



THE SPINAL FRONTIER

High-quality sequences for vertebrate species shed light on genome evolution

Health disparities
COVID-19 highlights the deadly toll of inequality

Melting point
Loss of glacier mass has accelerated since 2000

Microbial manoeuvre
High CO₂ levels reverse key metabolic cycle in extremophiles

Vol. 592, No. 7896
nature.com/nature

Nature.2021.05.01

[Sat, 01 May 2021]

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

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

Donation

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

This Week

- **[It's time to invite more people to join clinical trials](#)** [28 April 2021]
Editorial • Drug trials need more participants. Research shows the potential benefits of changing the criteria used to determine who can enrol.
- **[To remedy health disparities, more scientists must ‘get political’](#)** [28 April 2021]
Editorial • The pandemic has given scientists a more prominent voice in society. They need to use it to push for better health through equality.
- **[COVID vaccines: time to confront anti-vax aggression](#)** [27 April 2021]
World View • Halting the spread of the coronavirus will require a high-level counteroffensive against new destructive forces.
- **[A century of US data documents obesity’s racially skewed rise](#)** [23 April 2021]
Research Highlight • An analysis also finds that obesity is common at a much younger age among people born in the early 1980s than those born in the late 1950s.
- **[Animals around the world follow the ‘island rule’ to a curious fate](#)** [15 April 2021]
Research Highlight • Small animals that move to islands tend to evolve into giants, but big animals that do the same shrink.
- **[The world’s lightest uranium atom reveals nuclear secrets](#)** [22 April 2021]
Research Highlight • A flyweight isotope of uranium helps to shed light on a fundamental form of nuclear decay.
- **[Rivers give off stealth carbon at night](#)** [21 April 2021]
Research Highlight • Emissions made under cover of darkness account for much of the carbon flux from flowing waters.
- **[Material mimicking lobster belly cracks the code for toughness](#)** [23 April 2021]
Research Highlight • Who needs a shell? Crustacean’s lower surface spawns a synthetic substance that could lead to flexible armour.

- **Baboons show signs of prolonged stress — and the toll it takes** [21 April 2021]

Research Highlight • High levels of a stress hormone and truncated survival are linked for the first time in a free-ranging animal.

- **Babies' gut bacteria teem with antibiotic-resistance genes**

[21 April 2021]

Research Highlight • A young child's 'resistome' is shaped by the presence of older siblings and other environmental factors.

EDITORIAL

28 April 2021

It's time to invite more people to join clinical trials

Drug trials need more participants. Research shows the potential benefits of changing the criteria used to determine who can enrol.



Clinical trials participants tend to be younger, healthier people. More inclusive criteria are needed.Credit: BSIP/Universal Images Group/Getty

It took Patty Spears, a resident of North Carolina, three attempts to be allowed to participate in a clinical trial for a cancer vaccine to reduce the likelihood that her breast cancer would recur. For the first two trials that she applied for, Spears didn't meet the eligibility criteria — strict guidelines that determine who can participate in a trial. These criteria tend to favour

younger, healthier people. Even the third time around, Spears was nearly ruled out because her white blood cell count was barely above the study's minimum requirement.

That was more than 20 years ago. Today, Spears is a patient advocate at the University of North Carolina at Chapel Hill. Along with individuals at other organizations, including the US Food and Drug Administration (FDA) and the US National Cancer Institute, she is part of an effort that aims to expand eligibility for cancer clinical trials. They want more participants to find trials, and more trials to find participants.



[How Facebook and Twitter could be the next disruptive force in clinical trials](#)

The testing of therapies on a wider cohort of participants can increase the safety and efficacy of treatments, especially for those under-represented by medical research, such as older people and those from minority groups. For the funders and organizers of clinical trials, admitting a greater diversity of people potentially means more people taking part in trials. That could mean some trials get concluded more quickly — an important consideration, given that many clinical trials fail to meet their planned timeline for enrolling a full set of participants — and at lower cost.

The effort to expand eligibility is crucial. But it needs more support from funders and regulators around the world. Attempts to gain this support could be helped by the accumulating evidence — including a study published this month in *Nature* — showing the benefits of allowing more people to participate in clinical trials.

