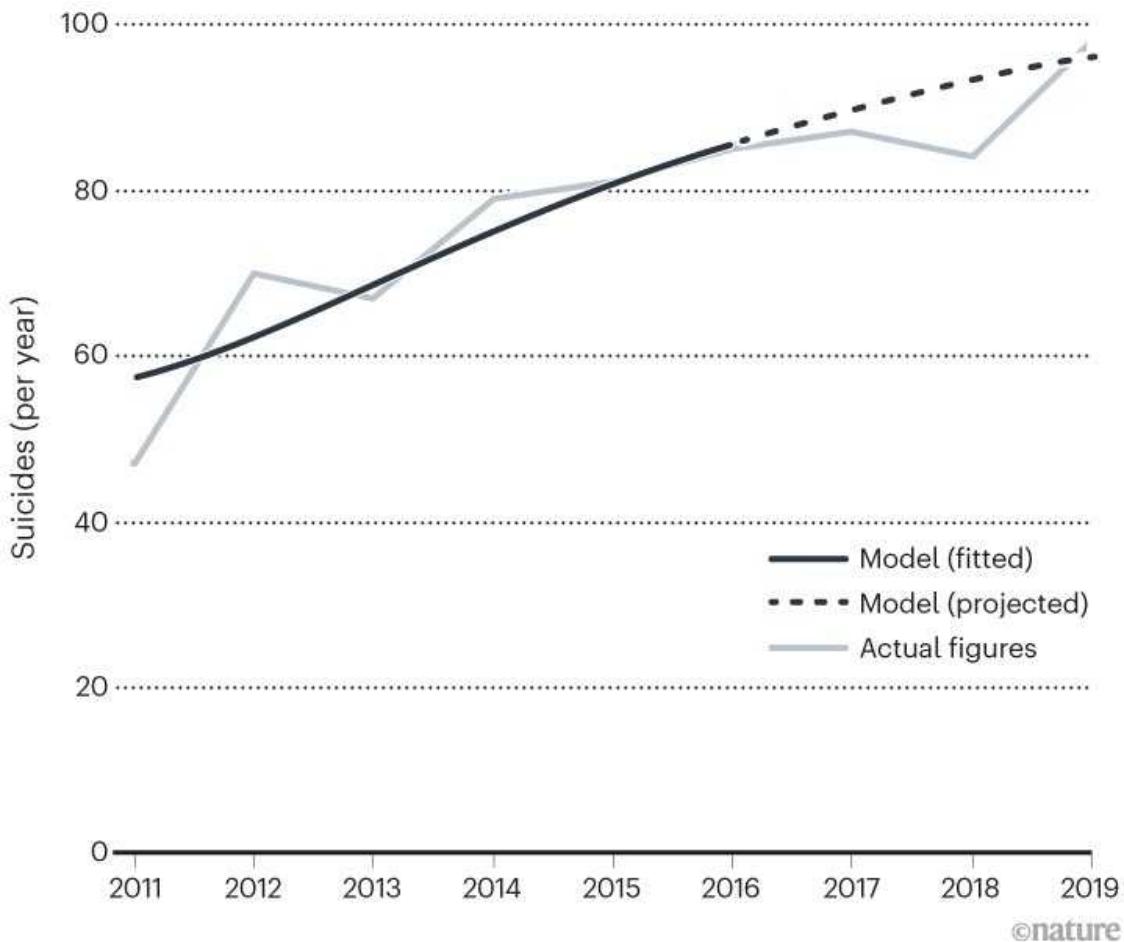
Lessons can be learned from infectious-disease research. Systems models allowed researchers to rapidly predict the spread of COVID-19 (see [*Nature* 580, 316–318; 2020](#)), integrating contact-tracing data based on commuting patterns and mobile-phone location trackers. Although imperfect, in some places these models provided a virtual testing ground for alternative assumptions and for the timing and scale of mitigation strategies, including lockdowns, mask wearing, school closures and vaccination. The models also accounted for the changing likelihood of people complying with such

measures as the pandemic persisted. Where decision makers worked closely and cooperatively with modellers, as in Australia, New Zealand (see go.nature.com/3kiw79n) or Taiwan, models were used to inform timely, decisive and effective responses to the pandemic³.

We argue that a similar systems-modelling approach should be used to tackle the mental-health challenge. Drawing together qualitative and quantitative evidence and data, models should capture changes triggered by the pandemic — such as education loss, job loss, domestic violence, social isolation, fear and uncertainty. Models should forecast demand for community mental-health services and acute care, including emergency-department presentations and psychiatric hospitalizations, as well as outcomes such as suicidal behaviour (see ‘Mental-health forecast’).

MENTAL-HEALTH FORECAST

The predictions of a model developed at the University of Sydney, Australia, to inform public-health decisions were close to the suicide rate for a region of New South Wales in the years before the COVID-19 pandemic.



Source: Australian Inst. Health and Welfare Natl Mortality Database

Credible projections of population mental-health outcomes are needed to stress-test new policies and mitigation strategies, from employment programmes to helplines and investments in education and retraining. Before allocating significant investments, alternative scenarios should be simulated to reveal the combination, scale, targeting, timing and duration of health, social and economic policies and initiatives that will deliver the greatest impacts. Understanding which combinations of interventions work best at which stage is key to reducing harm.

As experts in mental health and systems modelling, we outline here five challenges and four priorities to ensure that models are used to reliably guide policy and allocate resources.

Five challenges

Closing the care gap. In most societies, mental-health systems were underfunded and fragmented even pre-pandemic⁴. The gap between those that need care and those that receive it is widest in countries where there is conflict, such as Afghanistan, and in countries with higher rates of unemployment, unequal wealth distribution and budget cuts, such as Greece. Across low-, middle- and high-income countries there are significant treatment gaps between the most privileged people and those who are most marginalized, such as Indigenous populations in Canada, New Zealand, the United States and Australia⁵.



When is sorrow sickness? A history of depression

Allocating resources. Interest and investment in mental health is growing. Witness the work of the World Health Organization (WHO), UNICEF and global civil-society organizations such as the World Economic Forum's Global Shapers Community (see go.nature.com/3tpbtrz). In May 2020, the UN secretary-general António Guterres called for more urgent action on the

issue⁶, and the elevation of mental health and suicide prevention in the global development agenda in recent years will be important to the COVID-19 mental-health response⁷. Up to US\$160 billion has been committed by the World Bank Group to help developing countries tackle the health, social and economic impacts of the pandemic, and governments around the world are committing trillions to social and economic aid packages. However, without good planning tools, decision makers will continue to be challenged by the complexity of causal drivers, the quagmire of the known and the unknown, and the kaleidoscope of voices and choices.

Expanding the toolkit. The predominant approach to research on mental illness and suicidal behaviour uses retrospective data to identify independent risk factors, such as unemployment, substance abuse or childhood trauma. Half a century of such study has brought only partial progress towards population-level impact, as a recent meta-analysis noted⁸. In addition, analysing risk factors independently fails to account for their interactive effects. This makes robust projections of population mental-health outcomes difficult, if not impossible.

By contrast, infectious-disease epidemiology has matured into a robust interdisciplinary field through methodological expansion that makes routine use of the analytic techniques of complex-systems science. At the outset of the COVID-19 pandemic, this enabled global research teams to rapidly deploy and customize existing systems models. Researchers could analyse and forecast transmission trajectories under different conditions, allowing uncertainty to be quantified. As new information and data about the virus and its transmissibility became available, models were refined and provided governments with critical tools for testing and weighing the impact of responses, from mask wearing to travel quarantines.

Embracing ideas. Discourse on mitigating the mental-health impacts of COVID-19 has focused largely on mental-health education programmes, crisis helplines, improved access to virtual services, and tinkering with existing arrangements to enhance access to emergency care. This health-sector view fails to recognize that the most potent mental-health interventions can be social and economic. These could include employment support, eviction moratoriums, subsidized education and training, or

increased unemployment insurance^{9,10}. For instance, an income supplement that moved 14% of Native American households out of poverty in North Carolina saw a 32% decrease in psychiatric symptoms among children of those households¹¹.



A food-bank worker in California: economic programmes must be part of the mental-health policy toolkit.Credit: Brian L. Frank

Huge questions. It is not at all clear what combination of policies, initiatives or reforms are needed to respond effectively to the mental-health crisis. What impact will that combination have across different outcomes? Will the most effective combination be similar across different contexts? What targeting, timing, scale, frequency and intensity of investments are necessary? Will there be rebound effects when temporary mitigation strategies such as income support or eviction moratoriums are removed? What are the consequences of delayed actions? Will unintended

consequences arise from well-meant but ill-designed or ill-timed mitigation measures?

The scale of these challenges behoves us to take a more progressive research path^{8,12}. We must recognize the feedback loops, threshold effects, non-independence and non-linearity that characterize our subject of study^{13,14}. For example, an increase in unemployment increases both the prevalence of psychological distress (which can lead to increased substance misuse and suicidal behaviour) and domestic violence. Domestic violence increases rates of adverse childhood exposures and psychological distress, further driving up rates of substance misuse and so on¹⁴. Modelling and simulation can help us get a handle on such complexity.

Opportunity in crisis

A paradigm shift in population mental-health research is emerging at just the right time. The complexity and global scope of the crisis requires it, and the computational tools are sufficiently advanced and accessible to make it feasible. The past two decades have seen substantial investments in the data systems of low- and middle-income countries (LMICs), covering civil registration — the collection of statistics such as births, marriage, divorce, cause of death — and health information such as service coverage and capacity, and medical records^{15,16}. A 2021 report by the WHO¹⁷ on 183 of its member states highlighted that 86 have suicide data that are considered good quality (49 are high-income countries, and 37 are LMICs). And countries including India, China and the United Arab Emirates are collecting data that will contribute to estimating the burden of mental ill-health through national and regional surveys. These could be used as model inputs. Hence, the advanced modelling tools to support decisions are no longer the exclusive domain of high-income countries.

During the COVID-19 crisis, the Brain and Mind Centre at the University of Sydney, Australia, leveraged years of experience to develop a series of models to inform policy and planning for the country's regional, state and national mental-health systems. In youth and in the population as a whole, the models project the likely trajectories for a range of outcomes including the prevalence of psychological distress, rates of help-seeking, wait times,

emergency-department presentations, self-harm hospitalizations, and suicide deaths as a result of the pandemic^{[10,18](#)}. Such projections provide a test bed for probing the trade-offs and potential synergies of economic and social measures, among other strategies.



Science careers and mental health

For example, models from the Brain and Mind Centre suggested that among the smart choices for Australia in the recovery period (2021–25) were investments in childcare, employment programmes and job creation (particularly for women), active follow-up after suicide attempts, and expansion of digitally coordinated specialist mental-health services. These were forecast to prevent an estimated 6% of self-harm hospitalizations and 4.1% of emergency-department presentations relating to mental health^{[10](#)}. Investing in more psychiatric-hospital beds, awareness campaigns, helplines or stand-alone primary-care or specialist services were projected to deliver little impact, despite being considered ‘evidence based’.

Similar approaches are also under way in the United States. The US National Institute of Mental Health (NIMH) recently launched the Advanced Laboratories for Accelerating the Reach and Impact of Treatments for Youth and Adults with Mental Illness (ALACRITY) research-centres programme. These centres harness systems science, computational approaches, behavioural economics and digital health to test methods to synthesize data.

Still, there is progress to be made on several fronts. Pandemic modelling benefited from unprecedented levels of intergovernmental and intersectoral cooperation on the exchange of data and crucial information, for instance about movement patterns, that were shared across telecommunication, transport and health agencies. Unfortunately, similar cooperation is rare around the ethical sharing of data that would benefit efforts to model mental health during COVID, for instance across health, education, social services and the economic sector.

Four priorities

Several Australian members of the World Economic Forum's Global Future Council on Mental Health (for details, see Supplementary information) have spent five years applying systems modelling to mental health. This experience has highlighted the following four priorities for rapid and successful deployment of models to improve population mental health and prevent suicide. It is key to:

Use a technical blueprint. To construct systems models in diverse contexts, a clear picture of how social, economic and health factors interact to drive psychological distress and mental-health outcomes is needed. [An exemplar blueprint will be released this year¹⁹](#), based on COVID-19 research by the Brain and Mind Centre and supported by the World Economic Forum. It details how to develop a national or regional system-dynamics model of mental health, including key inputs, outputs and processes. The blueprint includes guidelines for analysing the projections of population mental-health outcomes. It also offers guidance on modelling the impact of policies such as investing in jobs, childcare, education or social connectedness.



Action on mental health needs global cooperation

Build multidisciplinary teams. Expertise in computational and systems science, epidemiology, psychology, psychiatry, social science, policy and economics will ensure that models have valid, robust, interdisciplinary underpinnings, and limit bias. The COVID-19 international modelling consortium (CoMo) is a notable example, drawing on a broad range of disciplines and contextual knowledge to provide relevant, trusted and tailored tools to more than 30 countries²⁰.

Strengthen data systems. Improving the range, quality and timeliness of data collected and continuously reappraising models as fresh data emerge will reduce uncertainty bounds, expand the insights that can come from systems modelling and improve decision making. In addition, a commitment to transparency and multidisciplinary co-design, and presenting the results of alternative assumptions helps to guard against the politicization of model outputs for public policy.

The data that are needed depends on the scope of the model and the outcome indicators of interest. These should be determined by national priorities, stakeholder inputs and the research question being posed. Time-series data used for model calibration range from demographic and labour-force statistics such as births, mortality, migration and unemployment rates, to estimates of the prevalence of distress and mental disorder and related data

such as emergency-department presentation, hospitalization and intentional-self-harm rates. Adjustments will need to be made depending on the local context. In Australia, for example, self-harm hospitalizations are recorded, but not data on suicide attempts; in Colombia, suicide attempts are reported. Model parameterization also draws on research evidence (including systematic reviews, randomized controlled trials and cohort studies) and expert consensus.

Commit to co-design. Models that are created with multidisciplinary stakeholders and, crucially, those with lived experience of mental ill-health have better credibility and validity. A participatory approach helps to identify the key referral pathways of the service system, bottlenecks and barriers. It highlights users' experiences of delay, disengagement, service gaps and interruptions to continuity of care¹⁴. Self-harm survivors, for example, identified that long wait times made it more likely that they would stop seeking care, prolonging their distress and increasing the likelihood of suicidal behaviours. Qualitative data, triangulated with quantitative service data, help to customize and ground models.

Recognize limitations

Of course modelling is no policy panacea — just look at the snail's pace of climate-change mitigation. There is necessary scepticism about the gulf between generalizations and granular detail. Useful forecasts in high-income countries might not work in low-income ones; and what helps the privileged might not aid the disadvantaged. Critics point out that if the social determinants of physical health aren't being addressed — as has been so obvious from disparate COVID-19 death and disability rates — it is naive to imagine they will suddenly be addressed for mental health once a good model is available. Critics also note that outputs are only as good as inputs, cautioning about a paucity of data and the perils of incorrect assumptions.

Nonetheless, as the pandemic rages on, alongside so many other global challenges, there has never been a more important time to strive for what we describe here.

doi: <https://doi.org/10.1038/d41586-021-02581-9>

This article has been written on behalf of World Economic Forum Global Future Council on Mental Health.

References

1. 1.
Tanaka, T. & Okamoto, S. *Nature Hum. Behav.* **5**, 229–238 (2021).
2. 2.
Czeisler, M. É. *et al. Morb. Mortal. Wkly Rep.* **69**, 1049–1057 (2020).
3. 3.
Chang, S. L., Harding, N., Zachreson, C., Cliff, O. M. & Prokopenko, M. *Nature Commun.* **11**, 5710 (2020).
4. 4.
World Health Organization. *The Impact of COVID-19 on Mental, Neurological and Substance Use Services: Results of a Rapid Assessment* (WHO, 2020).
5. 5.
Collins, P. Y. & Saxena, S. *Nature* **532**, 25–27, (2016).
6. 6.
United Nations. *Policy Brief: COVID-19 and the Need for Action on Mental Health* (United Nations, 2020).
7. 7.
Arensman, E. *Crisis* **38**, 1–6 (2017).

8. 8.

Franklin, J. C. *et al. Psychol. Bull.* **143**, 187–232 (2017).

9. 9.

Lund, C. *et al. Lancet Psychiatry* **5**, 357–369 (2018).

10. 10.

Occhipinti, J., Ho, N., Skinner, A., Song, Y. C. & Hickie, I. B. *Closing the Gender Gap: Supporting the Mental Health of Australian Women During the COVID-19 Era* (Brain and Mind Centre, University of Sydney, Sydney, Australia, 2021).

11. 11.

Costello, E. J., Compton, S. N., Keeler, G. & Angold, A. *JAMA* **290**, 2023–2029 (2003).

12. 12.

Atkinson, J. *et al. Front. Psychiatry* <https://doi.org/10.3389/fpsyg.2020.606035> (2020).

13. 13.

Atkinson, J. *et al. Front. Psychiatry* <https://doi.org/10.3389/fpsyg.2019.00448> (2019).

14. 14.

Occhipinti, J. *et al. BMC Med.* **19**, 61 (2021).

15. 15.

Hung, Y. W., Hoxha, K., Irwin, B. R., Law, M. R. & Grépin, K. A. *BMC Health Serv. Res.* **20**, 790 (2020).

16. 16.

Lopez, A. D., McLaughlin, D. & Richards, N. *BMC Med.* **18**, 58 (2020).

17. 17.

World Health Organization. *Suicide Worldwide in 2019: Global Health Estimates* (WHO, 2021).

18. 18.

Atkinson, J., Skinner, A., Lawson, K., Song, Y. C. & Hickie, I. B. *Road to Recovery: Restoring Australia's Mental Wealth* (Brain and Mind Centre, University of Sydney, 2020).

19. 19.

Occhipinti, J. *et al. Front. Psychiatry* <https://doi.org/10.3389/fpsyg.2021.759343> (2021).

20. 20.

Aguas, R. *et al. BMJ Glob. Health* **5**, e003126 (2020).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02581-9>

Work

- **How to shrink the gap that holds Black scientists back** [27

September 2021]

Career News • As UK universities prepare to welcome new undergraduates, a study suggests ways to level the playing field between white and minority-ethnic science students.

- **Stress testing avocados** [27 September 2021]

Where I Work • Savannah Braden monitors how fruits ripen and rot to reduce food waste.

- CAREER NEWS
- 27 September 2021
- Clarification [29 September 2021](#)

How to shrink the gap that holds Black scientists back

As UK universities prepare to welcome new undergraduates, a study suggests ways to level the playing field between white and minority-ethnic science students.

- [Ingrid Torjesen](#)



University College London, pictured during a fresher's fair, has encouraged a rise in the number of PhD students from minority ethnic groups. Credit:

Paul Davey/Alamy

Black students face barriers progressing to PhD programmes, and experience other career setbacks, because they are more likely to receive lower undergraduate degree classifications than their white peers, an analysis by a leading UK university has found.

The study, based on exam and coursework data, and published in the journal *eLife*, proposes strategies to tackle an unexplained “award gap” between undergraduate students who are white and those who are from minority ethnic groups ([L. Cramer *eLife* 10, e58971; 2021](#)).

The gap, first identified more than 25 years ago (see [go.nature.com/3czqjus](#)), refers to differences in academic performance between white students and those from Black, Asian and minority ethnic (BAME) communities. It persists, the paper finds, despite students from both groups having similar grades, on average, in their A-levels, the exams used by UK universities to select students for admission to undergraduate courses. Similar award gaps have been reported in other countries, the paper notes, including the Netherlands, the United States and Australia.



[Racism in science: the taint that lingers](#)

In March 2020, the Royal Society reported that, in 2018–19, white students were twice as likely (35.7% compared with 17.9%) to graduate with first-class honours degrees in science, technology, engineering and mathematics (STEM) than were their Black colleagues (see go.nature.com/3ehrhd).

UK honours degrees are classified as first class, for scores of 70% or higher; upper second (2:1), for 60–69%; lower second (2:2), for 50–59%; and third, for 40–49%. A first or upper second (equating to a score of 60% or higher) is considered a ‘good’ degree.

The paper warns that students without a first- or upper second-class degrees are likely to miss out on grants for graduate programmes, and are less likely to be able to access opportunities in highly sought-after research programmes at the most high-profile institutions. This disparity, says Louise Cramer, who conducted the study, could make those students’ paths to the upper echelons of academia more difficult, with the result that fewer will reach them.

Also, the implications of that award gap for pay and opportunities “snowball” as these students progress in their careers, says Cramer, who leads a team researching cytoskeleton mechanisms in cell polarity and cell migration at the MRC Laboratory for Molecular Cell Biology, University College London (UCL). “Not having that higher step at the beginning means your next step may not be so high.”



[Science diversified: Black researchers' perspectives](#)

“If graduates from minority ethnic groups want to stay in academia, the degree-award gap reduces their chances of getting a PhD position, particularly entry to the most competitive fully funded programmes,” she adds. “If they don’t want to stay in academia, they’re less likely to get a graduate-level job within one year after graduation, so immediately it impacts their first step and maybe their whole career and pension.”

Cramer assessed the results of 792 students who completed a cell biology module in academic years 2013–14 to 2018–19 at UCL, as part of the first, second or third year of their undergraduate programmes in biosciences or natural sciences; 47% of those students self-identified as members of a minority ethnic group, but, compared with their white counterparts, proportionately fewer of them were awarded good grades for cell biology (the average grade gap between the two groups was 8–13 percentage points).

Stereotype threat

It was not these students’ ability that was responsible for the award gap in these modules, but something that was happening in the university itself, says Cramer, who points out that all the students had entered university with similar A-level grades.

Cramer notes that the award gap in the cell biology modules was much greater for exams than for coursework, and suggests two possible reasons. The first is that the exams are essay-based, whereas the coursework is assessed in a range of ways. The second is that the exam process itself can be a form of ‘stereotype threat’ — in which negative perceptions about a specific group of people set them up to fail and thereby fulfil that stereotype.



Behavioural biologist Kevin Laland.Credit: Kevin Laland

When it comes to explaining the awards gap in undergraduate degrees as a whole, student experience might be a factor because BAME students often report experiencing racism on campus, and feeling that they do not fit in or that it's not safe to speak up about different backgrounds and experiences (see [*Nature* 597, 435–438; 2021](#)). These backgrounds are also often excluded from the curricula.



Racism in academia, and why the ‘little things’ matter

Kingston University in London, Brighton University and University of the Arts London have introduced measures to reduce the undergraduate award gap, although it is unclear which of these actions has had the most impact, Cramer says. They include taking steps to ‘decolonize’ curricula, by making the educational practice more inclusive; increasing training in unconscious bias; the blind marking of papers; giving students more feedback; and trying to nurture students’ feeling of belonging.

For the past 10 years, the undergraduate degree-award gap has narrowed by just 0.5% per year, she adds. “It’s going to take decades to close this gap. We can’t wait for that to happen — we’ve got to reduce the award gap and change recruitment criteria.”

Breaking barriers

The London Interdisciplinary Biosciences Consortium, a collaboration between eight leading universities including UCL, has removed the requirement for students to have a first-class degree to be accepted for PhD funding. And several PhD programmes in the United States, including Tetrad at the University of California, San Francisco, no longer require applicants to have passed the Graduate Record Examinations (GRE) test, a

standardized admissions test used by many US graduate schools. As a result of these and other measures, such as introducing summer research-training programmes for undergraduates, these institutions have seen a twofold to fourfold increase in students from minority ethnic groups in their PhD programmes, Cramer says.

Students from these groups fall out of the academic pathway early on, she adds. Nearly one-quarter of UK undergraduate-degree holders (23.69%) are from minority ethnic groups, but that proportion falls to 18.07% for PhD students and to 10.7% for postdocs, the paper points out.

The study notes that introducing steps to close the award gap — such as increasing the number of academics per student and hiring departmental staff to analyse performance data — will be expensive, and that “academics already work excessive hours just to get their ‘regular’ job done”.

A sense of belonging

Kevin Laland, a behavioural and evolutionary biologist and the only person of colour on the academic staff in the biology department at the University of St Andrews, UK, says: “I don’t teach a lot of students, so most who study biology here will not see a single Black or Asian lecturer or a lecturer from any other minority ethnic group” during their four-year degree. “It is easy for them to fall into the trap of thinking that they don’t belong, that they’re outsiders, that they don’t identify with what’s being taught — and if that is the case, that can be a very alienating and demotivating experience,” he adds.



Collection: Achieving diversity in research

Universities need to decolonize curricula to remove Western bias, and look at recruitment and promotion practices for academic faculty and staff members to ensure that students from minority ethnic groups are taught by more people with whom they can identify, he says.

In STEM subjects, only 3.5% of Black and 6.6% of Asian academic staff were professors in 2018–19, compared with 11.9% of white staff (see go.nature.com/3aa9r9). Faculty members from these groups were also less likely to be involved in teaching, with 48.3% having a research-only academic contract, compared with 32.3% of their white counterparts (see go.nature.com/3ehrhb).

Laland suggests that universities build on their successes in tackling gender bias. Since 2011, St Andrews has been a member of the Athena SWAN Charter (a framework used internationally to support and transform gender equality in higher education and research). The biology department implemented a number of actions, he says, including scrutinizing the wording of job advertisements, developing recruitment templates to ensure that guidelines were followed, implementing bias training, and offering flexible working and child care. “Collectively, we went from having extremely male-biased recruitment, to having no gender bias in our recruiting,” he says. “No marginal gain is too small to be dismissed.”

Nature **597**, 739-740 (2021)

doi: <https://doi.org/10.1038/d41586-021-02618-z>

Updates & Corrections

- **Clarification 29 September 2021:** An earlier version of this news story implied that the University of St Andrews made changes in response to the Athena SWAN Charter. In fact, the actions described were made by the biology department.

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02618-z>

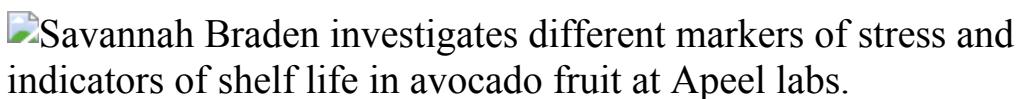
| [Section menu](#) | [Main menu](#) |

- WHERE I WORK
- 27 September 2021

Stress testing avocados

Savannah Braden monitors how fruits ripen and rot to reduce food waste.

- [Virginia Gewin](#) ⁰



Savannah Braden is senior biological scientist at Apeel in Goleta, California. Credit: Rocco Ceselin for *Nature*

My job here at Apeel, a food-systems technology company in Goleta, California, is to unlock how genes and metabolites, or products of food breakdown, control the degradation of fruits such as mango and avocado. The company uses plant-based coatings to double the shelf life of produce and reduce waste from a range of crops — including apples, mangos, cucumbers and citrus fruits.

As a senior biological scientist, I monitor fruits' basic physiology at harvest, including respiration and loss of mass. My work is very seasonal. From spring to summer, I visit local farms to sample avocados. In this picture, I'm cryofreezing slices of avocado to study markers of stress. Because avocados are alive, their cells respond to changes in their environment, including cutting, so I couldn't simply store the slices in a refrigerator. The slices have to be cryopreserved so that I can examine those specific stress markers, which can be strong predictors of whether the avocado will mould, turn brown or have disorders, which are often linked to how long produce will last post-harvest.

I've always been fascinated by how the natural world can inspire technology, and that is the focus of several laboratories at the University of California, Santa Barbara, where I finished my undergraduate degree in 2015 and my master's in 2016.

While there, I heard of Apeel, where I began working in 2016. James Rogers, Apeel's chief executive and founder, was researching how to reduce food waste and formed the company in 2012. Our goal is to make sure food gets turned into the most appropriate product that can reach a person who needs it. For example, if we can identify which avocados might ripen more rapidly, they can be targeted for guacamole.

We're at the forefront of understanding food in a different way. People want to know where it comes from and how it was grown. The goal of my job is to read the mind of avocados.

Nature **597**, 742 (2021)

doi: <https://doi.org/10.1038/d41586-021-02619-y>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02619-y>.

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

Research

- **[Heat management with a twist in layered materials](#)** [29 September 2021]
News & Views • The misalignment of crystal lattices in stacked monolayers of materials has been shown to prevent heat flow between the layers, while retaining flow within them. This finding opens up an inventive way to control heat at the nanoscale.
- **[Pollination advantage of rare plants unveiled](#)** [08 September 2021]
News & Views • An analysis of plant–pollinator interactions reveals that the presence of abundant plant species favours the pollination of rare species. Such asymmetric facilitation might promote the coexistence of species in diverse plant communities.
- **[Electron crystals come under the microscope](#)** [29 September 2021]
News & Views • In 2D materials, electrons at low densities can freeze into well-defined positions and form exotic structures called Wigner crystals. A non-invasive technique has been developed to image these crystals directly.
- **[Insights into the origins of pancreatic cancer](#)** [15 September 2021]
News & Views • The initial events that give rise to pancreatic cancer are not fully understood. Evidence from mice now implicates the enzyme Tert in setting the stage for the formation of this type of tumour.
- **[Protein condensates provide a platform for controlling chromatin](#)** [15 September 2021]
News & Views • The protein UTX regulates the DNA–protein complex chromatin to suppress tumour growth. Data suggest that the ability of UTX to condense into liquid-like droplets underlies its chromatin-regulating ability.
- **[From the archive](#)** [28 September 2021]
News & Views • Nature’s pages feature a report from a conference about the search for extraterrestrial life, and a reflection on scientific books.
- **[The importance of lake breach floods for valley incision on early Mars](#)** [29 September 2021]
Article • Lake breach flooding rapidly eroded almost a quarter of the volume of incised valleys on early Mars, influencing the topography of the wider Martian landscape.

- [**Imaging two-dimensional generalized Wigner crystals**](#) [29 September 2021]
Article • So far, only indirect evidence of Wigner crystals has been reported, but a specially designed scanning tunnelling microscope is used here to directly image them in a moiré heterostructure.
- [**Non-Hermitian topological whispering gallery**](#) [29 September 2021]
Article • An acoustic topological gallery insulator constructed from sonic crystals made of thermoplastic rods decorated with carbon nanotube films enables the out-coupling of amplified and focused sound at audible frequencies.
- [**Extremely anisotropic van der Waals thermal conductors**](#) [29 September 2021]
Article • Extremely anisotropic thermal conductors based on large-area van der Waals thin films with random interlayer rotations are reported here.
- [**The role of charge recombination to triplet excitons in organic solar cells**](#) [29 September 2021]
Article • A substantial pathway for energy loss in organic solar cells may be suppressed by engineering hybridization between non-fullerene acceptor triplet excitons and spin-triplet charge transfer excitons.
- [**Skilful precipitation nowcasting using deep generative models of radar**](#) [29 September 2021]
Article • A deep generative model using radar observations is used to create skilful precipitation predictions that are accurate and support real-world utility.
- [**Mercury stable isotopes constrain atmospheric sources to the ocean**](#) [29 September 2021]
Article • Mercury deposition pathways from the atmosphere to the ocean remain uncertain, but mercury stable isotope measurements from the Atlantic and Mediterranean show that ocean uptake of gaseous elemental mercury is more important than previously thought.
- [**Fine-root traits in the global spectrum of plant form and function**](#) [29 September 2021]
Article • The authors analyse the coordination and trade-off of the aboveground and fine-root traits of vascular plants using global trait databases.
- [**Pollinators contribute to the maintenance of flowering plant diversity**](#) [08 September 2021]
Article • The authors show that plant–pollinator interactions confer a fitness advantage to rare plant species in a biodiversity hotspot.

- **Genetic and epigenetic coordination of cortical interneuron development** [22 September 2021]
Article • The transcriptional programs and chromatin accessibility changes that cause two related interneuronal subtypes to diverge from common developmental origins and acquire specialized properties are elucidated.
- **Rational design of a new antibiotic class for drug-resistant infections** [15 September 2021]
Article • A lead-optimization strategy combining porin permeation properties and biochemical potency leads to development of a new class of antibiotic based on broad inhibition of penicillin-binding proteins from Gram-negative bacteria.
- **Emergence and expansion of SARS-CoV-2 B.1.526 after identification in New York** [24 August 2021]
Article • The dynamics of the spread of the SARS-CoV-2 variant B.1.526 suggest that resistance to neutralization by antibodies may evolve in other variants and contribute to the spread of COVID-19.
- **A lymphocyte–microglia–astrocyte axis in chronic active multiple sclerosis** [08 September 2021]
Article • Single-nucleus transcriptomics defines a diverse set of immune and glial cells at the chronically inflamed leading edge of demyelinated white matter lesions in patients with multiple sclerosis.
- **Acinar cell clonal expansion in pancreas homeostasis and carcinogenesis** [15 September 2021]
Article • A rare population of acinar cells expressing telomerase reverse transcriptase renew the acinar cell compartment during homeostasis, and are potential sources of premalignant cells in pancreatic carcinogenesis.
- **Programmable RNA targeting with the single-protein CRISPR effector Cas7-11** [06 September 2021]
Article • Cas7-11—the fusion of a putative Cas11 domain and four Cas7 subunits—cleaves RNA without detectable non-specific activity and, when optimized for RNA knockdown and editing in mammalian cells, has no effects on cell viability.
- **UTX condensation underlies its tumour-suppressive activity** [15 September 2021]
Article • Phase separation properties are a major determinant of UTX activity in chromatin regulation in tumour suppression, and are dependent on a core intrinsically disordered region of the protein.
- **Structure-based classification predicts drug response in EGFR-mutant NSCLC** [15 September 2021]

Article • Structural classification of mutations in the epidermal growth factor receptor causing non-small cell lung cancer is a better predictor of patient outcomes following drug treatment than traditional exon-based classification.

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

- NEWS AND VIEWS
- 29 September 2021

Heat management with a twist in layered materials

The misalignment of crystal lattices in stacked monolayers of materials has been shown to prevent heat flow between the layers, while retaining flow within them. This finding opens up an inventive way to control heat at the nanoscale.

- [Mariusz Zdrojek](#) [ORCID: http://orcid.org/0000-0002-8897-6205](#) ⁰

Any electrical device will eventually break if it overheats. Heat management is therefore a constant challenge in the design of electrical systems — especially if the miniaturization of electronic devices is to continue, because the downsizing of electronic components results in greater amounts of heat being generated per unit volume of a device. [Writing in Nature](#), Kim *et al.*¹ report findings that might enable heat management to be achieved in dense electronic circuitry by means of highly directional heat dissipation. The trick is to use materials that consist of stacked, atomically thin layers.

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 **597**, 637–638 (2021)

doi: <https://doi.org/10.1038/d41586-021-02572-w>

References

1. 1.

Kim, S. E. *et al.* *Nature* **597**, 660–665 (2021).

2. 2.

Ghosh, S. *et al.* *Nature Mater.* **9**, 555–558 (2010).

3. 3.

Minnich, A. J. *Nanoscale Microscale Thermophys. Eng.* **20**, 1–21 (2016).

4. 4.

Jiang, P., Qian, X., Yang, R. & Lindsay, L. *Phys. Rev. Mater.* **2**, 064005 (2018).

5. 5.

Cao, Y. *et al. Nature* **556**, 43–50 (2018).

6. 6.

Hennighausen, Z. & Kar, S. *Electron. Struct.* **3**, 014004 (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02572-w>

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

- NEWS AND VIEWS
- 08 September 2021

Pollination advantage of rare plants unveiled

An analysis of plant–pollinator interactions reveals that the presence of abundant plant species favours the pollination of rare species. Such asymmetric facilitation might promote the coexistence of species in diverse plant communities.

- [Marcelo A. Aizen](#) [ORCID: http://orcid.org/0000-0001-9079-9749](http://orcid.org/0000-0001-9079-9749) ⁰

Species diversification results from the balance between the formation of new species (speciation) and the loss of existing ones (extinction). The tremendous proliferation of different life forms on Earth can be attributed to both high rates of speciation and low rates of extinction. Flowering plants — a group called angiosperms — are one of the most diverse groups of non-mobile organism. There are approximately 352,000 plant species, nearly 90% of which depend, to various extents, on insects and other animals for pollination and seed production¹. These animal pollinators have been key to the unstoppable diversification of the angiosperms, starting at least 120 million years ago, with pollinators promoting speciation by acting as potent selection agents for a plethora of flower traits^{2,3}. Pollinators also aid species persistence by enabling pollen transfer between relatively distant individuals in sparse plant populations⁴. [Writing in Nature](#), Wei *et al.*⁵ report that, for plant species that flower at the same time, pollinators mediate interactions that might facilitate species coexistence in diverse plant communities.

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 **597**, 638-639 (2021)

doi: <https://doi.org/10.1038/d41586-021-02375-z>

References

1. 1.

Ollerton, J., Winfree, R. & Tarrant, S. *Oikos* **120**, 321–326 (2011).

2. 2.

Hernández-Hernández, T. & Wiens, J. J. *J. Am. Nat.* **195**, 948–963 (2020).

3. 3.

van der Niet, T. & Johnson, S. D. *Trends Ecol. Evol.* **27**, 353–361 (2012).

4. 4.

Bawa, K. S. *Trends Ecol. Evol.* **10**, 311–312 (1995).

5. 5.

Wei, N. *et al.* *Nature* **597**, 688–692 (2021).

6. 6.

Hegland, S. J., Grytnes, J.-A. & Totland, Ø. *Ecol. Res.* **24**, 929–936 (2009).

7. 7.

Sargent, R. D. & Ackerly, D. D. *Trends Ecol. Evol.* **23**, 123–130 (2008).

8. 8.

Ghazoul, J. *J. Ecol.* **94**, 295–304 (2006).

9. 9.

Tur, C., Sáez, A., Traveset, A. & Aizen, M. A. *Ecol. Lett.* **19**, 576–586 (2016).

10. 10.

Bergamo, P. J., Streher, N. S., Traveset, A., Wolowski, M. & Sazima, M. *Ecol. Lett.* **23**, 129–139 (2020).

11. 11.

Morales, C. L. & Traveset, A. *Crit. Rev. Plant Sci.* **27**, 221–238 (2008).

12. 12.

Silvertown, J., Franco, M., Pisanty, I. & Mendoza, A. *J. Ecol.* **81**, 465–476 (1993).

13. 13.

IPBES. *The Assessment Report on Pollinators, Pollination and Food Production of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services* (IPBES, 2016).

14. 14.

Knight, T. M. *et al. Annu. Rev. Ecol. Syst.* **36**, 467–497 (2005).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02375-z>

- NEWS AND VIEWS
- 29 September 2021

Electron crystals come under the microscope

In 2D materials, electrons at low densities can freeze into well-defined positions and form exotic structures called Wigner crystals. A non-invasive technique has been developed to image these crystals directly.

- [Carmen Rubio-Verdú](#) ⁹

In 1934, the physicist Eugene Wigner predicted that electrons could form crystals¹. Evidence of such Wigner crystals has been seen in liquid helium-4^{2,3}, in 2D electron gases under strong magnetic fields^{4,5} and in 2D semiconductors^{6–8}. However, these crystals have been difficult to observe directly because they are fragile and can be altered by any small perturbation. [Writing in Nature](#), Li *et al.*⁹ report the direct imaging of Wigner crystals in a 2D material.

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 **597**, 640-641 (2021)

doi: <https://doi.org/10.1038/d41586-021-02573-9>

References

1. 1.

Wigner, E. *Phys. Rev.* **46**, 1002–1011 (1934).

2. 2.

Williams, R., Crandall, R. S. & Willis, A. H. *Phys. Rev. Lett.* **26**, 7–9 (1971).

3. 3.

Grimes, C. C. & Adams, G. *Phys. Rev. Lett.* **42**, 795–798 (1979).

4. 4.

Andrei, E. Y., Deville, G., Glattli, D. C. & Williams, F. I. B. *Phys. Rev. Lett.* **60**, 2765–2768 (1988).

5. 5.

Jiang, H. W. *et al.* *Phys. Rev. Lett.* **65**, 633–636 (1990).

6. 6.

Zhou, Y. *et al.* *Nature* **595**, 48–52 (2021).

7. 7.

Smoleński, T. *et al.* *Nature* **595**, 53–57 (2021).

8. 8.

Regan, E. C. *et al.* *Nature* **579**, 359–363 (2020).

9. 9.

Li, H. *et al.* *Nature* **597**, 650–654 (2021).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02573-9>

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

- NEWS AND VIEWS
- 15 September 2021

Insights into the origins of pancreatic cancer

The initial events that give rise to pancreatic cancer are not fully understood. Evidence from mice now implicates the enzyme Tert in setting the stage for the formation of this type of tumour.

- [Laura D. Wood](#) ⁹ &
- [Anirban Maitra](#) ¹

The pancreas secretes enzymes that are required for the digestion of food in the intestine. Two main pancreatic cell types are involved in this — acinar cells, which secrete these enzymes, and ductal cells, which line channels to the intestine. The most common type of pancreatic cancer arises in a form that recapitulates such pancreatic ducts, both in its gland-like structure and in its pattern of protein expression. Yet, despite this similarity, substantial uncertainty remains regarding the cell type (cell of origin) that gives rise to pancreatic cancer.

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 **597**, 641-642 (2021)

doi: <https://doi.org/10.1038/d41586-021-02435-4>

References

1. 1.

Grimont, A., Leach, S. D. & Chadwani, R. *Cell. Mol. Gastroenterol. Hepatol.* <https://doi.org/10.1016/j.jcmgh.2021.07.014> (2021).

2. 2.

Neuhöfer, P. *et al.* *Nature* **597**, 715–719 (2021).

3. 3.

Fischer, C. G. & Wood, L. D. *J. Pathol.* **246**, 395–404 (2018).

4. 4.

Kipling, D. & Cooke, H. J. *Nature* **347**, 400–402 (1990).

5. 5.

van Heek, N. T. *et al.* *Am. J. Pathol.* **161**, 1541–1547 (2002).

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02435-4>

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

- NEWS AND VIEWS
- 15 September 2021

Protein condensates provide a platform for controlling chromatin

The protein UTX regulates the DNA–protein complex chromatin to suppress tumour growth. Data suggest that the ability of UTX to condense into liquid-like droplets underlies its chromatin-regulating ability.

- [David Lara-Astiaso](#)⁰ &
- [Brian J. P. Huntly](#)¹

In a process called liquid–liquid phase separation, proteins and nucleic acids self-organize into liquid-like droplets called condensates¹. These dynamic, membraneless compartments enable molecules to co-localize at the optimal concentrations needed for different reactions, excluding undesired molecules, and thereby controlling complex biochemical processes in space and time². A growing number of proteins have been described that contain intrinsically disordered regions (IDRs) and can therefore form condensates, and an increasing number of processes have been reported to be controlled by liquid–liquid phase separation. [Writing in Nature](#), Shi *et al.*³ add UTX to this list of proteins, and tumour suppression to the processes controlled by liquid–liquid phase separation.

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 **597**, 642–644 (2021)

doi: <https://doi.org/10.1038/d41586-021-02365-1>

References

1. 1.

Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. *Nature Rev. Mol. Cell Biol.* **18**, 285–298 (2017).

2. 2.

Feric, M. & Mistelli, T. *Trends Cell Biol.* **8**, 671–685 (2021).

3. 3.
Shi, B. *et al.* *Nature* **597**, 726–731 (2021).
4. 4.
Wang, L. & Shilatifard, A. *Cancer Cell* **35**, 168–176 (2019).
5. 5.
Tran, N., Broun, A. & Ge, K. *Mol. Cell Biol.* **40**, e00341-20 (2020).
6. 6.
Gozdecka, M. *et al.* *Nature Genet.* **50**, 883–894 (2018).
7. 7.
van Haaften, G. *et al.* *Nature Genet.* **41**, 521–523 (2009).
8. 8.
Dunford, A. *et al.* *Nature Genet.* **49**, 10–16 (2017).
9. 9.
Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K. & Sharp, P. A. *Cell* **169**, 13–23 (2017).
10. 10.
Fasciani, A. *et al.* *Nature Genet.* **52**, 1397–1411 (2020).
11. 11.
Peeples, W. & Rosen, M. K. *Nature Chem. Biol.* **17**, 693–702 (2021).

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

- NEWS AND VIEWS
- 28 September 2021

From the archive

Nature's pages feature a report from a conference about the search for extraterrestrial life, and a reflection on scientific books.

50 years ago

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 **597**, 643 (2021)

doi: <https://doi.org/10.1038/d41586-021-02574-8>

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-02574-8>

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

- Article
- [Published: 29 September 2021](#)

The importance of lake breach floods for valley incision on early Mars

- [Timothy A. Goudge](#) ORCID: [orcid.org/0000-0003-4297-9838^{1,2,3 na1}](https://orcid.org/0000-0003-4297-9838),
- [Alexander M. Morgan](#) ORCID: [orcid.org/0000-0003-2443-1676^{4,5 na1}](https://orcid.org/0000-0003-2443-1676),
- [Gaia Stucky de Quay](#) ORCID: [orcid.org/0000-0002-8598-543X^{1,2}](https://orcid.org/0000-0002-8598-543X) &
- [Caleb I. Fassett](#) ORCID: [orcid.org/0000-0001-9155-3804⁶](https://orcid.org/0000-0001-9155-3804)

[Nature](#) volume 597, pages 645–649 (2021)

- 389 Accesses
- 1214 Altmetric
- [Metrics details](#)

Subjects

- [Geomorphology](#)

Abstract

The surface environment of early Mars had an active hydrologic cycle, including flowing liquid water that carved river valleys^{1,2,3} and filled lake basins^{4,5,6}. Over 200 of these lake basins filled with sufficient water to breach the confining topography^{4,6}, causing catastrophic flooding and

incision of outlet canyons^{7,8,9,10}. Much past work has recognized the local importance of lake breach floods on Mars for rapidly incising large valleys^{7,8,9,10,11,12}; however, on a global scale, valley systems have often been interpreted as recording more persistent fluvial erosion linked to a distributed Martian hydrologic cycle^{1,2,3,13,14,15,16}. Here, we demonstrate the global importance of lake breach flooding, and find that it was responsible for eroding at least 24% of the volume of incised valleys on early Mars, despite representing only approximately 3% of total valley length. We conclude that lake breach floods were a major geomorphic process responsible for valley incision on early Mars, which in turn influenced the topographic form of many Martian valley systems and the broader landscape evolution of the cratered highlands. Our results indicate that the importance of lake breach floods should be considered when reconstructing the formative conditions for Martian valley systems.

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](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

Fig. 1: Valley networks and palaeolake outlet canyons on Mars.



Fig. 2: Cumulative distribution of depths for valley networks and palaeolake outlet canyons.

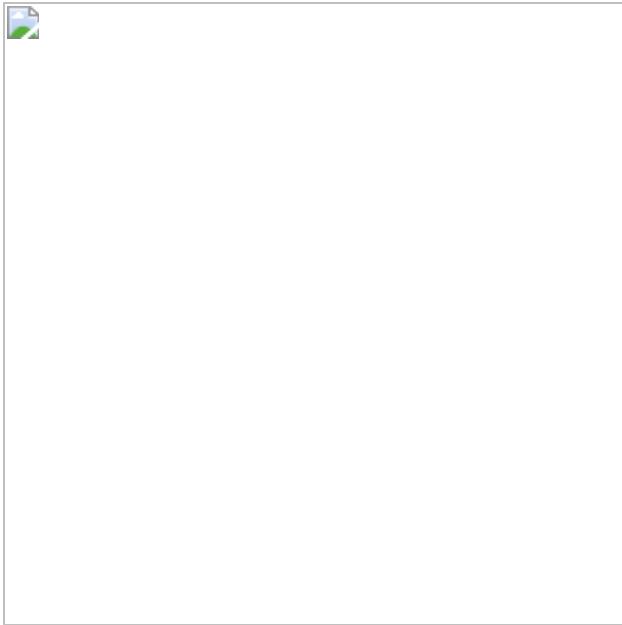
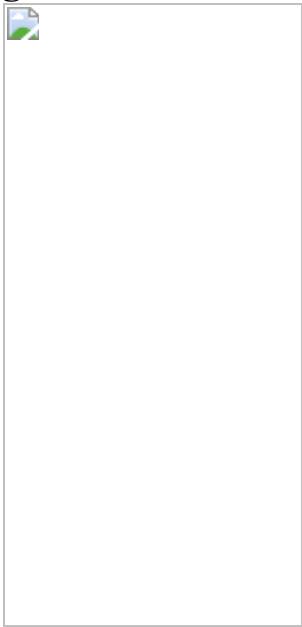


Fig. 3: Transverse valleys on Mars.



Data availability

All data used to conduct the analysis presented here have been archived through the Texas Data Repository and are available at <https://doi.org/10.18738/T8/STRFZH>. Archived data include georeferenced shapefiles (full valley catalogue, with classifications; open-basin palaeolake database; and masks) and rasters (PBTH output of valley depth).

References

1. 1.
Pieri, D. C. Martian valleys: morphology, distribution, age, and origin. *Science* **210**, 895–897 (1980).
2. 2.
Howard, A. D., Moore, J. M. & Irwin, R. P. An intense terminal epoch of widespread fluvial activity on early Mars: 1. Valley network incision and associated deposits. *J. Geophys. Res.* **110**, E12S14 (2005).
3. 3.
Hynek, B. M., Beach, M. & Hoke, M. R. T. Updated global map of Martian valley networks and implications for climate and hydrologic processes. *J. Geophys. Res.* **115**, E09008 (2010).
4. 4.
Cabrol, N. A. & Grin, E. A. Distribution, classification, and ages of Martian impact crater lakes. *Icarus* **142**, 160–172 (1999).
5. 5.
Irwin, R. P., Howard, A. D., Craddock, R. A. & Moore, J. M. An intense terminal epoch of widespread fluvial activity on early Mars: 2. Increased runoff and paleolake development. *J. Geophys. Res.* **110**, E12S15 (2005).
6. 6.
Fassett, C. I. & Head, J. W. Valley network-fed, open-basin lakes on Mars: distribution and implications for Noachian surface and subsurface hydrology. *Icarus* **198**, 37–56 (2008).
7. 7.

Irwin, R. P., Maxwell, T. A., Howard, A. D., Craddock, R. A. & Leverington, D. W. A large paleolake basin at the head of Ma'adim Vallis, Mars. *Science* **296**, 2209–2212 (2002).

8. 8.

Irwin, R. P., Howard, A. D. & Maxwell, T. A. Geomorphology of Ma'adim Vallis, Mars, and associated paleolake basins. *J. Geophys. Res.* **109**, E12009 (2004).

9. 9.

Goudge, T. A., Fassett, C. I. & Mohrig, D. Incision of paleolake outlet canyons on Mars from overflow flooding. *Geology* **47**, 7–10 (2019).

10. 10.

Irwin, R. P. & Grant, J. A. in *Megaflooding on Earth and Mars* (eds Burr, D. M. et al.) 209–224 (Cambridge Univ. Press, 2009).

11. 11.

Warner, N. H., Sowe, M., Gupta, S., Dumke, A. & Goddard, K. Fill and spill of giant lakes in the eastern Valles Marineris region of Mars. *Geology* **41**, 675–678 (2013).

12. 12.

Goudge, T. A. & Fassett, C. I. Incision of Licus Vallis, Mars from multiple lake overflow floods. *J. Geophys. Res.* **123**, 405–420 (2018).

13. 13.

Aharonson, O., Zuber, M. T., Rothman, D. H., Schorghofer, N. & Whipple, K. X. Drainage basins and channel incision on Mars. *Proc. Natl Acad. Sci.* **99**, 1780–1783 (2002).

14. 14.

Som, S. M., Montgomery, D. R. & Greenberg, H. M. Scaling relations for large Martian valleys. *J. Geophys. Res.* **114**, E02005 (2009).

15. 15.

Grau Galofre, A., Bahia, R. S., Jellinek, A. M., Whipple, K. X. & Gallo, R. Did martian valley networks substantially modify the landscape? *Earth Planet. Sci. Lett.* **547**, 116482 (2020).

16. 16.

Grau Galofre, A., Jellinek, A. M. & Osinski, G. R. Valley formation on early Mars by subglacial and fluvial erosion. *Nature Geosci.* **13**, 663–668 (2020).

17. 17.

Fassett, C. I. & Head, J. W. The timing of martian valley network activity: Constraints from buffered crater counting. *Icarus* **195**, 61–89 (2008).

18. 18.

Matsubara, Y., Howard, A. D. & Irwin, R. P. Constraints on the Noachian paleoclimate of the martian highlands from landscape evolution modeling. *J. Geophys. Res.* **123**, 2958–2979 (2018).

19. 19.

Goudge, T. A., Fassett, C. I., Head, J. W., Mustard, J. F. & Aureli, K. L. Insights into surface runoff on early Mars from paleolake basin morphology and stratigraphy. *Geology* **44**, 419–422 (2016).

20. 20.

Stucky de Quay, G., Goudge, T. A. & Fassett, C. I. Precipitation and aridity constraints from paleolakes on early Mars. *Geology* **48**, 1189–1193 (2020).

21. 21.

O'Connor, J. E. & Baker, V. R. Magnitudes and implications of peak discharges from glacial Lake Missoula. *Geol. Soc. Amer. Bull.* **104**, 267–279 (1992).

22. 22.

Gupta, S., Collier, J. S., Palmer-Felgate, A. & Potter, G. Catastrophic flooding origin of shelf valley systems in the English Channel. *Nature* **448**, 342–345 (2007).

23. 23.

Lamb, M. P. & Fonstad, M. A. Rapid formation of a modern bedrock canyon by a single flood event. *Nature Geosci.* **3**, 477–481 (2010).

24. 24.

Tanaka, K. L. et al. Geologic Map of Mars *US Geological Survey Scientific Investigations Map SIM 3292* <http://pubs.usgs.gov/sim/3292> (2014).

25. 25.

Luo, W., Pingel, T., Heo, J., Howard, A. & Jung, J. A progressive black top hat transformation algorithm for estimated valley volumes on Mars. *Comput. Geosci.* **75**, 17–23 (2015).

26. 26.

Luo, W., Cang, X. & Howard, A. D. New Martian valley network volume estimate consistent with ancient ocean and warm and wet climate. *Nat. Comm.* **8**, 15766 (2017).

27. 27.

Smith, D. E. et al. Mars Orbiter Laser Altimeter: experiment summary after the first year of global mapping of Mars. *J. Geophys. Res.* **106**,

689–23,722 (2001).

28. 28.

Mustard, J. F., Cooper, C. D. & Rifkin, M. K. Evidence for recent climate change on Mars from the identification of youthful near-surface ground ice. *Nature* **412**, 411–414 (2001).

29. 29.

Levy, J. S., Fassett, C. I., Head, J. W., Schwartz, C. & Watters, J. L. Martian water budget: geometric constraints on the volume of remnant, midlatitude debris-covered glaciers. *J. Geophys. Res.* **119**, 2188–2196 (2014).

30. 30.

Rosenberg, E. N. & Head, J. W. Late Noachian fluvial erosion on Mars: cumulative water volumes required to carve the valley networks and grain size of bed-sediment. *Planet. Space Sci.* **117**, 429–435 (2015).

31. 31.

Wordsworth, R. D. The climate of early Mars. *Annu. Rev. Earth Planet. Sci.* **44**, 381–408 (2016).

32. 32.

Stokes, M. & Mather, A. E. Tectonic origin and evolution of a transverse drainage: the Río Almanzora, Betic Cordillera, Southeast Spain. *Geomorphology* **50**, 59–81 (2003).

33. 33.

Douglass, J. & Schmeeckle, M. Analogue modeling of transverse drainage mechanisms. *Geomorphology* **84**, 22–43 (2007).

34. 34.

Hilgendorf, Z., Wells, G., Larson, P. H., Millett, J. & Kohout, M. From basins to rivers: understanding the revitalization and significance of top-down drainage integration mechanism in drainage basin evolution. *Geomorphology* **352**, 107020 (2020).

35. 35.

Irwin, R. P., Craddock, R. A., Howard, A. D. & Flemming, H. L. Topographic influences on development of Martian valley networks. *J. Geophys. Res.* **116**, E02005 (2011).

36. 36.

Black, B. A. et al. Global drainage patterns and the origins of topographic relief on Earth, Mars, and Titan. *Science* **356**, 727–731 (2017).

37. 37.

Douglass, J. C. et al. Evidence for the overflow origin of the Grand Canyon. *Geomorphology* **369**, 107361 (2020).

38. 38.

Geurts, A. H., Whittaker, A. C., Gawthorpe, R. L. & Cowie, P. A. Transient landscape and stratigraphic responses to drainage integration in the actively extending central Italian Apennines. *Geomorphology* **353**, 107013 (2020).

39. 39.

Howard, A. D., Dietrich, W. E. & Seidl, M. A. Modeling fluvial erosion on regional to continental scales. *J. Geophys. Res.* **99**, 971–12,986 (1994).

40. 40.

Berlin, M. M. & Anderson, R. S. Modeling of knickpoint retreat on the Roan Plateau, western Colorado. *J. Geophys. Res.* **112**, F03S06 (2007).

41. 41.

Fergason, R. L., Hare, T. M. & Laura, J. Mars MGS MOLA - MEX HRSC Blended DEM Global 200m v2. *Astrogeology PDS Annex, US Geological Survey* http://bit.ly/HRSC_MOLA_Bland_v0 (2018).

42. 42.

Neukum, G. et al. HRSC: the High Resolution Stereo Camera of Mars Express. *European Space Agency Special Publication ESA SP-1240*, 17–35 (2004).

43. 43.

Christensen, P. R. et al. The Thermal Emission Imaging System (THEMIS) for the Mars 2001 Odyssey mission. *Space Sci. Rev.* **110**, 85–130 (2004).

44. 44.

Edwards, C. S. et al. Mosaicking of global planetary image datasets: 1. Techniques and data processing for Thermal Emission Imaging System (THEMIS) multi-spectral data. *J. Geophys. Res.* **116**, E10008 (2011).

45. 45.

Malin, M. C. & Edgett, K. S. Evidence for recent groundwater seepage and surface runoff on Mars. *Science* **288**, 2330–2335 (2000).

46. 46.

Harrison, T. N., Osinski, G. R., Tornabene, L. L. & Jones, E. Global documentation of gullies with the Mars Reconnaissance Orbiter Context Camera and implications for their formation. *Icarus* **252**, 236–254 (2015).

47. 47.

Irwin, R. P., Watters, T. R., Howard, A. D. & Zimbelman, J. R. Sedimentary resurfacing and fretted terrain development along the crustal dichotomy boundary, Aeolis Mensae, Mars. *J. Geophys. Res.* **109**, E09011 (2004).

48. 48.

Irwin, R. P. & Watters, T. R. Geology of the Martian crustal dichotomy boundary: Age, modifications, and implications for modeling efforts. *J. Geophys. Res.* **115**, E11006 (2010).

49. 49.

Baker, V. R. & Milton, D. J. Erosion by catastrophic floods on Mars and Earth. *Icarus* **23**, 27–41 (1974).

50. 50.

Carr, M. H. Formation of martian flood features by release of water from confined aquifers. *J. Geophys. Res.* **84**, 2995–3007 (1979).

51. 51.

Coleman, N. M. & Baker, V. R. in *Megaflooding on Earth and Mars* (eds Burr, D. M. et al.) 172–193 (Cambridge Univ. Press, 2009).

52. 52.

Mest, S. C. & Crown, D. A. Geology of the Reull Vallis region, Mars. *Icarus* **153**, 89–110 (2001).

53. 53.

Kreslavsky, M. A. & Head, J. W. Kilometer-scale roughness of Mars: results from MOLA data analysis. *J. Geophys. Res.* **105**, 695–26,711 (2000).

54. 54.

Hoke, M. R. T., Hynek, B. M. & Tucker, G. E. Formation timescales of large Martian valley networks. *Earth Planet. Sci. Lett.* **312**, 1–12 (2011).

55. 55.

Luo, W. & Stepinski, T. F. Computer-generated global map of valley networks on Mars. *J. Geophys. Res.* **114**, E11010 (2009).

56. 56.

Malin, M. C. et al. Context Camera investigation on board the Mars Reconnaissance Orbiter. *J. Geophys. Res.* **112**, E05S04 (2007).

57. 57.

Shean, D. E. et al. An automated, open-source pipeline for mass production of digital elevation models (DEMs) from very-high-resolution commercial stereo satellite imagery. *ISPRS J. Photogramm. Remote Sens.* **116**, 101–117 (2016).

58. 58.

Beyer, R. A., Alexandrov, O. & McMichael, S. The Ames Stereo Pipeline: NASA’s open source software for deriving and processing terrain data. *Earth Space Sci.* **5**, 537–548 (2018).

59. 59.

Dickson, J. L., Kerber, L. A., Fassett, C. I. & Ehlmann, B. L. A global, blended CTX mosaic of Mars with vectorized seam mapping: a new mosaicking pipeline using principles of non-destructive image editing. In *49th Lunar and Planetary Science Conference* 2480 (2018).

Acknowledgements

This work was supported by funding through NASA MDAP grants 80NSSC17K0442 (T.A.G., G.S.d.Q. and C.I.F.) and 80NSSC17K0454

(A.M.M.). T.A.G. thanks E. Bamber for valuable discussions on transverse valley development on Mars, and J. Clarke for helpful manuscript discussions. This is the University of Texas at Austin Center for Planetary Systems Habitability (UT CPSH) contribution no. 0031.

Author information

Author notes

1. These authors contributed equally: Timothy A. Goudge, Alexander M. Morgan

Affiliations

1. Department of Geological Sciences, Jackson School of Geosciences, The University of Texas at Austin, Austin, TX, USA

Timothy A. Goudge & Gaia Stucky de Quay

2. Center for Planetary Systems Habitability, The University of Texas at Austin, Austin, TX, USA

Timothy A. Goudge & Gaia Stucky de Quay

3. CIFAR Azrieli Global Scholars Program, CIFAR, Toronto, Ontario, Canada

Timothy A. Goudge

4. Planetary Science Institute, Tucson, AZ, USA

Alexander M. Morgan

5. Center for Earth and Planetary Sciences, National Air and Space Museum, Smithsonian Institution, Washington, DC, USA

Alexander M. Morgan

6. NASA Marshall Space Flight Center, Huntsville, AL, USA

Caleb I. Fassett

Contributions

T.A.G. and C.I.F. conceived and designed the study, with input from A.M.M. and G.S.d.Q. T.A.G. conducted data collection, wrote the manuscript, and assisted with data analysis. A.M.M. conducted the data analysis and assisted with data collection. All authors contributed to data collection, interpretation of results, and assisted with writing of the manuscript.

Corresponding author

Correspondence to Timothy A. Goudge.

Ethics declarations

Competing interests

The authors declare no competing financial interests.

Additional information

Extended data

is available for this paper at <https://doi.org/10.1038/s41586-021-03860-1>.

Peer review information *Nature* thanks Alan Howard 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 Distribution of valleys removed from the original catalogue.

Final catalogue of valley networks in blue, removed valleys in yellow. Background is MOLA hillshade. See [Methods](#) for more complete description of each category. **a**, Inaccurate valley interpretations where the original mapping was not found to be robust. **b**, Valleys isolated to walls of craters with diameters < 300 km. **c**, Valleys associated with the crustal dichotomy boundary. **d**, Valleys associated with outflow channels. **e**, Valleys associated with Valles Marineris. **f**, Valleys associated with volcanic plains and edifices. **g**, Mawrth Vallis (yellow arrow points to valley). **h**, Uzboi Vallis (yellow arrow points to valley).

Extended Data Fig. 2 Example output from the progressive black top hat (PBTH) transformation.

Valley network depths outlined in black and palaeolake outlet canyon depths outlined in white. Palaeolake basins indicated in gold. Background is the THEMIS daytime infrared mosaic. **a**, Image centred at -9.4°N , 133.4°E . **b**, Image centred at -19.5°N , 344.4°E . See also Fig. [1a](#) for locations.

Extended Data Fig. 3 Cumulative distribution of depths for valley networks and palaeolake outlet canyons with Ma’adim Vallis split out.

Note the substantially deeper depths for Ma’adim Vallis (gold), but the consistently deeper depths for palaeolake outlet canyons even with Ma’adim Vallis removed (green). Total volume of each grouping is listed in the legend.

Extended Data Fig. 4 Palaeolake outlet canyon with hanging tributaries.

a, Mosaic of MOLA gridded topography and CTX stereo-derived DEMs B04_011272_1736-F05_037816_1709, B18_016507_1714-F20_043803_1714, and P22_009782_1707-J02_045425_1707 overlain on a mosaic of CTX images. Image centered at -9.0°N , 135.2°E . **b–d**, Topographic profiles of hanging tributaries entering the main outlet canyon. Data extracted from CTX stereo-derived DEMs. Raw data in grey, 5 point median filtered data in black. A–A' extracted from DEM P22_009782_1707-J02_045425_1707. B–B' extracted from DEM B18_016507_1714-F20_043803_1714. C–C' extracted from DEMs B18_016507_1714-F20_043803_1714 and B04_011272_1736-F05_037816_1709.

Extended Data Table 1 Total lengths of valley networks, palaeolake outlet canyons, and valleys removed from catalogue, as well as length of valley networks and palaeolake outlet canyons with the $\pm 30^{\circ}$ latitude and Early Hesperian age masks applied

Extended Data Table 2 Calculated volumes (m^3) with different age and/or latitude masks applied

Extended Data Table 3 Comparison of valley volume (m^3) estimates presented here with previous work

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Goudge, T.A., Morgan, A.M., Stucky de Quay, G. *et al.* The importance of lake breach floods for valley incision on early Mars. *Nature* **597**, 645–649 (2021). <https://doi.org/10.1038/s41586-021-03860-1>

- Received: 23 December 2020
- Accepted: 27 July 2021

- Published: 29 September 2021
- Issue Date: 30 September 2021
- DOI: <https://doi.org/10.1038/s41586-021-03860-1>

Share this article

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

[Get shareable link](#)

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

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03860-1>

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

- Article
- [Published: 29 September 2021](#)

Imaging two-dimensional generalized Wigner crystals

- [Hongyuan Li](#) ORCID: orcid.org/0000-0001-9119-5592^{1,2,3} na¹,
- [Shaowei Li](#) ORCID: orcid.org/0000-0002-4627-626X^{1,3,4,5} na¹,
- [Emma C. Regan](#) ORCID: orcid.org/0000-0002-9100-6031^{1,2,3},
- [Danqing Wang](#)^{1,2},
- [Wenyu Zhao](#)¹,
- [Salman Kahn](#) ORCID: orcid.org/0000-0002-0012-3305^{1,3},
- [Kentaro Yumigeta](#)⁶,
- [Mark Blei](#)⁶,
- [Takashi Taniguchi](#) ORCID: orcid.org/0000-0002-1467-3105⁷,
- [Kenji Watanabe](#) ORCID: orcid.org/0000-0003-3701-8119⁸,
- [Sefaattin Tongay](#) ORCID: orcid.org/0000-0001-8294-984X⁶,
- [Alex Zettl](#)^{1,3,4},
- [Michael F. Crommie](#) ORCID: orcid.org/0000-0001-8246-3444^{1,3,4}
&
- [Feng Wang](#) ORCID: orcid.org/0000-0001-8369-6194^{1,3,4}

Nature volume **597**, pages 650–654 (2021)

- 579 Accesses
- 1 Citations
- 92 Altmetric
- [Metrics details](#)

Subjects

- [Electronic properties and materials](#)
- [Quantum fluids and solids](#)
- [Scanning probe microscopy](#)
- [Two-dimensional materials](#)

Abstract

The Wigner crystal¹ has fascinated condensed matter physicists for nearly 90 years^{2,3,4,5,6,7,8,9,10,11,12,13,14}. Signatures of two-dimensional (2D) Wigner crystals were first observed in 2D electron gases under high magnetic field^{2,3,4}, and recently reported in transition metal dichalcogenide moiré superlattices^{6,7,8,9}. Direct observation of the 2D Wigner crystal lattice in real space, however, has remained an outstanding challenge. Conventional scanning tunnelling microscopy (STM) has sufficient spatial resolution but induces perturbations that can potentially alter this fragile state. Here we demonstrate real-space imaging of 2D Wigner crystals in WSe₂/WS₂ moiré heterostructures using a specially designed non-invasive STM spectroscopy technique. This employs a graphene sensing layer held close to the WSe₂/WS₂ moiré superlattice. Local STM tunnel current into the graphene layer is modulated by the underlying Wigner crystal electron lattice in the WSe₂/WS₂ heterostructure. Different Wigner crystal lattice configurations at fractional electron fillings of $n = 1/3$, $1/2$ and $2/3$, where n is the electron number per site, are directly visualized. The $n = 1/3$ and $n = 2/3$ Wigner crystals exhibit triangular and honeycomb lattices, respectively, to minimize nearest-neighbour occupations. The $n = 1/2$ state spontaneously breaks the original C3 symmetry and forms a stripe phase. Our study lays a solid foundation for understanding Wigner crystal states in WSe₂/WS₂ moiré heterostructures and provides an approach that is generally applicable for imaging novel correlated electron lattices in other systems.

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](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

Fig. 1: STM measurement of Wigner crystal states in a dual-gated WSe₂/WS₂ moiré superlattice.

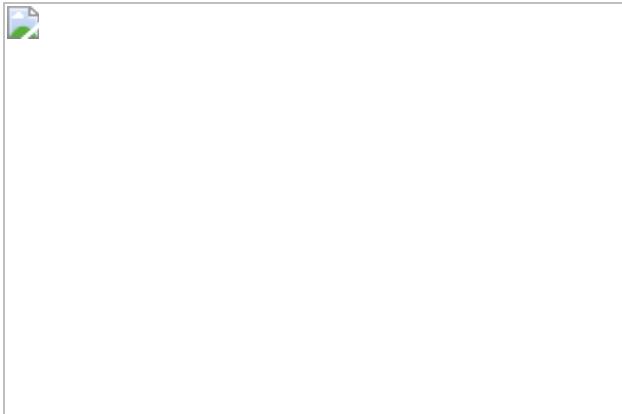


Fig. 2: Imaging Mott and generalized Wigner crystal states.

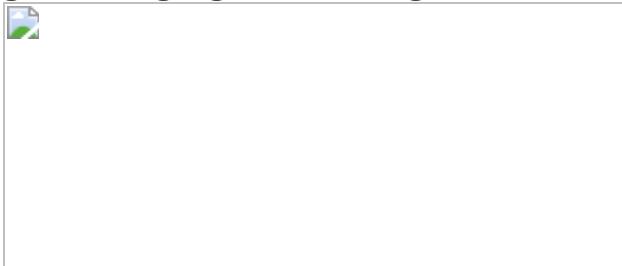
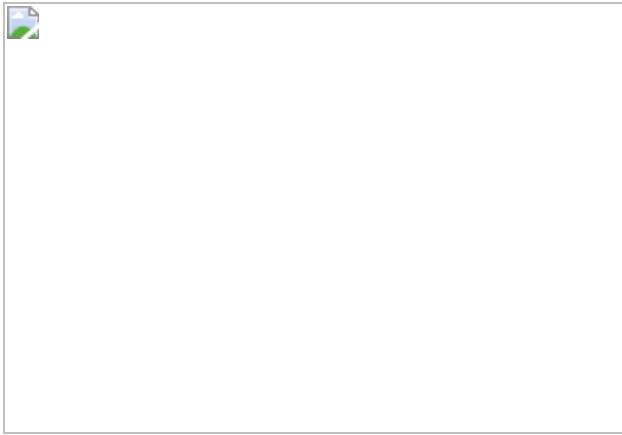


Fig. 3: Evolution of dI/dV maps for the $n = 2/3$ state with increased V_{bias} .



Data availability

The data supporting the findings of this study can be found at https://github.com/HongyuanLiCMP/Imaging_Generalized_Wigner_Crystals_data, and are also available from the corresponding authors upon reasonable request.

References

1. 1.

Wigner, E. On the interaction of electrons in metals. *Phys. Rev.* **46**, 1002–1011 (1934).

2. 2.

Goldman, V., Santos, M., Shayegan, M. & Cunningham, J. Evidence for two-dimentional quantum Wigner crystal. *Phys. Rev. Lett.* **65**, 2189–2192 (1990).

3. 3.

Jang, J., Hunt, B. M., Pfeiffer, L. N., West, K. W. & Ashoori, R. C. Sharp tunnelling resonance from the vibrations of an electronic Wigner crystal. *Nat. Phys.* **13**, 340–344 (2017).

4. 4.

Zhou, H., Polshyn, H., Taniguchi, T., Watanabe, K. & Young, A. Solids of quantum Hall skyrmions in graphene. *Nat. Phys.* **16**, 154–158 (2020).

5. 5.

Shapir, I. et al. Imaging the electronic Wigner crystal in one dimension. *Science* **364**, 870–875 (2019).

6. 6.

Regan, E. C. et al. Mott and generalized Wigner crystal states in WSe₂/WS₂ moiré superlattices. *Nature* **579**, 359–363 (2020).

7. 7.

Jin, C. et al. Stripe phases in WSe₂/WS₂ moiré superlattices. *Nat. Mater.* **20**, 940–944 (2021).

8. 8.

Xu, Y. et al. Correlated insulating states at fractional fillings of moiré superlattices. *Nature* **587**, 214–218 (2020).

9. 9.

Huang, X. et al. Correlated insulating states at fractional fillings of the WS₂/WSe₂ moiré lattice. *Nat. Phys.* **17**, 715–719 (2021).

10. 10.

Deshpande, V. V. & Bockrath, M. The one-dimensional Wigner crystal in carbon nanotubes. *Nat. Phys.* **4**, 314–318 (2008).

11. 11.

Crandall, R. & Williams, R. Crystallization of electrons on the surface of liquid helium. *Phys. Lett. A* **34**, 404–405 (1971).

12. 12.

Williams, R., Crandall, R. & Willis, A. Surface states of electrons on liquid helium. *Phys. Rev. Lett.* **26**, 7–9 (1971).

13. 13.

Grimes, C. & Adams, G. Evidence for a liquid-to-crystal phase transition in a classical, two-dimensional sheet of electrons. *Phy. Rev. Lett.* **42**, 795–798 (1979).

14. 14.

Williams, F. Collective aspects of charged-particle systems at helium interfaces. *Surface Sci.* **113**, 371–388 (1982).

15. 15.

Lam, P. K. & Girvin, S. Liquid-solid transition and the fractional quantum-Hall effect. *Phy. Rev. B* **30**, 473–475 (1984).

16. 16.

Levesque, D., Weis, J. & MacDonald, A. Crystallization of the incompressible quantum-fluid state of a two-dimensional electron gas in a strong magnetic field. *Phys. Rev. B* **30**, 1056–1058 (1984).

17. 17.

Tsui, D. C., Stormer, H. L. & Gossard, A. C. Two-dimensional magnetotransport in the extreme quantum limit. *Phys. Rev. Lett.s* **48**, 1559–1562 (1982).

18. 18.

Klitzing, K. V., Dorda, G. & Pepper, M. New method for high-accuracy determination of the fine-structure constant based on quantized Hall resistance. *Phys. Rev. Lett.* **45**, 494–497 (1980).

19. 19.

Pan, H., Wu, F. & Sarma, S. D. Quantum phase diagram of a moiré-Hubbard model. *Phys. Rev. B* **102**, 201104 (2020).

20. 20.

Hubbard, J. Generalized Wigner lattices in one dimension and some applications to tetracyanoquinodimethane (TCNQ) salts. *Phys. Rev. B* **17**, 494–505 (1978).

21. 21.

Li, H. et al. Imaging local discharge cascades for correlated electrons in WS₂/WSe₂ moiré superlattices. *Nat. Phys.* <https://doi.org/10.1038/s41567-021-01324-x> (2021).

22. 22.

Li, H. et al. Imaging moiré flat bands in three-dimensional reconstructed WSe₂/WS₂ superlattices. *Nat. Mater.* **20**, 945–950

(2021).

23. 23.

Zhang, Y. et al. Giant phonon-induced conductance in scanning tunnelling spectroscopy of gate-tunable graphene. *Nat. Phys.* **4**, 627–630 (2008).

24. 24.

Jung, S. et al. Evolution of microscopic localization in graphene in a magnetic field from scattering resonances to quantum dots. *Nat. Phys.* **7**, 245–251 (2011).

25. 25.

Decker, R. et al. Local electronic properties of graphene on a BN substrate via scanning tunneling microscopy. *Nano Lett.* **11**, 2291–2295 (2011).

26. 26.

Wong, D. et al. Spatially resolving density-dependent screening around a single charged atom in graphene. *Phys. Rev. B* **95**, 205419 (2017).

27. 27.

Yang, F. et al. Experimental determination of the energy per particle in partially filled Landau levels. *Phys. Rev. Lett.* **126**, 156802 (2021).

28. 28.

Li, T. et al. Charge-order-enhanced capacitance in semiconductor moiré superlattices. *Nat. Nanotechnol.* <https://doi.org/10.1038/s41565-021-00955-8> (2021).

29. 29.

Tomarken, S. L. et al. Electronic compressibility of magic-angle graphene superlattices. *Phys. Rev. Lett.* **123**, 046601 (2019).

30. 30.

Zondiner, U. et al. Cascade of phase transitions and Dirac revivals in magic-angle graphene. *Nature* **582**, 203–208 (2020).

31. 31.

Pierce, A. T. et al. Unconventional sequence of correlated Chern insulators in magic-angle twisted bilayer graphene. Preprint at <https://arxiv.org/abs/2101.04123> (2021).

32. 32.

Pradhan, N. A., Liu, N., Silien, C. & Ho, W. Atomic scale conductance induced by single impurity charging. *Phys. Rev. Lett.* **94**, 076801 (2005).

33. 33.

Brar, V. W. et al. Gate-controlled ionization and screening of cobalt adatoms on a graphene surface. *Nat. Phys.* **7**, 43–47 (2011).

34. 34.

Wong, D. et al. Characterization and manipulation of individual defects in insulating hexagonal boron nitride using scanning tunnelling microscopy. *Nat. Nanotechnol.* **10**, 949–953 (2015).

35. 35.

Teichmann, K. et al. Controlled charge switching on a single donor with a scanning tunneling microscope. *Phys. Rev. Lett.* **101**, 076103 (2008).

36. 36.

Wang, L. et al. One-dimensional electrical contact to a two-dimensional material. *Science* **342**, 614–617 (2013).

37. 37.

Schutte, W., De Boer, J. & Jellinek, F. Crystal structures of tungsten disulfide and diselenide. *J. Solid State Chem.* **70**, 207–209 (1987).

38. 38.

Kerelsky, A. et al. Maximized electron interactions at the magic angle in twisted bilayer graphene. *Nature* **572**, 95–100 (2019).

Acknowledgements

This work was primarily funded by the US Department of Energy, Office of Science, Office of Basic Energy Sciences, Materials Sciences and Engineering Division under contract no. DE-AC02-05-CH11231 (van der Waals heterostructure program KCFW16) (device electrode preparation and STM spectroscopy). Support was also provided by the US Army Research Office under MURI award W911NF-17-1-0312 (device layer transfer), and by the National Science Foundation Award DMR-1807233 (surface preparation). S.T. acknowledges support from DOE-SC0020653, NSF DMR 2111812, NSF DMR 1552220, NSF 2052527, DMR 1904716 and NSF CMMI 1933214 for WSe₂ and WS₂ bulk crystal growth and analysis. K.W. and T.T. acknowledge support from the Elemental Strategy Initiative conducted by the MEXT, Japan, grant number JPMXP0112101001, JSPS KAKENHI grant number JP20H00354 and the CREST(JPMJCR15F3), JST for bulk hBN crystal growth and analysis. E.C.R. acknowledges support from the Department of Defense (DoD) through the National Defense Science and Engineering Graduate Fellowship (NDSEG) Program. S.L. acknowledges support from Kavli ENSI Heising Simons Junior Fellowship. We also thank M. H. Naik for sharing unpublished theoretical simulation data on the WSe₂/WS₂ moiré superlattice.

Author information

Author notes

1. These authors contributed equally: Hongyuan Li, Shaowei Li

Affiliations

1. Department of Physics, University of California at Berkeley, Berkeley, CA, USA

Hongyuan Li, Shaowei Li, Emma C. Regan, Danqing Wang, Wenyu Zhao, Salman Kahn, Alex Zettl, Michael F. Crommie & Feng Wang

2. Graduate Group in Applied Science and Technology, University of California at Berkeley, Berkeley, CA, USA

Hongyuan Li, Emma C. Regan & Danqing Wang

3. Materials Sciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Hongyuan Li, Shaowei Li, Emma C. Regan, Salman Kahn, Alex Zettl, Michael F. Crommie & Feng Wang

4. Kavli Energy Nano Sciences Institute at the University of California Berkeley and the Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Shaowei Li, Alex Zettl, Michael F. Crommie & Feng Wang

5. Department of Chemistry and Biochemistry, University of California at San Diego, La Jolla, CA, USA

Shaowei Li

6. School for Engineering of Matter, Transport and Energy, Arizona State University, Tempe, AZ, USA

Kentaro Yumigeta, Mark Blei & Sefaattin Tongay

7. International Center for Materials Nanoarchitectonics, National Institute for Materials Science, Tsukuba, Japan

Takashi Taniguchi

8. Research Center for Functional Materials, National Institute for Materials Science, Tsukuba, Japan

Kenji Watanabe

Contributions

M.F.C. and F.W. conceived the project. H.L. and S.L. performed the STM measurement, H.L., E.C.R., D.W., W.Z. and S.K. fabricated the heterostructure device and performed the SHG measurement. K.Y., M.B. and S.T. grew WSe₂ and WS₂ crystals. K.W. and T.T. grew the hBN single crystal. All authors discussed the results and wrote the manuscript.

Corresponding authors

Correspondence to Shaowei Li or Michael F. Crommie or Feng Wang.

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work.

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

Extended data figures and tables

Extended Data Fig. 1 Comparison of single dI/dV spectra obtained at $V_{\text{TG}} = 0$ and $V_{\text{TG}} = 0.53$ V.

a, $V_{\text{TG}} = 0$. **b**, $V_{\text{TG}} = 0.53$ V. In **a** we display dI/dV spectra obtained when the graphene doping is near the CNP. A strong spectral change is observed when the graphene transitions from hole-doped to electron-doped. The Dirac point positions are denoted by vertical arrows. In **b** we display typical dI/dV spectra at $n = 1/3, 1/2, 2/3$ and 1 for correlated states (red) as well as for three other filling factors that lack correlated states (black). In each panel the dI/dV spectra are shifted vertically for clarity. The spectra indicate that the graphene sensing layer is more electron doped when the moiré heterostructure is in a correlated insulator state.

Extended Data Fig. 2 Moiré site dependence of the dI/dV spectra.

a, A typical STM topographic image of the moiré superlattice seen through the graphene sensing layer. **b–d**, Position dependent dI/dV spectra measured along the red linecut shown in **a** with $V_{\text{TG}} = 0.7$ V and $V_{\text{BG}} = 19$ V (**b**), 26.5 V and 35 V (**d**).

Extended Data Fig. 3 Raw images and FFT filtering of the dI/dV maps for the generalized Wigner crystal states.

a, Raw dI/dV map of the $n = 2/3$ state. **b**, FFT image of **a**. **c**, Real space dI/dV map after FFT filtering of **a**. In the filtering process, we removed the Fourier components within the six red circles indicated in **b**. This FFT filtering suppresses the periodic feature associated with the moiré superlattice. **d**, FFT image of **c**. **e**, Raw dI/dV map of the $n = 1/3$ state. **f**, FFT image of **e**. **g**, Real space dI/dV map after FFT filtering of **e**. The Fourier components within the red circles shown in **f** have been filtered out. **h**, FFT image of **g**. **i**, Raw dI/dV map of the $n = 1/2$ state. **j**, FFT image of **i**.

k, Real-space dI/dV map after FFT filtering of **i**. The Fourier components within the red circles shown in **j** have been filtered out. **I**, FFT image of **k**.

Extended Data Fig. 4 Uniaxial strain of the moiré superlattice.

a, Topography image shown in Fig. 1b. **b**, the corresponding FFT image. **c**, **d**, 1D height modulation along three directions (**c**; denoted by the red arrows in **a**), and the corresponding FFT results (**d**). The different moiré periods along the three different directions yields a uniaxial strain of 0.39% along the pink double-arrowed line (**a**, **b**). As a comparison, the stripe direction of the $n = 1/2$ Wigner crystal state is denoted by the yellow double-arrowed line.

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Li, H., Li, S., Regan, E.C. *et al.* Imaging two-dimensional generalized Wigner crystals. *Nature* **597**, 650–654 (2021).
<https://doi.org/10.1038/s41586-021-03874-9>

- Received: 03 April 2021
- Accepted: 04 August 2021
- Published: 29 September 2021
- Issue Date: 30 September 2021
- DOI: <https://doi.org/10.1038/s41586-021-03874-9>

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

- [**This is what a solid made of electrons looks like**](#)

- Davide Castelvecchi

Nature (2021)

[**Electron crystals come under the microscope**](#)

- Carmen Rubio-Verdú

News & Views 29 Sept 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03874-9>

- Article
- [Published: 29 September 2021](#)

Non-Hermitian topological whispering gallery

- [Bolun Hu](#) ORCID: [orcid.org/0000-0003-4810-7643¹](https://orcid.org/0000-0003-4810-7643),
- [Zhiwang Zhang](#) ORCID: [orcid.org/0000-0002-0287-0224¹](https://orcid.org/0000-0002-0287-0224),
- [Haixiao Zhang](#)¹,
- [Liyang Zheng](#)²,
- [Wei Xiong](#)¹,
- [Zichong Yue](#)¹,
- [Xiaoyu Wang](#)¹,
- [Jianyi Xu](#)¹,
- [Ying Cheng](#) ORCID: [orcid.org/0000-0002-9140-4742¹](https://orcid.org/0000-0002-9140-4742),
- [Xiaojun Liu](#) ORCID: [orcid.org/0000-0002-7826-9742¹](https://orcid.org/0000-0002-7826-9742) &
- [Johan Christensen](#) ORCID: [orcid.org/0000-0002-1604-250X²](https://orcid.org/0000-0002-1604-250X)

[Nature](#) volume **597**, pages 655–659 (2021)

- 382 Accesses
- 19 Altmetric
- [Metrics details](#)

Subjects

- [Acoustics](#)
- [Topological insulators](#)

Abstract

In 1878, Lord Rayleigh observed the highly celebrated phenomenon of sound waves that creep around the curved gallery of St Paul's Cathedral in London^{1,2}. These whispering-gallery waves scatter efficiently with little diffraction around an enclosure and have since found applications in ultrasonic fatigue and crack testing, and in the optical sensing of nanoparticles or molecules using silica microscale toroids. Recently, intense research efforts have focused on exploring non-Hermitian systems with cleverly matched gain and loss, facilitating unidirectional invisibility and exotic characteristics of exceptional points^{3,4}. Likewise, the surge in physics using topological insulators comprising non-trivial symmetry-protected phases has laid the groundwork in reshaping highly unconventional avenues for robust and reflection-free guiding and steering of both sound and light^{5,6}. Here we construct a topological gallery insulator using sonic crystals made of thermoplastic rods that are decorated with carbon nanotube films, which act as a sonic gain medium by virtue of electro-thermoacoustic coupling. By engineering specific non-Hermiticity textures to the activated rods, we are able to break the chiral symmetry of the whispering-gallery modes, which enables the out-coupling of topological 'audio lasing' modes with the desired handedness. We foresee that these findings will stimulate progress in non-destructive testing and acoustic sensing.

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](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

Fig. 1: Complex band diagram of a sonic topological insulator with acoustic gain.

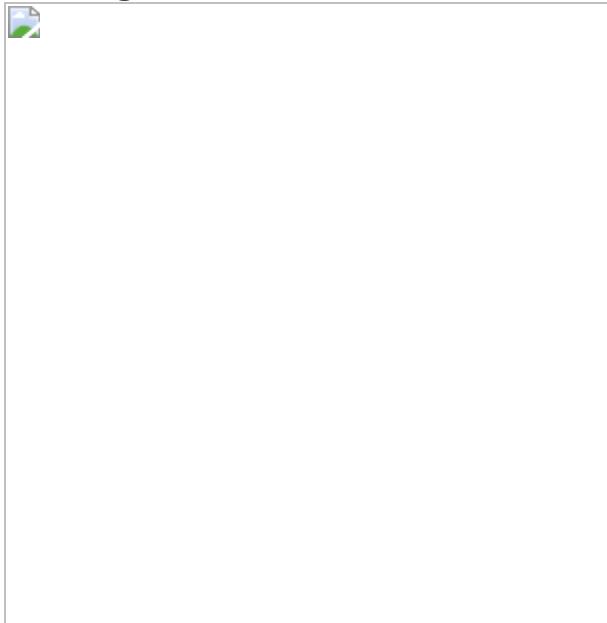


Fig. 2: Assembly and non-Hermitian phase engineering.

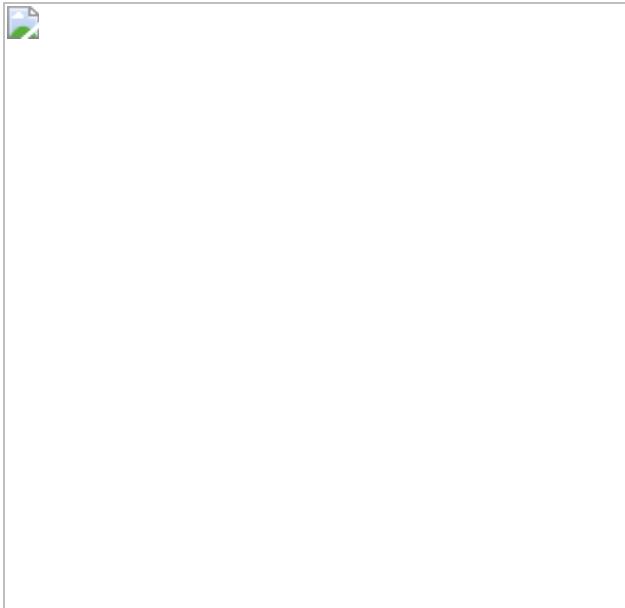


Fig. 3: Topological WG mode splitting.

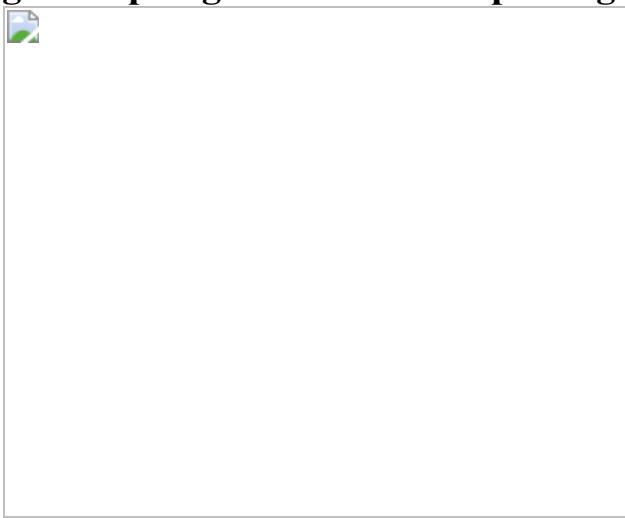
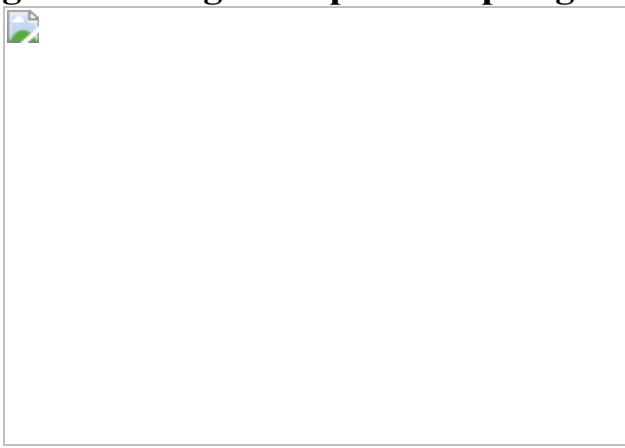


Fig. 4: Routing of amplified topological WG modes.



Data availability

The data that support the findings of this study are available from the corresponding authors on reasonable request.

Code availability

The code used to calculate the results for this work is available from the corresponding authors on reasonable request.

References

1. 1.

Lord Rayleigh *The Theory of Sound* Vol. II, 1st edn (MacMillan, 1878).

2. 2.

Lord Rayleigh CXII. The problem of the whispering gallery. *Phil. Mag.* **20**, 1001–1004 (1910).

3. 3.

Fleury, R., Sounas, D. L. & Alù, A. Parity-time symmetry in acoustics: theory, devices, and potential applications. *IEEE J. Sel. Top. Quantum Electron.* **22**, 121–129 (2016).

4. 4.

Gupta, S. K. et al. Parity-time symmetry in non-Hermitian complex optical media. *Adv. Mater.* **32**, 1903639 (2020).

5. 5.

Zhang, X., Xiao, M., Cheng, Y., Lu, M.-H. & Christensen, J. Topological sound. *Commun. Phys.* **1**, 97 (2018).

6. 6.

Khanikaev, A. B. et al. Photonic topological insulators. *Nat. Mater.* **12**, 233–239 (2013).

7. 7.

Gong, Z. et al. Topological phases of non-Hermitian systems. *Phys. Rev. X* **8**, 031079 (2018).

8. 8.

Foa Torres, L. E. F. Perspective on topological states of non-Hermitian lattices. *J. Phys. Mater.* **3**, 014002 (2019).

9. 9.

Lee, T. E. Anomalous edge state in a non-Hermitian lattice. *Phys. Rev. Lett.* **116**, 133903 (2016).

10. 10.

Wang, M., Ye, L., Christensen, J. & Liu, Z. Valley physics in non-Hermitian artificial acoustic boron nitride. *Phys. Rev. Lett.* **120**, 246601 (2018).

11. 11.

Zhang, Z., López, M. R., Cheng, Y., Liu, X. & Christensen, J. Non-Hermitian sonic second-order topological insulator. *Phys. Rev. Lett.* **122**, 195501 (2019).

12. 12.

Zhao, H. et al. Non-Hermitian topological light steering. *Science* **365**, 1163–1166 (2019).

13. 13.

Yao, S. & Wang, Z. Edge states and topological invariants of non-Hermitian systems. *Phys. Rev. Lett.* **121**, 086803 (2018).

14. 14.

Song, F., Yao, S. & Wang, Z. Non-Hermitian topological invariants in real space. *Phys. Rev. Lett.* **123**, 246801 (2019).

15. 15.

Okuma, N., Kawabata, K., Shiozaki, K. & Sato, M. Topological origin of non-Hermitian skin effects. *Phys. Rev. Lett.* **124**, 086801 (2020).

16. 16.

Xiao, L. et al. Non-Hermitian bulk–boundary correspondence in quantum dynamics. *Nat. Phys.* **16**, 761–766 (2020).

17. 17.

Helbig, T. et al. Generalized bulk–boundary correspondence in non-Hermitian topoelectrical circuits. *Nat. Phys.* **16**, 747–750 (2020).

18. 18.

Weidemann, S. et al. Topological funneling of light. *Science* **368**, 311–314 (2020).

19. 19.

Parto, M. et al. Edge-mode lasing in 1D topological active arrays. *Phys. Rev. Lett.* **120**, 113901 (2018).

20. 20.

St-Jean, P. et al. Lasing in topological edge states of a one-dimensional lattice. *Nat. Photon.* **11**, 651–656 (2017).

21. 21.

Zhao, H. et al. Topological hybrid silicon microlasers. *Nat. Commun.* **9**, 981 (2018).

22. 22.

Bandres, M. A. et al. Topological insulator laser: experiments. *Science* **359**, eaar4005 (2018).

23. 23.

Klempt, S. et al. Exciton-polariton topological insulator. *Nature* **562**, 552–556 (2018).

24. 24.

Bahari, B. et al. Nonreciprocal lasing in topological cavities of arbitrary geometries. *Science* **358**, 636–640 (2017).

25. 25.

Zeng, Y. et al. Electrically pumped topological laser with valley edge modes. *Nature* **578**, 246–250 (2020).

26. 26.

Hutson, A. R., McFee, J. H. & White, D. L. Ultrasonic amplification in CdS. *Phys. Rev. Lett.* **7**, 237–239 (1961).

27. 27.

Arnold, H. & Crandall, I. The thermophone as a precision source of sound. *Phys. Rev.* **10**, 22–38 (1917).

28. 28.

Xiao, L. et al. Flexible, stretchable, transparent carbon nanotube thin film loudspeakers. *Nano Lett.* **8**, 4539–4545 (2008).

29. 29.

Aliev, A. E., Lima, M. D., Fang, S. & Baughman, R. H. Underwater sound generation using carbon nanotube projectors. *Nano Lett.* **10**, 2374–2380 (2010).

30. 30.

Ma, T. & Shvets, G. All-Si valley-Hall photonic topological insulator. *New J. Phys.* **18**, 025012 (2016).

31. 31.

Ye, L. et al. Observation of acoustic valley vortex states and valley-chirality locked beam splitting. *Phys. Rev. B* **95**, 174106 (2017).

32. 32.

Lu, J. et al. Observation of topological valley transport of sound in sonic crystals. *Nat. Phys.* **13**, 369–374 (2017).

33. 33.

Ni, X., Gorlach, M. A., Alù, A. & Khanikaev, A. B. Topological edge states in acoustic kagome lattices. *New J. Phys.* **19**, 055002 (2017).

34. 34.

Zhang, Z. et al. Directional acoustic antennas based on valley-Hall topological insulators. *Adv. Mater.* **30**, 1803229 (2018).

35. 35.

Mei, J., Wu, Y., Chan, C. T. & Zhang, Z.-Q. First-principles study of Dirac and Dirac-like cones in phononic and photonic crystals. *Phys. Rev. B* **86**, 035141 (2012).

36. 36.

Makwana, M. P. & Craster, R. V. Geometrically navigating topological plate modes around gentle and sharp bends. *Phys. Rev. B* **98**, 184105

(2018).

37. 37.

Ochiai, T. Photonic realization of the (2+1)-dimensional parity anomaly. *Phys. Rev. B* **86**, 075152 (2012).

38. 38.

Vesterinen, V., Niskanen, A. O., Hassel, J. & Helisto, P. Fundamental efficiency of nanothermophones: modeling and experiments. *Nano Lett.* **10**, 5020–5024 (2010).

39. 39.

Anderson, P. W. Absence of diffusion in certain random lattices. *Phys. Rev.* **109**, 1492–1505 (1958).

40. 40.

Stützer, S. et al. Photonic topological Anderson insulators. *Nature* **560**, 461–465 (2018).

41. 41.

Liu, G.-G. et al. Topological Anderson insulator in disordered photonic crystals. *Phys. Rev. Lett.* **125**, 133603 (2020).

42. 42.

Zangeneh-Nejad, F. & Fleury, R. Disorder-induced signal filtering with topological metamaterials. *Adv. Mater.* **32**, 2001034 (2020).

Acknowledgements

This work was supported by the National Basic Research Program of China (2017YFA0303702), NSFC (12074183, 11922407, 11904035, 11834008, 11874215 and 12104226) and the Fundamental Research Funds

for the Central Universities (020414380181). Z.Z. acknowledges the support from the China National Postdoctoral Program for Innovative Talents (BX20200165) and the China Postdoctoral Science Foundation (2020M681541). L.Z. acknowledges support from the CONEX-Plus programme funded by Universidad Carlos III de Madrid and the European Union's Horizon 2020 research and innovation programme under Marie Skłodowska-Curie grant agreement 801538. J.C. acknowledges support from the European Research Council (ERC) through the Starting Grant 714577 PHONOMETA and from the MINECO through a Ramón y Cajal grant (grant number RYC-2015-17156).

Author information

Affiliations

1. Department of Physics, MOE Key Laboratory of Modern Acoustics, Collaborative Innovation Center of Advanced Microstructures, Nanjing University, Nanjing, China

Bolun Hu, Zhiwang Zhang, Haixiao Zhang, Wei Xiong, Zichong Yue, Xiaoyu Wang, Jianyi Xu, Ying Cheng & Xiaojun Liu

2. Department of Physics, Universidad Carlos III de Madrid, Madrid, Spain

Liyang Zheng & Johan Christensen

Contributions

Y.C. initiated the project and conceived the idea. Y.C., X.L. and J.C. guided the research. B.H., Z.Z., H.Z., L.Z. and J.C. carried out the theoretical analyses. Z.Z. and L.Z. developed the Hamiltonian model. B.H., Z.Z. and H.Z. conducted finite-element-method simulations, designed the experimental setup and conducted the measurements. W.X., Z.Y., X.W. and J.X. assisted with sample fabrication. Z.Z., Y.C. and J.C. wrote the manuscript. All the authors contributed to the discussions of the results and the manuscript preparation.

Corresponding authors

Correspondence to Zhiwang Zhang or Ying Cheng or Xiaojun Liu or Johan Christensen.

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Chunyin Qiu, Farzad Zangeneh-Nejad 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 Experimental setup for modulating the acoustic gain.](#)

The CNT films wrapped around cylinders are connected with the electrical input and thus play the role of acoustic gain thanks to electro-thermoacoustic coupling.

[Extended Data Fig. 2 Symmetry analysis and the dispersion relations calculated from the k·p method.](#)

a, Schematic of the unit cell under C_{3v} symmetry. **b**, **c**, Band diagrams of the C_{3v} -symmetric sonic crystal with $\beta = 0$ (**b**) and with $\beta = 0.05$ (**c**).

Coloured circles and solid curves epitomize the calculated results from the finite-element method and the $\mathbf{k} \cdot \mathbf{p}$ method, respectively. **d–f**, Same as **a–c**, but for the sonic crystal preserved under C_3 symmetry with the rotation angle $\theta = -\pi/6$.

Extended Data Fig. 3 Comparison between amplification through the meta-fluid and thermal-acoustic gain.

a, b, Simulated scattering pressure fields under the inward radiation of coaxial cylindrical waves by the meta-fluid model (**a**) and the thermal-acoustic gain model (**b**). **c**, Enhancement of the scattering pressure fields calculated by the meta-fluid model (purple line) and the thermal-acoustic gain model (orange dots) with different β at $f = 9.1$ kHz. **d**, Corresponding frequency dependence at $\beta = 0.05$.

Extended Data Fig. 4 Characterizations of the CNT film.

a, Experimental setup for measuring the sound pressure and the surface vibration displacement. **b**, Acoustic pressure amplitude spectra measured near the CNT film (solid curve) and the loudspeaker (dashed curve). **c**, Vibration displacements spectra measured by laser vibrometry. The solid and dashed curves represent surface displacement on the CNT film and the traditional paper basin loudspeaker, respectively. **d**, Photograph of the measurement setup for electrical impedance analysis. Inset: enlarged view of the single sample. **e**, Experimentally measured amplitude and phase of the impedance curve. **f**, Experimentally measured directivity pattern.

Extended Data Fig. 5 Physical model of the topological WG insulator.

The proposed topological domain wall can be regarded as a triangular acoustic waveguide, and the phase along each edge is labelled as α_j , with $j = 1, 2$ and 3 .

Extended Data Fig. 6 The influence of loss on the topological WG insulator.

a–c, Spectrally resolved amplification factors through simulations considering the inherent loss with three different gain-phase textures: $\phi = 0$ (**a**), $\phi = \pi$ (**b**) and $\phi = 2\pi$ (**c**). **d–f**, Pressure-field distributions and their chiralities of the three resonances at three different gain-phase textures corresponding to the frequencies $f_- = 8,889$ Hz (**d**), $f_0 = 8,999$ Hz (**e**) and $f_+ = 9,103$ Hz (**f**).

Extended Data Fig. 7 Valley chirality-selective sound emissions of the topological WG insulator.

a, Left: illustration of the designed device for the out-coupling of the chiral WG modes. Right: enlarged view of the router. The insets in the right panel show photographs of the cylinder trimers wrapped without or with CNT films. **b, c**, Momentum space analysis of the out-coupled K valley-projected topological WG mode of CW chirality at frequency $f_- = 8,889$ Hz (**b**) and K' valley-WG mode of CCW chirality at frequency $f_+ = 9,106$ Hz (**c**). The white solid hexagon represents the first Brillouin zone and the white dashed circle shows the equi-frequency contour in air. Ambient thermal colour represents the corresponding simulated sound energy fields.

Extended Data Fig. 8 Thermogram of the device and influence of temperature.

a, Photograph of the sample. **b, c**, Corresponding thermogram of the sample without (passive; **b**) and with (active; **c**) applied electric control. In **a–c**, the left column shows the entire sample and the right column shows the enlarged view of the partial sample outlined by the dashed frame. **d**, Temperature evolutions of air near the CNT film (blue curve) and in the background (orange curve) with time during the measurements. The shaded area corresponds to the temperature range of 21–22 °C in the experiments. **e**, The frequency shifts of the peaks corresponding to f_- , f_0 and f_+ under the variation of the temperature. Lines and dots represent the theoretical and simulated results.

Extended Data Fig. 9 Other types of active topological gallery and robustness against disturbances.

a, Schematic of the WG with a snowflake-shaped domain wall. **b**, Energy distributions of the CCW WG mode with $\phi = 2\pi$. **c**, Introducing phase disturbances. The solid curves in orange, light blue and purple represent the undistorted gain signals, and the dashed curves represent the distorted gain signals. **d**, Amplification spectrum including phase inhomogeneities with $\phi = 2\pi$. **e**, **f**, Same as **c**, **d**, but amplitude disturbances are introduced instead of phase disturbances.

Extended Data Fig. 10 Robustness of the topological WG insulator against the geometric defects.

a, b, Numerical defect analysis comprising one defective unit cell at each corner or side of the structure. At a gain-phase texture of $\phi = 2\pi$, we simulate the pressure fields of the system including defective units, that is, gainless, displaced or expanded cylinders (**a**) together with their corresponding spectral amplification factors (**b**). **c**, Schematic of the sample where the red and blue highlighted units label the perturbed rods located at the sides and corners, respectively. **d**, In the experiments, we chose three sets of perturbation displacements $\Delta d = 0.04a - 0.10a$ with $\phi = 2\pi$, whose measured amplification factors include both corner (top) and side (bottom) defects.

Supplementary information

Video 1

Distributions of the phase emanating the edges. Time evolution for the simulated distributions of the phase emanating the edges at $f = 8.80$ kHz within the topological band gap. The left panel shows the result when three gain assisted rods emit sound with zero delay ($\phi = 0$), while the right one shows the case when the phase increment acquires $2\pi/3$ to assume a full gain cycle of $\phi = 2\pi$.

Video 2

Topological whispering gallery modes splitting. Time evolution for the simulated pressure field distributions of the achiral/chiral topological whispering gallery modes. The middle panel shows the result at f_0 with the phase texture $\phi = 0$, behaving as the achiral topological whispering gallery mode. The left and right panels represent the cases with the phase texture $\phi = 2\pi$ at f_- and f_+ , respectively, which behave as the CW and CCW chiral modes, respectively.

Video 3

| Chiral topological whispering gallery mode along a snowflake-shaped domain wall. Time evolution for the simulated pressure field distributions along a complicated snowflake-shaped domain wall, where the CCW chiral propagations of sound waves tightly confined along the domain wall can be clearly observed.

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Hu, B., Zhang, Z., Zhang, H. *et al.* Non-Hermitian topological whispering gallery. *Nature* **597**, 655–659 (2021). <https://doi.org/10.1038/s41586-021-03833-4>

- Received: 29 January 2021
- Accepted: 16 July 2021
- Published: 29 September 2021

- Issue Date: 30 September 2021
- DOI: <https://doi.org/10.1038/s41586-021-03833-4>

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/s41586-021-03833-4>

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

- Article
- Open Access
- [Published: 29 September 2021](#)

Extremely anisotropic van der Waals thermal conductors

- [Shi En Kim¹](#),
- [Fauzia Mujid²](#),
- [Akash Rai](#) [ORCID: orcid.org/0000-0002-0288-3447³](#),
- [Fredrik Eriksson⁴](#),
- [Joonki Suh](#) [ORCID: orcid.org/0000-0002-0221-8447^{2,5}](#),
- [Preeti Poddar²](#),
- [Ariana Ray⁶](#),
- [Chibeam Park^{1,5}](#),
- [Erik Fransson](#) [ORCID: orcid.org/0000-0001-5262-3339⁴](#),
- [Yu Zhong²](#),
- [David A. Muller](#) [ORCID: orcid.org/0000-0003-4129-0473⁶](#),
- [Paul Erhart](#) [ORCID: orcid.org/0000-0002-2516-6061⁴](#),
- [David G. Cahill](#) [ORCID: orcid.org/0000-0001-5969-3460³](#) &
- [Jiwoong Park](#) [ORCID: orcid.org/0000-0002-0511-1442^{1,2,5}](#)

[Nature](#) volume **597**, pages 660–665 (2021)

- 618 Accesses
- 188 Altmetric
- [Metrics details](#)

Subjects

- [Two-dimensional materials](#)

Abstract

The densification of integrated circuits requires thermal management strategies and high thermal conductivity materials^{1,2,3}. Recent innovations include the development of materials with thermal conduction anisotropy, which can remove hotspots along the fast-axis direction and provide thermal insulation along the slow axis^{4,5}. However, most artificially engineered thermal conductors have anisotropy ratios much smaller than those seen in naturally anisotropic materials. Here we report extremely anisotropic thermal conductors based on large-area van der Waals thin films with random interlayer rotations, which produce a room-temperature thermal anisotropy ratio close to 900 in MoS₂, one of the highest ever reported. This is enabled by the interlayer rotations that impede the through-plane thermal transport, while the long-range intralayer crystallinity maintains high in-plane thermal conductivity. We measure ultralow thermal conductivities in the through-plane direction for MoS₂ ($57 \pm 3 \text{ mW m}^{-1} \text{ K}^{-1}$) and WS₂ ($41 \pm 3 \text{ mW m}^{-1} \text{ K}^{-1}$) films, and we quantitatively explain these values using molecular dynamics simulations that reveal one-dimensional glass-like thermal transport. Conversely, the in-plane thermal conductivity in these MoS₂ films is close to the single-crystal value. Covering nanofabricated gold electrodes with our anisotropic films prevents overheating of the electrodes and blocks heat from reaching the device surface. Our work establishes interlayer rotation in crystalline layered materials as a new degree of freedom for engineering-directed heat transport in solid-state systems.

[Download PDF](#)

Main

Anisotropic thermal conductors, in which heat flows faster in one direction than in another, can be characterized by the thermal conductivity anisotropy ratio ρ ($= \kappa_f/\kappa_s$) between the thermal conductivities along the fast axis (κ_f) and the slow axis (κ_s). One common way to engineer ρ in fully dense solids is via nanostructuring⁶, such as fabricating inorganic superlattices^{7,8,9,10,11} or designing symmetry-breaking crystal architectures in a single material¹². However, such engineered materials have relatively small ρ values of less than 20 at room temperature. Conversely, some natural crystalline materials have an intrinsically large ρ (for example, graphite¹ and hexagonal boron nitride (hBN)¹³, with $\rho \approx 340$ and 90 respectively), but they are often difficult to process scalably for thin film integration. Some of these films may also lack the electrical or optical properties necessary for functional device applications.

To design materials with higher ρ that are also suitable for real-world applications, an approach needs to be developed to include three key features: (1) a candidate material with intrinsically high κ_f , usually one with efficient phonon-mediated thermal

transport; (2) a method to substantially reduce κ_s without affecting κ_f , and (3) facile, scalable production and integration of such a material with precise control of the material dimensions (for example, film thickness). Layered van der Waals (vdW) materials such as graphite and transition metal dichalcogenides (TMDs) provide an ideal material platform for designing such high- ρ materials. They generally have excellent intrinsic in-plane thermal conductivities ($\kappa_{||}$) in single-crystalline form. Previous studies have also measured record-low thermal conductivities in nanocrystalline vdW films (for example, WSe₂)^{14,15,16,17} and heterostructures¹⁸. One currently missing capability, however, is an approach for significantly decreasing the out-of-plane thermal conductivity (κ_{\perp}) while maintaining high $\kappa_{||}$.

TMD films with interlayer rotations

Here we show that such capability is provided by interlayer rotations, as illustrated in Fig. 1a. Interlayer rotation breaks the through-plane translational symmetry at the atomic scale while retaining in-plane long-range crystallinity in each monolayer, thereby providing an effective means for suppressing only κ_{\perp} . For this, we produce large-area TMD films without interlayer registry (referred to here as r-TMD), which possess long-range crystallinity in-plane and relative lattice rotations at every interlayer interface (Fig. 1b). The films are produced in large-scale using two steps: wafer-scale growth of continuous TMD monolayers (polycrystalline; domain size D) and layer-by-layer stacking in vacuum using previously reported methods^{19,20} (see [Methods](#)).

Fig. 1: Structure of r-TMD films.



a, Conceptual strategy for engineering thermal anisotropy in a single material system, using random interlayer rotation in polycrystalline vdW layered materials. **b**, Schematic of an r-MoS₂ film with random crystalline orientation. **c**, Greyscale-

inverted TEM electron diffraction patterns probed from a $500 \text{ nm} \times 500 \text{ nm}$ area of a monolayer and an $N = 10$ r-MoS₂ film. Inset: darkfield TEM image of a monolayer; the scale bar denotes 400 nm and the colours denote different domain orientations from different crystal domains. **d**, HAADF-STEM image of a cross-section of an $N = 10$ r-MoS₂ film on AlO_x coated with Al, with an interlayer spacing of 6.4 Å. **e**, Large-area MoS₂ films transferred onto 1-inch diameter fused silica substrates.

The transmission electron microscopy (TEM) diffraction (Fig. 1c, left) and darkfield (inset) images from a representative MoS₂ monolayer show that it comprises large ($D \approx 1 \mu\text{m}$), randomly oriented crystalline domains, which connect laterally to form a continuous polycrystalline film. The vacuum stacking generates r-TMD films with a precise layer number (N) and high-quality interfaces²⁰ with interlayer rotation at every stacked interface. The TEM diffraction pattern of $N = 10$ r-MoS₂ (Fig. 1c, right) shows a ring-like pattern due to the significant increase in the number of diffraction spots, emphasizing the random crystalline orientation in the through-plane direction. Clean and well-defined interfaces can be seen from the cross-sectional high-angle annular darkfield scanning TEM (HAADF-STEM) images of r-MoS₂ (Fig. 1d and Extended Data Fig. 1; see [Methods](#)). The monolayers have a uniform interlayer spacing $d \approx 6.4 \text{ \AA}$ (see [Methods](#) and Extended Data Fig. 2), which is close to the expected value (6.5 Å) for twisted MoS₂ multilayers²¹. Both the growth and stacking steps are scalable, as shown by the optical images of $N = 1$ and $N = 10$ r-MoS₂ films ($\sim 1 \text{ cm}^2$) in Fig. 1e and as demonstrated later in Fig. 4. The large-scale uniformity of these films also enables precise and reproducible measurements with minimal spatial variation (Extended Data Fig. 3a–c). In our experiments, r-MoS₂ or r-WS₂ films with different N (up to 22) are transferred onto a sapphire wafer for the measurements of κ_{\perp} or suspended over a holey TEM grid (Fig. 3a) for the measurements of κ_{\parallel} .

Ultralow out-of-plane conductivity

In Fig. 2, we illustrate κ_{\perp} of r-TMD films, which is measured using time domain thermoreflectance (TDTR; Fig. 2a, inset; see [Methods](#)). A stream of laser pulses (pump) heats up the surface of an Al pad deposited on an r-TMD film on sapphire and produces a temperature-sensitive thermoreflectance signal ($-V_{\text{in}}/V_{\text{out}}$ in Fig. 2a), which is measured with a probe pulse after a varying time delay (for cooling). Figure 2a shows three representative curves measured from r-MoS₂ with $N = 1, 2$ and 10 . The curves flatten with increasing N , suggesting that heat dissipation slows down significantly. Fitting these curves using a heat diffusion model (solid lines, Fig. 2a) enables us to obtain R_{TDTR} , the total thermal resistance between the Al transducer layer and sapphire across the r-TMD film for different N .

Fig. 2: Through-plane thermal properties of r-MoS₂.

 [figure2](#)

a, TDTR heat dissipation curves of N -layer r-MoS₂ films. Inset: TDTR sample geometry. **b**, Measured thermal resistances across r-TMD films, where the error bars are the TDTR measurement uncertainties. The thermal conductivities for r-MoS₂ and r-WS₂ are calculated from the slope using the formula $R_{\text{TDTR}} = R_0 + Nd/\kappa_{\perp}$, whereby R_0 is the total interfacial thermal resistance. **c**, Experiment and MD simulation results of $\kappa(T)$ of MoS₂ and r-MoS₂ films. The error bars to the MD simulations originate from the simulation uncertainties. The dotted lines connecting the individual data points are guides to the eye. **d**, LA (top) and TA (bottom) phonon dispersion curves of r-MoS₂ along the Γ –A direction. The dotted lines denote the acoustic curves corresponding to bulk MoS₂. **e**, Lifetime of LA and TA phonons parallel to the Γ –A direction in bulk and r-MoS₂. The dashed line is the LA mode vibration period derived from the dispersion curve in **d**.

[Source data](#).

Figure 2b shows R_{TDTR} versus N for r-MoS₂ ($N \leq 22$; solid circles) and r-WS₂ ($N \leq 10$; open circles) measured under ambient conditions. We make two observations. First, R_{TDTR} monotonically increases with N . Second, R_{TDTR} varies linearly with N for $N \geq 2$. These observations confirm that the through-plane thermal transport in r-TMD films is diffusive in nature, in contrast to the ballistic transport reported in few-layer single-crystalline MoS₂ (as thick as 240 nm)^{22,23}. A single parameter κ_{\perp} characterizes the

thermal resistance across r-MoS₂ (or r-WS₂) using the equation $R_{\text{TDTR}} = R_0 + Nd/\kappa_{\perp}$ (1) where Nd is the total film thickness, and R_0 is a constant corresponding to the total interface resistance (r-TMD/Al and r-TMD/sapphire; see Extended Data Table 1). Therefore, we apply linear fitting to the data ($N \geq 2$) in Fig. 2b (solid lines) to determine κ_{\perp} of r-MoS₂ alone, regardless of the quality and chemical nature of the top and bottom interfaces (see [Methods](#) and Extended Data Fig. 3d), which can potentially be altered by metal deposition^{24,25}. We measure $\kappa_{\perp} = 57 \pm 3 \text{ mW m}^{-1} \text{ K}^{-1}$ for r-MoS₂ and $\kappa_{\perp} = 41 \pm 3 \text{ mW m}^{-1} \text{ K}^{-1}$ for r-WS₂, which are similar to the lowest value ever observed in a fully dense solid¹⁵ and comparable to the thermal conductivity of ambient air ($\sim 26 \text{ mW m}^{-1} \text{ K}^{-1}$). These values are approximately two orders of magnitude smaller than those of single-crystalline MoS₂ ($2\text{--}5 \text{ W m}^{-1} \text{ K}^{-1}$)^{26,27} or WS₂ ($\sim 3 \text{ W m}^{-1} \text{ K}^{-1}$)²⁷, despite the r-TMD films having the same chemical composition as their bulk counterparts as well as clean interfaces (Fig. 1d). This strongly suggests that the main difference, the interlayer rotation, is the principal cause for the ultralow κ_{\perp} in these r-TMD films. Furthermore, repeating similar TDTR experiments on r-MoS₂ at different temperatures (T) produces a relatively flat $\kappa_{\perp}(T)$ curve (green stars, Fig. 2c), a behaviour different from the decreasing κ_{\perp} with T seen in bulk MoS₂ (blue squares, lower Fig. 2c).

To understand the microscopic mechanisms that give rise to the dramatic reduction in κ_{\perp} , we carry out homogeneous non-equilibrium molecular dynamics (HNEMD) simulations for the model structures of r-MoS₂ and bulk MoS₂ (see [Methods](#) and Extended Data Table 2)^{28,29,30}. Figure 2c shows κ_{\parallel} and κ_{\perp} of r-MoS₂ (solid circles) and bulk MoS₂ (empty circles) calculated from our molecular dynamics (MD) simulations at different temperatures. The calculated κ_{\perp} drops by a factor of more than 20, from $3.7 \pm 0.5 \text{ W m}^{-1} \text{ K}^{-1}$ in bulk MoS₂ to $0.16 \pm 0.04 \text{ W m}^{-1} \text{ K}^{-1}$ in r-MoS₂ at 300 K, and also does not decrease with T , suggesting a transition away from the phonon-limited thermal transport mechanism observed in bulk MoS₂.

Further analysis of the vibrational spectrum of r-MoS₂ enables us to break down the reduction in κ_{\perp} in terms of the changes in the group velocities (v_g) and lifetimes (τ), which are the two factors that determine the thermal conductivity according to Boltzmann transport theory. Figure 2d shows that the v_g of the through-plane longitudinal acoustic (LA) mode in r-MoS₂ remains similar to that of bulk MoS₂ (dashed lines), but the transverse acoustic (TA) modes in r-MoS₂ undergo extreme softening with their v_g s practically vanishing³¹. This implies a loss of resistance with respect to lateral shear, consistent with the low-frequency Raman spectra of r-MoS₂

films (see [Methods](#) and Extended Data Fig. 4) and previous calculations^{32,33}. In addition, the τ of both the LA and the TA modes (Fig. 2e) in r-MoS₂ are more than one order of magnitude smaller than in bulk MoS₂, with the LA lifetimes being close to the period of the LA mode vibration (dashed line). From these results, the median mean free path $\langle \tilde{l} \rangle = v_g \tau$ for the LA modes is estimated to be 2 nm, suggesting that the heat-carrying LA modes are strongly scattered and that a larger D is unlikely to significantly affect κ_{\perp} since $D \gg \langle \tilde{l} \rangle$. Overall, the strongly suppressed TA modes, indicating a loss of resistance to lateral shear, and the overdamping of the LA modes as the main heat carriers, lead to extremely inefficient thermal transport along the through-plane direction in r-MoS₂. Along with the nearly temperature-independent κ_{\perp} , this result suggests a glass-like conduction mechanism.

In-plane conductivity and anisotropy

In contrast to κ_{\perp} , κ_{\parallel} remains high in our simulations with only a modest reduction compared to the ideal bulk crystal (less than a factor of two at 300 K; Fig. 2c). This is indeed what we observe in our Raman thermometry experiments as discussed in Fig. 3 (see [Methods](#)). We direct a focused laser spot ($\lambda = 532$ nm) at the centre of a suspended r-MoS₂ film (Fig. 3a; hole diameter of 5 μm , at 15 torr), which increases its temperature (ΔT) locally upon absorbing laser power P_{abs} . ΔT is then measured using the temperature-sensitive Raman peak shift ($\Delta\omega$) using a sensitivity factor ($|d\omega/dT|$) independently measured for each sample. Examples of Raman spectra measured for $N = 2$ are shown in Fig. 3b.

Fig. 3: In-plane thermal properties and thermal anisotropy of r-MoS₂ films.



a, 45° SEM micrograph of an $N=4$ r-MoS₂ film suspended on a TEM grid for Raman thermometry. **b**, Raman spectra of an $N=2$ r-MoS₂ film with different absorbed laser powers. Inset: Raman thermometry sample geometry. **c**, A_{1g} Raman peak shifts versus power absorbed by r-MoS₂ films of various N . Inset: layer-dependent thermal conductance values (absorbed power divided by temperature increase) in domain size $D = 1 \mu\text{m}$ and $D = 400 \text{ nm}$ r-MoS₂ films. The error bars are the propagated uncertainties from the calculation of the conductance value for each N . **d**, Comparison of ρ (y axis), κ_s (x axis), and κ_f (diagonal dashed lines) measured for different anisotropic thermal conductors. r-MoS₂ has an ultrahigh ρ close to 900, which is larger than bulk MoS₂, PG, and disordered layered WSe₂. The error bar for ρ of r-MoS₂ comes from the propagated uncertainties of the calculated κ_{\perp} and κ_{\parallel} values.

Source data.

Figure 3c plots $\Delta\omega$ versus P_{abs} for r-MoS₂ with different N (2 to 5). The slope of the linear fit ($|d(\Delta\omega)/dP_{\text{abs}}|$), which is inversely proportional to the in-plane thermal conductance of the film, is plotted in the inset (solid dots; $D \approx 1 \mu\text{m}$). We again observe a linear relation, which indicates that κ_{\parallel} is well defined for r-MoS₂ independent of N , similar to the case of κ_{\perp} . Using a simple diffusion model with radial symmetry (see [Methods](#) and Extended Data Fig. 5c for calculation details and other input measurements), we calculate a high κ_{\parallel} value of $50 \pm 6 \text{ W m}^{-1} \text{ K}^{-1}$. This value is similar to the predictions of our MD simulations (Fig. 2c) and consistent with previous reports of Raman thermometry on single-crystalline monolayer MoS₂ ($35\text{--}84 \text{ W m}^{-1} \text{ K}^{-1}$) at room temperature^{34,35,36,37}. The κ_{\parallel} of these r-MoS₂ films is close to the intrinsic phonon-limited value despite the films being made of polycrystalline monolayers. This result is further supported by our additional measurements on continuous r-MoS₂ films with a smaller $D \approx 400 \text{ nm}$ (open dots, dashed lines, Fig. 3c inset; Extended Data Fig. 5e). The measured value of $\kappa_{\parallel} \approx 44 \pm 6 \text{ W m}^{-1} \text{ K}^{-1}$ is within the margin of error of that of the $D \approx 1 \mu\text{m}$ films. This suggests that the phonon mean free path is smaller than 400 nm, which is consistent with previous reports^{23,38,39,40,41,42}. Furthermore, the measured in-plane conductance decreases with T (Extended Data Fig. 6a). This further confirms the phonon-mediated thermal transport mechanism in-plane, in contrast to the glass-like thermal conduction along the through-plane direction.

Our experiments and calculations confirm that interlayer rotation in r-TMD films results in highly directional thermal conductivity and a direction-dependent thermal conduction mechanism. The rotation significantly reduces κ_{\perp} while maintaining high κ_{\parallel} , leading to an ultrahigh value of ρ . We estimate $\rho \approx 880 \pm 110$ at room temperature

for the r-MoS₂ films, higher than that of pyrolytic graphite (PG), which is considered to be one of the most anisotropic thermal conductors ($\rho \approx 340$)⁴³. In Fig. 3d, we compare our result with other previously reported values of ρ in phonon-based solids^{15,27,43,44} (for a full comparison, see Extended Data Fig. 6b). Compared to a bulk MoS₂ crystal ($\rho \approx 20$)²⁷ or disordered layered WSe₂ ($\rho \approx 30$)^{15,44}, r-MoS₂ has a significantly larger ρ because interlayer rotation reduces only κ_{\perp} , as denoted by the grey arrow parallel to the equi- κ_f lines. This also suggests that ρ can be made even larger by starting with the monolayers of a layered vdW material with a higher $\kappa_{||}$ value such as graphene.

Anisotropic vdW heat diffuser

In Fig. 4, we show that the extreme anisotropy of our r-MoS₂ films can lead to excellent heat dissipation in-plane from a heat source and drastic thermal insulation in the through-plane direction. Using the COMSOL software, we perform thermal finite-element simulations of a 10-nm-thick r-MoS₂ film draped over a nanoscale Au electrode (15 nm tall, 100 nm wide) on a 50 nm SiO₂/Si substrate (Fig. 4a). Our simulation results show that for a fixed power of 8 mW supplied to the Au electrode (near thermal breakdown), the temperature rise ΔT of the Au electrode covered by r-MoS₂ is 50 K lower than that of the bare electrode, thereby demonstrating our film's effectiveness at spreading heat due to its excellent $\kappa_{||}$ (Fig. 4b, c). Interestingly, the extreme thermal anisotropy of our r-MoS₂ films provides thermal insulation in the through-plane direction, with much lower MoS₂ surface ΔT values that are only one-third of the value of the bare Au electrode. While single-crystal MoS₂ displays similar properties, the insulation effect is stronger in r-MoS₂ (Extended Data Fig. 7a). This implies that heat is efficiently directed away from the hot Au electrode laterally through r-MoS₂ but not to the surface of r-MoS₂, making the surface of the entire device significantly cooler.

Fig. 4: Temperature profiles and heat spreader efficiencies of r-MoS₂ films on Au electrodes.

 **figure4**

a, Schematic of the sample configuration of r-MoS₂ draped across a current-carrying Au electrode that is 100 nm wide, 15 nm thick and 10 μm long. **b**, Thermal finite element modelling results of Au electrodes (bare, covered with 10-nm-thick r-MoS₂) at constant heating power of 8 mW supplied through Joule heating. **c**, Lateral profiles of temperature increases across the Au/SiO₂ surface (solid dots) and on the r-MoS₂ top surface (open circles). Insets: cross-sectional temperature distribution of Au electrodes with and without r-MoS₂, using the same colour scale as in **b**. **d**, *I*–*V* curve of an Au electrode, with and without $N = 16$ r-MoS₂. Inset: optical micrograph of six fabricated Au electrodes. **e**, Histogram of I_c of Au electrodes with and without an $N = 16$ r-MoS₂ heat spreader and their median values.

[Source data](#).

Our experiments corroborate these simulation results. For this, we fabricate nanoscale Au electrodes with the same geometry and substrate as in our simulation (image

shown in Fig. 4d, inset) and transfer $N = 16$ r-MoS₂ (~ 10 nm thick) using the vacuum stacking process. Both bare and coated Au electrodes show similar resistance at low currents. At higher currents, current-induced Joule heating leads to the thermally activated electromigration process, which causes the electrodes to fail⁴⁵. Figure 4d compares representative current–voltage (I – V) curves measured from a bare and coated Au electrode, which shows that the Au electrode with r-MoS₂ can carry a larger current without breaking. The histogram of critical current I_c (maximum current a Au electrode sustains for at least 20 s) measured from 20 electrodes (10 bare and 10 with r-MoS₂) reveals a ~50% increase in the median I_c values (Fig. 4e). These results demonstrate our r-MoS₂ film’s ability to efficiently dissipate Joule heat and keep the electrodes cool, as our simulation predicts. As the electromigration process is dominated by the temperature, the observed increase of I_c and maximum power before breaking is in good agreement with our simulation in Fig. 4c. Furthermore, we note that the r-MoS₂ film can be integrated with the Au electrodes using mild conditions that do not affect their electrical properties (Extended Data Fig. 7b).

Outlook

We expect interlayer rotation to be an effective and generalizable way to reduce κ_{\perp} and potentially engineer anisotropic thermal properties in a variety of layered materials. Our results call for a systematic study of the exact relation between κ_{\perp} and rotation angle, which could reveal unexpected relationships analogous to the studies of electrical transport in twisted bilayer graphene⁴⁶. Interlayer rotations can be combined with other parameters (such as pressure or interlayer spacing^{47,48}) and advanced structures (superlattices and heterostructures¹⁸) to realize highly tunable ρ , allowing for the customization of thermal transport properties with an unprecedented level of directional and spatial control.

Methods

Sample preparation

Large-area, polycrystalline transition metal dichalcogenide (TMD) (MoS₂ and WS₂) monolayers were grown on SiO₂/Si substrates in a hot-walled tube furnace via metal-organic chemical vapour deposition adapted from a previously reported protocol¹⁹. The growth conditions were optimized to produce high-quality monolayer materials with structural characteristics necessary for thermal measurements. These characteristics include large domain size ($D \approx 1 \mu\text{m}$ and $0.4 \mu\text{m}$), full monolayer coverage, and laterally stitched grain boundaries.

Briefly, Mo(CO)₆ and W(CO)₆ (diluted in N₂ to 15 torr) were used as the metal precursors for the MoS₂ and WS₂ growths, respectively. (C₂H₅)₂S was used as the chalcogen source. All precursors were kept at room temperature. N₂ and H₂ were used as carrier gases. Typical growth times were 15–20 h for MoS₂ at a growth temperature of 525 °C. Typical growth times for WS₂ were 2 h at a temperature of 650 °C.

To make the r-TMD films, a TMD monolayer was spin-coated with PMMA A8 (poly-methyl methacrylate, 495 K, 4% diluted in anisole) at 2,800 rpm for 60 s, then baked at 180 °C for 3 min. The PMMA-coated monolayer was stacked onto TMD monolayers layer by layer to a target layer number (N) and transferred to the desired substrates using a previously reported programmed vacuum stacking method²⁰.

TDTR samples

The stacked r-TMD films were transferred onto sapphire substrates (Valley Design, C-plane), which were cleaned with Nanostrip solution for 20 min at 60 °C and then rinsed with deionized water. The PMMA layer on the film was removed by immersing the entire substrate in acetone at 60 °C for 1 h. The film was annealed under a 400/100 SCCM Ar/H₂ environment at 350 °C for 4 h. After cleaning, ~90-nm-thick, 90 μm × 90 μm Al pads were deposited onto the TMD films through a holey TEM grid shadow mask using electron-beam evaporation.

Raman thermometry samples

Raman experiments were performed on a different set of films from the TDTR-measured films. First, holey SiN_x transmission electron microscopy (TEM) grids were cleaned in a N₂/H₂ plasma at 100 °C and 180 mtorr for 3 min, followed by the transfer of stacked MoS₂ films onto the TEM grids. During the transfer process, the PMMA-coated r-MoS₂ was suspended on holey thermal release tape before contacting the TEM grid. The extra PMMA-MoS₂ not on the TEM grid was cut away at 180 °C so the PMMA layer was softened. PMMA was removed from r-MoS₂ on the TEM grid via annealing the film in 400/100 SCCM Ar/H₂ at 350 °C for 4 h.

Cross-sectional STEM

The $N=10$ films were coated with Al that was electron beam evaporated onto the surface, whereas the top surface of $N=20$ films was bare. The r-MoS₂ cross-section was prepared using a Thermo Scientific Strata 400 focused ion beam. Protective layers of carbon (~200 nm) and platinum (~1 μm) were deposited on the sample. A cross-

section was milled at a 90° angle from the sample using a Ga ion beam at 30 kV. The cross-section was then polished to ~150 nm thickness with the ion beam at 5 kV.

The cross-section was imaged in a Thermo Scientific Titan Themis scanning transmission electron microscope at 120 kV with a probe convergence angle of 21.4 mrad. The $N=10$ film was imaged at a beam voltage of 120 kV, whereas the $N=20$ film was imaged at 300 kV. All images were analysed using the open-source software Cornell Spectrum Imager⁴⁹. The high-angle annular dark field (HAADF) image of the sample (see 'TMD films with interlayer rotation' in the main text; Fig. 1d) shows, from top to bottom, the Al crystal lattice along the [110] zone axis, ten layers of MoS₂ (bright bands), followed by an AlO_x layer.

TDTR

We used TDTR to measure the thermal conductivity of our r-TMD films. We used a mode-locked Ti:sapphire laser, which produced a train of pulses at a repetition rate of 74.86 MHz, with wavelength centred at 785 nm and a total power of 18 mW. The steady-state temperature rise at the surface of the samples was <4 K for all temperatures. For the low temperature TDTR measurements, an INSTECH stage was used with liquid nitrogen cooling; the other beam conditions were the same. The laser beam was split into pump and probe beams. A mechanical delay stage was used to delay the arrival of the probe with respect to the pump on the sample surface by changing their optical path difference, before they were focused onto the sample surface through an objective lens. The $1/e^2$ radius of the focused laser beams was 10.7 μm. For our measurements, we modulated the pump beam at a frequency of 9.3 MHz so that the thermoreflectance change at the sample surface could be detected by the probe beam through lock-in detection. The ratio of the in-phase and out-of-phase signals from the lock-in was fitted to a thermal diffusion model. The full details of the TDTR measurement can be found elsewhere^{50,51}.

Calculation of κ_{\perp}

The modelling required material parameters such as heat capacity (C), thickness (h), interface conductance (G) and thermal conductivity (κ) for each layer. Our TDTR samples have three chemically distinct layers with the following structure (from the top): Al/r-TMD/sapphire. In our fitting process, the heat capacities of all materials were adopted from literature⁵². The thickness of Al layer was obtained from picosecond acoustics using a longitudinal speed of sound of 6.42 nm ps⁻¹ (Extended Data Fig. 3e). The thickness of the r-TMD film was calculated from the product of N and the interlayer spacing (d). The latter was measured by performing grazing-incidence wide-angle X-ray spectroscopy (GIWAXS; see GIWAXS section below in

the Methods and Extended Data Fig. 2) on the r-TMD films, which gave $d \approx 0.64$ nm. The total thicknesses of the r-MoS₂ films were <15 nm; thus, this layer was treated as part of the Al–sapphire interface as a single thermal layer characterized by a single thermal conductance value G_1 . We used the bulk value of the volumetric heat capacity of 1.89 J K⁻¹ cm⁻³ for the r-MoS₂ layer. The thermal conductivity of the Al layer was calculated from the Wiedemann–Franz law using the electrical resistance of a transducer layer deposited on a bare sapphire substrate as a reference sample. The thermal conductivity of the sapphire substrate, 35 W m⁻¹ K⁻¹, was measured using the same reference sample. Thus, the only remaining free parameter to fit for was G_1 . To obtain κ_{TMD} from G_1 , we perform TDTR on various N -layer TMD films, then perform a linear fit on the effective thermal resistance (R_{TDTR} , equal to $1/G_1$) versus N data points; the slope of the linear fit is inversely proportional to the thermal conductivity, whereas the y intercept yields the total interfacial thermal resistances (R_0) of the top and bottom interfaces. In Extended Data Table 1, our R_0 values match the values reported in literature^{22,27,53}. We note that, although R_0 changes depending on the chemical nature of the metal–TMD interface, the slope of the $R_{\text{TDTR}}–N$ plot (which is used to extract κ_{\perp}) remains constant, despite the use of different transducer metals, as illustrated in Extended Data Fig. 3d.

For highly anisotropic materials, the anisotropy ratio of an in-plane thermal conductivity to a through-plane conductivity should be included in the thermal model. Despite the ultrahigh thermal anisotropy expected of our r-TMD films, our through-plane thermal conductivity measurements were probably not sensitive to the thermal conductivity anisotropy given the thinness of our r-TMD films. Hence, we assumed a one-dimensional thermal transport model and neglected the in-plane thermal transport in our calculations. We found that the effect of the anisotropy was significant only at a smaller modulation frequency ($f = 1.12$ MHz) and $1/e^2$ beam radius of ~ 3.2 μm , and so we deliberately chose a larger f and a $1/e^2$ beam radius to reduce the sensitivity of our TDTR signal to the in-plane thermal transport.

Raman thermometry

We followed a similar procedure from previous reports^{34,54,55} with the modification of lower pressures during measurement. All the Raman measurements were performed using a Horiba Raman spectrometer with a laser excitation wavelength of 532 nm and a long-working distance, 50× objective lens (numerical aperture (NA) = 0.5). The r-MoS₂ A_{1g} peak shift (ω) versus temperature (T) relation was calibrated using a temperature-controlled, low-vacuum-compatible Linkam stage. For all our Raman measurements, we used the A_{1g} peak since this out-of-plane vibrational mode is less

sensitive to in-plane strain⁵⁶. The $\omega-T$ calibration measurements were performed at atmospheric pressure and with low laser powers. The stage was purged with dry N₂ gas throughout the calibration step to prevent oxidative damage to the film at high temperatures. Extended Data Fig. 5d shows representative $\omega-T$ calibration curves for $N=2$ and $N=4$ r-MoS₂ films, where a linear fit was performed to obtain the temperature-dependent Raman coefficients. This process was repeated for r-MoS₂ films with different domain sizes D (400 nm and 1 μm) for $N=1-3$ (Extended Data Fig. 5e).

To measure the in-plane thermal conductivity ($\kappa_{||}$) of our films, the laser power (P) was varied and the corresponding $\Delta\omega$ values were recorded. The in-plane thermal conductance was obtained from the reciprocal of the slope of the $\Delta\omega-P$ linear fit, which is illustrated for r-MoS₂ films with $N=2-5$ and $D=1 \mu\text{m}$ in Fig. 3c and for r-MoS₂ films with $N=1-3$ and $D=400 \text{ nm}$ in Extended Data Fig. 5b. As thermal conductivity changes with temperature, laser powers were kept below 250 μW to induce a relatively small ΔT in the film and ensure that the value of $\kappa_{||}$ remained relatively constant. This was verified from the observation of a linear $\Delta\omega-P$ regime for $P < 250 \mu\text{W}$. Any higher laser powers caused the $\Delta\omega-P$ curve to deviate from the linear regime with $(\frac{\partial \Delta\omega}{\partial P})^2 < 0$. This indicates that the local film temperature increased faster at higher $P > 250 \mu\text{W}$, which signified that the thermal conductivity could no longer be assumed to be constant. Instead, the thermal conductivity decreased with increasing temperature, consistent with the T -dependent Raman measurements.

The $\Delta\omega-P$ measurements were conducted at a pressure of 15 torr to eliminate any heat loss to air. We verified that a lower pressure down to 4 mtorr gave rise to similar $\Delta\omega$ values as the measurements at 15 torr (Extended Data Fig. 5a), weighted by the beam spot size.

The other relevant input quantities for our thermal calculations were obtained as follows: the beam spot radius (r_0) was estimated using the knife-edge method, whereby a one-dimensional Raman map was taken across a gold step edge on an Au-patterned silicon chip, and the spatial distribution of the integrated peak intensities was fitted to an error function. We measured $r_0 = 0.71 \pm 0.09 \mu\text{m}$. The laser powers were measured using a Thorlabs standard silicon photodiode power sensor. The r-MoS₂ absorbance ($A = \frac{\text{Absorbed light intensity}}{\text{Incident light intensity}}$) was measured at room temperature on a white-light microscope with a 532 nm band-pass filter and a low-NA condenser aperture. We measured the light intensity transmitted through and reflected from a r-MoS₂ film suspended on a TEM grid, then compared it against a blank TEM grid. The data were collected using a 12-bit SensiCam QE CCD camera. The pixel intensities were analysed using ImageJ. The

values for A were calculated using the formula $A = 1 - T - R$. We measured $A(N)$ for $N = 1-5$, then fitted A to a power law. $A(N)$ was found to follow the relation $\backslash(A=1-(0.92)^{N})$ (Extended Data Fig. 5c), which matched previous reports²⁰. We use the value A measured at room temperature for our Raman analysis.

Calculation of κ_{\parallel}

To obtain the value of κ_{\parallel} , we used the two-dimensional thermal diffusion equation with a radial symmetry, following previous reports of Raman thermometry of two-dimensional films^{54,55}. We assumed a Gaussian laser profile $\backslash(q(r)=\frac{1}{\sqrt{\pi}}e^{-r^2/4R^2})$. We solved for κ_{\parallel} numerically using the following equations:

$$\begin{aligned} \frac{\partial}{\partial r} \left(\kappa_{\parallel} \frac{\partial T}{\partial r} \right) + \frac{q(r)}{d} - \frac{G}{d} (T - T_{\infty}) &= 0 \quad r < R \\ \frac{\partial}{\partial r} \left(\kappa_{\parallel} \frac{\partial T}{\partial r} \right) + \frac{G}{d} (T - T_{\infty}) &= 0 \quad r > R \end{aligned}$$

applying the boundary conditions

$$\begin{aligned} T_{\text{susp}}(R) &= T_{\text{supp}}(R) \\ \frac{\partial T}{\partial r}(R) &= \frac{\partial T}{\partial r}(R) \end{aligned}$$

where r is the distance from TEM hole centre, P is the laser power, t is the film thickness, R is the TEM hole radius, T is the film temperature where $r \leq R$ and $r \geq R$, T_{∞} is the ambient temperature, A is the fraction of laser power absorbed, and $G = 10 \text{ MW m}^{-2} \text{ K}^{-1}$ is the interfacial thermal conductance between r-MoS₂ and SiN_x.

We solved for the expression of $T(r)$ and obtained an expression for the average temperature measured by the Raman shift

$$T_{\text{avg}} = \frac{1}{2\pi R} \int_0^{2\pi R} T(r) r dr$$

κ_{\parallel} was obtained by substituting the experimentally measured value for T_{avg} and solving the above equation numerically for κ_{\parallel} . We calculated

κ_{\parallel} for each N , and we reported the average value in the main text.

The total measurement uncertainty reported in the main text was calculated based on the error assessment for individual parameters. We used an approximate analytical solution

$$\begin{aligned} & \$\$ \left(\frac{\partial \kappa_{\parallel}}{\partial T} \right) = \frac{\partial \kappa_{\parallel}}{\partial m} \frac{\partial m}{\partial a} \approx - \\ & \left(\frac{\partial \kappa_{\parallel}}{\partial P} \right) = \frac{\partial \kappa_{\parallel}}{\partial \ln r_0} \left(\frac{\partial \ln r_0}{\partial R} \right)^{-1} \left(\frac{\partial R}{\partial A_0} \right) \frac{2 \pi d}{A_0} \left(\frac{\partial A_0}{\partial \ln r_0} \right) \end{aligned}$$

where ω is the Raman frequency of A_{1g} peak, A_0 is the absorption of the monolayer, and d is the thickness of a monolayer. The difference between the full numerical solution and this analytical form is below 3%. We identified the following independent quantities that carry uncertainty for consideration in our overall uncertainty estimation of κ_{\parallel} .

1. 1.

$\left(\frac{\partial \kappa_{\parallel}}{\partial P} \right)$: the associated uncertainty was derived from the error in the linear fit of $\left(\frac{\partial \kappa_{\parallel}}{\partial P} \right)$ and $\left(\frac{\partial \kappa_{\parallel}}{\partial T} \right)$ for every sample measured. The total uncertainty in the average $\left(\frac{\partial \kappa_{\parallel}}{\partial P} \right)$ value was 9% for $D = 1 \mu\text{m}$ and 8% for $D = 400 \text{ nm}$.

2. 2.

$\left(\frac{\partial \kappa_{\parallel}}{\partial A_0} \right)$: the uncertainty in A_0 from the $A(N)$ fit was 4%.

3. 3.

$\left(\frac{\partial \kappa_{\parallel}}{\partial r_0} \right)$: the uncertainty originated from the uncertainty in r_0 . From 14 repeated measurements of the beam spot size using the knife-edge method, we calculated the standard deviation of r_0 to be 12%, which translated to an uncertainty in the expression $\left(\frac{\partial \kappa_{\parallel}}{\partial r_0} \right)$ to be 9%.

Total uncertainty in κ_{\parallel} : 13% (for both $D = 1 \mu\text{m}$ and $D = 400 \text{ nm}$).

Variable pressure Raman thermometry measurements

Previous Raman thermometry measurements on graphene films⁵⁷ and carbon nanotubes⁵⁸ had shown an appreciable difference between measurements performed in air and at lower pressures, as well as in different gaseous environments. We extended the same precaution and repeated our Raman thermometry measurements at low pressures to reduce heat dissipation to air, an extra heat loss channel that would lead to an overestimation of the thermal conductivity of the r-MoS₂ films.

Our $\Delta\omega-P$ measurements in Fig. 3 were conducted at a pressure of 15 torr. In Extended Data Fig. 5a, we compared the $\Delta\omega$ values of $N=2$ r-MoS₂ at three different P (1 atm, 15 torr and 4 mtorr), after correcting for the different laser spot sizes.

Temperature-dependent Raman thermometry for κ_{\parallel}

We performed Raman thermometry while varying the ambient temperature T_a using a Linkam stage. No oxidation or sample damage was detected for any of the temperatures used. We performed the same $\Delta\omega-P$ measurements and calculated the κ_{\parallel} value of for each T_a . We plot a κ_{\parallel} versus T curve, where the x axis is $T = T_a$ (Extended Data Fig. 6a).

We note that the measured values of κ_{\parallel} here were lower than the room temperature values reported in the main text. We ascribe this to the sub-optimal growth conditions for the constituent monolayers used for this sample.

r-MoS₂ heat spreader experiments (electromigration of Au nanoelectrodes)

All Au electrodes were fabricated on Si substrates with 50 nm dry SiO₂ in three nanopatterning and deposition layers: (A) the nanoelectrodes (10 μm long, 100 nm wide, 15 nm thick); (B) the contact pads that would interface with the external electronics (200 μm long, 300 μm wide, 100 nm thick); and (C) the leads connecting the nanoelectrodes and the contact pads ($\sim 1,000 \mu\text{m}$ long, 50 μm wide, 15 nm thick).

We first defined the leads (B) and then the contact pads (C) using standard photolithography, and electron-beam evaporation of Ti (1 nm)/Au and lift-off. The final step was defining the nanoelectrodes (A) using electron-beam lithography, deposition of 15 nm Au, and lift-off.

Electron-beam lithography

We used a bilayer of resists: copolymer P(MMA-MAA 11%) in ethyl lactate and 950 K PMMA A4. The writing was executed with a Raith EPG 5000 Plus E-beam writer with the beam conditions of 25 nA current, dose of $1,200 \mu\text{C cm}^{-2}$, 300 μm aperture size, 100 kV accelerating voltage.

Film transfer

After the nanoelectrodes, leads and pads were fabricated, the device was cleaned with an O_2 plasma for 30 s to remove any resist residue and to promote adhesion of the r-MoS₂ film to the Au electrodes and the SiO₂ surface. A PMMA coated $N=16$ r-MoS₂ film was transferred onto the electrodes using the same process as the stacking method as outlined above. The PMMA on the r-MoS₂ film was removed by immersing the entire chip in toluene at 60 °C for 1 h.

Electrical measurements

All measurements were performed in ambient conditions with a home-built probe station in a two-probe geometry. To measure I_c in Fig. 4d, we swept the voltage bias in only one direction at a rate such that the rate in current increase is 0.05 mA per 20 s.

For comparison, we deposited SiN_x onto Au electrodes (10 μm long, 10 nm thick, 100 nm wide) using plasma-enhanced chemical vapour deposition with the following conditions: 10 s deposition at 90 °C and 10 torr and 1,000 W plasma power, with 25 SCCM/35 SCCM SiH₄ and N₂ as the precursors. The film thickness was measured via ellipsometry to be 16 nm.

Computational methodology

Structural models

Structural models were created according to an algorithm previously described in literature⁵⁹, which was implemented in Python using the atomic simulation environment package⁶⁰. The structure models were subsequently relaxed using an analytic bond-order potential⁶¹ and implemented in the LAMMPS package⁶².

The main r-MoS₂ model used in the simulations described here comprised 10 randomly stacked layers with a total of 10,152 atoms. The 10 layers came in pairs; each pair was related by a 60° rotation. The four primitive angles present in the stack

are 16.1, 25.28, 34.72 and 43.9°. Due to strain, each layer contained a different number of atoms in accordance with strains of around 10%.

The bulk structure used in the MD simulations comprised 40 layers (20 conventional unit-cells) with a total of 26,880 atoms and cell vectors of 44.44, 43.98 and 243.57 Å.

HNEMD simulations

The interatomic potential in our simulations produces the expected slight increase in the interlayer spacing in r-MoS₂ and yields thermal conductivities and phonon dispersions of bulk MoS₂ that agree with previous experimental observations and Boltzmann transport calculations based on density functional theory³⁸, confirming our MD model's suitability for this study. The structures described above were driven by an optimized driving force (Extended Data Fig. 8a) and subsequently relaxed, after which the thermal conductivity was computed using HNEMD simulations⁶³ and implemented in the graphics processing units molecular dynamics (GPUMD) package²⁹. We also included the effects of thermal expansion in the simulations (Extended Data Fig. 8b). The calculated κ_{\perp} values of r-MoS₂ are higher than the experimental values. We attribute the discrepancy to our neglecting any quantum effects and all boundary scattering in our simulations. Including such effects could further improve results from simulations. Statistics and averages were gathered from ten independent simulation runs for each system and temperature. The other parameters used in these simulations are compiled in Extended Data Table 2.

Phonon dispersion and lifetimes

We first generated the bulk MoS₂ phonon dispersion in the harmonic limit as a reference to phonon dispersion calculations using MD simulations. We computed the harmonic (0 K) phonon dispersion using the PHONOPY package⁶⁴. Forces were computed for $6 \times 6 \times 2$ supercells using the LAMMPS code. Lifetimes were calculated using the lowest applicable order of perturbation theory using the PHONO3PY package⁶⁵, which also provided us with the thermal conductivity as obtained from a direct solution of the Boltzmann transport equation⁶⁶. In these calculations, the Brillouin zone was sampled using a $10 \times 10 \times 10$ Γ -centred q-point mesh, which was chosen for consistency with the supercell used in the HNEMD simulations.

Next, we compared the dispersion of bulk MoS₂ to that calculated using MD simulations to verify the accuracy of our MD simulations for calculating the phonon dispersion of r-MoS₂. For both bulk and r-MoS₂, we extracted the phonon dispersions and lifetimes at 300 K by analysing the longitudinal and transverse current correlation

functions generated by MD simulations in the microcanonical (NVE) ensemble using dynasor⁶⁷. The MD simulations details were otherwise identical to the HNEMD simulations. The obtained correlation functions were Fourier transformed and fitted to peak shape functions corresponding to (over)damped harmonic oscillators using the full expressions given in the dynasor paper⁶⁷ to obtain phonon frequencies and lifetimes.

Finite-element analysis

We used the COMSOL software to simulate the steady-state temperature distribution in a Au electrode on a SiO₂/Si substrate. Our geometry contains an Au electrode that is 100 nm wide, 15 nm thick, and 10 μm long on a 50-nm-thick SiO₂ layer on a Si substrate. We layer a 10-nm-thick MoS₂ film onto the Au electrode and the SiO₂ layer. For the thermal anisotropy consideration, we define the thermal conductivity slow axis direction to always be perpendicular to the film's bottom surface in contact with the substrate or the Au electrode, including the Au electrode side walls.

We supply the Au electrode with 8 mW uniformly over the entire volume as the heat source, matching the power conditions at which the Au electrode fails in our experiments. As the boundary condition, we set the bottom surface on the Si substrate to be at 293.15 K. We also account for all the interfacial thermal resistances between heterogeneous surfaces in our calculations, which include r-MoS₂/Au, r-MoS₂/SiO₂, Au/SiO₂, SiO₂/Si (refs. [68,69,70,71](#)). All effects of radiation are neglected as they do not affect the temperature values in our simulations.

Low-frequency Raman measurements

The low-frequency Raman spectra of $N=2, 3$ and 4 r-MoS₂ films, along with the spectrum for MoS₂, are shown in Extended Data Fig. [4a](#). From the layer-dependence of the peak positions, we assigned these to be the breathing modes of MoS₂ (ref. [72](#)). We did not observe any peaks corresponding to shear modes in r-MoS₂. Our findings agree with theoretical studies of low-frequency Raman modes of twisted MoS₂ bilayers, which showed that the shear mode peaks redshift to below the detection capabilities (2 cm^{-1})⁷³. The positions of the breathing mode peaks of the r-MoS₂ films were close to those in exfoliated few layer MoS₂ (Extended Data Fig. [4b](#)) as reported in literature^{72,74}. This observation agreed with our MD simulations that suggested that the transverse vibrational mode was suppressed by interlayer rotation, while the longitudinal vibrational mode was retained.

GIWAXS

The GIWAXS measurement was performed using SAXSLAB (XENOCS)'s GANESHA (lab-source Cu K α , photon flux \sim 108 photons s $^{-1}$) to characterize the interlayer spacing of r-MoS₂ films. An $N=10$ r-MoS₂ film was prepared on a SiO₂/Si substrate. The incidence angle of the X-ray beam was 0.2° and the integration time was \sim 60 s. Radially integrating the two-dimensional diffraction images along the out-of-plane direction produced the diffraction spectrum along the *c* axis shown in Extended Data Fig. 2.

The peak position of 14° corresponded to an interlayer spacing of 6.4 Å in the [0 0 1] direction, which matched previous reports of r-TMD films²⁰.

Data availability

The data that support the findings of this study are available from the corresponding authors on reasonable request. [Source data](#) are provided with this paper.

Code availability

All code used in this work is available from the corresponding authors on reasonable request.

References

1. 1. Minnich, A. J. Exploring the extremes of heat conduction in anisotropic materials. *Nanoscale Microscale Thermophys. Eng.* **20**, 1–21 (2016).
2. 2. Cui, Y., Li, M. & Hu, Y. Emerging interface materials for electronics thermal management: experiments, modeling, and new opportunities. *J. Mater. Chem. C* **8**, 10568–10586 (2020).
3. 3. Moore, A. L. & Shi, L. Emerging challenges and materials for thermal management of electronics. *Mater. Today* **17**, 163–174 (2014).

4. 4.

Norley, J., Tzeng, J. J.-W., Getz, G., Klug, J. & Fedor, B. The development of a natural graphite heat-spreader. In *Seventeenth Annual IEEE Semiconductor Thermal Measurement and Management Symposium* 107–110 (IEEE, 2001).

5. 5.

Chung, D. D. L. & Takizawa, Y. Performance of isotropic and anisotropic heat spreaders. *J. Electron. Mater.* **41**, 2580–2587 (2012).

6. 6.

Termentzidis, K. Thermal conductivity anisotropy in nanostructures and nanostructured materials. *J. Phys. Appl. Phys.* **51**, 094003 (2018).

7. 7.

Lee, S.-M., Cahill, D. G. & Venkatasubramanian, R. Thermal conductivity of Si-Ge superlattices. *Appl. Phys. Lett.* **70**, 2957–2959 (1997).

8. 8.

Luckyanova, M. N. et al. Anisotropy of the thermal conductivity in GaAs/AlAs superlattices. *Nano Lett.* **13**, 3973–3977 (2013).

9. 9.

Borca-Tasciuc, T. et al. Thermal conductivity of symmetrically strained Si/Ge superlattices. *Superlattices Microstruct.* **28**, 199–206 (2000).

10. 10.

Lee, W.-Y. et al. Anisotropic temperature-dependent thermal conductivity by an Al₂O₃ interlayer in Al₂O₃/ZnO superlattice films. *Nanotechnology* **28**, 105401 (2017).

11. 11.

Rawat, V., Koh, Y. K., Cahill, D. G. & Sands, T. D. Thermal conductivity of (Zr,W)N/ScN metal/semiconductor multilayers and superlattices. *J. Appl. Phys.* **105**, 024909 (2009).

12. 12.

Sun, B. et al. Dislocation-induced thermal transport anisotropy in single-crystal group-III nitride films. *Nat. Mater.* **18**, 136–140 (2019).

13. 13.

Jiang, P., Qian, X., Yang, R. & Lindsay, L. Anisotropic thermal transport in bulk hexagonal boron nitride. *Phys. Rev. Mater.* **2**, 64005 (2018).

14. 14.

Chiritescu, C. et al. Low thermal conductivity in nanoscale layered materials synthesized by the method of modulated elemental reactants. *J. Appl. Phys.* **104**, 033533 (2008).

15. 15.

Chiritescu, C. et al. Ultralow thermal conductivity in disordered, layered WSe₂ crystals. *Science* **315**, 351–353 (2007).

16. 16.

Hadland, E. C. et al. Ultralow thermal conductivity of turbostratically disordered MoSe₂ ultra-thin films and implications for heterostructures. *Nanotechnology* **30**, 285401 (2019).

17. 17.

Hadland, E. et al. Synthesis, characterization, and ultralow thermal conductivity of a lattice-mismatched SnSe₂ (MoSe₂)_{1.32} heterostructure. *Chem. Mater.* **31**, 5699–5705 (2019).

18. 18.

Vaziri, S. et al. Ultrahigh thermal isolation across heterogeneously layered two-dimensional materials. *Sci. Adv.* **5**, eaax1325 (2019).

19. 19.

Kang, K. et al. High-mobility three-atom-thick semiconducting films with wafer-scale homogeneity. *Nature* **520**, 656–660 (2015).

20. 20.

Kang, K. et al. Layer-by-layer assembly of two-dimensional materials into wafer-scale heterostructures. *Nature* **550**, 229–233 (2017).

21. 21.

Liu, K. et al. Evolution of interlayer coupling in twisted molybdenum disulfide bilayers. *Nat. Commun.* **5**, 4966 (2014).

22. 22.

Sood, A. et al. Quasi-ballistic thermal transport across MoS₂ thin films. *Nano Lett.* **19**, 2434–2442 (2019).

23. 23.

Li, Z. et al. Size dependence and ballistic limits of thermal transport in anisotropic layered two-dimensional materials. Preprint at *arXiv* <https://arxiv.org/abs/1711.02772> (2017).

24. 24.

Schauble, K. et al. Uncovering the effects of metal contacts on monolayer MoS₂. *ACS Nano* **14**, 14798–14808 (2020).

25. 25.

Liu, Y. et al. Approaching the Schottky–Mott limit in van der Waals metal–semiconductor junctions. *Nature* **557**, 696–700 (2018).

26. 26.

Liu, J., Choi, G.-M. & Cahill, D. G. Measurement of the anisotropic thermal conductivity of molybdenum disulfide by the time-resolved magneto-optic Kerr effect. *J. Appl. Phys.* **116**, 233107 (2014).

27. 27.

Jiang, P., Qian, X., Gu, X. & Yang, R. Probing anisotropic thermal conductivity of transition metal dichalcogenides MX₂ (M = Mo, W and X = S, Se) using time-domain thermoreflectance. *Adv. Mater.* **29**, 1701068 (2017).

28. 28.

Evans, D. J. Homogeneous NEMD algorithm for thermal conductivity—application of non-canonical linear response theory. *Phys. Lett. A* **91**, 457–460 (1982).

29. 29.

Fan, Z., Chen, W., Vierimaa, V. & Harju, A. Efficient molecular dynamics simulations with many-body potentials on graphics processing units. *Comput. Phys. Commun.* **218**, 10–16 (2017).

30. 30.

Xu, K. et al. Thermal transport in MoS₂ from molecular dynamics using different empirical potentials. *Phys. Rev. B* **99**, 054303 (2019).

31. 31.

Li, D., Schleife, A., Cahill, D. G., Mitchson, G. & Johnson, D. C. Ultralow shear modulus of incommensurate [SnSe]_n[MoSe₂]_n layers synthesized by the method of modulated elemental reactants. *Phys. Rev. Mater.* **3**, 043607 (2019).

32. 32.

Huang, S. et al. Probing the interlayer coupling of twisted bilayer MoS₂ using photoluminescence spectroscopy. *Nano Lett.* **14**, 5500–5508 (2014).

33. 33.

Erhart, P., Hyldgaard, P. & Lindroth, D. O. Microscopic origin of thermal conductivity reduction in disordered van der Waals solids. *Chem. Mater.* **27**, 5511–5518 (2015).

34. 34.

Yan, R. et al. Thermal conductivity of monolayer molybdenum disulfide obtained from temperature-dependent Raman spectroscopy. *ACS Nano* **8**, 986–993 (2014).

35. 35.

Sahoo, S., Gaur, A. P. S., Ahmadi, M., Guinel, M. J.-F. & Katiyar, R. S. Temperature-dependent Raman studies and thermal conductivity of few-layer MoS₂. *J. Phys. Chem. C* **117**, 9042–9047 (2013).

36. 36.

Zhang, X. et al. Measurement of lateral and interfacial thermal conductivity of single- and bilayer MoS₂ and MoSe₂ using refined optothermal Raman technique. *ACS Appl. Mater. Interfaces* **7**, 25923–25929 (2015).

37. 37.

Taube, A., Judek, J., Łapińska, A. & Zdrojek, M. Temperature-dependent thermal properties of supported MoS₂ monolayers. *ACS Appl. Mater. Interfaces* **7**, 5061–5065 (2015).

38. 38.

Lindroth, D. O. & Erhart, P. Thermal transport in van der Waals solids from first-principles calculations. *Phys. Rev. B* **94**, 115205 (2016).

39. 39.

Jo, I., Pettes, M. T., Ou, E., Wu, W. & Shi, L. Basal-plane thermal conductivity of few-layer molybdenum disulfide. *Appl. Phys. Lett.* **104**, 201902 (2014).

40. 40.

Wei, X. et al. Phonon thermal conductivity of monolayer MoS₂: a comparison with single layer graphene. *Appl. Phys. Lett.* **105**, 103902 (2014).

41. 41.

Cai, Y., Lan, J., Zhang, G. & Zhang, Y.-W. Lattice vibrational modes and phonon thermal conductivity of monolayer MoS₂. *Phys. Rev. B* **89**, 035438 (2014).

42. 42.

Liu, X., Zhang, G., Pei, Q.-X. & Zhang, Y.-W. Phonon thermal conductivity of monolayer MoS₂ sheet and nanoribbons. *Appl. Phys. Lett.* **103**, 133113 (2013).

43. 43.

Ho, C. Y., Powell, R. W. & Liley, P. E. Thermal conductivity of the elements. *J. Phys. Chem. Ref. Data* **1**, 279–421 (1972).

44. 44.

Mavrokefalos, A., Nguyen, N. T., Pettes, M. T., Johnson, D. C. & Shi, L. In-plane thermal conductivity of disordered layered WSe₂ and (W)_x(WSe₂)_y superlattice films. *Appl. Phys. Lett.* **91**, 171912 (2007).

45. 45.

Ho, P. & Kwok, T. Electromigration in metals. *Rep. Prog. Phys.* **52**, 301–348 (1989).

46. 46.

Cao, Y. et al. Unconventional superconductivity in magic-angle graphene superlattices. *Nature* **556**, 43–50 (2018).

47. 47.

Sood, A. et al. An electrochemical thermal transistor. *Nat. Commun.* **9**, 4510 (2018).

48. 48.

Chen, S., Sood, A., Pop, E., Goodson, K. E. & Donadio, D. Strongly tunable anisotropic thermal transport in MoS₂ by strain and lithium intercalation: first-principles calculations. *2D Mater.* **6**, 025033 (2019).

49. 49.

Cueva, P., Hovden, R., Mundy, J. A., Xin, H. L. & Muller, D. A. Data processing for atomic resolution electron energy loss spectroscopy. *Microsc. Microanal.* **18**, 667–675 (2012).

50. 50.

Cahill, D. G. Analysis of heat flow in layered structures for time-domain thermoreflectance. *Rev. Sci. Instrum.* **75**, 5119–5122 (2004).

51. 51.

Liu, J. et al. Simultaneous measurement of thermal conductivity and heat capacity of bulk and thin film materials using frequency-dependent transient thermoreflectance method. *Rev. Sci. Instrum.* **84**, 034902 (2013).

52. 52.

Ditmars, D. A., Plint, C. A. & Shukla, R. C. Aluminum. I. Measurement of the relative enthalpy from 273 to 929 K and derivation of thermodynamic functions for Al(s) from 0 K to its melting point. *Int. J. Thermophys.* **6**, 499–515 (1985).

53. 53.

Yue, X. F. et al. Measurement of interfacial thermal conductance of few-layer MoS₂ supported on different substrates using Raman spectroscopy. *J. Appl. Phys.* **127**, 104301 (2020).

54. 54.

Lee, J.-U., Yoon, D., Kim, H., Lee, S. W. & Cheong, H. Thermal conductivity of suspended pristine graphene measured by Raman spectroscopy. *Phys. Rev. B* **83**, 081419 (2011).

55. 55.

Cai, W. et al. Thermal transport in suspended and supported monolayer graphene grown by chemical vapor deposition. *Nano Lett.* **10**, 1645–1651 (2010).

56. 56.

Wang, Y., Cong, C., Qiu, C. & Yu, T. Raman spectroscopy study of lattice vibration and crystallographic orientation of monolayer MoS₂ under uniaxial strain. *Small* **9**, 2857–2861 (2013).

57. 57.

Chen, S. et al. Raman measurements of thermal transport in suspended monolayer graphene of variable sizes in vacuum and gaseous environments. *ACS Nano* **5**, 321–328 (2011).

58. 58.

Hsu, I.-K., Pettes, M. T., Aykol, M., Shi, L. & Cronin, S. B. The effect of gas environment on electrical heating in suspended carbon nanotubes. *J. Appl. Phys.* **108**, 084307 (2010).

59. 59.

Hermann, K. Periodic overlayers and moiré patterns: theoretical studies of geometric properties. *J. Phys. Condens. Matter* **24**, 314210 (2012).