Exclusions apply

Most clinical trials have a list of eligibility criteria that must be met before a participant can enrol. These requirements vary from trial to trial and can be designated by investigators, study sponsors and, when they are involved in study design, patient groups. Criteria are devised to protect the safety of participants, so trials might exclude people who are unwell, older or pregnant. Exclusion criteria might also yield ‘clean’ data — that is, data on people who are more like each other. But it also means that trial participants are less representative of patients — who come in all ages and have a spectrum of health conditions.

Too often, the requirements are selected simply because the list of exclusion criteria has become a template, carried forwards without scrutiny from one trial to the next. Restricting eligibility in this way can disproportionately affect groups that are already under-represented in medical research. For example, in the United States, diabetes is more common among Black people than white people, and can lead to reduced kidney function. As a result, trials that exclude people with reduced kidney function could disproportionately exclude Black participants.



Why emergency COVID-vaccine approvals pose a dilemma for scientists

A more systematic, scientific approach to crafting eligibility criteria could help. In a study published in *Nature* on 7 April, researchers studied the electronic medical records of more than 60,000 people in the United States with advanced non-small-cell lung cancer¹. The team compared the survival outcomes of people who had participated in clinical trials of drugs for this type of cancer² and people who would have been excluded from participating in clinical trials but who had taken the same drugs outside the studies. The results showed that if a more-diverse group of people had been allowed to take part in the trials, the overall survival outcomes would have been almost the same — but that the pool of eligible trial participants would have more than doubled.

In a separate study, pharmacologist Donald Harvey at Emory University in Atlanta, Georgia, also showed that the widening of eligibility criteria is beneficial to trials for non-small-cell lung cancer drugs. According to data presented at a 9 April meeting held by Friends of Cancer Research, a think tank and advocacy group in Washington DC, and the American Society of Clinical Oncology, allowing people with cancers and those with impaired kidney function to take part in trials increased the proportion of participants aged 75 or older from 16% to 22%. This is important, because the majority of people with cancer are older, yet older people with cancer are under-represented in clinical trials.



How researchers are addressing the racial imbalance in multiple myeloma

These studies follow the publication in March of a fresh set of recommendations from the two organizations. Both have been working to re-evaluate commonly used eligibility criteria since 2016³. They are recommending guidelines for making science-based decisions about whether people who are taking or have recently taken other medications should be enrolled in studies. Now that clinical-trial investigators, researchers and funders are taking the first steps towards changing standard practice, regulators must show support. In 2020, the FDA issued guidance to clinical-trial designers regarding criteria such as HIV status and the presence of brain metastases.

This is impressive progress, but it is time for the effort to broaden its reach — beyond cancer and beyond the United States. Explicit endorsement from other regulators and trial sponsors could propel the movement internationally, and further analyses of electronic medical records could help to establish which requirements should be kept and which are superfluous for studies of various conditions. Together, these changes could foster trials that are faster and more meaningful for the patients they are ultimately meant to serve.

Widening the criteria for trial participation will take a concerted international effort from investigators, trial sponsors and drug regulators. A

more systematic approach, driven by data and greater involvement of patient groups, can and should be used to select participants — not only for cancer clinical trials, but also for studies for other diseases.

Nature **592**, 659-660 (2021)

doi: <https://doi.org/10.1038/d41586-021-01099-4>

References

1. 1.

Liu, R. *et al. Nature* **592**, 629–633 (2021).

2. 2.

Fehrenbacher, L., Ackerson, L. & Somkin, C. *J. Clin. Oncol.* **27** (suppl.), 6538 (2009).

3. 3.

Kim, E. S. *et al. Clin. Cancer Res.* <https://doi.org/10.1158/1078-0432.CCR-20-3852> (2021).