60. 60.
Hjorth Larsen, A. et al. The atomic simulation environment—a Python library for working with atoms. *J. Phys. Condens. Matter* **29**, 273002 (2017).
61. 61.
Liang, T., Phillpot, S. R. & Sinnott, S. B. Parametrization of a reactive many-body potential for Mo–S systems. *Phys. Rev. B* **79**, 245110 (2009).
62. 62.
Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **117**, 1–19 (1995).
63. 63.
Evans, D. J. Homogeneous NEMD algorithm for thermal conductivity—application of non-canonical linear response theory. *Phys. Lett. A* **91**, 457–460 (1982).
64. 64.
Togo, A. & Tanaka, I. First principles phonon calculations in materials science. *Scr. Mater.* **108**, 1–5 (2015).
65. 65.
Togo, A., Chaput, L. & Tanaka, I. Distributions of phonon lifetimes in Brillouin zones. *Phys. Rev. B* **91**, 094306 (2015).
66. 66.
Chaput, L. Direct solution to the linearized phonon Boltzmann equation. *Phys. Rev. Lett.* **110**, 265506 (2013).
67. 67.
Fransson, E., Slabanja, M., Erhart, P. & Wahnström, G. dynasor—a tool for extracting dynamical structure factors and current correlation functions from molecular dynamics simulations. *Adv. Theory Simul.* **4**, 2000240 (2021).
68. 68.

Yalon, E. et al. Energy dissipation in monolayer MoS₂ electronics. *Nano Lett.* **17**, 3429–3433 (2017).

69. 69.

Freedy, K. M., Olson, D. H., Hopkins, P. E. & McDonnell, S. J. Titanium contacts to MoS₂ with interfacial oxide: interface chemistry and thermal transport. *Phys. Rev. Mater.* **3**, 104001 (2019).

70. 70.

Lombard, J., Detcheverry, F. & Merabia, S. Influence of the electron–phonon interfacial conductance on the thermal transport at metal/dielectric interfaces. *J. Phys. Condens. Matter* **27**, 015007 (2015).

71. 71.

Kimling, J., Philippi-Kobs, A., Jacobsohn, J., Oepen, H. P. & Cahill, D. G. Thermal conductance of interfaces with amorphous SiO₂ measured by time-resolved magneto-optic Kerr-effect thermometry. *Phys. Rev. B* **95**, 184305 (2017).

72. 72.

Zhao, Y. et al. Interlayer breathing and shear modes in few-trilayer MoS₂ and WSe₂. *Nano Lett.* **13**, 1007–1015 (2013).

73. 73.

Huang, S. et al. Low-frequency interlayer Raman modes to probe interface of twisted bilayer MoS₂. *Nano Lett.* **16**, 1435–1444 (2016).

74. 74.

Zhang, X. et al. Raman spectroscopy of shear and layer breathing modes in multilayer MoS₂. *Phys. Rev. B* **87**, 115413 (2013).

Acknowledgements

We thank D. F. Ogletree, J. Jureller, A. J. Mannix, J.-U. Lee, K.-H. Lee and M. Lee for their helpful discussions. We also acknowledge J.-H. Kang, A. Ye and C. Liang for their help with materials preparation. Primary funding for this work comes from Air Force Office of Scientific Research MURI projects (FA9550–18–1–0480 and FA9550–

16-1-0031). Material growths done by F.M. and C.P. are partially supported by the University of Chicago MRSEC (NSF DMR-2011854) and Samsung Advanced Institute of Technology. This work makes use of the characterization facilities of the University of Chicago MRSEC (NSF DMR-2011854) and the Pritzker Nanofabrication Facility at the University of Chicago, which receives support from SHyNE Resource (NSF ECCS-1542205), a node of NSF's NNCI network. TDTR measurements are supported by Office of Naval Research MURI grant N00014-16-1-2436 and are carried out in Frederick Seitz Materials Research Laboratory at the University of Illinois at Urbana-Champaign. F.E., E.F. and P.E. are funded by the Knut and Alice Wallenberg Foundation (2014.0226), the Swedish Research Council (2015-04153 and 2018-06482), and the FLAG-ERA JTC-2017 project MECHANIC funded by the Swedish Research Council (VR 2017-06819); they acknowledge the computer time allocations by the Swedish National Infrastructure for Computing at NSC (Linköping) and C3SE (Gothenburg). A.R. and electron microscopy at the Cornell Center for Materials Research are supported by the NSF MRSEC grant DMR-1719875. The Titan microscope was acquired with the NSF MRI grant DMR-1429155. F.M. acknowledges support by the NSF Graduate Research Fellowship Program under grant no. DGE-1746045. Y.Z. acknowledges support by the Camille and Henry Dreyfus Foundation, Inc., under the Dreyfus Environmental Postdoc award EP-16-094.

Author information

Affiliations

1. Pritzker School of Molecular Engineering, University of Chicago, Chicago, IL, USA
Shi En Kim, Chibeom Park & Jiwoong Park
2. Department of Chemistry, University of Chicago, Chicago, IL, USA
Fauzia Mujid, Joonki Suh, Preeti Poddar, Yu Zhong & Jiwoong Park
3. Department of Materials Science and Engineering and Materials Research Laboratory, University of Illinois at Urbana-Champaign, Urbana, IL, USA
Akash Rai & David G. Cahill
4. Department of Physics, Chalmers University of Technology, Gothenburg, Sweden
Fredrik Eriksson, Erik Fransson & Paul Erhart

5. James Franck Institute, University of Chicago, Chicago, IL, USA

Joonki Suh, Chibeom Park & Jiwoong Park

6. School of Applied and Engineering Physics, Cornell University, Ithaca, NY, USA

Ariana Ray & David A. Muller

Contributions

S.E.K., J.S., and J.P. conceived the experiments. S.E.K. produced and performed the structural characterizations of the r-TMD films, and ran the finite-element simulations. F.M. and C.P. grew the TMD monolayers. A. Rai and D.G.C. performed the TDTR measurements. F.E., E.F. and P.E. conducted the atomistic simulations. P.P. fabricated the Au electrodes for the electromigration experiments. A. Ray and D.A.M. performed the STEM imaging. S.E.K. performed the Raman measurements and analysed the data with the help of D.G.C. and Y.Z. S.E.K., F.M. and J.P. wrote the manuscript. J.P., D.G.C. and P.E. oversaw the project. All authors discussed the manuscript and provided feedback.

Corresponding authors

Correspondence to Paul Erhart or David G. Cahill or Jiwoong Park.

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Davide Donadio, M. Muñoz Rojo and Mariusz Zdrojek 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 Cross-sectional TEM images of $N=20$ and 10 r-MoS₂ on AlO_x.

Each set of N images are taken from the same sample at different locations. The $N=10$ films are coated with Al that was electron beam evaporated onto the surface.

Extended Data Fig. 2 GIWAXS data of $N=10$ r-MoS₂.

The peak position corresponds to a 2θ value of 14° , which translates to an interlayer spacing of 6.4 \AA (scattering direction).

Extended Data Fig. 3 Additional TDTR measurements and details.

a, Microscope image of an $N=10$ r-MoS₂ film coated with a square grid of Al pads. **b**, 4×4 TDTR map of R_{TDTR} of an $N=10$ r-MoS₂ film. **c**, Histogram of R_{TDTR} array measurements. **d**, TDTR measurements of $N \leq 10$ r-TMD films coated with Au or Al. The error bars denote s.d.; number of TDTR measurements per film sample, $n = 3$ for Au samples; $n = 3–5$ for the Al samples. **e**, Picosecond acoustics of a MoS₂ monolayer on thick sapphire substrate, coated with an Al transducer layer. The y axis V_{in} is the in-phase signal of the lock-in amplifier. The red arrows indicate the acoustic waves reflected at the Al/MoS₂ interface.

Extended Data Fig. 4 Low-frequency Raman modes of r-MoS₂.

a, Raman spectra reflecting the breathing modes (BM) of r-MoS₂ (blue) and the shear mode (SM) for MoS₂. **b**, The low-frequency Raman peak positions of r-MoS₂ and exfoliated MoS₂. The filled squares indicate the BM peak positions of r-MoS₂. The open squares indicate the BM peak positions of exfoliated MoS₂, and the open circles indicate the SM peak positions of exfoliated MoS₂⁷².

Extended Data Fig. 5 Raman thermometry on r-MoS₂ films.

a, $\Delta\omega-P_{\text{abs}}$ curves of representative $N=2$ r-MoS₂ films at different pressures. The P_{abs} values along the x axis are normalized to account for the slight differences in beam spot sizes ($\Delta r = 20\%$). The results at 15 torr and 4 mtorr signify no effect of reducing the pressure to below 15 torr. $\Delta\omega$ was approximately five-fold smaller at atmospheric pressure due to the extra heat loss channel by air. **b**, $\Delta\omega-P_{\text{abs}}$ curves of r-MoS₂ films made up of $D = 400 \text{ nm}$ (grain size) monolayers. **c**, Optical absorption of suspended r-

MoS_2 films, which follows the trend $(A=1-\{(1-A)_0\}^N)$, where A_0 is the monolayer absorptance. From the fit, $A_0 = 0.08 \pm 0.003$. **d**, $\omega-T$ calibration measurements of suspended r- MoS_2 films ($D = 1 \mu\text{m}$), with the $N=2$ and $N=4$ data as the representative curves. **e**, $\omega-T$ slopes versus layer number for all films, with $D = 400 \text{ nm}$ or $1 \mu\text{m}$.

Extended Data Fig. 6 $\kappa(T)$ and ρ of r- MoS_2 .

a, $\kappa(T)$ of r- MoS_2 , with $\kappa_{||}$ measured using Raman thermometry of $N=4$ r- MoS_2 , and κ_{\perp} measured via TDTR as reported in Fig. 2c. The error bars are the propagated uncertainties from the calculation of the conductance value for each N . The error bars denote the propagated uncertainty of the calculations from the input parameters. We observed that $\kappa_{||}$ decreased with T , alluding to phonon-mediated heat transport and attesting to the long-range crystallinity of the r- MoS_2 films in-plane. This was in contrast with the $\kappa_{\perp}(T)$ behaviour (Fig. 2c), which showed a slightly increasing trend. **b**, Catalogue of experimentally measured anisotropy ratios at room temperature versus slow-axis thermal conductivity (κ_s) of thermally anisotropic materials from literature, by category.

Extended Data Fig. 7 r- MoS_2 efficacy as a heat spreader.

a, Finite element simulations of the linear temperature profiles of Au electrodes covered with MoS_2 and r- MoS_2 . **b**, SiN_x as heat spreaders for Au electrodes. Electrical properties of 10-nm-thick, 100-nm-wide and 10- μm -long Au electrodes before and after 16 nm SiN_x film deposition onto the electrodes using plasma-enhanced chemical vapour deposition. In contrast to r- MoS_2 , the direct deposition of an ultrathin inorganic film such as SiN_x with a comparable κ to $\kappa_{||}$ of r- MoS_2 negatively affects the performance of the Au electrodes.

Extended Data Fig. 8 Optimization of the MD simulations for κ calculations.

a, Optimization of the driving force of the system, where the grey zone denotes the error. **b**, Effect of thermal expansion on κ .

Extended Data Table 1 Comparison of literature thermal boundary resistances with experimentally measured values

Extended Data Table 2 Parameters used in the calculations of the thermal conductivity using GPUMD

Source data

[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

Kim, S.E., Mujid, F., Rai, A. *et al.* Extremely anisotropic van der Waals thermal conductors. *Nature* **597**, 660–665 (2021). <https://doi.org/10.1038/s41586-021-03867-8>

- Received: 05 February 2021
- Accepted: 29 July 2021
- Published: 29 September 2021
- Issue Date: 30 September 2021
- DOI: <https://doi.org/10.1038/s41586-021-03867-8>

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

Heat management with a twist in layered materials

- Mariusz Zdrojek

News & Views 29 Sept 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03867-8>

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

- Article
- [Published: 29 September 2021](#)

The role of charge recombination to triplet excitons in organic solar cells

- [Alexander J. Gillett](#) [ORCID: orcid.org/0000-0001-7572-7333](#)¹,
- [Alberto Privitera](#) [ORCID: orcid.org/0000-0002-7062-8077](#)²,
- [Rishat Dilmurat](#)³,
- [Akchheta Karki](#)⁴,
- [Deping Qian](#) [ORCID: orcid.org/0000-0001-8637-9178](#)⁵,
- [Anton Pershin](#) [ORCID: orcid.org/0000-0002-2414-6405](#)^{3,6},
- [Giacomo Londi](#) [ORCID: orcid.org/0000-0001-7777-9161](#)³,
- [William K. Myers](#) [ORCID: orcid.org/0000-0001-5935-9112](#)⁷,
- [Jaewon Lee](#)^{4,8},
- [Jun Yuan](#)^{5,9},
- [Seo-Jin Ko](#)^{4,10},
- [Moritz K. Riede](#)²,
- [Feng Gao](#) [ORCID: orcid.org/0000-0002-2582-1740](#)⁵,
- [Guillermo C. Bazan](#)⁴,
- [Akshay Rao](#) [ORCID: orcid.org/0000-0003-4261-0766](#)¹,
- [Thuc-Quyen Nguyen](#) [ORCID: orcid.org/0000-0002-8364-7517](#)⁴,
- [David Beljonne](#) [ORCID: orcid.org/0000-0001-5082-9990](#)³ &
- [Richard H. Friend](#) [ORCID: orcid.org/0000-0001-6565-6308](#)¹

Nature volume 597, pages 666–671 (2021)

- 308 Accesses

- 104 Altmetric
- [Metrics details](#)

Subjects

- [Solar cells](#)
- [Solar energy and photovoltaic technology](#)

Abstract

The use of non-fullerene acceptors (NFAs) in organic solar cells has led to power conversion efficiencies as high as 18%¹. However, organic solar cells are still less efficient than inorganic solar cells, which typically have power conversion efficiencies of more than 20%². A key reason for this difference is that organic solar cells have low open-circuit voltages relative to their optical bandgaps³, owing to non-radiative recombination⁴. For organic solar cells to compete with inorganic solar cells in terms of efficiency, non-radiative loss pathways must be identified and suppressed. Here we show that in most organic solar cells that use NFAs, the majority of charge recombination under open-circuit conditions proceeds via the formation of non-emissive NFA triplet excitons; in the benchmark PM6:Y6 blend⁵, this fraction reaches 90%, reducing the open-circuit voltage by 60 mV. We prevent recombination via this non-radiative channel by engineering substantial hybridization between the NFA triplet excitons and the spin-triplet charge-transfer excitons. Modelling suggests that the rate of back charge transfer from spin-triplet charge-transfer excitons to molecular triplet excitons may be reduced by an order of magnitude, enabling re-dissociation of the spin-triplet charge-transfer exciton. We demonstrate NFA systems in which the formation of triplet excitons is suppressed. This work thus provides a design pathway for organic solar cells with power conversion efficiencies of 20% or more.

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](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

Fig. 1: Triplet formation pathways and organic solar cell materials.

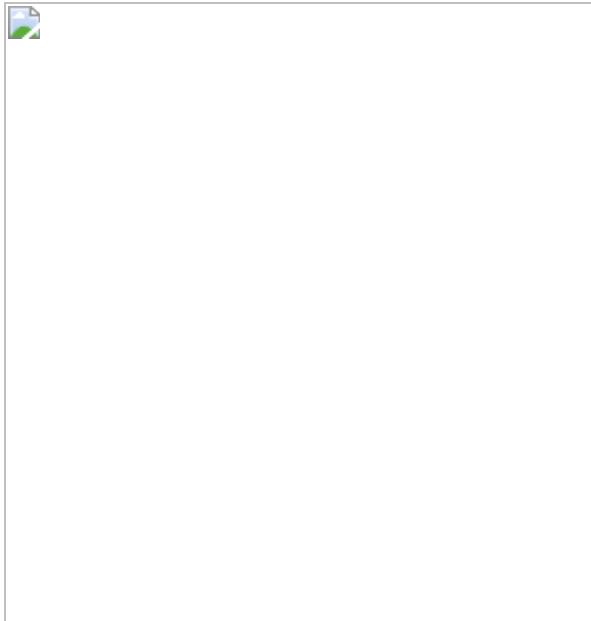


Fig. 2: Spectroscopic investigations of triplet formation in model NFA blends.

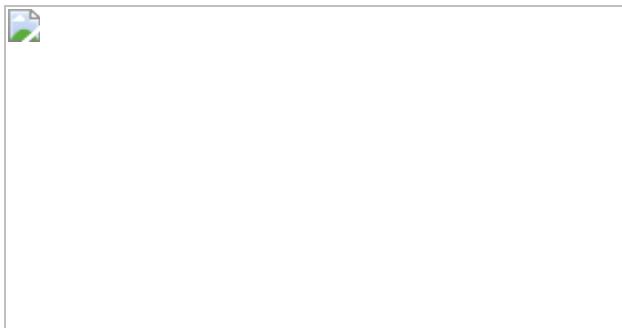
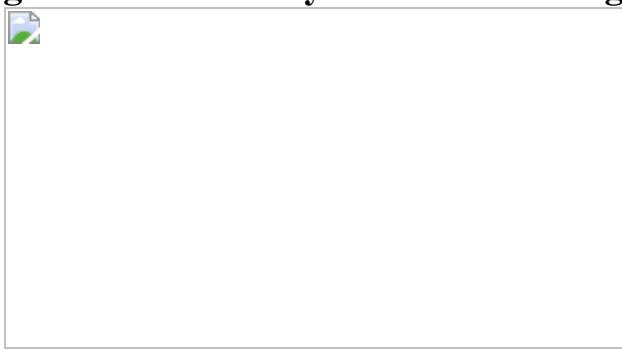


Fig. 3: The role of hybridization in organic solar cell blends.



Data availability

The data that support the plots within this paper and other findings of this study are available at the University of Cambridge Repository

(<https://doi.org/10.17863/CAM.75316>).

References

1. 1.

Liu, Q. et al. 18% efficiency organic solar cells. *Sci. Bull.* **65**, 272–275 (2020).

2. 2.

Green, M. A. et al. Solar cell efficiency tables (version 55). *Prog. Photovolt. Res. Appl.* **28**, 3–15 (2020).

3. 3.

Liu, S. et al. High-efficiency organic solar cells with low non-radiative recombination loss and low energetic disorder. *Nat. Photon.* **14**, 300–305 (2020).

4. 4.

Menke, S. M., Ran, N. A., Bazan, G. C. & Friend, R. H. Understanding energy loss in organic solar cells: toward a new efficiency regime. *Joule* **2**, 25–35 (2018).

5. 5.

Yuan, J. et al. Single-junction organic solar cell with over 15% efficiency using fused-ring acceptor with electron-deficient core. *Joule* **3**, 1140–1151 (2019).

6. 6.

Shockley, W. & Queisser, H. J. Detailed balance limit of efficiency of p–n junction solar cells. *J. Appl. Phys.* **32**, 510–519 (1961).

7. 7.

Ross, R. T. Some thermodynamics of photochemical systems. *J. Chem. Phys.* **46**, 4590–4593 (1967).

8. 8.

Rau, U. Reciprocity relation between photovoltaic quantum efficiency and electroluminescent emission of solar cells. *Phys. Rev. B* **76**, 085303 (2007).

9. 9.

Lee, J. et al. Design of nonfullerene acceptors with near-infrared light absorption capabilities. *Adv. Energy Mater.* **8**, 1801209 (2018).

10. 10.

Cui, Y. et al. Over 16% efficiency organic photovoltaic cells enabled by a chlorinated acceptor with increased open-circuit voltages. *Nat. Commun.* **10**, 2515 (2019).

11. 11.

Qian, D. et al. Design rules for minimizing voltage losses in high-efficiency organic solar cells. *Nat. Mater.* **17**, 703–709 (2018).

12. 12.

Zhou, Z. et al. Subtle molecular tailoring induces significant morphology optimization enabling over 16% efficiency organic solar cells with efficient charge generation. *Adv. Mater.* **32**, 1906324 (2020).

13. 13.

Li, S., Li, C., Shi, M. & Chen, H. New phase for organic solar cell research: emergence of Y-series electron acceptors and their perspectives. *ACS Energy Lett.* **5**, 1554–1567 (2020).

14. 14.

Yuan, J. et al. Reducing voltage losses in the A-DA'D-A acceptor-based organic solar cells. *Chem* **6**, 2147–2161 (2020).

15. 15.

Vandewal, K., Mertens, S., Benduhn, J. & Liu, Q. The cost of converting excitons into free charge carriers in organic solar cells. *J. Phys. Chem. Lett.* **11**, 129–135 (2020).

16. 16.

Geffroy, B., le Roy, P. & Prat, C. Organic light-emitting diode (OLED) technology: materials, devices and display technologies. *Polym. Int.* **55**, 572–582 (2006).

17. 17.

Classen, A. et al. The role of exciton lifetime for charge generation in organic solar cells at negligible energy-level offsets. *Nat. Energy* **5**, 711–719 (2020).

18. 18.

Eisner, F. D. et al. Hybridization of local exciton and charge-transfer states reduces nonradiative voltage losses in organic solar cells. *J. Am. Chem. Soc.* **141**, 6362–6374 (2019).

19. 19.

Chen, X.-K., Coropceanu, V. & Brédas, J.-L. Assessing the nature of the charge-transfer electronic states in organic solar cells. *Nat. Commun.* **9**, 5295 (2018).

20. 20.

Wang, J., Chepelianskii, A., Gao, F. & Greenham, N. C. Control of exciton spin statistics through spin polarization in organic optoelectronic devices. *Nat. Commun.* **3**, 1191 (2012).

21. 21.

Chen, X.-K., Wang, T. & Brédas, J.-L. Suppressing energy loss due to triplet exciton formation in organic solar cells: the role of chemical structures and molecular packing. *Adv. Energy Mater.* **7**, 1602713 (2017).

22. 22.

Rao, A. et al. The role of spin in the kinetic control of recombination in organic photovoltaics. *Nature* **500**, 435–439 (2013).

23. 23.

Kraffert, F. et al. Charge separation in PCPDTBT:PCBM blends from an EPR perspective. *J. Phys. Chem. C* **118**, 28482–28493 (2014).

24. 24.

Köhler, A. & Beljonne, D. The singlet–triplet exchange energy in conjugated polymers. *Adv. Funct. Mater.* **14**, 11–18 (2004).

25. 25.

Hodgkiss, J. M. et al. Exciton-charge annihilation in organic semiconductor films. *Adv. Funct. Mater.* **22**, 1567–1577 (2012).

26. 26.

Benduhn, J. et al. Impact of triplet excited states on the open-circuit voltage of organic solar cells. *Adv. Energy Mater.* **8**, 1800451 (2018).

27. 27.

Cohen, A. E. Nanomagnetic control of intersystem crossing. *J. Phys. Chem. A* **113**, 11084–11092 (2009).

28. 28.

Shoaee, S. et al. Decoding charge recombination through charge generation in organic solar cells. *Sol. RRL* **3**, 1900184 (2019).

29. 29.

Dimitrov, S. D. et al. Polaron pair mediated triplet generation in polymer/fullerene blends. *Nat. Commun.* **6**, 6501 (2015)

30. 30.

Salvadori, E. et al. Ultra-fast spin-mixing in a diketopyrrolopyrrole monomer/fullerene blend charge transfer state. *J. Mater. Chem. A* **5**, 24335–24343 (2017).

31. 31.

Menke, S. M. et al. Limits for recombination in a low energy loss organic heterojunction. *ACS Nano* **10**, 10736–10744 (2016).

32. 32.

Xue, L. et al. Side chain engineering on medium bandgap copolymers to suppress triplet formation for high-efficiency polymer solar cells. *Adv. Mater.* **29**, 1703344 (2017).

33. 33.

Chow, P. C. Y., Gélinas, S., Rao, A. & Friend, R. H. Quantitative bimolecular recombination in organic photovoltaics through triplet exciton formation. *J. Am. Chem. Soc.* **136**, 3424–3429 (2014).

34. 34.

Di Nuzzo, D. et al. Improved film morphology reduces charge carrier recombination into the triplet excited state in a small bandgap polymer-fullerene photovoltaic cell. *Adv. Mater.* **22**, 4321–4324 (2010).

35. 35.

Karuthedath, S. et al. Buildup of triplet-state population in operating TQ1:PC₇₁BM devices does not limit their performance. *J. Phys. Chem. Lett.* **11**, 2838–2845 (2020).

36. 36.

Wang, R. et al. Charge separation from an intra-moiety intermediate state in the high-performance PM6:Y6 organic photovoltaic blend. *J. Am. Chem. Soc.* **142**, 12751–12759 (2020).

37. 37.

Gelinas, S. et al. Ultrafast long-range charge separation in organic semiconductor photovoltaic diodes. *Science* **343**, 512–516 (2014).

38. 38.

Jakowetz, A. C. et al. Visualizing excitations at buried heterojunctions in organic semiconductor blends. *Nat. Mater.* **16**, 551–557 (2017).

39. 39.

Menke, S. M. et al. Order enables efficient electron-hole separation at an organic heterojunction with a small energy loss. *Nat. Commun.* **9**, 277 (2018).

40. 40.

Burke, T. M., Sweetnam, S., Vandewal, K. & McGehee, M. D. Beyond langevin recombination: how equilibrium between free carriers and charge transfer states determines the open-circuit voltage of organic solar cells. *Adv. Energy Mater.* **5**, 1500123 (2015).

41. 41.

Hosseini, S. M. et al. Putting order into PM6:Y6 solar cells to reduce the Langevin recombination in 400 nm thick junction. *Sol. RRL* **4**, 2000498 (2020).

42. 42.

Karki, A. et al. Understanding the high performance of over 15% efficiency in single-junction bulk heterojunction organic solar cells. *Adv. Mater.* **31**, 1903868 (2019).

43. 43.

Niklas, J. et al. Highly-efficient charge separation and polaron delocalization in polymer-fullerene bulk-heterojunctions: a comparative multi-frequency EPR and DFT study. *Phys. Chem. Chem. Phys.* **15**, 9562–9574 (2013).

44. 44.

Richert, S., Tait, C. E. & Timmel, C. R. Delocalisation of photoexcited triplet states probed by transient EPR and hyperfine spectroscopy. *J. Magn. Reson.* **280**, 103–116 (2017).

45. 45.

Thomson, S. A. J. et al. Charge separation and triplet exciton formation pathways in small molecule solar cells as studied by time-resolved EPR spectroscopy. *J. Phys. Chem. C* **121**, 22707–22719 (2017).

46. 46.

Hintze, C., Steiner, U. E. & Drescher, M. Photoexcited triplet state kinetics studied by electron paramagnetic resonance spectroscopy. *ChemPhysChem* **18**, 6–16 (2017).

47. 47.

Benduhn, J. et al. Intrinsic non-radiative voltage losses in fullerene-based organic solar cells. *Nat. Energy* **2**, 17053 (2017).

48. 48.

Kubas, A. et al. Electronic couplings for molecular charge transfer: benchmarking CDFT, FODFT and FODFTB against high-level ab initio calculations. II. *Phys. Chem. Chem. Phys.* **17**, 14342–14354 (2015).

49. 49.

Chang, W. et al. Spin-dependent charge transfer state design rules in organic photovoltaics. *Nat. Commun.* **6**, 6415 (2015).

50. 50.

Street, R. A., Song, K. W., Northrup, J. E. & Cowan, S. Photoconductivity measurements of the electronic structure of organic solar cells. *Phys. Rev. B* **83**, 165207 (2011).

51. 51.

Rasaiah, J. C., Hubbard, J. B., Rubin, R. J. & Lee, S. H. Kinetics of bimolecular recombination processes with trapping. *J. Phys. Chem.* **94**, 652–662 (1990).

52. 52.

Lee, J. et al. Bandgap narrowing in non-fullerene acceptors: single atom substitution leads to high optoelectronic response beyond 1000 nm. *Adv. Energy Mater.* **8**, 1801212 (2018).

53. 53.

de Mello, J. C., Wittmann, H. F. & Friend, R. H. An improved experimental determination of external photoluminescence quantum efficiency. *Adv. Mater.* **9**, 230–232 (1997).

54. 54.

Lee, C.-L., Yang, X. & Greenham, N. C. Determination of the triplet excited-state absorption cross section in a polyfluorene by energy

transfer from a phosphorescent metal complex. *Phys. Rev. B* **76**, 245201 (2007).

55. 55.

Biskup, T. Structure-function relationship of organic semiconductors: detailed insights from time-resolved EPR spectroscopy. *Front Chem.* **7**, 10 (2019).

56. 56.

Weber, S. Transient EPR. *eMagRes* **6**, 255–270 (2017).

57. 57.

Niklas, J. & Poluektov, O. G. Charge transfer processes in OPV materials as revealed by EPR spectroscopy. *Adv. Energy Mater.* **7**, 1602226 (2017).

58. 58.

Righetto, M. et al. Engineering interactions in QDs-PCBM blends: a surface chemistry approach. *Nanoscale* **10**, 11913–11922 (2018).

59. 59.

Franco, L. et al. Time-resolved EPR of photoinduced excited states in a semiconducting polymer/PCBM blend. *J. Phys. Chem. C* **117**, 1554–1560 (2013).

60. 60.

Buckley, C. D., Hunter, D. A., Hore, P. J. & McLauchlan, K. A. Electron spin resonance of spin-correlated radical pairs. *Chem. Phys. Lett.* **135**, 307–312 (1987).

61. 61.

Hore, P. J., Hunter, D. A., McKie, C. D. & Hoff, A. J. Electron paramagnetic resonance of spin-correlated radical pairs in photosynthetic reactions. *Chem. Phys. Lett.* **137**, 495–500 (1987).

62. 62.

Stoll, S. & Schweiger, A. EasySpin, a comprehensive software package for spectral simulation and analysis in EPR. *J. Magn. Reson.* **178**, 42–55 (2006).

Acknowledgements

A.J.G. and R.H.F. acknowledge support from the Simons Foundation (grant number 601946) and the EPSRC (EP/M01083X/1 and EP/M005143/1). This project has received funding from the ERC under the European Union’s Horizon 2020 research and innovation programme (grant agreement number 670405). A.K. and T.-Q.N. were supported by the Department of the Navy, Office of Naval Research award number N00014-21-1-2181. A.K. acknowledges funding by the Schlumberger foundation. A.Privitera, R.D., A.Pershin, G.L., M.K.R. and D.B. were supported by the European Union’s Horizon 2020 research and innovation programme under Marie Skłodowska Curie grant agreement number 722651 (SEPOMO project). Computational resources in Mons were provided by the Consortium des Équipements de Calcul Intensif (CÉCI), funded by the Fonds de la Recherche Scientifiques de Belgique (FRS-FNRS) under grant number 2.5020.11, as well as the Tier-1 supercomputer of the Fédération Wallonie-Bruxelles, infrastructure funded by the Walloon Region under grant agreement number 1117545. D.B. is a FNRS Research Director. F.G. acknowledges the Stiftelsen för Strategisk Forskning through a Future Research Leader programme (FFL18-0322). Transient electron paramagnetic resonance measurements were performed in the Centre for Advanced ESR (CAESR) in the Department of Chemistry at the University of Oxford, and this work was supported by the EPSRC (EP/L011972/1). We thank T. Biskup and A. Sperlich for their assistance with simulation and interpretation of the transient electron paramagnetic resonance data.

Author information

Affiliations

1. Cavendish Laboratory, University of Cambridge, Cambridge, UK
Alexander J. Gillett, Akshay Rao & Richard H. Friend
2. Clarendon Laboratory, University of Oxford, Oxford, UK
Alberto Privitera & Moritz K. Riede
3. Laboratory for Chemistry of Novel Materials, Université de Mons, Mons, Belgium
Rishat Dilmurat, Anton Pershin, Giacomo Londi & David Beljonne
4. Centre for Polymers and Organic Solids, Department of Chemistry and Biochemistry, University of California at Santa Barbara, Santa Barbara, CA, USA
Akchheta Karki, Jaewon Lee, Seo-Jin Ko, Guillermo C. Bazan & Thuc-Quyen Nguyen
5. Department of Physics, Chemistry and Biology (IFM), Linköping University, Linköping, Sweden
Deping Qian, Jun Yuan & Feng Gao
6. Wigner Research Centre for Physics, Budapest, Hungary
Anton Pershin
7. Centre for Advanced Electron Spin Resonance, Inorganic Chemistry Laboratory, University of Oxford, Oxford, UK
William K. Myers

8. Department of Chemical Engineering and Applied Chemistry,
Chungnam National University, Daejeon, South Korea

Jaewon Lee

9. College of Chemistry and Chemical Engineering, Central South
University, Changsha, China

Jun Yuan

10. Division of Advanced Materials, Korea Research Institute of Chemical
Technology, Daejeon, South Korea

Seo-Jin Ko

Contributions

A.J.G., T.-Q.N. and R.H.F. conceived the work. A.J.G. performed the transient absorption measurements. A.Privitera and W.K.M. conducted the transient electron paramagnetic resonance measurements. R.D., A.Pershin and G.L. carried out the quantum chemical calculations. A.K., D.Q., J.Y. and S.-J.K. fabricated and tested the organic solar cell devices. A.J.G. and J.Y. performed the photoluminescence quantum efficiency measurements. J.L. synthesized SiOTIC-4F and the IEICO derivatives. M.K.R., F.G., G.C.B., T.-Q.N, D.B. and R.H.F. supervised their group members involved in the project. A.J.G., A.R. and R.H.F. wrote the manuscript, with input from all authors.

Corresponding authors

Correspondence to Alexander J. Gillett or Thuc-Quyen Nguyen or David Beljonne or Richard H. Friend.

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Till Biskup, Yingping Zou 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 contains Supplementary Text and Data, including Supplementary Figs. 1–70, Supplementary Tables 1–3 and Supplementary References – see contents page for details.

Reporting Summary

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Gillett, A.J., Privitera, A., Dilmurat, R. *et al.* The role of charge recombination to triplet excitons in organic solar cells. *Nature* **597**, 666–671 (2021). <https://doi.org/10.1038/s41586-021-03840-5>

- Received: 16 June 2020

- Accepted: 20 July 2021
- Published: 29 September 2021
- Issue Date: 30 September 2021
- DOI: <https://doi.org/10.1038/s41586-021-03840-5>

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/s41586-021-03840-5>

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

- Article
- Open Access
- [Published: 29 September 2021](#)

Skilful precipitation nowcasting using deep generative models of radar

- [Suman Ravuri](#) ORCID: [orcid.org/0000-0002-7481-7633¹](https://orcid.org/0000-0002-7481-7633)^{nal},
- [Karel Lenc](#) ORCID: [orcid.org/0000-0001-6119-0045¹](https://orcid.org/0000-0001-6119-0045)^{nal},
- [Matthew Willson](#) ORCID: [orcid.org/0000-0002-8730-1927¹](https://orcid.org/0000-0002-8730-1927)^{nal},
- [Dmitry Kangin](#) ORCID: [orcid.org/0000-0001-9769-7585^{2,3}](https://orcid.org/0000-0001-9769-7585),
- [Remi Lam¹](#),
- [Piotr Mirowski¹](#),
- [Megan Fitzsimons²](#),
- [Maria Athanassiadou²](#),
- [Sheleem Kashem¹](#),
- [Sam Madge²](#),
- [Rachel Prudden^{2,3}](#),
- [Amol Mandhane](#) ORCID: [orcid.org/0000-0002-3412-2634¹](https://orcid.org/0000-0002-3412-2634),
- [Aidan Clark¹](#),
- [Andrew Brock¹](#),
- [Karen Simonyan¹](#),
- [Raia Hadsell¹](#),
- [Niall Robinson^{2,3}](#),
- [Ellen Clancy¹](#),
- [Alberto Arribas^{2,4}](#) &
- [Shakir Mohamed](#) ORCID: [orcid.org/0000-0002-1184-5776¹](https://orcid.org/0000-0002-1184-5776)

Nature volume 597, pages 672–677 (2021)

- 308 Altmetric
- [Metrics details](#)

Subjects

- [Computer science](#)
- [Environmental sciences](#)

Abstract

Precipitation nowcasting, the high-resolution forecasting of precipitation up to two hours ahead, supports the real-world socioeconomic needs of many sectors reliant on weather-dependent decision-making^{1,2}. State-of-the-art operational nowcasting methods typically advect precipitation fields with radar-based wind estimates, and struggle to capture important non-linear events such as convective initiations^{3,4}. Recently introduced deep learning methods use radar to directly predict future rain rates, free of physical constraints^{5,6}. While they accurately predict low-intensity rainfall, their operational utility is limited because their lack of constraints produces blurry nowcasts at longer lead times, yielding poor performance on rarer medium-to-heavy rain events. Here we present a deep generative model for the probabilistic nowcasting of precipitation from radar that addresses these challenges. Using statistical, economic and cognitive measures, we show that our method provides improved forecast quality, forecast consistency and forecast value. Our model produces realistic and spatiotemporally consistent predictions over regions up to 1,536 km × 1,280 km and with lead times from 5–90 min ahead. Using a systematic evaluation by more than 50 expert meteorologists, we show that our generative model ranked first for its accuracy and usefulness in 89% of cases against two competitive methods. When verified quantitatively, these nowcasts are skillful without resorting to blurring. We show that generative nowcasting can provide probabilistic predictions that improve forecast value and support operational utility, and at resolutions and lead times where alternative methods struggle.

[Download PDF](#)

Main

The high-resolution forecasting of rainfall and hydrometeors zero to two hours into the future, known as precipitation nowcasting, is crucial for weather-dependent decision-making. Nowcasting informs the operations of a wide variety of sectors, including emergency services, energy management, retail, flood early-warning systems, air traffic control and marine services^{1,2}. For nowcasting to be useful in these applications the forecast must provide accurate predictions across multiple spatial and temporal scales, account for uncertainty and be verified probabilistically, and perform well on heavier precipitation events that are rarer, but more critically affect human life and economy.

Ensemble numerical weather prediction (NWP) systems, which simulate coupled physical equations of the atmosphere to generate multiple realistic precipitation forecasts, are natural candidates for nowcasting as one can derive probabilistic forecasts and uncertainty estimates from the ensemble of future predictions⁷. For precipitation at zero to two hours lead time, NWPs tend to provide poor forecasts as this is less than the time needed for model spin-up and due to difficulties in non-Gaussian data assimilation^{8,9,10}. As a result, alternative methods that make predictions using composite radar observations have been used; radar data is now available (in the UK) every five minutes and at $1\text{ km} \times 1\text{ km}$ grid resolution¹¹. Established probabilistic nowcasting methods, such as STEPS and PySTEPS^{3,4}, follow the NWP approach of using ensembles to account for uncertainty, but model precipitation following the advection equation with a radar source term. In these models, motion fields are estimated by optical flow, smoothness penalties are used to approximate an advection forecast, and stochastic perturbations are added to the motion field and intensity model^{3,4,12}. These stochastic simulations allow for ensemble nowcasts from which both probabilistic and deterministic forecasts can be derived and are applicable and consistent at multiple spatial scales, from the kilometre scale to the size of a catchment area¹³.

Approaches based on deep learning have been developed that move beyond reliance on the advection equation^{5,6,14,15,16,17,18,19}. By training these models on large corpora of radar observations rather than relying on in-built physical assumptions, deep learning methods aim to better model traditionally difficult non-linear precipitation phenomena, such as convective initiation and heavy precipitation. This class of methods directly predicts precipitation rates at each grid location, and models have been developed for both deterministic and probabilistic forecasts. As a result of their direct optimization and fewer inductive biases, the forecast quality of deep learning methods—as measured by per-grid-cell metrics such as critical success index (CSI)²⁰ at low precipitation levels (less than 2 mm h^{-1})—has greatly improved.

As a number of authors have noted^{5,6}, forecasts issued by current deep learning systems express uncertainty at increasing lead times with blurrier precipitation fields, and may not include small-scale weather patterns that are important for improving forecast value. Furthermore, the focus in existing approaches on location-specific predictions, rather than probabilistic predictions of entire precipitation fields, limits their operational utility and usefulness, being unable to provide simultaneously consistent predictions across multiple spatial and temporal aggregations. The ability to make skilful probabilistic predictions is also known to provide greater economic and decision-making value than deterministic forecasts^{21,22}.

Here we demonstrate improvements in the skill of probabilistic precipitation nowcasting that improves their value. To create these more skilful predictions, we

develop an observations-driven approach for probabilistic nowcasting using deep generative models (DGMs). DGMs are statistical models that learn probability distributions of data and allow for easy generation of samples from their learned distributions. As generative models are fundamentally probabilistic, they have the ability to simulate many samples from the conditional distribution of future radar given historical radar, generating a collection of forecasts similar to ensemble methods. The ability of DGMs to both learn from observational data as well as represent uncertainty across multiple spatial and temporal scales makes them a powerful method for developing new types of operationally useful nowcasting. These models can predict smaller-scale weather phenomena that are inherently difficult to predict due to underlying stochasticity, which is a critical issue for nowcasting research. DGMs predict the location of precipitation as accurately as systems tuned to this task while preserving spatiotemporal properties useful for decision-making. Importantly, they are judged by professional meteorologists as substantially more accurate and useful than PySTEPS or other deep learning systems.

Generative models of radar

Our nowcasting algorithm is a conditional generative model that predicts N future radar fields given M past, or contextual, radar fields, using radar-based estimates of surface precipitation \mathbf{X}_T at a given time point T . Our model includes latent random vectors \mathbf{Z} and parameters $\boldsymbol{\theta}$, described by

$$\begin{aligned} & \text{\$\$}\backslash\begin{array}{c} P(\{\{\bf X\}\})_{_M+1\{\rm :}\}M+N\}| \\ & \{\{\bf X\}\}_{_1\{\rm :}\}M)=\int P(\{\{\bf X\}\}_{_M+1\{\rm :}\}M+N)\{\rm d}\} \\ & \{\{\bf Z\}\},\{\{\bf X\}\}_{_1\{\rm :}\}M},\{\boldsymbol{\theta}\})P(\{\bf Z\}\{\rm d}\} \\ & \{\{\bf X\}\}_{_1\{\rm :}\}M)\{\rm d}\}\{\bf Z\}. \end{array}\text{\$\$} \\ (1) \end{aligned}$$

The integration over latent variables ensures that the model makes predictions that are spatially dependent. Learning is framed in the algorithmic framework of a conditional generative adversarial network (GAN)^{[23,24,25](#)}, specialized for the precipitation prediction problem. Four consecutive radar observations (the previous 20 min) are used as context for a generator (Fig. [1a](#)) that allows sampling of multiple realizations of future precipitation, each realization being 18 frames (90 min).

Fig. 1: Model overview and case study of performance on a challenging precipitation event starting on = 24 June 2019 at 16:15 UK, showing convective cells over eastern Scotland.

 **figure1**

DGMR is better able to predict the spatial coverage and convection compared to other methods over a longer time period, while not over-estimating the intensities, and is significantly preferred by meteorologists (93% first choice, $n = 56$, $P < 10^{-4}$). **a**, Schematic of the model architecture showing the generator with spatial latent vectors **Z**. **b**, Geographic context for the predictions. **c**, A single prediction at $T + 30$, $T + 60$ and $T + 90$ min lead time for different models. Critical success index (CSI) at thresholds 2 mm h^{-1} and 8 mm h^{-1} and continuous ranked probability score (CRPS) for an ensemble of four samples shown in the bottom left corner. For axial attention we show the mode prediction. Images are $256 \text{ km} \times 256 \text{ km}$. Maps produced with Cartopy and SRTM elevation data⁴⁶.

Learning is driven by two loss functions and a regularization term, which guide parameter adjustment by comparing real radar observations to those generated by the model. The first loss is defined by a spatial discriminator, which is a convolutional neural network that aims to distinguish individual observed radar fields from generated fields, ensuring spatial consistency and discouraging blurry predictions. The

second loss is defined by a temporal discriminator, which is a three-dimensional (3D) convolutional neural network that aims to distinguish observed and generated radar sequences, imposes temporal consistency and penalizes jumpy predictions. These two discriminators share similar architectures to existing work in video generation²⁶. When used alone, these losses lead to accuracy on par with Eulerian persistence. To improve accuracy, we introduce a regularization term that penalizes deviations at the grid cell resolution between the real radar sequences and the model predictive mean (computed with multiple samples). This third term is important for the model to produce location-accurate predictions and improve performance. In the [Supplementary Information](#), we show an ablation study supporting the necessity of each loss term. Finally, we introduce a fully convolutional latent module for the generator, allowing for predictions over precipitation fields larger than the size used at training time, while maintaining spatiotemporal consistency. We refer to this DGM of rainfall as DGMR in the text.

The model is trained on a large corpus of precipitation events, which are 256×256 crops extracted from the radar stream, of length 110 min (22 frames). An importance-sampling scheme is used to create a dataset more representative of heavy precipitation ([Methods](#)). Throughout, all models are trained on radar observations for the UK for years 2016–2018 and evaluated on a test set from 2019. Analysis using a weekly train–test split of the data, as well as data of the USA, is reported in Extended Data Figs. 1–9 and the [Supplementary Information](#). Once trained, this model allows fast full-resolution nowcasts to be produced, with a single prediction (using an NVIDIA V100 GPU) needing just over a second to generate.

Intercomparison case study

We use a single case study to compare the nowcasting performance of the generative method DGMR to three strong baselines: PySTEPS, a widely used precipitation nowcasting system based on ensembles, considered to be state-of-the-art^{3,4,13}; UNet, a popular deep learning method for nowcasting¹⁵; and an axial attention model, a radar-only implementation of MetNet¹⁹. For a meteorologically challenging event, Figs. [1b](#), [c](#) and [4b](#) shows the ground truth and predicted precipitation fields at $T + 30$, $T + 60$ and $T + 90$ min, quantitative scores on different verification metrics, and comparisons of expert meteorologist preferences among the competing methods. Two other cases are included in Extended Data Figs. [2](#) and [3](#).

The event in Fig. [1](#) shows convective cells in eastern Scotland with intense showers over land. Maintaining such cells is difficult and a traditional method such as PySTEPS overestimates the rainfall intensity over time, which is not observed in reality and does not sufficiently cover the spatial extent of the rainfall. The UNet and axial attention models roughly predict the location of rain, but owing to aggressive

blurring, over-predict areas of rain, miss intensity and fail to capture any small-scale structure. By comparison, DGMR preserves a good spatial envelope, represents the convection and maintains heavy rainfall in the early prediction, although with less accurate rates at $T + 90$ min and at the edge of the radar than at previous time steps. When expert meteorologists judged these predictions against ground truth observations, they significantly preferred the generative nowcasts, with 93% of meteorologists choosing it as their first choice (Fig. 4b).

The figures also include two common verification scores. These predictions are judged as significantly different by experts, but the scores do not provide this insight. This study highlights a limitation of using existing popular metrics to evaluate forecasts: while standard metrics implicitly assume that models, such as NWPs and advection-based systems, preserve the physical plausibility of forecasts, deep learning systems may outperform on certain metrics by failing to satisfy other needed characteristics of useful predictions.

Forecast skill evaluation

We verify the performance of competing methods using a suite of metrics as is standard practice, as no single verification score can capture all desired properties of a forecast. We report the CSI²⁷ to measure location accuracy of the forecast at various rain rates. We report the radially averaged power spectral density (PSD)^{28,29} to compare the precipitation variability of nowcasts to that of the radar observations. We report the continuous ranked probability score (CRPS)³⁰ to determine how well the probabilistic forecast aligns with the ground truth. For CRPS, we show pooled versions, which are scores on neighbourhood aggregations that show whether a prediction is consistent across spatial scales. Details of these metrics, and results on other standard metrics, can be found in Extended Data Figs. 1–9 and the [Supplementary Information](#). We report results here using data from the UK, and results consistent with these showing generalization of the method on data from the USA in Extended Data Figs. 1–9.

Figure 2a shows that all three deep learning systems produce forecasts that are significantly more location-accurate than the PySTEPS baseline when compared using CSI. Using paired permutation tests with alternating weeks as independent units to assess statistical significance, we find that DGMR has significant skill compared to PySTEPS for all precipitation thresholds ($n = 26$, $P < 10^{-4}$) ([Methods](#)).

Fig. 2: Deterministic verification scores for the UK in 2019.

 **figure2**

a, CSI across 20 samples for precipitation thresholds at 1 mm h^{-1} (left), 4 mm h^{-1} (middle) and 8 mm h^{-1} (right). We also report results for the axial attention mode prediction. UNet generates a single deterministic prediction. **b**, Radially averaged power spectral density for full-frame 2019 predictions for all models at $T + 30 \text{ min}$ (left) and $T + 90 \text{ min}$ (middle and right). At $T + 90 \text{ min}$, UNet (middle) has an effective resolution of 32 km; both axial attention (right) sample and mode predictions have an effective resolution of 16 km.

[Source data](#).

The PSD in Fig. 2b shows that both DGMR and PySTEPS match the observations in their spectral characteristics, but the axial attention and UNet models produce forecasts with medium- and small-scale precipitation variability that decreases with increasing lead time. As they produce blurred predictions, the effective resolution of the axial attention and UNet nowcasts is far less than the $1 \text{ km} \times 1 \text{ km}$ resolution of the data. At $T + 90 \text{ min}$, the effective resolution for UNet is 32 km and for axial attention is 16 km, reducing the value of these nowcasts for meteorologists.

For probabilistic verification, Fig. 3a,b shows the CRPS of the average and maximum precipitation rate aggregated over regions of increasing size³¹. When measured at the grid-resolution level, DGMR, PySTEPS and axial attention perform similarly; we also show an axial attention model with improved performance obtained by rescaling its

output probabilities³² (denoted ‘axial attention temp. opt.’). As the spatial aggregation is increased, DGMR and PySTEPS provide consistently strong performance, with DGMR performing better on maximum precipitation. The axial attention model is significantly poorer for larger aggregations and underperforms all other methods at scale four and above. Using alternating weeks as independent units, paired permutation tests show that the performance differences between DGMR and the axial attention temp. opt. are significant ($n = 26$, $P < 10^{-3}$).

Fig. 3: Probabilistic verification scores for the UK in 2019.

 figure3



Graphs show CRPS scores at the grid resolution (left), 4-km aggregations (middle) and 16-km aggregations (right). **a**, Pooled CRPS using the average rain rate. **b**, Pooled CRPS using the maximum rain rate.

[Source data.](#)

NWP and PySTEPS methods include post-processing that is used by default in their evaluation to improve reliability. We show a simple post-processing method for DGMR in Figs. 2 and 3 (denoted ‘recal’) ([Methods](#)), which further improves its skill scores over the uncalibrated approach. Post-processing improves the reliability diagrams and rank histogram to be as or more skilful than the baseline methods (Extended Data Fig. 4). We also show evaluation on other metrics, performance on a data split over weeks rather than years, and evaluation recapitulating the inability of NWPs to make predictions at nowcasting timescales (Extended Data Figs. 4–6). We show results on a US dataset in Extended Data Figs. 7–9.

Together, these results show that the generative approach verifies competitively compared to alternatives: it outperforms (on CSI) the incumbent STEPS nowcasting approach, provides probabilistic forecasts that are more location accurate, and preserves the statistical properties of precipitation across spatial and temporal scales without blurring whereas other deep learning methods do so at the expense of them.

Forecast value evaluation

We use both economic and cognitive analyses to show that the improved skill of DGMR results in improved decision-making value.

We report the relative economic value of the ensemble prediction to quantitatively evaluate the benefit of probabilistic predictions using a simple and widely used decision-analytic model²², see the [Supplementary Information](#) for a description. Figure 4a shows that DGMR provides the highest economic value relative to the baseline methods (has highest peak and greater area under the curve). We use 20 member ensembles and show three accumulation levels used for weather warnings by Met Éireann (the Irish Meteorological service uses warnings defined directly in mm h⁻¹; <https://www.met.ie/weather-warnings>). This analysis shows the ability of the generative ensemble to capture uncertainty, and we show the improvement with samples in Extended Data Figs. 4 and 9, and postage stamp plots to visualize the ensemble variability in Supplementary Data 1–3.

Fig. 4: DGMR provides greater decision-making value when assessed using both economic and cognitive analyses.

 **figure4**

a, Relative economic value analysis across 20 samples for three 90-min rainfall accumulations, using 4-km aggregations. UNet generates a single deterministic prediction. **b**, Meteorologist preferences for the case study in Fig. 1. **c**, Meteorologist rankings for medium rain (5 mm h^{-1}) cases. **d**, Meteorologist rankings for heavy rain (10 mm h^{-1}) cases. Horizontal bars show the percentage of meteorologists who chose each method as their first choice. Whisker lines show the Clopper–Pearson 95% confidence interval. Meteorologists significantly preferred DGMR to alternatives ($n = 56$, $P < 10^{-4}$).

[Source data](#).

Importantly, we ground this economic evaluation by directly assessing decision-making value using the judgments of expert meteorologists working in the 24/7 operational centre at the Met Office (the UK's national meteorology service). We conducted a two-phase experimental study to assess expert judgements of value, involving a panel of 56 experts. In phase 1, all meteorologists were asked to provide a ranked preference assessment on a set of nowcasts with the instruction that 'preference is based on [their] opinion of accuracy and value'. Each meteorologist assessed a unique set of nowcasts, which, at the population level, allows for uncertainty characteristics and meteorologist idiosyncrasies to be averaged out in reporting the

statistical effect. We randomly selected 20% of meteorologists to participate in a phase 2 retrospective recall interview³³.

Operational meteorologists seek utility in forecasts for critical events, safety and planning guidance. Therefore, to make meaningful statements of operational usefulness, our evaluation assessed nowcasts for high-intensity events, specifically medium rain (rates above 5 mm h^{-1}) and heavy rain (rates above 10 mm h^{-1}). Meteorologists were asked to rank their preferences on a sample of 20 unique nowcasts (from a corpus of 2,126 events, being all high-intensity events in 2019). Data were presented in the form shown in Fig. 1b,c, showing clearly the initial context at $T + 0 \text{ min}$, the ground truth at $T + 30 \text{ min}$, $T + 60 \text{ min}$, and $T + 90 \text{ min}$, and nowcasts from PySTEPS, axial attention and DGMR. The identity of the methods in each panel was anonymized and their order randomized. See the Methods for further details of the protocol and of the ethics approval for human subjects research.

The generative nowcasting approach was significantly preferred by meteorologists when asked to make judgments of accuracy and value of the nowcast, being their most preferred 89% (95% confidence interval (CI) [0.86, 0.92]) of the time for the 5 mm h^{-1} nowcasts (Fig. 4c; $P < 10^{-4}$), and 90% (95% CI [0.87, 0.92]) for the 10 mm h^{-1} nowcasts (Fig. 4d, $P < 10^{-4}$). We compute the P value assessing the binary decision whether meteorologists chose DGMR as their first choice using a permutation test with 10,000 resamplings. We indicate the Clopper–Pearson CI. This significant meteorologist preference is important as it is strong evidence that generative nowcasting can provide meteorologists with physical insight not provided by alternative methods, and provides a grounded verification of the economic value analysis in Fig. 4a.

Meteorologists were not swayed by the visual realism of the predictions, and their responses in the subsequent structured interviews showed that they approached this task by making deliberate judgements of accuracy, location, extent, motion and rainfall intensity, and reasonable trade-offs between these factors ([Supplementary Information](#), section C.6). In the phase 2 interviews, PySTEPS was described as “being too developmental which would be misleading”, that is, as having many “positional errors” and “much higher intensity compared with reality”. The axial attention model was described as “too bland”, that is, as being “blocky” and “unrealistic”, but had “good spatial extent”. Meteorologists described DGMR as having the “best envelope”, “representing the risk best”, as having “much higher detail compared to what [expert meteorologists] are used to at the moment”, and as capturing “both the size of convection cells and intensity the best”. In the cases where meteorologists chose PySTEPS or the axial attention as their first choice, they pointed out that DGMR showed decay in the intensity for heavy rainfall at $T + 90 \text{ min}$ and had difficulty predicting isolated showers, which are important future improvements for the method.

See the [Supplementary Information](#) for further reports from this phase of the meteorologist assessment.

Conclusion

Skilful nowcasting is a long-standing problem of importance for much of weather-dependent decision-making. Our approach using deep generative models directly tackles this important problem, improves on existing solutions and provides the insight needed for real-world decision-makers. We showed—using statistical, economic and cognitive measures—that our approach to generative nowcasting provides improved forecast quality, forecast consistency and forecast value, providing fast and accurate short-term predictions at lead times where existing methods struggle.

Yet, there remain challenges for our approach to probabilistic nowcasting. As the meteorologist assessment demonstrated, our generative method provides skilful predictions compared to other solutions, but the prediction of heavy precipitation at long lead times remains difficult for all approaches. Critically, our work reveals that standard verification metrics and expert judgments are not mutually indicative of value, highlighting the need for newer quantitative measurements that are better aligned with operational utility when evaluating models with few inductive biases and high capacity. Whereas existing practice focuses on quantitative improvements without concern for operational utility, we hope this work will serve as a foundation for new data, code and verification methods—as well as the greater integration of machine learning and environmental science in forecasting larger sets of environmental variables—that makes it possible to both provide competitive verification and operational utility.

Methods

We provide additional details of the data, models and evaluation here, with references to extended data that add to the results provided in the main text.

Datasets

A dataset of radar for the UK was used for all the experiments in the main text. Additional quantitative results on a US dataset are available in [Supplementary Information section A](#).

UK dataset

To train and evaluate nowcasting models over the UK, we use a collection of radar composites from the Met Office RadarNet4 network. This network comprises more than 15 operational, proprietary C-band dual polarization radars covering 99% of the UK (see figure 1 in ref. ³⁴). We refer to ref. ¹¹ for details about how radar reflectivity is post-processed to obtain the two-dimensional radar composite field, which includes orographic enhancement and mean field adjustment using rain gauges. Each grid cell in the $1,536 \times 1,280$ composite represents the surface-level precipitation rate (in mm h^{-1}) over a $1 \text{ km} \times 1 \text{ km}$ region in the OSGB36 coordinate system. If a precipitation rate is missing (for example, because the location is not covered by any radar, or if a radar is out of order), the corresponding grid cell is assigned a negative value which is used to mask the grid cell at training and evaluation time. The radar composites are quantized in increments of $1/32 \text{ mm h}^{-1}$.

We use radar collected every five minutes between 1 January 2016 and 31 December 2019. We use the following data splits for model development. Fields from the first day of each month from 2016 to 2018 are assigned to the validation set. All other days from 2016 to 2018 are assigned to the training set. Finally, data from 2019 are used for the test set, preventing data leakage and testing for out of distribution generalization. For further experiments testing in-distribution performance using a different data split, see [Supplementary Information section C](#).

Training set preparation

Most radar composites contain little to no rain. Supplementary Table ² shows that approximately 89% of grid cells contain no rain in the UK. Medium to heavy precipitation (using rain rate above 4 mm h^{-1}) comprises fewer than 0.4% of grid cells in the dataset. To account for this imbalanced distribution, the dataset is rebalanced to include more data with heavier precipitation radar observations, which allows the models to learn useful precipitation predictions.

Each example in the dataset is a sequence of 24 radar observations of size $1,536 \times 1,280$, representing two continuous hours of data. The maximum rain rate is capped at 128 mm h^{-1} , and sequences that are missing one or more radar observations are removed. 256×256 crops are extracted and an importance sampling scheme is used to reduce the number of examples containing little precipitation. We describe this importance sampling and the parameters used in [Supplementary Information section A.1](#). After subsampling and removing entirely masked examples, the number of examples in the training set is roughly 1.5 million.

Model details and baselines

Here, we describe the proposed method and the three baselines to which we compare performance. When applicable, we describe both the architectures of the models and the training methods. There is a wealth of prior work, and we survey them as additional background in [Supplementary Information section E](#).

DGMR

A high-level description of the model was given in the main text and in Fig. [1a](#), and we provide some insight into the design decisions here.

Architecture design

The nowcasting model is a generator that is trained using two discriminators and an additional regularization term. Extended Data Fig. [1](#) shows a detailed schematic of the generative model and the discriminators. More precise descriptions of these architectures are given in Supplement B and corresponds to the code description; pseudocode is also available in the [Supplementary Information](#).

The generator in Fig. [1a](#) comprises the conditioning stack which processes past four radar fields that is used as context. Making effective use of such context is typically a challenge for conditional generative models, and this stack structure allows information from the context data to be used at multiple resolutions, and is used in other competitive video GAN models, for example, in ref. [26](#). This stack produces a context representation that is used as an input to the sampler. A latent conditioning stack takes samples from $N(0, 1)$ Gaussian distribution, and reshapes into a second latent representation. The sampler is a recurrent network formed with convolutional gated recurrent units (GRUs) that uses the context and latent representations as inputs. The sampler makes predictions of 18 future radar fields (the next 90 min). This architecture is both memory efficient and has had success in other forecasting applications. We also made comparisons with longer context using the past 6 or 8 frames, but this did not result in appreciable improvements.

Two discriminators in Fig. [1b](#) are used to allow for adversarial learning in space and time. The spatial and temporal discriminator share the same structure, except that the temporal discriminator uses 3D convolutions to account for the time dimension. Only 8 out of 18 lead times are used in the spatial discriminator, and a random 128×128 crop used for the temporal discriminator. These choices allow the models to fit within memory. We include a spatial attention block in the latent conditioning stack since it allows the model to be more robust across different types of regions and events, and provides an implicit regularization to prevent overfitting, particularly for the US dataset.

Both the generator and discriminators use spectrally normalized convolutions throughout, similar to ref. 35, since this is widely established to improve optimization. During model development, we initially found that including a batch normalization layer (without variance scaling) prior to the linear layer of the two discriminators improved training stability. The results presented use batch normalization, but we later were able to obtain nearly identical quantitative and qualitative results without it.

Objective function

The generator is trained with losses from the two discriminators and a grid cell regularization term (denoted $\|\mathbf{L} - \mathbf{R}\|_{\text{theta}}$). The spatial discriminator D_ϕ has parameters ϕ , the temporal discriminator T_ψ has parameters ψ , and the generator G_θ has parameters θ . We indicate the concatenation of two fields using the notation $\{\mathbf{X}; G\}$. The generator objective that is maximized is

```

$ \$\begin{array}{c} \{\{\mathscr{L}\}\}_\_\{\rm G\}(\{\boldsymbol{\theta}\})= \\ \{\mathbb{E}\}_\_\{\bf X\}_\_{1\{\rm M+N\}}[\{\mathbb{E}\}_\_\{\bf Z\}] \\ [D(\{G\}_\_\{\boldsymbol{\theta}\})(\{\bf Z\}\{\rm M\}) \\ \{\bf X\}_\_{1\{\rm M\}})+T(\{\rm M\}\{\bf X\}_\_{1\{\rm M\}}\{\rm M\}) \\ \{G\}_\_\{\boldsymbol{\theta}\}](\{\bf Z\}\{\rm M\})\{\bf X\}_\_{1\{\rm M\}} \\ \{\rm M\}) \\ \{\rm M\})]-\lambda \{\mathscr{L}\}_\_R(\{\boldsymbol{\theta}\}) \\ \{\rm M\}) \end{array} \$\$

(2)

\$ \$\begin{array}{c} \{\mathscr{L}\}_\_\{\rm R\}(\{\boldsymbol{\theta}\})=\frac{1}{\{HWN\}\{\rm M\}}[\{\mathbb{E}\}_\_\{\bf Z\}]\{G\}_\_\{\boldsymbol{\theta}\} \\ (\{\bf Z\}\{\rm M\})\{\bf X\}_\_{1\{\rm M\}})- \\ \{\bf X\}_\_{M+1\{\rm M+N\}}]\cdot w(\{\bf X\}_\_{M+1\{\rm M+N\}}) \\ \{\rm M\})_1 \end{array} \$\$

(3)

```

We use Monte Carlo estimates for expectations over the latent \mathbf{Z} in equations (2) and (3). These are calculated using six samples per input $\mathbf{X}_{1:M}$, which comprises $M=4$ radar observations. The grid cell regularizer ensures that the mean prediction remains close to the ground truth, and is averaged across all grid cells along the height H , width W and lead-time N axes. It is weighted towards heavier rainfall targets using the function $w(y) = \max(y + 1, 24)$, which operate element-wise for input vectors, and is clipped at 24 for robustness to spuriously large values in the radar. The GAN spatial discriminator loss $\langle \{\{\mathbf{L}\}\} \rangle_{\{\mathbf{D}\}}(\phi)$ and temporal discriminator loss $\langle \{\{\mathbf{L}\}\} \rangle_{\{\mathbf{T}\}}(\psi)$ are minimized with respect to parameters ϕ and ψ , respectively; $\text{ReLU}(x) = \max(0, x)$. The discriminator losses use a hinge loss formulation²⁶:

```
 $$\begin{array}{c} \{\mathscr{L}\}\}_{\{\rm D\}}\}(\varphi) = \\ \{\mathbb{E}\}_{\{\bf X\}}_{\{1\rm M+N\},\{\bf Z\}}[\rm ReLU](1- \\ \{D\}_{\varphi}(\{\bf X\}_{\{M+1\rm M+N\}})+\rm ReLU)(1+\{D\}_{\varphi} \\ \{G(\{\bf Z\}\rm M)\{\bf X\}}_{\{1\rm M\}})). \end{array} $$
```

(4)

```

$ \$\begin{array}{c} \{\mathscr{L}\}\}_{\rm T}(\psi) = \\ \{\mathbb{E}\}_{\bf X}[\rm ReLU(T-M+N)\bf Z] \\ + \rm ReLU(1+\psi(\rm ReLU(M+N))) \\ G(\bf Z)(M-\psi(\rm ReLU(M))) \end{array} $$

```

(3)

Evaluation

During evaluation, the generator architecture is the same, but unless otherwise noted, full radar observations of size $1,536 \times 1,280$, and latent variables with height and width $1/32$ of the radar observation size ($48 \times 40 \times 8$ of independent draws from a normal distribution), are used as inputs to the conditioning stack and latent conditioning stack, respectively. In particular, the latent conditioning stack allows for spatiotemporally consistent predictions for much larger regions than those on which the generator is trained.

For operational purposes and decision-making, the most important aspect of a probabilistic prediction is its resolution³⁶. Specific applications will require different requirements on reliability that can often be addressed by post-processing and calibration. We develop one possible post-processing approach to improve the reliability of the generative nowcasts. At prediction time, the latent variables are samples from a Gaussian distribution with standard deviation 2 (rather than 1), relying on empirical insights on maintaining resolution while increasing sample diversity in generative models^{24,37}. In addition, for each realization we apply a stochastic perturbation to the input radar by multiplying a single constant drawn from a unit-mean gamma distribution $G(\alpha = 5, \beta = 5)$ to the entire input radar field. Extended Data Figures 4 (UK) and 9 (US) shows how the post-processing improves the reliability diagram and rank histogram compared to the uncorrected approach.

Training

The model is trained for 5×10^5 generator steps, with two discriminator steps per generator step. The learning rate for the generator is 5×10^{-5} , and is 2×10^{-4} for the discriminator and uses Adam optimizer³⁸ with $\beta_1 = 0.0$ and $\beta_2 = 0.999$. The scaling parameter for the grid cell regularization is set to $\lambda = 20$, as this produced the best

continuous ranked probability score results on the validation set. We train on 16 tensor processing unit cores (<https://cloud.google.com/tpu>) for one week on random crops of the training dataset of size 256×256 measurements using a batch size of 16 per training step. The [Supplementary Information](#) contains additional comparisons showing the contributions of the different loss components to overall performance. We evaluated the speed of sampling by comparing speed on both CPU (10 core AMD EPYC) and GPU (NVIDIA V100) hardware. We generate ten samples and report the median time: for CPU the median time per sample was 25.7 s, and 1.3 s for the GPU.

UNet baseline

We use a UNet encoder–decoder model as strong baseline similarly to how it was used in related studies^{5,15}, but we make architectural and loss function changes that improve its performance at longer lead times and heavier precipitation. First, we replace all convolutional layers with residual blocks, as the latter provided a small but consistent improvement across all prediction thresholds. Second, rather than predicting only a single output and using autoregressive sampling during evaluation, the model predicts all frames in a single forward pass. This somewhat mitigates the excessive blurring found in ref. ⁵ and improves results on quantitative evaluation. Our architecture consists of six residual blocks, where each block doubles the number of channels of the latent representation followed by spatial down-sampling by a factor of two. The representation with the highest resolution has 32 channels which increases up to 1,024 channels.

Similar to ref. ⁶, we use a loss weighted by precipitation intensity. Rather than weighting by precipitation bins, however, we reweight the loss directly by the precipitation to improve results on thresholds outside of the bins specified by ref. ⁶. Additionally, we truncate the maximum weight to 24 mm h^{-1} as an error in reflectivity of observations leads to larger error in the precipitation values. We also found that including a mean squared error loss made predictions more sensitive to radar artefacts; as a result, the model is only trained with precipitation weighted mean average error loss.

The model is trained with batch size eight for 1×10^6 steps, with learning rate 2×10^{-4} with weight decay, using the Adam optimizer with default exponential rates. We select a model using early stopping on the average area under the precision–recall curve on the validation set. The UNet baselines are trained with 4 frames of size 256×256 as context.

Axial attention baseline

As a second strong deep learning-based baseline, we adapt the MetNet model¹⁹, which is a combination of a convolutional long short-term memory (LSTM) encoder¹⁷ and an axial attention decoder³⁹, for radar-only nowcasting. MetNet was demonstrated to achieve strong results on short-term (up to 8 h) low precipitation forecasting using radar and satellite data of the continental USA, making per-grid-cell probabilistic predictions and factorizing spatial dependencies using alternating layers of axial attention.

We modified the axial attention encoder–decoder model to use radar observations only, as well as to cover the spatial and temporal extent of data in this study. We rescaled the targets of the model to improve its performance on forecasts of heavy precipitation events. After evaluation on both UK and US data, we observed that additional satellite or topographical data as well as the spatiotemporal embeddings did not provide statistically significant CSI improvement. An extended description of the model and its adaptations is provided in [Supplementary Information section D](#).

The only prediction method described in ref. ¹⁹ is the per-grid-cell distributional mode, and this is considered the default method for comparison. To ensure the strongest baseline model, we also evaluated other prediction approaches. We assessed using independent samples from the per-grid-cell marginal distributions, but this was not better than using the mode when assessed quantitatively and qualitatively. We also combined the marginal distributions with a Gaussian process copula, in order to incorporate spatiotemporal correlation similar to the stochastically perturbed parametrization tendencies (SPPT) scheme of ref. ⁴⁰. We used kernels and correlation scales chosen to minimize spatiotemporally pooled CRPS metrics. The best performing was the product of a Gaussian kernel with 25 km spatial correlation scale, and an AR(1) kernel with 60 min temporal correlation scale. Results, however, were not highly sensitive to these choices. All settings resulted in samples that were not physically plausible, due to the stationary and unconditional correlation structure. These samples were also not favoured by external experts. Hence, we use the mode prediction throughout.

PySTEPS baseline

We use the PySTEPS implementation from ref. ⁴ using the default configuration available at <https://github.com/pySTEPS/pysteps>. Refs. ^{3,4} provide more details of this approach. In our evaluation, unlike other models evaluated that use inputs of size 256×256 , PySTEPS is given the advantage of being fed inputs of size 512×512 , which was found to improve its performance. PySTEPS includes post-processing using probability matching to recalibrate its predictions and these are used in all results.

Performance evaluation

We evaluate our model and baselines using commonly used quantitative verification measures, as well as qualitatively using a cognitive assessment task with expert meteorologists. Unless otherwise noted, models are trained on years 2016–2018 and evaluated on 2019 (that is, a yearly split).

Expert meteorologist study

The expert meteorologist study described is a two-phase protocol consisting of a ranked comparison task followed by a retrospective recall interview. The study was submitted for ethical assessment to an independent ethics committee and received favourable review. Key elements of the protocol involved consent forms that clearly explained the task and time commitment, clear messaging on the ability to withdraw from the study at any point, and that the study was not an assessment of the meteorologist's skills and would not affect their employment and role in any way. Meteorologists were not paid for participation, since involvement in these types of studies is considered part of the broader role of the meteorologist. The study was anonymized, and only the study lead had access to the assignment of experimental IDs. The study was restricted to meteorologists in guidance-related roles, that is, meteorologists whose role is to interpret weather forecasts, synthesize forecasts and provide interpretations, warnings and watches. Fifty-six meteorologists agreed to participate in the study.

Phase 1 of the study, the rating assessment, involved each meteorologist receiving a unique form as part of their experimental evaluation. The axial attention mode prediction is used in the assessment, and this was selected as the most appropriate prediction during the pilot assessment of the protocol by the chief meteorologist. The phase 1 evaluation comprised an initial practice phase of three judgments for meteorologists to understand how to use the form and assign ratings, followed by an experimental phase that involved 20 trials that were different for every meteorologist, and a final case study phase in which all meteorologists rated the same three scenarios (the scenarios in Fig. 1a, and Extended Data Figs. 2 and 3); these three events were chosen by the chief meteorologist—who is independent of the research team and also did not take part in the study—as difficult events that would expose challenges for the nowcasting approaches we compare. Ten meteorologists participated in the subsequent retrospective recall interview. This interview involved an in-person interview in which experts were asked to explain the reasoning for their assigned rating and what aspects informed their decision-making. These interviews all used the same script for consistency, and these sessions were recorded with audio only. Once all the audio was transcribed, the recordings were deleted.

The 20 trials of the experimental phase were split into two parts, each containing ten trials. The first ten trials comprised medium rain events (rainfall greater than 5 mm

h^{-1}) and the second 10 trials comprised heavy rain events (rainfall greater than 10 mm h^{-1}). 141 days from 2019 were chosen by the chief meteorologist as having medium-to-heavy precipitation events. From these dates, radar fields were chosen algorithmically according to the following procedure. First, we excluded from the crop selection procedure the 192 km that forms the image margins of each side of the radar field. Then, the crop over 256 km regions, containing the maximum fraction of grid cells above the given threshold, 5 or 10 mm h^{-1} , was selected from the radar image. If there was no precipitation in the frame above the given threshold, the selected crop was the one with the maximum average intensity. We use predictions without post-processing in the study. Each meteorologist assessed a unique set of predictions, which allows us to average over the uncertainty in predictions and individual preference to show statistical effect.

Extended Data Figure 2 shows a high-intensity precipitation front with decay and Extended Data Fig. 3 shows a cyclonic circulation event (low-pressure area), both of which are difficult for current deep learning models to predict. These two cases were also assessed by all expert meteorologists as part of the evaluative study, and in both cases, meteorologists significantly preferred the generative approach ($n = 56$, $P < 10^{-4}$) to competing methods. For the high-intensity precipitation front in Extended Data Fig. 2, meteorologists ranked first the generative approach in 73% of cases. Meteorologists reported that DGMR has “decent accuracy with both the shape and intensity of the feature … but loses most of the signal for embedded convection by $T + 90$ ”. PySTEPS is “too extensive with convective cells and lacks the organisation seen in the observations”, and the axial attention model as “highlighting the worst areas” but “looks wrong”.

For the cyclonic circulation in Extended Data Fig. 3, meteorologists ranked first the generative approach in 73% of cases. Meteorologists reported that it was difficult to judge this case between DGMR and PySTEPS. When making their judgment, they chose DGMR since it has “best fit and rates overall”. DGMR “captures the extent of precipitation overall [in the] area, though slightly overdoes rain coverage between bands”, whereas PySTEPS “looks less spatially accurate as time goes on”. The axial attention model “highlights the area of heaviest rain although its structure is unrealistic and too binary”. We provide additional quotes in [Supplementary Information section C.6](#).

Quantitative evaluation

We evaluate all models using established metrics²⁰: CSI, CRPS, Pearson correlation coefficient, the relative economic value^{22,41,42}, and radially averaged PSD. These are described in [Supplementary Information section F](#).

To make evaluation computationally feasible, for all metrics except PSD, we evaluate the models on a subsampled test set, consisting of 512×512 crops drawn from the full radar images. We use an importance sampling scheme (described in [Supplementary Information section A.1](#)) to ensure that this subsampling does not unduly compromise the statistical efficiency of our estimators of the evaluation metrics. The subsampling reduces the size of the test set to 66,851 and [Supplementary Information section C.3](#) shows that results obtained when evaluating CSI are not different when using the dataset with or without subsampling. All models except PySTEPS are given the centre 256×256 crop as input. PySTEPS is given the entire 512×512 crop as input as this improves its performance. The predictions are evaluated on the centre 64×64 grid cells, ensuring that models are not unfairly penalized by boundary effects. Our statistical significance tests use every other week of data in the test set (leaving $n = 26$ weeks) as independent units. We test the null hypothesis that performance metrics are equal for the two models, against the two-sided alternative, using a paired permutation test⁴³ with 10^6 permutations.

Extended Data Figure 4 shows additional probabilistic metrics that measure the calibration of the evaluated methods. This figure shows a comparison of the relative economic value of the probabilistic methods, showing DGMR providing the best value. We also show how the uncertainty captured by the ensemble increases as the number of samples used is increased from 1 to 20.

Extended Data Figure 5 compares the performance to that of an NWP, using the UKV deterministic forecast⁴⁴, showing that NWPs are not competitive in this regime. See [Supplementary Information section C.2](#) for further details of the NWP evaluation.

To verify other generalization characteristics of our approach—as an alternative to the yearly data split that uses training data of 2016–2018 and tests on 2019—we also use a weekly split: where the training, validation and test sets comprise Thursday through Monday, Tuesday, and Wednesday, respectively. The sizes of the training and test datasets are 1.48 million and 36,106, respectively. Extended Data Figure 6 shows the same competitive verification performance of DGMR in this generalization test.

To further assess the generalization of our method, we evaluate on a second dataset from the USA using the multi-radar multi-sensitivity (MRMS) dataset, which consists of radar composites for years 2017–2019⁴⁵. We use two years for training and one year for testing, and even with this more limited data source, our model still shows competitive performance relative to the other baselines. Extended Data Figs. 7–9 compares all methods on all metrics we have described, showing both the generalization and skilful performance on this second dataset. The [Supplementary Information](#) contains additional comparisons on performance with different initializations and performance of different loss function components.

Data availability

Processed radar data for the UK yearly data split is released under a creative commons licence. A smaller dataset for exploratory analysis is freely available, and the full dataset (around 1 TB) is also available; for details, see github.com/deepmind/deepmind-research/tree/master/nowcasting. The associated datasets contain public sector information licensed by the Met Office under the UK Open Government Licence 3.0. For the raw data, other licences, and alternative time periods, the data from the UK can be obtained with appropriate agreements from the Met Office; see <https://www.metoffice.gov.uk/research/weather/observations-research/radar-products> or contact the Met Office Data Provisioning Team using dpt@metoffice.gov.uk. The multi-radar multi-sensor (MRMS) dataset is available with appropriate agreements from NOAA; see <https://www.nssl.noaa.gov/projects/mrms/> or contact the MRMS data teams using mrms@noaa.gov. [Source data](#) are provided with this paper.

Code availability

We rely on several open-source code frameworks including Iris (scitools-iris.readthedocs.io), Cartopy (scitools.org.uk/cartopy), TensorFlow (www.tensorflow.org), and Colab (colab.sandbox.google.com). We have also used open-source tools for PySTEPS (pysteps.github.io), and for Axial Attention (github.com/google-research/google-research/tree/master/axial). The pseudocode for the generative algorithm can be found in the file `pseudocode.py` in the [Supplementary Information](#). All the neural architecture details and hyperparameters are described in Methods and Supplement. Alongside this model pseudocode, we have also released a pretrained version of the generative model available at github.com/deepmind/deepmind-research/tree/master/nowcasting.

References

1. 1.

Wilson, J. W., Feng, Y., Chen, M. & Roberts, R. D. Nowcasting challenges during the Beijing Olympics: Successes, failures, and implications for future nowcasting systems. *Weather Forecast.* **25**, 1691–1714 (2010).

2. 2.

Schmid, F., Wang, Y. & Harou, A. *Guidelines for Nowcasting Techniques* vol. 1198 (World Meteorological Organization, 2017).

3. 3.

Bowler, N. E., Pierce, C. E. & Seed, A. W. STEPS: a probabilistic precipitation forecasting scheme which merges an extrapolation nowcast with downscaled NWP. *Quart. J. Roy. Meteorol. Soc.* **132**, 2127–2155 (2006).

4. 4.

Pulkkinen, S. et al. PySTEPS: an open-source Python library for probabilistic precipitation nowcasting (v1. 0). *Geosci. Mod. Dev.* **12**, 4185–4219 (2019).

5. 5.

Ayzel, G., Scheffer, T. & Heistermann, M. Rainnet v1.0: a convolutional neural network for radar-based precipitation nowcasting. *Geosci. Mod. Dev.* **13**, 2631–2644 (2020).

6. 6.

Shi, X. et al. Deep learning for precipitation nowcasting: a benchmark and a new model. In *Advances in Neural Information Processing Systems* vol. 30, 5617–5627 (NeurIPS, 2017).

7. 7.

Toth, Z. & Kalnay, E. Ensemble forecasting at NCEP and the breeding method. *Mon. Weather Rev.* **125**, 3297–3319 (1997).

8. 8.

Pierce, C., Seed, A., Ballard, S., Simonin, D. & Li, Z. In *Doppler Radar Observations: Weather Radar, Wind Profiler, Ionospheric Radar, and Other Advanced Applications* (eds Bech, J. & Chau, J. L.) 97–142 (IntechOpen, 2012).

9. 9.

Sun, J. Convective-scale assimilation of radar data: progress and challenges. *Quart. J. Roy. Meteorol. Soc.* **131**, 3439–3463 (2005).

10. 10.

Buehner, M. & Jacques, D. Non-Gaussian deterministic assimilation of radar-derived precipitation accumulations. *Mon. Weather Rev.* **148**, 783–808 (2020).

11. 11. Harrison, D. et al. The evolution of the Met Office radar data quality control and product generation system: Radarnet. In *AMS Conference on Radar Meteorology* 14–18 (AMS, 2015).
12. 12. Germann, U. & Zawadzki, I. Scale dependence of the predictability of precipitation from continental radar images. Part II: probability forecasts. *J. Appl. Meteorol.* **43**, 74–89 (2004).
13. 13. Imhoff, R., Brauer, C., Overeem, A., Weerts, A. & Uijlenhoet, R. Spatial and temporal evaluation of radar rainfall nowcasting techniques on 1,533 events. *Water Resour. Res.* **56**, e2019WR026723 (2020).
14. 14. Lebedev, V. et al. Precipitation nowcasting with satellite imagery. In *International Conference on Knowledge Discovery & Data Mining* 2680–2688 (ACM, 2019).
15. 15. Agrawal, S. et al. Machine learning for precipitation nowcasting from radar images. Preprint at <https://arxiv.org/abs/1912.12132> (2019).
16. 16. Trebing, K., Stańczyk, T. & Mehrkanon, S. SmaAt-UNet: precipitation nowcasting using a small attention-UNet architecture. *Pattern Recog. Lett.* **145**, 178–186 (2021).
17. 17. Xingjian, S. et al. Convolutional LSTM network: a machine learning approach for precipitation nowcasting. In *Advances in Neural Information Processing Systems* vol. 28, 802–810 (NeurIPS, 2015).
18. 18. Foresti, L., Sideris, I. V., Nerini, D., Beusch, L. & Germann, U. Using a 10-year radar archive for nowcasting precipitation growth and decay: a probabilistic

machine learning approach. *Weather Forecast.* **34**, 1547–1569 (2019).

19. 19.

Sønderby, C. K. et al. MetNet: a neural weather model for precipitation forecasting. Preprint at <https://arxiv.org/abs/2003.12140> (2020).

20. 20.

Jolliffe, I. T. & Stephenson, D. B. *Forecast Verification: A Practitioner's Guide in Atmospheric Science* (John Wiley & Sons, 2012).

21. 21.

Palmer, T. & Räisänen, J. Quantifying the risk of extreme seasonal precipitation events in a changing climate. *Nature* **415**, 512–514 (2002).

22. 22.

Richardson, D. S. Skill and relative economic value of the ECMWF ensemble prediction system. *Quart. J. Roy. Meteorol. Soc.* **126**, 649–667 (2000).

23. 23.

Goodfellow, I. et al. Generative adversarial nets. In *Advances in Neural Information Processing Systems* vol. 27, 2672–2680 (NeurIPS, 2014).

24. 24.

Brock, A., Donahue, J. & Simonyan, K. Large scale GAN training for high fidelity natural image synthesis. In *International Conference on Learning Representations* (ICLR, 2019).

25. 25.

Mirza, M. & Osindero, S. Conditional generative adversarial nets. Preprint at <https://arxiv.org/abs/1411.1784> (2014).

26. 26.

Clark, A., Donahue, J. & Simonyan, K. Adversarial video generation on complex datasets. Preprint at <https://arxiv.org/abs/1907.06571> (2019).

27. 27.

Schaefer, J. T. The critical success index as an indicator of warning skill. *Weather Forecast.* **5**, 570–575 (1990).

28. 28.

Harris, D., Foufoula-Georgiou, E., Drogemeier, K. K. & Levit, J. J. Multiscale statistical properties of a high-resolution precipitation forecast. *J. Hydrol.* **2**, 406–418 (2001).

29. 29.

Sinclair, S. & Pegram, G. Empirical mode decomposition in 2-D space and time: a tool for space-time rainfall analysis and nowcasting. *Hydrol. Earth Sys. Sci.* **9**, 127–137 (2005).

30. 30.

Gneiting, T. & Raftery, A. E. Strictly proper scoring rules, prediction, and estimation. *J. Am. Stat. Assoc.* **102**, 359–378 (2007).

31. 31.

Gilleland, E., Ahijevych, D., Brown, B. G., Casati, B. & Ebert, E. E. Intercomparison of spatial forecast verification methods. *Weather Forecast.* **24**, 1416–1430 (2009).

32. 32.

Guo, C., Pleiss, G., Sun, Y. & Weinberger, K. Q. On calibration of modern neural networks. In *International Conference on Machine Learning* vol. 34, 1321–1330 (ICLR, 2017).

33. 33.

Crandall, B. W. & Hoffman, R. R. In *The Oxford Handbook of Cognitive Engineering* (ed. Lee, J. D.) 229–239 (Oxford Univ. Press, 2013).

34. 34.

Fairman, J., Schultz, D., Kirshbaum, D., Gray, S. & Barrett, A. Climatology of size, shape, and intensity of precipitation features over Great Britain and Ireland. *J. Hydrometeorol.* **18**, 1595–1615 (2017).

35. 35.

Zhang, H., Goodfellow, I., Metaxas, D. & Odena, A. Self-attention generative adversarial networks. In *International Conference on Machine Learning* vol. 36, 7354–7363 (ICLR, 2019).

36. 36.

Atger, F. The skill of ensemble prediction systems. *Mon. Weather Rev.* **127**, 1941–1953 (1999).

37. 37.

Ravuri, S. & Vinyals, O. Classification accuracy score for conditional generative models. In *Advances in Neural Information Processing Systems* vol. 32 (NeurIPS, 2019).

38. 38.

Kingma, D. P. & Ba, J. Adam: a method for stochastic optimization. In *International Conference on Learning Representations* (ICLR, 2015).

39. 39.

Ho, J., Kalchbrenner, N., Weissenborn, D. & Salimans, T. Axial attention in multidimensional transformers. Preprint at <https://arxiv.org/abs/1912.12180> (2019).

40. 40.

Palmer, T. et al. *Stochastic Parametrization and Model Uncertainty* (ECMWF, 2009).

41. 41.

Roberts, N. M. & Lean, H. W. Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Weather Rev.* **136**, 78–97 (2008).

42. 42.

Schwartz, C. S. et al. Toward improved convection-allowing ensembles: Model physics sensitivities and optimizing probabilistic guidance with small ensemble membership. *Weather Forecast.* **25**, 263–280 (2010).

43. 43.

Edgington, E. & Onghena, P. *Randomization Tests* (CRC, 2007).

44. 44.

Bush, M. et al. The first Met Office unified model–JULES regional atmosphere and land configuration, RAL1. *Geosci. Mod. Dev.* **13**, 1999–2029 (2020).

45. 45.

Smith, T. M. et al. Multi-radar multi-sensor (MRMS) severe weather and aviation products: Initial operating capabilities. *Bull. Amer. Meteorol. Soc.* **97**, 1617–1630 (2016).

46. 46.

Jarvis, A., Reuter, H., Nelson, A. & Guevara, E. *Hole-Filled Seamless SRTM Data V4* (International Centre for Tropical Agriculture, 2008).

Acknowledgements

We acknowledge the expertise and contributions of the anonymous expert meteorologists that are central to the findings of this study. We thank our DeepMind colleagues, including A. Muldal, A. Pierce, D. Hassabis, D. Smith, E. White, J. Donahue, K. McKee, K. Kavukcuoglu, L. Bennet, L. Deason, M. Grimes, O. Vinyals, P. Luc and R. Ahamed; A. Banki-Horvath and L. Chumakova; and our colleagues from the Met Office, including C. Bartholomew, D. Suri, K. Norman, S. Adams, P. Davies and T. McCaie.

Author information

Author notes

1. These authors contributed equally: Suman Ravuri, Karel Lenc, Matthew Willson

Affiliations

1. DeepMind, London, UK

Suman Ravuri, Karel Lenc, Matthew Willson, Remi Lam, Piotr Mirowski, Sheleem Kashem, Amol Mandhane, Aidan Clark, Andrew Brock, Karen Simonyan, Raia Hadsell, Ellen Clancy & Shakir Mohamed

2. Met Office, Exeter, UK

Dmitry Kangin, Megan Fitzsimons, Maria Athanassiadou, Sam Madge, Rachel Prudden, Niall Robinson & Alberto Arribas

3. University of Exeter, Exeter, UK

Dmitry Kangin, Rachel Prudden & Niall Robinson

4. University of Reading, Reading, UK

Alberto Arribas

Contributions

S.M., A.A., E.C., S.R., N. R, K.S. and R.H. managed the research. S. Madge, M.A., D.K. and K.L. collected and prepared the raw data. R.L., M.W., S.K., K.L., A.M., D.K., R.P. and M.A. created data sets and pipelines. S.R., K.L., P.M., M.W., A.M., R.L., D.K., R.P., A.C. and A.B. wrote the software and conducted experiments. K.L., S.R., M.W., R.L., D.K. and S.K. produced the figures and plots. M.F., S.M., D.K. A.A., E.C. and S.R. established and ran the meteorologist evaluation. S.M., S.R., A.A., N.R., K.L., D.K., R.P., N.R., E.C., P.M., R.L., M.W. and M.A. wrote the paper. E.C., A.A., N.R., S.R. and S.M., managed licensing and legal agreements.

Corresponding author

Correspondence to Shakir Mohamed.

Ethics declarations

Competing interests

S.R., K.L., M.W., R.L., P.M., S.K., A.M., A.C., A.B., K.S., R.H., E.C. and S.M., are employees of DeepMind, a subsidiary of Alphabet Inc., and own Alphabet stock. D.K., M.F., R.P., M.A., S.M., S.A., N.R. and A.A. were employees of the UK Met Office during the entirety of this research. A.A. contributed to this research while at the Met Office and is now at Microsoft. Provisional patent 63/150,509 was filed covering the generative algorithm described in this paper, listing the authors S.R., K.L., M.J.W., R.L. and P.M as inventors. The authors declare no other competing interests related to the manuscript.

Additional information

Peer review information *Nature* thanks Imme Ebert-Uphoff 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 Detailed architecture of DGMR.

a, Generator architecture. S2D is space- to-depth operation; D2S is depth to space. **b**, Temporal discriminator architecture (top left), spatial discriminator (middle left), and latent conditioning stack (bottom left) of the generator. On the right are architectures for the G block (top), D and 3D block (middle), and L block (right). For all panels, (\uparrow) or (\downarrow) indicates spatial up-sampling or down-sampling, respectively.

Extended Data Fig. 2 Case study of performance on a challenging precipitation event starting at 2019-07-24 at 03:15 UK, showing two separate banded structures of intense rainfall in the north-east and south-west over northern England, DGMR is better able to predict the spatial coverage and convection compared to other methods over a longer time period, while not over-estimating the intensities, and is significantly preferred by meteorologists (73% first choice, $N=56$, $p < 2 \times 10^{-4}$).

a, Geographic context for the predictions. **b**, A single prediction at $T+30$, $T+60$, and $T+90$ min lead time for different models. CSI at thresholds 2 mm h^{-1} and 8 mm h^{-1} and CRPS for an ensemble of four samples shown in a bottom left corner. For axial attention we show the mode prediction and the single sample. Images are $256 \text{ km} \times 256 \text{ km}$. **c**, Expert meteorologist preference for the visualized prediction (axial attention uses the mode prediction; we report the percentage of meteorologists for their first-choice rating as well as the Clopper–Pearson 95% confidence interval). Maps produced with Cartopy and SRTM elevation data⁴⁶.

Extended Data Fig. 3 Case study of performance on a challenging precipitation event starting 2019-07-30 at 15:15 UK, showing a pattern of precipitation around a low-pressure area which is slow moving, resulting in the cyclonic banded structures over England, DGMR captures extent of precipitation overall over the area, though slightly

overdoes rain coverage between bands, and is significantly preferred by meteorologists (73% first choice, $N=56$, $p < 2 \times 10^{-4}$).

a, Geographic context for the predictions. **b**, A single prediction at $T+30$, $T+60$, and $T+90$ min lead time for different models. CSI at thresholds 2 mm h^{-1} and 8 mm h^{-1} and CRPS for an ensemble of four samples shown in a bottom left corner. For axial attention we show the mode prediction and the single sample. Images are $256 \text{ km} \times 256 \text{ km}$. **c**, Expert meteorologist preference for the visualized prediction (axial attention uses the mode prediction; we report the percentage of meteorologists for their first-choice rating as well as the Clopper–Pearson 95% confidence interval). Maps produced with Cartopy and SRTM elevation data⁴⁶.

Extended Data Fig. 4 Further verification scores for the UK in 2019.

a, Comparison of relative economic value across 20 samples of different models for different rain accumulations. UNet generates a single deterministic prediction. **b**, Effect of larger ensemble in increasing economic value. **c**, Pearson correlation coefficient of various models at grid-resolution (left), rain rates averaged over a 4 km aggregation (middle) and averaged over a 16 km aggregation (right). **d**, Reliability diagrams and sharpness plots for two precipitation thresholds for $T+60$ min predictions. **e**, Rank histogram at $T+60$ min.

Extended Data Fig. 5 Verification scores for the UK by yearly splits aligned with NWP initialization times.

a, CSI across 20 samples of different models across precipitation thresholds 1 mm h^{-1} (left), 4 mm h^{-1} , 8 mm h^{-1} (right). UNet generates a single deterministic prediction. **b**, CRPS of various models for original predictions (left), average rain rate over a 4 km aggregation (middle), and average rain rate over a 16 km aggregation (right). **c**, Radially averaged power spectral density for full-frame 2019 predictions for different models. Please note that these results are computed with significantly fewer examples of the UK yearly dataset due to the NWP lead time alignment.

Extended Data Fig. 6 Verification scores for the UK (weekly split).

a, CSI for 1 mm h^{-1} (left), 4 mm h^{-1} , 8 mm h^{-1} (right) precipitation thresholds. **b**, Radially averaged power spectral density for full-frame predictions at $T+30$ (left), $T+60$ (middle), and $T+90$ min (right). **c**, CRPS at grid-scale (left), rain rates averaged over a 4 km aggregation (middle), rain rates averaged over a 16 km aggregation (right). **d**, CRPS at grid scale (left), maximum rain rate over a 4 km aggregation (middle), and maximum rain rate over a 16 km aggregation (right). **e**, Relative

economic value analysis using 20 samples for three 90 min rainfall accumulations, using 4 km aggregation. UNet generates a single deterministic prediction.

Extended Data Fig. 7 Verification scores for the United States in 2019.

a, CSI for 1 mm h^{-1} (left), 4 mm h^{-1} , 8 mm h^{-1} (right) precipitation thresholds. **b**, CRPS at grid-resolution (left), CRPS for rain rates averaged over a $4 \text{ km} \times 4 \text{ km}$ area (middle), CRPS for rain rates averaged over a $16 \text{ km} \times 16 \text{ km}$ area (right). **c**, CRPS at grid-resolution (left), maximum rain rate over a $4 \text{ km} \times 4 \text{ km}$ area (middle), and maximum rain rate over a $16 \text{ km} \times 16 \text{ km}$ area (right). **d**, Relative economic value analysis across 20 samples of different models for three 90 min rainfall accumulations, using 4 km aggregation. UNet generates a single deterministic prediction.

Extended Data Fig. 8 Radially averaged power spectral density for the United States in 2019.

a, Map of United States with three $1,536 \times 1,536$ regions: Pacific Northwest (left), Midwest (middle), Northeast (right). **b**, Radially averaged power spectral density for Pacific Northwest region for different models at $T + 30$ (left), $T + 60$ (middle), and $T + 90 \text{ min}$ (right). **c**, Radially averaged power spectral density for Midwest region for different models at $T + 30$ (left), $T + 60$ (middle) and $T + 90 \text{ min}$ (right). **d**, Radially averaged power spectral density for Northeast region for different models at $T + 30$ (left), $T + 60$ (middle), and $T + 90 \text{ min}$ (right). Map produced with Cartopy.

Extended Data Fig. 9 Further verification scores for the United States in 2019.

a, Comparison of relative economic value for 20 samples for different rain accumulations. UNet generates a single deterministic prediction. **b**, Effect of larger ensemble in increasing economic value. **c**, Pearson correlation coefficient of various models at grid-resolution (left), rain rates averaged over a 4 km aggregation (middle), and rain rates averaged over a 16 km aggregation (right). **d**, Reliability diagrams and sharpness plots for two precipitation thresholds for $T + 60 \text{ min}$ predictions. **e**, Rank histogram at $T + 60 \text{ min}$.

Supplementary information

Supplementary Information

The Supplementary Information contains six sections: section A provides more details about the datasets used; section B gives more details of the generative model

architecture; section C provides additional experiments mentioned in the methods; section D gives a more detailed description of the re-implemented baselines; section E provides context of the related work in nowcasting research; section F describes the precise definitions of the metrics used and their variants.

Source data

[**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

Ravuri, S., Lenc, K., Willson, M. *et al.* Skilful precipitation nowcasting using deep generative models of radar. *Nature* **597**, 672–677 (2021).
<https://doi.org/10.1038/s41586-021-03854-z>

- Received: 17 February 2021
- Accepted: 27 July 2021
- Published: 29 September 2021

- Issue Date: 30 September 2021
- DOI: <https://doi.org/10.1038/s41586-021-03854-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/s41586-021-03854-z>

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

- Article
- [Published: 29 September 2021](#)

Mercury stable isotopes constrain atmospheric sources to the ocean

- [Martin Jiskra](#) ORCID: [orcid.org/0000-0003-4991-8122](#)^{1,2 na1},
- [Lars-Eric Heimbürger-Boavida](#) ORCID: [orcid.org/0000-0003-0632-5183](#)^{2,3 na1},
- [Marie-Maëlle Desgranges](#)³,
- [Mariia V. Petrova](#)³,
- [Aurélie Dufour](#)³,
- [Beatriz Ferreira-Araujo](#)²,
- [Jérémy Masbou](#)^{2,4},
- [Jérôme Chmeleff](#)²,
- [Melilotus Thyssen](#)³,
- [David Point](#) ORCID: [orcid.org/0000-0002-5218-7781](#)² &
- [Jeroen E. Sonke](#) ORCID: [orcid.org/0000-0001-7146-3035](#)²

Nature volume **597**, pages 678–682 (2021)

- 142 Accesses
- 144 Altmetric
- [Metrics details](#)

Subjects

- [Marine chemistry](#)

Abstract

Human exposure to toxic mercury (Hg) is dominated by the consumption of seafood^{1,2}. Earth system models suggest that Hg in marine ecosystems is supplied by atmospheric wet and dry Hg(ii) deposition, with a three times smaller contribution from gaseous Hg(0) uptake^{3,4}. Observations of marine Hg(ii) deposition and Hg(0) gas exchange are sparse, however⁵, leaving the suggested importance of Hg(ii) deposition⁶ ill-constrained. Here we present the first Hg stable isotope measurements of total Hg (tHg) in surface and deep Atlantic and Mediterranean seawater and use them to quantify atmospheric Hg deposition pathways. We observe overall similar tHg isotope compositions, with median $\Delta^{200}\text{Hg}$ signatures of 0.02‰, lying in between atmospheric Hg(0) and Hg(ii) deposition end-members. We use a $\Delta^{200}\text{Hg}$ isotope mass balance to estimate that seawater tHg can be explained by the mixing of 42% (median; interquartile range, 24–50%) atmospheric Hg(ii) gross deposition and 58% (50–76%) Hg(0) gross uptake. We measure and compile additional, global marine Hg isotope data including particulate Hg, sediments and biota and observe a latitudinal $\Delta^{200}\text{Hg}$ gradient that indicates larger ocean Hg(0) uptake at high latitudes. Our findings suggest that global atmospheric Hg(0) uptake by the oceans is equal to Hg(ii) deposition, which has implications for our understanding of atmospheric Hg dispersal and marine ecosystem recovery.

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](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

Fig. 1: Depth profiles of seawater Hg species concentrations and total and particulate Hg stable isotope composition at station K2 in the Mediterranean Sea and the North Atlantic.

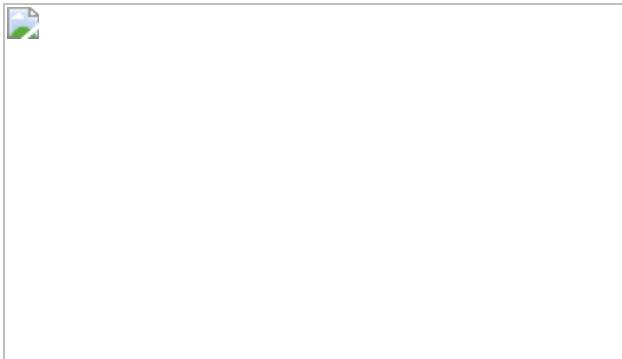


Fig. 2: Hg stable isotope composition of atmospheric Hg deposition sources (gaseous Hg(0) and Hg(ii) in rainfall), and seawater (total (tHg) and particulate (pHg)).

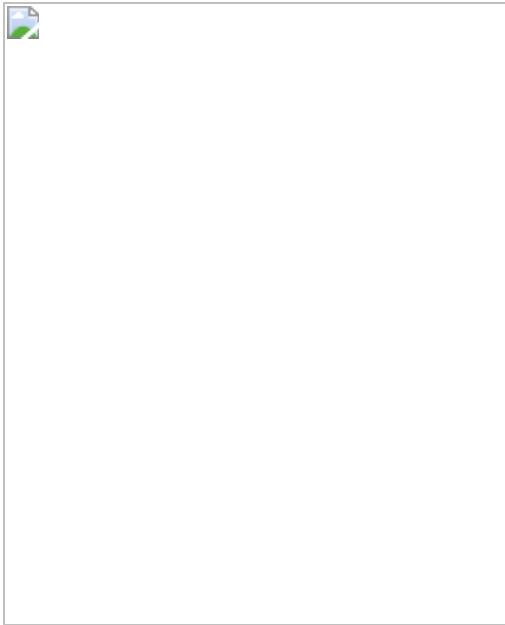
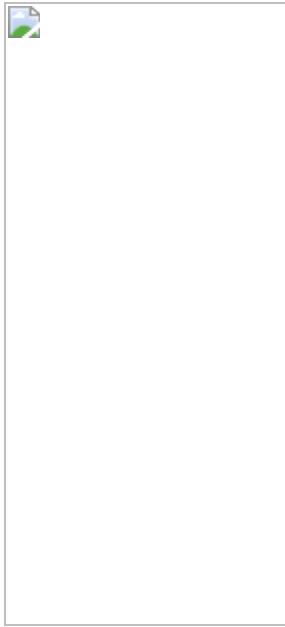


Fig. 3: Hg stable isotope composition in different ocean basins for total Hg (tHg) and particulate Hg (pHg) in seawater, marine sediments and marine fish.



Fig. 4: Latitudinal variation in atmospheric Hg sources and deposition fluxes to the global ocean.



Data availability

Hg stable isotope and Hg speciation data that support the findings of this study are available from <https://doi.org/10.5281/zenodo.4740464>. Source data are provided with this paper.

References

1. 1.

Sunderland, E. M. Mercury exposure from domestic and imported estuarine and marine fish in the U.S. seafood market. *Environ. Health Perspect.* **115**, 235–242 (2007).

2. 2.

Lavoie, R. A., Bouffard, A., Maranger, R. & Amyot, M. Mercury transport and human exposure from global marine fisheries. *Sci. Rep.* **8**, 6705 (2018).

3. 3.

Horowitz, H. M. et al. A new mechanism for atmospheric mercury redox chemistry: implications for the global mercury budget. *Atmos. Chem. Phys.* **17**, 6353–6371 (2017).

4. 4.

Travnikov, O. et al. Multi-model study of mercury dispersion in the atmosphere: atmospheric processes and model evaluation. *Atmos. Chem. Phys.* **17**, 5271–5295 (2017).

5. 5.

Zhang, L., Zhou, P., Cao, S. & Zhao, Y. Atmospheric mercury deposition over the land surfaces and the associated uncertainties in observations and simulations: a critical review. *Atmos. Chem. Phys.* **19**, 15587–15608 (2019).

6. 6.

Zhang, Y. et al. A coupled global atmosphere-ocean model for air-sea exchange of mercury: insights into wet deposition and atmospheric redox chemistry. *Environ. Sci. Technol.* **53**, 5052–5061 (2019).

7. 7.

Sheehan, M. C. et al. Global methylmercury exposure from seafood consumption and risk of developmental neurotoxicity: a systematic review. *Bull. World Health Organ.* **92**, 254–269 (2014).

8. 8.

Mason, R. P. & Fitzgerald, W. F. Alkylmercury species in the Equatorial Pacific. *Nature* **347**, 457–459 (1990).

9. 9.

Selin, N. E. Global biogeochemical cycling of mercury: a review. *Ann. Rev. Environ. Resour.* **34**, 43–63 (2009).

10. 10.

Fitzgerald, W. F., Lamborg, C. H. & Hammerschmidt, C. R. Marine biogeochemical cycling of mercury. *Chem. Rev.* **107**, 641–662 (2007).

11. 11.

Li, C. et al. Unequal anthropogenic enrichment of mercury in Earth's northern and southern hemispheres. *ACS Earth Space Chem.* **4**, 2073–2081 (2020).

12. 12.

Outridge, P. M., Mason, R. P., Wang, F., Guerrero, S. & Heimbürger-Boavida, L. E. Updated global and oceanic mercury budgets for the United Nations Global Mercury Assessment 2018. *Environ. Sci. Technol.* **52**, 11466–11477 (2018).

13. 13.

Lamborg, C. H. et al. A global ocean inventory of anthropogenic mercury based on water column measurements. *Nature* **512**, (2014).

14. 14.

Schartup, A. T. et al. Climate change and overfishing increase neurotoxicant in marine predators. *Nature* **572**, 648–650 (2019).

15. 15.

Sunderland, E. M. & Mason, R. P. Human impacts on open ocean mercury concentrations. *Global Biogeochem. Cycles* **21**, GB4022 (2007).

16. 16.

Amos, H. M. et al. Global biogeochemical implications of mercury discharges from rivers and sediment burial. *Environ. Sci. Technol.* **48**, 9514–9522 (2014).

17. 17.

Zhang, Y. et al. Biogeochemical drivers of the fate of riverine mercury discharged to the global and Arctic oceans. *Global Biogeochem. Cycles* **29**, 854–864 (2015).

18. 18.

Zhang, Y., Jaegle, L., Thompson, L. & Streets, D. G. Six centuries of changing oceanic mercury. *Global Biogeochem. Cycles* **28**, 1251–1261 (2014).

19. 19.

Soerensen, A. L. et al. Elemental mercury concentrations and fluxes in the tropical atmosphere and ocean. *Environ. Sci. Technol.* **48**, 11312–11319 (2014).

20. 20.

Kuss, J., Zülicke, C., Pohl, C. & Schneider, B. Atlantic mercury emission determined from continuous analysis of the elemental mercury sea-air concentration difference within transects between 50°N and 50°S. *Global Biogeochem. Cycles* **25**, GB3021 (2011).

21. 21.

Wang, J., Xie, Z., Wang, F. & Kang, H. Gaseous elemental mercury in the marine boundary layer and air-sea flux in the Southern Ocean in austral summer. *Sci. Total Environ.* **603–604**, 510–518 (2017).

22. 22.

Krabbenhoft, D. P. & Sunderland, E. M. Global change and mercury. *Science* **341**, 1457–1458 (2013).

23. 23.

Enrico, M. et al. Atmospheric mercury transfer to peat bogs dominated by gaseous elemental mercury dry deposition. *Environ. Sci. Technol.* **50**, 2405–2412 (2016).

24. 24.

Demers, J. D., Blum, J. D. & Zak, D. R. Mercury isotopes in a forested ecosystem: implications for air-surface exchange dynamics and the global mercury cycle. *Global Biogeochem. Cycles* **27**, 222–238 (2013).

25. 25.

Obrist, D. et al. Tundra uptake of atmospheric elemental mercury drives Arctic mercury pollution. *Nature* **547**, 201–204 (2017).

26. 26.

Chen, J., Hintelmann, H., Feng, X. & Dimock, B. Unusual fractionation of both odd and even mercury isotopes in precipitation from Peterborough, ON, Canada. *Geochim. Cosmochim. Acta* **90**, 33–46 (2012).

27. 27.

Fu, X. W. et al. Mass-independent fractionation of even and odd mercury isotopes during atmospheric mercury redox reactions. *Environ. Sci. Technol.* **55**, 10164–10174 (2021).

28. 28.

Enrico, M. et al. Holocene atmospheric mercury levels reconstructed from peat bog mercury stable isotopes. *Environ. Sci. Technol.* **51**, 5899–5906 (2017).

29. 29.

Štrok, M., Baya, P. A. & Hintelmann, H. The mercury isotope composition of Arctic coastal seawater. *C.R. Geosci.* **347**, 368–376 (2015).

30. 30.

Motta, L. C. et al. Mercury cycling in the North Pacific subtropical gyre as revealed by mercury stable isotope ratios. *Global Biogeochem. Cycles* **33**, 777–794 (2019).

31. 31.

Bowman, K. L., Lamborg, C. H. & Agather, A. M. A global perspective on mercury cycling in the ocean. *Sci. Total Environ.* **710**, 136166 (2020).

32. 32.

Cossa, D., Averyt, B. & Pirrone, N. The origin of methylmercury in open Mediterranean waters. *Limnol. Oceanogr.* **54**, 3 (2009).

33. 33.

Heimbürger, L.-E. et al. Methyl mercury distributions in relation to the presence of nanoand picophytoplankton in an oceanic water column (Ligurian Sea, north-western Mediterranean). *Geochim. Cosmochim. Acta* **74**, 5549–5559 (2010).

34. 34.

Cossa, D. et al. Mercury distribution and transport in the North Atlantic Ocean along the GEOTRACES-GA01 transect. *Biogeosciences* **15**, 2309–2323 (2018).

35. 35.

Gehrke, G. E., Blum, J. D. & Meyers, P. A. The geochemical behavior and isotopic composition of Hg in a mid-Pleistocene western Mediterranean sapropel. *Geochim. Cosmochim. Acta* **73**, 1651–1665 (2009).

36. 36.

Ogrinc, N., Hintelmann, H., Kotnik, J., Horvatl, M. & Pirrone, N. Sources of mercury in deep-sea sediments of the Mediterranean Sea as revealed by mercury stable isotopes. *Sci. Rep.* **9**, 11626 (2019).

37. 37.

Zhang, Y. et al. Biogeochemical drivers of the fate of riverine mercury discharged to the global and Arctic oceans. *Global Biogeochem. Cycles* **29**, 854–864 (2015).

38. 38.

Shah, V. & Jaeglé, L. Subtropical subsidence and surface deposition of oxidized mercury produced in the free troposphere. *Atmos. Chem. Phys.* **17**, 8999–9017 (2017).

39. 39.

Rolison, J. M., Landing, W. M., Cohen, M. D., Luke, W. & Salters, V. J. M. Isotopic composition of species-specific atmospheric Hg in a coastal environment. *Chem. Geol.* **336**, 37–49 (2012).

40. 40.

Washburn, S. J., Blum, J. D., Motta, L. C., Bergquist, B. A. & Weiss-Penzias, P. Isotopic composition of Hg in fogwaters of coastal California. *Environ. Sci. Technol. Lett.* **8**, 3–8 (2020).

41. 41.

Yu, B. et al. New evidence for atmospheric mercury transformations in the marine boundary layer from stable mercury isotopes. *Atmos. Chem. Phys.* **20**, 9713–9723 (2020).

42. 42.

Demers, J. D., Sherman, L. S., Blum, J. D., Marsik, F. J. & Dvonch, J. T. Coupling atmospheric mercury isotope ratios and meteorology to identify sources of mercury impacting a coastal urban-industrial region

near Pensacola, Florida, USA. *Global Biogeochem. Cycles* **29**, 1689–1705 (2015).

43. 43.

Fu, X. et al. Isotopic composition of gaseous elemental mercury in the marine boundary layer of East China Sea. *J. Geophys. Res. Atmos.* **123**, 7656–7669 (2018).

44. 44.

Semeniuk, K. & Dastoor, A. Development of a global ocean mercury model with a methylation cycle: outstanding issues. *Global Biogeochem. Cycles* **31**, 400–433 (2017).

45. 45.

Soerensen, A. L. et al. An improved global model for air-sea exchange of mercury: high concentrations over the North Atlantic. *Environ. Sci. Technol.* **44**, 8574–8580 (2010).

46. 46.

Method 1631, Revision E: Mercury in Water by Oxidation, Purge and Trap, and Cold Vapor Atomic Fluorescence Spectrometry (EPA, 2002).

47. 47.

Cutter, G. J. et al. Sampling and Sample-handling Protocols for GEOTRACES Cruises. *Version 3.0.* (2017).

48. 48.

Sun, R., Enrico, M., Heimbürger, L.-E., Scott, C. & Sonke, J. E. A double-stage tube furnace-acid-trapping protocol for the pre-concentration of mercury from solid samples for isotopic analysis. *Anal. Bioanal. Chem.* **405**, 6771–6781 (2013).

49. 49.

Heimbürger, L. E. et al. Shallow methylmercury production in the marginal sea ice zone of the central Arctic Ocean. *Sci. Rep.* **5**, 10318 (2015).

50. 50.

Jiskra, M., Sonke, J. E., Agnan, Y., Helmig, D. & Obrist, D. Insights from mercury stable isotopes on terrestrial-atmosphere exchange of Hg(0) in the Arctic tundra. *Biogeosciences* **16**, 4051–4064 (2019).

51. 51.

Blum, J. D. & Bergquist, B. A. Reporting of variations in the natural isotopic composition of mercury. *Anal. Bioanal. Chem.* **388**, 353–359 (2007).

52. 52.

Jiskra, M. et al. Mercury deposition and re-emission pathways in boreal forest soils investigated with Hg isotope signatures. *Environ. Sci. Technol.* **49**, 7188–7196 (2015).

53. 53.

Bergquist, B. A. & Blum, J. D. Mass-dependent and -independent fractionation of Hg isotopes by photoreduction in aquatic systems. *Science* **318**, 417–420 (2007).

54. 54.

Vermeesch, P. IsoplotR: a free and open toolbox for geochronology. *Geosci. Front.* **9**, 1479–1493 (2018).

55. 55.

Blum, J. D. & Johnson, M. W. Recent developments in mercury stable isotope analysis. *Rev. Mineral. Geochem.* **82**, 733–757 (2017).

Acknowledgements

This work was supported by research grants ANR-17-CE34-0010 MERTOX to D.P., FP7-IDEAS-ERC grant no. 258537, and H2020 ERA-PLANET grant no. 689443 via the iCUPE and iGOSP project to J.E.S., Chantier Arctique Francais funding via the Pollution in the Arctic System Project to J.E.S. and L.E.H.B., H2020 Marie Skłodowska-Curie grant no. 657195 and Swiss National Science Foundation grant PZ00P2_174101 to M.J., APOG DECOMAR, MISTRALS AT P&C and the AXA RF grants to L.E.H.B., and the French National Research Agency (ANR-13-BS06-0014, ANR-12-PDOC-0025-01), the French National Centre for Scientific Research (CNRS-LEFE-CYBER), the LabexMER (ANR-10-LABX-19), and Ifremer. We are grateful to G. Sarthou and P. Lherminier, chief scientists of the 2014 GEOVIDE cruise, and to H. Planquette for coordinating clean sampling. We thank M. Rutgers van der Loeff, T. Kanzow and the Alfred-Wegener-Institute for Polar and Marine Research for organizing the 2016 GRIFF cruise. We thank E. de Saint-Léger and F. Péault of the technical division of INSU for support with operations at sea. We thank L. Laffont for laboratory management and O. Grossot and D. Malengros for technical assistance. We thank L. Metral from MARBEC and F. Ménard from MIO for providing tuna fish samples from the Mediterranean Sea. We thank J. Kuss and H. Horowitz for discussion on gas exchange model parameterization. We thank the captains, crew and sampling teams onboard the RV Antedon II, RV Pourquoi Pas? and FS Polarstern for their support at sea. Thanks also go to the shipboard participants, captain and crew of the N/O l'Atalante for obtaining sediment samples from the 2015 VESPA cruise. VESPA was funded by the French Ministry of Research and Higher Education, with support from the governments of New Zealand and New Caledonia.

Author information

Author notes

1. These authors contributed equally: Martin Jiskra, Lars-Eric Heimbürger-Boavida

Affiliations

1. Environmental Geosciences, University of Basel, Basel, Switzerland

Martin Jiskra

2. Géosciences Environnement Toulouse, CNRS/IRD/Université Paul Sabatier Toulouse III, Toulouse, France

Martin Jiskra, Lars-Eric Heimbürger-Boavida, Beatriz Ferreira-Araujo, Jérémy Masbou, Jérôme Chmeleff, David Point & Jeroen E. Sonke

3. Aix Marseille Université, CNRS/INSU, Université de Toulon, IRD, Mediterranean Institute of Oceanography (MIO) UM 110, Marseille, France

Lars-Eric Heimbürger-Boavida, Marie-Maëlle Desgranges, Mariia V. Petrova, Aurélie Dufour & Melilotus Thyssen

4. Institut Terre et Environnement de Strasbourg, Université de Strasbourg/EOST/ENGEES/CNRS, Strasbourg, France

Jérémy Masbou

Contributions

L.E.H.B., J.E.S., M.J. and D.P. conceived the study. L.E.H.B., M.J., D.P., M.P., M.V.P., M.M.D. and J.E.S. performed sampling. J.E.S., M.J. and L.E.H.B. developed and applied the tHg isotope pre-concentration methods. J.E.S., M.J., J.M. and J.C. performed isotope measurements. M.M.D., M.V.P., A.D., L.E.H.B., M.J., J.M., D.P. and M.T. performed additional laboratory work. M.J., J.E.S. and L.E.H.B. analysed the data. J.E.S. and M.J. wrote the draft paper, which was improved by contributions from L.E.H.B. and D.P., and commented on by all authors.

Corresponding authors

Correspondence to Martin Jiskra or Lars-Eric Heimbürger-Boavida or Jeroen E. Sonke.

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 Summary of marine Hg\(ii\) deposition and Hg\(0\) air-sea exchange fluxes.](#)

Gross fluxes (solid arrows, $Mg\text{ y}^{-1}$) are based on published model estimates³. Hg(0) exchange is bidirectional, meaning that despite surface ocean Hg(0) supersaturation and large Hg(0) evasion, Hg(0) invasion is substantial. Marine $\Delta^{200}\text{Hg}$ signatures of 0.04‰ indicate a relatively more important contribution of the atmospheric Hg(0) end-member to marine Hg than current 3D models suggest. This indicates that either 3D model Hg(ii) deposition is overestimated or that Hg(0) invasion is underestimated (black dotted arrows, indicating 2–3 times lower or 2–3 times higher fluxes, required to fit $\Delta^{200}\text{Hg}$ data).

[Extended Data Fig. 2 Maps of sampling locations for total and particulate Hg isotopes.](#)

Top: sampling locations K2 in the Mediterranean Sea (purple), Atlantic Ocean (yellow) and Fram Strait (green). Bottom: magnification of the four Mediterranean locations, with main station K2 (large purple circle), and pHg station K1 and Julio (small purple circles), and Endoume pier in Marseille Bay (grey square). Maps were made with Ocean Data View (Schlitzer, Reiner, Ocean Data View, odv.awi.de, 2021).

[Extended Data Fig. 3 Latitudinal distribution of Hg\(ii\) wet deposition.](#)

Annual mean Hg(ii) wet deposition ($\mu\text{g m}^{-2} \text{y}^{-1}$) at oceanic locations in the northern and southern hemispheres (NH, SH), binned in 5° latitude. Mean values (\pm standard deviation, SD) were calculated when sufficient data was available per 5° latitude band, and interpolated using polynomial fitting when no data were available (in which case a mean observed SD of 30% was applied). MDN, mercury deposition network; GMOS, global mercury observation system; USA, CAN, PR, United States of America, Canada, Puerto Rico.

[Source data](#)

[Extended Data Fig. 4 Latitudinal distribution of dissolved gaseous Hg \(DGM\) concentrations.](#)

Mean (\pm standard deviation) DGM are binned in 5° latitude bands, and equal weight was given to each study. Polar studies, affected by sea ice show unusually high concentrations (mean 219 fM in the Arctic, mean 138 fM around Antarctica) for high latitude waters and were excluded in 5° latitude binning (replaced in calculations by ‘open water only’ DGM data at 55–60°S and 75–80°N).

[Source data](#)

[Extended Data Fig. 5 Atmospheric deposition pathways of the zonal reference model.](#)

a, Marine Hg(ii) gross deposition, Hg(0) gross invasion, Hg(0) gross evasion, and net Hg flux [Hg(ii) deposition + Hg(0) invasion – Hg(0) evasion]; all in $\mu\text{g m}^{-2} \text{y}^{-1}$ with evasion shown as negative numbers. Hg(0) invasion is driven by observed atmospheric Hg(0) and wind speed. Hg(ii) deposition is dominated by Hg(ii) wet deposition. Hg(0) evasion is driven by DGM concentrations and wind speed. The net Hg evasion trends shows important net deposition in the northern hemisphere, and net evasion in the southern hemisphere. **b**, Reference model Hg gross deposition fluxes ($\mu\text{g m}^{-2} \text{y}^{-1}$) as a function of latitude used in estimating marine $\Delta^{200}\text{Hg}$ in Fig. 4a (main text). Hg(ii) wet deposition observations as in Extended Data Fig. 3; Hg(ii) dry deposition was fixed at $5 \mu\text{g m}^{-2} \text{y}^{-1}$, and constrained as 40% of total Hg(ii) deposition³⁸, since no dry deposition observations over oceans exist. Hg(0) invasion (ocean uptake, same as in top panel) is calculated from the observed inter-hemispheric atmospheric gaseous Hg(0) gradient³, wind and sea surface temperature (Copernicus), and the latest Hg(0) air–sea gas exchange model (see Supplementary Information).

[Source data](#)

[Extended Data Fig. 6 Estimated latitudinal variation in \$\Delta^{200}\text{Hg}\$ of atmospheric Hg\(ii\) deposition.](#)

The small variation is caused by the variable contributions (Extended Data Figure 5) of Hg(ii) wet deposition with $\Delta^{200}\text{Hg}$ of 0.16‰, and Hg(ii) dry deposition with $\Delta^{200}\text{Hg}$ of 0.10‰ (Extended Data Table 1). The dashed line represents the median and the shaded area the interquartile range (IQR).

[Source data](#)

[Extended Data Fig. 7 Variation of \$\Delta^{204}\text{Hg}\$ in marine samples.](#)

a, $\Delta^{204}\text{Hg}$ versus $\Delta^{200}\text{Hg}$. The dashed line represents the York regression using IsoplotR⁵⁴ for all marine samples ($\Delta^{200}\text{Hg} = -0.32(\pm 0.06) \Delta^{204}\text{Hg} + (0.03 \pm 0.004)$, (\pm se), MSWD = 0.213). **b**, $\Delta^{204}\text{Hg}$ boxplot for 5° latitudinal intervals. Marine samples are shown in boxes, where the bold horizontal

line represents the median, the boxes the interquartile range, the whiskers 1.5 times the IQR and outliers are represented by dots. The solid line represents the predicted $\Delta^{204}\text{Hg}$ based on the observational relationship between $\Delta^{204}\text{Hg}$ and $\Delta^{200}\text{Hg}$ in terrestrial samples by Blum and Johnson, 2017⁵⁵. The dashed line represents the predicted $\Delta^{204}\text{Hg}$ derived from the York regression shown in panel **a**. $\Delta^{204}\text{Hg}$ data are available for 339 out of 791 marine samples. Note that for pHg and tHg samples presented here, $\Delta^{204}\text{Hg}$ was not measured because of the low abundance of the ^{204}Hg isotope, and unavailability of a second $10^{13} \Omega$ amplifier.

[Source data](#)

Extended Data Table 1 Summary of Hg stable isotope data

Extended Data Table 2 Modelled contribution of Hg(ii) from wet and dry deposition

Supplementary information

[Supplementary Information](#)

Supplementary Methods, Figs. 1–3 and references.

[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. 3](#)

[**Source Data Extended Data Fig. 4**](#)

[**Source Data Extended Data Fig. 5**](#)

[**Source Data Extended Data Fig. 6**](#)

[**Source Data Extended Data Fig. 7**](#)

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Jiskra, M., Heimbürger-Boavida, LE., Desgranges, MM. *et al.* Mercury stable isotopes constrain atmospheric sources to the ocean. *Nature* **597**, 678–682 (2021). <https://doi.org/10.1038/s41586-021-03859-8>

- Received: 08 July 2020
- Accepted: 28 July 2021
- Published: 29 September 2021
- Issue Date: 30 September 2021
- DOI: <https://doi.org/10.1038/s41586-021-03859-8>

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/s41586-021-03859-8>

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

- Article
- [Published: 29 September 2021](#)

Fine-root traits in the global spectrum of plant form and function

- [Carlos P. Carmona](#) [ORCID: orcid.org/0000-0001-6935-4913](#)^{1 na1},
- [C. Guillermo Bueno](#)^{1 na1},
- [Aurele Toussaint](#) [ORCID: orcid.org/0000-0002-5738-4637](#)^{1 na1},
- [Sabrina Träger](#)^{1,2 na1},
- [Sandra Díaz](#) [ORCID: orcid.org/0000-0003-0012-4612](#)^{3,4},
- [Mari Moora](#) [ORCID: orcid.org/0000-0002-4819-7506](#)¹,
- [Alison D. Munson](#)⁵,
- [Meelis Pärtel](#) [ORCID: orcid.org/0000-0002-5874-0138](#)¹,
- [Martin Zobel](#) [ORCID: orcid.org/0000-0001-7957-6704](#)¹ &
- [Riin Tamme](#) [ORCID: orcid.org/0000-0002-4998-1394](#)^{1 na1}

Nature volume **597**, pages 683–687 (2021)

- 394 Accesses
- 89 Altmetric
- [Metrics details](#)

Subjects

- [Biodiversity](#)
- [Natural variation in plants](#)

Abstract

Plant traits determine how individual plants cope with heterogeneous environments. Despite large variability in individual traits, trait coordination and trade-offs^{1,2} result in some trait combinations being much more widespread than others, as revealed in the global spectrum of plant form and function (GSPFF³) and the root economics space (RES⁴) for aboveground and fine-root traits, respectively. Here we combine the traits that define both functional spaces. Our analysis confirms the major trends of the GSPFF and shows that the RES captures additional information. The four dimensions needed to explain the non-redundant information in the dataset can be summarized in an aboveground and a fine-root plane, corresponding to the GSPFF and the RES, respectively. Both planes display high levels of species aggregation, but the differentiation among growth forms, families and biomes is lower on the fine-root plane, which does not include any size-related trait, than on the aboveground plane. As a result, many species with similar fine-root syndromes display contrasting aboveground traits. This highlights the importance of including belowground organs to the GSPFF when exploring the interplay between different natural selection pressures and whole-plant trait integration.

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](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

Fig. 1: The aboveground and fine-root planes.

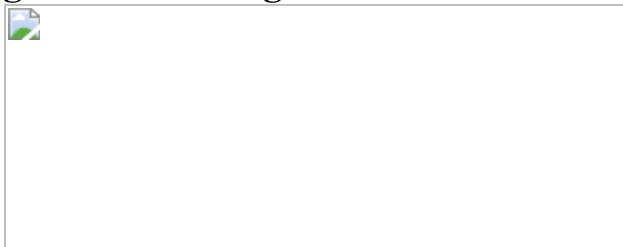


Fig. 2: Patterns of dissimilarity on the aboveground and fine-root planes.

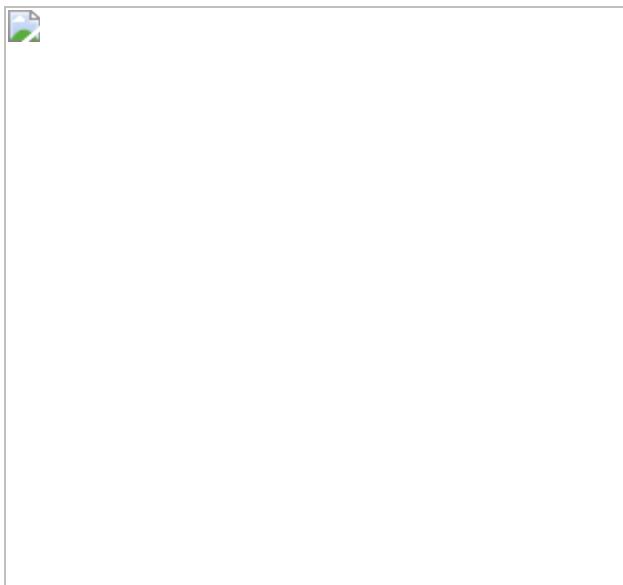


Fig. 3: Patterns of redundancy on the aboveground and fine-root planes.



Data availability

The datasets generated and analysed during the current study are available in the Figshare repository: <https://doi.org/10.6084/m9.figshare.13140146>.

Code availability

The R code used in the current study is available in the Figshare repository: <https://doi.org/10.6084/m9.figshare.13140146>.

References

1. 1.

Grime, J. P. *Plant Strategies, Vegetation Processes, and Ecosystem Properties* (John Wiley and Sons, 2001).

2. 2.

Reich, P. B. et al. The evolution of plant functional variation: traits, spectra, and strategies. *Int. J. Plant Sci.* **164**, S143–S164 (2003).

3. 3.

Díaz, S. et al. The global spectrum of plant form and function. *Nature* **529**, 167–171 (2016).

4. 4.

Bergmann, J. et al. The fungal collaboration gradient dominates the root economics space in plants. *Sci. Adv.* **6**, eaba3756 (2020).

5. 5.

Wright, I. J. et al. The worldwide leaf economics spectrum. *Nature* **428**, 821–827 (2004).

6. 6.

Kattge, J. et al. TRY plant trait database — enhanced coverage and open access. *Glob. Chang. Biol.* **26**, 119–188 (2020).

7. 7.

Iversen, C. M. et al. A global Fine-Root Ecology Database to address below-ground challenges in plant ecology. *New Phytol.* **215**, 15–26 (2017).

8. 8.

Guerrero-Ramírez, N. R. et al. Global root traits (GRoT) database. *Glob. Ecol. Biogeogr.* **30**, 25–37 (2021).

9. 9.

McCormack, M. L. et al. Redefining fine roots improves understanding of below-ground contributions to terrestrial biosphere processes. *New Phytol.* **207**, 505–518 (2015).

10. 10.

Rasse, D. P., Rumpel, C. & Dignac, M. F. Is soil carbon mostly root carbon? Mechanisms for a specific stabilisation. *Plant Soil* **269**, 341–

356 (2005).

11. 11.

Eissenstat, D. M. Costs and benefits of constructing roots of small diameter. *J. Plant Nutr.* **15**, 763–782 (1992).

12. 12.

Freschet, G. T., Cornelissen, J. H. C., van Logtestijn, R. S. P. & Aerts, R. Evidence of the ‘plant economics spectrum’ in a subarctic flora. *J. Ecol.* **98**, 362–373 (2010).

13. 13.

Reich, P. B. The world-wide ‘fast–slow’ plant economics spectrum: a traits manifesto. *J. Ecol.* **102**, 275–301 (2014).

14. 14.

Shen, Y. et al. Linking aboveground traits to root traits and local environment: implications of the plant economics spectrum. *Front. Plant Sci.* **10**, 1412 (2019).

15. 15.

Kramer-Walter, K. R. et al. Root traits are multidimensional: specific root length is independent from root tissue density and the plant economic spectrum. *J. Ecol.* **104**, 1299–1310 (2016).

16. 16.

Bergmann, J., Ryo, M., Prati, D., Hempel, S. & Rillig, M. C. Root traits are more than analogues of leaf traits: the case for diaspore mass. *New Phytol.* **216**, 1130–1139 (2017).

17. 17.

Weemstra, M. et al. Towards a multidimensional root trait framework: a tree root review. *New Phytol.* **211**, 1159–1169 (2016).

18. 18.

Ma, Z. et al. Evolutionary history resolves global organization of root functional traits. *Nature* **555**, 94–97 (2018).

19. 19.

de la Riva, E. G. et al. Root traits across environmental gradients in Mediterranean woody communities: are they aligned along the root economics spectrum? *Plant Soil* **424**, 35–48 (2018).

20. 20.

Craine, J. M., Lee, W. G., Bond, W. J., Williams, R. J. & Johnson, L. C. Environmental constraints on a global relationship among leaf and root traits of grasses. *Ecology* **86**, 12–19 (2005).

21. 21.

Liese, R., Alings, K. & Meier, I. C. Root branching is a leading root trait of the plant economics spectrum in temperate trees. *Front. Plant Sci.* **8**, 315 (2017).

22. 22.

Carmona, C. P. et al. Erosion of global functional diversity across the tree of life. *Sci. Adv.* **7**, eabf2675 (2021).

23. 23.

Niklas, K. J. Modelling below- and above-ground biomass for non-woody and woody plants. *Ann. Bot.* **95**, 315–321 (2005).

24. 24.

Liu, G. et al. Coordinated variation in leaf and root traits across multiple spatial scales in Chinese semi-arid and arid ecosystems. *New Phytol.* **188**, 543–553 (2010).

25. 25.

Galland, T., Carmona, C. P., Götzenberger, L., Valencia, E. & de Bello, F. Are redundancy indices redundant? An evaluation based on parameterized simulations. *Ecol. Indic.* **116**, 106488 (2020).

26. 26.

Valverde-Barrantes, O. J., Maherli, H., Baraloto, C. & Blackwood, C. B. Independent evolutionary changes in fine-root traits among main clades during the diversification of seed plants. *New Phytol.* **228**, 541–553 (2020).

27. 27.

Cornwell, W. K. et al. Plant species traits are the predominant control on litter decomposition rates within biomes worldwide. *Ecol. Lett.* **11**, 1065–1071 (2008).

28. 28.

Freschet, G. T. et al. Climate, soil and plant functional types as drivers of global fine-root trait variation. *J. Ecol.* **105**, 1182–1196 (2017).

29. 29.

De Deyn, G. B. & Van der Putten, W. H. Linking aboveground and belowground diversity. *Trends Ecol. Evol.* **20**, 625–633 (2005).

30. 30.

Pausas, J. G. & Bond, W. J. Humboldt and the reinvention of nature. *J. Ecol.* **107**, 1031–1037 (2019).

31. 31.

Poorter, H. et al. Biomass allocation to leaves, stems and roots: meta-analyses of interspecific variation and environmental control. *New Phytol.* **193**, 30–50 (2012).

32. 32.

Moora, M. Mycorrhizal traits and plant communities: perspectives for integration. *J. Veg. Sci.* **25**, 1126–1132 (2014).

33. 33.

Freschet, G. T. et al. Root traits as drivers of plant and ecosystem functioning: current understanding, pitfalls and future research needs. *New Phytol.* <https://doi.org/10.1111/nph.17072> (2021).

34. 34.

McCormack, M. L. & Iversen, C. M. Physical and functional constraints on viable belowground acquisition strategies. *Front. Plant Sci.* **10**, 1215 (2019).

35. 35.

Wells, C. E. & Eissenstat, D. M. Beyond the roots of young seedlings: the influence of age and order on fine root physiology. *J. Plant Growth Regul.* **21**, 324–334 (2002).

36. 36.

Zanne, A. E. et al. Three keys to the radiation of angiosperms into freezing environments. *Nature* **506**, 89–92 (2014).

37. 37.

USDA. USDA PLANTS Database (accessed 3rd July 2020);
<https://plants.sc.egov.usda.gov>

38. 38.

Engemann, K. et al. A plant growth form dataset for the New World. *Ecology* **97**, 3243 (2016).

39. 39.

BGCI. GlobalTreeSearch online database (accessed 3rd July 2020);
https://www.bgci.org/globaltree_search.php

40. 40.

The Plant List. The Plant List (accessed 17th February 2020);
<http://www.theplantlist.org>

41. 41.

Cayuela, L., Macarro, I., Stein, A. & Oksanen, J. Taxonstand: Taxonomic Standardization of Plant Species Names. R package version 2.2. <https://CRAN.R-project.org/package=Taxonstand> (2019).

42. 42.

Stekhoven, D. J. & Buhlmann, P. MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics* **28**, 112–118 (2012).

43. 43.

Oliveira, B. F., Sheffers, B. R. & Costa, G. C. Decoupled erosion of amphibians' phylogenetic and functional diversity due to extinction. *Glob. Ecol. Biogeogr.* **29**, 309–319 (2020).

44. 44.

Penone, C. et al. Imputation of missing data in life-history trait datasets: which approach performs the best? *Methods Ecol. Evol.* **5**, 961–970 (2014).

45. 45.

Jin, Y. & Qian, H. VPhyloMaker: an R package that can generate very large phylogenies for vascular plants. *Ecography* **42**, 1353–1359 (2019).

46. 46.

Smith, S. A. & Brown, J. W. Constructing a broadly inclusive seed plant phylogeny. *Am. J. Bot.* **105**, 302–314 (2018).

47. 47.

Whittaker, R. H. *Communities and Ecosystems* (Macmillan, 1975).

48. 48.

Stefan, V. & Levin, S. plotbiomes: Plot Whittaker biomes with ggplot2. R package version 0.0.0.9001
<https://github.com/valentinitnelav/plotbiomes> (2021).

49. 49.

Ricklefs, R. E. *The Economy of Nature* (W. H. Freeman and Company, 2008).

50. 50.

GBIF. GBIF Occurrence Download (accessed 15 December 2019);
<https://doi.org/10.15468/dl.thlxph>

51. 51.

South, A. rworldmap: a new R package for mapping global data. *R J.* **3**, 35–43 (2011).

52. 52.

Dinno, A. paran: Horn's Test of Principal Components/Factors. R package version 1.5.2. <https://CRAN.R-project.org/package=paran> (2018).

53. 53.

Dray, S. & Dufour, A.-B. The ade4 package: implementing the duality diagram for ecologists. *J. Stat. Softw.* <https://doi.org/10.18637/jss.v022.i04> (2007).

54. 54.

Duong, T. ks: kernel density estimation and kernel discriminant analysis for multivariate data in R. *J. Stat. Softw.* <https://doi.org/10.18637/jss.v021.i07> (2015).

55. 55.

Duong, T. ks: Kernel smoothing. R package version 1.11.5 <https://CRAN.R-project.org/package=ks> (2019).

56. 56.

Carmona, C. P., Bello, F., Mason, N. W. H. & Lepš, J. Trait probability density (TPD): measuring functional diversity across scales based on TPD with R. *Ecology* **100**, e02876 (2019).

57. 57.

Carmona, C. P. TPD: methods for measuring functional diversity based on Trait Probability Density. R package version 1.1.0. <https://CRAN.R-project.org/package=TPD> (2019).

58. 58.

Duong, T. & Hazelton, M. L. Plug-in bandwidth matrices for bivariate kernel density estimation. *J. Nonparametr. Stat.* **15**, 17–30 (2003).

59. 59.

Carmona, C. P., de Bello, F., Mason, N. W. H. & Lepš, J. Traits without borders: integrating functional diversity across scales. *Trends Ecol. Evol.* **31**, 382–394 (2016).

60. 60.
- Mason, N. W. H., Mouillot, D., Lee, W. G. & Wilson, J. B. Functional richness, functional evenness and functional divergence: the primary components of functional diversity. *Oikos* **111**, 112–118 (2005).
61. 61.
- Villéger, S., Mason, N. W. H. & Mouillot, D. New multidimensional functional diversity indices for a multifaceted framework in functional ecology. *Ecology* **89**, 2290–2301 (2008).
62. 62.
- Oksanen, J. et al. vegan: Community Ecology Package. R package version 2.5-5 <https://CRAN.R-project.org/package=vegan> (2019).
63. 63.
- Carmona, C. P. et al. Taxonomical and functional diversity turnover in Mediterranean grasslands: interactions between grazing, habitat type and rainfall. *J. Appl. Ecol.* **49**, 1084–1093 (2012).
64. 64.
- Micó, E. et al. Contrasting functional structure of saproxylic beetle assemblages associated to different microhabitats. *Sci. Rep.* **10**, 1520 (2020).
65. 65.
- Blonder, B. et al. New approaches for delineating n -dimensional hypervolumes. *Methods Ecol. Evol.* **9**, 305–319 (2018).
66. 66.
- Carmona, C. P., de Bello, F., Mason, N. W. H. & Lepš, J. The density awakens: a reply to Blonder. *Trends Ecol. Evol.* **31**, 667–669 (2016).

67. 67.

Mouillot, D. et al. Niche overlap estimates based on quantitative functional traits: a new family of non-parametric indices. *Oecologia* **145**, 345–353 (2005).

68. 68.

de Bello, F., Carmona, C. P., Mason, N. W. H., Sebastià, M.-T. & Lepš, J. Which trait dissimilarity for functional diversity: trait means or trait overlap? *J. Veg. Sci.* **24**, 807–819 (2013).

69. 69.

Traba, J., Iranzo, E. C., Carmona, C. P. & Malo, J. E. Realised niche changes in a native herbivore assemblage associated with the presence of livestock. *Oikos* **126**, 1400–1409 (2017).

70. 70.

Cornwell, W. K., Schwilk, D. W. & Ackerly, D. D. A trait-based test for habitat filtering: Convex Hull Volume. *Ecology* **87**, 1465–1471 (2006).

71. 71.

Blonder, B., Lamanna, C., Violle, C. & Enquist, B. J. The n -dimensional hypervolume. *Glob. Ecol. Biogeogr.* **23**, 595–609 (2014).

72. 72.

Blonder, B. Hypervolume concepts in niche- and trait-based ecology. *EcoGraphy* **41**, 1441–1455 (2018).

73. 73.

Ricotta, C. et al. Measuring the functional redundancy of biological communities: a quantitative guide. *Methods Ecol. Evol.* **7**, 1386–1395 (2016).

74. 74.

Mouillot, D. et al. Functional over-redundancy and high functional vulnerability in global fish faunas on tropical reefs. *Proc. Natl. Acad. Sci. USA* **111**, 13757–13762 (2014).

75. 75.

Carmona, C. P., de Bello, F., Sasaki, T., Uchida, K. & Pärtel, M. Towards a common toolbox for rarity: a response to Viole et al. *Trends Ecol. Evol.* **32**, 889–891 (2017).

76. 76.

Viole, C. et al. Functional rarity: the ecology of outliers. *Trends Ecol. Evol.* **32**, 356–367 (2017).

77. 77.

Borcard, D., Legendre, P. & Drapeau, P. Partialling out the spatial component of ecological variation. *Ecology* **73**, 1045–1055 (1992).

78. 78.

Gower, J. C. General coefficient of similarity and some of its properties. *Biometrics* **27**, 857–871 (1971).

79. 79.

Carmona, C. P. et al. Agriculture intensification reduces plant taxonomic and functional diversity across European arable systems. *Funct. Ecol.* **34**, 1448–1460 (2020).

80. 80.

Gherardi, L. A. & Sala, O. E. Global patterns and climatic controls of belowground net carbon fixation. *Proc. Natl Acad. Sci. USA* **117**, 20038–20043 (2020).

Acknowledgements

This study has been supported by the TRY initiative on plant traits (<http://www.try-db.org>). The TRY initiative and database is hosted, developed and maintained by J. Kattge and G. Boenisch (Max Planck Institute for Biogeochemistry, Jena, Germany). TRY is currently supported by Future Earth/bioDISCOVERY and the German Centre for Integrative Biodiversity Research (iDiv) Halle–Jena–Leipzig. This study was financed by the Estonian Ministry of Education and Research (PSG293 for C.P.C., C.G.B., S.T. and R.T.; PRG609 for M.P. and R.T.; PSG505 for A.T.; and PRG1065 for M.M., M.Z. and C.G.B.), by the European Union through the European Regional Development Fund (Centre of Excellence EcolChange), the University of Tartu (PLTOM20903). S.D. received partial funding from CONICET, FONCyT and Universidad Nacional de Córdoba (Argentina), IAI SGP and the Newton Fund. Partial funding for A.D.M. is from the Natural Sciences and Engineering Research Council of Canada.

Author information

Author notes

1. These authors contributed equally: Carlos P. Carmona, C. Guillermo Bueno, Aurele Toussaint, Sabrina Träger, Riin Tamme

Affiliations

1. Institute of Ecology and Earth Sciences, University of Tartu, Tartu, Estonia

Carlos P. Carmona, C. Guillermo Bueno, Aurele Toussaint, Sabrina Träger, Mari Moora, Meelis Pärtel, Martin Zobel & Riin Tamme

2. Institute of Biology/Geobotany and Botanical Garden, Martin Luther University Halle-Wittenberg, Halle (Saale), Germany

Sabrina Träger

3. Consejo Nacional de Investigaciones Científicas y Técnicas, Instituto Multidisciplinario de Biología Vegetal (IMBIV), CONICET, Universidad Nacional de Córdoba, Córdoba, Argentina

Sandra Díaz

4. Facultad de Ciencias Exactas, Físicas y Naturales, Universidad Nacional de Córdoba, Córdoba, Argentina

Sandra Díaz

5. Centre for Forest Research, Département des Sciences du bois et de la forêt, Université Laval, Quebec, Quebec, Canada

Alison D. Munson

Contributions

C.P.C., C.G.B., A.T., S.T., M.M. and R.T. conceived the study. R.T. and C.P.C. collected and processed the trait data. A.T. and S.T. collected and processed the biogeographical and climate data. C.P.C. and A.T. analysed the data with input from all authors. C.G.B., S.T. and R.T. performed a literature search. C.P.C. wrote a first draft of the manuscript, assisted by C.G.B., A.T., S.T. and R.T. S.D., M.M., A.M., M.P. and M.Z. contributed to the design of analyses, interpretation of results and article writing.

Corresponding author

Correspondence to Carlos P. Carmona.

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Ian Wright 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 Correlations between traits.

Pairwise correlations between the considered traits in the different datasets (black: full dataset with 1,719 species, blue: imputed dataset with 1,218 species, orange: complete dataset with 301 species). The lower-left triangle of the matrix contains scatterplots of traits (after log-10 transformation) showing the relationship (including regression lines) between each pair of traits. The diagonal includes a probability density function showing the distribution of each individual trait. The upper-right triangle includes the value of the correlation coefficients and, in the case of the full dataset, the number of species with empirical data for both traits (imputed and complete dataset always considered the same numbers of species). Lines for each dataset have different thickness to allow visualization of the correlation and probability density function even when there is high overlap between lines corresponding to different datasets.

Extended Data Fig. 2 Geographical, climatic and phylogenetic cover of the datasets.

a, Global map (Robinson projection) showing the occurrences (according to GBIF: <http://www.gbif.org>) of the species in the imputed dataset (1,218 species with empirical information for at least three aboveground and two fine-root traits). **b**, Number of species present in the major biomes⁴⁷ in the imputed dataset (in parentheses, number of species in the complete dataset). **c**, Distribution of species across the phylogeny of vascular plants excluding

ferns (Polypodiopsida) and lycopods (Lycopodiopsida) in the complete (301 species) and imputed datasets.

[Extended Data Fig. 3 Individual aboveground and fine-root functional spaces.](#)

Probabilistic distributions of the 2,630 and 748 species with complete empirical information for aboveground (**a**) and fine-root (**b**) traits in the functional spaces defined by a PCA on the corresponding traits followed by varimax rotation. The colour gradient (red-yellow-white) depicts different density of species in the defined space (red areas are more densely populated). Arrow length is proportional to the loadings of the considered traits in the resulting space. Aboveground traits are represented in green tones and fine-root traits in brown tones. Thick contour lines indicate the 0.5 and 0.99 quantiles, and thinner lines indicate quantiles 0.6, 0.7, 0.8 and 0.9.

[Extended Data Fig. 4 Functional space using the complete dataset.](#)

Probabilistic distributions of the 301 species with complete empirical information in the functional space defined by a PCA followed by varimax rotation based on both aboveground and fine-root traits. Each panel shows a combination of two of the four components that define the full plant spectrum. The colour gradient (red-yellow-white) depicts different density of species in the defined space (red areas are more densely populated). Arrow length is proportional to the loadings of the considered traits in the resulting space. Only those traits that had a loading of at least 0.3 in any of the represented components are shown to improve visualization (see loadings of all components in table S2). Aboveground traits are represented in green tones and fine-root traits in brown tones. Thick contour lines indicate the 0.5 and 0.99 quantiles, and thinner lines indicate quantiles 0.6, 0.7, 0.8 and 0.9.

[Extended Data Fig. 5 Comparison of the occupation of the functional space with multivariate normal distributions.](#)

Functional richness profile (amount of functional space occupied by quantiles of the functional spectra), difference in % of functional space occupied with respect to the null models, and functional divergence (representing the degree to which the density of species in the trait space is distributed towards the extremes of the distribution of species; right column) considering the first and second (**a**), third and fourth (**b**) and all components (**c**). In the functional profile plots (top of the left column in each case) green lines represent the mean, 2.5% and 97.5% quantiles of the functional richness profiles of null models ($n = 499$) representing multivariate normal distributions with equivalent parameters (means and standard deviations) than the observed data; orange lines represent the functional richness profile of the observed spectra. The values of functional richness for the 0.5 and 0.99 quantiles of all profiles are shown for comparison. The difference plots (bottom of the left column in each case), represent the percentage of functional space occupied by each quantile in relation to the mean of the null models; negative percentages mean that the considered quantile of the observed distribution occupies less space than the average of the null models, and vice versa. Each repetition of the null model ($n = 499$) is represented with a thin green line, whereas thicker green lines represent the mean, 2.5% and 97.5% quantiles of the 499 null models, and the orange line represents the observed distribution. The right column of each case represents the observed and null values of functional divergence; two-sided p values were estimated by confronting the value of the Standardize Effect Size (SES) with the cumulative normal distribution with mean = 0 and standard deviation = 1. The centre, bounds of box, and whiskers of the boxplot represent the median, 25th and 75th percentiles, and 1.5 times the interquartile range, respectively.

[Extended Data Fig. 6 Functional space using the imputed dataset.](#)

Probabilistic distributions of the 1,218 species with information for at least three aboveground and two fine-root traits (imputed dataset) in the functional space defined by a PCA followed by varimax rotation based on both aboveground and fine-root traits of the subset of species with complete empirical information. Each panel shows a combination of two of the four components that define the full plant spectrum. The colour gradient (red-

yellow-white) depicts different density of species in the defined space (red areas are more densely populated). Arrow lengths are proportional to the loadings of the considered traits in the resulting space. Only those traits that had a loading of at least 0.4 in any of the represented components are shown to improve visualization (see loadings of all components in Extended Data Table 2). Aboveground traits are represented in green tones and fine-root traits in brown tones. Thick contour lines indicate the 0.5 and 0.99 quantiles, and thinner lines indicate quantiles 0.6, 0.7, 0.8 and 0.9.

Extended Data Table 1 Traits considered in the study

Extended Data Table 2 Functional spaces considering different datasets

Extended Data Table 3 Angle between eigenvectors in the (non-rotated) PCA based on the complete dataset

Extended Data Table 4 Functional redundancy patterns

Supplementary information

[**Supplementary Information**](#)

This file contains Supplementary Application 1, Supplementary Note 1, Supplementary Methods 1, Supplementary Methods 2 and Supplementary references.

[**Reporting Summary**](#)

[**Peer Review File**](#)

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Carmona, C.P., Bueno, C.G., Toussaint, A. *et al.* Fine-root traits in the global spectrum of plant form and function. *Nature* **597**, 683–687 (2021). <https://doi.org/10.1038/s41586-021-03871-y>

- Received: 01 October 2020
- Accepted: 30 July 2021
- Published: 29 September 2021
- Issue Date: 30 September 2021
- DOI: <https://doi.org/10.1038/s41586-021-03871-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/s41586-021-03871-y>

- Article
- [Published: 08 September 2021](#)

Pollinators contribute to the maintenance of flowering plant diversity

- [Na Wei](#) [ORCID: orcid.org/0000-0002-7345-501X^{1,2}](#),
- [Rainee L. Kaczorowski](#) [ORCID: orcid.org/0000-0002-6414-7533¹](#),
- [Gerardo Arceo-Gómez^{1,3}](#),
- [Elizabeth M. O'Neill¹](#),
- [Rebecca A. Hayes¹](#) &
- [Tia-Lynn Ashman](#) [ORCID: orcid.org/0000-0002-9884-5954¹](#)

[Nature](#) volume **597**, pages 688–692 (2021)

- 3988 Accesses
- 137 Altmetric
- [Metrics details](#)

Subjects

- [Biodiversity](#)
- [Plant ecology](#)

Abstract

Mechanisms that favour rare species are key to the maintenance of diverse communities^{1,2,3}. One of the most critical tasks for conservation of

flowering plant biodiversity is to understand how plant–pollinator interactions contribute to the maintenance of rare species^{4,5,6,7}. Here we show that niche partitioning in pollinator use and asymmetric facilitation confer fitness advantage of rarer species in a biodiversity hotspot using phylogenetic structural equation modelling that integrates plant–pollinator and interspecific pollen transfer networks with floral functional traits. Co-flowering species filtered pollinators via floral traits, and rarer species showed greater pollinator specialization leading to higher pollination-mediated male and female fitness than more abundant species. When plants shared pollinator resources, asymmetric facilitation via pollen transport dynamics benefitted the rarer species at the cost of more abundant species, serving as an alternative diversity-promoting mechanism. Our results emphasize the importance of community-wide plant–pollinator interactions that affect reproduction for biodiversity maintenance.

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](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

Fig. 1: Schematic framework of how pollinator niche partitioning, asymmetric facilitation and automatic assurances confer rare species advantage.

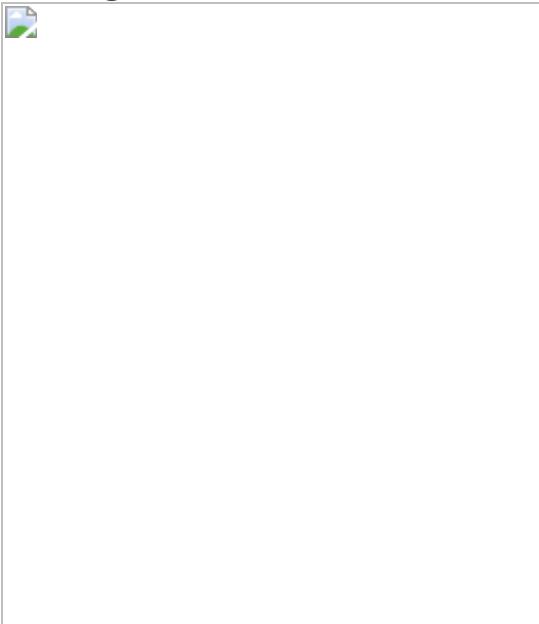
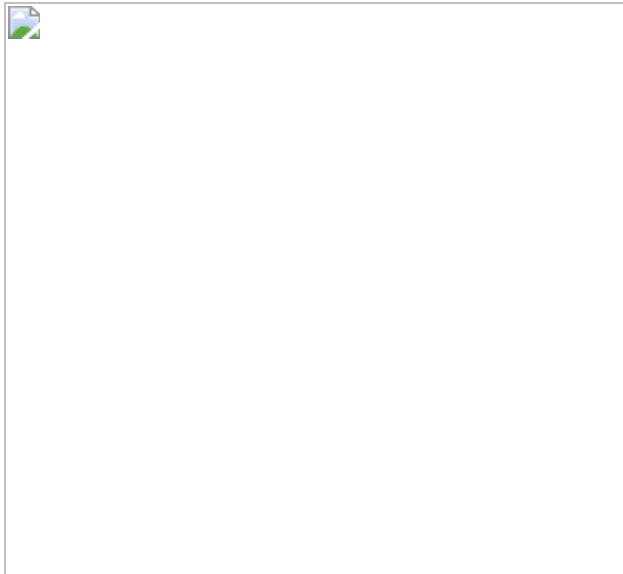


Fig. 2: Plant–pollinator interactions favour rare species through pollinator niche partitioning and asymmetric facilitation, which are mediated by floral traits.



Data availability

All data that support the findings of this study are included in this published article and its [Supplementary Information](#) files and source data files. [Source data](#) are provided with this paper.

Code availability

All software used in this study are provided in the Methods, [Supplementary Information](#) and the accompanying Reporting Summary.

References

1. 1.

Hubbell, S. P. *The Unified Neutral Theory of Biodiversity and Biogeography* (Princeton Univ. Press, 2001).

2. 2.

Wills, C. et al. Nonrandom processes maintain diversity in tropical forests. *Science* **311**, 527–531 (2006).

3. 3.

Chesson, P. Mechanisms of maintenance of species diversity. *Annu. Rev. Ecol. Evol. Syst.* **31**, 343–366 (2000).

4. 4.

Ollerton, J. Pollinator diversity: distribution, ecological function, and conservation. *Annu. Rev. Ecol. Evol. Syst.* **48**, 353–376 (2017).

5. 5.

Vamosi, J. C. et al. Pollination decays in biodiversity hotspots. *Proc. Natl Acad. Sci. USA* **103**, 956–961 (2006).

6. 6.

Bennett, J. M. et al. Land use and pollinator dependency drives global patterns of pollen limitation in the Anthropocene. *Nat. Commun.* **11**, 3999 (2020).

7. 7.

Potts, S. G. et al. Global pollinator declines: trends, impacts and drivers. *Trends Ecol. Evol.* **25**, 345–353 (2010).

8. 8.

Vamosi, J. C., Magallon, S., Mayrose, I., Otto, S. P. & Sauquet, H. Macroevolutionary patterns of flowering plant speciation and extinction. *Annu. Rev. Plant Biol.* **69**, 685–706 (2018).

9. 9.

Ollerton, J., Winfree, R. & Tarrant, S. How many flowering plants are pollinated by animals? *Oikos* **120**, 321–326 (2011).

10. 10.

Rodger, J. G. et al. 2021 Widespread vulnerability of plant seed production to pollinator decline. *Sci. Adv.* (in the press).

11. 11.

Pimm, S. L., Jones, H. L. & Diamond, J. On the risk of extinction. *Am. Nat.* **132**, 757–785 (1988).

12. 12.

Sargent, R. D. & Ackerly, D. D. Plant–pollinator interactions and the assembly of plant communities. *Trends Ecol. Evol.* **23**, 123–130 (2008).

13. 13.

Benadi, G. & Pauw, A. Frequency dependence of pollinator visitation rates suggests that pollination niches can allow plant species coexistence. *J. Ecol.* **106**, 1892–1901 (2018).

14. 14.

Bruno, J. F., Stachowicz, J. J. & Bertness, M. D. Inclusion of facilitation into ecological theory. *Trends Ecol. Evol.* **18**, 119–125 (2003).

15. 15.

Benadi, G., Bluthgen, N., Hovestadt, T. & Poethke, H. J. Population dynamics of plant and pollinator communities: stability reconsidered. *Am. Nat.* **179**, 157–168 (2012).

16. 16.

Moeller, D. A. Facilitative interactions among plants via shared pollinators. *Ecology* **85**, 3289–3301 (2004).

17. 17.

Bergamo, P. J., Susin Streher, N., Traveset, A., Wolowski, M. & Sazima, M. Pollination outcomes reveal negative density-dependence coupled with interspecific facilitation among plants. *Ecol. Lett.* **23**, 129–139 (2020).

18. 18.

Barrett, S. C. H. The evolution of plant sexual diversity. *Nat. Rev. Genet.* **3**, 274–284 (2002).

19. 19.

Ashman, T. L. & Arceo-Gómez, G. Toward a predictive understanding of the fitness costs of heterospecific pollen receipt and its importance in co-flowering communities. *Am. J. Bot.* **100**, 1061–1070 (2013).

20. 20.

Moreira-Hernández, J. I. & Muchhalá, N. Importance of pollinator-mediated interspecific pollen transfer for angiosperm evolution. *Annu. Rev. Ecol. Evol. Syst.* **50**, 191–217 (2019).

21. 21.

Ashman, T. L. et al. Pollen limitation of plant reproduction: ecological and evolutionary causes and consequences. *Ecology* **85**, 2408–2421 (2004).

22. 22.

Tur, C., Saez, A., Traveset, A. & Aizen, M. A. Evaluating the effects of pollinator-mediated interactions using pollen transfer networks: evidence of widespread facilitation in south Andean plant communities. *Ecol. Lett.* **19**, 576–586 (2016).

23. 23.

Levin, D. A. & Anderson, W. W. Competition for pollinators between simultaneously flowering species. *Am. Nat.* **104**, 455–467 (1970).

24. 24.

Ashman, T. L., Alonso, C., Parra-Tabla, V. & Arceo-Gómez, G. Pollen on stigmas as proxies of pollinator competition and facilitation: complexities, caveats and future directions. *Ann. Bot.* **125**, 1003–1012 (2020).

25. 25.

Lloyd, D. G. Some reproductive factors affecting the selection of self-fertilization in plants. *Am. Nat.* **113**, 67–79 (1979).

26. 26.

Sargent, R. D. & Otto, S. P. The role of local species abundance in the evolution of pollinator attraction in flowering plants. *Am. Nat.* **167**, 67–80 (2006).

27. 27.

Adler, P. B., Fajardo, A., Kleinhesselink, A. R. & Kraft, N. J. B. Trait-based tests of coexistence mechanisms. *Ecol. Lett.* **16**, 1294–1306 (2013).

28. 28.

Armbruster, W. S. The specialization continuum in pollination systems: diversity of concepts and implications for ecology, evolution and conservation. *Funct. Ecol.* **31**, 88–100 (2017).

29. 29.

Minnaar, C., Anderson, B., de Jager, M. L. & Karron, J. D. Plant–pollinator interactions along the pathway to paternity. *Ann. Bot.* **123**, 225–245 (2019).

30. 30.

Kantsa, A. et al. Disentangling the role of floral sensory stimuli in pollination networks. *Nat. Commun.* **9**, 1041 (2018).

31. 31.

Fang, Q. & Huang, S. Q. A directed network analysis of heterospecific pollen transfer in a biodiverse community. *Ecology* **94**, 1176–1185 (2013).

32. 32.

Baldwin, B. G. Origins of plant diversity in the California floristic province. *Annu. Rev. Ecol. Evol. Syst.* **45**, 347–369 (2014).

33. 33.

Bascompte, J., Jordano, P., Melian, C. J. & Olesen, J. M. The nested assembly of plant–animal mutualistic networks. *Proc. Natl Acad. Sci. USA* **100**, 9383–9387 (2003).

34. 34.

Thomson, J. D., Fung, H. F. & Ogilvie, J. E. Effects of spatial patterning of co-flowering plant species on pollination quantity and purity. *Ann. Bot.* **123**, 303–310 (2019).

35. 35.

Rezende, E. L., Lavabre, J. E., Guimaraes, P. R., Jordano, P. & Bascompte, J. Non-random coextinctions in phylogenetically structured mutualistic networks. *Nature* **448**, 925–928 (2007).

36. 36.

Song, C. L., Rohr, R. P. & Saavedra, S. Why are some plant–pollinator networks more nested than others? *J. Anim. Ecol.* **86**, 1417–1424 (2017).

37. 37.

Hegland, S. J., Nielsen, A., Lazaro, A., Bjerknes, A. L. & Totland, O. How does climate warming affect plant–pollinator interactions? *Ecol. Lett.* **12**, 184–195 (2009).

38. 38.

Ohlemuller, R. et al. The coincidence of climatic and species rarity: high risk to small-range species from climate change. *Biol. Lett.* **4**, 568–572 (2008).

39. 39.

Arceo-Gómez, G., Kaczorowski, R. L. & Ashman, T.-L. A network approach to understanding patterns of coflowering in diverse communities. *Int. J. Plant Sci.* **179**, 569–582 (2018).

40. 40.

Koski, M. H. et al. Plant–flower visitor networks in a serpentine metacommunity: assessing traits associated with keystone plant species. *Arthropod Plant Interact.* **9**, 9–21 (2015).

41. 41.

Arceo-Gómez, G. et al. Patterns of among- and within-species variation in heterospecific pollen receipt: the importance of ecological generalization. *Am. J. Bot.* **103**, 396–407 (2016).

42. 42.

Chao, A. et al. Rarefaction and extrapolation with Hill numbers: a framework for sampling and estimation in species diversity studies. *Ecol. Monogr.* **84**, 45–67 (2014).

43. 43.

R Core Team. *R: A Language and Environment for Statistical Computing*, <https://www.R-project.org/> (R Foundation for Statistical Computing, 2019).

44. 44.

Arceo-Gómez, G., Alonso, C., Ashman, T.-L. & Parra-Tabla, V. Variation in sampling effort affects the observed richness of plant–plant interactions via heterospecific pollen transfer: implications for interpretation of pollen transfer networks. *Am. J. Bot.* **105**, 1601–1608 (2018).

45. 45.

Hayes, R. A., Cullen N., Kaczorowski R. L., O’Neill E. M. & Ashman T-L. A community-wide description and key of pollen from co-flowering plants of the serpentine seeps of McLaughlin Reserve. *Madrono* (in the press).

46. 46.

Dafni, A. *Pollination Ecology: a Practical Approach* (Oxford Univ. Press, 1992).

47. 47.

McMurdie, P. J. & Holmes, S. Waste NOT, want not: why rarefying microbiome data is inadmissible. *PLoS Comp. Biol.* **10**, e1003531 (2014).

48. 48.

Qian, H. & Jin, Y. An updated megaphylogeny of plants, a tool for generating plant phylogenies and an analysis of phylogenetic community structure. *J. Plant Ecol.* **9**, 233–239 (2016).

49. 49.

Zanne, A. E. et al. Three keys to the radiation of angiosperms into freezing environments. *Nature* **506**, 89–92 (2014).

50. 50.

Hinchliff, C. E. et al. Synthesis of phylogeny and taxonomy into a comprehensive tree of life. *Proc. Natl Acad. Sci. USA* **112**, 12764–12769 (2015).

51. 51.

Paradis, E. & Schliep, K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* **35**, 526–528 (2019).

52. 52.

Michonneau, F., Brown, J. W. & Winter, D. J. rotl: an R package to interact with the Open Tree of Life data. *Methods Ecol. Evol.* **7**, 1476–1481 (2016).

53. 53.

Revell, L. J. phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol. Evol.* **3**, 217–223 (2012).

54. 54.

Le, S., Josse, J. & Husson, F. FactoMineR: an R package for multivariate analysis. *J. Stat. Softw.* **25**, 1–18 (2008).

55. 55.

Dormann, C. F., Gruber, B. & Fruend, J. Introducing the bipartite package: analysing ecological networks. *R News* **8**, 8–11 (2008).

56. 56.

Feinsinger, P., Spears, E. E. & Poole, R. W. A simple measure of niche breadth. *Ecology* **62**, 27–32 (1981).

57. 57.

Horn, H. S. Measurement of "overlap" in comparative ecological studies. *Am. Nat.* **100**, 419–424 (1966).

58. 58.

Almeida-Neto, M., Guimaraes, P., Guimaraes, P. R., Loyola, R. D. & Ulrich, W. A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement. *Oikos* **117**, 1227–1239 (2008).

59. 59.

Csardi, G. & Nepusz, T. The igraph software package for complex network research. *InterJournal* **1695**, 1–9 (2006).

60. 60.

Patefield, W. Algorithm AS 159: an efficient method of generating random $R \times C$ tables with given row and column totals. *Appl. Stat.* **30**, 91–97 (1981).

61. 61.

Oksanen, J. et al. vegan: Community Ecology Package. R package version 2.5–5, <https://CRAN.R-project.org/package=vegan> (2019).

62. 62.

Wickham, H. *ggplot2: Elegant Graphics for Data Analysis* (Springer, 2016).