Jobs from Nature Careers

- - - [All jobs](#)
 - - [Research Associate / Postdoc](#)

[Technische Universität Dresden \(TU Dresden\)](#)

[Dresden, Germany](#)

JOB POST

- **PhD position in evolutionary ecology of insect immune priming**

Université Bourgogne Franche-Comté (UBFC)

DIJON, France

JOB POST

- **PhD contract Doctoral contract in sociology-demography : Analysis of differential socio-demographic behaviors by sex among 50-74 year olds**

Université Bourgogne Franche-Comté (UBFC)

DIJON, France

JOB POST

- **PhD - Dark ecological networks and land management: methodological and strategic contributions for reducing ecological impacts of artificial light**

Université Bourgogne Franche-Comté (UBFC)

BESANCON, France

JOB POST

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

EDITORIAL

28 April 2021

To remedy health disparities, more scientists must ‘get political’

The pandemic has given scientists a more prominent voice in society. They need to use it to push for better health through equality.



People attend a drive-through COVID-19 vaccination event at Fresno City College, in one of California’s poorest regions. Credit: Brian L. Frank for *Nature*

For more than 150 years, scholarship and research have revealed how poor and marginalized communities are disproportionately affected by disease. People are more likely to become unwell if they earn low wages, have few employment protections, live in unsafe environments, receive poor-quality

education, or are discriminated against. Whether Prussia's typhus epidemic of 1847–48, tuberculosis outbreaks in the United States in the 1930s or chronic diseases today, researchers conclude that people would live longer, healthier lives if a society's collective wealth could be shared more equally ([M. Marmot *Lancet* 365, 1099–1104; 2005](#)).

Scholars in disciplines ranging from economics to epidemiology and sociology have proposed ideas for how to share the world's wealth ([R. G. Wilkinson and K. E. Pickett *Soc. Sci. Med.* 65, 1965–1978; 2007](#)). But their advice has mostly been disregarded by politicians. This is in part because the idea that the public and private sectors need to have a greater role in reducing inequality has been at odds with the thrust of global politics [for at least four decades](#).

During the COVID-19 pandemic, the successes that scientists have scored with drugs, vaccines and other interventions have given researchers a voice in decision-making. They need to use that position to advocate policies that would improve social determinants of better health, such as living wages, employment protections and high-quality educational opportunities. In this way, scientists need to 'get political'.



[Will COVID force public health to confront America's epic inequality?](#)

That will require, among other things, scientists to consider how they can best achieve political impact and policy engagement. But advice is on hand. A [News Feature this week](#) describes how community organizations in one of the poorest regions of the United States, California's San Joaquin Valley, tried to curb COVID-19 in communities of colour by tackling some of the disease's underlying determinants, in part through political engagement.

Hundreds of thousands of people in the valley — mainly immigrants — work on farms and in food-processing or meat-packing plants. Compared with California's more affluent regions, wages in the valley are low and labour protections weak. And neighbourhoods of agricultural workers often have poor-quality schools, insufficient clinics and few markets selling healthy food. Some areas even lack clean, running water. A child born in San Francisco, California, is expected to live for at least ten years longer than children born in many parts of the valley.

State and county public-health officials know this, but are often unable to push local leaders for the necessary policy changes. This is because they are generally hired to carry out the wishes of elected politicians, and their budgets and jurisdictions are therefore determined by those politicians.

But academic scientists are not tied by these constraints. During the pandemic, researchers in the San Joaquin Valley have partnered with grass-roots groups to try to address inequities and push agriculture companies to report COVID-19 outbreaks and protect their employees with face masks and physical distancing. They have also distributed free tests, and provided outreach and financial assistance for under-served communities.

Funding gaps

The COVID-19 pandemic has laid bare society's inequities more starkly than most people today will have witnessed. In the United States, Black, Hispanic and Latinx people are about three times as likely as white Americans to be hospitalized with COVID-19. In the United Kingdom, Bangladeshi-Britons have [twice the risk of death from COVID-19](#) as white British people. In Singapore, most COVID-19 cases have occurred in migrant workers, including those from Malaysia, Bangladesh and India.



The dangers of fringe economics in government

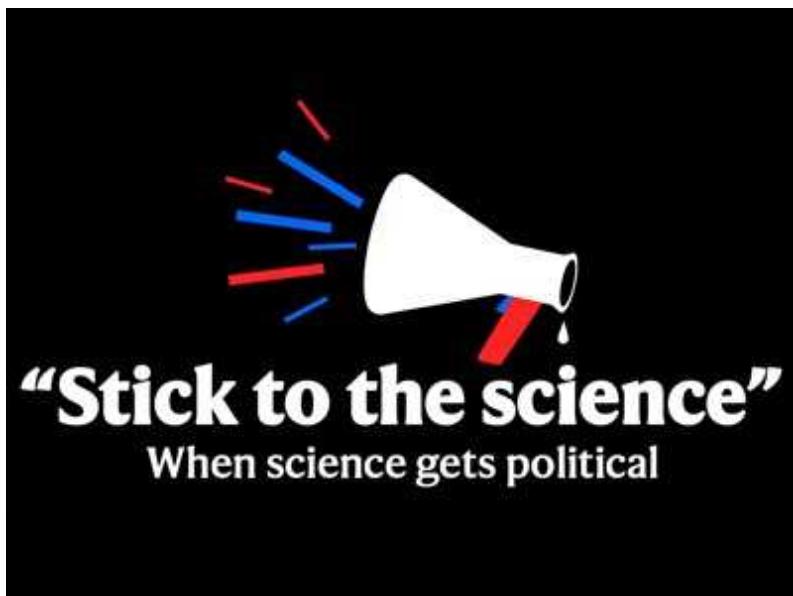
And the gap between rich and poor is rising in much of the world. In the United States alone, billionaires have increased their wealth by 44% — to a total of US\$4.3 trillion — in the past year, and 1.5 million jobs were created in the top wage bracket. And yet, 7.9 million low-wage workers lost their jobs. The number of families reporting hunger and homelessness climbed in 2020.

In the United States and Europe, funders have responded with grants to explore and address health disparities and their social determinants. Researchers should push themselves to go further than simply collecting data on the social and economic roots of poor health, and perhaps even deeper than devising interventions to address immediate needs, such as programmes to teach communities about nutrition.

Getting involved

But there are few long-term funding opportunities for such work, or even for researchers whose main objective is evidence-based policy — let alone systemic reform — and that, too, needs to change. Funders and research leaders must place a higher value on these types of impact in research-evaluation criteria. Then scientists would have a greater incentive to

collaborate with economists and political scientists to devise ways to share wealth and turn around rising inequality. Those who study racism could work with epidemiologists to better understand why economic and political systems have marginalized certain groups of people for decades, and how reparations or other reforms could begin to turn the tide.



['Stick to the science': Nature's podcast series on science and politics](#)

Researchers can also work with think tanks to write the short, research-informed reports that are required reading for politicians and policymakers, but are not part of conventional journal publishing. And they could co-design their studies with grass-roots groups who advocate for — and work with — communities in need. Scientists should also consider standing for elected bodies such as local, regional or national legislatures — which often have a dearth of representatives with research expertise.

Scientific discoveries and inventions made during the pandemic have led to progress in diagnostics, therapies and, of course, vaccine production. But the pandemic is far from over, and, combined with economic inequality and climate change, the world is in a precarious era. Now that the pandemic has elevated scientists' voice in society, more must learn how best to use that voice to advance the cause of economic, racial and social justice. Without such change, the essential research that is scientists' main focus will

ultimately fall short of achieving its goal of building healthier, more resilient, more equal and more just societies.

Nature **592**, 660 (2021)

doi: <https://doi.org/10.1038/d41586-021-01100-0>

Jobs from Nature Careers

- - - [All jobs](#)
 - - [**Research Associate / Postdoc**](#)
 - [Technische Universität Dresden \(TU Dresden\)](#)
 - [Dresden, Germany](#)
 - [JOB POST](#)
 - [**PhD position in evolutionary ecology of insect immune priming**](#)
 - [Université Bourgogne Franche-Comté \(UBFC\)](#)
 - [DIJON, France](#)
 - [JOB POST](#)
 - [**PhD contract Doctoral contract in sociology-demography : Analysis of differential socio-demographic behaviors by sex among 50-74 year olds**](#)
 - [Université Bourgogne Franche-Comté \(UBFC\)](#)

DIJON, France

JOB POST

- **PhD - Dark ecological networks and land management: methodological and strategic contributions for reducing ecological impacts of artificial light**

Université Bourgogne Franche-Comté (UBFC)

BESANCON, France

JOB POST

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01100-0>

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

WORLD VIEW

27 April 2021

COVID vaccines: time to confront anti-vax aggression



Halting the spread of the coronavirus will require a high-level counteroffensive against new destructive forces.