63. 63.

Bastian, M., Heymann, S. & Jacomy, M. Gephi: an open source software for exploring and manipulating networks. Presented at the Third international AAAI Conference on Weblogs and Social Media (2009).

64. 64.

Arceo-Gómez, G., Kaczorowski, R. L., Patel, C. & Ashman, T. L. Interactive effects between donor and recipient species mediate fitness costs of heterospecific pollen receipt in a co-flowering community. *Oecologia* **189**, 1041–1047 (2019).

65. 65.

Keck, F., Rimet, F., Bouchez, A. & Franc, A. phylosignal: an R package to measure, test, and explore the phylogenetic signal. *Ecol. Evol.* **6**, 2774–2780 (2016).

66. 66.

Orme, D. et al. caper: comparative analyses of phylogenetics and evolution in R. R package version 1.0.1, <https://CRAN.R-project.org/package=caper> (2018).

67. 67.

Barrett, S. C. H. The evolution of plant sexual diversity. *Nat. Rev. Genet.* **3**, 274–284 (2002).

68. 68.

Fort, H., Vazquez, D. P. & Lan, B. L. Abundance and generalisation in mutualistic networks: solving the chicken-and-egg dilemma. *Ecol. Lett.* **19**, 4–11 (2016).

69. 69.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D. & R Core Team. nlme: linear and nonlinear mixed effects models. R package version 3.1-143, <https://CRAN.R-project.org/package=nlme> (2019).

70. 70.

Lefcheck, J. S. & Freckleton, R. piecewiseSEM: piecewise structural equation modelling in R for ecology, evolution, and systematics. *Methods Ecol. Evol.* **7**, 573–579 (2015).

71. 71.

Fox, J. & Weisberg, S. *An R companion to Applied Regression*, 3rd edition (Sage, 2019).

72. 72.

Blüthgen, N., Menzel, F. & Blüthgen, N. Measuring specialization in species interaction networks. *BMC Ecol.* **6**, 9 (2006).

73. 73.

Shipley, B. The AIC model selection method applied to path analytic models compared using a d-separation test. *Ecology* **94**, 560–564 (2013).

74. 74.

van der Bijl, W. phylopath: easy phylogenetic path analysis in R. *PeerJ* **6**, e4718 (2018).

Acknowledgements

We thank J. Baker, D. Chang, A. M. Arters, U. Meenakshinathan, K. Doleski, S. Barratt-Boyes, R. A. Ashman and M. Holden for assistance with stigma pollen identification, floral trait measurements and insect specimen processing; J. Rawlins, J. Pawelek, R. Androw and B. Coulter for insect identification; J. Hyland and V. Verdecia for logistic support at the Carnegie Museum of Natural History; McLaughlin field station staff for logistical support of field work; and the members of the Ashman, Wood and Turcotte laboratories for discussion. This work was supported by the National Science Foundation (DEB 1452386) to T.-L.A.

Author information

Affiliations

1. Department of Biological Sciences, University of Pittsburgh,
Pittsburgh, PA, USA

Na Wei, Rainee L. Kaczorowski, Gerardo Arceo-Gómez, Elizabeth M. O'Neill, Rebecca A. Hayes & Tia-Lynn Ashman

2. The Holden Arboretum, Kirtland, OH, USA

Na Wei

3. Department of Biological Sciences, East Tennessee State University,
Johnson City, TN, USA

Gerardo Arceo-Gómez

Contributions

T.-L.A. conceived the study. N.W. and T.-L.A. led the conceptual development. N.W. analysed the data. N.W., T.-L.A. and R.L.K. wrote the manuscript. N.W., T.-L.A., R.L.K. and G.A.-G. contributed to manuscript revisions. R.L.K., E.M.O., R.A.H., G.A.-G. and T.-L.A. collected the data.

Corresponding authors

Correspondence to Na Wei or Tia-Lynn Ashman.

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Marcelo Aizen, Christopher Kaiser-Bunbury 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 Community-wide plant–pollinator network.](#)

a, Plant species ($n = 79$) coloured by families are arranged on the left according to phylogeny. The numbers of pollinator species that plants interacted with are shown as black bars and numbers within parentheses. **b**, Pollinator species ($n = 416$) are arranged along the top according to the size and similarity of plant assemblages that they interacted with. **c**, The observed numbers of interactions are denoted as frequency ('Freq') by the colour scale

[Source data.](#)

[Extended Data Fig. 2 Rarefaction shows that the majority of pollinator diversity was captured with our sampling intensity.](#)

Rarefaction curves of each of the 79 plant species (**a, b**) that were observed for plant–pollinator interactions (Supplementary Table 1) and the 64 plant species (**c, d**) that were included in the phylogenetic structural equation models (Fig. 2). The observed number of pollinators is represented by the solid portion of each coloured line, whereas the dashed portion indicates extrapolation in the rarefaction analysis using the R package iNEXT⁴². Lines colours are randomly assigned. Pollinator diversity, especially Chao's Shannon diversity (**b, d**), started to level off at the observed number of pollinators for most plant species, reflecting sufficient sampling to capture pollinator diversity

[Source data.](#)

Extended Data Fig. 3 Floral trait variation and abundance.

a, Multivariate analysis (factor analysis of mixed data, FAMD) of 20 floral traits (Supplementary Table 3). Plant species ($n = 73$, abbreviated as the first two letters of genus and species names and coloured by plant family) are segregated along the first two dimensions, representing mainly size-related and other (shape/colour/inflorescence) floral traits, respectively. These traits vary independently from species floral abundance (symbol size). **b**, Species rarity based on floral abundance (log-transformed) was correlated with rarity based on occurrence in the number of surveyed plots (see [Methods](#), two-sided Pearson's correlation test, $r = 0.64$, $t = 6.9$, d.f. = 70, $P = 1.8 \times 10^{-9}$). The 95% confidence intervals of the mean are shown

[Source data](#).

Extended Data Fig. 4 Multivariate analysis of floral traits associated with pollinator attraction.

a, b, In the first four dimensions of the factor analysis of mixed data (FAMD), the centroid of each category within a qualitative trait is indicated, with symbol shape representing different qualitative traits. Quantitative traits are represented by arrows. Individual plant species ($n = 73$) are shown in the background with colours indicating plant families and symbol sizes indicating floral abundances (Extended Data Fig. 3). **c**, The traits that contributed to $\geq 15\%$ of variation of the first three dimensions are highlighted in colour

[Source data](#).

Extended Data Fig. 5 Multivariate analysis of floral traits associated with male organ.

a, b, In the first four dimensions of the factor analysis of mixed data (FAMD), the centroid of each category within a qualitative trait is indicated, with symbol shape representing different qualitative traits. Quantitative traits are represented by arrows. Individual plant species ($n = 73$) are shown

in the background with colours indicating plant families and symbol sizes indicating floral abundances (Extended Data Fig. 3). c, The traits that contributed to $\geq 15\%$ of variation of the first three dimensions are highlighted in colour

[Source data.](#)

Extended Data Fig. 6 Multivariate analysis of floral traits associated with female organ.

a, b, In the first four dimensions of the factor analysis of mixed data (FAMD), the centroid of each category within a qualitative trait is indicated, with symbol shape representing different qualitative traits. Quantitative traits are represented by arrows. Individual plant species ($n = 73$) are shown in the background with colours indicating plant families and symbol sizes indicating floral abundances (Extended Data Fig. 3). c, The traits that contributed to $\geq 15\%$ of variation of the first four dimensions are highlighted in colour

[Source data.](#)

Extended Data Fig. 7 Pollen transfer network.

The network was constructed based on pollen deposited on 54 stigmas of 66 individual plant species (Supplementary Table 5). Plant species (nodes) are abbreviated as the first two letters of genus and species names (Supplementary Table 3), with unidentified species denoted with ‘U’. Node size indicates the number of flowering plant species that pollen is received from, and node colour darkness indicates the number of flowering plant species that pollen is donated to. That is, larger nodes represent better recipients and darker nodes better donors. Arrows and their sizes indicate the direction and amount (counts) of pollen transfer, respectively

[Source data.](#)

Extended Data Fig. 8 Validation of fractional identity approach and rarefaction of pollen received by stigmas.

a, There was a strong relationship between heterospecific pollen (HP) richness when fractionally identified pollen grains were excluded (*y*-axis, ‘no ambiguity’) and included (*x*-axis, ‘fractional’): $n = 66$ plant species, general linear model, slope = 0.73, $t = 17.1$, $P < 2 \times 10^{-9}$. The dotted 95% confidence intervals of the mean are shown. **b, c**, Rarefaction analysis using the R package iNEXT⁴² showed that the majority of pollen species richness (**b**) and Chao’s Shannon diversity (**c**) were captured by the sampled styles ($n = 54$ on average) for each plant species ($n = 66$, coloured lines). The observed (solid) and extrapolated (dashed) portion of each rarefaction line are indicated

[Source data.](#)

[Extended Data Fig. 9 Phylogenetic structural equation models \(PSEMs\).](#)

a–h, The PSEMs considered pollinator niche partitioning, asymmetric facilitation, pollination assurance and numeric assurance (orange arrows). Pollination assurance and numeric assurance are collectively referred to as automatic assurances. **i**, Model fitting and nested model selection used the R packages piecewiseSEM⁷⁰ and phylopath⁷⁴. Sample size was 64 plant species. df and P , degree of freedom and P value of the two-sided Fisher’s C statistic; AIC, the Akaike’s information criterion; CICc, the C statistic information criterion corrected for small sample sizes; w , CICc weights. Standardized regression coefficients of paths and model averaging are in Supplementary Table 6.

Extended Data Table 1 Serpentine seep system within the McLaughlin Natural Reserve study area

Supplementary information

[Supplementary Information](#)

This file contains Supplementary Methods, Supplementary Notes and Supplementary References.

Reporting Summary

Supplementary Table 1

Plant–pollinator interactions observed across sites and years.

Supplementary Table 2

Plant–pollinator network metrics and comparisons to null models.

Supplementary Table 3

Species-level floral traits.

Supplementary Table 4

Phylogenetic signals of plant–pollinator interactions, interspecific pollen transfer network metrics and floral traits.

Supplementary Table 5

Stigma pollen data.

Supplementary Table 6

Standardized regression coefficients and model averaging of phylogenetic structural equation models.

Source data

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. 5**](#)

[**Source Data Extended Data Fig. 6**](#)

[**Source Data Extended Data Fig. 7**](#)

[**Source Data Extended Data Fig. 8**](#)

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Wei, N., Kaczorowski, R.L., Arceo-Gómez, G. *et al.* Pollinators contribute to the maintenance of flowering plant diversity. *Nature* **597**, 688–692 (2021). <https://doi.org/10.1038/s41586-021-03890-9>

- Received: 17 September 2020
- Accepted: 09 August 2021
- Published: 08 September 2021
- Issue Date: 30 September 2021
- DOI: <https://doi.org/10.1038/s41586-021-03890-9>

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

Pollination advantage of rare plants unveiled

- Marcelo A. Aizen

News & Views 08 Sept 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03890-9>

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

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

Genetic and epigenetic coordination of cortical interneuron development

- [Kathryn C. Allaway](#)^{1,2,3} na1,
- [Mariano I. Gabitto](#)⁴ na1,
- [Orly Wapinski](#)³ na1,
- [Giuseppe Saldi](#) [ORCID: orcid.org/0000-0002-9994-7354](#)^{2,3,5},
- [Chen-Yu Wang](#) [ORCID: orcid.org/0000-0002-3912-5113](#)^{2,3},
- [Rachel C. Bandler](#)^{1,2,3},
- [Sherry Jingjing Wu](#) [ORCID: orcid.org/0000-0001-9625-1661](#)^{2,3},
- [Richard Bonneau](#) [ORCID: orcid.org/0000-0003-4354-7906](#)^{4,5,6} &
- [Gord Fishell](#) [ORCID: orcid.org/0000-0002-9640-9278](#)^{2,3}

[Nature](#) volume 597, pages 693–697 (2021)

- 4865 Accesses
- 88 Altmetric
- [Metrics details](#)

Subjects

- [Development of the nervous system](#)
- [Genetics of the nervous system](#)

Abstract

One of the hallmarks of the cerebral cortex is the extreme diversity of interneurons^{1,2,3}. The two largest subtypes of cortical interneurons, parvalbumin- and somatostatin-positive cells, are morphologically and functionally distinct in adulthood but arise from common lineages within the medial ganglionic eminence^{4,5,6,7,8,9,10,11}. This makes them an attractive model for studying the generation of cell diversity. Here we examine how developmental changes in transcription and chromatin structure enable these cells to acquire distinct identities in the mouse cortex. Generic interneuron features are first detected upon cell cycle exit through the opening of chromatin at distal elements. By constructing cell-type-specific gene regulatory networks, we observed that parvalbumin- and somatostatin-positive cells initiate distinct programs upon settling within the cortex. We used these networks to model the differential transcriptional requirement of a shared regulator, *Mef2c*, and confirmed the accuracy of our predictions through experimental loss-of-function experiments. We therefore reveal how a common molecular program diverges to enable these neuronal subtypes to acquire highly specialized properties by adulthood. Our methods provide a framework for examining the emergence of cellular diversity, as well as for quantifying and predicting the effect of candidate genes on cell-type-specific development.

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](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

Fig. 1: Distal chromatin elements provide the earliest markers of interneuron identity.

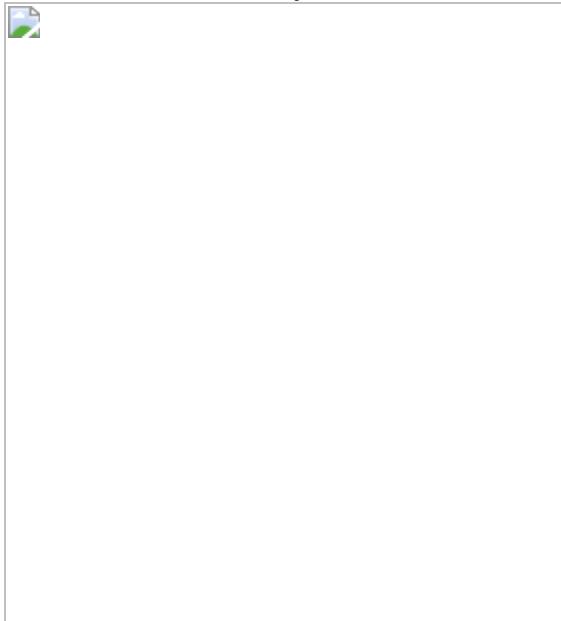


Fig. 2: Remodelling of interneuron chromatin architecture during migration and post-settling.

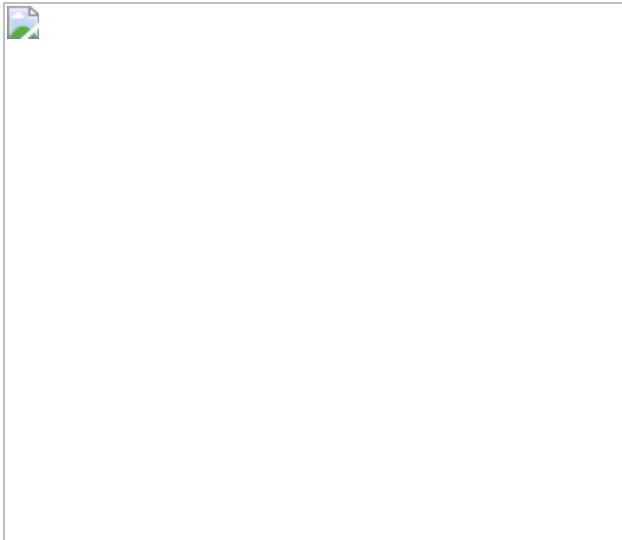


Fig. 3: The maturation of gene networks is characterized by the emergence of cell-type-specific regulatory interactions.

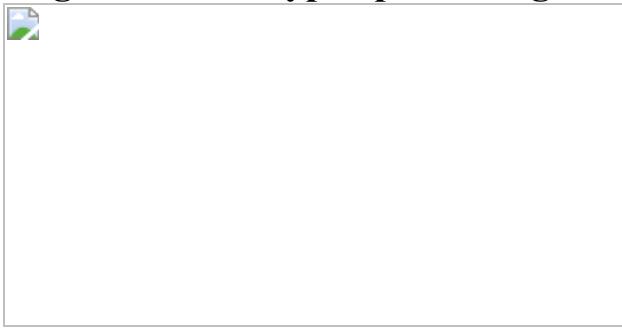
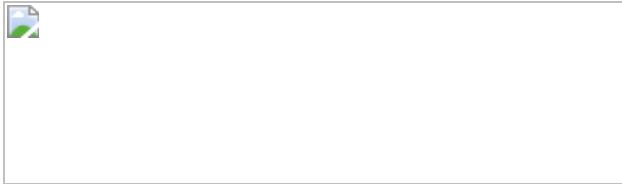


Fig. 4: Loss of *Mef2c* disproportionately affects the gene regulatory landscape of PV cells.



Data Availability

All sequencing data produced for this study are available at the Gene Expression Omnibus (GEO) at accession number [GSE165233](#).

Code Availability

The ChromA package, used for peak calling in all scATAC-seq datasets, is available at <https://github.com/marianogabitto/ChromA>. The CellRanger and CellRanger ATAC pipelines, used for preprocessing of scRNA-seq and scATAC-seq libraries, respectively, are available at <https://support.10xgenomics.com/single-cell-gene-expression/software/downloads/latest> and <https://support.10xgenomics.com/single-cell-atac/software/downloads/latest>. The snapATAC package, used for downstream analysis of scATAC-seq datasets, is available at <https://github.com/r3fang/SnapATAC>. The Seurat package, used for scRNA-seq analysis and integration of scRNA-seq and scATAC-seq datasets, is available at <https://github.com/satijalab/seurat>. Scripts for analysis are available at <https://github.com/marianogabitto/ChromatinDynamics2020-Analysis>.

References

1. 1.
Ascoli, G. A. et al. Petilla terminology: Nomenclature of features of GABAergic interneurons of the cerebral cortex. *Nat. Rev. Neurosci.* **9**, 557–568 (2008).
2. 2.
Defelipe, J. et al. New insights into the classification and nomenclature of cortical GABAergic interneurons. *Nat. Rev. Neurosci.* **14**, 202–216 (2013).
3. 3.
Kepcs, A. & Fishell, G. Interneuron cell types are fit to function. *Nature* **505**, 318–326 (2014).
4. 4.

Mayer, C. et al. Clonally related forebrain interneurons disperse broadly across both functional areas and structural boundaries. *Neuron* **87**, 989–998 (2015).

5. 5.

Harwell, C. C. et al. Wide dispersion and diversity of clonally related inhibitory interneurons. *Neuron* **87**, 999–1007 (2015).

6. 6.

Freund, T. F. Interneuron diversity series: rhythm and mood in perisomatic inhibition. *Trends Neurosci.* **26**, 489–495 (2003).

7. 7.

Tremblay, R., Lee, S. & Rudy, B. GABAergic interneurons in the neocortex: from cellular properties to circuits. *Neuron* **91**, 260–292 (2016).

8. 8.

Mayer, C. et al. Developmental diversification of cortical inhibitory interneurons. *Nature* **555**, 457–462 (2018).

9. 9.

Mi, D. et al. Early emergence of cortical interneuron diversity in the mouse embryo. *Science* **360**, 81–85 (2018).

10. 10.

Wonders, C. P. et al. A spatial bias for the origins of interneuron subgroups within the medial ganglionic eminence. *Dev. Biol.* **314**, 127–136 (2008).

11. 11.

Batista-Brito, R. & Fishell, G. The developmental integration of cortical interneurons into a functional network. *Curr. Top. Dev. Biol.* **87**, 81–118 (2009).

12. 12.

Wang, Y. et al. *Dlx5* and *Dlx6* regulate the development of parvalbumin-expressing cortical interneurons. *J. Neurosci.* **30**, 5334–5345 (2010).

13. 13.

Pai, E. L.-L. et al. *Mafb* and *c-Maf* have prenatal compensatory and postnatal antagonistic roles in cortical interneuron fate and function. *Cell Rep.* **26**, 1157–1173.e5 (2019).

14. 14.

Pai, E. L.-L. et al. *Maf* and *Mafb* control mouse pallial interneuron fate and maturation through neuropsychiatric disease gene regulation. *eLife* **9**, e54903 (2020).

15. 15.

Ma, S. et al. Chromatin potential identified by shared single-cell profiling of RNA and chromatin. *Cell* **183**, 1103–1116.e20 (2020).

16. 16.

Sinnamon, J. R. et al. The accessible chromatin landscape of the murine hippocampus at single-cell resolution. *Genome Res.* **29**, 857–869 (2019).

17. 17.

Preissl, S. et al. Single-nucleus analysis of accessible chromatin in developing mouse forebrain reveals cell-type-specific transcriptional regulation. *Nat. Neurosci.* **21**, 432–439 (2018).

18. 18.

Fang, R. et al. Comprehensive analysis of single cell ATAC-seq data with SnapATAC. *Nat. Commun.* **12**, 1337 (2021).

19. 19.

Stuart, T. et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902.e21 (2019).

20. 20.

Barkas, N. et al. Joint analysis of heterogeneous single-cell RNA-seq dataset collections. *Nat. Methods* **16**, 695–698 (2019).

21. 21.

Miyoshi, G. & Fishell, G. GABAergic interneuron lineages selectively sort into specific cortical layers during early postnatal development. *Cereb. Cortex* **21**, 845–852 (2011).

22. 22.

Batista-Brito, R. et al. The cell-intrinsic requirement of *Sox6* for cortical interneuron development. *Neuron* **63**, 466–481 (2009).

23. 23.

Azim, E., Jabaudon, D., Fame, R. M. & MacKlis, J. D. SOX6 controls dorsal progenitor identity and interneuron diversity during neocortical development. *Nat. Neurosci.* **12**, 1238–1247 (2009).

24. 24.

Au, E. et al. A modular gain-of-function approach to generate cortical interneuron subtypes from ES cells. *Neuron* **80**, 1145–1158 (2013).

25. 25.

Castro, D. M., de Veaux, N. R., Miraldi, E. R. & Bonneau, R. Multi-study inference of regulatory networks for more accurate models of gene regulation. *PLoS Comput. Biol.* **15**, 1–22 (2019).

26. 26.

Yadav, T., Quivy, J.-P. & Almouzni, G. Chromatin plasticity: a versatile landscape that underlies cell fate and identity. *Science* **361**, 1332–1336 (2018).

27. 27.

Ladstätter, S. & Tachibana, K. Genomic insights into chromatin reprogramming to totipotency in embryos. *J. Cell Biol.* **218**, 70–82 (2019).

28. 28.

Harrington, A. J. et al. MEF2C regulates cortical inhibitory and excitatory synapses and behaviors relevant to neurodevelopmental disorders. *eLife* **5**, e20059 (2016).

29. 29.

Cosgrove, D. et al. Genes influenced by MEF2C contribute to neurodevelopmental disease via gene expression changes that affect multiple types of cortical excitatory neurons. *Hum. Mol. Genet.* **30**, 961–970 (2021).

30. 30.

Monory, K. et al. The endocannabinoid system controls key epileptogenic circuits in the hippocampus. *Neuron* **51**, 455–466 (2006).

31. 31.

Madisen, L. et al. A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat. Neurosci.* **13**,

133–140 (2010).

32. 32.

Mo, A. et al. Epigenomic signatures of neuronal diversity in the mammalian brain. *Neuron* **86**, 1369–1384 (2015).

33. 33.

Vong, L. H., Ragusa, M. J. & Schwarz, J. J. Generation of conditional *Mef2c*^{loxP/loxP} mice for temporal- and tissue-specific analyses. *Genesis* **43**, 43–48 (2005).

34. 34.

Fogarty, M. et al. Spatial genetic patterning of the embryonic neuroepithelium generates GABAergic interneuron diversity in the adult cortex. *J. Neurosci.* **27**, 10935–10946 (2007).

35. 35.

Tirosh, I. et al. Single-cell RNA-seq supports a developmental hierarchy in human oligodendrogloma. *Nature* **539**, 309–313 (2016).

36. 36.

Setty, M. et al. Characterization of cell fate probabilities in single-cell data with Palantir. *Nat. Biotechnol.* **37**, 451–460 (2019).

37. 37.

Street, K. et al. Slingshot: cell lineage and pseudotime inference for single-cell transcriptomics. *BMC Genomics* **19**, 477 (2018).

38. 38.

Gabitto, M. I. et al. Characterizing chromatin landscape from aggregate and single-cell genomic assays using flexible duration modeling. *Nat. Commun.* **11**, 1–10 (2020).

39. 39.

Heinz, S. et al. Simple combinations of lineage-determining transcription factors prime *cis*-regulatory elements required for macrophage and B cell identities. *Mol. Cell* **38**, 576–589 (2010).

40. 40.

Stroud, H. et al. An activity-mediated transition in transcription in early postnatal neurons. *Neuron* **107**, 874–890.e8 (2020).

41. 41.

Liu, N. et al. Direct promoter repression by BCL11A controls the fetal to adult hemoglobin switch. *Cell* **173**, 430–442.e17 (2018).

42. 42.

Zhu, Q., Liu, N., Orkin, S. H. & Yuan, G.-C. CUT&RUNTools: a flexible pipeline for CUT&RUN processing and footprint analysis. *Genome Biol.* **20**, 192 (2019).

43. 43.

Pliner, H. et al. Chromatin accessibility dynamics of myogenesis at single cell resolution. Preprint at <https://doi.org/10.1101/155473> (2017).

44. 44.

Jackson, C. & Skok Gibbs, C. Inferelator example data and scripts. <https://doi.org/10.5281/zenodo.3355524> (2019).

Acknowledgements

We thank J. Dimidschstein, E. Favuzzi, D. Ginty, E. Macosko and E. Miraldi for feedback and discussion; N. Yusuf and M. Fernandez-Otero for technical assistance; R. Batista-Brito for providing mice for the *Mef2c* loss-

of-function experiments; and C. Davis for sharing the CUT&RUN protocol and reagents. G.F. and K.C.A. acknowledge support from the following sources: National Institutes of Health (NIH) 1UG3MH120096-01 (G.F.), NIH 5R01MH071679-15 (G.F.) and NIH F31NS103398 (K.C.A.). R.B. and M.I.G. acknowledge support from the following sources: NIH R01DK103358, Simons Foundation, NSF-IOS 1546218, NIH R35GM122515, NSF-CBET 1728858 and NIH R01AI130945.

Author information

Author notes

1. These authors contributed equally: Kathryn C. Allaway, Mariano I. Gabitto, Orly Wapinski

Affiliations

1. Neuroscience Institute, New York University, New York, NY, USA

Kathryn C. Allaway & Rachel C. Bandler

2. Department of Neurobiology, Harvard Medical School, Boston, MA, USA

Kathryn C. Allaway, Giuseppe Saldi, Chen-Yu Wang, Rachel C. Bandler, Sherry Jingjing Wu & Gord Fishell

3. Stanley Center for Psychiatric Research, Broad Institute, Cambridge, MA, USA

Kathryn C. Allaway, Orly Wapinski, Giuseppe Saldi, Chen-Yu Wang, Rachel C. Bandler, Sherry Jingjing Wu & Gord Fishell

4. Center for Computational Biology, Flatiron Institute, Simons Foundation, New York, NY, USA

Mariano I. Gabitto & Richard Bonneau

5. Department of Biology, New York University, New York, NY, USA

Giuseppe Saldi & Richard Bonneau

6. Center for Data Science, New York University, New York, NY, USA

Richard Bonneau

Contributions

K.C.A., M.I.G. and O.W. contributed equally to this work and are featured in alphabetical order in the author list. K.C.A., M.I.G., O.W., R.B. and G.F. conceived the original project. K.C.A. led experimental work, with assistance from O.W., C.-Y.W., R.C.B. and S.J.W., and supervision from G.F. M.I.G. led computational analysis, with assistance from G.S. and K.C.A., and supervision from R.B. K.C.A., M.I.G., G.S., R.B. and G.F. interpreted the data and wrote the manuscript, with input from O.W. and C.-Y.W.

Corresponding authors

Correspondence to Richard Bonneau or Gord Fishell.

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

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

Extended data figures and tables

Extended Data Fig. 1 Quality control of embryonic scRNA-seq samples and mitotic versus postmitotic discrimination.

a, Number of cells in Dlx6a– and Dlx6a+ scRNA-seq datasets collected from E13 MGE in Dlx6a-Cre;Ai9 mice and multiome dataset in E13 MGE wild type mice. **b**, Mean reads per cell in Dlx6a–, Dlx6a+ and multiome scRNA-seq datasets. **c**, Median genes detected per cell in Dlx6a–, Dlx6a+ and multiome scRNA-seq datasets. **d**, Fraction of cells scored to be in G2/M or S phase of the cycle cycle at each maturation score. Blue line indicates cells from the Dlx6– dataset, red line indicates cells from the Dlx6+ dataset. **e**, Diffusion map of E13 Dlx6a– and Dlx6a+ MGE scRNA-seq data color-coded by assignment to a mitotic (red) or postmitotic (blue) state. **f**, Percentage of cycling cells as a function of the position along the maturation trajectory color-coded by assignment to a mitotic (red) or postmitotic (blue) state in E13 Dlx6a– and Dlx6a+ MGE scRNA-seq datasets. **g**, Diffusion map of E13 multiome MGE scRNA-seq data color-coded by assignment to a mitotic (red) or postmitotic (blue) state. **h**, Percentage of cycling cells as a function of the position along the maturation trajectory color-coded by assignment to a mitotic (red) or postmitotic (blue) state in E13 multiome MGE scRNA-seq dataset. **i**, Line plots indicating the promoter accessibility (blue) and gene expression (red) for six developmentally-regulated genes.

Extended Data Fig. 2 Developmental characterization of embryonic E13 MGE cells surveyed using scRNA-seq, scATAC-seq, and multiomic methods.

a, Analysis of Dlx6a–, Dlx6a+ and multiome scRNA-seq datasets collected from E13 MGE in Dlx6a-Cre;Ai9 mice and multiome dataset in E13 MGE wild type mice. **b**, Analysis of Dlx6a–, Dlx6a+ and multiome scATAC-seq datasets collected from E13 MGE in Dlx6a-Cre;Ai9 mice and multiome

dataset in E13 MGE wild type mice. **i**, Dlx6 $^{+/-}$ FAC-sorted and multiome cells. **ii**, Unbiased cluster annotation. **iii**, Dlx6 $^{+}$ and Dlx6 $^{-}$ annotation. **iv**, Cell cycling phase annotation. **v**, Pseudotime annotation. **vi**, Mitotic and postmitotic cell annotation. In **ii-vi**, Annotations are performed on scRNA-seq datasets and transferred to scATAC-seq datasets through the multiome dataset. scRNA- and scATAC-seq low dimensional representation reflects UMAP embedding. **vii**, Average gene expression and promoter accessibility for unbiased clusters.

Extended Data Fig. 3 Orthogonal maturation trajectory methods reveal branching fates in postmitotic MGE cells at E13.

a, Palantir Analysis of Analysis of Dlx6a $^{-}$, Dlx6a $^{+}$ and multiome scRNA-seq postmitotic cells. **b**, UMAP Analysis of Analysis of Dlx6a $^{-}$, Dlx6a $^{+}$ and multiome scRNA-seq postmitotic cells. **c**, Diffusion Maps Analysis of Analysis of Dlx6a $^{-}$, Dlx6a $^{+}$ and multiome scRNA-seq postmitotic cells. For **a**, **b**, **c**:**i**, Unbiased clustering annotation. **ii**, Branch trajectories color-coded by slingshot pseudotime. **iii**, Gene expression for mitotic marker (Fabp7) and branch-specific marker (Zic1, Meis2, Maf). **iv**, For Palantir, cells are color coded by Palantir Pseudotime or Differentiation Potential. **d**, Confusion matrix reveals agreement of branch labels between different trajectory methods. **e**, TFs with binding motif highly enriched in a branch specific manner. **f**, Average gene expression across branches for TFs in **e**.

Extended Data Fig. 4 Chromatin accessibility precedes gene expression in branch 1- specific genes.

a, Heat map depicting gene expression, Promoter and distal elements accessibility throughout branch 1 pseudotime for loci \pm 500kb around six branch 1 - specific genes. Distal elements are selected based on the relevance for classifying branch 1 cells. For each gene, row 1 shows gene expression, row 2 promoter accessibility, row 3 aggregated accessibility and the remaining rows are distal branch 1 classifying elements. Each trace has been smoothed using the lowess function in R. **b**, Gene expression, Promoter and Aggregated accessibility throughout branch 1 pseudotime for

10 branch 1 - specific genes. **c**, Only postmitotic cells are classified as branch neurons by supervised classification methods. **c-i)** Classification of E13 cells into mitotic or postmitotic cells based on cell cycle RNA score. **c-ii)** Classification of E13 cells into mitotic or branch 1,2,3 lineages based on chromatin accessibility.

Extended Data Fig. 5 Analysis of MGE-derived cortical interneuron scRNA-seq datasets from E18 through P28.

a-d, UMAP showing individual scRNA-seq datasets from Dlx6a+ labeled cortical neurons, subsetted for PV+ and SST+ interneurons, collected at E18 (**a**), P2 (**b**), P10 (**c**), and P28 (**d**). Left UMAP in each panel is color-coded by cluster identity (individually determined for each dataset). Right UMAP in each panel is color-coded by cardinal class. Bottom right UMAP in (**c**) and (**d**) is color-coded by cortical region of origin (anterior lateral motor cortex - ALM, primary visual cortex - V1). In (**d**), the right-most boxes show each cluster linked to its subtype identity, determined by expression of marker genes. **e**, Integration of E18, P2, P10, and P28 scRNA-seq datasets using Seurat CCA. Left UMAP is color-coded by timepoint, right is color-coded by cardinal class. **f**, Prediction score indicating the confidence of label transfer between each timepoint. Cluster labels were transferred from P2 to E18 (top), from P10 to P2 (middle) and from P28 and P10 (bottom). Label transfer prediction scores were lowest between P10 and P2 timepoints.

Extended Data Fig. 6 Analysis of MGE-derived interneuron scATAC-seq datasets from E18 through P28.

a-d, UMAP showing individual scATAC-seq datasets from Dlx6a+ labeled cortical neurons, subsetted for PV+ and SST+ interneurons, collected at E18 (**a**), P2 (**b**), P10 (**c**), and P28 (**d**). Left UMAP in each panel is color-coded by cluster identity (individually determined for each dataset). Top right of each panel shows gene body accessibility of Gad2 (interneuron marker gene), Lhx6 (MGE-derived interneuron marker gene), Tac1 or PV (PV cIN marker gene), and Sst (SST cIN marker gene). Bottom right UMAP in each panel is color-coded by cardinal class, determined by

accessibility of cardinal-class specific loci (identified at P28) and confirmed by marker gene accessibility. See [methods](#) for more information on cardinal class assignment in ATAC-seq data. Additional bottom right UMAP in **(c)** and **(d)** is color-coded by cortical region of origin (anterior lateral motor cortex - ALM, primary visual cortex - V1). **e**, Integration of E18, P2, P10, and P28 scATAC-seq datasets using Seurat CCA. Left UMAP is color-coded by timepoint, right is color-coded by cardinal class. **f**, Prediction score indicating the confidence of label transfer between each timepoint. Cluster labels were transferred from P2 to E18 (top), from P10 to P2 (middle) and from P28 and P10 (bottom). Label transfer prediction scores progressively increased over developmental time.

Extended Data Fig. 7 Integration of MGE-derived interneuron scRNA-and scATAC-seq datasets.

a–d, UMAPs showing scRNA- and scATAC-seq datasets integrated using Seurat CCA at E18 **(a)**, P2 **(b)**, P10 **(c)**, and P28 **(d)**. Top panel shows UMAP color-coded by RNA-seq cluster identity - i.e., the labels used for label transfer and calculation of prediction scores in Figure [2a–d](#). Bottom panel UMAP is color-coded by cardinal class identity. **e–h**, UMAPs showing scRNA- and scATAC-seq datasets integrated using CONOS E18 **(e)**, P2 **(f)**, P10 **(g)**, and P28 **(h)**. Top panel shows UMAP color-coded by dataset of origin (RNA or ATAC). Bottom panel UMAP is color-coded by cardinal class identity. Cells from RNA and ATAC datasets integrated relatively well at later developmental timepoints, but were mostly segregated at earlier timepoints (E18, P2).

Extended Data Fig. 8 Distally located loci open in specific cardinal classes earlier in development than proximal elements.

a, Average signal (top) and signal at each locus (bottom) for peaks specifically called for PV cells at P28 that are within promoters or gene bodies (i.e., proximal elements). Promoter regions were defined as TSS+2 kb upstream. Signal within PV and SST cells at P28, P10, P2, and E18 are shown over a window 5 kb up- and down-stream of peak center (compare to Figure [2f](#), which zooms in to a 1 kb +/- window). **b**, Average signal (top)

and signal at each locus (bottom) for peaks specifically called for PV cells at P28 that are distally located at each timepoint. **c**, Average signal (top) and signal at each locus (bottom) for peaks specifically called for SST cells at P28 that are within promoters or gene bodies (proximal) at each timepoint. **d**, Average signal (top) and signal at each locus (bottom) for peaks specifically called for SST cells at P28 that are distally located at each timepoint. **e**, Top five enriched motifs for P28 PV-specific proximal peaks. **f**, Top five enriched motifs for P28 PV-specific distal peaks. **g**, Top five enriched motifs for P28 SST-specific proximal peaks. **h**, Top five enriched motifs for P28 SST-specific distal peaks.

Extended Data Fig. 9 MEF2C CUT&RUN in PV and SST interneurons at P28.

a, Number of peaks called in each CUT&RUN replicate from PV (blue) and SST (red) cells, and the number of peaks present in both replicates from each cell type (intersect). **b**, Percentage of peaks containing a canonical Mef2c motif in each replicate. **c**, Significance (-log p-value) of enrichment of Mef2c motif in each replicate. **d**, The distribution of CUT&RUN peaks that were also present in P28 ATAC peak sets containing MEF2C motifs. For this analysis, the CUT&RUN replicate intersect peak sets of PV and SST were subsetted for those peaks that were also present in the ATAC peaks called for each cell type at P28 and found to contain a canonical MEF2C motif. Of those, they were categorized for those that were present in both PV and SST ATAC peak sets or unique to one cell type. **e**, Compares peaks identified in CUT&RUN (orange) or jointly in CUT & RUN and ATAC-seq analyses (green). ATAC-seq peaks used in this analysis were those identified in each cell type at P28 regardless of presence or absence of a canonical Mef2c motif. Compare with Fig. 2j which performs the same analysis but only includes ATAC peaks with a Mef2c motif. **f**, **g**, HOMER results for de novo motif discovery in replicate intersect peak sets for PV and SST cINs. **h**, Genomic location of each CUT&RUN peak in replicate intersect peak sets for PV and SST cINs. **i**, Integrative Genomics Viewer (IGV) snapshots showing bigwig files for each CUT&RUN replicate and associated peaks called in both replicates (intersect) for PV and SST cINs. Genomic loci shows are examples of genes with nearby peaks for commonly expressed genes (Rbfox1, Grin2a) and PV-enriched

genes (Erbb4, Pthlh, Plxcd3). **j** and **k**, Gene Ontology (GO) term analysis. Each CUT&peak was associated with the nearest gene TSS. These lists of genes for PV and SST cells were then used as input for GO term analysis, revealing an enrichment of genes associated with synapse development.

Extended Data Fig. 10 Single-cell RNA- and ATAC-seq analysis of Mef2c cKO interneurons compared to wild-type (WT) cells at P2.

a, UMAP of P2 WT snRNA-seq and P2 *Mef2c* cKO snRNA-seq data integrated using Seurat and color-coded by cluster identity. WT dataset here was prepared using single nuclei to match cKO rather than the whole cell dataset (see Fig. 4). **b**, UMAP in (a) segregated according to timepoint and color-coded by cell type. **c**, Cluster composition delineated by (i) marker gene expression (ii) cell type (iii) cell number. Compare a–c here with Fig. 4 a–c — this figure contains single nucleus data for both WT and cKO while Fig 4 WT data is whole cell. **d**, UMAP of Mef2c cKO and WT cells color-coded by cluster. **e**, UMAP of Mef2c cKO and WT cells color-coded by genotype. **f**, UMAP of Mef2c cKO and WT cells color-coded by cardinal class identity. Mef2c cKO identity was determined by the accessibility of marker genes (see **d**) and assignment of clusters to the appropriate cardinal class. **g**, Gene body accessibility of SST and PV cIN marker genes (SST: Sst, Grin3a, Elfn1, Cacng3, Grm1, Satb1, Tmem91. PV: Tac1, Erbb4). **h**, Pie chart representation of scATAC-seq data showing the total number of peaks in E18 WT, P2 WT, and P2 Mef2c cKO, subdivided into peaks that are PV or SST cell specific or shared across both cell types. Compare to Figure 4d, which shows similar pie charts but only for peaks with Mef2c motifs.

Supplementary information

Supplementary Information

This file contains a full guide for Supplementary Tables 1–4 (Supplementary Tables supplied as separate Excel files).

Reporting Summary

Supplementary Table 1

Quality control metrics and statistics for scATAC- and scRNA-seq collected for this study – see Supplementary Information document for full description.

Supplementary Table 2

Genomic coordinates of distal elements associated branch 1 genes in E13 postmitotic precursors – see Supplementary Information document for full description.

Supplementary Table 3

Motif enrichment analysis of scATAC-seq data from PV and SST cells across development – see Supplementary Information document for full description.

Supplementary Table 4

Predicted direct and second-order *Mef2c* targets and DE genes from *Mef2c* LOF experiments – see Supplementary Information document for full description.

Peer Review File

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Allaway, K.C., Gabitto, M.I., Wapinski, O. *et al.* Genetic and epigenetic coordination of cortical interneuron development. *Nature* **597**, 693–697 (2021). <https://doi.org/10.1038/s41586-021-03933-1>

- Received: 06 April 2020
- Accepted: 18 August 2021
- Published: 22 September 2021
- Issue Date: 30 September 2021
- DOI: <https://doi.org/10.1038/s41586-021-03933-1>

Share this article

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

[Get shareable link](#)

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

[Copy to clipboard](#)

Provided by the Springer Nature SharedIt content-sharing initiative

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03933-1>

- Article
- [Published: 15 September 2021](#)

Rational design of a new antibiotic class for drug-resistant infections

- [Thomas F. Durand-Reville](#)¹,
- [Alita A. Miller](#) ORCID: orcid.org/0000-0002-7110-0493¹,
- [John P. O'Donnell](#)¹,
- [Xiaoyun Wu](#)²,
- [Mark A. Sylvester](#)¹,
- [Satenig Guler](#)^{1,3},
- [Ramkumar Iyer](#)¹,
- [Adam B. Shapiro](#) ORCID: orcid.org/0000-0002-2662-6146¹,
- [Nicole M. Carter](#)¹,
- [Camilo Velez-Vega](#)^{1,2},
- [Samir H. Moussa](#)¹,
- [Sarah M. McLeod](#)¹,
- [April Chen](#)¹,
- [Angela M. Tanudra](#)¹,
- [Jing Zhang](#)^{1,4},
- [Janelle Comita-Prevoir](#)¹,
- [Jan A. Romero](#)^{1,5},
- [Hoan Huynh](#)⁶,
- [Andrew D. Ferguson](#)⁷,
- [Peter S. Horanyi](#) ORCID: orcid.org/0000-0003-2218-9986⁸,
- [Stephen J. Mayclin](#)⁸,
- [Henry S. Heine](#)⁹,
- [George L. Drusano](#)⁹,
- [Jason E. Cummings](#)¹⁰,

- [Richard A. Slayden¹⁰](#) &
- [Ruben A. Tommasi](#) ORCID: [orcid.org/0000-0002-1573-7422¹](https://orcid.org/0000-0002-1573-7422)

[*Nature*](#) volume **597**, pages 698–702 (2021)

- 8073 Accesses
- 347 Altmetric
- [Metrics details](#)

Subjects

- [Bacterial infection](#)
- [Drug discovery and development](#)

Abstract

The development of new antibiotics to treat infections caused by drug-resistant Gram-negative pathogens is of paramount importance as antibiotic resistance continues to increase worldwide¹. Here we describe a strategy for the rational design of diazabicyclooctane inhibitors of penicillin-binding proteins from Gram-negative bacteria to overcome multiple mechanisms of resistance, including β -lactamase enzymes, stringent response and outer membrane permeation. Diazabicyclooctane inhibitors retain activity in the presence of β -lactamases, the primary resistance mechanism associated with β -lactam therapy in Gram-negative bacteria^{2,3}. Although the target spectrum of an initial lead was successfully re-engineered to gain in vivo efficacy, its ability to permeate across bacterial outer membranes was insufficient for further development. Notably, the features that enhanced target potency were found to preclude compound uptake. An improved optimization strategy leveraged porin permeation properties concomitant with biochemical potency in the lead-optimization stage. This resulted in ETX0462, which has potent in vitro and in vivo activity against *Pseudomonas aeruginosa* plus all other Gram-negative ESKAPE pathogens, *Stenotrophomonas maltophilia* and biothreat pathogens. These attributes, along with a favourable preclinical safety

profile, hold promise for the successful clinical development of the first novel Gram-negative chemotype to treat life-threatening antibiotic-resistant infections in more than 25 years.

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](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

Fig. 1: Rational design of DBOs for multi-PBP inhibition.



Fig. 2: Distinct phenotypes for DBO analogues.

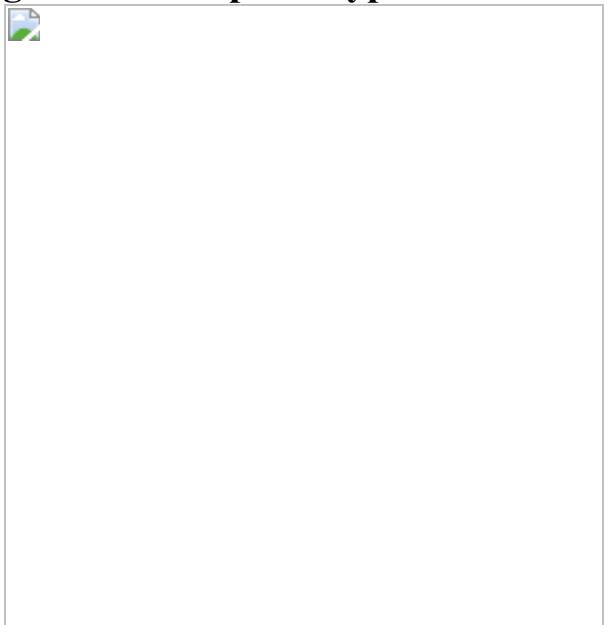


Fig. 3: Crystal structure of *Pa*PBP3–ETX0462.

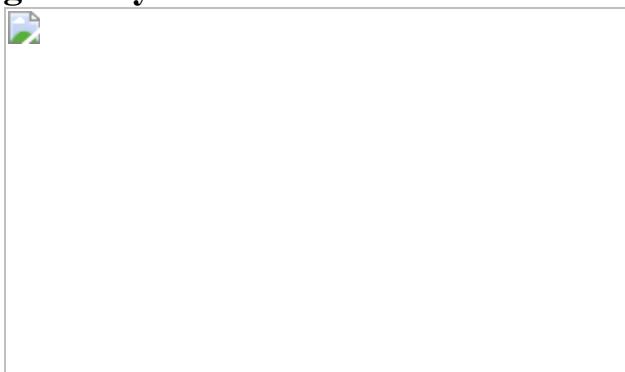
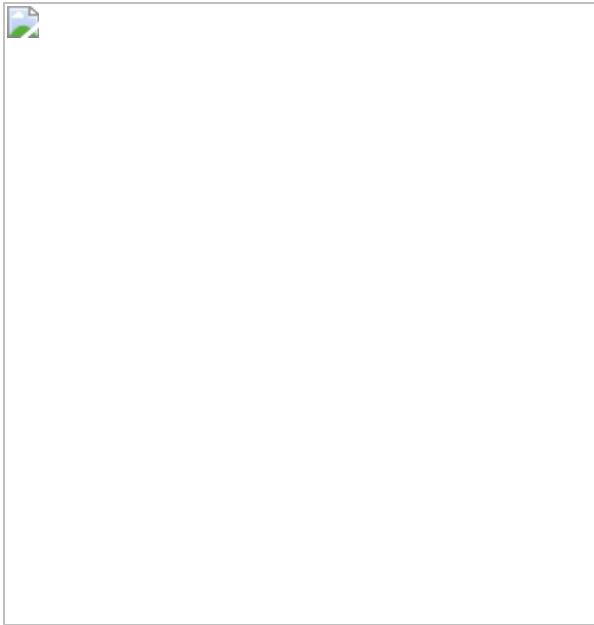


Fig. 4: In vivo efficacy of ETX0462.



Data availability

Atomic coordinates and structure factors of the *PaPBP3–ETX0462* structure have been deposited in the Protein Data Bank under accession code [7JWL](#). [Source data](#) are provided with this paper.

References

1. 1.

Tacconelli, E. et al. Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. *Lancet Infect. Dis.* **18**, 318–327 (2018).

2. 2.

Bonomo, R. A. Beta-lactamases: a focus on current challenges. *Cold Spring Harb. Perspect. Med.* **7**, a025239 (2016).

3. 3.

Bush, K. & Bradford, P. A. Interplay between β -lactamases and new β -lactamase inhibitors. *Nat. Rev. Microbiol.* **17**, 295–306 (2019).

4. 4.

Ehmann, D. E. et al. Avibactam is a covalent, reversible, non- β -lactam β -lactamase inhibitor. *Proc. Natl Acad. Sci. USA* **109**, 11663–11668 (2012).

5. 5.

Livermore, D. M., Mushtaq, S., Warner, M., Vickers, A. & Woodford, N. In vitro activity of cefepime/zidebactam (WCK 5222) against Gram-negative bacteria. *J. Antimicrob. Chemother.* **72**, 1373–1385 (2017).

6. 6.

Morinaka, A. et al. OP0595, a new diazabicyclooctane: mode of action as a serine β -lactamase inhibitor, antibiotic and β -lactam 'enhancer'. *J. Antimicrob. Chemother.* **70**, 2779–2786 (2015).

7. 7.

Durand-Reville, T. F. et al. ETX2514 is a broad-spectrum β -lactamase inhibitor for the treatment of drug-resistant Gram-negative bacteria including *Acinetobacter baumannii*. *Nat. Microbiol.* **2**, 17104 (2017).

8. 8.

Miller, A. A. et al. In vitro characterization of ETX1317, a broad-spectrum β -lactamase inhibitor that restores and enhances β -lactam activity against multi-drug-resistant *Enterobacteriales*, including carbapenem-resistant strains. *ACS Infect. Dis.* **6**, 1389–1397 (2020).

9. 9.

Wang, D. Y., Abboud, M. I., Markoulides, M. S., Brem, J. & Schofield, C. J. The road to avibactam: the first clinically useful non-

β -lactam working somewhat like a β -lactam. *Future Med. Chem.* **8**, 1063–1084 (2016).

10. 10.

Levy, N. et al. Structural basis for *E. coli* penicillin binding protein (PBP) 2 inhibition, a platform for drug design. *J. Med. Chem.* **62**, 4742–4754 (2019).

11. 11.

Kidd, J. M., Abdelraouf, K. & Nicolau, D. P. Efficacy of human-simulated bronchopulmonary exposures of cefepime, zidebactam and the combination (WCK 5222) against MDR *Pseudomonas aeruginosa* in a neutropenic murine pneumonia model. *J. Antimicrob. Chemother.* **75**, 149–155 (2020).

12. 12.

O'Donnell, J. et al. Pharmacokinetic/pharmacodynamic determination and preclinical pharmacokinetics of the β -lactamase inhibitor ETX1317 and its orally available prodrug ETX0282. *ACS Infect. Dis.* **6**, 1378–1388 (2020).

13. 13.

Doumith, M., Mushtaq, S., Livermore, D. M. & Woodford, N. New insights into the regulatory pathways associated with the activation of the stringent response in bacterial resistance to the PBP2-targeted antibiotics, mecillinam and OP0595/RG6080. *J. Antimicrob. Chemother.* **71**, 2810–2814 (2016).

14. 14.

Levasseur, P., Pace, J., Lee Coleman, K. & Lowther, J. *Novel Combinations of Antibacterial Nitrogenous Heterocyclic Compounds with Other Antibacterial Compounds, and Use Thereof as Drugs*. Patent Cooperation Treaty WO/2010/041112 (2010).

15. 15.

Aszodi, J., Lampilas, M., Musicki, B., Rowlands, D. A. & Colette, P. *Heterocyclic Compounds, Method for Preparing Same and Use Thereof as Medicines, in Particular as Antibacterial Agents*. Patent Cooperation Treaty WO/2002/100860 (2001).

16. 16.

Nonejuie, P., Burkart, M., Pogliano, K. & Pogliano, J. Bacterial cytological profiling rapidly identifies the cellular pathways targeted by antibacterial molecules. *Proc. Natl Acad. Sci. USA* **110**, 16169–16174 (2013).

17. 17.

Lomovskaya, O. et al. Identification and characterization of inhibitors of multidrug resistance efflux pumps in *Pseudomonas aeruginosa*: novel agents for combination therapy. *Antimicrob. Agents Chemother.* **45**, 105–116 (2001).

18. 18.

Viljanen, P. & Vaara, M. Susceptibility of Gram-negative bacteria to polymyxin B nonapeptide. *Antimicrob. Agents Chemother.* **25**, 701–705 (1984).

19. 19.

Iyer, R. et al. Whole-cell-based assay to evaluate structure permeation relationships for carbapenem passage through the *Pseudomonas aeruginosa* porin OprD. *ACS Infect. Dis.* **3**, 310–319 (2017).

20. 20.

Chevalier, S. et al. Structure, function and regulation of *Pseudomonas aeruginosa* porins. *FEMS Microbiol. Rev.* **41**, 698–722 (2017).

21. 21.

Hancock, R. E. & Brinkman, F. S. Function of *Pseudomonas* porins in uptake and efflux. *Ann. Rev. Microbiol.* **56**, 17–38 (2002).

22. 22.

Nikaido, H. Molecular basis of bacterial outer membrane permeability revisited. *Microbiol. Mol. Biol. Rev.* **67**, 593–656 (2003).

23. 23.

Kos, V. N. et al. The resistome of *Pseudomonas aeruginosa* in relationship to phenotypic susceptibility. *Antimicrob. Agents Chemother.* **59**, 427–436 (2015).

24. 24.

Dötsch, A. et al. The *Pseudomonas aeruginosa* transcriptome in planktonic cultures and static biofilms using RNA sequencing. *PLoS ONE* **7**, e31092 (2012).

25. 25.

Turner, K. H., Everett, J., Trivedi, U., Rumbaugh, K. P. & Whiteley, M. Requirements for *Pseudomonas aeruginosa* acute burn and chronic surgical wound infection. *PLoS Genet.* **10**, e1004518 (2014).

26. 26.

Richter, M. F. et al. Predictive compound accumulation rules yield a broad-spectrum antibiotic. *Nature* **545**, 299–304 (2017).

27. 27.

Eren, E. et al. Substrate specificity within a family of outer membrane carboxylate channels. *PLoS Biol.* **10**, e1001242 (2012).

28. 28.

Bajaj, H. et al. Bacterial outer membrane porins as electrostatic nanosieves: exploring transport rules of small polar molecules. *ACS Nano* **11**, 5465–5473 (2017).

29. 29.

Cadwell, J. J. S. The hollow fiber infection model for antimicrobial pharmacodynamics and pharmacokinetics. *Adv. Pharmacoepidem. Drug Safety* **S1**, 1–5 (2012).

30. 30.

Sastry, G. M., Adzhigirey, M., Day, T., Annabhimoju, R. & Sherman, W. Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comp. Aid. Mol. Des.* **27**, 221–234 (2013).

31. 31.

Hess, B., Kutzner, C., van der Spoel, D. & Lindahl, E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J. Chem. Theor. Comput.* **4**, 435–447 (2008).

32. 32.

Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).

33. 33.

Sindhikara, D. J., Yoshida, N. & Hirata, F. Placevent: an algorithm for prediction of explicit solvent atom distribution-application to HIV-1 protease and F-ATP synthase. *J. Comp. Chem.* **33**, 1536–1543 (2012).

34. 34.

Velez-Vega, C., McKay, D. J., Aravamuthan, V., Pearlstein, R. & Duca, J. S. Time-averaged distributions of solute and solvent motions:

exploring proton wires of GFP and PfM2DH. *J. Chem. Inform. Model.* **54**, 3344–3361 (2014).

35. 35.

M07: Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria that Grow Aerobically 11th edn (Clinical and Laboratory Standards Institute, 2018).

36. 36.

M26: Methods for Determining Bactericidal Activity 19th edn (Clinical and Laboratory Standards Institute, 1999).

37. 37.

Iyer, R., Moussa, S. H., Durand-Réville, T. F., Tommasi, R. & Miller, A. *Acinetobacter baumannii* OmpA is a selective antibiotic permeant porin. *ACS Infect. Dis.* **4**, 373–381 (2018).

38. 38.

Penwell, W. F. et al. Molecular mechanisms of sulbactam antibacterial activity and resistance determinants in *Acinetobacter baumannii*. *Antimicrob. Agents Chemother.* **59**, 1680–1689 (2015).

39. 39.

Shapiro, A. B., Gao, N., Gu, R. F. & Thresher, J. Fluorescence anisotropy-based measurement of *Pseudomonas aeruginosa* penicillin-binding protein 2 transpeptidase inhibitor acylation rate constants. *Anal. Biochem.* **463**, 15–22 (2014).

40. 40.

Shapiro, A. B., Gu, R. F., Gao, N., Livchak, S. & Thresher, J. Continuous fluorescence anisotropy-based assay of BOCILLIN FL penicillin reaction with penicillin binding protein 3. *Anal. Biochem.* **439**, 37–43 (2013).

41. 41.

Murphy-Benenato, K. E. et al. SAR and structural analysis of siderophore-conjugated monocarbam inhibitors of PBP3. *ACS Med. Chem. Lett.* **6**, 537–542 (2015).

42. 42.

Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).

43. 43.

Gerber, A. U. et al. Impact of dosing intervals on activity of gentamicin and ticarcillin against *Pseudomonas aeruginosa* in granulocytopenic mice. *J. Infect. Dis.* **147**, 910–917 (1983).

44. 44.

Gudmundsson, S., Vogelman, B. & Craig, W. A. The in-vivo postantibiotic effect of imipenem and other new antimicrobials. *J. Antimicrob. Chemother.* **18**, 67–73 (1986).

45. 45.

Goutelle, S. et al. The Hill equation: a review of its capabilities in pharmacological modelling. *Fundam. Clin. Pharmacol.* **22**, 633–648 (2008).

46. 46.

Institute of Laboratory Animal Resources Commission of Life Sciences. *A Guide for the Care and Use of Laboratory Animals* (National Academy Press, 1996).

47. 47.

Byrne, W. R. et al. Antibiotic treatment of experimental pneumonic plague in mice. *Antimicrob. Agents Chemother.* **42**, 675–681 (1998).

48. 48.

May, K. R. The Collison nebulizer: description, performance and application. *J. Aerosol Sci.* **4**, 235–243 (1973).

49. 49.

Hartings, J. M. & Roy, C. J. The automated bioaerosol exposure system: preclinical platform development and a respiratory dosimetry application with nonhuman primates. *J. Pharmacol. Toxicol. Methods* **49**, 39–55 (2004).

50. 50.

Guyton, A. C. Measurement of the respiratory volumes of laboratory animals. *Am. J. Physiol.* **150**, 70–77 (1947).

51. 51.

Morris, J. et al. Neurotropic threat characterization of *Burkholderia pseudomallei* strains. *Emerg. Infect Dis.* **21**, 58–63 (2015).

Acknowledgements

We thank D. McKinney, K. Thakur, S. Narayan and Pharmaron for their chemistry support; IHMA and JMI laboratories for microbiological surveillance studies; Neosome Life Sciences for in vivo efficacy studies; MPI Research for preclinical toxicology studies and J. Abendroth for help with crystallography. This research programme was partially supported by the Cooperative Agreement No. 4500002444 from ASPR/BARDA and by awards from Wellcome Trust and Germany's Federal Ministry of Education and Research (BMBF), as administrated by CARB-X. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Department of Health and Human Services Office of the Assistant Secretary for Preparedness and Response, other funders, or

CARB-X. 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. Use of the LS-CAT Sector 21 was supported by the Michigan Economic Development Corporation and the Michigan Technology Tri-Corridor (grant 085P1000817).

Author information

Affiliations

1. Entasis Therapeutics, Waltham, MA, USA

Thomas F. Durand-Reville, Alita A. Miller, John P. O'Donnell, Mark A. Sylvester, Satenig Guler, Ramkumar Iyer, Adam B. Shapiro, Nicole M. Carter, Camilo Velez-Vega, Samir H. Moussa, Sarah M. McLeod, April Chen, Angela M. Tanudra, Jing Zhang, Janelle Comita-Prevoir, Jan A. Romero & Ruben A. Tommasi

2. Novartis Institute for Biomedical Research, Cambridge, MA, USA

Xiaoyun Wu & Camilo Velez-Vega

3. EMD Serono, Billerica, MA, USA

Satenig Guler

4. Janssen Pharmaceuticals, Raritan, NJ, USA

Jing Zhang

5. Pfizer, Cambridge, MA, USA

Jan A. Romero

6. Alkermes, Waltham, MA, USA

Hoan Huynh

7. Takeda, Cambridge, MA, USA

Andrew D. Ferguson

8. UCB Boston, Bedford, MA, USA

Peter S. Horanyi & Stephen J. Mayclin

9. Institute for Therapeutic Innovation, Orlando, FL, USA

Henry S. Heine & George L. Drusano

10. Colorado State University, Fort Collins, CO, USA

Jason E. Cummings & Richard A. Slayden

Contributions

R.A.T. oversaw execution of all aspects of the project. T.F.D.-R., A.A.M. and J.P.O. wrote the manuscript and directed chemistry, biology and drug metabolism and pharmacokinetics/toxicology studies, respectively. X.W., M.A.S., S.G., J.Z., J.C.-P., J.A.R. and H.H. performed medicinal chemistry experiments. N.M.C., S.M.M., S.H.M. and R.I. performed microbiological experiments. A.C. and A.M.T. performed drug metabolism and pharmacokinetics experiments. S.H.M. performed microscopy, whole-genome sequencing and analysis. R.I. performed TOMAS. A.B.S. performed biochemical experiments. C.V.-V. generated and analysed MD simulations. A.D.F., P.S.H. and S.J.M. performed crystallography experiments. H.S.H., G.L.D., J.E.C. and R.A.S. performed microbiological and *in vivo* efficacy experiments using biothreat pathogens.

Corresponding author

Correspondence to Ruben A. Tommasi.

Ethics declarations

Competing interests

T.F.D.-R., A.A.M., J.P.O., X.W., M.A.S., S.G., R.I., A.B.S., N.M.C., C.V.-V., S.H.M., S.M.M., A.C., A.M.T., J.Z., J.C.-P., J.A.R. and R.A.T. are current or former employees of Entasis Therapeutics. H.H., A.D.F., P.S.H., S.J.M., H.S.H., G.L.D., J.E.C. and R.A.S. have no competing interests to declare.

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 Compound-dependent bacterial morphologies and evidence of porin overexpression in TOMAS.](#)

a, Morphology of *P. aeruginosa* PAO1 without drug treatment (left panel, 5 µm scale bar) or treated with 0.5X MIC of either NXL-105 (middle panel, 5 µm scale bar) or Compound 2 (right panel, 10 µm scale bar). Micrographs are representative of results obtained for three biologically independent studies. **b**, SDS-PAGE analysis of overexpressed porins in TOMAS¹⁹ after overnight induction with 120 µM L-arabinose. 2 µg outer membrane extracts were loaded onto a 13% polyacrylamide-SDS gel in 6 M urea. Expression levels are comparable to those observed for OprD in TOMAS¹⁹. Gel is representative of two biologically independent studies. Bands corresponding to the overexpressed porins are highlighted in boxes. Upper panel = *P. aeruginosa* porins; lower panel = *K. pneumoniae* porins.

[Extended Data Fig. 2 In vitro and in vivo activity of ETX0462.](#)

a, Percent cumulative growth inhibition of 40 MDR *P. aeruginosa* clinical isolates from the CDC MDR *P. aeruginosa* panel at increasing concentrations of ETX0462 (blue), meropenem-vaborbactam (MEM-VAB, red), imipenem-relebactam (IMI-REL, orange, ceftolozane-tazobactam (TOL-TAZ, purple) and ceftazidime-avibactam (CAZ-AVI, green). MICs were measured in singleton. **b**, A composite Emax fitting of the %fT>MIC vs change in CFU burden for 7 strains (MIC range of 0.25 – 4 mg/L, $R^2 = 0.87$) evaluated in a murine neutropenic thigh model *in vivo* suggests > 1-log kill when unbound systemic concentrations of ETX0462 exceed the MIC for 60% of the dosing interval. **c**, Kaplan-Meier survival plot of *Y. pestis* CO92 infected mice using an aerosolized dose (20 X LD₅₀ of 6.8 x 10⁴ CFU) followed by treatment with ETX0462 vs. vehicle control, ciprofloxacin (CIP) or ceftazidime (CTZ) shows equivalent *in vivo* efficacy (Log-rank test p < 0.0001 vs. control).

[Source data](#)

Extended Data Table 1 The structure, spectrum of β -lactamase inhibitory activity and *in vitro* antibacterial spectrum of activity of selected DBO analogs

Extended Data Table 2 Antibacterial activity, frequency of resistance and PBP acylation rates for direct-acting DBO analogs

Extended Data Table 3 Antibacterial activity of Compound 2 against recent clinical isolates of Gram-negative pathogens

Extended Data Table 4 Antibacterial activity of Compound 2 and ETX0462 against an isogenic *P. aeruginosa* panel of strains overexpressing individual β -lactamases

Extended Data Table 5 Genotypes of *P. aeruginosa* clinical isolates with varying susceptibilities to Compound 2 and ETX0462, using PAO1 as the reference strain (highlighted in gray)

Extended Data Table 6 Antibacterial activity of Compound 2 alone or in the presence of the efflux inhibitor phenylalanine arginine naphthylamide (PA β N, 10 mg/L) or the outer membrane permeabilizer polymyxin B nonapeptide (PMBN, 2 mg/L) against PAO1 and representative, less-susceptible *P. aeruginosa* clinical isolates

Extended Data Table 7 Inhibition of serine β -lactamases by ETX0462

Extended Data Table 8 Antibacterial activity of ETX0462 and comparators against global clinical isolates of Gram-negative and biothreat pathogens

Supplementary information

Supplementary Information

This file contains Supplementary Methods, details regarding the synthetic scheme and experimental procedure for ETX0462 and compound 2, HPLC and NMR spectra data, Tables 1, 2 and Supplementary Fig. 1.

Reporting Summary

Source data

Source Data Fig. 2

Source Data Fig. 4

Source Data Extended Data Fig. 2

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Durand-Reville, T.F., Miller, A.A., O'Donnell, J.P. *et al.* Rational design of a new antibiotic class for drug-resistant infections. *Nature* **597**, 698–702 (2021). <https://doi.org/10.1038/s41586-021-03899-0>

- Received: 20 October 2020
- Accepted: 11 August 2021
- Published: 15 September 2021
- Issue Date: 30 September 2021
- DOI: <https://doi.org/10.1038/s41586-021-03899-0>

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/s41586-021-03899-0>

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

- Article
- Open Access
- [Published: 24 August 2021](#)

Emergence and expansion of SARS-CoV-2 B.1.526 after identification in New York

- [Medini K. Annavajhala](#) ORCID: [orcid.org/0000-0002-9229-8849^{1 na1}](https://orcid.org/0000-0002-9229-8849),
- [Hiroshi Mohri](#) ORCID: [orcid.org/0000-0003-3101-7625^{2 na1}](https://orcid.org/0000-0003-3101-7625),
- [Pengfei Wang²](#),
- [Manoj Nair](#) ORCID: [orcid.org/0000-0002-5994-3957²](https://orcid.org/0000-0002-5994-3957),
- [Jason E. Zucker](#) ORCID: [orcid.org/0000-0001-6987-6412¹](https://orcid.org/0000-0001-6987-6412),
- [Zizhang Sheng²](#),
- [Angela Gomez-Simmonds¹](#),
- [Anne L. Kelley¹](#),
- [Maya Tagliavia¹](#),
- [Yaoxing Huang](#) ORCID: [orcid.org/0000-0001-6270-1644²](https://orcid.org/0000-0001-6270-1644),
- [Trevor Bedford³](#),
- [David D. Ho](#) ORCID: [orcid.org/0000-0003-1627-149X^{1,2,4 na2}](https://orcid.org/0000-0003-1627-149X) &
- [Anne-Catrin Uhlemann](#) ORCID: [orcid.org/0000-0002-9798-4768^{1 na2}](https://orcid.org/0000-0002-9798-4768)

Nature volume 597, pages 703–708 (2021)

- 12k Accesses
- 44 Altmetric
- [Metrics details](#)

Subjects

- [Microbiology](#)
- [SARS-CoV-2](#)
- [Virology](#)

Abstract

SARS-CoV-2 infections have surged across the globe in recent months, concomitant with considerable viral evolution^{1,2,3}. Extensive mutations in the spike protein may threaten the efficacy of vaccines and therapeutic monoclonal antibodies⁴. Two signature spike mutations of concern are E484K, which has a crucial role in the loss of neutralizing activity of antibodies, and N501Y, a driver of rapid worldwide transmission of the B.1.1.7 lineage. Here we report the emergence of the variant lineage B.1.526 (also known as the Iota variant⁵), which contains E484K, and its rise to dominance in New York City in early 2021. This variant is partially or completely resistant to two therapeutic monoclonal antibodies that are in clinical use and is less susceptible to neutralization by plasma from individuals who had recovered from SARS-CoV-2 infection or serum from vaccinated individuals, posing a modest antigenic challenge. The presence of the B.1.526 lineage has now been reported in all 50 states in the United States and in many other countries. B.1.526 rapidly replaced earlier lineages in New York, with an estimated transmission advantage of 35%. These transmission dynamics, together with the relative antibody resistance of its E484K sub-lineage, are likely to have contributed to the sharp rise and rapid spread of B.1.526. Although SARS-CoV-2 B.1.526 initially outpaced B.1.1.7 in the region, its growth subsequently slowed concurrently with the rise of B.1.1.7 and ensuing variants.

[Download PDF](#)

Main

Evolution of SARS-CoV-2 was slow at the beginning of the global pandemic⁶; however, multiple major variants of concern have emerged over the past year^{1,2,3,7}. These lineages are characterized by mutations in the spike protein, raising concerns that they may escape from therapeutic monoclonal and vaccine-induced antibodies. The hallmark mutation of B.1.1.7—a SARS-CoV-2 variant of concern first identified in the UK—is N501Y, located in the receptor-binding domain (RBD) of spike¹. This mutation appears to render the virus more transmissible and virulent^{8,9,10}, perhaps owing to a higher binding affinity of N501Y for the ACE2 receptor¹¹ or a greater propensity to evade host innate immune responses¹². Two other variants of concern, B.1.351² and P.1³, also harbour the N501Y mutation, in addition to an E484K substitution in the RBD^{2,3}. P.1 was identified as part of a second surge in Manaus, Brazil, despite a high pre-existing SARS-CoV-2 seroprevalence in the population^{13,14}. Reinfections with P.1 and another related Brazilian variant P.2 harbouring E484K have been documented^{15,16}. A previous study on B.1.351 demonstrated that this variant is

refractory to neutralization by a number of monoclonal antibodies directed to the top of the RBD, including several that have received emergency use authorization⁴. B.1.351 was markedly more resistant to neutralization by plasma from individuals who had recovered from SARS-CoV-2 infection and sera from vaccinated individuals. Of note, these effects were mediated in part by the E484K mutation. These findings are concerning in light of recent reports that three vaccine trials in South Africa showed a substantial drop in efficacy^{17,18}. Similarly, P.1 was also relatively resistant to antibody neutralization, although to a lesser degree¹⁹. In this study, we have implemented rapid molecular screening for signature mutations implicated in the success of these early variants of concern.

Rapid screening for SARS-CoV-2 mutations

We developed rapid PCR-based single-nucleotide-polymorphism (SNP) assays (Extended Data Fig. 1) to identify N501Y and E484K mutations in SARS-CoV-2-positive clinical samples stored at the Columbia University Biobank. We genotyped 1,533 samples between 1 November 2020 and 15 March 2021; 169 (11%) contained E484K, 43 (2.8%) contained N501Y and 1 sample contained both mutations. The earliest sample containing E484K was collected in mid-November 2020. The proportion of samples containing E484K increased substantially from 1.8% between 1 and 15 December 2020 to 26.1% between 1 and 15 March 2021 (Fig. 1a). Targeted PCR genotyping was continued through 1 May 2021 but was supplemented and subsequently replaced by whole-genome sequencing beginning in mid-March 2021. The frequency of viruses harbouring N501Y also increased over time, from the earliest detection in mid-January to 5.3% of screened isolates by the beginning of March.

Fig. 1: Prevalence of E484K-harbouring SARS-CoV-2 and B.1.526.



a, Detection of viruses with key signature mutations in spike protein over time. The earliest detected variant with the E484K mutation was collected in mid-November 2020. The prevalence of E484K (samples with E484K/total PCR-genotyped samples)

subsequently increased over time, from 1.8% between 1 and 15 December 2020 to 26.1% between 1 and 15 March 2021. Throughout late 2020 and early 2021, we identified fewer isolates with N501Y than with E484K, with a maximum of 5.9% of isolates containing N501Y in mid-February 2021. **b**, Distribution of different viral lineages identified by whole-genome sequencing. Within our collection of SARS-CoV-2 genomes ($n = 1,507$), the B.1.526 lineage rapidly increased in prevalence in early 2021, replacing the majority of other lineages (the blank space) present during this timeframe. This was followed by a steady rise in B.1.1.7 by mid-2021. The line below the x axis denotes the time period used to calculate the growth advantage of B.1.526 over other viruses that appeared earlier.

Genomic surveillance of SARS-CoV-2

We next performed untargeted whole-genome nanopore sequencing of nasopharyngeal swab samples collected throughout the study period with a cycle threshold (C_t) ≤ 35 . We obtained 1,507 SARS-CoV-2 whole genomes (59% of samples with $C_t \leq 35$; Extended Data Fig. 2). Sequencing results verified the E484K and N501Y substitutions in all samples identified by PCR screening. Of the sequenced N501Y isolates, 31 out of 41 (76%) were consistent with the B.1.1.7 lineage. Samples that harboured both N501Y and E484K were genotyped as P.1 ($n = 6$), B.1.351 ($n = 1$) and B.1.623 ($n = 1$). However, unexpectedly, the large majority of PCR-screened cases with E484K (98 out of 128 (77%)) were from the B.1.526 lineage²⁰.

Analysis of this genomic collection (Fig. 1b) showed that by May 2021, SARS-CoV-2 variants (including B.1.526, B.1.1.7 and, more recently, P.1) comprised two-thirds of all sequenced isolates, replacing the vast majority of earlier lineages (Fig. 1b). The proportion of infections caused by B.1.526 rose rapidly from late 2020 to February 2021, and remained at approximately 40–50% of all sequenced cases from March to May 2021, despite a concurrent increase in B.1.1.7. Indeed, during December and January, when the prevalence of B.1.1.7 was still negligible (Fig. 1b), the frequency of all viruses in the B.1.526 lineage increased from less than 5% to 50%, while the frequency of other lineages declined from more than 95% to 50% (Fig. 1b, where white blank space represents other lineages). Calculations using these numbers in a head-to-head comparison and an established mathematical method²¹ indicate that B.1.526 has a growth advantage of approximately 5% per day. Similarly, fitting a logistic regression model to 478 individual observations from the extended timeframe of November 2020 to January 2021 shows that B.1.526 had a similar growth advantage of 4.6% per day (95% confidence interval 2.8–6.5% per day). Given that the serial interval for SARS-CoV-2 transmission is about 7 days in the absence of any intervention²², these results suggest that B.1.526 is about 35% more transmissible than non-variant viruses.

Demographic and clinical features, including clinical outcomes, were largely similar in patients infected with viruses containing E484K versus those without the signature E484K or N501Y mutations, and between patients with B.1.526-E484K versus those with non-variant lineages²³ (Extended Data Table 1). However, significantly lower C_t values were associated with both E484K (29.49 versus 30.71, $P = 0.013$) and B.1.526-E484K (27.65 versus 28.81 in non-variant lineages, $P = 0.015$), indicating a modestly higher viral load in these variant samples. A significantly higher proportion of patients infected with B.1.526-E484K were admitted to hospital or presented to the emergency department ($P = 0.037$).

Signature B.1.526 lineage mutations

We identified signature spike protein mutations in the B.1.526 lineage by comparing all genomes generated in this study (Extended Data Fig. 3). Phylogenetic examination showed that the B.1.526 lineage comprises two closely related sub-lineages harbouring either E484K (B.1.526-E484K; defined as Pangolin lineage B.1.526) or S477N (B.1.526-S477N; Pangolin lineage B.1.526.2), and the additional sub-lineage B.1.526.1, harbouring the L452R substitution (B.1.526-L452R). B.1.526-E484K and B.1.526-S477N share the characteristic spike protein mutations L5F, T95I, D253G, D614G and either A701V or Q957R, along with either E484K or S477N. Non-spike mutations widely shared by B.1.526-E484K and B.1.526-S477N isolates include: T85I in ORF1a-nsp2; L438P in ORF1a-nsp4, a 9-base pair (bp) deletion ($\Delta 106–108$) in ORF1a-nsp6; P323L in ORF1b-nsp12; Q88H in ORF1b-nsp13; Q57H in ORF3a; and P199L and M234I in the N gene. While B.1.526-L452R isolates shared a number of mutations across the genome in ORF-1ab, ORF-3ab, ORF-8 and N, they did not share characteristic spike mutations with B.1.526-E484K and B.1.526-S477N.

To further investigate the evolutionary history of B.1.526, we performed phylogenetic analyses on genomes in this collection and in the GISAID collection harbouring the ORF1a-nsp6 deletion $\Delta 106–108$, along with the mutation A20262G that uniquely defines the parent clade containing B.1.526 and related viruses (Fig. 2a). We observed a stepwise emergence of the key lineage-defining mutations, with T95I, D253G and L5F appearing in the earliest phylogenetic nodes. Isolates subsequently branched into four sub-lineages, with two major groups B.1.526-E484K and B.1.526-S477N containing A701V, and a smaller sub-lineage B.1.526-S477N containing Q957R. The B.1.526-L452R lineage—which emerged in parallel with these—is related to B.1.526-E484K and B.1.526-S477N, but forms a distinct phylogenetic branch (Extended Data Fig. 3).

Fig. 2: Spike protein amino acid substitutions and structural changes represented in sequenced isolates.

 **figure2**

a, Maximum-likelihood phylogenetic tree of 2,309 SARS-CoV-2 viruses coloured according to spike protein haplotype. Spike protein mutations are labelled on the tree, showing the stepwise accumulation of signature B.1.526 mutations T95I, D253G and L5F, and branching of B.1.526-E484K (orange) and two B.1.526-S477N sub-lineages (yellow, blue). The B.1.526-L452R sub-lineage (green) emerged in parallel. An interactive version of this figure is available at <https://nextstrain.org/groups/blab/ncov/ny/B.1.526>. **b**, Key mutations of B.1.526 displayed on the spike trimer. The D253G mutation resides in the antigenic supersite within the N-terminal domain (NTD), a target for neutralizing antibodies, E484K and S477N at the RBD interface with the cellular receptor ACE2, and A701V near the furin cleavage site.

Fig. 2b shows the localization of signature B.1.526-E484K and B.1.526-S477N mutations within the spike protein. D253G resides in the antigenic supersite in the N-terminal domain²⁴, which is a target for neutralizing antibodies²⁵, whereas E484K is situated at the RBD interface with the cellular receptor ACE2. The A701V mutation near the furin cleavage site is also shared with variant B.1.351.

Antibody neutralization of B.1.526

The effects of the signature spike protein mutations in B.1.526 on antibody neutralization were first assessed using vesicular stomatitis virus (VSV)-based pseudoviruses, as previously described^{4,25}. Pseudoviruses were constructed containing S477N or E484K alone, or containing all five signature mutations (L5F, T95I, D253G, A701V and E484K or S477N) (NYΔ5(E484K) or NYΔ5(S477N)), and analysed in a neutralization assay with 12 monoclonal antibodies (including 5 with emergency use

authorization), 20 plasma samples from patients who had recovered from SARS-CoV-2 infection and 22 sera from vaccinated individuals⁴. The neutralizing activities of 12 monoclonal antibodies covering a range of epitopes on RBD were essentially unaltered against the S477N and NYΔ5(S477N) pseudoviruses (Extended Data Fig. 4a), showing that this mutation has no discernible antigenic impact, as was validated using convalescent plasma and vaccinee sera (Extended Data Fig. 4b). However, the activities of several antibodies—including REGN10933 and LY-CoV555, which are already in clinical use—were either impaired or lost when tested against E484K and NYΔ5(E484K) pseudoviruses (Fig. 3a). Similarly, neutralizing activities of convalescent plasma or vaccinee sera were reduced by 4.1-fold or 3.3- to 3.6-fold, respectively, against NYΔ5(E484K) (Fig. 3b). Neutralization studies on the authentic B.1.526-E484K virus yielded similar results, although the magnitudes of resistance to convalescent plasma and vaccinee sera were slightly lower at 2.6-fold and 1.8- to 2.0-fold, respectively (Fig. 3b). A comparative analysis with other variants of concern (Fig. 3c) showed that antibody resistance of B.1.526-E484K is probably lower than that of B.1.351 and closer to that of P.1. Overall, these results demonstrate the need to modify the antibody therapies currently in use and to monitor the efficacy of current vaccines in regions where B.1.526-E484K is prevalent.

Fig. 3: Neutralization studies of B.1.526-E484K and comparative analyses.

 **figure3**

a. Neutralizing activities of 12 monoclonal antibodies against pseudoviruses containing E484K alone or all five signature B.1.526 mutations (L5F, T9I, D253G, A701V and E484K) (NYΔ5(E484K)), as well as against the authentic B.1.526-E484K. Antibodies with emergency use authorization are shown with bold solid lines. Data are mean ± s.e.m. of technical triplicates and represent one of two independent

experiments. IC₅₀, half-maximal inhibitory concentration; WT, wild type. **b**, Fold change in virus-neutralizing activity of plasma from patients who have recovered from SARS-CoV-2 infection (convalescent plasma) ($n = 20$) and sera from vaccinated individuals (vaccinee sera) ($n = 22$) against the NYΔ5(E484K) pseudovirus compared with wild-type pseudovirus, as well as against authentic B.1.526-E484K and wild-type virus (WA1) (numbers shown above P -values). ID₅₀, antibody dose required to reduce viral count by 50%. **c**, Fold change in neutralization ID₅₀ for plasma from patients who have recovered from SARS-CoV-2 infection and sera from vaccinated individuals against different variant pseudoviruses and live viruses compared with wild-type counterparts. The data for B.1.1.7, B.1.351 and P.1 were derived from previous studies^{4,19}. Data from 20 recovered patients or 22 vaccinated individuals were averaged and are presented as arithmetic mean \pm s.e.m. (individual data points are also shown). Statistical comparisons were made using the Wilcoxon matched-pairs signed rank test; two-tailed P -values are reported.

The spread of B.1.526 across New York and the US

Prevalence of the novel variant B.1.526 surged rapidly in the CUIMC catchment area (Fig. 4a) and throughout New York state (Fig. 4b) following its emergence in late 2020, replacing other lineages and initially outpacing B.1.1.7. A multinomial logistic regression model describing the concurrent growth rates of these two lineages shows that starting in mid-April 2021, B.1.1.7 surpassed B.1.526 owing to a slightly higher fitness, with estimated growth rates in New York state of 5.3% per day for B.1.1.7 (95% confidence interval 5.0–5.7%) and 3.4% per day for B.1.526 (95% confidence interval 3.2–3.6%) (Fig. 4b). These estimates suggest a fitness advantage of B.1.526 over existing non-variant lineages^{21,22} of 22–25% over a serial interval of 7 days during a period when multiple variants were competing simultaneously. Furthermore, the estimates also suggest a fitness advantage of B.1.1.7 over existing non-variant lineages of 35–40%, as well as a fitness advantage of B.1.1.7 over B.1.526 of 12–15%. Both lineages grew quickly (Fig. 4a,b), but once they reached a high frequency of circulating viruses, the competition between them caused the growth of B.1.1.7 to slow and that of B.1.526 to decline.

Fig. 4: Spread of B.1.1.7 and B.1.526 lineages in New York and the United States.

 **figure4**

a, b, Frequencies of B.1.1.7 (blue) and B.1.526 (yellow) lineages in the CUIMC catchment area (**a**) and New York state (**b**) during January to May 2021, with dots representing daily seven-day sliding window averages and lines representing fit to a multinomial logistic regression model. **c**, Ternary plot of state-level frequency trajectories for 42 US states, separating frequencies of B.1.1.7, B.1.526 and other lineages. Each state-level trajectory is shown as a line from lower left in January 2021—when both B.1.1.7 and B.1.526 were rare—to right, as B.1.1.7 and B.1.526 increase in frequency. The trajectory for New York state is highlighted in purple. **d**, The same data presented in **c**, except frequency of B.1.1.7 is plotted against logistic growth rate of B.1.526. **e**, Phylogenetic tree of 933 B.1.526 samples from across the United States. Branch tips are coloured on the basis of location of sampling and branches are coloured by inferred ancestral location. **f**, Phylogeographic view of data from **e**. Each sampling location is represented as a circle with area proportional to sample count and each inferred transition event across the phylogeny is drawn as an arc connecting the inferred origin and destination. Most migration events are inferred to be direct dispersals from New York state. Data in **e, f** were visualized using NextStrain (<https://nextstrain.org>) and made available through a CC-BY-4.0 license.

Trajectories of the frequencies of B.1.1.7 and B.1.526 across states (Fig. 4c, Extended Data Fig. 5) show two general patterns: (1) an initial rapid increase of both lineages until the proportion of other lineages had been eclipsed, followed by a decline of B.1.526, as seen in New York and in several neighbouring states; and (2) rapid growth and resulting dominance of B.1.1.7, preventing the further rise of B.1.526. The dynamics between these two lineages is described further in Fig. 4d, which plots the logistic growth rate of B.1.526 against the frequency of B.1.1.7, again at the state level. At lower frequencies of B.1.1.7, all states show a similarly rapid growth of B.1.526 as it replaces non-variant lineages. However, as B.1.1.7 increases in

frequency, the growth of B.1.526 slows, again indicative of the slightly higher fitness of B.1.1.7. At a minimum, the proportion of B.1.526 increased rapidly where B.1.1.7 was not already dominant and continued to grow at a similar pace as B.1.1.7 in several states (Extended Data Fig. 5).

Phylogeographic analysis of the B.1.526 lineage revealed ancestral viruses originating in New York in August 2020, diversifying within the state, and then dispersing to other states (Fig. 4e,f). State-level genomic data showed that B.1.526 was concentrated primarily in New York and surrounding states, including New Jersey and Rhode Island (Extended Data Fig. 5). This suggests that B.1.526—and B.1.526-E484K in particular—became widespread in the region, the original epicenter of COVID-19 in the United States^{26,27}, although the lineage has also grown in states outside the northeastern United States (for example, North Carolina). By the end of April 2021, B.1.526 was widely distributed within the United States, and the lineage had emerged and expanded in multiple states across the country (Fig. 4f). This ability of B.1.526 to spread rapidly across the United States (Extended Data Fig. 5) and internationally is notable.

Discussion

Here we report the emergence of the SARS-CoV-2 lineage B.1.526 and the surge of B.1.526 infections in New York during the second wave of the COVID-19 pandemic. Neutralization studies on B.1.526-E484K demonstrate that the activities of several antibodies are either impaired or lost with this variant, including two antibodies (Ly-CoV555 and REGN10933) that are already in clinical use. Furthermore, neutralizing activities of plasma from individuals who have recovered from SARS-CoV-2 infection or sera from vaccinated individuals were lower against B.1.526-E484K. By contrast, the S477N mutation, a key signature of another B.1.526 sub-lineage, did not have an impact on antibody neutralization.

This study has several limitations. This was a single-centre genomic survey of patients attending hospital and may not have fully captured patients with milder disease. However, our results are comparable to genomic data released by public health laboratories in the region and incorporate all publicly available data for phylogeographic context and growth rate calculations. As in all genomic surveillance studies, we predominantly sequenced samples with a C_t less than 30, but this included a high proportion of samples throughout the study period. In addition, our PCR screen enabled us to obtain unbiased estimates of E484K and N501Y prevalence early on in the study. PCR approaches may be increasingly warranted for continued surveillance during non-surge periods, during which C_t values trend higher. Finally, transmissibility estimates based on observed prevalence are imperfect as they reflect observed growth rates rather than intrinsic transmissibility of the virus.

Together, our findings underscore the importance of the E484K mutation, which has emerged in at least 246 different lineages²⁸ of SARS-CoV-2, a powerful illustration of convergent evolution. This highlights that E484K can rapidly emerge in multiple clonal backgrounds and may warrant targeted screening for this key mutation in addition to robust genomic surveillance programs. However, B.1.526 is one of the few lineages with E484K that has risen to prominence. The greatest threat of B.1.526 appears to be the ease with which it spreads, with an estimated transmissibility approximately 35% higher than non-variant viruses when competing head to head. Despite the notable transmissibility of B.1.1.7, B.1.526 was able to spread rapidly in the United States, replacing other lineages and continuing to increase in frequency in several states where both B.1.526 and B.1.1.7 were predominant. Similarly, although B.1.351 may pose the greatest antigenic challenge to antibodies and vaccines, the B.1.526-E484K sub-lineage also exhibits resistance to antibody neutralization. Our findings present a clear-cut example of SARS-CoV-2 evolution in real time. B.1.526, with its higher transmissibility, appeared suddenly and rose to dominance, only to wane as variants (B.1.1.7 and, more recently, B.1.617.2) with even higher fitness emerged. These observations are a stark reminder that increasingly concerning variants are expected to emerge if SARS-CoV-2 is allowed to continue its spread.

Methods

Clinical cohort

This observational study took place at an academic quaternary care centre in New York City. Nasopharyngeal swabs obtained as part of routine clinical care were tested by the Clinical Microbiology laboratory, and positive specimens were transferred to the Columbia University Biobank for inactivation and storage. Electronic health records data extracted for this analysis included demographics, laboratory results, admission, discharge and transfer dates, current and historical international classification of disease codes (ICD 9 and ICD 10) extracted from the clinical data warehouse. The study was reviewed and approved by the Columbia University Institutional Review Board (IRB) (AAAT0123). The IRB waived consent for the entirety of this observational study, including for the collection and sequencing of the viral samples as well as the abstraction of the clinical metadata, as this observational study met the requirements for this exception. These include minimal risks to subjects, not adversely affecting the rights and welfare of the subjects, and that the research could not be carried out without the waiver.

PCR screening

Extended Data Fig. 1 describes the overall protocol for variant screening. To enable rapid PCR-based screening, we prepared RNA using the heat inactivation method in place of RNA-isolation methods²⁹. First, 50 µl of nasal swab sample in viral transport medium was transferred to 96-well PCR plates, covered with an adhesive aluminium foil (VWR 60941-076) and incubated at 95 °C for 5 min using the PCR instrument. After the centrifugation of the plate at >2,100g for 5 min, 5 µl of the supernatant from each sample, which contains viral RNA, was used for the SNP assay.

The SNP assay consists of four steps as follows: reverse transcription of viral RNA, pre-read of the SNP assay, real-time PCR and post-read of the SNP assay. 5 µl of RNA from the supernatant was added to 15 µl of the single step quantitative PCR with reverse transcription (RT–qPCR) reaction mix, which consists of 5 µl of TaqPath 1-step RT–qPCR Master Mix, CG (4×) (ThermoFisher Scientific), 500 nM of forward and reverse primers, 120 nM of VIC-MGB probe, 50 nM of FAM-MGB probe, 1/2000 volume of ROX Reference Dye (Invitrogen) as the final concentration, and nuclease-free water to adjust the total reaction volume of 20 µl. Each reaction plate included 8 control wells, 5×10^6 and 5×10^3 copies of WA-1 (wild type), B.1.1.7 and B.1.351, which were generated by PCR to match the variant sequences, and 2 wells with water as no template controls (NTC).

The primer pairs and probes used are as follows. For the SNP assay for position 501, a primer pair of 501.F: 5'- GGT TTT AAT TGT TAC TTT CCT TTA CA-3' and 501.R: 5'-AGT TCA AAA GAA AGT ACT ACT ACT CTG TAT G-3' were used with two TaqMan probes (ThermoFisher Scientific), one for wild type, VIC.N501MGB: [VIC]-AA CCC ACT AAT GGT-MGBNFQ and the other for variant type, FAM.Y501MGB: [FAM]-AAC CCA CTT ATG GT-MGBNFQ. For position 484, a primer pair of 484.F: 5'-AGA GAG ATA TTT CAA CTG AAA TCT ATCAGG-3' and 484.R: 5'-GAA ACC ATA TGA TTG TAA AGG AAA GTA AC-3' were used with two probes, one for wild type, VIC.E484MGB: [VIC]-ATG GTG TTG AAG GT-MGBNFQ and the other for variant type, FAM.K484MGB: [FAM]-ATG GTG TTA AAG GT-MGBNFQ.

The reaction plate was subjected to: (1) reverse transcription reaction at 25 °C for 2 min, 50 °C for 15 min and a hold at 4 °C; (2) SNP assay (pre-read) at 60 °C for 30 s; (3) real-time PCR at 95 °C for 20 s followed by 50 cycles of two-step PCR, at 95 °C for 3 s and at 60 °C for 30 s with the fast 7500 mode; followed by (4) SNP assay (post-read) at 60 °C for 30 s using ABI 7500 Fast Dx real-time PCR instrument with SDS Software (ThermoFisher Scientific). The genotype at each key position for each sample was determined by reading the component signal of the amplification and the allelic discrimination analysis software in the program.

Whole-genome sequencing

Extended Data Fig. 2 displays a flowchart outlining samples available for this study. Isolates with C_t values below 35 were selected for sequencing using the ARTIC v3 low-cost protocol targeting 400-bp amplicons³⁰ or Rapid Barcoding kit protocol targeting 1,200-bp amplicons³¹. In brief, RNA was extracted using the Qiagen RNeasy Mini kit or Zymo DNA/RNA Mini kit. Reverse transcription was performed using LunaScript RT SuperMix (NEB). Tiling PCR was performed on the cDNA, and amplicons were barcoded using the Oxford Nanopore Native Barcoding Expansion 96 kit. Pooled barcoded libraries were then sequenced on an Oxford Nanopore MinION sequencer using R9.4.1 flow cells. Base calling was performed in MinKNOW software v21.02.1. Sequencing runs were monitored in real time using RAMPART (<https://artic-network.github.io/rampart/>) to ensure sufficient genomic coverage with minimal runtime. Consensus sequence generation was performed using the ARTIC bioinformatics pipeline (<https://github.com/artic-network/artic-ncov2019>). Genomes were manually curated by visually inspecting sequencing alignment files for verification of key residues in Geneious v10.2.6.

Phylogenetic analysis

Phylogenetic reconstruction of amino acid changes (Fig. 2a) was conducted using the Nextstrain³² workflow at <https://github.com/nextstrain/ncov>, which aligns sequences against the Wuhan Hu-1 reference using nextalign (<https://github.com/nextstrain/nextclade>), constructs a maximum-likelihood phylogenetic tree via IQ-TREE³³, estimates molecular clock branch lengths via TreeTime³⁴ and reconstructs nucleotide and amino acid changes (also via TreeTime). This workflow was applied to 2,309 SARS-CoV-2 genomes with the 9-bp deletion Δ106–108 in ORF1a-nsp6 along with mutation A20262G, which demarcates the parent clade to lineage B.1.526 alongside 688 global reference viruses. This analysis was conducted on data downloaded³⁵ from GISAID (<https://gisaid.org/>) on 5 April 2021. Phylogeographic reconstruction of spread from New York state (Fig. 4e,f) was similarly conducted using the same Nextstrain workflow with the addition of performing ancestral trait reconstruction of the geographic ‘division’ attribute of 933 SARS-CoV-2 genomes downloaded from GISAID on 6 Jun 2021.

Neutralization studies of pseudoviruses

We assayed the neutralizing activity of monoclonal antibodies, convalescent plasma and vaccinee sera against E484K, S477N and wild-type (D614G) pseudoviruses, as well as pseudovirus NYΔ5 containing all five signature mutations of B.1.526-E484K (L5F, T95I, D253G, E484K, D614G and A701V), as previously described²⁵. We examined four monoclonal antibodies with emergency use authorization (CB6, REGN10987, REGN10933 and LY-CoV555), and eight additional RBD monoclonal

antibodies, including from our own collection (2-15, 2-7, 1-57 and 2-36)²⁵ as well as S309³⁶, COV2-2196 and COV2-2130³⁷, and C121³⁸. We also examined convalescent plasma collected in March and April 2020 ($n = 20$ patients), and sera from individuals who had received Moderna or Pfizer vaccine⁴ ($n = 22$). In brief, Vero E6 cells (ATCC) were seeded in 96-well plates (2×10^4 cells per well). Cell lines were negative for *Mycoplasma*, as assessed using the *Mycoplasma* PCR ELISA (Sigma). Pseudoviruses were incubated with serial dilutions of the test samples in triplicate for 30 min at 37 °C. The mixture was added to cultured cells and incubated for an additional 24 h. Luminescence was measured using a Britelite plus Reporter Gene Assay System (PerkinElmer), and IC₅₀ was defined as the dilution at which the relative light units were reduced by 50% compared with the virus control wells (virus + cells) after subtraction of the background in the control groups with cells only. The IC₅₀ values were calculated using nonlinear regression in GraphPad Prism 8.0. Statistical analysis was performed using a Wilcoxon matched-pairs signed rank test. Two-tailed *P*-values are reported.

Neutralization of infectious SARS-CoV-2

Infectious SARS-CoV-2 isolate hCoV-19/USA/NY-NP-DOH1/2021 was isolated at the Aaron Diamond AIDS Center (Columbia University Medical Center) from a nasopharyngeal swab and propagated for one passage in Vero E6 cells (ATCC). Infectious titre of the resulting virus was determined by an end-point dilution and cytopathic effect (CPE) assay on Vero-E6 cells as described previously²⁵. The virus has since been deposited at BEI Resources (catalogue (cat.) no. NR-55359). SARS-CoV-2 virus USA-WA1/2020 (WA1), obtained from BEI Resources (cat. no. NR-52281) served as the control in experiments.

An end-point dilution microplate neutralization assay was performed to measure the neutralization activity of twenty convalescent patient plasma samples and twelve purified monoclonal antibodies. In brief, plasma samples were subjected to successive fivefold dilutions starting from 1:100. Similarly, antibodies were serially diluted (fivefold dilutions) starting at 50 µg ml⁻¹. Triplicates of each dilution were incubated with SARS-CoV-2 at an multiplicity of infection of 0.1 in Eagle's minimum essential medium (EMEM; ATCC) with 7.5% inactivated fetal calf serum (FCS) for 1 h at 37 °C. After incubation, the virus–antibody mixture was transferred to a monolayer of Vero-E6 cells grown overnight. The cells were incubated with the mixture for about 70 h. CPE of viral infection was visually scored for each well in a blinded fashion by two independent observers. The results were then converted into percentage neutralization at a given sample dilution or antibody concentration, and the mean ± s.e.m. was plotted using a five-parameter dose-response curve in GraphPad Prism v8.4.

Growth dynamics

Growth dynamics of B.1.1.7 and B.1.526 were obtained through by downloading ‘metadata’ from GISAID on 6 June 2021 for all 422,760 viruses sampled from the United States collected after 1 January 2021. These metadata have PANGO lineages³⁹ already assigned to each genome sequence. Daily state-level frequencies (and frequencies for CUIMC) were extracted for plotting using seven-day sliding window averages of the prevalence of B.1.1.7 and B.1.526, calculated as the number of sequence-verified samples from each strain divided by the total number of positive samples with C_t values below 35, as this threshold value was used for sequencing.

Separately, a multinomial logistic regression model was fit directly to the observation data consisting of individual genomes, their dates of sampling (independent variable X in days since 1 January 2021) and their category labels (dependent variable Y , “B.1.1.7”, “B.1.526” and “other”). This results in a 4-parameter model where both B.1.1.7 and B.1.526 have parameters specified for frequency at day 0 (1 January 2021) and logistic growth rate. This model was fit to the data using the Classify package of Mathematica v12.2.

Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

Data availability

All genomes and associated metadata generated as a part of this study have been uploaded to GISAID (<https://gisaid.org>) and NCBI GenBank (BioProject Accession [PRJNA751551](#)). Biological materials (that is, variant pseudoviruses) generated as a part of this study will be made available but may require execution of a materials transfer agreement.

Code availability

Data processing and visualization were performed using publicly available software and packages, primarily RStudio v1.2.5033, GraphPad Prism v8.4, and iTOL (<https://itol.embl.de/>). The exact workflows used for phylogenetic (Fig. 2a) and phylogeographic analysis of public GISAID data (Fig. 4e,f) are available at <https://github.com/blab/ncov-ny>. Frequency dynamics were modelled using Mathematica in notebooks also available at <https://github.com/blab/ncov-ny>.

References

1. 1.

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. *Virological.org* <https://virological.org/t/preliminary-genomic-characterisation-of-an-emergent-sars-cov-2-lineage-in-the-uk-defined-by-a-novel-set-of-spoke-mutations/563> (2020).

2. 2.

Tegally, H. et al. Emergence and rapid spread of a new severe acute respiratory syndrome-related coronavirus 2 (SARS-CoV-2) lineage with multiple spike mutations in South Africa. Preprint at <https://doi.org/10.1101/2020.12.21.20248640> (2020).

3. 3.

Faria, N. R. et al. Genomics and epidemiology of the P.1 SARS-CoV-2 lineage in Manaus, Brazil. *Science* **372**, 815-821(2021).

4. 4.

Wang, P. et al. Antibody resistance of SARS-CoV-2 variants B.1.351 and B.1.1.7. *Nature* **593**, 130–135 (2021).

5. 5.

Tracking SARS-CoV-2 variants. *World Health Organisation* <https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/> (2021).

6. 6.

Duchene, S. et al. Temporal signal and the phylodynamic threshold of SARS-CoV-2. *Virus Evol.* **6**, veaa061 (2020).

7. 7.

Cherian, S. et al. Convergent evolution of SARS-CoV-2 spike mutations, L452R, E484Q and P681R, in the second wave of COVID-19 in Maharashtra, India. Preprint at <https://doi.org/10.1101/2021.04.22.440932> (2021).

8. 8.

Iacobucci, G. Covid-19: new UK variant may be linked to increased death rate, early data indicate. *BMJ* **372**, n230 (2021).

9. 9.

Volz, E. et al. Transmission of SARS-CoV-2 lineage B.1.1.7 in England: insights from linking epidemiological and genetic data. Preprint at <https://doi.org/10.1101/2020.12.30.20249034> (2021).

10. 10.

Washington, N. L. et al. Emergence and rapid transmission of SARS-CoV-2 B.1.1.7 in the United States. *Cell* **184**, 2587–2594.e7 (2021).

11. 11.

Greaney, A. J. et al. Comprehensive mapping of mutations in the SARS-CoV-2 receptor-binding domain that affect recognition by polyclonal human plasma antibodies. *Cell Host Microbe* **29**, 463–476.e466 (2021).

12. 12.

Thorne, L. G. et al. Evolution of enhanced innate immune evasion by the SARS-CoV-2 B.1.1.7 UK variant. Preprint at <https://doi.org/10.1101/2021.06.06.446826> (2021).

13. 13.

Naveca, F. et al. SARS-CoV-2 reinfection by the new variant of concern (VOC) P.1 in Amazonas, Brazil. *Virological.org* <https://virological.org/t/sars-cov-2-reinfection-by-the-new-variant-of-concern-voc-p-1-in-amazonas-brazil/596> (2021).

14. 14.

Sabino, E. C. et al. Resurgence of COVID-19 in Manaus, Brazil, despite high seroprevalence. *Lancet* **397**, 452–455 (2021).

15. 15.

Zucman, N., Uhel, F., Descamps, D., Roux, D. & Ricard, J. D. Severe reinfection with South African SARS-CoV-2 variant 501Y.V2: a case report. *Clin. Infect. Dis.* <https://doi.org/10.1093/cid/ciab129> (2021).

16. 16.

Nonaka, C. K. V. et al. Genomic evidence of SARS-CoV-2 reinfection involving E484K spike mutation, Brazil. *Emerg. Infect. Dis.* **27**, 1522–1524 (2021).

17. 17.

Callaway, E. & Mallapaty, S. Novavax offers first evidence that COVID vaccines protect people against variants. *Nature* **590**, 17 (2021).

18. 18.

Madhi, S. A. et al. Efficacy of the ChAdOx1 nCoV-19 Covid-19 vaccine against the B.1.351 variant. *N. Engl. J. Med.* **384**, 1885–1898 (2021).

19. 19.

Wang, P. et al. Increased resistance of SARS-CoV-2 variant P.1 to antibody neutralization. *Cell Host Microbe* **29**, 747–751.e4 (2021).

20. 20.

West, A. P., Barnes, C. O., Yang, Z. & Bjorkman, P. J. Detection and characterization of the SARS-CoV-2 lineage B.1.526 in New York. *Nat. Commun.* **12**, 4886 (2021).

21. 21.

Goudsmit, J., De Ronde, A., Ho, D. D. & Perelson, A. S. Human immunodeficiency virus fitness in vivo: calculations based on a single zidovudine resistance mutation at codon 215 of reverse transcriptase. *J. Virol.* **70**, 5662–5664 (1996).

22. 22.

Ali, S. T. et al. Serial interval of SARS-CoV-2 was shortened over time by nonpharmaceutical interventions. *Science* **369**, 1106–1109 (2020).

23. 23.

SARS-CoV-2 variant classifications and definitions. *Centers for Disease Control and Prevention* <https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-info.html> (2021).

24. 24.

Cerutti, G. et al. Potent SARS-CoV-2 neutralizing antibodies directed against spike N-terminal domain target a single supersite. *Cell Host Microbe* **29**, 819–833.e7 (2021).

25. 25.

Liu, L. et al. Potent neutralizing antibodies against multiple epitopes on SARS-CoV-2 spike. *Nature* **584**, 450–456 (2020).

26. 26.

COVID-19: Data. *New York City Department of Health* <https://www1.nyc.gov/site/doh/covid/covid-19-data-trends.page#antibody> (2021).

27. 27.

Lasek-Nesselquist, E., Lapierre, P., Schneider, E., George, K. S. & Pata, J. The localized rise of a B.1.526 SARS-CoV-2 variant containing an E484K mutation in New York State. Preprint at <https://doi.org/10.1101/2021.02.26.21251868> (2021).

28. 28.

Alaa Abdel Latif, K. G. et al. B.1.526 lineage report. *Outbreak.info* <https://outbreak.info/situation-reports/S-E484K> (2021).

29. 29.

Smyrlaki, I. et al. Massive and rapid COVID-19 testing is feasible by extraction-free SARS-CoV-2 RT-PCR. *Nat. Commun.* **11**, 4812 (2020).

30. 30.

Quick, J. Artic protocol. *Protocols.io* <https://www.protocols.io/view/ncov-2019-sequencing-protocol-v3-locost-bh42j8ye> (2021).

31. 31.

Freed, N., Vlkova, M., Faisal, M. B. & Silander, O. Rapid and inexpensive whole-genome sequencing of SARS-CoV2 using 1200 bp tiled amplicons and Oxford Nanopore rapid barcoding. *Biol. Methods Protoc.* **5**, bpaa014 (2020).

32. 32.

Hadfield, J. et al. Nextstrain: real-time tracking of pathogen evolution. *Bioinformatics* **34**, 4121–4123 (2018).

33. 33.

Minh, B. Q. et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* **37**, 1530–1534 (2020).

34. 34.

Sagulenko, P., Puller, V. & Neher, R. A. TreeTime: maximum-likelihood phylodynamic analysis. *Virus Evol.* **4**, vex042 (2018).

35. 35.

Shu, Y. & McCauley, J. GISAID: global initiative on sharing all influenza data—from vision to reality. *Euro Surveill.* **22**, 30494 (2017).

36. 36.

Pinto, D. et al. Cross-neutralization of SARS-CoV-2 by a human monoclonal SARS-CoV antibody. *Nature* **583**, 290–295 (2020).

37. 37.

Zost, S. J. et al. Rapid isolation and profiling of a diverse panel of human monoclonal antibodies targeting the SARS-CoV-2 spike protein. *Nat Med* **26**, 1422–1427 (2020).

38. 38.

Robbiani, D. F. et al. Convergent antibody responses to SARS-CoV-2 in convalescent individuals. *Nature* **584**, 437–442 (2020).

39. 39.

Rambaut, A. et al. A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat. Microbiol.* **5**, 1403–1407 (2020).

Acknowledgements

Biospecimens used for this research were obtained from the Columbia University Biobank (CUB) with technical support from V. J. Mahadeva, S. Fernando and S. T.

Parker-Jones. CUB is supported by the Irving Institute for Clinical and Translational Research (NCATS UL1TR001873). In particular, we thank M. Reilly, E. Hod and the CUB COVID-19 Genomics Consortium for facilitating this effort. We also thank L. Liu and S. Iketani for technical support, and A. Perelson for mathematical input. We gratefully acknowledge all the authors, the originating laboratories responsible for obtaining the specimens and the submitting laboratories for generating the genetic sequence and metadata and sharing via the GISAID Initiative, on which part of the presented research is based. This work was in part funded by NIH/NIDA grant U01 DA053949 (A.-C.U. and M.K.A.) and by support from Andrew and Peggy Cherng, Samuel Yin, Barbara Picower and the JBP Foundation, Brii Biosciences, Roger and David Wu, and the Bill and Melinda Gates Foundation. T.B. is a Pew Biomedical Scholar and is supported by NIH grant no. R35 GM119774-01. Funders and funding agencies had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author information

Author notes

1. These authors contributed equally: Medini K. Annavajhala, Hiroshi Mohri
2. These authors jointly supervised this work: David D. Ho, Anne-Catrin Uhlemann

Affiliations

1. Division of Infectious Diseases, Department of Medicine, Columbia University Vagelos College of Physicians and Surgeons, New York, NY, USA
Medini K. Annavajhala, Jason E. Zucker, Angela Gomez-Simmonds, Anne L. Kelley, Maya Tagliavia, David D. Ho & Anne-Catrin Uhlemann
2. Aaron Diamond AIDS Research Center, Columbia University Vagelos College of Physicians and Surgeons, New York, NY, USA
Hiroshi Mohri, Pengfei Wang, Manoj Nair, Zizhang Sheng, Yaoxing Huang & David D. Ho
3. Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA, USA
Trevor Bedford

4. Department of Microbiology and Immunology, Columbia University Irving Medical Center, New York, NY, USA

David D. Ho

Contributions

Conceptualization: A.-C.U., D.D.H., M.K.A. and H.M. Data curation: M.K.A., H.M., J.E.Z., P.W., M.N., Z.S., T.B., A.G.-S., Y.H., A.L.K., M.T. and A.-C.U. Formal analysis: M.K.A., P.W., J.E.Z., T.B. and A.G.-S. Funding acquisition: A.-C.U., D.D.H. and M.K.A. Investigation: M.K.A., H.M., J.E.Z., P.W., M. N., A.L.K., M.T., T.B. and Y.H. Methodology: M.K.A., H.M., P.W., M.N., T.B. and Y.H. Supervision: A.-C.U. and D.D.H. Visualization: M.K.A., P.W. and T.B.; Writing, original draft: A.-C.U., M.K.A., H.M. and D.D.H. Writing, review and editing: all authors.

Corresponding authors

Correspondence to David D. Ho or Anne-Catrin Uhlemann.

Ethics declarations

Competing interests

P.W., M.N., Y.H. and D.D.H. are inventors on a provisional patent application on monoclonal antibodies against SARS-CoV-2. D.D.H. is a member of the scientific advisory board of Brii Biosciences, which has provided a grant to Columbia University to support this and other studies on SARS-CoV-2. A.-C.U. and D.D.H. have received funding from Merck & Co. unrelated to this study.

Additional information

Peer review information *Nature* thanks Tulio De Oliveira, Tommy Tsan-Yuk Lam 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 Rapid PCR-based screening assay protocol to identify samples harboring key substitutions.

a, Viral RNA is prepared by heat inactivation and centrifugation. The supernatant is then used for the SNP assay, which entails four steps: the reverse transcription (RT) reaction, pre-PCR reading of the plate to assess background fluorescence (SNP pre-read), real-time PCR, and post-PCR reading of the plate to measure fluorescence (SNP post-read). The runtime for this entire protocol is approximately two hours. **b**, Genotype at targeted sites in COVID-19 viral RNA can be determined with two MGB probes, one for wild type (conjugated with VIC) and the other for variant type (conjugated with FAM). **c**, Example signals for the variant type (K484; blue), the wild type (E484; red) and samples with no signal (black) are shown.

Extended Data Fig. 2 Flowchart for SARS-CoV-2-positive nasopharyngeal swabs included in this study.

Top, during the study period of 1 November 2020 to 1 May 2021, 6,680 patients tested positive for SARS-CoV-2 at our hospital centre and affiliated hospitals. From these 9,174 positive nasopharyngeal swabs, 3,433 swabs were stored as part of the Columbia University Biobank COVID-19 sample repository and available for this study. Left, PCR-based genotyping assays for E484K and N501Y (see Extended Data Fig. 1) were performed on 2,345 samples. We identified a significant proportion of samples with E484K (11%), later confirmed through sequencing to primarily fall within the B.1.526 lineage, and a number of samples with N501Y (3.9%), primarily within the B.1.1.7 lineage. Right, we performed whole-genome sequencing on 1,507 samples. Of these, 32% belonged to B.1.526 and the sublineages B.1.526.1 and B.1.526.2, while B.1.1.7 constituted a much smaller proportion of samples at our centre (7.5%).

Extended Data Fig. 3 Unique patterns of spike gene mutations.

Phylogenetic tree based on whole-genome alignment of genomes sequenced from our hospital centre with at least one mutation of interest or concern (E484K, N501Y, S477N, or L452R) and unique spike protein mutation constellations ($n=64$). Branches are labelled according to Pangolin-assigned lineage identifications. Residues at which at least one sample harboured a mutation are displayed above the S-protein schematic. Residues highlighted in colour represent mutations when compared to the Wuhan-Hu-1 strain.

Extended Data Fig. 4 Neutralization studies of B.1.526-S477N.

a, Neutralizing activities of 12 monoclonal antibodies against pseudoviruses containing S477N alone or all five signature B.1.526-S477N mutations (L5F, T95I, D253G, A701V, and S477N), termed NYΔ5(S477N). Antibodies with emergency use authorization are shown in bold solid lines. Data are represented as mean ± s.e.m. of technical triplicates and represent one of two independent experiments. **b**, Neutralizing activities of convalescent plasma ($n = 20$) against NYΔ5(S477N) as well as against the authentic B.1.526 virus with S477N, and neutralizing activities of vaccinee sera ($n = 22$) against the NYΔ5(S477N) pseudovirus, compared to wild-type counterparts. Statistical comparisons were made using the Wilcoxon matched-pairs signed rank test; two-tailed P -values are reported.

[Extended Data Fig. 5 State-level growth dynamics of B.1.526 and B.1.1.7.](#)

Daily state-level frequencies of B.1.526 (in yellow) and B.1.1.7 (in blue), based on GISAID data downloaded on 6 June 2021, were used to plot 7-day sliding window averages of the prevalence of each lineage (shown as dots in the figure). A 4-parameter multinomial logistic regression model was fit directly to the observation data, in which both B.1.1.7 and B.1.526 have parameters specified for frequency at day 0 (1 January 2021) and logistic growth rate (shown as lines in the figure). States are ordered according to frequency of B.1.526 at the final timepoint of 8 May 2021.

Extended Data Table 1 Clinical characteristics of patients infected with SARS-CoV-2 based on viral genotype

Supplementary information

[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

Annavajhala, M.K., Mohri, H., Wang, P. *et al.* Emergence and expansion of SARS-CoV-2 B.1.526 after identification in New York. *Nature* **597**, 703–708 (2021). <https://doi.org/10.1038/s41586-021-03908-2>

- Received: 07 April 2021
- Accepted: 13 August 2021
- Published: 24 August 2021
- Issue Date: 30 September 2021
- DOI: <https://doi.org/10.1038/s41586-021-03908-2>

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/s41586-021-03908-2>

- Article
- [Published: 08 September 2021](#)

A lymphocyte–microglia–astrocyte axis in chronic active multiple sclerosis

- [Martina Absinta](#) [ORCID: orcid.org/0000-0003-0276-383X](#)^{1,2,3},
- [Dragan Maric](#) [ORCID: orcid.org/0000-0003-2912-7921](#)⁴,
- [Marjan Gharagozloo](#) [ORCID: orcid.org/0000-0002-2166-4012](#)¹,
- [Thomas Garton](#) [ORCID: orcid.org/0000-0001-9352-8530](#)¹,
- [Matthew D. Smith](#)¹,
- [Jing Jin](#) [ORCID: orcid.org/0000-0003-1508-8036](#)¹,
- [Kathryn C. Fitzgerald](#) [ORCID: orcid.org/0000-0003-3137-0322](#)¹,
- [Anya Song](#) [ORCID: orcid.org/0000-0002-0885-7530](#)⁵,
- [Poching Liu](#)⁶,
- [Jing-Ping Lin](#) [ORCID: orcid.org/0000-0003-0686-0215](#)²,
- [Tianxia Wu](#)⁷,
- [Kory R. Johnson](#)⁸,
- [Dorian B. McGavern](#) [ORCID: orcid.org/0000-0001-9568-545X](#)⁹,
- [Dorothy P. Schafer](#) [ORCID: orcid.org/0000-0003-2201-6276](#)⁵,
- [Peter A. Calabresi](#)¹ &
- [Daniel S. Reich](#) [ORCID: orcid.org/0000-0002-2628-4334](#)²

[Nature](#) volume 597, pages 709–714 (2021)

- 10k Accesses
- 194 Altmetric
- [Metrics details](#)

Subjects

- [Complement cascade](#)
- [Multiple sclerosis](#)
- [Neuroimmunology](#)
- [Prognostic markers](#)

Abstract

Multiple sclerosis (MS) lesions that do not resolve in the months after they form harbour ongoing demyelination and axon degeneration, and are identifiable *in vivo* by their paramagnetic rims on MRI scans^{1,2,3}. Here, to define mechanisms underlying this disabling, progressive neurodegenerative state^{4,5,6} and foster development of new therapeutic agents, we used MRI-informed single-nucleus RNA sequencing to profile the edge of demyelinated white matter lesions at various stages of inflammation. We uncovered notable glial and immune cell diversity, especially at the chronically inflamed lesion edge. We define ‘microglia inflamed in MS’ (MIMS) and ‘astrocytes inflamed in MS’, glial phenotypes that demonstrate neurodegenerative programming. The MIMS transcriptional profile overlaps with that of microglia in other neurodegenerative diseases, suggesting that primary and secondary neurodegeneration share common mechanisms and could benefit from similar therapeutic approaches. We identify complement component 1q (C1q) as a critical mediator of MIMS activation, validated immunohistochemically in MS tissue, genetically by microglia-specific C1q ablation in mice with experimental autoimmune encephalomyelitis, and therapeutically by treating chronic experimental autoimmune encephalomyelitis with C1q blockade. C1q inhibition is a potential therapeutic avenue to address chronic white matter inflammation, which could be monitored by longitudinal assessment of its dynamic biomarker, paramagnetic rim lesions, using advanced MRI methods.

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](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

Fig. 1: snRNA-seq demonstrates pathological stage- and location-specific population diversity in MS lesions versus control white matter.

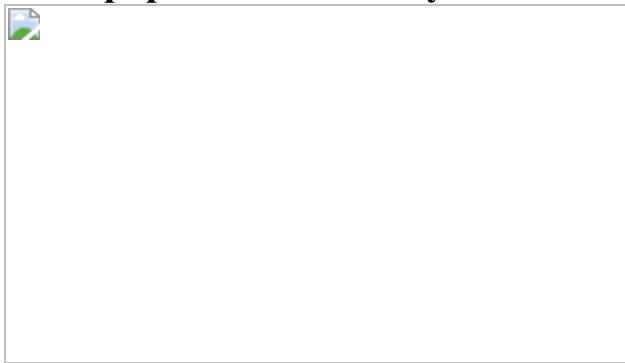


Fig. 2: Two MIMS clusters at the chronic active lesion edge have divergent effector functions.

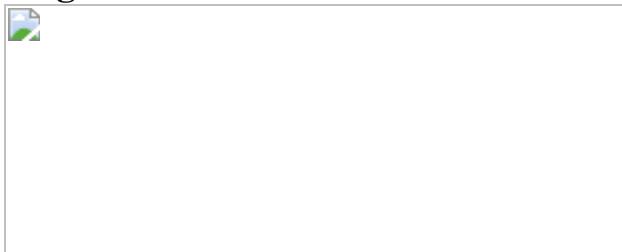


Fig. 3: MIMS–AIMS interactions anchor the glial interactome, and C1q and C3 activity is increased in chronic active white matter lesions.



Data availability

All raw snRNA-seq data (fastq files) were deposited to GEO (<https://www.ncbi.nlm.nih.gov/geo/> under accession number [GSE180759](#)). [Source data](#) are provided with this paper.

Code availability

Codes used to analyse the snRNA-seq data in the current study are available from the corresponding authors upon request.

References

1. 1.

Bagnato, F., et al. Tracking iron in multiple sclerosis: a combined imaging and histopathological study at 7 tesla. *Brain* **134**, 3602–3615 (2011).

2. 2.

Absinta, M., et al. Persistent 7-tesla phase rim predicts poor outcome in new multiple sclerosis patient lesions. *J. Clin. Invest.* **126**, 2597–2609 (2016).

3. 3.

Dal-Bianco, A., et al. Slow expansion of multiple sclerosis iron rim lesions: pathology and 7 T magnetic resonance imaging. *Acta Neuropathol.* **133**, 25–42 (2017).

4. 4.

Absinta, M., et al. Association of chronic active multiple sclerosis lesions with disability in vivo. *JAMA Neurol.* **76**, 1474–1483 (2019).

5. 5.

Elliott, C., et al. Chronic white matter lesion activity predicts clinical progression in primary progressive multiple sclerosis. *Brain* **142**, 2787–2799 (2019).

6. 6.

Maggi, P., et al. Chronic white matter inflammation and neurofilament levels in patients with multiple sclerosis. *Neurology* **97**, e543–e553 (2021).

7. 7.

Kuhlmann, T., et al. An updated histological classification system for multiple sclerosis lesions. *Acta Neuropathol.* **133**, 13–24 (2017).

8. 8.

Reich, D. S., Lucchinetti, C. F. & Calabresi, P. A. Multiple sclerosis. *N. Engl. J. Med.* **378**, 169–180 (2018).

9. 9.

Krishnaswami, S. R., et al. Using single nuclei for RNA-seq to capture the transcriptome of postmortem neurons. *Nat. Protoc.* **11**, 499–524 (2016).

10. 10.

Hendrickx, D. A. E., et al. Gene expression profiling of multiple sclerosis pathology identifies early patterns of demyelination surrounding chronic active lesions. *Front. Immunol.* **8**, 1810 (2017).

11. 11.

Elkjaer, M. L., et al. Molecular signature of different lesion types in the brain white matter of patients with progressive multiple sclerosis. *Acta Neuropathol. Commun.* **7**, 205 (2019).

12. 12.

Jackle, K., et al. Molecular signature of slowly expanding lesions in progressive multiple sclerosis. *Brain* **143**, 2073–2088 (2020).

13. 13.

Schirmer, L., et al. Neuronal vulnerability and multilineage diversity in multiple sclerosis. *Nature* **573**, 75–82 (2019).

14. 14.

Maric, D., et al. Whole-brain tissue mapping toolkit using large-scale highly multiplexed immunofluorescence imaging and deep neural networks. *Nat. Commun.* **12**, 1550 (2021).

15. 15.

Comabella, M., Montalban, X., Munz, C. & Lunemann, J. D. Targeting dendritic cells to treat multiple sclerosis. *Nat. Rev. Neurol.* **6**, 499–507 (2010).

16. 16.

Clayton, B. L. L. & Popko, B. Endoplasmic reticulum stress and the unfolded protein response in disorders of myelinating glia. *Brain Res.* **1648**, 594–602 (2016).

17. 17.

Krasemann, S., et al. The TREM2-APOE pathway drives the transcriptional phenotype of dysfunctional microglia in neurodegenerative diseases. *Immunity* **47**, 566–581.e569 (2017).

18. 18.

Keren-Shaul, H., et al. A unique microglia type associated with restricting development of Alzheimer's disease. *Cell* **169**, 1276–1290.e1217 (2017).

19. 19.

Masuda, T., et al. Spatial and temporal heterogeneity of mouse and human microglia at single-cell resolution. *Nature* **566**, 388–392 (2019).

20. 20.

Sankowski, R., et al. Mapping microglia states in the human brain through the integration of high-dimensional techniques. *Nat. Neurosci.* **22**, 2098–2110 (2019).

21. 21.

Jakel, S., et al. Altered human oligodendrocyte heterogeneity in multiple sclerosis. *Nature* **566**, 543–547 (2019).

22. 22.

Miron, V. E., et al. M2 microglia and macrophages drive oligodendrocyte differentiation during CNS remyelination. *Nat. Neurosci.* **16**, 1211–1218 (2013).

23. 23.

Cignarella, F., et al. TREM2 activation on microglia promotes myelin debris clearance and remyelination in a model of multiple sclerosis. *Acta Neuropathol.* **140**, 513–534 (2020).

24. 24.

Breij, E. C., et al. Homogeneity of active demyelinating lesions in established multiple sclerosis. *Ann. Neurol.* **63**, 16–25 (2008).

25. 25.

van der Poel, M., Hoepel, W., Hamann, J., Huitinga, I. & Dunnen, J. D. IgG immune complexes break immune tolerance of human microglia. *J. Immunol.* **205**, 2511–2518 (2020).

26. 26.

Browaeys, R., Saelens, W. & Saeys, Y. NicheNet: modeling intercellular communication by linking ligands to target genes. *Nat. Methods* **17**, 159–162 (2020).

27. 27.

Rusinova, I., et al. Interferome v2.0: an updated database of annotated interferon-regulated genes. *Nucleic Acids Res.* **41**, D1040–D1046 (2013).

28. 28.

Kirby, L., et al. Oligodendrocyte precursor cells present antigen and are cytotoxic targets in inflammatory demyelination. *Nat. Commun.* **10**, 3887 (2019).

29. 29.

Falcao, A. M., et al. Disease-specific oligodendrocyte lineage cells arise in multiple sclerosis. *Nat. Med.* **24**, 1837–1844 (2018).

30. 30.

Liddelow, S. A., et al. Neurotoxic reactive astrocytes are induced by activated microglia. *Nature* **541**, 481–487 (2017).

31. 31.

Zhang, B. & Horvath, S. A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.* **4**, 17 (2005).

32. 32.

Dal-Bianco, A., et al. Long-term evolution of multiple sclerosis iron rim lesions in 7 T MRI. *Brain* **144**, 833–847 (2021).

33. 33.

International Multiple Sclerosis Genetics Consortium. Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. *Science* **365**, eaav7188 (2019).

34. 34.

Fitzgerald, K. C., et al. Early complement genes are associated with visual system degeneration in multiple sclerosis. *Brain* **142**, 2722–2736 (2019).

35. 35.

Kaiser, T. & Feng, G. Tmem119-EGFP and Tmem119-CreERT2 transgenic mice for labeling and manipulating microglia. *eNeuro* **6**, 0448–18.2019 (2019).

36. 36.

Werneburg, S., et al. Targeted complement inhibition at synapses prevents microglial synaptic engulfment and synapse loss in demyelinating disease. *Immunity* **52**, 167–182.e167 (2020).

37. 37.

Fonseca, M. I., et al. Cell-specific deletion of C1qa identifies microglia as the dominant source of C1q in mouse brain. *J. Neuroinflammation* **14**, 48 (2017).

38. 38.

Stephan, A. H., et al. A dramatic increase of C1q protein in the CNS during normal aging. *J. Neurosci.* **33**, 13460–13474 (2013).

39. 39.

Michailidou, I., et al. Complement C1q–C3-associated synaptic changes in multiple sclerosis hippocampus. *Ann. Neurol.* **77**, 1007–1026 (2015).

40. 40.

van der Valk, P. & De Groot, C. J. Staging of multiple sclerosis (MS) lesions: pathology of the time frame of MS. *Neuropathol. Appl. Neurobiol.* **26**, 2–10 (2000).

41. 41.

Matson, K. J. E., et al. Isolation of adult spinal cord nuclei for massively parallel single-nucleus RNA sequencing. *J. Vis. Exp.* **140**, 58413 (2018).

42. 42.

Butler, A., Hoffman, P., Smibert, P., Papalex, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**, 411–420 (2018).

43. 43.

Stuart, T., et al. Comprehensive integration of single-cell data. *Cell* **177**, 1888–1902.e1821 (2019).

44. 44.

Korsunsky, I., et al. Fast, sensitive and accurate integration of single-cell data with Harmony. *Nat. Methods* **16**, 1289–1296 (2019).

45. 45.

Fabregat, A., et al. The Reactome Pathway Knowledgebase. *Nucleic Acids Res.* **46**, D649–D655 (2018).

46. 46.

Raudvere, U., et al. gProfiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res.* **47**, W191–W198 (2019).

47. 47.

Zhou, Y., et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523 (2019).

48. 48.

Bonnardel, J., et al. Stellate cells, hepatocytes, and endothelial cells imprint the Kupffer cell identity on monocytes colonizing the liver macrophage niche. *Immunity* **51**, 638–654.e639 (2019).

49. 49.

Krzywinski, M., et al. Circos: an information aesthetic for comparative genomics. *Genome Res.* **19**, 1639–1645 (2009).

50. 50.

Schindelin, J., et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).

51. 51.

Absinta, M., et al. Postmortem magnetic resonance imaging to guide the pathologic cut: individualized, 3-dimensionally printed cutting boxes for fixed brains. *J. Neuropathol. Exp. Neurol.* **73**, 780–788 (2014).

52. 52.

Polman, C. H., et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Ann. Neurol.* **69**, 292–302 (2011).

53. 53.

Absinta, M., et al. Identification of chronic active multiple sclerosis lesions on 3T MRI. *Am. J. Neuroradiol.* **39**, 1233–1238 (2018).

54. 54.

Yao, B., et al. Chronic multiple sclerosis lesions: characterization with high-field-strength MR imaging. *Radiology* **262**, 206–215 (2012).

55. 55.

Sati, P., et al. Rapid, high-resolution, whole-brain, susceptibility-based MRI of multiple sclerosis. *Mult. Scler.* **20**, 1464–1470 (2014).

Acknowledgements

This study was supported by the Intramural Research Program of NINDS, the Adelson Medical Research Foundation (to D.S.R. and D.P.S.), the Conrad N. Hilton Foundation (grant no. 17313 to M.A.) and the Cariplo Foundation (grant no. 2019-1677 to M.A.). We thank Y. Li and Y. Luo from the DNA Sequencing and Genomics Core (NHLBI/NIH); G. Trivellin (NICHD/NIH), who helped with some technical aspects of single nucleus RNA-seq processing; C. Covino from the Advanced Light and Electron Microscopy BioImaging Center (ALEMBIC, Vita-Salute San Raffaele University, Milan), who helped with visualization of multiplex immunofluorescence; T. Yednock and V. Mehta at Annexon Biosciences for providing the anti-C1q monoclonal antibody (ANX-

M1.21); G. Nair, P. Sati, E. S. Beck, M. Donadieu, the NINDS Translational Neuroradiology Section and the NIMH Functional MRI Facility for assistance with MRI scanning; I. Cortese, J. Ohayon and the NINDS Neuroimmunology Clinic for outstanding patient care; and D. Bergles, R. Franklin, J. Kipnis and N. Mihelson for fruitful discussion of the results.

Author information

Affiliations

1. Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD, USA

Martina Absinta, Marjan Gharagozloo, Thomas Garton, Matthew D. Smith, Jing Jin, Kathryn C. Fitzgerald & Peter A. Calabresi

2. Translational Neuroradiology Section, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD, USA

Martina Absinta, Jing-Ping Lin & Daniel S. Reich

3. IRCCS San Raffaele Hospital and Vita-Salute San Raffaele University, Milan, Italy

Martina Absinta

4. Flow and Imaging Cytometry Core Facility, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD, USA

Dragan Maric

5. Department of Neurobiology and the Brudnik Neuropsychiatry Institute, University of Massachusetts Medical School, Worcester, MA, USA

Anya Song & Dorothy P. Schafer

6. DNA Sequencing and Genomics Core, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD, USA

Poching Liu

7. Clinical Trials Unit, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD, USA

Tianxia Wu

8. Bioinformatics Section, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD, USA

Kory R. Johnson

9. Viral Immunology and Intravital Imaging Section, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD, USA

Dorian B. McGavern

Contributions

M.A., J.-P.L. and P.L. performed the nucleic extraction and snRNA-seq experiments; M.A. and K.R.J. performed the bioinformatic analysis; M.A. and D.M. performed the immunostainings and multiplex immunofluorescence; A.S., M.G., T.G., M.D.S., J.J., D.P.S. and P.A.C. performed the animal studies; M.A., D.S.R. and T.W. performed the MRI studies; P.A.C. and K.C.F. performed the genotyping study; M.A., D.S.R., P.A.C., D.P.S. and D.B.M. discussed and interpreted findings; M.A. and D.S.R. designed the study and wrote the manuscript. All co-authors read, revised and approved the manuscript.

Corresponding authors

Correspondence to Martina Absinta or Daniel S. Reich.

Ethics declarations

Competing interests

The authors declare no competing interests. M.A. received consulting fees from Sanofi, unrelated to this study. P.A.C. received grant support from Annexon Biosciences for testing the anti-C1q-blocking antibody in EAE. The funders of the study had no role in the collection, analysis, or interpretation of data, in the writing of the manuscript, or in the decision to submit the paper for publication.

Additional information

Peer review information *Nature* thanks Oleg Butovsky 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 Chronic active/slowly expanding rim lesions can be visualized with MRI in vivo.](#)

a, Susceptibility-based axial MRI at three levels showing the chronic active MS lesions with paramagnetic rims (magnified view in the insets, arrows) of a 38-year-old man with progressive MS, who requires assistance to walk 100 meters without resting. Scale bar = 10 mm. **b**, MRI-pathology of a frontal periventricular chronic active/slowly expanding lesion in a man with progressive MS who died at age 59. On the serial *in vivo* T1-weighted coronal images (clinical MR images sensitive to both demyelination and axonal loss), the rim lesion clearly expanded over a period of 7 years. The lesion was classified as chronic active by histological analysis. Accumulation of iron-laden phagocytes (CD68 and Turnbull iron staining)

was seen at the lesion edge. Smouldering demyelination can also be inferred by the co-presence of early (LFB+, blue) and late (PAS+, purple) myelin degradation products within phagocytes at the lesion edge. Most reactive GFAP+ astrocytes do not contain iron. Magnified views are shown in the insets. Scale bar = 20 μ m (myelin PLP); 10 μ m (LFB-PAS); 50 μ m (CD68, Turnbull iron/CD68, Turnbull iron/GFAP). **c**, In vivo long-term evolution of PRLs: representative examples of persistent/stable, faded, and disappeared rims. **d**, Bar graph showing the evolution of PRLs in each MS case followed over time. All cases had >5 years of yearly MRI follow up, except for cases 2–4, where follow-up was between 3.5 and 5 years. **e**, Survival analysis showed that the median PRL survival time is about 7 years.

Extended Data Fig. 2 Cell cluster proportions vary across tissue types.

a, Bar graphs showing the distribution of nuclei percentages by cell type (mean, standard error). See Fig. 1 for corresponding UMAP plots and distribution of nuclei count per sample. **b**, Bar graph showing the percentage of MIMS-foamy and MIMS-iron nuclei for each sample and location. WM, white matter; CA, chronic active; CI, chronic inactive; MIMS, microglia inflamed in multiple sclerosis; OPC, oligodendrocyte progenitor cells; mono/moDC, monocytes and monocyte-derived dendritic cells; oligo, oligodendrocytes; ox, oxidative stress.

Extended Data Fig. 3 Subclustering markers for astrocytes, oligodendrocytes and OPC.

a–c, Subclustering UMAP plots for astrocytes, oligodendrocytes and OPC, and associated stacked violin plots depicting selected differentially expressed genes for each subcluster. **d**, Bar graph showing A1-marker³⁰ expression (Z-scores relative to all astrocytes) for each astrocyte cluster. Most A1 markers were highly expressed in AIMS. UMAP, uniform manifold approximation and projection; OPC, oligodendrocyte progenitor cells; oligo, oligodendrocytes; AIMS, astrocytes inflamed in MS.

Extended Data Fig. 4 Lymphocytes, monocyte-derived dendritic cells, MIMS and AIMS are identified at the chronic active lesion edge.

Chronic active lesion overview. Multiplex immunostaining of a brain tissue block from a 48-year-old woman with progressive MS (case MS4 in Supplementary Table 1), including a chronic active demyelinated lesion (devoid of myelin proteolipid protein (PLP) staining). Numbers indicate areas magnified in subsequent panels for validation of immune cell and inflamed astrocytes and represent the chronic active lesion edge, lesion centre and periplaque white matter. Lymphocytes and microglia at the lesion edge (panels 1 and 2). At the chronic active lesion edge, there are groups of CD8 T cells within the perivascular space and sparsely within the parenchyma (arrows). CD20 B cells are fewer than CD8 T cells and are located prevalently within the perivascular space (dashed arrows).

Transition from myelination to demyelination is shown with staining for CNPase, an oligodendrocyte and myelin marker. Residual myelin fragments can be seen at the edge (arrowheads), presumably not yet removed by phagocytes. IBA1-microglia/macrophages are frequent, and they have an activated morphology (round shape without ramifications). Bar graphs (second row) show the gene expression Z-scores of markers implemented for the identification in tissue of the most relevant glial cell populations at the chronic active lesion edge. Homeostatic microglia vs. MIMS: different spatial locations (insets 3–5). In the periplaque white matter, most microglia are P2RY12+ (a homeostatic marker) with short and thick processes, whereas at the chronic active lesion edge, most are CD68+ (indicating upregulation of antigen and lipid processing) with round, activated morphology. At the lesion core, fewer microglia can be identified, and these show a round morphology consistent with activation. Interestingly, some of them are P2RY12+, potentially suggesting the return of some homeostatic markers. MIMS-iron (inset 6). Accumulation of iron-laden phagocytes (CD68 and ferritin light chain, FTL) is seen at the lesion edge. FTL is within the first 100 top differentially expressed genes in MIMS, especially in MIMS-iron (Fig. 2c). Iron retention in phagocytes can be seen by MRI at the chronic active lesion edge as a paramagnetic rim (MRI biomarker; Fig. 1a). MIMS-foamy (inset 7). Colocalization of PPARG and CD68. PPARG and CD68 double positive (white, arrows) microglia are especially seen at

the lesion edge, suggesting their involvement in energy metabolism and modulation of inflammation as well as clearance of myelin debris. Monocytes/monocyte-derived dendritic cells and MIMS (inset 8). CD68 microglia outnumber CD83 monocyte-derived mature dendritic cells (arrows) at the chronic active lesion edge. AIMS and MIMS (insets 9, 10a, and 10b). In addition to IBA1+/CD68+ microglia, the lesion edge is enriched for inflamed astrocytes (positive for VIM and APOE but negative for IBA1 and CD68), sometimes in close proximity (dashed white box). Compared to activated microglia, inflamed astrocytes are bigger and show radial processes. moDC, monocyte-derived dendritic cells; MIMS, microglia inflamed in MS; AIMS, astrocytes inflamed in MS. Scale bar: 20 μ m.

Extended Data Fig. 5 Pathway analysis of the two MIMS populations and comparison with DAM.

a, b, Gene enrichment and pathway analysis of the top 100 differentially expressed genes are shown for MIMS-foamy and MIMS-iron populations, separately. Using the gProfileR package, different sources were compared (GO:BP, GO:CC, GO:MF, KEGG). Only significant pathway terms are plotted in the graphs ($p < 0.05$, correction methods "g_SCS"). See the text for interpretation. **c,** Homeostatic microglia vs. MIMS vs. DAM: direct comparison of two single-cell RNA-seq datasets (MS [current work] and a mouse AD model [5XFAD]). Data for the AD model are derived from supplementary table 2 of ref. [18](#). In each dataset, only significant differentially expressed genes in the comparison between homeostatic vs. MIMS and homeostatic vs. DAM were retained ($p < 0.001$). Volcano plots report overlapping genes for each microglia phenotype. MS, multiple sclerosis; MIMS, microglia inflamed in MS; AD, Alzheimer disease; DAM, degenerative disease-associated microglia.

Extended Data Fig. 6 Two MIMS populations were identified through re-analysis of two published snRNA-seq MS datasets.

a, snRNA-seq initial mapping (UMAP plot) and annotations based on the top differentially expressed genes in each cluster after re-analysis of raw

data from Schirmer et al. ¹³ and Jäkel et al. ²¹. **b**, Immune cell nuclei were mapped onto our immune subclustering map. Most of the immune cell populations identified were seen also in the other datasets, including the two MIMS populations. UMAP, uniform manifold approximation and projection; MIMS, microglia inflamed in MS.

Extended Data Fig. 7 Multiplex immunostaining of human chronic active lesions.

a, Overview of the multiplex immunostaining method (see the text for details). **b**, A representative example of the chronic active edge and myeloid subpopulations. In the magnified view, identification of MIMS-foamy, MIMS-iron, and macrophages using 7 primary antibodies (combined or separate channels). **c**, Within the lesion core, most cells (both myeloid cells and astrocytes) are positive for the senescence marker p16^{INK4a}. Separate channels are shown to facilitate the visualization of different markers. **d**, Quantification of the proportion of p16-positive nuclei at different locations. The lesion core showed the higher percentage of p16-positive nuclei (ANOVA $p < 0.0001$, Tukey's multiple comparison post-hoc analysis * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). Scale bar: 20 μ m. Blue: DAPI (nuclei).

Extended Data Fig. 8 MIMS-AIMS gene modules and correlation.

a, b, Hierarchical clustering dendrograms of genes (**a**) and module colours (**b**) based on weighted correlation network analysis (WGCNA) of 918 variable genes from immune cells and astrocyte clusters 5, 9, 10. We identified 2 MIMS gene modules (pink: C1QB, CD74, CEBPD, HLA-DRA, ITM2B, RPS19; black: ACTB, APOE, CD81, EEF1A1, FTH1, FTL, PSAP) and an AIMS gene module (magenta: CALM1, CLU, CRYAB, CST3, GAPDH, GFAP). The complete list of modules is shown in Supplementary Table 9. **c**, Heatmap showing the eigengene adjacency matrix that represents the relationships among the identified gene modules. MIMS gene modules (pink and black) were highly correlated with other myeloid gene modules and with the AIMS module, but not with the

nonreactive astrocyte module. AST, astrocytes; IMM, immune cells; MIMS, microglia inflamed in MS; AIMS, astrocytes inflamed in MS; ME, module eigengene; MT, mitochondrial genes.

Extended Data Fig. 9 Mapping complement and MS susceptibility genes onto the snRNA-seq dataset.

a, Heatmap showing the expression of genes involved in the classical complement cascade, including complement genes, receptors, and regulators. Z-scores are relative to all cells. Most of these genes map onto immune cell, astrocyte, and vascular cell populations. Cluster identifiers correspond to Fig. 1c, d. **b**, For each cell population, the number of complement cascade-related genes with Z-score >1.5 is plotted on the UMAP. **c**, Heatmaps showing the expression of genes involved in the classical complement cascade for the immune cell (left, pink) and astrocyte (right, green) subclusters. Z-scores are relative to immune cells and astrocytes, respectively. Among microglia, C1Q, C3 and CFD (C3 activator) were expressed mainly by MIMS-iron, whereas negative regulators of complement activation (LAIR1, LAIR2, CR1) were expressed mainly by homeostatic microglia and perivascular macrophages. Among astrocytes, C1 complex (C1Q, C1R, C1S, C1QBP), C3, C4 and CALR were expressed mainly by AIMS, whereas C2 and C6 were expressed by some other reactive astrocytes. **d**, Mapping MS-susceptibility risk genes onto the snRNA-seq dataset: the list of 558 prioritized MS susceptibility genes was obtained from a recent genome-wide association study (GWAS)³³ and mapped onto all three snRNA-seq datasets (the current dataset as well as that of Schirmer et al.¹³ and Jäkel et al.²¹). Low-expressed genes (within the 25th percentile average expression) were excluded. MS susceptibility genes were then assigned to clusters if the z-scored average gene expression was >2 . Clusters were classified based on the number of MS susceptibility genes (>80 , $50-80$, $30-50$, $1-30$ and <10 genes). Results were colour-mapped onto each snRNA-seq UMAP. Most MS susceptibility genes mapped onto the immune and vascular cell clusters. Interestingly, an excitatory neuronal population (“py”) expressed some MS susceptibility genes as well. oligo, oligodendrocytes; imm, immune cells; OPC, oligodendrocyte progenitor cells; lym, lymphocytes; ast, astrocytes; endo, endothelial cells; peri, pericytes; neu, neurons; MIMS, microglia inflamed

in MS; mono/moDC, monocytes and monocyte-derived dendritic cells; AIMS, astrocytes inflamed in MS; exc, excitatory; in, inhibitory.

Extended Data Fig. 10 C1q mediates microglia activation in mouse EAE.

Iba1+ cells in microglia-specific C1q cKO with EAE appear less reactive. **a**, Visual thalamus was immunostained at PID12–14 for Iba1 (microglia/macrophages) and Clec7a (disease-associated microglia). Yellow arrows denote Iba1+Clec7a+ cells, white arrows Iba1+Clec7a- cells. Scale bar = 100 μ m. **b, c**, Clec7a decreased in cKO mice with EAE compared to Ctrl-EAE littermates. The density of Iba1+ cells and Iba1 MFI was attenuated to control values in cKO-EAE mice. * $p < 0.05$, by one-way ANOVA and Tukey's posthoc test (**b**) or Kruskal-Wallis test and Dunn posthoc test (**c**). **d**, Iba1+ cells were morphologically characterized into 3 categories (representative images in the panel below the quantification, scale bar = 20 μ m). Iba1+ cells in cKO-EAE mice were indistinguishable from CFA controls. cKO-EAE showed more ramified Iba1+ cells and fewer amoeboid cells compared to Ctrl-EAE littermates. ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$ by two-way ANOVA and Tukey's posthoc test. Error bars: SEM. Anti-C1q treatment reduces expression of FTL and Iba1 in chronic EAE. **e**, Experimental paradigm (twice weekly treatment with isotype control or C1q-blocking antibody (M1.21) from EAE onset until PID42). **f**, Representative images of FTL and Iba1 immunostaining. **g, h**, Higher magnifications of **f**. White arrows: Iba1+ cells. **i**, EAE scores for each treatment arm. **j, k**, Quantification of the expression of C1q, FTL and Iba1 in hippocampal white matter (outlined with the dashed line in **f**) and count of Iba1+FTL+ cells. Student t-test, $p^* \leq 0.05$, *** ≤ 0.001 . Error bars: SEM. TAM, tamoxifen; FTL, ferritin light chain; MFI, mean fluorescent intensity; PID, post-immunization day; SEM, standard error of the mean.

[Source data](#)

Supplementary information

[Supplementary Information](#)

This file contains supplementary text and supplementary references

Reporting Summary

Peer Review File

Supplementary Tables 1–13

This file contains Supplementary Tables 1–13 and their accompanying legends.

Source data

Source Data Extended Data Fig. 10

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Absinta, M., Maric, D., Gharagozloo, M. *et al.* A lymphocyte–microglia–astrocyte axis in chronic active multiple sclerosis. *Nature* **597**, 709–714 (2021). <https://doi.org/10.1038/s41586-021-03892-7>

- Received: 30 November 2020
- Accepted: 10 August 2021
- Published: 08 September 2021
- Issue Date: 30 September 2021
- DOI: <https://doi.org/10.1038/s41586-021-03892-7>

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/s41586-021-03892-7>

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

- Article
- [Published: 15 September 2021](#)

Acinar cell clonal expansion in pancreas homeostasis and carcinogenesis

- [Patrick Neuhöfer](#) ORCID: [orcid.org/0000-0003-3936-383X^{1,2,3}](https://orcid.org/0000-0003-3936-383X),
- [Caitlin M. Roake](#)^{1,2,3},
- [Stewart J. Kim](#) ORCID: [orcid.org/0000-0001-7196-0541^{1,2,3}](https://orcid.org/0000-0001-7196-0541),
- [Ryan J. Lu](#)^{1,2,3},
- [Robert B. West](#)⁴,
- [Gregory W. Charville](#) ORCID: [orcid.org/0000-0002-2774-2704⁴](https://orcid.org/0000-0002-2774-2704) &
- [Steven E. Artandi](#) ORCID: [orcid.org/0000-0001-5580-7512^{1,2,3}](https://orcid.org/0000-0001-5580-7512)

[Nature](#) volume 597, pages 715–719 (2021)

- 4587 Accesses
- 52 Altmetric
- [Metrics details](#)

Subjects

- [Pancreatic cancer](#)
- [Regeneration](#)

Abstract

Pancreatic ductal adenocarcinoma (PDAC) is one of the leading causes of cancer deaths worldwide¹. Studies in human tissues and in mouse models have suggested that for many cancers, stem cells sustain early mutations driving tumour development^{2,3}. For the pancreas, however, mechanisms underlying cellular renewal and initiation of PDAC remain unresolved. Here, using lineage tracing from the endogenous telomerase reverse transcriptase (*Tert*) locus, we identify a rare TERT-positive subpopulation of pancreatic acinar cells dispersed throughout the exocrine compartment. During homeostasis, these TERT^{high} acinar cells renew the pancreas by forming expanding clones of acinar cells, whereas randomly marked acinar cells do not form these clones. Specific expression of mutant *Kras* in TERT^{high} acinar cells accelerates acinar clone formation and causes transdifferentiation to ductal pre-invasive pancreatic intraepithelial neoplasms by upregulating Ras–MAPK signalling and activating the downstream kinase ERK (phospho-ERK). In resected human pancreatic neoplasms, we find that foci of phospho-ERK-positive acinar cells are common and frequently contain activating *KRAS* mutations, suggesting that these acinar regions represent an early cancer precursor lesion. These data support a model in which rare TERT^{high} acinar cells may sustain *KRAS* mutations, driving acinar cell expansion and creating a field of aberrant cells initiating pancreatic tumorigenesis.

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](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

Fig. 1: Identification of an acinar cell subpopulation with elevated *Tert* expression.

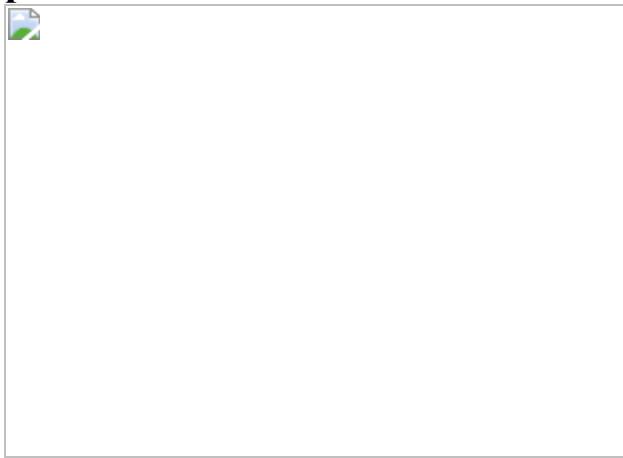


Fig. 2: TERT^{high} acinar cells repopulate the pancreas during homeostasis and regeneration.

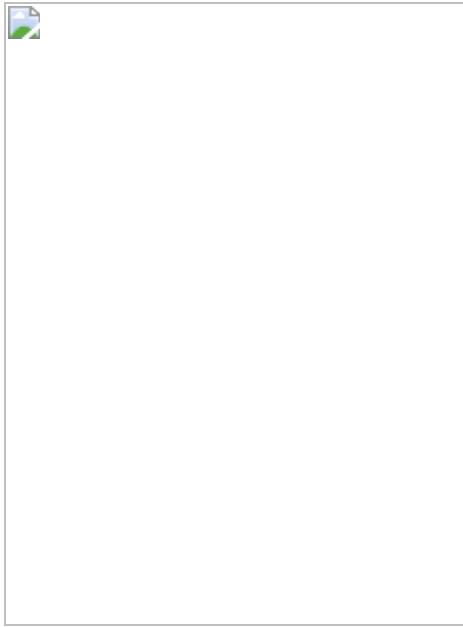


Fig. 3: TERT^{high} acinar cells initiate pancreatic neoplasia upon Kras activation and injury.

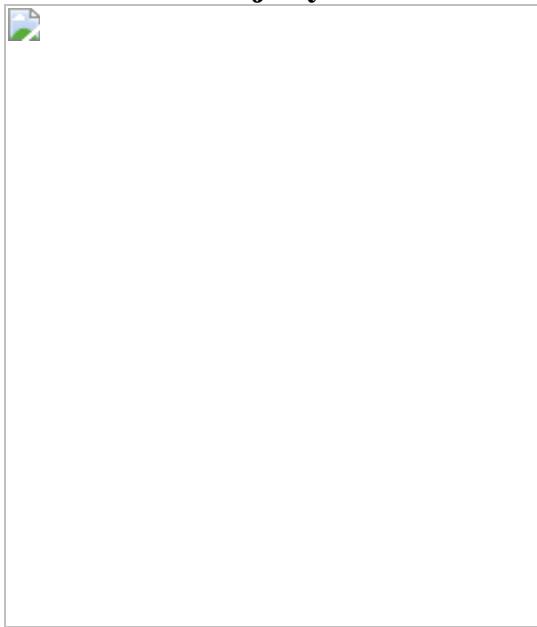
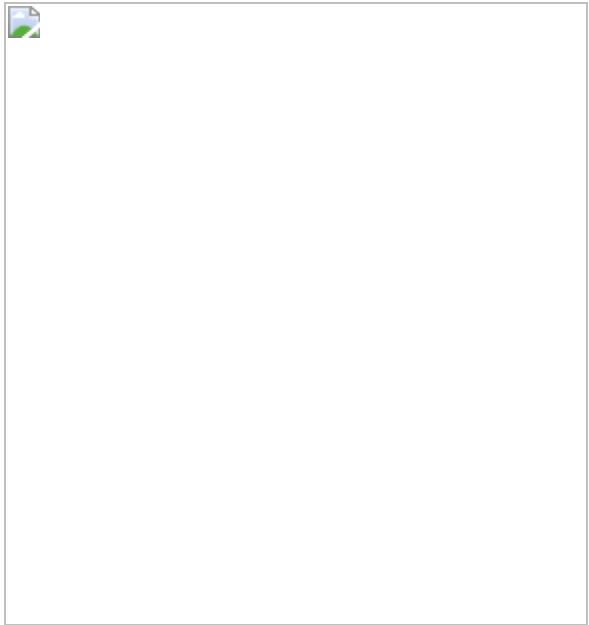


Fig. 4: Identification of pERK⁺ human acinar cells harbouring *KRAS* mutations.



Data availability

Sequencing data for *KRAS* variants is available with the SRA accession: [PRJNA598774](#). All other data are available from the authors upon reasonable request. [Source data](#) are provided with this paper.

Code availability

Custom Python scripts used in this manuscript are available at (https://github.com/cmroake/KRAS_scripts).

References

1. 1.

Rawla, P., Sunkara, T. & Gaduputi, V. Epidemiology of pancreatic cancer: global trends, etiology and risk factors. *World J. Oncol.* **10**, 10–27 (2019).

2. 2.

Jan, M. et al. Clonal evolution of preleukemic hematopoietic stem cells precedes human acute myeloid leukemia. *Sci. Transl. Med.* **4**, 149ra118 (2012).

3. 3.

Barker, N. et al. Crypt stem cells as the cells-of-origin of intestinal cancer. *Nature* **457**, 608–611 (2009).

4. 4.

Murtaugh, L. C. & Keefe, M. D. Regeneration and repair of the exocrine pancreas. *Annu. Rev. Physiol.* **77**, 229–249 (2015).

5. 5.

Desai, B. M. et al. Preexisting pancreatic acinar cells contribute to acinar cell, but not islet β cell, regeneration. *J. Clin. Invest.* **117**, 971–977 (2007).

6. 6.

Westphalen, C. B. et al. Dclk1 defines quiescent pancreatic progenitors that promote injury-induced regeneration and tumorigenesis. *Cell Stem Cell* **18**, 441–455 (2016).

7. 7.

Sangiorgi, E. & Capecchi, M. R. Bmi1 lineage tracing identifies a self-renewing pancreatic acinar cell subpopulation capable of maintaining pancreatic organ homeostasis. *Proc. Natl Acad. Sci. USA* **106**, 7101–7106 (2009).

8. 8.

Hruban, R. H., Goggins, M., Parsons, J. & Kern, S. E. Progression model for pancreatic cancer. *Clin. Cancer Res.* **6**, 2969–2972 (2000).

9. 9.

Kopp, J. L. et al. Identification of Sox9-dependent acinar-to-ductal reprogramming as the principal mechanism for initiation of pancreatic ductal adenocarcinoma. *Cancer Cell* **22**, 737–750 (2012).

10. 10.

Habbe, N. et al. Spontaneous induction of murine pancreatic intraepithelial neoplasia (mPanIN) by acinar cell targeting of oncogenic Kras in adult mice. *Proc. Natl Acad. Sci. USA* **105**, 18913–18918 (2008).

11. 11.

De La, O. J. et al. Notch and Kras reprogram pancreatic acinar cells to ductal intraepithelial neoplasia. *Proc. Natl Acad. Sci. USA* **105**, 18907–18912 (2008).

12. 12.

Flowers, B. M. et al. Cell of origin influences pancreatic cancer subtype. *Cancer Discov.* **11**, 660–677 (2021).

13. 13.

Boj, S. F. et al. Organoid models of human and mouse ductal pancreatic cancer. *Cell* **160**, 324–338 (2015).

14. 14.

Lee, J. et al. Reconstituting development of pancreatic intraepithelial neoplasia from primary human pancreas duct cells. *Nat. Commun.* **8**, 14686 (2017).

15. 15.

Kanda, M. et al. Presence of somatic mutations in most early-stage pancreatic intraepithelial neoplasia. *Gastroenterology* **142**, 730–733.e739 (2012).

16. 16.

Shi, C. et al. *KRAS2* mutations in human pancreatic acinar-ductal metaplastic lesions are limited to those with PanIN: implications for the human pancreatic cancer cell of origin. *Mol. Cancer Res.* **7**, 230–236 (2009).

17. 17.

Jones, S. et al. Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. *Science* **321**, 1801–1806 (2008).

18. 18.

Hingorani, S. R. et al. Preinvasive and invasive ductal pancreatic cancer and its early detection in the mouse. *Cancer Cell* **4**, 437–450 (2003).

19. 19.

Ying, H. et al. Genetics and biology of pancreatic ductal adenocarcinoma. *Genes Dev.* **30**, 355–385 (2016).

20. 20.

Roake, C. M. & Artandi, S. E. Regulation of human telomerase in homeostasis and disease. *Nat. Rev. Mol. Cell Biol.* **21**, 384–397 (2020).

21. 21.

Pech, M. F. et al. High telomerase is a hallmark of undifferentiated spermatogonia and is required for maintenance of male germline stem cells. *Genes Dev.* **29**, 2420–2434 (2015).

22. 22.

Petersen, G. M. et al. A genome-wide association study identifies pancreatic cancer susceptibility loci on chromosomes 13q22.1, 1q32.1 and 5p15.33. *Nat. Genet.* **42**, 224–228 (2010).

23. 23.

Lin, S. et al. Distributed hepatocytes expressing telomerase repopulate the liver in homeostasis and injury. *Nature* **556**, 244–248 (2018).

24. 24.

Morris, J. P. T., Cano, D. A., Sekine, S., Wang, S. C. & Hebrok, M. Beta-catenin blocks Kras-dependent reprogramming of acini into pancreatic cancer precursor lesions in mice. *J. Clin. Invest.* **120**, 508–520 (2010).

25. 25.

Pylayeva-Gupta, Y., Grabocka, E. & Bar-Sagi, D. RAS oncogenes: weaving a tumorigenic web. *Nat. Rev. Cancer* **11**, 761–774 (2011).

26. 26.

Collins, M. A., Yan, W., Sebolt-Leopold, J. S. & Pasca di Magliano, M. MAPK signaling is required for dedifferentiation of acinar cells and development of pancreatic intraepithelial neoplasia in mice. *Gastroenterology* **146**, 822–834.e827 (2014).

27. 27.

Bryant, K. L., Mancias, J. D., Kimmelman, A. C. & Der, C. J. KRAS: feeding pancreatic cancer proliferation. *Trends Biochem. Sci.* **39**, 91–100 (2014).

28. 28.

Martincorena, I. et al. Somatic mutant clones colonize the human esophagus with age. *Science* **362**, 911–917 (2018).

29. 29.

Martincorena, I. et al. Tumor evolution. High burden and pervasive positive selection of somatic mutations in normal human skin. *Science*

348, 880–886 (2015).

30. 30.

Busque, L. et al. Recurrent somatic TET2 mutations in normal elderly individuals with clonal hematopoiesis. *Nat. Genet.* **44**, 1179–1181 (2012).

31. 31.

Madisen, L. et al. A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat. Neurosci.* **13**, 133–140 (2010).

32. 32.

Jackson, E. L. et al. Analysis of lung tumor initiation and progression using conditional expression of oncogenic K-ras. *Genes Dev.* **15**, 3243–3248 (2001).

33. 33.

Kopinke, D. et al. Ongoing Notch signaling maintains phenotypic fidelity in the adult exocrine pancreas. *Dev. Biol.* **362**, 57–64 (2012).

34. 34.

Hegyi, P. et al. l-arginine-induced experimental pancreatitis. *World J. Gastroenterol.* **10**, 2003–2009 (2004).

Acknowledgements

This work was supported by grants from the NIH (AG056575, CA244114 and CA197563; S.E.A.), the Emerson Collective (S.E.A.), anonymous support (S.E.A.), the DFG (NE 2006/1-1; P.N.), California TRDRP (25FT-0011; P.N.) and the California Institute for Regenerative Medicine (R.J.L.). We thank members of the Artandi laboratory for critical comments and P. Chu from the Stanford Human Pathology/Histology Service Center for

technical assistance, and acknowledge the Stanford Functional Genomics Facility and Stanford Shared FACS Facility.

Author information

Affiliations

1. Stanford Cancer Institute, Stanford University School of Medicine, Stanford, CA, USA

Patrick Neuhöfer, Caitlin M. Roake, Stewart J. Kim, Ryan J. Lu & Steven E. Artandi

2. Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA

Patrick Neuhöfer, Caitlin M. Roake, Stewart J. Kim, Ryan J. Lu & Steven E. Artandi

3. Department of Biochemistry, Stanford University School of Medicine, Stanford, CA, USA

Patrick Neuhöfer, Caitlin M. Roake, Stewart J. Kim, Ryan J. Lu & Steven E. Artandi

4. Department of Pathology, Stanford University School of Medicine, Stanford, CA, USA

Robert B. West & Gregory W. Charville

Contributions

P.N. and S.E.A. conceived the study. P.N. and S.E.A. designed the experiments. P.N., S.J.K. and R.J.L. performed histological analysis. P.N. and R.B.W. performed the LCM experiments. G.W.C. analysed and evaluated the human patient samples. C.M.R. performed the sequencing analyses. P.N. and S.E.A. analysed the data and wrote the paper.

Corresponding author

Correspondence to Steven E. Artandi.

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Laura Wood 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 Identification of a TERT-high subpopulation in acinar cells and beta cells but not duct cells using lineage tracing.](#)

a, *Tert*^{CreERT2/+}; *Rosa26*^{LSL-tdTomato/+} mouse, which responds to tamoxifen-induced Cre-mediated activation of the tdTomato reporter allele. **b**, Immunohistochemistry analysis of Tomato in a *Tert*^{CreERT2/+}; *Rosa26*^{LSL-tdTomato/+} (*TertCreER*) pancreas after oil and tamoxifen injection (n=4 mice for oil ctrl; n=11 mice for tamoxifen). Mo, month. **c, d**, Co-immunofluorescence stains 10 days after single dose of tamoxifen (5mg) in *TertCreER* pancreas (n=4 mice). Tomato (red, **c, d**), DAPI (blue, **c, d**), Insulin (grey, **c, d**) Sox9 (green, **c**) and Amylase (green, **d**) are shown. **e, f**, Single-molecule RNA FISH for *Tert* mRNA (**e**) on FACS-purified bulk acinar cells (**e, f**) from *TertCreER* mice 10d after Tamoxifen (5mg) injection

(n=4 mice). **g**, Quantification of Tomato-positive cells in *TertCreER* mice in the acinar cell compartment 10 days and within pancreatic ductal cells 10 days and 12 months post tamoxifen activation (n=5 mice). **h**, qRT-PCR for indicated genes in FACS-sorted Tom^{Neg} (white bars) and Tom^{Pos} (grey bars) acinar cells. **i, j**, Immunofluorescence analysis of a *TertCreER* pancreas 10 days after a single dose of 5mg tamoxifen (n=5 mice). Tomato (red), Dclk1 (**i**, green) or BMI1 (**j**, green) and DAPI (blue) are shown. Scale bar, 100μm (**b**), 50μm (**e, f**) and 25μm (**c, d, i, j**). All data are mean± SD. *P* values calculated by two-sided Mann-Whitney test (**h**).

[Source data](#)

Extended Data Fig. 2 Randomly marked acinar cells show limited clonogenicity and TERT-high acinar cells renew to yield TERT-low cells.

a, Schematic of lineage tracing in *TertCreER* mice treated with a single dose of tamoxifen on postnatal day 60 and analysis at indicated time points. Mo, month. **b–d**, Immunohistochemistry for Tomato in *PtflaCreER* mice 10d post tamoxifen (5mg (**b**), 0.5mg (**c**) and 0.05mg (**d**)) injection (n=4 mice). **e**, Distribution of Tomato-positive acinar cells per clone for each time point in *PtflaCreER* mice (each color represents one mouse; n=4 for each time point) **f**, *TertCreER* mice are injected with Tamoxifen (5mg) on postnatal day 60 and put on BrdU water (1mg/ml) for 30 days. **g**, Immunofluorescence analysis of a *TertCreER* pancreas after 30 days of BrdU water. Tomato (red), BrdU (green) and DAPI (blue) are shown. TERT^{High} Tomato/BrdU double positive cells (green arrow) and TERT^{Low} BrdU positive cells (orange arrow) shown. **h**, Quantification of TERT^{Low} and TERT^{High} BrdU+ cells (n=5 mice) (mean ± SD). **i–k**, Single-molecule RNA FISH for *Tert* on FACS-purified bulk (**i**), TERT^{High} (**j**) and TERT^{Low} derived (**k**) acinar cells of *TertCreER* mice(n=3). Scale bars, 25μm (**g**), 50μm (**i–k**), 100 μm (**b–d**). *P* values calculated by two-sided Mann-Whitney test (**h**).