- [Peter Hotez](#) 0

1. [Peter Hotez](#)

1. Peter Hotez is a vaccine scientist and author at Baylor College of Medicine and Texas Children's Hospital in Houston, Texas. His latest book is *Preventing the Next Pandemic: Vaccine Diplomacy in a Time of Anti-Science*. Twitter: [@PeterHotez](#)

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Nearly one billion COVID-19 vaccine doses have been delivered in less than six months, but anti-vaccine disinformation and targeted attacks on scientists are undermining progress. These threats must be confronted directly, and the authority and expertise of the health community alone aren't enough to do this.

Even before the pandemic, I had a front-row seat to all of this. I have co-led efforts to develop vaccines in programmes, including a COVID-19 vaccine currently being tested in India. I also have an adult daughter with autism; my 2018 book, *Vaccines Did Not Cause Rachel's Autism*, became a dog whistle for anti-vaccine activists.

The World Health Organization recognized vaccine hesitancy as a top threat to global health before the pandemic. As COVID-19 vaccines moved through development, the public-health communities anticipated considerable vaccine hesitancy. Teams of experts, including me, began meeting regularly online to discuss how best to amplify evidence-based messages, deliver public-service announcements and address concerns around COVID-19 immunization.



The race to curb the spread of COVID vaccine disinformation

I experienced sinking feelings during our Zoom calls. Although certainly worthwhile, I knew that messaging alone would be inadequate. We'd

already seen this inadequacy in our efforts to prevent measles from returning to the United States and Europe in 2019, and to bolster vaccination rates for human papillomavirus to prevent cervical and other cancers. With COVID-19, our pro-vaccine messages would be drops in a vast sea of misinformation, much of it poured in deliberately by anti-vaccine forces.

I have a long-standing disagreement with many of my US public-health colleagues. I admire their commitment to disease prevention, but when I ask for a more direct way to counter anti-vaccine aggression, I'm told, "that's not our approach; confrontation gives them a platform and oxygen." In my opinion, this attitude reflects a time when we had dial-up modems. Today, the anti-vaccine empire has hundreds of websites and perhaps 58 million followers on social media. The bad guys are winning, in part because health agencies either underestimate or deny the reach of anti-science forces, and are ill-equipped to counter it.

Investigations by the US State Department and the UK Foreign Office have described how Russian intelligence organizations seek to discredit Western COVID-19 vaccines. One campaign implies that it could turn people into monkeys. This builds on a longer, well-documented history of Russia-sponsored disinformation, presumably to destabilize the United States and other democratic countries. The administration of US President Joe Biden has warned Russian media groups to halt their anti-vaccine aggression, and announced sanctions tied to disinformation and other behaviour, but we need much more.



[Anti-vaccine movement could undermine efforts to end coronavirus pandemic, researchers warn](#)

The United States hosts the world's largest and best-organized anti-vaccine groups. According to the London-based Center for Countering Digital Hate, these are influential groups, not a spontaneous grass-roots movement. Many far-right extremist groups that spread false information about last year's US presidential election are doing the same about vaccines. Anti-vaccine groups also target Black communities; an anti-vaccine documentary released in March vilifies COVID-19 vaccine testing among African Americans, calling it "medical racism".

Global anti-vaccine messaging around the adenovirus vaccines means that more people will die and the pandemic will be prolonged. Extremely rare but life-threatening blood clots caused the United States to pause roll-out of the Johnson & Johnson vaccine, and many European nations have stopped or restricted use of the Oxford–AstraZeneca vaccine for similar reasons. However, those regions have other vaccine options, which is not the case for many countries. In March, Cameroon and the Democratic Republic of the Congo halted use of the Oxford–AstraZeneca vaccine, and the African Union has stopped procurement.

Many people in Africa are tapping into anti-vaccine messaging. A rumour-tracking program from the analytics company Novetta in McLean, Virginia, finds that Russia specifically targets African countries to discredit Western vaccines in favour of its own Sputnik V. US-based anti-vaccine groups invoke colonialism and eugenics. Now, tens of thousands of vaccine doses are going unused. Anti-vaccine disinformation has turned reasonable questions and concerns over rare side effects into conspiracy worries, exaggerated fears, and outrage at being treated like 'guinea pigs'.

Accurate, targeted counter-messaging from the global health community is important but insufficient, as is public pressure on social-media companies. The United Nations and the highest levels of governments must take direct, even confrontational, approaches with Russia, and move to dismantle anti-vaccine groups in the United States.

Efforts must expand into the realm of cyber security, law enforcement, public education and international relations. A high-level inter-agency task force reporting to the UN secretary-general could assess the full impact of anti-vaccine aggression, and propose tough, balanced measures. The task force should include experts who have tackled complex global threats such as terrorism, cyber attacks and nuclear armament, because anti-science is now approaching similar levels of peril. It is becoming increasingly clear that advancing immunization requires a counteroffensive.

Nature **592**, 661 (2021)

doi: <https://doi.org/10.1038/d41586-021-01084-x>

Jobs from Nature Careers

- - - [All jobs](#)
 -
 - [Research Associate / Postdoc](#)
 - [Technische Universität Dresden \(TU Dresden\)](#)
 - [Dresden, Germany](#)
 - [JOB POST](#)
 - [PhD position in evolutionary ecology of insect immune priming](#)
 - [Université Bourgogne Franche-Comté \(UBFC\)](#)
 - [DIJON, France](#)
 - [JOB POST](#)

- **PhD contract Doctoral contract in sociology-demography : Analysis of differential socio-demographic behaviors by sex among 50-74 year olds**

Université Bourgogne Franche-Comté (UBFC)

DIJON, France

JOB POST

- **PhD - Dark ecological networks and land management: methodological and strategic contributions for reducing ecological impacts of artificial light**

Université Bourgogne Franche-Comté (UBFC)

BESANCON, France

JOB POST

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01084-x>



Midshipmen in the United States in around 1900. A study found that body-mass index, a gauge of obesity, has increased with the generations during the twentieth century. Credit: Buyenlarge/Getty

Metabolism

23 April 2021

A century of US data documents obesity's racially skewed rise

An analysis also finds that obesity is common at a much younger age among people born in the early 1980s than those born in the late 1950s.

In the United States, the rise in obesity over the past 40 years has disproportionately affected Black and Hispanic people. But it has been difficult to discern how this disparity has changed historically.

To fill the gap, Yang Claire Yang and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill and their colleagues combined data from four large studies whose participants had been born throughout the twentieth century. The researchers divided the participants into cohorts based on birth year and analysed participants' body mass index (BMI), a proxy for obesity.

The team found that BMI rose more steeply from early adolescence to late middle age in more recent cohorts than in earlier ones. Within individual cohorts, average BMI was higher for Black and Hispanic people than for white people, and the BMI disparity between Black and white people was greater for more recent cohorts.

Socioeconomic inequality did not fully account for the racial or ethnic disparities in BMI. The stress of discrimination might partly drive these disparities, the authors say.

[Proc. Natl Acad. Sci. USA \(2021\)](#)

- [Metabolism](#)
-

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01049-0>



The Komodo dragon, which weighs on average some 70 kilograms, is an example of an island species that evolved outsize proportions. Credit: Alamy
Evolution

15 April 2021

Animals around the world follow the ‘island rule’ to a curious fate

Small animals that move to islands tend to evolve into giants, but big animals that do the same shrink.

An analysis of more than 1,000 species shows that birds, mammals and reptiles on islands tend to be either miniature or gigantic versions of their

mainland counterparts — evidence that an evolutionary tenet called the ‘island rule’ applies to a wide variety of vertebrates.