[Source data](#)

Extended Data Fig. 3 During injury, TERT^{High} acinar show enhanced repopulation compared with randomly marked acinar cells.

a, Schema of cerulein-induced pancreatitis. **b**, Immunofluorescence analysis of *TertCreER* mice 30 days after cerulein injections. Tomato (red), Ki67 (green) and DAPI (blue) are shown. TERT^{High}, Ki67⁺ positive cells (green arrow) and TERT^{Low}, Ki67⁺ cells (orange arrows) are shown. **c**, Quantification of TERT^{High}, Ki67⁺ and TERT^{Low}, Ki67⁺ cells at indicated time points (n=5 mice) (mean ± SD). **d**, Schematic of L-Arginine induced pancreatitis. **e**, Immunofluorescence analysis of *TertCreER* mice 30 days after L-Arginine induced pancreatitis. **f**, Quantification of Tomato-positive cells per clone with and without injury (n=5 mice) (mean ± SD). Scale bars, 25μm (**d**), 250 μm (**e**). *P* values calculated by two-sided Mann-Whitney test (**c**, **f**).

[Source data](#)

Extended Data Fig. 4 Enhanced PanIN formation from TERT-High acinar cells compared with sparsely labeled randomly marked acinar cells.

a, Schema for cerulein-induced pancreatitis. **b**, Immunohistochemistry for Tomato in *PtflaCreER; Kras* mice and *TertCreER; Kras* mice one month after cerulein-induced pancreatitis (AP). **c**, Quantification of number of clones per 20x high-power field (HPF) in *PtflaCreER; Kras* mice (n=7) and *TertCreER; Kras* mice (n=7) one month after cerulein-induced pancreatitis (AP) (mean ± SD). **d**, **e**, Representative metaplastic area (**d**, dashed line) and PanIN lesion (**e**, black arrow) in H&E staining of *TertCreER; Kras* mice one-month post injury (AP) (n=9 mice). **f**, **g**, Representative H&E staining (**f**) and Immunohistochemistry for Tomato (**g**) of *PtflaCreER; Kras* mice one-month post injury (AP) with high dose tamoxifen (5mg) (n=9 mice). **h–k**, Immunofluorescence analysis of a *TertCreER* (**h**, **i**) and *TertCreER; Kras* (**j**, **k**) mice one month after sham (**h**, **j**) or cerulein (**i**, **k**) treatment. Acinar cell expansion (red arrows) and PanINs (white arrow) are shown. Scale bar 100μm (**a**, **c–f**) and 50μm (**g–j**).

[Source data](#)

Extended Data Fig. 5 MEK inhibitor prevents PanIN formation from TERT-High cells expressing mutant Kras.

a, Schema of cerulein-induced pancreatitis with MEKi treatment (Trametinib). **b, c**, H&E (**b**) and Tomato (**c**) staining of *TertCreER; Kras* mice one month after cerulein induced pancreatitis with (Trametinib) and without (Vehicle) MEKi treatment. **d**, Quantification of Tomato-positive cells per clone with (n=4 mice) and without (n=3 mice) MEK inhibition (mean ± SD). Scale bars, 100μm.

[Source data](#)

Extended Data Fig. 6 Specificity of pERK-positive acinar regions in human samples.

pERK 1/2 staining of human pancreas cancer specimen using indicated antibodies and IgG control (n=7 for #4376, n=44 for #4370 and n=3 for IgG control. Scale bars, 100μm.

Extended Data Table 1 Disease status of non-PDAC and autopsy samples

Extended Data Table 2 pERK status and Kras mutation sequences

Supplementary information

Supplementary Information

FACS gating strategy

Reporting Summary

Peer Review File

Source data

[Source Data Fig. 1](#)

[Source Data Fig. 2](#)

[Source Data Fig. 3](#)

[Source Data Fig. 4](#)

[Source Data Extended Data Fig. 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](#)

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Neuhöfer, P., Roake, C.M., Kim, S.J. *et al.* Acinar cell clonal expansion in pancreas homeostasis and carcinogenesis. *Nature* **597**, 715–719 (2021).
<https://doi.org/10.1038/s41586-021-03916-2>

- Received: 17 January 2020

- Accepted: 16 August 2021
- Published: 15 September 2021
- Issue Date: 30 September 2021
- DOI: <https://doi.org/10.1038/s41586-021-03916-2>

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

[Insights into the origins of pancreatic cancer](#)

- Laura D. Wood
- Anirban Maitra

News & Views 15 Sept 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03916-2>

- Article
- [Published: 06 September 2021](#)

Programmable RNA targeting with the single-protein CRISPR effector Cas7-11

- [Ahsen Özcan¹](#),
- [Rohan Krajeski¹](#),
- [Eleonora Ioannidi^{1,2}](#),
- [Brennan Lee¹](#),
- [Apolonia Gardner^{1,3}](#),
- [Kira S. Makarova](#) [ORCID: orcid.org/0000-0002-8174-2844⁴](#),
- [Eugene V. Koonin](#) [ORCID: orcid.org/0000-0003-3943-8299⁴](#),
- [Omar O. Abudayyeh](#) [ORCID: orcid.org/0000-0002-7979-3220^{1 na1}](#)
&
- [Jonathan S. Gootenberg](#) [ORCID: orcid.org/0000-0003-2757-2175^{1 na1}](#)

[Nature](#) volume 597, pages 720–725 (2021)

- 16k Accesses
- 399 Altmetric
- [Metrics details](#)

Subjects

- [Biotechnology](#)
- [Microbiology](#)

- [Molecular biology](#)

Abstract

CRISPR–Cas interference is mediated by Cas effector nucleases that are either components of multisubunit complexes—in class 1 CRISPR–Cas systems—or domains of a single protein—in class 2 systems^{1,2,3}. Here we show that the subtype III-E effector Cas7-11 is a single-protein effector in the class 1 CRISPR–Cas systems originating from the fusion of a putative Cas11 domain and multiple Cas7 subunits that are derived from subtype III-D. Cas7-11 from *Desulfonema ishimotonii* (*DiCas7-11*), when expressed in *Escherichia coli*, has substantial RNA interference effectivity against mRNAs and bacteriophages. Similar to many class 2 effectors—and unique among class 1 systems—*DiCas7-11* processes pre-CRISPR RNA into mature CRISPR RNA (crRNA) and cleaves RNA at positions defined by the target:spacer duplex, without detectable non-specific activity. We engineered Cas7-11 for RNA knockdown and editing in mammalian cells. We show that Cas7-11 has no effects on cell viability, whereas other RNA-targeting tools (such as short hairpin RNAs and Cas13) show substantial cell toxicity^{4,5}. This study illustrates the evolution of a single-protein effector from multisubunit class 1 effector complexes, expanding our understanding of the diversity of CRISPR systems. Cas7-11 provides the basis for new programmable RNA-targeting tools that are free of collateral activity and cell toxicity.

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](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

Fig. 1: Cas7-11 single-protein effectors possess both pre-crRNA processing and interference activity.

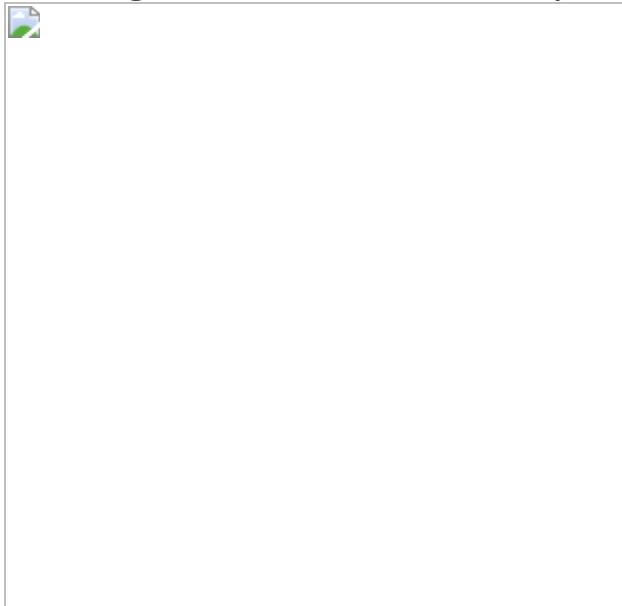


Fig. 2: Mechanism of programmable RNA cleavage by DiCas7-11.

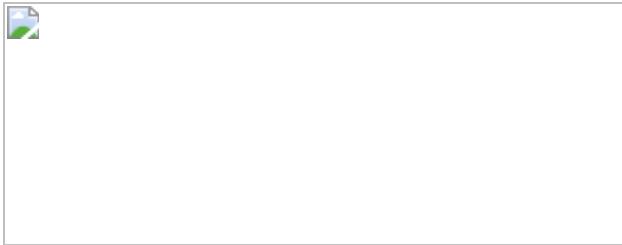
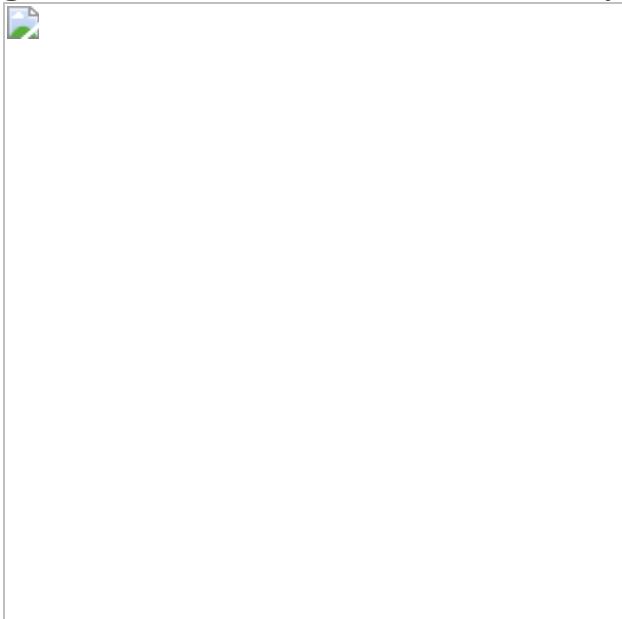


Fig. 3: Engineered Cas7-11 orthologues can be used for mammalian transcript targeting.



Fig. 4: Cas7-11 lacks collateral activity in mammalian cells.



Data availability

Sequencing data are available from the Sequence Read Archive under BioProject accession number [PRJNA657647](#).

Code availability

Supporting information and computational tools are available on the Abudayyeh–Gootenberg laboratory website (<https://www.abugootlab.org/>).

References

1. 1.
Wright, A. V., Nuñez, J. K. & Doudna, J. A. Biology and applications of CRISPR systems: harnessing nature's toolbox for genome engineering. *Cell* **164**, 29–44 (2016).
2. 2.
Hsu, P. D., Lander, E. S. & Zhang, F. Development and applications of CRISPR–Cas9 for genome engineering. *Cell* **157**, 1262–1278 (2014).
3. 3.
Makarova, K. S. et al. Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.* **18**, 67–83 (2020).
4. 4.
Wang, Q. et al. The CRISPR–Cas13a gene-editing system induces collateral cleavage of RNA in glioma cells. *Adv. Sci.* **1**, 1901299 (2019).
5. 5.
Wang, L., Zhou, J., Wang, Q., Wang, Y. & Kang, C. Rapid design and development of CRISPR–Cas13a targeting SARS-CoV-2 spike protein. *Theranostics* **11**, 649–664 (2021).
6. 6.
Shmakov, S. et al. Discovery and functional characterization of diverse class 2 CRISPR–Cas systems. *Mol. Cell* **60**, 385–397 (2015).

7. 7.

Samai, P. et al. Co-transcriptional DNA and RNA cleavage during type III CRISPR–Cas immunity. *Cell* **161**, 1164–1174 (2015).

8. 8.

Staals, R. H. J. et al. RNA targeting by the type III-A CRISPR–Cas Csm complex of *Thermus thermophilus*. *Mol. Cell* **56**, 518–530 (2014).

9. 9.

Tamulaitis, G. et al. Programmable RNA shredding by the type III-A CRISPR–Cas system of *Streptococcus thermophilus*. *Mol. Cell* **56**, 506–517 (2014).

10. 10.

Hale, C. R. et al. RNA-guided RNA cleavage by a CRISPR RNA–Cas protein complex. *Cell* **139**, 945–956 (2009).

11. 11.

Abudayyeh, O. O. et al. C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science* **353**, aaf5573 (2016).

12. 12.

East-Seletsky, A. et al. Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection. *Nature* **538**, 270–273 (2016).

13. 13.

Anantharaman, V., Makarova, K. S., Burroughs, A. M., Koonin, E. V. & Aravind, L. Comprehensive analysis of the HEPN superfamily:

identification of novel roles in intra-genomic conflicts, defense, pathogenesis and RNA processing. *Biol. Direct* **8**, 15 (2013).

14. 14.

Konermann, S. et al. Transcriptome engineering with RNA-targeting type VI-D CRISPR effectors. *Cell* **173**, 665–676.e14 (2018).

15. 15.

Yan, W. X. et al. Cas13d is a compact RNA-targeting type VI CRISPR effector positively modulated by a WYL-domain-containing accessory protein. *Mol. Cell* **70**, 327–339.e5 (2018).

16. 16.

Smargon, A. A. et al. Cas13b is a type VI-B CRISPR-associated RNA-guided RNase differentially regulated by accessory proteins Csx27 and Csx28. *Mol. Cell* **65**, 618–630.e7 (2017).

17. 17.

Gootenberg, J. S. et al. Nucleic acid detection with CRISPR–Cas13a/C2c2. *Science* **356**, 438–442 (2017).

18. 18.

Gootenberg, J. S. et al. Multiplexed and portable nucleic acid detection platform with Cas13, Cas12a, and Csm6. *Science* **360**, 439–444 (2018).

19. 19.

Myhrvold, C. et al. Field-deployable viral diagnostics using CRISPR–Cas13. *Science* **360**, 444–448 (2018).

20. 20.

Abudayyeh, O. O., Gootenberg, J. S., Kellner, M. J. & Zhang, F. nucleic acid detection of plant genes using CRISPR–Cas13. *CRISPR J.* **2**, 165–171 (2019).

21. 21.

Kellner, M. J., Koob, J. G., Gootenberg, J. S., Abudayyeh, O. O. & Zhang, F. SHERLOCK: nucleic acid detection with CRISPR nucleases. *Nat. Protoc.* **14**, 2986–3012 (2019).

22. 22.

Abudayyeh, O. O. et al. RNA targeting with CRISPR–Cas13. *Nature* **550**, 280–284 (2017).

23. 23.

Cox, D. B. T. et al. RNA editing with CRISPR–Cas13. *Science* **358**, 1019–1027 (2017).

24. 24.

Wilson, C., Chen, P. J., Miao, Z. & Liu, D. R. Programmable m⁶A modification of cellular RNAs with a Cas13-directed methyltransferase. *Nat. Biotechnol.* **38**, 1431–1440 (2020).

25. 25.

Koonin, E. V. & Zhang, F. Coupling immunity and programmed cell suicide in prokaryotes: life-or-death choices. *Bioessays* **39**, 1–9 (2017).

26. 26.

Makarova, K. S., Grishin, N. V., Shabalina, S. A., Wolf, Y. I. & Koonin, E. V. A putative RNA-interference-based immune system in prokaryotes: computational analysis of the predicted enzymatic machinery, functional analogies with eukaryotic RNAi, and hypothetical mechanisms of action. *Biol. Direct* **1**, 7 (2006).

27. 27.

Silas, S. et al. Direct CRISPR spacer acquisition from RNA by a natural reverse transcriptase–Cas1 fusion protein. *Science* **351**, aad4234 (2016).

28. 28.

Shmakov, S. A. et al. The CRISPR spacer space is dominated by sequences from species-specific mobilomes. *mBio* **8**, e01397-17 (2017).

29. 29.

Abudayyeh, O. O. et al. A cytosine deaminase for programmable single-base RNA editing. *Science* **365**, 382–386 (2019).

30. 30.

Putzbach, W. et al. Many si/shRNAs can kill cancer cells by targeting multiple survival genes through an off-target mechanism. *eLife* **6**, e29702 (2017).

31. 31.

Uhlén, M. et al. Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).

32. 32.

Zhu, W., Lomsadze, A. & Borodovsky, M. Ab initio gene identification in metagenomic sequences. *Nucleic Acids Res.* **38**, e132 (2010).

33. 33.

Couvin, D. et al. CRISPRCasFinder, an update of CRISRFinder, includes a portable version, enhanced performance and integrates search for Cas proteins. *Nucleic Acids Res.* **46**, W246–W251 (2018).

34. 34.

Edgar, R. C. PILER-CR: fast and accurate identification of CRISPR repeats. *BMC Bioinf.* **8**, 18 (2007).

35. 35.

Joung, J. et al. Genome-scale CRISPR–Cas9 knockout and transcriptional activation screening. *Nat. Protoc.* **12**, 828–863 (2017).

36. 36.

Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

37. 37.

Law, C. W., Chen, Y., Shi, W. & Smyth, G. K. voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* **15**, R29 (2014).

38. 38.

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).

39. 39.

Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47 (2015).

Acknowledgements

We thank M. Terns, E. Sontheimer and M. Mittens for discussions; A. Badran for discussions and plate-reader assistance; P. Reginato, D. Weston and E. Boyden for MiSeq instrumentation; S. Tonegawa and D. King for

centrifuge assistance; G. Feng and D. Wang for gel imager support; G. Paradis and M. Griffin for flow cytometry assistance; A. Sejr Hansen for providing mES cells; and R. Desimone, J. Crittenden and S. Lall for support and discussions. O.O.A. and J.S.G. are supported by NIH grant 1R21-AI149694; The McGovern Institute Neurotechnology (MINT) program; and the McGovern Institute. K.S.M. and E.V.K. are supported by the Intramural Research Program of the National Institutes of Health (National Library of Medicine). The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author information

Author notes

1. These authors jointly supervised this work: Omar O. Abudayyeh, Jonathan S. Gootenberg

Affiliations

1. McGovern Institute for Brain Research at MIT, Massachusetts Institute of Technology, Cambridge, MA, USA

Ahsen Özcan, Rohan Krajeski, Eleonora Ioannidi, Brennan Lee, Apolonia Gardner, Omar O. Abudayyeh & Jonathan S. Gootenberg

2. ETH Zurich, Zurich, Switzerland

Eleonora Ioannidi

3. Harvard University, Harvard Medical School, Virology Program, Boston, MA, USA

Apolonia Gardner

4. National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, Bethesda, MD, USA

Kira S. Makarova & Eugene V. Koonin

Contributions

O.O.A. and J.S.G. conceived the study; O.O.A. and J.S.G designed and participated in all experiments. A.Ö. participated in small-RNA-sequencing analyses, protein purifications, and in vitro cleavage and bacterial assays. E.I. and A.G. assisted with mammalian Cas7-11 assays; R.K. helped with cloning and plasmid sequencing; B.L. helped with genome and metagenome mining of new orthologues; K.S.M. and E.V.K. contributed to sequence and phylogeny analyses; O.O.A., E.V.K., J.S.G and A.Ö. wrote the manuscript with help from all authors.

Corresponding authors

Correspondence to Omar O. Abudayyeh or Jonathan S. Gootenberg.

Ethics declarations

Competing interests

O.O.A. and J.S.G. are co-inventors on a patent application (US 63/073,898) filed by MIT relating to work in this manuscript. O.O.A. and J.S.G. are co-founders of Sherlock Biosciences, Proof Diagnostics, Moment Biosciences and Tome Biosciences. O.O.A. and J.S.G. were advisors for Beam Therapeutics during the performance of the described research.

Additional information

Peer review information *Nature* thanks Malcolm White 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 Additional type III-E family members and novel type III-D2 loci architecture and multiple alignment of representative orthologues.

a, Loci corresponding to listed accession IDs (see Supplementary Table 1), with key genes highlighted. CRISPR array symbols are not representative of the number of spacers in the array. **b**, Multiple alignment of representative Cas7-11 orthologues showing conservation of the residues involved in catalysis: D177, D429, D654, D745, D758, and E959. For the D177 alignment, members of the Csm3 family, which are Cas7-like proteins, are included for analysis of conserved residues (highlighted in grey).

Extended Data Fig. 2 Heterologous expression of the *Desulfonema ishimotonii* Type III-E CRISPR–Cas system and associated CRISPR array.

a, Heterologously expressed *Desulfonema ishimotonii* Type III-E full locus matures crRNAs. **b**, Heterologously expressed single effector protein *DiCas7-11* matures crRNAs in *E. coli*. **c**, Expression of the *DiCas7-11* CRISPR array without any effector or accessory proteins. **d**, Schematic of the DR secondary structure. Scissors denote the cleavage site for processing and maturation. **e**, Processing of crRNAs occurs when the *CJcCas7-11* locus is heterologously expressed in *E. coli*. Mature 37 nt crRNAs are generated, containing a 15nt DR.

Extended Data Fig. 3 The RNA-guided RNA-targeting Cas7-11 is capable of defence against ssRNA MS2 phage and RNA knockdown in bacteria.

a, Schematic of CRISPR array screen of all spacers targeting the MS2 genome. **b**, Results of the MS2 interference screen shown as box plots. Enrichment of *DiCas7-11* spacers in the phage targeting condition denote

survival of bacteria and enhanced representation of specific active spacers. Boxes denotes 25th and 75th percentiles with the median marked by the middle line. The whiskers are calculated via the Tukey method (1.5 times the inter-quartile range). Outliers are denoted by blue plus symbols.

c, Results of the MS2 interference screen showing enrichment of *DiCas7-11a* non-targeting spacers across varying phage dilution amounts as box plots. Boxes denotes 25th and 75th percentiles with the median marked by the middle line. The whiskers are calculated via the Tukey method (1.5 times the inter-quartile range). Outliers are denoted by blue plus symbols.

d, Number of *DiCas7-11a* spacers that display survival enrichment over a threshold of 1.7 across different phage dilution conditions. **e**, Quantification of resistance conferred by top MS2-targeting *DiCas7-11* spacers compared against a panel of 4 non-targeting spacers. Resistance is quantified as the highest surviving titer of MS2 phage that generates plaques in the dilution assay. **f**, Quantification of resistance conferred by two top MS2-targeting *DiCas7-11* spacers compared against a panel of 6 non-targeting spacers. Resistance is quantified as the highest surviving titer of MS2 phage that generates plaques in the dilution assay. Data are mean ± s.e.m.; $n = 3$.

Extended Data Fig. 4 Protospacer flanking site (PFS) sequences for *DiCas7-11* and effect of accessory proteins on interference.

a, Gating strategy used for all bacterial RFP knockdown experiments in Fig. 1e,g and Extended Data Fig. 6l. **b**, Schematic for targeting of *DiCas7-11* DNA target in a plasmid with resulting sequence motifs. **c**, Schematic for targeting of *DiCas7-11* RNA target in β-lactamase (ampicillin resistance gene) and the resulting PFS determined by 20 depleted targets. **d**, PFS analysis of top spacers from the MS2 phage screen (e-5 condition) in the 8 bp flanking the target region to the left. **e**, PFS analysis of top spacers from the MS2 phage screen (e-5 condition) in the 8 bp flanking the target region to the right. **f**, Schematic of the Type III-E locus of the *Desulfonema ishimotonii*. **g**, Phage plaque assay of the Type III-E *DiCas7-11* effector alone and as part of the entire locus. Two top MS2 targeting guides along with four non-targeting guides are used to assess target interference and survival against phage. Resistance is quantified as the highest surviving titer of MS2 phage that generates plaques in the dilution assay. **h**, Phage plaque assay of *E. coli* transformed with the *DiCas7-11* locus containing

corresponding accessory gene knockouts. A top MS2 targeting guide along with a non-targeting guide is used. Resistance is quantified as the highest surviving titer of MS2 phage that generates plaques in the dilution assay. Data are mean \pm s.e.m.; $n = 3$. **i**, Phage plaque assay of *E. coli* transformed with either the complete *DiCas7-11* locus or *DiCas7-11* supplemented with corresponding accessory gene knock-ins. A top MS2 targeting guide along with a non-targeting guide is used to assess target interference and survival against phage. Resistance is quantified as the highest surviving titer of MS2 phage that generates plaques in the dilution assay. Data are mean \pm s.e.m.; $n = 3$.

Extended Data Fig. 5 Additional characterization of *DiCas7-11* cleavage activity.

- a**, Purified *DiCas7-11* product showing purity at different concentrations. **b**, *DiCas7-11a* cleavage of synthetic 31nt MS2 ssRNA with a 31nt crRNA completely duplexed to the target showing cleavage fragments that are generated in the targeting condition with protein (indicated by red triangles). **c**, *DiCas7-11* incubated with a crRNA targeting a single-stranded DNA (ssDNA) target with 5' labelling of the ssDNA strand. **d**, *DiCas7-11a* incubated with a crRNA targeting a double-stranded (dsDNA) target with 5' labelling of the top strand (left) or bottom strand (right). **e**, *DiCas7-11* incubated with either ssRNA or dsRNA targeting and crRNAs tiling the MS2 target. **f**, In vitro cleavage of MS2 ssRNA at 37 °C with varying concentrations of *DiCas7-11* incubated with a crRNA against the target with and without Csx29 protein. MS2 ssRNA target is 5' labelled with Cy5. Cleavage bands are marked by asterisks. **g**, Protospacer flanking site (PFS) sequence screen showing in vitro cleavage of randomized PFS targets and lack of sequence preference flanking the target site for *DiCas7-11* cleavage. **h**, Incubation of a long body-labelled MS2 ssRNA target with *DiCas7-11* and targeting or non-targeting crRNA in the presence of different ions or chelating agents, demonstrating dependence on magnesium, manganese, or calcium. **i**, In vitro cleavage of ssRNA 1 target containing a region of complementarity against the mature *DiCas7-11* DR. The crRNA and *DiCas7-11* are incubated with targets containing different length regions of complementarity to the DR. The ssRNA target is body labelled with Cy5. **j**, Cleavage of ssRNA 1 target with increasing amounts of *DiCas7-11*-

crRNA complex from 0 nM to 233 nM. The ssRNA target is body labelled with Cy5. **k**, Cleavage of ssRNA 1 target at increasing incubation periods from 0 min to 180 min. The ssRNA target is body labelled with Cy5.

Extended Data Fig. 6 Additional biochemical characterization of Cas7-11.

a, Schematic showing sequence of *DiCas7-11* crRNA 1, targeting the ssRNA 1 target. **b**, Cleavage of ssRNA 1 target with crRNA 1 of varying DR and spacer lengths. **c**, Schematic showing sequence of *DiCas7-11* crRNA 1, targeting the long MS2 target. **d**, Cleavage of long MS2 target with crRNA 1 with mutagenized DRs. **e**, *DiCas7-11*:crRNA complex binding to a complementary MS2 ssRNA target is determined by electrophoretic mobility shift assay (EMSA). EMSAs are performed for wild-type (WT) *DiCas7-11*, dead *DiCas7-11*, a non-targeting crRNA, and a MS2-targeting guide alone. **f**, *DiCas7-11*:crRNA complex binding to a complementary MS2 ssDNA target is determined by EMSA. **g**, Quantification of band intensities for non-targeting guide EMSA in **e** showing fitted curve (mean \pm s.e.m.; $n = 3$). **h**, Quantification of band intensities for active *DiCas7-11* with MS2 targeting guide EMSA in **e** showing fitted curve (mean \pm s.e.m.; $n = 3$). **i**, Quantification of band intensities for MS2 targeting guide EMSA gel in **e** showing fitted curve (mean \pm s.e.m.; $n = 3$). **j**, Quantification of band intensities for dead *DiCas7-11* EMSA in **e** showing fitted curve (mean \pm s.e.m.; $n = 3$). **k**, Quantification of band intensities for ssDNA EMSA gel in **f** showing fitted curve (mean \pm s.e.m.; $n = 3$). **l**, RFP knockdown in *E. coli* is assayed via flow cytometry with wild-type and D429A/D654A *DiCas7-11* protein expression. Data are mean \pm s.e.m.; $n = 3$. **m**, Quantification of resistance conferred by top MS2-targeting *DiCas7-11* spacers with active *DiCas7-11*, dead *DiCas7-11*, and no protein. Resistance is quantified as the highest surviving titer of MS2 phage that generates plaques (mean \pm s.e.m.; $n = 3$).

Extended Data Fig. 7 Additional characterization of *DiCas7-11* processing.

a, CRISPR array processing assay showing robust processing activity by *DiCas7-11* at concentrations ranging from 0 to 233 nM. **b**, Schematic of in vitro transcription of pre-crRNA and processing by *DiCas7-11*. **c**, Sequence of single spacer pre-crRNA, showing locations of mutated bases. **d**, Processing of transcribed pre-crRNA, showing the effect of single position mutations on processing. **e**, Processing of transcribed pre-crRNA at different concentrations of *DiCas7-11*. **f**, Processing of pre-crRNA by *DiCas7-11* in the presence of different ions or chelating agents. CRISPR array processing by *DiCas7-11* is ion independent. **g**, Processing activity of a synthetic *DiCas7-11* CRISPR array by *DiCas7-11* protein with predicted catalytic processing mutants in the protein insert region with a characteristic KxYxH catalytic triad motif.

Extended Data Fig. 8 Further characterization of RNA knockdown in mammalian cells by Cas7-11.

a, Knockdown of Gluc mRNA in mammalian cells by *DiCas7-11* with a panel of guides tiled across the transcript. **b**, Knockdown of Gluc mRNA in mammalian cells by *DiCas7-11* with N- and C-terminus fusions of nuclear export signal (NES) tags (Adeno type 5 E1B-55K) with guides tiled across the transcript. **c**, Knockdown activity in HEK293FT cells of *DiCas7-11*, shRNA, LwaCas13a, PspCas13b, and RfxCas13d against the Gluc transcript normalized to corresponding non-targeting controls. All guides are designed to target the same region of the Gluc transcript. **d**, Comparison of knockdown activity of Gluc mRNA in mammalian cells between active *DiCas7-11*, catalytically inactive D429A/D654A *DiCas7-11*, and GFP. Guides have the full DR sequence. **e**, Knockdown activity of *DiCas7-11* against the multiple endogenous transcripts normalized to two non-targeting conditions. Dotted lines represent background expression and 65% knockdown by *DiCas7-11*. **f**, Knockdown of EGFP and mCherry mRNA in mammalian cells by *DiCas7-11* with a panel of tiling guides. **g**, Knockdown of Gluc mRNA in mammalian cells by *DiCas7-11* with a panel of guides of different lengths. **h**, Knockdown of Gluc mRNA in mammalian cells by hvsCas7-11 with a panel of guides containing the full DR sequence. Guides are designed to be tiled across the Gluc transcript. **i**, Knockdown of Gluc mRNA in HepG2 cells with hvsCas7-11 and no protein conditions using guides 1 and 11 as well as a non-targeting guide. **j**, Knockdown of Gluc

mRNA in HEK293FT cells with hvsCas7-11 modified with an NES sequence (Adeno type 5 E1B-55K) and using Gluc guide 1. The NES tag is either C-terminally fused or fused on both ends of the hvsCas7-11 protein. All data in this figure are mean \pm s.e.m.; $n = 2$ or $n = 3$, as shown.

Extended Data Fig. 9 Development of d DiCas7-11 for programmable RNA-editing in mammalian cells and further characterization of specificity.

a, RNA editing via d DiCas7-11-NES-ADAR2 fusions programmed to target a specific adenosine via a cytidine mismatch. **b**, DiCas7-11 guide design for programmable A-to-I editing. The mismatch distance is defined as the position of the A-C mismatch relative to the 3' end of the guide. **c**, RNA A-to-I editing of *Cypridina* luciferase (Cluc) mRNA W85X mutation in mammalian cells by dDiCas7-11-NES-ADAR2. Guides have mismatch distances between 2–50nt. Editing is measured either by restoration of Cluc luciferase activity normalized to the non-targeting guide condition (top) or percentage conversion of the adenosine to inosine (bottom). **d**, RNA A-to-I editing of Cluc mRNA W85X mutation in mammalian cells with different nuclear localization tags and linker architectures compared to ADAR deaminase without *DiCas7-11* fusion. Data are mean \pm s.e.m.; $n = 3$. **e**, End point readout of collateral activity of either *DiCas7-11* or LwaCas13a targeting and non-targeting crRNA against MS2 ssRNA target. **f**, RFP knockdown by *DiCas7-11* or LwaCas13a in *E. coli*. Corresponding growth curves are shown in Fig. 4b. **g-h**, Effects on cell health by Gluc targeting with CRISPR effectors in (**g**) HEK293FT cells and (**h**) U87 glioblastoma cells. **i**, Number of significant off-targets when targeting the Gluc transcript in HEK293FT and U87 cells by *DiCas7-11*, shRNA, LwaCas13a, PspCas13b, and RfxCas13d using transcriptome-wide RNA-sequencing data. Off-targets are determined by significance testing of differentially expressed transcripts between targeting and non-targeting guide conditions. **j**, Number of significant off-targets when targeting the Gluc transcript in HEK293FT and U87 cells by *DiCas7-11*, shRNA, LwaCas13a, PspCas13b, and RfxCas13d using transcriptome-wide RNA-sequencing data. Off-targets are determined by significance testing of differentially expressed transcripts between the targeting guide condition and cells only expressing EGFP. Data are mean \pm s.e.m.; $n = 3$.

Extended Data Fig. 10 Model of Cas7-11 expression, processing, and interference.

a, Schematic of Cas7-11 expression, processing, and interference against ssRNA viruses and other ssRNA targets. **b**, Comparison of Cas7-11 features to characteristics of Cas13 and type III-A/B systems. **c**, Hypothetical scenario for the evolution of a single-protein CRISPR–Cas effector from a multi-subunit effector complex. Cas7-11, a single-effector programmable RNase combining the pre-crRNA processing and target-cleaving RNase activities, apparently originated from two type III-D variants. Three Cas7 domains (domains 3, 4 and 5) are derived from subtype III-D2 that contains a Cas7x3 effector protein along with Cas10 and another Cas7-like domain fused to a Cas5-like domain. The origin of the N-terminal Cas7 and putative Cas11 domain of Cas7-11 is not entirely clear, but most likely, these domains are derived from a III-D1 variant, where both genes are stand-alone³.

Supplementary information

Supplementary Tables

This file contains Supplementary Tables 1–8.

Reporting Summary

Supplementary Data 1

Multiple sequence alignment of multiple Cas7-11 and Cas7x3 orthologues are shown.

Supplementary Data 2

A list of all off-targets for Cas7-11, Cas13 and RNA interference knockdown in mammalian cells.

Rights and permissions

[Reprints and Permissions](#)

About this article

Cite this article

Özcan, A., Krajeski, R., Ioannidi, E. *et al.* Programmable RNA targeting with the single-protein CRISPR effector Cas7-11. *Nature* **597**, 720–725 (2021). <https://doi.org/10.1038/s41586-021-03886-5>

- Received: 12 November 2020
- Accepted: 09 August 2021
- Published: 06 September 2021
- Issue Date: 30 September 2021
- DOI: <https://doi.org/10.1038/s41586-021-03886-5>

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

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

- Article
- [Published: 15 September 2021](#)

UTX condensation underlies its tumour-suppressive activity

- [Bi Shi¹](#),
- [Wei Li](#) [ORCID: orcid.org/0000-0003-3566-5793¹](#),
- [Yansu Song¹](#),
- [Zhenjia Wang²](#),
- [Rui Ju¹](#),
- [Aleksandra Ulman](#) [ORCID: orcid.org/0000-0002-8077-6890¹](#),
- [Jing Hu¹](#),
- [Francesco Palomba](#) [ORCID: orcid.org/0000-0002-2249-2961³](#),
- [Yanfang Zhao⁴](#),
- [John Philip Le¹](#),
- [William Jarrard¹](#),
- [David Dimoff¹](#),
- [Michelle A. Digman](#) [ORCID: orcid.org/0000-0003-4611-7100³](#),
- [Enrico Gratton³](#),
- [Chongzhi Zang^{1,2,5,6}](#) &
- [Hao Jiang](#) [ORCID: orcid.org/0000-0003-1890-9813^{1,4,6}](#)

[Nature](#) volume **597**, pages 726–731 (2021)

- 7722 Accesses
- 32 Altmetric
- [Metrics details](#)

Subjects

- [Epigenetics](#)
- [Histone post-translational modifications](#)
- [Intrinsically disordered proteins](#)
- [Tumour-suppressor proteins](#)

Abstract

UTX(also known as *KDM6A*) encodes a histone H3K27 demethylase and is an important tumour suppressor that is frequently mutated in human cancers¹. However, as the demethylase activity of UTX is often dispensable for mediating tumour suppression and developmental regulation^{2,3,4,5,6,7,8}, the underlying molecular activity of UTX remains unknown. Here we show that phase separation of UTX underlies its chromatin-regulatory activity in tumour suppression. A core intrinsically disordered region (cIDR) of UTX forms phase-separated liquid condensates, and cIDR loss caused by the most frequent cancer mutation of UTX is mainly responsible for abolishing tumour suppression. Deletion, mutagenesis and replacement assays of the intrinsically disordered region demonstrate a critical role of UTX condensation in tumour suppression and embryonic stem cell differentiation. As shown by reconstitution in vitro and engineered systems in cells, UTX recruits the histone methyltransferase MLL4 (also known as KMT2D) to the same condensates and enriches the H3K4 methylation activity of MLL4. Moreover, UTX regulates genome-wide histone modifications and high-order chromatin interactions in a condensation-dependent manner. We also found that UTY, the Y chromosome homologue of UTX with weaker tumour-suppressive activity, forms condensates with reduced molecular dynamics. These studies demonstrate a crucial biological function of liquid condensates with proper material states in enabling the tumour-suppressive activity of a chromatin regulator.

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](#)
- [Access through your institution](#)
- [Learn about institutional subscriptions](#)

Fig. 1: cIDR-dependent UTX phase separation is critical for tumour suppression.

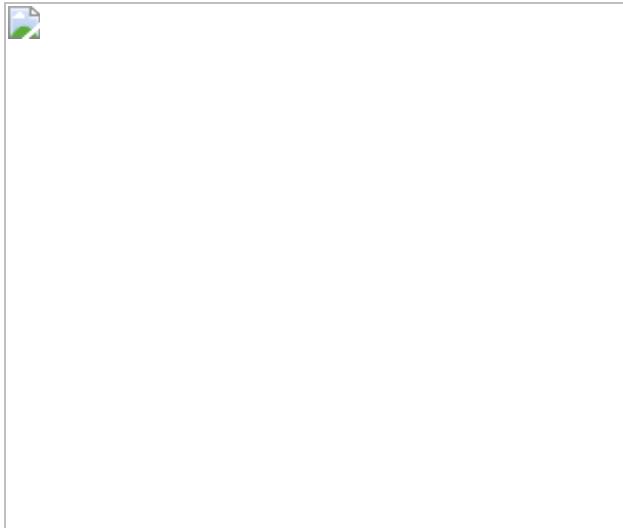


Fig. 2: Endogenous UTX condensation regulates ES cell differentiation.

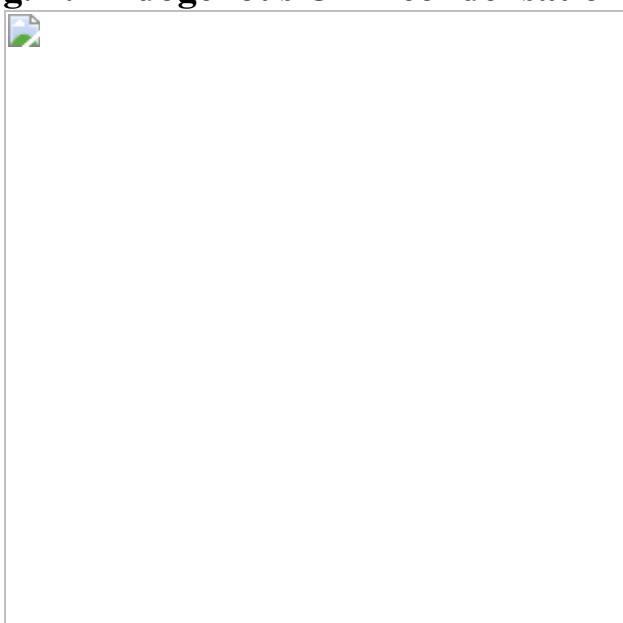


Fig. 3: Co-condensation of UTX and MLL4 enriches histone modification activities.



Fig. 4: UTX condensates regulate genome-wide chromatin activities.

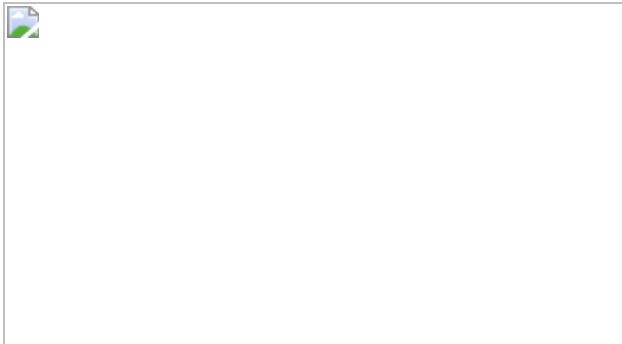
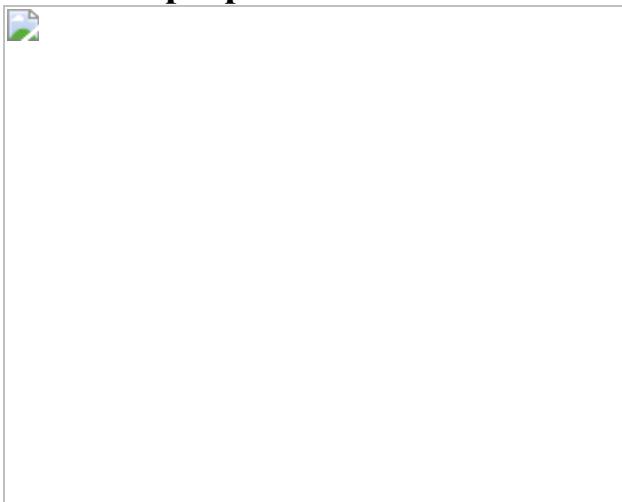


Fig. 5: UTY is a weaker tumour suppressor than UTX, with altered condensate properties.



Data availability

The high-throughput sequencing data, including RNA-seq, ChIP-seq, PRO-seq and HiChIP, have been deposited in the Gene Expression Omnibus with the accession number [GSE149420](#). Cancer mutation data are available from cBioPortal (<http://www.cbioportal.org/>) and the TCGA database. [Source data](#) are provided with this paper.

Code availability

The scripts used to analyse the data from this study are freely available at: https://github.com/zanglab/utx_code.

References

1. 1.

Wang, L. & Shilatifard, A. UTX mutations in human cancer. *Cancer Cell* **35**, 168–176 (2019).

2. 2.

Gozdecka, M. et al. UTX-mediated enhancer and chromatin remodeling suppresses myeloid leukemogenesis through noncatalytic inverse regulation of ETS and GATA programs. *Nat. Genet.* **50**, 883–894 (2018).

3. 3.

Andricovich, J. et al. Loss of KDM6A activates super-enhancers to induce gender-specific squamous-like pancreatic cancer and confers sensitivity to BET inhibitors. *Cancer Cell* **33**, 512–526.e8 (2018).

4. 4.

Morales Torres, C., Laugesen, A. & Helin, K. Utx is required for proper induction of ectoderm and mesoderm during differentiation of embryonic stem cells. *PLoS ONE* **8**, e60020 (2013).

5. 5.

Wang, C. et al. UTX regulates mesoderm differentiation of embryonic stem cells independent of H3K27 demethylase activity. *Proc. Natl Acad. Sci. USA* **109**, 15324–15329 (2012).

6. 6.

Shpargel, K. B., Sengoku, T., Yokoyama, S. & Magnuson, T. UTX and UTY demonstrate histone demethylase-independent function in mouse embryonic development. *PLoS Genet.* **8**, e1002964 (2012).

7. 7.

Shpargel, K. B., Starmer, J., Wang, C., Ge, K. & Magnuson, T. UTX-guided neural crest function underlies craniofacial features of Kabuki syndrome. *Proc. Natl. Acad. Sci. USA* **114**, E9046–E9055 (2017).

8. 8.

Miller, S. A., Mohn, S. E. & Weinmann, A. S. Jmjd3 and UTX play a demethylase-independent role in chromatin remodeling to regulate T-box family member-dependent gene expression. *Mol. Cell* **40**, 594–605 (2010).

9. 9.

Banani, S. F., Lee, H. O., Hyman, A. A. & Rosen, M. K. Biomolecular condensates: organizers of cellular biochemistry. *Nat. Rev. Mol. Cell Biol.* **18**, 285–298 (2017).

10. 10.

Li, W. et al. Biophysical properties of AKAP95 protein condensates regulate splicing and tumorigenesis. *Nat. Cell Biol.* **22**, 960–972 (2020).

11. 11.

Lin, Y., Protter, D. S., Rosen, M. K. & Parker, R. Formation and maturation of phase-separated liquid droplets by RNA-binding proteins. *Mol. Cell* **60**, 208–219 (2015).

12. 12.

Alam, H. et al. KMT2D deficiency impairs super-enhancers to confer a glycolytic vulnerability in lung cancer. *Cancer Cell* **37**, 599–617 (2020).

13. 13.

Mansour, A. A. et al. The H3K27 demethylase Utx regulates somatic and germ cell epigenetic reprogramming. *Nature* **488**, 409–413 (2012).

14. 14.

Tran, N., Broun, A. & Ge, K. Lysine demethylase KDM6A in differentiation, development, and cancer. *Mol. Cell. Biol.* **40**, e00341-20 (2020).

15. 15.

Shin, Y. et al. Spatiotemporal control of intracellular phase transitions using light-activated optoDroplets. *Cell* **168**, 159–171.e14 (2017).

16. 16.

Kim, J. H. et al. UTX and MLL4 coordinately regulate transcriptional programs for cell proliferation and invasiveness in breast cancer cells. *Cancer Res.* **74**, 1705–1717 (2014).

17. 17.

Janicki, S. M. et al. From silencing to gene expression: real-time analysis in single cells. *Cell* **116**, 683–698 (2004).

18. 18.

Wang, S. P. et al. A UTX–MLL4–p300 transcriptional regulatory network coordinately shapes active enhancer landscapes for eliciting transcription. *Mol. Cell* **67**, 308–321.e6 (2017).

19. 19.

Tie, F., Banerjee, R., Conrad, P. A., Scacheri, P. C. & Harte, P. J. Histone demethylase UTX and chromatin remodeler BRM bind directly to CBP and modulate acetylation of histone H3 lysine 27. *Mol. Cell. Biol.* **32**, 2323–2334 (2012).

20. 20.

Mumbach, M. R. et al. HiChIP: efficient and sensitive analysis of protein-directed genome architecture. *Nat. Methods* **13**, 919–922

(2016).

21. 21.

Fang, R. et al. Mapping of long-range chromatin interactions by proximity ligation-assisted ChIP-seq. *Cell Res.* **26**, 1345–1348 (2016).

22. 22.

Wang, Z., Zhang, Y. & Zang, C. BART3D: inferring transcriptional regulators associated with differential chromatin interactions from Hi-C data. *Bioinformatics* <https://doi.org/10.1093/bioinformatics/btab173> (2021).

23. 23.

Dunford, A. et al. Tumor-suppressor genes that escape from X-inactivation contribute to cancer sex bias. *Nat. Genet.* **49**, 10–16 (2017).

24. 24.

Gazova, I., Lengeling, A. & Summers, K. M. Lysine demethylases KDM6A and UTY: the X and Y of histone demethylation. *Mol. Genet. Metabol.* **127**, 31–44 (2019).

25. 25.

Li, X. et al. UTX is an escape from X-inactivation tumor-suppressor in B cell lymphoma. *Nat. Commun.* **9**, 2720 (2018).

26. 26.

Sze, C. C. & Shilatifard, A. MLL3/MLL4/COMPASS family on epigenetic regulation of enhancer function and cancer. *Cold Spring Harb. Perspect. Med.* **6**, a026427 (2016).

27. 27.

Herz, H. M. et al. The H3K27me3 demethylase dUTX is a suppressor of Notch- and Rb-dependent tumors in *Drosophila*. *Mol. Cell. Biol.* **30**, 2485–2497 (2010).

28. 28.

Ma, L. et al. Co-condensation between transcription factor and coactivator p300 modulates transcriptional bursting kinetics. *Mol. Cell* **81**, 1682–1697.e7 (2021).

29. 29.

Fasciani, A. et al. MLL4-associated condensates counterbalance Polycomb-mediated nuclear mechanical stress in Kabuki syndrome. *Nat. Genet.* **52**, 1397–1411 (2020).

30. 30.

Boija, A. et al. Transcription factors activate genes through the phase-separation capacity of their activation domains. *Cell* **175**, 1842–1855.e16 (2018).

31. 31.

Benyoucef, A. et al. UTX inhibition as selective epigenetic therapy against TAL1-driven T-cell acute lymphoblastic leukemia. *Genes Dev.* **30**, 508–521 (2016).

32. 32.

Faralli, H. et al. UTX demethylase activity is required for satellite cell-mediated muscle regeneration. *J. Clin. Invest.* **126**, 1555–1565 (2016).

33. 33.

Beyaz, S. et al. The histone demethylase UTX regulates the lineage-specific epigenetic program of invariant natural killer T cells. *Nat. Immunol.* **18**, 184–195 (2017).

34. 34.

Bogershausen, N. et al. Mutation update for Kabuki syndrome genes *KMT2D* and *KDM6A* and further delineation of X-linked Kabuki syndrome subtype 2. *Hum. Mutat.* **37**, 847–864 (2016).

35. 35.

Yang, Z. et al. The DPY30 subunit in SET1/MLL complexes regulates the proliferation and differentiation of hematopoietic progenitor cells. *Blood* **124**, 2025–2033 (2014).

36. 36.

Jiang, H. et al. Role for Dpy-30 in ES cell-fate specification by regulation of H3K4 methylation within bivalent domains. *Cell* **144**, 513–525 (2011).

37. 37.

Schulz, W. A., Lang, A., Koch, J. & Greife, A. The histone demethylase UTX/KDM6A in cancer: progress and puzzles. *Int. J. Cancer* **145**, 614–620 (2019).

38. 38.

Wang, J. et al. A molecular grammar governing the driving forces for phase separation of prion-like RNA binding proteins. *Cell* **174**, 688–699.e16 (2018).

39. 39.

Niaki, A. G. et al. Loss of dynamic RNA interaction and aberrant phase separation induced by two distinct types of ALS/FTD-linked FUS mutations. *Mol. Cell* **77**, 82–94.e4 (2020).

40. 40.

Holehouse, A. S., Das, R. K., Ahad, J. N., Richardson, M. O. & Pappu, R. V. CIDER: resources to analyze sequence-ensemble relationships of intrinsically disordered proteins. *Biophys. J.* **112**, 16–21 (2017).

41. 41.

Dignam, J. D., Lebovitz, R. M. & Roeder, R. G. Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **11**, 1475–1489 (1983).

42. 42.

Fujioka, A. et al. Dynamics of the Ras/ERK MAPK cascade as monitored by fluorescent probes. *J. Biol. Chem.* **281**, 8917–8926 (2006).

43. 43.

[No authors listed] Illuminating the dark proteome. *Cell* **166**, 1074–1077 (2016).

44. 44.

Rossow, M. J., Sasaki, J. M., Digman, M. A. & Gratton, E. Raster image correlation spectroscopy in live cells. *Nat. Protoc.* **5**, 1761–1774 (2010).

45. 45.

Digman, M. A. & Gratton, E. Analysis of diffusion and binding in cells using the RICS approach. *Microsc. Res. Tech.* **72**, 323–332 (2009).

46. 46.

Day, C. A., Kraft, L. J., Kang, M. & Kenworthy, A. K. Analysis of protein and lipid dynamics using confocal fluorescence recovery after photobleaching (FRAP). *Curr. Protoc. Cytometry* Ch. 2, Unit2.19 (2012).

47. 47.

Kwak, H., Fuda, N. J., Core, L. J. & Lis, J. T. Precise maps of RNA polymerase reveal how promoters direct initiation and pausing. *Science* **339**, 950–953 (2013).

48. 48.

Judd, J. et al. A rapid, sensitive, scalable method for precision run-on sequencing (PRO-seq). Preprint at *bioRxiv* <https://doi.org/10.1101/2020.05.18.102277> (2020).

49. 49.

Dobin, A. et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).

50. 50.

Li, B. & Dewey, C. N. RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).

51. 51.

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).

52. 52.

Huang da, W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**, 44–57 (2009).

53. 53.

Zhou, Y. et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523 (2019).

54. 54.

Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

55. 55.

Li, H. et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

56. 56.

Zhang, Y. et al. Model-based analysis of ChIP-seq (MACS). *Genome Biol.* **9**, R137 (2008).

57. 57.

Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

58. 58.

Zang, C. et al. A clustering approach for identification of enriched domains from histone modification ChIP-seq data. *Bioinformatics* **25**, 1952–1958 (2009).

59. 59.

Juric, I. et al. MAPS: model-based analysis of long-range chromatin interactions from PLAC-seq and HiChIP experiments. *PLoS Comput. Biol.* **15**, e1006982 (2019).

60. 60.

Li, D., Hsu, S., Purushotham, D., Sears, R. L. & Wang, T. WashU Epigenome Browser update 2019. *Nucleic Acids Res.* **47**, W158–W165 (2019).

61. 61.

Wang, S. et al. Modeling cis-regulation with a compendium of genome-wide histone H3K27ac profiles. *Genome Res.* **26**, 1417–1429 (2016).

62. 62.

Servant, N. et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol.* **16**, 259 (2015).

Acknowledgements

We thank K. Kelly for MiaPaca2 cells; J.-B. Shin for COS-7 cells; T. Stukenberg for LacO array-containing cells; M. Mayo for the luciferase plasmid; T. Stukenberg and C. Brangwynne for the Cry2-containing plasmid; B. Ren for advice on HiChIP assays; M. Guertin and K. M. Sathyan for advice on PRO-seq assays; K. Ge for advice on immunoblotting for MLL3 and MLL4 and providing MLL3 antibody; Y. Bao and UVa Genome Analysis and Technology Core for deep sequencing services; and K. Siller, G. A. Bautista, J. Huband, R. Sun and E. Hall at the University of Virginia Research Computing Center for technical support. H.J. was supported by NIH grant 1 R21 CA257936-01 and start-up funds from the University of Virginia. The confocal microscopy system at the Keck Center of University of Virginia was supported by grants from NIH (OD016446 and S10 OD025156). Sequencing in this research was supported by NCI Cancer Center Support Grant 5P30CA044579. H.J. is a recipient of the American Cancer Society Research Scholar Award (RSG-15-166-01-DMC) and the Leukemia & Lymphoma Society Scholar Award (1354-19). C.Z. was supported by NIH grant R35 GM133712. Z.W. was supported by the Farrow Fellowship and by the NCI Cancer Center Support Grant P30 CA44579. F.P. and M.A.D. were supported in part by a grant from the NSF (MCB-1615701). M.A.D. and E.G. were funded by NIH grant number P41-GM103540.

Author information

Author notes

1. These authors contributed equally: Bi Shi, Wei Li, Yansu Song, Zhenjia Wang

Affiliations

1. Department of Biochemistry and Molecular Genetics, University of Virginia School of Medicine, Charlottesville, VA, USA

Bi Shi, Wei Li, Yansu Song, Rui Ju, Aleksandra Ulman, Jing Hu, John Philip Le, William Jarrard, David Dimoff, Chongzhi Zang & Hao Jiang

2. Center for Public Health Genomics, University of Virginia School of Medicine, Charlottesville, VA, USA

Zhenjia Wang & Chongzhi Zang

3. Laboratory of Fluorescence Dynamics, The Henry Samueli School of Engineering, University of California, Irvine, CA, USA

Francesco Palomba, Michelle A. Digman & Enrico Gratton

4. Department of Biochemistry and Molecular Genetics, University of Alabama at Birmingham School of Medicine, Birmingham, AL, USA

Yanfang Zhao & Hao Jiang

5. Department of Public Health Sciences, University of Virginia School of Medicine, Charlottesville, VA, USA

Chongzhi Zang

6. UVA Cancer Center, University of Virginia School of Medicine, Charlottesville, VA, USA

Chongzhi Zang & Hao Jiang

Contributions

B.S., W.L. and Y.S. designed and performed most experiments and analysed the results, including phase separation, functional assays in cancer and stem cells, and genomic assays. Z.W. conducted bioinformatic analyses under the guidance of C.Z. R.J. and A.U. designed and performed substantial amount of phase separation and genomic editing works. J.H. and Y.Z. performed some experiments on molecular biology and protein work with the assistance of J.P.L., W.J. and D.D. F.P. performed and analysed the RICS experiments under the guidance of M.A.D. and E.G. H.J. conceived and supervised the project, designed the experiments, analysed the results and wrote the paper.

Corresponding author

Correspondence to Hao Jiang.

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Brian Huntly, Tanja Mittag and Panagiotis Ntziachristos 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 UTX undergoes cIDR-dependent phase separation.**](#)

a. Relative expression of UTX 555* mutant in 10 cancer patient samples that harbor this mutation and for which the mRNA level data are available. From cBioportal. Each bar is a patient sample. z-score indicates the number of standard deviations away from the mean of expression in the reference. It shows that 3 samples showed increase in mRNA level, and 4 samples showed reduction within 1 standard deviation of expression in reference samples, suggesting that this nonsense mutation does not usually lead to transcript decay and should express a truncated UTX (1–554). **b.** Top, a schematic of the full-length UTX. Bottom, a plot for number of non-synonymous mutations (type indicated in legend) within each 100 amino acids-long region on UTX, from 394 patients in TCGA. The plot is aligned with the top schematic. **c.** Top, a schematic showing division of the full-length UTX into 5 regions that were individually purified as fusion with Maltose-binding protein (MBP, to enhance solubility). See Constructs in Methods for rationale for region division. Bottom, phase contrast images of 50 μ M MBP-UTX fragments corresponding to each indicated UTX region in 30 mM NaCl after treatment with TEV protease for 1 hr to remove the MBP tag. **d.** Coomassie blue staining image of purified EGFP and EGFP-cIDR. **e.** Fluorescence (left) and DIC (right) microscopy images of 30 μ M EGFP-cIDR at increasing concentrations of NaCl at indicated times after dilution in the condensation buffer. Note the increase of droplet size and the occasional appearance of irregular shape at 45 min, suggesting growth and aging phenomena that are typical for phase separation. **f.** Fluorescence (left) and DIC (right) images showing wetting of 30 μ M EGFP-cIDR at the surface of the coverslip immediately after dilution. **g.** Coomassie blue staining image of purified EGFP-UTX (419–848). **h.** Fluorescence (up) and DIC (bottom) images of EGFP-UTX (419–848) at indicated concentrations. **i.** Coomassie blue staining image of purified mEGFP and mEGFP-UTX (419–848). **j.** Fluorescence (up) and DIC (bottom) images of mEGFP and mEGFP-UTX (419–848) at indicated concentrations and time points after dilution in condensation buffer. **k.** Coomassie blue staining image of purified mCherry and mCherry-UTX (419–848). Note that, throughout this manuscript, a band of unknown identity, indicated by the asterisk, always appeared at ~15kD smaller than the purified mCherry or mCherry-tagged protein of interest, regardless of the identity of the protein of interest. We thus speculate that it is either a degradation product of mCherry or an unrelated protein tightly associated with mCherry. **l.** Fluorescence (up) and

DIC (bottom) images of 40 μ M mCherry and mCherry-UTX (419–848) at indicated time points after dilution. **m**. Immunoblotting by indicated antibody of total lysates from untransfected control or 293T cells transfected with indicated constructs that are fused to EGFP. **n**. Quantification of nuclear UTX concentration by anti-UTX immunoblotting. Total lysates from indicated number of 293T and mouse ES cells were loaded, along with indicated ng of purified EGFP-UTX (1–848). We determined that the nuclear concentration of endogenous UTX to be ~153 nM in 293T cancer cells and ~383 nM in mouse ESCs. **o**. Transfected UTX forms nuclear foci at a concentration that is on the same order of magnitude as the endogenous level. Left, immunoblotting (top, by anti-UTX antibody) and Ponceau S staining (bottom) of total lysates from untransfected control or 293T cells transfected with a very low dose of plasmid for UTX-EGFP, with a transfection efficiency of ~30% as examined under microscope. Three different amounts (1:2 serial dilution) of the lysates and purified EGFP-UTX (1–848) were loaded as indicated. UTX-EGFP amount was roughly 2 times that of the endogenous UTX, and estimated to be < 1 μ M in the transfected cells, assuming that the negative cell under microscope had little contribution to the UTX-EGFP immunoblotting signal. The fluorescence image of the corresponding cells is shown on the right.

Extended Data Fig. 2 UTX cIDR and the key residues in cIDR are important for forming condensates and its tumour suppressive activity.

- a.** Growth curves of THP-1 cells electroporated with indicated constructs, shown as mean \pm SD. n = 3 independent experiments. P value is between MT2 and Δ cIDR at the last time point. **b.** Immunoblotting by indicated antibody of total lysates from THP-1 cells electroporated with empty vector or indicated UTX WT or mutants. **c.** Fluorescence images of THP-1 cells electroporated with indicated UTX constructs fused to EGFP. **d.** Analysis of amino acid enrichment for UTX cIDR. By Composition Profiler, using indicated background sample. **e.** Fluorescence images of 293T cells transfected with the corresponding UTX WT or mutants fused with EGFP. **f.** Immunoblotting by indicated antibody of total lysates from 293T cells transfected with empty vector or indicated UTX WT or mutants fused to EGFP. **g.** Immunoblotting by indicated antibody of total lysates from THP-1

cells electroporated with empty vector or indicated UTX WT or mutants. **h.** Representative fluorescence microscopy images of foci in FRAP assays for full-length UTX WT and HY fused to mEGFP in nuclei following transfection into 293T cells. The photobleached focus is indicated by the dashed box and its images at indicated time points are amplified. **i–n** all used 20 µM mEGFP, mEGFP-UTX (419–848) WT or HY as indicated. **i.** Pictures and OD₆₀₀ plotted as mean ± SD. n = 3 independent experiments. **j.** Representative fluorescence microscopy and DIC images at indicated times in condensation. Scale bar, 10 µm. **k.** FRAP recovery curves as mean ± SD at indicated early and late times in condensation. n = 13 and 12 droplets each for WT and HY, respectively, for early times, and n = 8 and 12 droplets each for WT and HY, respectively, for late times. **l–n.** Line RICS assays for WT (n = 32 droplets) and HY (n = 12) over the time range of 0–100 min in condensation. **l.** Fluorescence Images and line carpets. **m.** Line RICS autocorrelation curves. **n.** Plot showing the mean ± SD of the diffusion coefficients. *P<0.05, ***P<0.001, and ****P<0.0001, by two-sided t-test for all.

Extended Data Fig. 3 Tumour suppressive activity of UTX can be maintained by replacing its cIDR with unrelated protein IDRs.

a. Immunoblotting by anti-UTX or GAPDH of total lysates from THP-1 cells electroporated with empty vector or indicated UTX WT or mutants. **b.** Immunoblotting by anti-UTX (top) or GAPDH (bottom) of total lysates from THP-1 or MiaPaca2 cells that were transduced with empty vector or indicated UTX WT or mutants and induced with doxycycline. **c.** Confocal images of 293T cells transfected with indicated UTX chimeric constructs all fused to EGFP. **d.** Growth curves of THP-1 cells electroporated with indicated vector or UTX constructs, shown as average ± SD. n = 3 independent experiments. **e.** Representative results of colony formation assay of THP-1 cells transduced with indicated UTX WT or mutant construct. Bottom left, the number of colonies are plotted as mean ± SD (n = 3 independent experiments). Bottom right, the size of colonies are plotted as mean ± SD. n = 100 colonies for each except 73 colonies for UTX-eIF_{IDR}. **f.** Kaplan Meier Curves for survival of the animals grafted with

THP-1 cells stably transduced with indicated vector or UTX constructs. P = 0.0024 for WT vs. Δ cIDR, and 0.015 for Δ cIDR vs. UTX-eIF_{IDR}. n = 8 mice each, except n = 10 for Vector. **g.** Representative crystal violet staining images of MiaPaca2 cells transduced with indicated UTX WT or mutant construct. Two thousand cells were seeded and cultured for 7 days in the absence or presence of doxycycline (to induce transgene expression) before staining. Right, staining absorbance for cells with doxycycline are plotted as mean \pm SD from n = 3 independent repeats. **h.** MiaPaca2 cells stably transduced with indicated vector or UTX constructs were grafted into immune-deficient mice. Tumours from day 20 were imaged and their weights are plotted as average \pm SD. n = 5 mice each for Vector and UTX-FUS(IDR), or 6 mice each for the others. **i.** Immunoblotting of total lysates (Input) and anti-FLAG antibody-mediated immunoprecipitation (IP) from 293T cells transfected with empty vector, or FLAG-tagged UTX WT or Δ cIDR, using indicated antibodies. **j–o.** RNA-seq analysis for gene expression in the THP-1 cells transduced with indicated UTX constructs (two samples each). **j.** Volcano plot. Blue and red dots represent genes that were significantly (adj. P<0.05) up- or down-regulated (over 1.25 fold), respectively, by WT versus vector control. The adjusted P values on y axis were from DESeq2. The detailed gene information is in Supplementary Table 1. **k.** Gene ontology analysis for the indicated gene clusters from the heat map in Fig. 1j. Red plot shows the significantly enriched gene functions for the 300 genes in the red bracket in Fig. 1j, which were strongly induced by WT and UTX-eIF_{IDR}, but not as strongly by Δ cIDR, and not affected by Δ TPR. Blue shows the significantly enriched gene functions for the 200 genes in the blue bracket in Fig. 1j, which were markedly repressed by WT and UTX-eIF_{IDR}, but not Δ cIDR and Δ TPR. **l.** Principal component analysis of quantile normalized gene expression (log2 of TPM, pseudo-count=1) using 1,271 DEGs in WT versus vector control. **m.** Scatter plots for gene expression changes between indicated samples. Black dashed line indicates the linear fit with correlation coefficient (r) labeled. Red and blue dots are genes significantly up- or down-regulated (over 1.25 fold) in WT over vector, respectively, as in j. **n.** Box plots for gene expression changes between indicated samples for the up- and down-regulated genes (defined in m) in WT over vector. Data are median (horizontal line), 25–75th percentiles (box) and 1.5 times the interquartile range recorded (whiskers). P values are shown for data significantly greater

or less than 0 by two-sided one sample *t*-test. Left, n = 675 genes each; right, n = 596 genes each. **o.** Heatmap showing relative expression levels of genes that were significantly (adj. P<0.05) up- or down-regulated (1.25 fold each) by WT compared to vector control in the THP-1 cells transduced with vector control, UTX WT or catalytically inactive (MT2) mutant (two samples each). Each row represents a gene. P values by one-way ANOVA followed by Tukey's post hoc test for all analyses except by two-sided log-rank test for g, two-sided modified Fisher's exact test for k, and two-sided one sample *t*-test for n.

[Source data](#)

Extended Data Fig. 4 Condensation of the endogenous *Utx* is important for ESC differentiation.

a. Schematic of FLAG-mEGFP Knock-in (KI) at *Utx* in ESCs. **b.** Genomic PCR products by indicated primers for different clones. The predicted sizes are 961 bp (by F1+R1) and 954 bp (by F2+R2) for the KI clones. **c.** Immunoblotting for KI clones by green circles (neighbouring lanes are clones of mixed population). **d.** Schematic for generating a *Utx* KO ESC clone. **e.** Sequencing results showing the 4 bp loss in *Utx* in the KO clone. **f.** Immunoblotting for the *Utx* KO clone. **g.** Schematic for generating *Utx* ΔcIDR ESC clones. These clones were generated by *cIDR* deletion followed with selection for in-frame rejoined *Utx*. **h.** Genomic PCR products for different clones. The bottom right gel used the primer pair in yellow shown in g, with predicted size of 205 bp for ΔcIDR. The other gels used the primer pair in black in g, with predicted size of 1040 bp for ΔcIDR. **i.** Sequencing of the indicated ΔcIDR clones. The in-frame deleted genomic regions are shown in the square brackets. The deletion of amino acids additional to the cIDR was due to the design of the gRNAs to ensure efficient recognition of CRISPR/Cas9. These residues are all in the disordered region. A phenylalanine was also added (in red) by the repair process in clone #72. **j.** Immunoblotting for the *Utx* ΔcIDR clones. **k.** Schematic of FLAG-mEGFP KI at *Utx* in the *Utx* ΔcIDR ESC clones. **l.** Genomic PCR products for different clones using the F2 and R2 primer pair shown in a. **m.** Immunoblotting for mEGFP-tagged WT and *Utx*Δ cIDR clones. **n.** Representative confocal microscopic images of ESCs that did not

have mEGFP KI (Negative) or had mEGFP KI in the *Utx* WT or Δ cIDR clones. Two different samples of Δ cIDR clone #25 that had relatively high and low mEGFP signal levels are here to show that the signal levels did not affect the fuzziness or how dispersed UTX was. **o.** Granularity of the mEGFP signal was plotted as mean \pm SD. n = 28 *Utx* WT and 38 Δ cIDR cells, each containing two different clones shown in panel n. Granularity, or mEGFP signal fluctuation, was calculated as SD of the relative signal intensity of a line plot (normalized to the mean signal intensity of the entire line) at each pixel on the line through an individual cell. P value by two-sided *t*-test. The line plot was generated by ImageJ, and an example is shown as the yellow line on a mEGFP- Δ cIDR clone in n. **p.** Venn diagrams for genes that, compared to the corresponding WT embryoid bodies, were downregulated (left) or upregulated (right) by KO (cyan) or Δ cIDR (salmon, two clones) in late embryoid bodies (day 8 or 9 in differentiation). Fold changes are indicated. From RNA-seq of two 2 independent differentiation assays. P values by two-sided Fisher's exact test. **q.** Significantly enriched Biological Processes in embryonic development for genes downregulated (↓) or upregulated (↑) over 2 fold by *Utx* KO or Δ cIDR in embryoid bodies late in differentiation, compared to the corresponding WT embryoid bodies. By Metascape. P values by two-sided hypergeometric test.

Extended Data Fig. 5 UTX co-condenses with and enhance chromatin recruitment of MLL4 and p300 in cells.

a. A schematic showing the domain compositions of MLL4, MLL4_f, and MLL4C(N). **b.** Left, a schematic showing the optoDroplet system used for the right, which are confocal images of 293T cells expressing MLL4C(N)-mCherry-Cry2, before and after light activation. **c.** Left, a schematic showing the optoDroplet system used for the right, which are confocal images of 293T cells expressing UTX-mCherry-Cry2, before and after light activation. **d.** Left, a schematic showing the optoDroplet system used for the right, which are confocal images of 293T cells expressing both UTX (WT or indicated mutant)-mCherry-Cry2 and MLL4C(N)-EGFP, before and after light activation. **e.** Left, a schematic showing the LacO-array system. Middle, fluorescence microscopy images of LacO-cells transfected with EGFP-LacI (control) or UTX (WT or mutant)-EGFP-LacI, each together

with p300₍₂₅₁₋₅₅₀₎-mCherry. Right, recruitment coefficients of p300₍₂₅₁₋₅₅₀₎ relative to the LacO-bound control or UTX (WT or mutant) are determined as described in Methods and plotted as average \pm SD. n = 11, 19, 17, 24 cells for control, WT, ΔcIDR, and ΔTPR, respectively. P values by one-way ANOVA followed by Tukey's post hoc test.

Extended Data Fig. 6 Purified UTX and MLL4 or p300 regions form co-condensates through specific interaction mediated by TPR.

a. Coomassie blue staining image of purified EGFP-UTX (419-848), EGFP-UTX (1-848), and mCherry-MLL4C(N). **b.** Fluorescence images of purified proteins individually diluted to indicated concentrations. **c.** Fluorescence images of purified proteins mixed and diluted to the final concentration as indicated, equal to the concentrations in b. **d.** Fluorescence images of purified proteins individually diluted to indicated concentrations. **e.** Fluorescence images of purified proteins mixed and diluted to the final concentration equal to that of the corresponding protein in d. The exposure time for GFP in the top row was longer than those in the other two rows to adjust the GFP brightness due to the much lower concentration of the GFP-fused proteins in the top row. **f.** Recruitment coefficient for MLL4C(N) was determined as the relative ratio of the mCherry signal amount inside the GFP condensates over that inside the mCherry condensates, and plotted as mean \pm SD. n = 3 independent fields each. Note that each field had ~100 total droplets. **P<0.01 by one-way ANOVA followed by Tukey's post hoc test. **g.** Domain schematic of p300. P300 (251-550) containing ZF1 domain is labeled. **h.** Fluorescence images of purified proteins mixed to the indicated final concentrations.

Extended Data Fig. 7 UTX condensates enrich histone modification activities.

a. Schematic of the procedures to isolate the UTX (WT or ΔcIDR) and MLL4_f proteins for in vitro histone methylation (HMT) and imaging analyses. Note that all HMT assays involved endogenous MLL3/4 associated with isolated UTX. **b.** Immunoblotting for UTX in the UTX (WT

or ΔcIDR) pulldowns, along with indicated ng of purified EGFP-UTX (1-848). This allowed an estimate of the concentration of WT-GFP or ΔcIDR-GFP to be ~40 nM in the pulldowns. **c.** Fluorescence images of the UTX (WT or ΔcIDR) pulldowns in the condensation buffer, with ~25 nM UTX protein and 150 mM NaCl. **d.** Immunoblotting for the indicated proteins in the isolated UTX (WT or ΔcIDR) complexes. **e.** Representative results of in vitro HMT assays using the UTX (WT or ΔcIDR)-GFP pulldown. UTX and indicated histone modifications at indicated time points of HMT assay were determined by Immunoblotting. Total Histones are shown by Ponceau S staining. Right, Relative H3K4me1 increase (from time 0) is plotted as mean from 2 independent experiments. **f.** Representative results of in vitro HMT assays using the UTX (WT or 555*) pulldown. Top, immunoblotting by indicated antibodies and Ponceaus S staining for total histones in HMT assay. Bottom, H3K4me1 activity was determined as relative H3K4me1 increase from preexisting methylation (in control) and plotted as mean ± SD. n = 3 independent experiments. P values by two-sided *t*-test. **g.** Immunoblotting using anti-FLAG antibody for the isolated UTX (WT or ΔcIDR) complexes, MLL4_f-mCherry, or their mixture as indicated. UTX (WT or ΔcIDR) and MLL4_f all had FLAG tag. **h.** Fluorescence microscopy images of indicated protein complex, either as individual protein (top row) or mixed as indicated (bottom two rows) in the condensation buffer. The concentration of each protein in the mixture was the same as that of the corresponding individual protein shown on top row. **i.** Immunostaining for Myc (for Myc-tagged UTX), H3K27me3, and DAPI in COS-7 cells transfected with indicated UTX WT or mutant construct. Cells with UTX overexpression (indicated by arrows) were scored for the signal intensity of H3K27me3 and quantified in **j.** **j.** Quantification of the cells showing Decrease (↓), no change (-), or increase (↑) of H3K27me3. Shown are the percentage of cells showing the indicated changes in one out of a total of five biological repeats. In parenthesis is the number of cells showing the changes/total number of cells scored. In each repeat, around 100 cells were scored for each sample.

Extended Data Fig. 8 UTX condensates regulate genomic histone modifications and enhancer transcription.

a. A Venn diagram for genomic sites bound by UTX WT and Δ cIDR. The overlap number refers to the number of WT peaks that overlapped with Δ cIDR peaks. **b.** Genomic features of all 15,554 UTX WT-bound sites. The number of sites for each feature is shown in the parenthesis. **c.** Number and percentage of the genomic sites bound by either UTX WT (normal) or specifically by Δ cIDR (aberrant, not bound by WT) that overlap (within indicated distance) with any of the activating histone marks (H3K4me1, 2, 3, and H3K27ac) merged across all cell samples. **d.** Heatmap (left) and composite plot (right) for all 15,554 UTX WT-binding peaks in indicated cells. Heatmap rows were ranked by UTX peak levels (MACS2 peaks q-value, high to low). UTX ChIP-seq signals per 20bp non-overlapped bins covering 4kb regions centered at UTX peaks were shown. Details in Supplementary Table 3. **e, f.** Heatmap (left) and composite plot (right) for H3K4me1 (e) or H3K27ac (f) peaks in indicated cells at the UTX-binding sites (number indicated) with higher H3K4me1 (e) or H3K27ac (f) intensity in WT than vector expressing cells. Heatmap rows were ranked by UTX peak significance (MACS2 q-value, high to low). ChIP-seq signals from indicated histone modification per 20bp non-overlapped bins covering 4kb regions centered at UTX peaks were shown. **g–j.** Left, scatter plot using the indicated H3K4me3 (g), H3K4me2 (h), H3K27me3 (i), and MLL4 binding (j) signal level change in cells indicated on the x and y axes, for all 15,554 UTX WT-bound sites. Black dashed line indicates the linear fit with correlation coefficient (r) labeled. Right, composite plot for indicated modification or binding signals centered at UTX peaks in indicated cells. **k.** Scatter plots using the indicated histone modification signal level change in cells indicated on the x and y axes, for all 15,554 UTX WT-bound sites. Black dashed line indicates the linear fit with correlation coefficient (r) labeled. **l.** Top, scatter plot for PRO-seq signal change at all 4,806 UTX WT-bound enhancers in indicated cells. Black dashed line indicates the linear fit with correlation coefficient (r) labeled. Bottom, composite plot for the PRO-seq signals centered at PRO-seq peaks in indicated cells. **m.** Heatmap (left) and composite plot (right) for PRO-seq signals at UTX WT-bound enhancers that had increased (top row) or decreased (bottom row) PRO-seq signals over 1.2 fold in WT compared to vector expressing cells, in indicated cells. PRO-seq signals centered at PRO-seq peaks were shown.

Extended Data Fig. 9 UTX condensates regulate long-range chromatin interactions.

a. Distribution of types of genomic interactions. P, promoter ($TSS \pm 2\text{kb}$); E, enhancer (promoter non-overlapped, union DNase I hypersensitive sites); Desert: all other interactions, <1% in all. **b.** Distribution of the chromatin loop size in indicated cells. A loop is the connection of two different 5 kb chromatin regions by high-confidence contact reads. **c.** Number of chromatin loops identified in indicated cells. For simplicity, Venn diagrams for only Vector, WT, and $\Delta c\text{IDR}$ cells are shown. **d.** Box plots for the number of normalized HiChIP reads per loop for identified loops in indicated cells. The normalized read counts in a loop represents the relative intensity of the interaction. For H3K4me3 HiChIP (left), $n = 57632, 67755, 88993, 95696, 90566$ loops in Vector, WT, $\Delta c\text{IDR}$, UTX-eIF_{IDR}, $\Delta c\text{TPR}$ samples, respectively. For H3K27ac HiChIP (right), $n = 20912, 27794, 37811, 31232, 26556$ loops in Vector, WT, $\Delta c\text{IDR}$, UTX-eIF_{IDR}, $\Delta c\text{TPR}$ samples, respectively. **e.** Scatter plots for all 26,237 genes using the promoter H3K27ac HiChIP DCI of cells indicated on the x and y axes. Dashed line shows the linear fit with indicated correlation coefficient (r). Red and blue dots represent 479 and 112 genes with DCI of WT over vector > 5 and < -5 , respectively. **f.** Box plots for DCI scores comparing cells indicated on x axis for genes represented by the red (left) and blue (right) dots from e. P values are shown for data significantly greater or less than 0 by two-sided one sample t -test. Left, $n = 479$ genes each; right, $n = 112$ genes each. **g.** Box plots for DCI scores at promoter ($TSS \pm 2\text{kb}$) of cell samples (indicated on top of each plot) for genes whose expression was up- or downregulated (indicated on x axis) in the corresponding cells (indicated on top of each plot). Genes were selected as those both have expression and DCI data. The trend of change and the P value between indicated two groups of genes are shown. $n =$ indicated number on x axis. **h.** Gene pathway analysis using DCI scores from indicated HiChIP data. Red shows pathways relevant to cancer or leukaemia inhibition among the top 20 significantly enriched Biological Processes for genes with promoter chromatin interactions substantially increased in WT over vector expressing cells. Blue shows pathways particularly relevant to cell growth and survival among the top 20 significantly enriched Biological Processes for genes with

promoter chromatin interactions substantially decreased in WT over vector expressing cells. All by Metascape. Details in Supplementary Table 5. **i, j.** BART3D results showing transcriptional regulators that were likely to be associated with genomic regions that had increased (i) or decreased (j) chromatin interactions in WT over vector expressing cells from indicated HiChIP datasets. Top 10 ranked transcriptional regulators are labeled in red and their names are shown. For all box plots, data are median (horizontal line), 25–75th percentiles (box) and 1.5 times the interquartile range recorded (whiskers). P values by two-sided t-test for all except by hypergeometric test for h, and from BART3D for **i, j.**

Extended Data Fig. 10 Integrated genomic analysis for example genes.

a. Normalized RNA-seq reads of indicated genes that are shown for integrated genomic analysis, in Vector (V), WT (W), ΔcIDR (c), UTX-eIF_{IDR} (e), ΔcTPR (T) samples. One repeat of the Vector sample was set to 1. *HK2* gene (boxed) differs from the other genes in that it was repressed by UTX WT but not ΔcIDR. **b.** Integrated genomic analysis for *MX2-MX1* cluster, *OAS1-3-2* cluster, *OASL*, *UBE2W-LY96* cluster, *KCNJ2*, and *HK2*. as examples for UTX condensates-regulated gene expression associated with chromatin interactions. Top, PRO-seq 3' end coverage tracks (red, + strand; blue, - strand), with the same scale for all samples. A higher *MX2* transcriptional activity in WT than ΔcIDR cells is shown by the denser WT PRO-seq signals when IGV-zoomed in in the dashed box on top. Also note the higher transcriptional activity of *LY96* in UTX WT than vector and ΔcIDR cells, as shown by the PRO-seq signal density and the - strand eRNA (in green oval) at both *LY96* promoter and gene-downstream enhancer. Green arrows show the higher H3K4me1 or H3K27ac signal in WT than the other two cell samples at *OASL* promoter. PRO-seq confirmed transcriptional changes and, together with the H3K4me1 and H3K27ac peaks, revealed actively and bi-directionally transcribing enhancers. Note that H3K27ac and/or H3K4me1 were increased at *MX2*, *OASL*, and *KCNJ2* promoters or enhancers upon WT but not ΔcIDR expression. Bottom, H3K27ac HiChIP arc plots of loops are shown on WashU EpiGenome Browser. All arc plots have the same scale for statistical significance (arc darkness) of the loop calls. Black arrows indicate normal loops in WT-

expressing cells that were lost in vector, Δ cIDR, or Δ TPR cells. Cyan arrows indicate loops normally suppressed by UTX WT but aberrantly formed in vector, Δ cIDR, or Δ TPR cells.

Extended Data Fig. 11 UTY is a weaker tumour suppressor than UTX due to its altered physical property of condensates.

a. Immunoblotting by indicated antibody of total lysates from THP-1 cells electroporated with empty vector or indicated Myc-tagged UTX WT or mutants. **b.** Amino acid sequence alignment of UTX and UTY cIDR. Middle row shows identical residues (by letter) and conservative mutations (by “+”). Y and F are labeled with green box. Negatively charged residues (E and D) are in red font, and positively charged residues (R and K) are in blue font. Charged residues that are unique in UTX or UTY cIDR are bold and underlined. **c.** Net charge per residue for cIDR of UTX and UTY, determined by CIDER. **d.** Images of Ni resin after binding with bacterial lysates during purification of indicated proteins. Red arrows point to the unusual aggregation of the resin bound to UTY cIDR. **e.** Coomassie blue staining image of purified UTX cIDR and UTY cIDR with indicated fluorescence tag. A non-specific band is indicated by the asterisk. **f.** Indicated proteins all at 30 μ M protein and in 150 mM NaCl were subjected to centrifugation. The unspun mixture (“mixed”), supernatant (“super”), and pellets were resolved on SDS-PAGE gel followed with coomassie blue staining. Partition Percentage is plotted (right) as described in Methods as mean \pm SD of 3 repeats. *** $P<0.001$ by two-sided *t*-test. **g.** Fluorescence and DIC images of mCherry-UTX cIDR or mCherry-UTY cIDR at indicated concentrations and time points after dilution into the condensation buffer. **h.** Fluorescence microscopy images of mEGFP-UTX cIDR or mEGFP-UTY cIDR at increasing concentrations of protein and NaCl. A shorter exposure time was used for all images of 20 μ M protein. **i.** Fluorescence Image and line carpet of mEGFP-UTX cIDR and mEGFP-UTY cIDR droplets. **j.** Line RICS autocorrelation curves of 50 μ M mEGFP-UTX cIDR and mEGFP-UTY cIDR at early (< or = 100 min) or late (> 170 min) time of phase separation after dilution into the condensation buffer, shown as mean \pm SD. n_{UTX early} = 8, n_{UTY early} = 9, n_{UTX late} = 8, n_{UTY late} = 9 droplets. **k.** Representative fluorescence

microscopy images for FRAP assays of UTX or UTX-UTY_{IDR} fused to EGFP in nuclei following transfection into 293T cells. The photobleached focus was boxed and also amplified in the images on the right. Photobleached at time 0.