Big mainland animals often evolve smaller bodies on islands, whereas small mainland species become larger. A prominent example is the island-dwelling Komodo dragon (*Varanus komodoensis*), which is the world’s largest lizard and can exceed 150 kilograms.

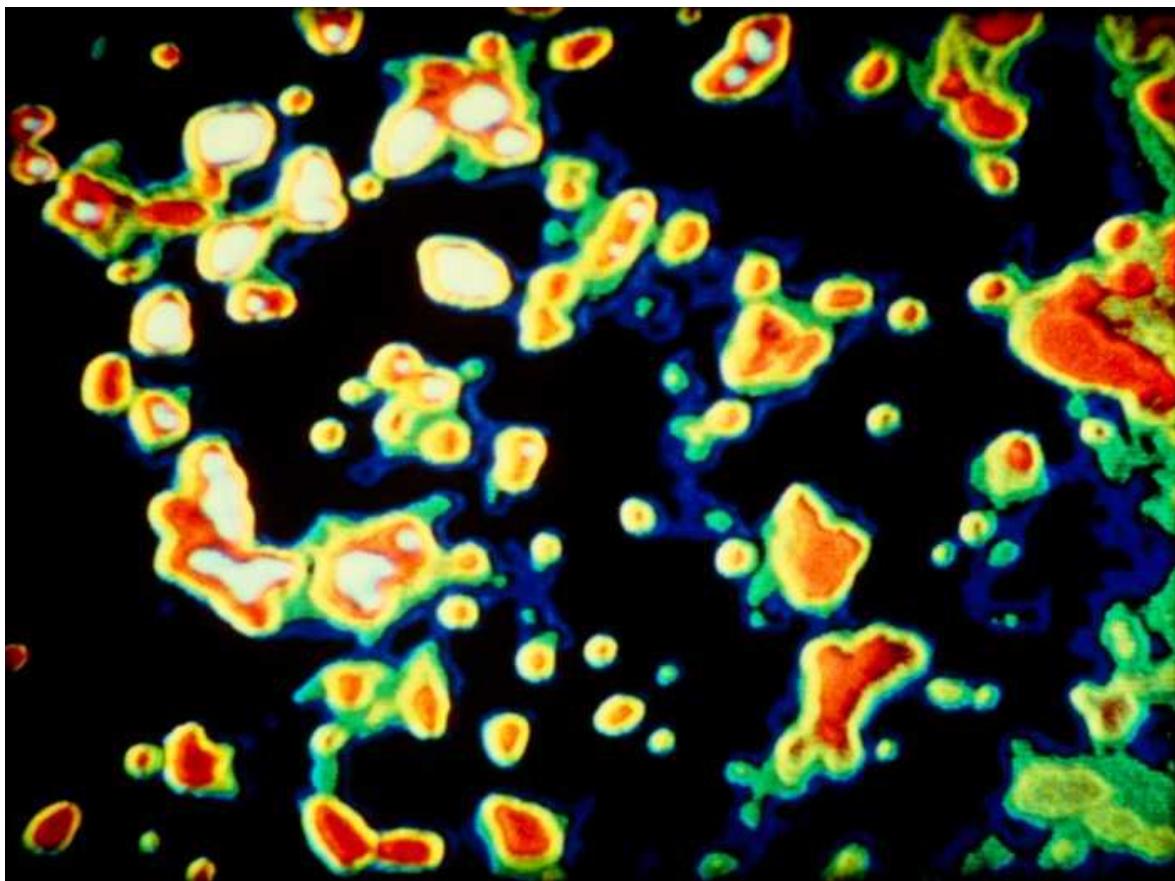
Ana Benítez-López at the Doñana Biological Station in Seville, Spain, and her colleagues set out to settle the debate about whether the dragon and other species are mere flukes or part of a broader evolutionary pattern. The team gathered data from multiple studies to examine 1,166 island-dwelling species and 886 of their mainland counterparts. In a departure from some previous studies, the researchers drew data from diverse sources, such as museum specimens and studies unrelated to the island rule.