Extended Data Fig. 12 UTX cancer-associated missense mutations alter condensate properties.

All panels in this figure used 20 μM (unless otherwise indicated in a) mEGFP-UTX (419-848) WT or mutants as indicated. S781Y and S814F mutation was found in one and two patients, respectively, and both were predicted by ICGC to have high functional impact. 5M is a compound missense mutation containing S674Y, S781Y, H808Y, S814F, and S818L, all found in one cancer patient. See [Methods](#) “Cancer database analysis for mutations” for details. **a.** Pictures and relative turbidity as OD600 normalized to the value of WT and plotted as mean ± SD. n = 3 independent experiments. Bottom, coomassie blue staining of the corresponding proteins, sometimes with finely adjusted (10%) different concentrations (for S814F) to ensure fair comparison with WT. **b.** Representative fluorescence microscopy and DIC images at indicated times in condensation. Scale bar, 10 μm. **c.** FRAP recovery curves as mean ± SD at indicated early and late times in condensation. Top two panels, n = 13 droplets for WT and S781Y each at early times, and n = 9 for each of them at late times. Bottom two panels, n = 9, 10, 10 for WT, 5M, S814F, respectively, at early times, and n = 8 for each of them at late times. The statistics is for the last time points, but, for WT and S814F at late condensation times, is for 0-143 seconds after photobleaching. **d–f.** Line RICS assays for WT (n = 21 droplets), S781Y (n = 12), S814F (n = 10), and 5M (n = 11) over the time range of 0-100 min in condensation. **d.** Fluorescence Images and line carpets. **e.** Line RICS autocorrelation curves. **f.** Plot showing the mean ± SD of the diffusion coefficients. ***P<0.001, and ****P<0.0001 by two-sided t-test for all.

Supplementary information

Supplementary Figure 1

This file contains gel source images

Reporting Summary

Supplementary Table 1 Changes of gene expression by expression of WT UTX and mutants in leukaemia cells

RNA-seq analyses (TPM) for selected genes in THP-1 cells that were stably transduced with vector, UTX WT, Δ cIDR, UTX(eIF_{IDR}), Δ TPR or UTX(FUS_{IDR}) (in sheet 1), or with vector, UTX WT or MT2 (in sheet 2). Two biological repeats each. These genes were significantly (adjusted $P < 0.05$, from DESeq2) up- or down-regulated (1.25-fold each) by WT compared with vector control.

Supplementary Table 2 Changes of gene expression by deletion of endogenous UTX or cIDR in mouse ES cells and embryoid bodies

Normalized reads by RNA-seq analyses for all genes in the mouse genome in the WT, genomically edited *Utx*-KO and Δ cIDR ES cells and late embryoid body from the differentiation assays. Sorted by the mean reads in all four ES cell samples. The differential expression details including the log₂ fold change and adjusted P values were from DESeq2.

Supplementary Table 3 Genome-wide binding of UTX

Based on UTX ChIP-seq for THP-1 cells that were stably transduced with vector, UTX WT or Δ cIDR (DEL in the table). UTX peaks were called using vector as control. The ‘ChIP-seq binding at WT UTX’ sheet also shows the histone modification and UTX and MLL4 binding signals at the WT-bound sites in different cells. The peak details including q -values were from MACS2.

Supplementary Table 4 HiChIP analysis

HiChIP analysis in THP-1 cells that were stably transduced with vector, UTX WT, ΔcIDR, UTX(eIF_{IDR}) or ΔTPR. There are four sheets in this file. Sheet ‘QC’ lists numbers of reads of indicated characteristics. The other three sheets show called chromatin loops in HiChIP assays indicated by the sheet name. The raw count and FDR were from MAPS results, and the raw count was normalized by per 100 million mapped pairs after duplicate removal (from MAPS QC file).

Supplementary Table 5 Integrated genomic analysis for genes

Integrated genomic analysis for all genes in the genome in THP-1 cells that were stably transduced with vector, UTX WT, ΔcIDR (DEL), UTX(eIF_{IDR}) (EIF) or ΔTPR (TPR). Each gene is shown with its expression levels (from RNA-seq) and DCI scores. The differential expression details including the log₂ fold change and adjusted *P* values were from DESeq2.

Supplementary Table 6 Oligonucleotide information

Information for oligonucleotides used for CRISPR–Cas9-mediated genome editing in this work.

Supplementary Video 1 UTX cIDR droplets in vitro are highly dynamic and can fuse

Purified UTX cIDR formed droplets immediately upon dilution into the condensation buffer and reached 30 μM protein and 150 mM NaCl. The red arrows point to fusion events immediately prior to the fusion at different times throughout the video. Played at 100× speed.

Supplementary Video 2 FRAP assays for intracellular UTX

Representative FRAP assays for UTX–EGFP transfected into 293T cells. Played at 5× speed.

Supplementary Video 3 OptoDroplets by MLL4C(N)–mCherry–Cry2 and UTX–EGFP

Example 1 for OptoDroplets formed by MLL4C(N)–mCherry–Cry2 and UTX (WT, ΔcIDR or ΔTPR)–EGFP in 293T cells. Played at 5× speed.

Supplementary Video 4 OptoDroplets by MLL4C(N)–mCherry–Cry2 and UTX–EGFP

Example 2 for OptoDroplets formed by MLL4C(N)–mCherry–Cry2 and UTX (WT, ΔcIDR or ΔTPR)–EGFP in 293T cells. Played at 5× speed.

Supplementary Video 5 OptoDroplets by UTX–mCherry–Cry2 and MLL4C(N)–EGFP

Example for OptoDroplets formed by UTX (WT, ΔcIDR or ΔTPR)–mCherry–Cry2 and MLL4C(N)–EGFP in 293T cells. Played at 5× speed.

Supplementary Video 6 FRAP assays for UTX and UTY cIDR in vitro

Representative FRAP assays for purified mEGFP–UTX or UTY cIDR in vitro. Played at 5× speed.

Source data

Source Data Fig. 1

Source Data Extended Data Fig. 3

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Shi, B., Li, W., Song, Y. *et al.* UTX condensation underlies its tumour-suppressive activity. *Nature* **597**, 726–731 (2021).
<https://doi.org/10.1038/s41586-021-03903-7>

- Received: 27 April 2020
- Accepted: 12 August 2021
- Published: 15 September 2021
- Issue Date: 30 September 2021
- DOI: <https://doi.org/10.1038/s41586-021-03903-7>

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

[Protein condensates provide a platform for controlling chromatin](#)

- David Lara-Astiaso
- Brian J. P. Huntly

News & Views 15 Sept 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03903-7>

- Article
- Open Access
- [Published: 15 September 2021](#)

Structure-based classification predicts drug response in *EGFR*-mutant NSCLC

- [Jacqulyne P. Robichaux¹](#),
- [Xiuning Le¹](#),
- [R. S. K. Vijayan²](#),
- [J. Kevin Hicks³](#),
- [Simon Heeke ORCID: orcid.org/0000-0002-5916-534X¹](#),
- [Yasir Y. Elamin ORCID: orcid.org/0000-0002-5312-909X¹](#),
- [Heather Y. Lin⁴](#),
- [Hibiki Udagawa¹](#),
- [Ferdinandos Skoulidis¹](#),
- [Hai Tran¹](#),
- [Susan Varghese¹](#),
- [Junjin He¹](#),
- [Fahao Zhang¹](#),
- [Monique B. Nilsson¹](#),
- [Lemei Hu¹](#),
- [Alissa Poteete¹](#),
- [Waree Rinsurongkawong⁵](#),
- [Xiaoshan Zhang⁶](#),
- [Chenghui Ren⁶](#),
- [Xiaoke Liu^{1,7}](#),
- [Lingzhi Hong¹](#),
- [Jianjun Zhang ORCID: orcid.org/0000-0001-7872-3477¹](#),
- [Lixia Diao ORCID: orcid.org/0000-0001-6995-4492⁸](#),
- [Russell Madison⁹](#),
- [Alexa B. Schrock⁹](#),
- [Jennifer Saam¹⁰](#),
- [Victoria Raymond¹⁰](#),
- [Bingliang Fang⁶](#),

- [Jing Wang](#) [ORCID: orcid.org/0000-0002-5398-0802⁸](#),
- [Min Jin Ha⁴](#),
- [Jason B. Cross²](#),
- [Jhanelle E. Gray¹¹](#) &
- [John V. Heymach](#) [ORCID: orcid.org/0000-0001-9068-8942¹](#)

Nature volume **597**, pages 732–737 (2021)

- 11k Accesses
- 175 Altmetric
- [Metrics details](#)

Subjects

- [Non-small-cell lung cancer](#)
- [Targeted therapies](#)

Abstract

Epidermal growth factor receptor (*EGFR*) mutations typically occur in exons 18–21 and are established driver mutations in non-small cell lung cancer (NSCLC)^{1,2,3}. Targeted therapies are approved for patients with ‘classical’ mutations and a small number of other mutations^{4,5,6}. However, effective therapies have not been identified for additional *EGFR* mutations. Furthermore, the frequency and effects of atypical *EGFR* mutations on drug sensitivity are unknown^{1,3,7,8,9,10}. Here we characterize the mutational landscape in 16,715 patients with *EGFR*-mutant NSCLC, and establish the structure–function relationship of *EGFR* mutations on drug sensitivity. We found that *EGFR* mutations can be separated into four distinct subgroups on the basis of sensitivity and structural changes that retrospectively predict patient outcomes following treatment with *EGFR* inhibitors better than traditional exon-based groups. Together, these data delineate a structure-based approach for defining functional groups of *EGFR* mutations that can effectively guide treatment and clinical trial choices for patients with *EGFR*-mutant NSCLC and suggest that a structure–function-based approach may improve the prediction of drug sensitivity to targeted therapies in oncogenes with diverse mutations.

[Download PDF](#)

Main

Patients with classical EGFR mutations (L858R or exon 19 deletions (Ex19del)) show marked improvements in clinical outcomes when treated with first-, second- or third-generation tyrosine kinase inhibitors (TKIs)^{4,5,6,11}. Other *EGFR* mutations in the kinase domain (exons 18–21) have also been established as oncogenic drivers of NSCLC⁸. Patients with atypical *EGFR* mutations show heterogeneous and reduced responses to EGFR inhibitors^{1,3,7,8,9,10}, including osimertinib¹². Atypical *EGFR* mutations with US Food and Drug Administration (FDA)-approved treatments are *EGFR* S768I, L861Q and G719X, for which afatinib was deemed effective on the basis of retrospective studies^{13,14,15}, and the EGFR/MET bispecific antibody amivantamab for exon 20 insertions (Ex20ins)¹⁶. There are no clear established guidelines for EGFR TKI treatment for patients with atypical *EGFR* mutations without an FDA-approved TKI, often resulting in patients receiving chemotherapy. Clinical trial design and treatment of patients with atypical *EGFR* mutations often rely on mutated-exon location to predict treatment, although heterogeneity in drug sensitivity across a single exon has been observed^{1,8,17}. Therefore, there is an unmet clinical need to establish a system for classifying *EGFR* mutations that is predictive of drug sensitivity and for more robust clinical trial design.

Clinical outcomes for atypical mutations

To characterize the molecular landscape of *EGFR*-mutant NSCLC, we used 5 patient databases with genomic profiling ([Methods](#)), representing 16,715 patients with *EGFR*-mutant NSCLC. There were 11,619 patients in whom primary and/or co-occurring mutations were recorded on a per-patient basis (Extended Data Fig. [1a](#)). Among those patients, 67.1% had classical *EGFR* mutations (L858R and/or Ex19del with or without T790M); 30.8% had atypical *EGFR* mutations, including Ex20ins (9.1%), atypical mutations (12.6%), or a complex mutation including an atypical mutation (9.1%); and 2.2% had a classical mutation with T790M and an atypical mutation (Fig. [1a](#), Extended Data Fig. [1b](#)). Atypical *EGFR* mutations ($n = 7,199$) occurred primarily in exons 18 (23.7%) and 20 (20.9% insertions and 19.2% point mutations; Fig. [1b](#)). Prevalent hotspots for atypical mutations were the P-loop (L718–V726, 13.6%) and the C-terminal loop of the α C-helix (A767–G779, 29.4%, Fig. [1c](#)).

Fig. 1: Atypical *EGFR* mutations are associated with worse patient outcomes.

 **figure1**

a, Percentage of patients with NSCLC containing classical and atypical *EGFR* mutations ($n = 11,619$ patients). Classical *EGFR* mutations are L858R, T790M and various Ex19dels ([Methods](#)). **b**, Percentage of atypical *EGFR* mutations observed in patients with NSCLC ($n = 7,199$ mutations). Atypical *EGFR* mutations are defined as

non-classical, non-synonymous mutations. **c**, Lollipop plot of frequency of atypical *EGFR* mutations in patients with NSCLC ($n = 7,199$ mutations). *EGFR* mutations associated with acquired drug resistance are highlighted in red. **d**, Kaplan–Meier plot of time to treatment failure (TTF) (time from TKI commencement until radiologic progression, discontinuation, or death) of patients with NSCLC tumours containing classical ($n = 245$ patients) or atypical ($n = 109$ patients) *EGFR* mutations after EGFR TKI treatment. **e**, Forest plot of HR calculated from Kaplan–Meier plots of patients with various subsets of atypical mutations or classical *EGFR* mutations. In **d**, **e**, HR and P value were calculated using two-sided Mantel–Cox log-rank tests. Data are HR $\pm 95\%$ confidence interval. All atypical, $n = 109$; all atypical without Ex20ins, $n = 97$; exon 18, $n = 29$; exon 19, $n = 22$; exon 20, $n = 41$; exon 21, $n = 18$. NS, not significant.

Source data.

To assess the effect of atypical *EGFR* mutations on patient outcomes, we determined the time to treatment failure¹⁸ (TTF) of patients with NSCLC containing classical or atypical *EGFR* mutations. When treated with an EGFR TKI, patients with atypical *EGFR* mutations had a shorter TTF compared with patients with classical *EGFR* mutations (Fig. [1d](#), hazard ratio (HR) = 1.8, $P < 0.0001$), even when patients with Ex20ins were excluded from the analysis (Fig. [1e](#), HR = 1.6, $P < 0.0001$) or when patients were stratified by mutation exon location (Fig. [1e](#), Extended Data Fig. [1c](#)). When patients were stratified by TKI treatment, those with classical *EGFR* mutations had a longer TTF than those with atypical *EGFR* mutations when treated with first-generation (HR = 1.9, $P = 0.0005$) or third-generation TKIs (HR = 3.0, $P < 0.0001$) (Extended Data Fig. [1d,e](#)). A similar trend was observed for second-generation TKIs; however, the difference was not statistically significant (HR = 1.7, $P = 0.10$) (Extended Data Fig. [1f](#)). Validating these findings in the cBioPortal database, patients with atypical *EGFR* mutations had a shorter progression free interval¹⁹ and overall survival, irrespective of treatment (Extended Data Fig. [1g,h](#)).

Structural groups predict drug response

We generated a panel of 76 cell lines expressing *EGFR* mutations spanning exons 18–21 and screened these cell lines against 18 EGFR inhibitors representing first- (non-covalent), second (covalent) and third- (covalent, T790M targeting) generation and Ex20ins-active TKIs (Supplementary Table [1](#)). Using hierarchical clustering of in vitro selectivity over WT EGFR and mutational mapping of *EGFR* mutations, we observed four distinct subgroups of *EGFR* mutations: classical-like mutations that were distant from the ATP-binding pocket (Extended Data Fig. [2a](#)), T790M-like mutations in the hydrophobic core (Extended Data Fig. [2b](#)), insertions in the loop at the C-terminal end of the α C-helix in exon 20 (Ex20ins-L; Extended Data Fig. [2c](#)), and mutations on the interior surface of the ATP-binding pocket or C-terminal end of the α C-helix, which

were predicted to be P-loop and α C-helix compressing (PACC) (Fig. 2a, Extended Data Fig. 2d). Supervised heat maps of mutant/wild-type ratios by exon location (Extended Data Fig. 3a) and structure–function groups (Extended Data Fig. 3b) showed distinct differences, suggesting that structure–function-based groups better defined groups of mutations by drug sensitivity than exon-based classification. To test this hypothesis, we calculated the correlations of drug sensitivity and selectivity for each mutation to the predicted drug sensitivity by exon or structure–function groups (Extended Data Fig. 4a) and then compared the median rho value of each correlation for both groups. We found that structure–function-based groups were more predictive of mutation sensitivity than exon-based groups ($P < 0.0001$) (Fig. 2b). We used a secondary approach employing machine learning to analyse data by classification and regression trees (CART) and determine variable importance^{20,21} (Extended Data Fig. 4b). Structure–function-based groups had a higher variable importance than exon-based groups, suggesting that structure–function-based groups were more predictive of which mutational groups would be most sensitive to a particular drug compared with exon-based groups ($P < 0.0001$) (Fig. 2c). We validated these findings without T790M mutations (Extended Data Fig. 4c, d), and structure–function-based groups remained more predictive of mutation and drug sensitivity than exon-based groups ($P = 0.0034$ and $P < 0.0001$, respectively) (Extended Data Fig. 4e, f).

Fig. 2: EGFR mutations can be separated into four distinct subgroups.

 figure2

a, Heat map with unsupervised hierarchical clustering of log(mutant/wild type (WT)) ratios from Ba/F3 cells expressing indicated mutations after drug treatment. To determine the mutant/WT ratio, half-maximal inhibitory concentration (IC_{50}) values for each drug and cell line were calculated and then compared to the average IC_{50} values for Ba/F3 cells expressing wild-type *EGFR* (+10 ng ml⁻¹ EGF). Squares are representative of the median of $n = 3$ replicates. The order of co-occurring mutations

was assigned arbitrarily. Groups were assigned on the basis of structural predictions ([Methods](#)). Gen, generation. **b**, Dot plot of Spearman's rho values for correlations of mutations versus exon-based group averages or structure–function-based averages for each drug. Dots are representative of rho value of each mutation; bars show mean \pm s.d., $n = 77$ cell lines or mutations. **c**, Dot plot of variable importance calculated from CART. Dots are representative of variable importance for each drug; bars show median + 95% confidence interval of variable importance for all drugs ($n = 18$ drugs) (Supplementary Table [2](#)). In **b**, **c**, P value was determined using a paired two-sided t -test.

[Source data](#).

Classical-like, atypical *EGFR* mutations were predicted to have little effect on the overall structure of EGFR compared with wild-type EGFR (Extended Data Fig. [5a–d](#)) and were sensitive and selective for all classes of EGFR TKIs, particularly third-generation TKIs, *in vitro* and *in vivo* (Extended Data Fig. [5e–g](#)). Mutations and assigned groups are in Supplementary Table [4](#).

Exon 20 mutations are heterogenous

Studies showed that exon 20 mutations are heterogenous in their response to TKIs^{[10,22](#)}. Insertions in the α C-helix (for example, an FQEA insertion at A763 (A763insFQEA)) were pan-sensitive to EGFR TKIs^{[23,24](#)}, whereas those in the loop following the α C-helix (A767–C775) were not^{[25,26](#)}, and the T790M mutation was sensitive to third- but not first- or second-generation TKIs. We found that most exon 20 point mutations were PACC mutations; that exon 20 insertions in the α C-helix were classical-like mutations; and, that the remainder of exon 20 insertions occurring in the C-terminal loop of the α C-helix were a distinct subgroup: exon 20 loop insertions (Ex20ins-L) (Fig. [2a](#)). Ex20ins-L were sensitive only to select second-generation TKIs (that is, poziotinib and tarlox-TKI) and Ex20ins-active TKIs *in vitro* and *in vivo* (Fig. [2a](#), Extended Data Fig. [6a–c](#)). However, even within Ex20ins-L mutations, some degree of heterogeneity in drug sensitivity was observed (Fig. [2a](#), Extended Data Fig. [6a](#)). Using a panel of an additional 15 Ba/F3 cell lines expressing Ex20ins-L mutations spanning A767–V774, we found that Ex20ins-L mutations could be subdivided into two subgroups: near- and far-loop Ex20ins (Extended Data Fig. [6d](#)). Exon 20 near-loop insertions (Ex20ins-NL) were more sensitive to second-generation and Ex20ins-active TKIs compared with exon 20 far-loop insertions (Ex20ins-FL) ($P = 0.0025$ and $P = 0.027$, respectively) (Extended Data Fig. [6e](#)). These data exemplify that mutations within an exon are heterogenous and that an exon-based classification is unlikely to be optimal for guiding treatment decisions.

Drug repurposing for resistant mutations

Although all T790M-like mutants had at least one mutation in the hydrophobic core, there were two distinct subgroups of T790M-like mutants—third-generation TKI sensitive (T790M-like-3S) and third-generation TKI resistant (T790M-like-3R) (Extended Data Fig. 7a). Previous reports have shown that protein kinase C²⁷ (PKC) and anaplastic lymphoma kinase^{28,29} (ALK) inhibitors exhibit off-target activity for EGFR mutations including T790M, and the non-covalent nature of these compounds predict that they retain activity in mutations that interrupt covalent binding. T790M-like-3S mutants had high selectivity for third-generation TKIs and some Ex20ins-active inhibitors and moderate selectivity for ALK and PKC inhibitors (Extended Data Fig. 7b). T790M-like-3R mutants, complex mutations comprising T790M and a known drug-resistance mutation (that is, C797S³⁰, L718X³¹ or L792H^{18,31}), were resistant to classical EGFR TKIs but retained selectivity for select ALK and PKC inhibitors such as brigatinib or midostaurin (Extended Data Fig. 7c). These data support expanding testing of ALK and/or PKC inhibitors or development of novel non-covalent inhibitors for the broader group of T790M-like-3R mutations.

Second-generation TKIs inhibit PACC mutations

PACC mutations comprise mutations spanning exons 18–21 including G719X, L747X, S768I, L792X and T854I and were predicted to alter the orientation of the P-loop or α C-helix (Extended Data Fig. 8a,b). In silico analysis of the interaction of osimertinib with PACC mutations G719S and L718Q predicted that changes in the orientation of the P-loop alter the position of TKI stabilization points tilting the indole ring of osimertinib away from the P-loop, destabilizing drug binding (Extended Data Fig. 8c, d). By contrast, second-generation TKIs do not interact with the P-loop of EGFR and maintain interaction points in the hydrophobic cleft (Extended Data Fig. 8d,e). When we compared the selectivity of EGFR TKIs for PACC mutations, we found that second-generation TKIs were significantly more selective for PACC mutations than any other TKI class (Fig. 3a). In vivo, we also observed that NSCLC patient-derived xenografts (PDXs) containing G719A mutations were resistant to the third-generation TKI osimertinib, but most sensitive to the second-generation TKI poziotinib (Fig. 3b, Extended Data Fig. 8f). Notably, a patient with a complex PACC mutation, E709K/G719S, saw clinical benefit and tumour shrinkage with afatinib treatment after progressing on osimertinib (Extended Data Fig. 8g).

Fig. 3: PACC mutations are robustly sensitive to second-generation TKIs.

 **figure3**

a, Dot plot of mutant/WT IC₅₀ values of Ba/F3 cells expressing PACC mutations. **b**, Tumour growth curves for PDXs containing EGFR G719A PACC mutation treated with TKIs five days per week. Symbols show mean tumour volume \pm s.e.m., $n = 5$ mice. Osi 5, osimertinib 5 mg kg⁻¹; osi 25, osimertinib 25 mg kg⁻¹. **c**, Heat map with unsupervised hierarchical clustering of log(mutant/WT) ratios from Ba/F3 cells expressing indicated mutations after drug treatment. Squares represent the median of $n = 3$ replicates. Mutation order was assigned arbitrarily; groups were assigned on the basis of predicted mutational impact. **d**, Dot plot of mutant/WT IC₅₀ values of Ba/F3 cells expressing classical EGFR mutations (white bars) with or without PACC mutations (coloured bars). In **a**, **d**, *P* values were determined by one-way analysis of variance (ANOVA) with unequal s.d. and Holm–Sidak’s multiple comparisons test. **e**, Average mutant/WT ratio of Ba/F3 cells expressing classical EGFR mutations (white

bars), and classical EGFR mutations plus C797S (shaded bar), T790M (hashed bars) or T790M and C797S (shaded and hashed bars). *P* values were determined by one-way ANOVA with repeated measures and post hoc Fishers' multiple comparisons test. In **a, f, h**, bars show mean \pm s.e.m. of mutant/WT ratio for all mutations and drugs; dots show representative average mutant/WT of $n = 3$ replicates.

Source data.

Similarly, acquired PACC mutations co-occurring with primary classical *EGFR* mutations retained sensitivity to second-generation TKIs while acquiring resistance to third-generation TKIs in an allele-specific manner (Fig. [3c, d](#)). In silico analysis of acquired PACC mutation, G796S, co-occurring with Ex19del was predicted to confer resistance to third-generation TKIs such as osimertinib by shifting the hinge region of the receptor, preventing stabilization of osimertinib at M793 and displacing the acrylamide group of osimertinib away from C797 (Extended Data Fig. [8h](#)). However, second-generation TKIs were less affected by shifts in the hinge region of the receptor and were predicted to maintain the orientation of the acrylamide group (Extended Data Fig. [8f](#)). Previous studies have reported that C797S mutations confer resistance to third-generation TKIs even without the presence of T790M³⁰ (Fig. [3e](#)). Similarly, C797S mutations without T790M conferred resistance to Ex20ins-active inhibitors, but not first- or second-generation TKIs unless T790M was present (Fig. [3e](#)). Retrospectively, we identified three patients with NSCLC containing *EGFR* L858R mutations that received first-line osimertinib and subsequently developed an *EGFR*-dependent mechanism of resistance. In all patients, a PACC mutation was identified upon biopsy at progression (Extended Data Fig. [9a–c](#)). Two patients acquired a L718V mutation, and one acquired two PACC mutations (V765L and C797S). All patients were treated with a second-generation TKI and experienced clinical benefits of stable disease and tumour shrinkage (Extended Data Fig. [9a–c](#)). These data demonstrate that both primary and acquired PACC mutations are sensitive to second-generation TKIs, and structure–function-based groupings could identify a novel class of mutations, PACC mutations, for which second-generation TKIs had higher selectivity and efficacy than third-generation drugs.

Structure-based groups predict outcomes

To determine whether structure–function-based groups could identify patients who are most likely to benefit from a treatment better than exon-based groups, we used a publicly available database of outcomes for patients harbouring atypical EGFR mutations treated with afatinib³² and determined overall response rate (ORR) and duration of treatment (DOT). Structure–function-based grouping showed clear differences between sensitive (classical-like and PACC) and resistant (T790M-like and Ex20ins-L) subgroups (ORR of 63% versus 20%), whereas exon-based groups had

less variation between groups (Extended Data Fig. 10a, b). Structure–function-based groups identified that patients with PACC mutations ($n = 156$) had a significantly longer DOT for afatinib than other structure-based groups (DOT: 17.1 months, $P < 0.0001$) (Fig. 4a, b). Using exon-based groups, we also identified that patients with exon 18 mutations ($n = 87$) had a longer DOT than patients with mutations in exons 19–21 (DOT: 17.4 months, $P < 0.0001$) (Fig 4b, Extended Data Fig. 10c); however, the structure-based approach identified nearly twice as many individuals who benefited from afatinib treatment.

Fig. 4: Structure–function groups better predict patient outcomes than exon-based groups.

 figure4

a, Kaplan–Meier plot of DOT of patients with NSCLC tumours containing atypical *EGFR* mutations ($n = 358$ patients) stratified by structure-based groups treated with afatinib. **b**, Forest plot of HRs calculated from Kaplan–Meier plots in **a** and Extended Data Fig. 10c. In **a**, **b**, classical-like, $n = 58$; T790M-like, $n = 68$; Ex20ins-L, $n = 76$; PACC, $n = 156$; exon 18, $n = 87$; exon 19, $n = 19$; exon 20, $n = 195$; exon 21, $n = 63$. **c**,

Kaplan–Meier plot of TTF of patients with PACC mutations treated with first-, second- or third-generation EGFR TKIs. **d**, Forest plot of HRs calculated from Kaplan–Meier plots in **c** and Extended Data Fig. [10d–h](#). In **c, d**, PACC, $n = 53$; non-PACC, $n = 56$; exon 18, $n = 40$; exon 19, $n = 19$; exon 20, $n = 24$; exon 21, $n = 26$. **e**, Kaplan–Meier plot of TTF of patients with PACC mutations ($n = 25$) or non-PACC mutations ($n = 13$) treated with second-generation TKIs. In **a–e**, HRs and P values were calculated using two-sided Mantel–Cox log-rank tests. In **b, d**, data are representative of $\text{HR} \pm 95\%$ confidence interval.

[Source data](#).

To determine whether structure-based groups could identify which class of inhibitors would provide the most benefit to patients with atypical EGFR mutations, we performed retrospective analyses of TTF of patients with atypical *EGFR* mutations treated with EGFR TKIs in MD Anderson Cancer Center GEMINI and Moffitt Cancer Center databases, and TTF was determined for the first EGFR TKI for which patients were treated. Most patients (80%) were stage IV at diagnosis, and there were no statistical differences in patient characteristics (Supplementary Tables [5, 6](#)). When stratified by structure–function-based groups, patients with PACC mutations treated with second-generation TKIs had a significantly longer TTF than patients treated with either first- or third-generation TKIs (21.7 months versus 10.0 months and 4.1 months, respectively; $P < 0.0001$, $\text{HR} = 0.23$) (Fig. [4c, d](#)). By contrast, TTF was not significantly different between classes of EGFR TKIs in patients with non-PACC mutations (Fig. [4d](#), Extended Data Fig. [10d](#)). Further, patients with PACC mutations had a longer TTF than patients with non-PACC mutations when treated with second-generation TKIs (21.7 months versus 10.0 months, respectively; $\text{HR} = 2.6$, $P = 0.0068$) (Fig. [4e](#)). When patients were stratified by exon and TTF was calculated for first-, second- and third-generation TKIs, significant differences were observed only in patients with exon 18 mutations treated with second-generation TKIs compared with third-generation TKIs (20.9 months versus 5.5 months; $P = 0.001$, $\text{HR} = 0.29$) (Fig [4d](#), Extended Data Fig. [10e–h](#)). Therefore, structure–function classification identified not only a larger subgroup of patients, but also a subgroup with greater benefit from second-generation TKIs than the exon-based classification.

Discussion

The diversity and higher than previously appreciated prevalence of atypical *EGFR* mutations shown here highlights the necessity of comprehensive next-generation sequencing (NGS) for patients with NSCLC. We show that *EGFR* mutations, including atypical mutations, can be divided into four distinct subgroups based on structure and function (Fig. [5](#)), and that structure–function-based groups can predict drug sensitivity and patient outcomes better than exon-based groups. While previous

studies have shown activity of second-generation TKIs in patients with specific exon 18 mutations^{[33,34](#)}, structure–function-based grouping identified a larger subgroup of *EGFR* mutations, PACC mutants, for which second-generation TKIs were more selective than third-generation TKIs. Clinically, second-generation TKIs have been associated with inhibition of wild-type EGFR and related adverse events^{[15,35,36](#)}; however, most second-generation TKIs are dosed at the maximum tolerated doses, resulting in plasma concentrations 10–100 fold greater than concentrations necessary for inhibiting PACC mutations. Unlike osimertinib, second-generation TKIs have limited activity in the central nervous system, demonstrating the need for novel TKIs with reduced inhibition of wild-type EGFR inhibition and CNS activity that can inhibit PACC mutants.

Fig. 5: EGFR mutations can be divided into four distinct subgroups.

 figure5

Representative space-filling models of each EGFR subgroup demonstrate changes in overall shape of drug-binding pocket. The P-loop (blue), hinge region (ATP-binding site (orange)), hydrophobic cleft (green), and α C-helix (yellow) are shown. Red dots represent location of mutations. Arrows indicate location of structural changes

compared with wild-type EGFR. The most common mutations are shown for each group, and drug sensitivity or selectivity is colour-coded and listed from most selective or sensitive (green) to resistant (red). PKCi, PKC inhibitor; ALKi, ALK inhibitor.

These findings demonstrate that structure–function-based groups can identify classes of drugs that may be effective for whole groups of mutations, reflecting the observation that mutations in different regions of the gene may induce similar changes in protein structure. For example, L718Q, S768I, and T854I correspond to exons 18, 20, and 21, respectively, but are all PACC mutations with similar structural effects on drug binding. Conversely, mutations within the same exon may induce quite disparate changes. L747_K754del-insATSPE, L747P and E746-A750del mutations occur in exon 19 but are T790M-like, PACC, and classical mutations, respectively, with distinct drug sensitivity and structural effects. A clinical challenge for physicians treating patients with *EGFR*-mutant cancers is to appropriately identify and match patient mutations with the best EGFR TKI. While a separate classification for each individual mutation could provide more precision than the groups described here, validating the clinical activities of different drugs for each mutation is not feasible. The classification presented here provides a framework through which clinicians, informed by internet-based tools or companies providing NGS reports, could more effectively personalize EGFR TKI therapy. Finally, these findings support the notion that for cancers containing oncogenes with diverse mutations, adopting a structure–function-based approach may improve clinical trial design and drug development.

Methods

No sample size calculations were done to predetermine group sizes, and investigators were not blinded during randomization and outcome assessments.

Analysis of EGFR variants in MD Anderson Cancer Center GEMINI, Foundation Medicine, Guardant Health and cBioPortal databases

To analyse the numbers and frequencies of different *EGFR* mutations among patients with NSCLC in the MD Anderson Cancer Center GEMINI database, the database was queried for patients with *EGFR* mutations ($n = 1,054$) and manually curated as classical or atypical *EGFR* mutations. The MD Anderson Cancer Center GEMINI database is prospectively collected from patients consented and enrolled on protocol number PA13-0589 in accordance with the MD Anderson Institutional Review Board.

EGFR mutations were determined from formalin-fixed paraffin-embedded tumours or digital-droplet PCR for blood samples by CLIA-certified methods as previously described^{18,37}. In brief, samples from MD Anderson Cancer Center were collected through molecular pathology and mutations were determined by next-generation

sequencing panels of tumour tissue DNA (MD Anderson Cancer Center Molecular Diagnostics Laboratory). MD Anderson Molecular Diagnostics Laboratory is a tissue molecular profiling method for NGS-based analysis to detect mutations in hotspot regions of 50 genes, and in April 2016, it was expanded to analyse 134 unique genes for the detection of somatic mutations in coding sequences of 128 genes and selected copy number variations (amplifications) in 49 genes. Moffitt Cancer Center used diagnostic methods such as Clarinet (bi-directional sequencing of exons 18–21 of *EGFR*), Pyrosequencing of *EGFR* gene (exons 18–21), and Moffitt Illumina TruSight Tumor 26 (TST26). Moffitt Trusight is a NGS Illumina sequencing platform with a panel of 170 genes. Commercial NGS platforms including FoundationOne and Guardant360 were used by both MD Anderson and Moffitt Cancer Center as described below.

To identify patients with *EGFR* mutations in the Foundation Medicine database, patient samples taken between November 2011 and May 2020 previously subjected to hybrid-capture based comprehensive genomic profiling using formalin-fixed paraffin-embedded tissue or plasma using previously validated assays^{38,39}, were analysed for *EGFR* mutations ($n = 10,221$). Patients were stratified by *EGFR* mutation, and *EGFR* mutations were manually curated as atypical or classical *EGFR* mutations. Classical *EGFR* mutations were defined as L858R point mutations, T790M mutations, and various exon 19 deletions including any deletion in exon 19 beginning at amino acid E746 or L747 and ending at amino acid A755. Deletions also including insertions were allowed and still considered classical exon 19 deletions. Atypical *EGFR* mutations were defined as non-synonymous mutations that were not defined as classical mutations. Patients with *EGFR* mutations where the sequence of the mutation was unknown were excluded from the analysis.

To determine the frequency of individual *EGFR* variants reported across the MD Anderson GEMINI database, cBioPortal, Foundation Medicine and the Guardant Health database, each database was analysed separately, and the average of all databases was determined. To determine the frequency of atypical mutations in the MD Anderson GEMINI and Foundation Medicine databases, atypical mutations were identified as described above and total number of known EGFR mutations across all patients was tabulated. For the analysis of cBioPortal, all non-overlapping studies were selected and exported. For overlapping studies, only the largest dataset was used, and all known EGFR mutations were tabulated. To determine the frequencies of *EGFR* variants from Guardant Health, a database of sequenced circulating free DNA (cfDNA), the Guardant360 clinical database was searched for NSCLC samples tested between November 2016 and November 2019 harbouring *EGFR* mutations ($n = 5,026$ patients). Guardant360 is a CLIA-certified, CAP/NYSDOH accredited comprehensive cfDNA NGS test that reports on SNVs, indels, fusions and SNVs in up to 73 genes. The Guardant360 clinical database, and the four datasets reported here, are enriched in

North American patients with NSCLC; the frequency of atypical EGFR mutations may differ in Asia or other regions.

Analysis of TTF in MD Anderson Cancer Center GEMINI and Moffitt Cancer Center

To determine TTF after EGFR TKI treatment, patients with NSCLC harbouring an EGFR mutation in the tyrosine kinase domain (exons 18–22) were identified in the MD Anderson GEMINI and Moffitt Cancer Center databases. Data collection for Moffitt Cancer Center (MCC) patients was performed under the protocol (MCC 19161), which was formally reviewed and granted approval by MCC in accordance with the Declaration of Helsinki and the 21st Century Cures Act. Outcomes were recorded for patients for only first EGFR TKI. Patients were stratified by classical (L858R or Ex19del, as defined above) or atypical (non-classical). There were 333 patients with NSCLC identified in the MD Anderson GEMINI database who had tumours expressing atypical mutations. Of these patients, 88 patients received at least one line of EGFR TKI treatment. In addition, at Moffitt Cancer Center, there were 21 patients with NSCLC with tumours harbouring atypical EGFR mutations. Clinical parameters were extracted from the respective databases. Patients previously receiving chemotherapy were included, and TTF was calculated for the first EGFR TKI received. TTF was determined as previously described¹⁸ and defined as time from commencement of EGFR TKI to radiologic progression, TKI discontinuation, or death, and was not based on RECIST criteria. For patients treated beyond progression, radiologic progression was recorded as the end point, and data cut-off was May 2021. Median TTF was calculated using the Kaplan–Meier method. HR and *P* values were determined using GraphPad Prism software and two-sided Mantel–Cox log-rank tests.

Analysis of OS and PFI from cBioPortal Database

For overall survival (OS) and progression-free interval (PFI), analysis of patients in cBioportal was determined as previously described¹⁹ for patients receiving any treatment with survival information and qualifying EGFR mutation. This information was curated from cBioportal by selecting all non-overlapping studies of NSCLC. For overlapping studies, the largest database was selected. PFI and OS analysis were restricted to the tyrosine kinase domain. Median OS and median PFI were calculated using the Kaplan–Meier method. HR and *P* values were determined using GraphPad Prism software and two-sided Mantel–Cox log-rank tests.

Ba/F3 cell generation, drug screening and IC₅₀ approximations

Ba/F3 cells were obtained as a gift from G. Mills (MD Anderson Cancer Center) and maintained in RPMI (Sigma) containing 10% FBS, 1% penicillin-streptomycin and 10 ng ml⁻¹ recombinant mIL-3 (R&D Biosystems). To establish stable Ba/F3 cell lines, Ba/F3 cells were transduced with retroviruses containing mutant EGFR plasmids for 12–24 h. Retroviruses were generated using Lipofectamine 2000 (Invitrogen) transfections of Phoenix 293T-ampho cells (Orbigen) with pBabe-Puro based vectors listed in Supplementary Table 7. Vectors were generated by GeneScript or Bioinnovatise using parental vectors from Addgene listed in Supplementary Table 7. After 48–72 h of transduction, 2 µg ml⁻¹ puromycin (Invitrogen) was added to Ba/F3 cell lines in complete RPMI. To select for EGFR-positive cell lines, cells were stained with PE-EGFR (Biolegend) and sorted by fluorescence-activated cell sorting. After sorting, EGFR-positive cells were maintained in RPMI containing 10% FBS, 1% penicillin-streptomycin, and 1 ng ml⁻¹ EGF to support cell viability. Drug screening was performed as previously described^{22,36}. Shortly, cells were plated in 384-well plates (Greiner Bio-One) at 2,000–3,000 cells per well in technical triplicate. Seven different concentrations of TKIs or DMSO vehicle were added to reach a final volume of 40 µl per well. After 72 h, 11 µl of Cell Titer Glo (Promega) was added to each well. Plates were incubated for a minimum of 10 min, and bioluminescence was determined using a FLUOstar OPTIMA plate reader (BMG LABTECH). Raw bioluminescence values were normalized to DMSO control-treated cells, and values were plotted in GraphPad Prism. Non-linear regressions were used to fit the normalized data with a variable slope, and IC₅₀ values were determined by GraphPad prism by interpolation of concentrations at 50% inhibition. Drug screens were performed in technical triplicate on each plate and either duplicate or triplicate biological replicates. Mutant to WT ratios for each drug were calculated by dividing the IC₅₀ values of mutant cell lines by the average IC₅₀ value of Ba/F3 cells expressing WT EGFR supplemented with 10 ng ml⁻¹ EGF for each drug. Statistical differences between groups were determined by one-way ANOVA as described in the figure legends.

In silico mutational mapping and docking experiments

X-ray structures of wild type EGFR in complex with AMP-PNP (2ITX) and osimertinib (4ZAU), and EGFR L858R mutant in complex with AMP-PNP (Protein Data Bank (PDB) ID: 2ITV) were retrieved from the Protein Data Bank. Molecular Operating Environment (2019.01; Chemical Computing Group CCCG) was used to generate mutant homology models, construct protein–ligand models and for visualization. Pymol was used for visualization of mutation location on WT EGFR (PDB ID: 2ITX) and structural alignment with EGFR D770insNPG (PDB ID: 4LRM) or EGFR G719S (PDB ID: 2ITN).

Heat map generation

Heat maps and hierarchical clustering were generated by plotting the median log (Mut/WT) value for each cell line and each drug using R and the ComplexHeatmap package⁴⁰ 2.6.2 (R Foundation for Statistical Computing). Hierarchical clustering was determined by Euclidean distance between Mut/WT ratios. For co-occurring mutations, mutation order was assigned arbitrarily, and for acquired mutations, mutations were assigned in the order mutations are observed clinically. Structure–function groups were assigned based on predicted impact of mutation on receptor conformation.

Determination of EGFR groups and subgroups

Mutational mapping was used to separate EGFR mutations into distinct groups based on predicted drug sensitivity. Structural features of EGFR mutations with known drug sensitivity (that is, classical EGFR mutations^{41,42}, T790M^{43,44,45} and exon 20 insertions^{22,25}) were used as the basis for predicting the impact of mutations on drug sensitivity. Using mutational mapping there were four distinct groups: (1) no obvious effect on the drug binding pocket (similar to L858R); (2) a mutation in the hydrophobic core (similar to T790M); (3) a large inward shift of both the α C-helix and P-loop (similar to exon 20 insertions); and (4) a slight inward shift of the α C-helix and/or P-loop due to direct changes to the either the α C-helix and/or P-loop or indirectly through alterations of the β -pleated sheets that are predicted to effect the position either the α C-helix and/or P-loop. Groups were validated by hierarchical clustering of in vitro sensitivity of Ba/F3 cells expressing the various EGFR mutations. Subgroups such as T790M-like-3S/T790M-like-3R and Ex20ins-NL/Ex20ins-FL were defined based on cell line sensitivity data.

Statistical analyses of structure-function groups

Correlations for mutations were determined using Spearman’s rho by correlating the median log (Mut/WT) value for each mutation and drug versus the average of the median log (Mut/WT) value for the structure–function-based group or exon-based group for which the mutation belongs. For each correlation, the mutation tested was removed from the average structure function and exon-based groups. Average rho values were compared by two-sided Student’s *t*-test. To determine whether structure function groups or exon groups were better predictor of drug sensitivity, we performed recursive-partitioning analyses to construct a decision tree for each drug using structure function group and mutation data on exons 18, 19, 20, and 21 as predictors. The decision tree classified samples by posing a series of decision rules based on predictors. Each decision rule was constrained in an internal node, and every internal node points to yes-or-no questions that result in a ‘yes’ or ‘no’ branch. We applied the

CART algorithm^{20,21} using the rpart R package. We calculated variable importance as the sum of the goodness of split measures for each split. These were scaled to sum to 100 for a tree. Median SAS version 9.4 and R version 3.5.6 were used to carry out the computations for all analyses. The structure function group variable was involved in the first and second splits in all of the 18 regression trees of drug sensitivity. The variable importance of this variable was in a range of 66–94%. Both the order of the split and variable importance indicate that the structure function group variable was more predictive than the exon-based variables in evaluation of drug sensitivity. Code for this analysis can be found at <https://github.com/MD-Anderson-Bioinformatics/EGFR-Structure-Function-Nature-Manuscript>.

PDX generation and in vivo experiments

As part of the MD Anderson Cancer Center Lung Cancer Moon Shots program, PDXs harbouring EGFR G719A and EGFR L858R/E709K were generated and maintained in accordance with Good Animal Practices and with approval from MD Anderson Cancer Center Institutional Animal Care and Use Committee on protocol number PA140276 as previously described⁴⁶. Surgical samples were rinsed with serum-free RPMI supplemented with 1% penicillin-streptomycin then implanted into the right flank of 5- to 6-week-old NSG female mice within 2 h of resection. Tumours were validated for EGFR mutations by DNA fingerprinting and quantitative PCR as described⁴⁶. PDXs harbouring EGFR S768dupSVD were purchased from Jackson Laboratories (J100672). To propagate tumours, 5- to 6-week-old female NSG mice (NOD.Cg-Prkdcscid IL2rgtmWjl/Szj) were purchased from Jackson Laboratories (005557). Fragments of NSCLC tumours expressing EGFR S768dupSVD, G719A or L858R/E709K were implanted into 6- to 8-week-old female NSG mice. Once tumours reached 2,000 mm³, they were collected and re-implanted into the right flank of 6- to 8-week-old female NSG mice. Tumours were measured 3 times per week and were randomized into treatment groups when tumors reached a volume of 275–325 mm³ for the EGFR G719A and S768dupSVD models, and 150–175 mm³ for the L858R/E709K model. Treatment groups included vehicle control (0.5% methylcellulose, 0.05% Tween-80 in dH₂O), 100 mg kg⁻¹ erlotinib, 20 mg kg⁻¹ afatinib, 2.5 mg kg⁻¹ poziotinib, 5 mg kg⁻¹ osimertinib, and 25 mg kg⁻¹ osimertinib. During treatment, body weight and tumour volumes were measured three times per week, and mice received treatment five days per week (Monday to Friday). Dosing holidays were given if mouse body weight decreased by more than 10% or overall body weight dropped below 20 g. Maximum allowed tumour burden by approved IACUC protocol was a volume of 2,000 mm³. Mice were humanely euthanized if tumour sizes exceeded the maximum size.

Case studies of patients treated with second-generation TKIs

Patients were consented under the GEMINI protocol (PA13-0589) which was approved in accordance with the MD Anderson Institutional Review Board, or protocol MCC 19161, which was formally reviewed and granted approval by Moffitt Cancer Center in accordance with the Declaration of Helsinki and the 21st Century Cures Act for retrospective analysis of patient outcomes and treatment course for case studies of patients presented. Both protocols include informed consent for publication of deidentified data.

Retrospective analysis of ORR and duration of treatment with afatinib

Response to afatinib and duration of afatinib treatment was tabulated from 803 patients in the Uncommon EGFR Database (www.uncommonegfrmutations.com). Objective response rate was reported in 529 patients. Patients were stratified by either structure–function-based groups or exon-based groups and ORR was determined by counting the number of patients reported to have complete response or partial response. Fisher’s exact test was used to determine statistical differences between subgroups (structure based or exon-based). Duration of treatment was provided in the Uncommon EGFR Database for 746 patients. Patients were stratified by structure–function-based groups and exon-based groups and median DOT was calculated using the Kaplan–Meier method. Statistical differences in Kaplan–Meier plots, HR and *P* values were generated using GraphPad Prism software and the Mantel–Cox log-rank method. When mutations were not explicitly stated (that is, exon 19 mutation) those patients were excluded from the structure–function-based analysis but included in the exon-based analysis.

Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

Data availability

Source data for all figures can be found at <https://github.com/MD-Anderson-Bioinformatics/EGFR-Structure-Function-Nature-Manuscript>. Public datasets used in this study include non-overlapping studies including NSCLC in cBioportal (www.cbioportal.org) and the Uncommon EGFR Database (www.uncommonegfrmutations.com). Details of specific studies accessed can be found in the Reporting Summary. Data from Foundation Medicine and Guardant Health were provided under data use agreements; however, summarized data used in Fig. 1 and Extended Data Fig. 1 are provided at <https://github.com/MD-Anderson>.

[Bioinformatics/EGFR-Structure-Function-Nature-Manuscript](#). [Source data](#) are provided with this paper.

Code availability

All code used in this study can be found at <https://github.com/MD-Anderson-Bioinformatics/EGFR-Structure-Function-Nature-Manuscript>.

References

1. 1.

Russo, A. et al. Heterogeneous responses to epidermal growth factor receptor (EGFR) tyrosine kinase inhibitors (TKIs) in patients with uncommon EGFR mutations: new insights and future perspectives in this complex clinical scenario. *Int. J. Mol. Sci.* **20**, 1431 (2019).

2. 2.

Kobayashi, Y. et al. EGFR exon 18 mutations in lung cancer: molecular predictors of augmented sensitivity to afatinib or neratinib as compared with first- or third-generation TKIs. *Clin. Cancer Res.* **21**, 5305–5313 (2015).

3. 3.

Klughammer, B. et al. Examining treatment outcomes with erlotinib in patients with advanced non-small cell lung cancer whose tumors harbor uncommon EGFR mutations. *J. Thorac. Oncol.* **11**, 545–555 (2016).

4. 4.

Rosell, R. et al. Erlotinib versus standard chemotherapy as first-line treatment for European patients with advanced EGFR mutation-positive non-small-cell lung cancer (EURTAC): a multicentre, open-label, randomised phase 3 trial. *Lancet Oncol.* **13**, 239–246 (2012).

5. 5.

Sequist, L. V. et al. Phase III study of afatinib or cisplatin plus pemetrexed in patients with metastatic lung adenocarcinoma with EGFR mutations. *J. Clin. Oncol.* **31**, 3327–3334, <https://doi.org/10.1200/JCO.2012.44.2806> (2013).

6. 6.

Soria, J. C. et al. Osimertinib in untreated EGFR-mutated advanced non-small-cell lung cancer. *N. Engl. J. Med.* **378**, 113–125 (2018).

7. 7.

Yoshikawa, S. et al. Structural basis for the altered drug sensitivities of non-small cell lung cancer-associated mutants of human epidermal growth factor receptor. *Oncogene* **32**, 27–38 (2013).

8. 8.

Massarelli, E., Johnson, F. M., Erickson, H. S., Wistuba, I. I. & Papadimitrakopoulou, V. Uncommon epidermal growth factor receptor mutations in non-small cell lung cancer and their mechanisms of EGFR tyrosine kinase inhibitors sensitivity and resistance. *Lung Cancer* **80**, 235–241 (2013).

9. 9.

Yasuda, H., Kobayashi, S. & Costa, D. B. EGFR exon 20 insertion mutations in non-small-cell lung cancer: preclinical data and clinical implications. *Lancet Oncol.* **13**, e23–e31 (2012).

10. 10.

Kosaka, T. et al. Response heterogeneity of EGFR and HER2 exon 20 insertions to covalent EGFR and HER2 inhibitors. *Cancer Res.* **77**, 2712–2721 (2017).

11. 11.

Ramalingam, S. S. et al. Overall survival with osimertinib in untreated, EGFR-mutated advanced NSCLC. *N. Engl. J. Med.* **382**, 41–50 (2020).

12. 12.

Cho, J. H. et al. Osimertinib for patients with non-small-cell lung cancer harboring uncommon EGFR mutations: a multicenter, open-label, phase II trial (KCSG-LU15-09). *J. Clin. Oncol.* **38**, 488–495 (2020).

13. 13.

Wu, Y. L. et al. Afatinib versus cisplatin plus gemcitabine for first-line treatment of Asian patients with advanced non-small-cell lung cancer harbouring EGFR mutations (LUX-Lung 6): an open-label, randomised phase 3 trial. *Lancet Oncol.* **15**, 213–222 (2014).

14. 14.
- Yang, J. C. et al. Clinical activity of afatinib in patients with advanced non-small-cell lung cancer harbouring uncommon EGFR mutations: a combined post-hoc analysis of LUX-Lung 2, LUX-Lung 3, and LUX-Lung 6. *Lancet Oncol.* **16**, 830–838 (2015).
15. 15.
- Yang, J. C. et al. Afatinib versus cisplatin-based chemotherapy for EGFR mutation-positive lung adenocarcinoma (LUX-Lung 3 and LUX-Lung 6): analysis of overall survival data from two randomised, phase 3 trials. *Lancet Oncol.* **16**, 141–151 (2015).
16. 16.
- Administration, U. S. F. a. D. (ed FDA) (2021).
17. 17.
- Shi, J. et al. Uncommon EGFR mutations in a cohort of Chinese NSCLC patients and outcomes of first-line EGFR-TKIs and platinum-based chemotherapy. *Chin. J. Cancer Res.* **29**, 543–552 (2017).
18. 18.
- Le, X. et al. Landscape of EGFR -dependent and -independent resistance mechanisms to osimertinib and continuation therapy post-progression in EGFR-mutant NSCLC. *Clin. Cancer Res.* **24**, 6195–6203 (2018).
19. 19.
- Liu, J. et al. An integrated TCGA pan-cancer clinical data resource to drive high-quality survival outcome analytics. *Cell* **173**, 400–416.e411 (2018).
20. 20.
- Strobl, C., Malley, J. & Tutz, G. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychol. Methods* **14**, 323–348 (2009).
21. 21.

Breiman, L. *Classification and Regression Trees* (Wadsworth International Group, 1984).

22. 22.

Robichaux, J. P. et al. Mechanisms and clinical activity of an EGFR and HER2 exon 20-selective kinase inhibitor in non-small cell lung cancer. *Nat. Med.* **24**, 638–646 (2018).

23. 23.

Vasconcelos, P., Kobayashi, I. S., Kobayashi, S. S. & Costa, D. B. Preclinical characterization of mobocertinib highlights the putative therapeutic window of this novel EGFR inhibitor to EGFR exon 20 insertion mutations. *JTO Clin. Res. Rep.* **2**, 100105 (2021).

24. 24.

Vasconcelos, P. et al. EGFR-A763_Y764insFQEA is a unique exon 20 insertion mutation that displays sensitivity to approved and in-development lung cancer EGFR tyrosine kinase inhibitors. *JTO Clin. Res. Rep.* **1**, 100051 (2020).

25. 25.

Yasuda, H. et al. Structural, biochemical, and clinical characterization of epidermal growth factor receptor (EGFR) exon 20 insertion mutations in lung cancer. *Sci. Transl. Med.* **5**, 216ra177 (2013).

26. 26.

Yang, M. et al. NSCLC harboring EGFR exon-20 insertions after the regulatory C-helix of kinase domain responds poorly to known EGFR inhibitors. *Int. J. Cancer* **139**, 171–176 (2016).

27. 27.

Lee, H. J. et al. Noncovalent wild-type-sparing inhibitors of EGFR T790M. *Cancer Discov.* **3**, 168–181 (2013).

28. 28.

Wang, X. et al. Lung adenocarcinoma harboring EGFR 19del/C797S/T790M triple mutations responds to brigatinib and anti-EGFR antibody combination therapy. *J. Thorac. Oncol.* **14**, e85–e88 (2019).

29. 29.

Uchibori, K. et al. Brigatinib combined with anti-EGFR antibody overcomes osimertinib resistance in EGFR-mutated non-small-cell lung cancer. *Nat. Commun.* **8**, 14768 (2017).

30. 30.

Thress, K. S. et al. Acquired EGFR C797S mutation mediates resistance to AZD9291 in non-small cell lung cancer harboring EGFR T790M. *Nat. Med.* **21**, 560–562 (2015).

31. 31.

Ou, S. I. et al. Emergence of novel and dominant acquired EGFR solvent-front mutations at Gly796 (G796S/R) together with C797S/R and L792F/H mutations in one EGFR (L858R/T790M) NSCLC patient who progressed on osimertinib. *Lung Cancer* **108**, 228–231 (2017).

32. 32.

Yang, J. C. et al. Afatinib for the treatment of NSCLC harboring uncommon EGFR mutations: a database of 693 cases. *J. Thorac. Oncol.* **15**, 803–815 (2020).

33. 33.

Tamiya, M. et al. Successful treatment of an osimertinib-resistant lung adenocarcinoma with an exon 18 EGFR mutation (G719S) with afatinib plus bevacizumab. *Invest. New Drugs* **39**, 232–236 (2021).

34. 34.

Ma, C. et al. Effective treatment with afatinib of lung adenocarcinoma with leptomeningeal metastasis harboring the exon 18 p.G719A mutation in the EGFR gene was detected in cerebrospinal fluid: a case report. *Front. Oncol.* **10**, 1635 (2020).

35. 35.

Mok, T. S. et al. Improvement in overall survival in a randomized study that compared dacomitinib with gefitinib in patients with advanced non-small-cell lung cancer and EGFR-activating mutations. *J. Clin. Oncol.* **36**, 2244–2250 (2018).

36. 36.

Robichaux, J. P. et al. Pan-cancer landscape and analysis of ERBB2 mutations identifies pozotinib as a clinically active inhibitor and enhancer of T-DM1 activity. *Cancer Cell* **36**, 444–457.e447 (2019).

37. 37.

Liu, X. et al. Concurrent use of aspirin with osimertinib is associated with improved survival in advanced EGFR-mutant non-small cell lung cancer. *Lung Cancer* **149**, 33–40 (2020).

38. 38.

Frampton, G. M. et al. Development and validation of a clinical cancer genomic profiling test based on massively parallel DNA sequencing. *Nat. Biotechnol.* **31**, 1023–1031 (2013).

39. 39.

Clark, T. A. et al. Analytical validation of a hybrid capture-based next-generation sequencing clinical assay for genomic profiling of cell-free circulating tumor DNA. *J. Mol. Diagn.* **20**, 686–702 (2018).

40. 40.

Galili, T., O'Callaghan, A., Sidi, J. & Sievert, C. heatmaply: an R package for creating interactive cluster heatmaps for online publishing. *Bioinformatics* **34**, 1600–1602 (2018).

41. 41.

Tamirat, M. Z., Koivu, M., Elenius, K. & Johnson, M. S. Structural characterization of EGFR exon 19 deletion mutation using molecular dynamics simulation. *PLoS ONE* **14**, e0222814 (2019).

42. 42.

Tsigelny, I. F. et al. Molecular determinants of drug-specific sensitivity for epidermal growth factor receptor (EGFR) exon 19 and 20 mutants in non-small cell lung cancer. *Oncotarget* **6**, 6029–6039 (2015).

43. 43.

Yun, C. H. et al. The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP. *Proc. Natl Acad. Sci. USA* **105**, 2070–2075 (2008).

44. 44.

Finlay, M. R. et al. Discovery of a potent and selective EGFR inhibitor (AZD9291) of both sensitizing and T790M resistance mutations that spares the wild type form of the receptor. *J. Med. Chem.* **57**, 8249–8267 (2014).

45. 45.

Cross, D. A. et al. AZD9291, an irreversible EGFR TKI, overcomes T790M-mediated resistance to EGFR inhibitors in lung cancer. *Cancer Discov* **4**, 1046–1061 (2014).

46. 46.

Chen, Y. et al. Tumor characteristics associated with engraftment of patient-derived non-small cell lung cancer xenografts in immunocompromised mice. *Cancer* **125**, 3738–3748 (2019).

Acknowledgements

We thank the patients, families and caregivers involved in the GEMINI protocol and the patient advocacy groups who have an essential role in patient care and research, including the Exon 20 Group. We also thank I. G. Munoz for critical review and editorial assistance. This work was supported by the generous philanthropic contributions to The University of Texas MD Anderson Lung Cancer Moon Shots Program and the MD Anderson Cancer Center Support Grant P30 CA016672 including animal housing and care in the Research Animal Support Facility (RASF). Additional support was contributed by NIH-R01CA247975, CPRIT-IIRA RP200150, NIH/NCI R01CA234183, NIH/NCI R01CA190628, Lung SPORE P50 CA070907-20, the David Bruton Jr Endowment, the Rexanna Foundation for Fighting Lung Cancer, 1U54CA224065-01, research support from Spectrum Pharmaceuticals (J.H.), the Hallman Fund, the Stading Fund for EGFR inhibitor resistance, the Gil and Dody Weaver Foundation, the Hanlon Fund, the Richardson Fund for EGFR mutant lung cancer research, and ASCO CDA-57112 (Y.Y.E.).

Author information

Affiliations

1. Department of Thoracic/Head and Neck Medical Oncology, MD Anderson Cancer Center, Houston, TX, USA

Jacqulyne P. Robichaux, Xiuning Le, Simon Heeke, Yasir Y. Elamin, Hibiki Udagawa, Ferdinandos Skoulidis, Hai Tran, Susan Varghese, Junqin He, Fahao Zhang, Monique B. Nilsson, Lemei Hu, Alissa Poteete, Xiaoke Liu, Lingzhi Hong, Jianjun Zhang & John V. Heymach

2. Institute for Applied Cancer Science, MD Anderson Cancer Center, Houston, TX, USA

R. S. K. Vijayan & Jason B. Cross

3. Department of Individualized Cancer Management, Moffitt Cancer Center, Tampa, FL, USA

J. Kevin Hicks

4. Department of Biostatistics, MD Anderson Cancer Center, Houston, TX, USA

Heather Y. Lin & Min Jin Ha

5. Quantitative Research Computing, MD Anderson Cancer Center, Houston, TX, USA

Waree Rinsurongkawong

6. Department of Thoracic and Cardiovascular Surgery, MD Anderson Cancer Center, Houston, TX, USA

Xiaoshan Zhang, Chenghui Ren & Bingliang Fang

7. Department of Thoracic Oncology, West China Medical School, West China Hospital, Sichuan University, Sichuan, China

Xiaoke Liu

8. Department of Bioinformatics and Computational Biology, MD Anderson Cancer Center, Houston, TX, USA

Lixia Diao & Jing Wang

9. Foundation Medicine, Cambridge, MA, USA

Russell Madison & Alexa B. Schrock

10. Guardant Health, Redwood City, CA, USA

Jennifer Saam & Victoria Raymond

11. Department of Thoracic Oncology, Moffitt Cancer Center, Tampa, FL, USA

Jhanelle E. Gray

Contributions

The study was conceptualized by J.P.R., X. Le and J.V.H.; and experiments were completed by J.P.R., R.S.K.V., L. Hu, J.H., H.U., X.Z., C.R., F.Z. and A.P. Patient analyses including tabulation and analysis of patient outcomes was completed by X. Le, W.R., J.P.R., H.T., J.K.H., J.E.G., Y.Y.E., X. Liu, J.Z., F.S., S.V., R.M., A.B.S. and J.S. Statistical analyses were completed by S.H., J.P.R., L.D., M.J.H., H.Y.L., L. Hong, and J.W. Mutational mapping, crystal visualization and in silico modelling was completed by J.P.R., R.S.K.V. and J.B.C., and figures were prepared by J.P.R., S.H. and R.S.K.V. The original draft of the manuscript was completed by J.P.R., J.B.C. and J.V.H.; and editing was completed by J.P.R., R.S.K.V., J.B.C., X. Le, Y.Y.E., J.K.H., A.B.S., J.E.G. and M.B.N. Funding was provided by J.V.H., X. Le., J.P.R. and Y.Y.E. J.P.R., X. Le, J.B.C., J.E.G. and J.V.H. supervised the experiments and analysis of the manuscripts. All co-authors approved the final manuscript.

Corresponding author

Correspondence to John V. Heymach.

Ethics declarations

Competing interests

The research being reported in this publication is research in which The University of Texas MD Anderson Cancer Center has an institutional financial conflict of interest. Because MD Anderson is committed to the protection of human subjects and the effective management of its financial conflicts of interest in relation to its research activities, MD Anderson has implemented an Institutional Conflict of Interest Management and Monitoring Plan to manage and monitor the conflict of interest with respect to MD Anderson's conduct of this research. MD Anderson, including J.P.R., M.B.N. and J.V.H. have filed a patent for the use of poziotinib for treating EGFR- and HER2-mutant cancers and licensed the technology to Spectrum Pharmaceuticals.

J.P.R. and J.V.H. receive research support from Spectrum Pharmaceuticals, Takeda and Enliven Therapeutics. MD Anderson, including M.B.N., J.P.R. and J.V.H. have a pending patent submitted for treatment of EGFR TKI resistant NSCLC, and another, including J.P.R., S.H. and J.V.H., for the classification of EGFR mutations. J.P.R. and M.B.N. have no non-financial competing interests. J.V.H. also receives grant or research support from AstraZeneca and GlaxoSmithKline and has served on advisory committees for AstraZeneca, Boehringer Ingelheim, Bristol-Myers Squibb, Catalyst, EMD Serono, Foundation Medicine (FMI), Hengrui Therapeutics, Genentech, GlaxoSmithKline, Guardant Health, Eli Lilly, Merck, Novartis, Pfizer, Roche, Sanofi, Seattle Genetics, Spectrum Pharmaceuticals, Takeda. As non-financial competing interests, J.V.H. serves as scientific advisor for Rexanna Foundation and the EGFR resisters. X. Le receives consulting/advisory fees from EMD Serono (Merck KGaA), AstraZeneca, Spectrum Pharmaceuticals, Eli Lilly, Boehringer Ingelheim, Bristol-Myers Squibb, Novartis, Hengrui Therapeutics, Daiichi Sankyo, and Celgene, and research funding to the institution from Eli Lilly and Boehringer Ingelheim, and no non-financial competing interests. Y.Y.E. discloses research support from Spectrum Pharmaceuticals, AstraZeneca, Takeda, Eli Lilly, Xcovery, Turning Point Therapeutics; advisory role for AstraZeneca, Eli Lilly and Turning Point Therapeutics; and accommodation expenses from Eli Lilly, and no non-financial competing interests. J.Z. reports research funding and consulting fees from Bristol-Myers Squibb, AstraZeneca, Geneplus, OrigMed, Innovent, Merck, Johnson and Johnson, and no non-financial competing interests. F.S. receives consulting/advisory fees from Amgen, Navire Pharma, Intellisphere LLC; research funding to the institution from Amgen, Mirati Therapeutics, Boehringer Ingelheim, Merck&Co, Novartis, Pfizer; speaker's fees from Bristol-Myers Squibb and RV Mais Promocao Eventos LTDS; other support from AstraZeneca Pharmaceuticals; travel expenses from Tango Therapeutics and reports past stock ownership of Moderna Inc and BioNTech SE. As non-financial competing interests F.S. serves as scientific advisor for Tango Therapeutics and has previously served as advisor for Calithera Biosciences. J.E.G. reports research support from AstraZeneca, Boehringer Ingelheim, Bristol-Myers Squibb, Genentech, G1 Therapeutics, Merck, Novartis, Pfizer, and the Lurie Institute of Cancer Research, and serves as a consultant/advisor for Daiichi Sankyo, Janssen, Novartis, Merck, Invata, EMD Serono, Bristol-Myers Squibb, Blueprint Medicines, and AstraZeneca. H.T. reports research support from Bayer AS, Guardant Health, Ziopharm, and no non-financial interests. J.S. and V.R. are shareholders/full time employees of Guardant Health and report no non-financial interests. R.M. and A.B.S. are full time employees of Foundation medicine, a wholly-owned subsidiary of Roche, and have Roche stock ownership, with no non-financial competing interests. J.K.H. receives research support from OneOme, is a consultant for Quest Diagnostics and 23andMe, and has served on an advisory committee for Novartis, and has provided educational content for the American College of Clinical Pharmacy, Florida Pharmacy Association and

HorizonCME. All other authors report no financial or non-financial competing interests.

Additional information

Peer review information *Nature* thanks Daniel Costa, Helena Yu 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 Patients with atypical EGFR mutations have worse clinical outcomes than those with classical EGFR mutations.](#)

a, Diagram of patient sample sources and types across databases. **b**, Lollipop plot of frequency of all EGFR mutations observed in patients with NSCLC (N=24,934 mutations). *EGFR* mutations associated with acquired drug resistance, as described by the literature, are highlighted in red. **c–f**, Kaplan-Meier plot of TTF of patients with NSCLC tumors harboring classical (N=264 patients) or atypical *EGFR* mutations stratified by (**c**) exon after treatment with an EGFR TKI (Exon 18 N= 40, Exon 19 N=19, Exon 20 N=15, Exon 21 N=26), or EGFR TKI class including 1st- (**d**), 2nd- (**e**), or 3rd- (**f**) gen TKIs in MD Anderson GEMINI and Moffitt Cancer center databases. Patients that received prior chemotherapy or immunotherapy were included, but TTF was calculated for first EGFR TKI received. **g, h**, Kaplan-Meier plot of PFI (**g**) and OS (**h**) of patients with NSCLC tumors harboring classical (N=50 for PFI and N=52 for OS) or atypical (N=35 for PFI and N=39 for OS) *EGFR* mutations from cBioPortal. Atypical EGFR mutations were limited to mutations in the tyrosine kinase domain, and treatment and stage were unknown. **c–h**, HRs and p-values were calculated using two-sided Mantel-Cox, Log-Rank tests

[Source data.](#)

[Extended Data Fig 2 Mutational mapping of EGFR mutations.](#)

a–d, Mutational mapping of classical-like (**a**), T790M-like (**b**), exon 20 loop insertion (red/blue) and WT (grey/green) (**c**), and PACC mutants (**d**) onto WT EGFR crystal (PDB 2ITX).

Extended Data Fig 3 Heat maps generated through supervised clustering by structure-function-based groups cluster drug sensitivity better than exon-based groups.

a, b, Heat maps supervised clustering by exon-based (**a**) or structure-function-based (**b**) groups of log (Mutant/WT) ratios from Ba/F3 cells expressing indicated mutations after 72 h of indicated drug treatment. To determine the mutant/WT ratio, IC₅₀ values for each drug and cell line were calculated and then compared to the average IC₅₀ values of Ba/F3 cells expressing WT EGFR (+10ng/ml EGF to maintain viability). Squares are representative of the median of n=3 replicates. For co-occurring mutations, the order of mutations 1, 2, and 3 were assigned arbitrarily. Groups were assigned based on structural predictions

[Source data.](#)

Extended Data Fig. 4 Structure-function-based groupings are more predictive of drug and mutation sensitivity compared to exon-based groupings.

a, Bar plot of Spearman rho values for indicated mutations compared to exon-based groups (yellow) or structure-function-based groups (green). The delta of the two rho values is shown as an overlapped grey bar. When the delta bar shifts to the right, the spearman rho value was higher for structure-function-based groups, and when the grey bar shifts to the left, the spearman rho value was higher for the exon-based groups. **b,** Representative classification and regression trees for each indicated drug. Colors represent drug sensitivity (green) or resistance (red) as defined by log (mutant IC₅₀/WT EGFR IC₅₀). **c,** Bar plot of Spearman rho values for indicated mutations (excluding T790M mutations) compared to exon-based groups (yellow) or structure-function-based groups (green). The delta of the two rho values is shown as an overlapped grey bar. **d,** Representative classification and regression trees for each indicated drug excluding T790M from the analysis. Colors represent drug sensitivity (green) or resistance (red) as defined by log (mutant IC₅₀/WT EGFR IC₅₀). **e,** Dot plot of rho values from Spearman correlations of mutations vs exon-based group averages or structure-function based averages for each drug excluding T790M mutations. Dots are representative of each mutation; bars are representative of the average rho value ± standard deviation (SD). p-value was determined using a paired two-sided t-test, and n = 59 cell lines/mutations. **f,** Dot plot of variable importance calculated as sum of the goodness of split for each split in the classification and regression trees (CART). Dots are representative of variable importance for each drug in the exon and structure-function-based groups as indicated and excluding T790M mutations. Bars are representative of the median ± 95% confidence interval of variable

importance for all drugs (Supplementary Table 3). p-value was determined using a paired two-sided *t*-test, and n = 18 drugs

[Source data](#).

Extended Data Fig. 5 Classical-like EGFR mutations are not predicted to alter the drug-binding pocket and are most sensitive to 3rd gen EGFR TKIs.

a, b, Rendering of crystal structure of WT EGFR (PDB 2ITX) visualized as both a ribbon (**a**) and space filling (**b**) models. Residues important in receptor signaling and drug binding are highlighted. **c, d,** Overlapped rendering of WT crystal (grey) (**c**) and L861R (blue) and space filing model (**d**) of L861Q demonstrate the R861 substitution is distal from the drug binding pocket and has minimal impact on the overall structure of EGFR compared to WT. **e,** Dot plot of mutant/WT IC₅₀ values of Ba/F3 cells expressing classical-like EGFR mutations and treated with indicated classes of EGFR TKIs. Dots are representative of average of n=3 replicate mutant/WT IC₅₀ values of individual cell lines expressing classical-like mutations with individual drugs. Bars are representative of average mutant/WT IC₅₀ values ± SEM for each class of EGFR TKI and all classical-like cell lines. p-values were determined by one-way ANOVA with unequal SD as determined by Brown-Forsythe test to determine differences in SD. Holm-Sidak's multiple comparisons test was used to determine differences between groups. **f,** Tumor growth curves for PDXs harboring EGFR L858R E709K complex mutation treated with indicated inhibitors. Tumors were measured three times per week and symbols are average of tumor volumes ± SEM. Mice were randomized into six groups: vehicle (N=6), poziotinib 2.5mg/kg (N=7), erlotinib 100mg/kg (N=6), afatinib 20mg/kg (N=6), osimertinib 5mg/kg (N=6), and osimertinib 25mg/kg (N=6). Mice received drug 5 days per week, and mice were euthanized at day 28 to harvest tumors. **g,** Dot plot of percent change in tumor volume on day 28 of tumors described in **f**. Dots are representative of each tumor, and bars are representative of average ± SEM for each group. Statistical differences were determined by ordinary one-way ANOVA with post-hoc Tukey's multiple comparisons test to determine differences between groups vehicle, N=6 mice, poziotinib 2.5mg/kg, N=7 mice, erlotinib 100mg/kg, N=6 mice, afatinib 20mg/kg, N=6 mice, osimertinib 5mg/kg, N=6, and osimertinib 25mg/kg, N=6 mice.

[Source data](#)

Extended Data Fig. 6 Exon 20 loop insertions are a distinct class of EGFR mutations.

a, Dot plot of mutant/WT IC₅₀ values of Ba/F3 cells expressing exon 20 loop insertion mutations and treated with indicated classes of EGFR TKIs. Dots are representative of average of n=3 replicate mutant/WT IC₅₀ values of individual cell lines expressing exon 20 insertion mutations with individual drugs. Bars are representative of average mutant/WT IC₅₀ values ± SEM for each class of EGFR TKI and all Ba/F3 cell lines. p-values were determined by one-way ANOVA with unequal SD as determined by Brown-Forsythe test to determine differences in SD. Holm-Sidak's multiple comparisons test was used to determine differences between groups. **b**, Tumor growth curves for PDXs harboring EGFR S768dupSVD exon 20 insertion mutation treated with indicated inhibitors. Tumors were measured three times per week and symbols are average of tumor volumes ± SEM. Mice were randomized into four groups: vehicle (N=4), poziotinib 2.5mg/kg (N=5), osimertinib 5mg/kg (N=4), and osimertinib 25mg/kg (N=5). Mice received drug 5 days per week, and mice were euthanized at day 21 to harvest tumors. **c**, Dot plot of percent change in tumor volume on day 21 of tumors described in panel **c** (vehicle, N=4 mice, poziotinib 2.5mg/kg, N=5 mice, osimertinib 5mg/kg, N=4 mice, and osimertinib 25mg/kg, N=5 mice). Dots are representative of each tumor, and bars are representative of average ± SEM for each group. Statistical differences were determined by ordinary one-way ANOVA with post-hoc Tukey's multiple comparisons test to determine differences between groups. **d**, Heat map with unsupervised hierarchical clustering of log (Mutant/WT) ratios from Ba/F3 cells expressing indicated exon 20 loop insertions after 72 h of indicated drug treatment. Squares are representative of the median of n=3 replicates. **e**, Bar plot of average mutant/WT ratio of Ba/F3 cells expressing exon 20 loop insertions separated by mutational groups for indicated drug classes. Bars are representative of average mutant/WT ratio for all mutations and drugs in the indicated groups, dots are representative of average (n=3) mutant/WT ratio of individual mutations and drugs. Error bars are representative of SEM for each bar. Statistical differences between near- and far-loop mutants was determined by two-sided unpaired t-test with unequal variance

[Source data](#).

[Extended Data Fig. 7 Drug repurposing can overcome T790M-like resistance mutations.](#)

a, Heat map with unsupervised hierarchical clustering of log (mutant/WT) ratios from Ba/F3 cells expressing indicated mutations after 72 h of indicated drug treatment. Squares are representative of the median of n=3 replicates. For co-occurring mutations, the order of mutations 1, 2, and 3 were assigned arbitrarily. Groups were assigned based on hierarchical clustering and known resistance mutations. **b, c**, Dot plot of mutant/WT IC₅₀ values of Ba/F3 cells expressing T790M-like-3S (3rd-gen TKI

sensitive) (**b**) and T790M-like-3R (3rd-gen TKI resistant) (**c**) treated with indicated classes of EGFR TKIs. Dots are representative of average of n=3 replicate mutant/WT IC₅₀ values of individual cell lines with individual drugs and drugs of each class are grouped together. Bars are representative of average mutant/WT IC₅₀ values ± SEM for each class of EGFR TKI and all cell lines. p-values were determined by one-way ANOVA with unequal variance as determined by Brown-Forsythe test to determine differences in variance. Holm-Sidak's multiple comparisons test was used to determine differences between groups

[Source data.](#)

Extended Data Fig. 8 PACC mutations alter the orientation of the P-loop and/or α C-helix and are sensitive to 2nd-gen TKIs.

a, Overlap of G719S (PDB 2ITN, green) and WT EGFR (PDB 2ITX, grey) crystal structures demonstrate a robust shift of F723 (red arrow) in the P-loop orienting the benzyl ring in a downward position condensing the P-loop in the drug binding pocket. Further, G719S displays an inward shift of the α C-helix compared to the WT crystal structure. **b**, Surface representation of G719S (PDB 2ITN) with P-loop (red), α C-helix (blue), hinge region (orange), C797 (yellow), and DFG motif (green) highlighted to demonstrate steric hindrance of drug binding pocket caused by shifted P-loop. **c**, Comparison of osimertinib bound to wild-type EGFR (PDB 4ZAU, green) or EGFR G719S (PDB 2ITN, purple) demonstrates destabilization of TKI-protein interactions. **d**, In silico homology model of EGFR L718Q (pink) with predicted binding modes of osimertinib and poziotinib structures demonstrates that Q718 hinders the interaction of osimertinib (green) with M793 and shifts the Michael acceptor (reactive group, green arrow) out of alignment with C797 (yellow arrow). In contrast, poziotinib (blue) is less effected by Q718 and is still positioned to react with C797, even in the context of L718Q mutations. **e**, In silico modeling of EGFR G719S (purple) with poziotinib (blue) shows no predicted changes in poziotinib binding or TKI-protein interactions. **f**, Dot plot of percent change in tumor volume on day 28 of tumors described in Fig. 3c. Dots are representative of each tumor, and bars are representative of average ± SEM for each group, N=5 mice per treatment group. Statistical differences were determined by ordinary one-way ANOVA with post-hoc Tukey's multiple comparisons test to determine differences between groups. **g**, Representative CT images from a patient harbouring E709K G719S complex mutation before and after four weeks of afatinib treatment. Red arrow indicates resolved pleural effusion on the left lobe and reduced pleural effusion and tumor in the right lobe. **h**, In silico modeling of EGFR Ex19del G796S (purple) with osimertinib in the reactive (blue) and predicted (orange) conformations demonstrate destabilization of TKI-protein interactions in the hinge region (yellow), displacing the reactive group (arrow). **i**, In silico modeling of EGFR

Ex19del G796S (purple) with the reactive conformation of poziotinib (blue) and the predicted conformation of poziotinib (orange) predicted minimal changes in poziotinib binding and similar TKI-protein interactions.

[Source data](#)

Extended Data Fig. 9 2nd-gen TKIs confer durable clinical benefit in patients with acquired osimertinib-resistant NSCLC.

a, Representative CT images from a patient after 5.5 months of osimertinib and 6 months after afatinib treatment. Red arrow indicates lesion harboring L858R/L718V. **b**, CT scan of a patient after 17 months of osimertinib treatment showed new pleural lesion that tested positive for both EGFR L858R and L718V mutations (red arrow), and CT image of patient four weeks after beginning poziotinib treatment shows reduction in size of the pleural lesion (red arrow). Blue arrow indicates resolved pleural effusion. **a, b**, Schematic below CT images shows timeline of patient treatments and outcomes. **c**, Schematic representations of a patient treatments and outcomes that acquired two PACC mutations (V765L and C797S) after 18 months of osimertinib treatment. PR = partial response, PD = progressive disease, SD = stable disease, SRS = stereotactic radiosurgery Chemo/IO = carboplatin/ pemetrexed+ pembrolizumab.

Extended Data Fig. 10 Structure-function groups identify patients with greater benefit to 2nd-gen TKIs than exon-based groups.

a, b, Overall response rate to afatinib stratified by structure-function-based groups (N= 507: Classical-like N=91, T790M-like N=103, Ex20ins-L N=120, and PACC N=193) (**a**) or exon based groups (N= 528: Exon 18 N=133, Exon 19 N=22, Exon 20 N=294, Exon 21 N=79) (**b**). When mutations were not explicitly stated (N=21), those patients were excluded from the structure-function based analysis. **c**, Kaplan-Meier plot of duration of afatinib treatment of patients with NSCLC tumors harboring atypical EGFR mutations (N= 364 patients) stratified by exon-based groups. Exon 18 N=87, Exon 19 N=19, Exon 20 N=195, and Exon 21 N=63. **d**, Kaplan-Meier plot of TTF of patients with NSCLC harboring non-PACC atypical EGFR mutations (N= 56) treated with 1st- (N=25), 2nd- (N=13), or 3rd-gen (N=18) EGFR TKIs. Forrest plot comparing PACC and non-PACC mutants can be found in Fig. [4d](#). **e-h**, Kaplan-Meier plots of TTF of patients atypical EGFR mutations stratified by EGFR TKI class for exons 18 (N=40) (**e**), 19 (N=19) (**f**), 20 (N=24) (**g**), and 21 (N=26) (**h**). Forrest plot comparing HRs and p-values across exons can be found in Fig. [4d](#).

[Source data](#)

Supplementary information

Supplementary Tables

This file contains Supplementary Tables 1–7.

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. 8

Source Data Extended Data Fig. 10

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the

original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

Reprints and Permissions

About this article

Cite this article

Robichaux, J.P., Le, X., Vijayan, R.S.K. *et al.* Structure-based classification predicts drug response in EGFR-mutant NSCLC. *Nature* **597**, 732–737 (2021).
<https://doi.org/10.1038/s41586-021-03898-1>

- Received: 13 April 2021
- Accepted: 11 August 2021
- Published: 15 September 2021
- Issue Date: 30 September 2021
- DOI: <https://doi.org/10.1038/s41586-021-03898-1>

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

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

Collections