The authors found widespread evidence for the island rule around the world. Shifts in body size were most extreme for mammals and reptiles on smaller, more remote islands.

[Nature Ecol. Evol. \(2021\)](#)

- [Evolution](#)

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-00981-5>



A mixture of solitary and clustered uranium atoms (artificially coloured). Scientists have produced an ultralight uranium atom with only 122 neutrons.
Credit: Dr Mitsuo Ohtsuki/Science Photo Library

Atomic and molecular physics

22 April 2021

The world's lightest uranium atom reveals nuclear secrets

A flyweight isotope of uranium helps to shed light on a fundamental form of nuclear decay.

The creation of the lightest uranium atom ever gives scientists a better understanding of a fundamental type of radioactive decay.

All elements have one or more isotopes, which differ from each other in the number of neutrons in their nuclei. Almost all naturally occurring uranium atoms contain either 143 or 146 neutrons.

Zai-Guo Gan at the Chinese Academy of Sciences in Lanzhou and his colleagues have produced a uranium isotope with only 122 neutrons by firing a beam of argon at a tungsten target until atoms of each element fused together — an extremely rare event that formed uranium atoms. The team then extracted the 122-neutron isotope using a magnetic device called a separator.

All uranium isotopes undergo α -decay, a process whereby an atom loses two protons and two neutrons. Unexpectedly, the authors found that their ultralight isotope and a previously detected uranium isotope containing 124 neutrons decay more easily than do light isotopes of other elements. This suggests that interactions between protons and neutrons in atomic nuclei can have a greater role in α -decay than previously thought.

[Phys. Rev. Lett. \(2021\)](#)

- [Atomic and molecular physics](#)

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01050-7>



Algae fills the Paktajokka River in northern Sweden. Photosynthesis in the river means that its daytime and night-time carbon emissions differ substantially. Credit: Gerard Rocher-Ros

Biogeochemistry

21 April 2021

Rivers give off stealth carbon at night

Emissions made under cover of darkness account for much of the carbon flux from flowing waters.

Rivers and streams might be giving off more carbon dioxide than had previously been thought — because they're doing it at night when nobody had been looking.

Flowing waters are an important source of carbon to the atmosphere, but researchers haven't had a good handle on how much carbon rivers emit. That's in part because a river's carbon emissions vary greatly over a 24-hour period, and scientists typically collect samples of river water only during the day — meaning they miss carbon fluxes at night.

Lluís Gómez-Gener at the Swiss Federal Institute of Technology in Lausanne, Gerard Rocher-Ros at Umeå University in Sweden and their colleagues analysed round-the-clock data gathered over the past few years by sensors in 66 rivers around the world. The sensors indicated that, on average, rivers emitted 27% more carbon dioxide at night than during the day.

The authors estimate that, globally, carbon emissions from rivers are roughly 35% higher than the previous estimates that were based on only daytime sampling.

[Nature Geosci. \(2021\)](#)

- [Biogeochemistry](#)

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01032-9>



The tough, flexible tissue covering a lobster's underbelly serves as a model for a new synthetic substance. Credit: D. Lada/Alamy

Biomaterials

23 April 2021

Material mimicking lobster belly cracks the code for toughness

Who needs a shell? Crustacean's lower surface spawns a synthetic substance that could lead to flexible armour.

A lobster's underbelly is protected by a membrane that is strong but flexible, allowing the animal to manoeuvre. Now, that membrane has inspired a strong, flexible synthetic material — work that could help pave the way for protective fabrics.

Xuanhe Zhao at the Massachusetts Institute of Technology in Cambridge and his colleagues constructed their material as the lobster does: from a hydrogel, a network of water-loving polymer chains. To better mimic the biological version, the researchers welded their polymer threads together with water and partially crystallized them with heat, creating small, scattered crystals inside the fibres, to add structural complexity.

Tearing apart the resulting film requires more than ten times the energy needed to rip up an uncrosslinked control sample. Further improvement came when the team stacked layers of polymer in the same manner as the sheets of chitin fibres in the lobster's soft armour: with each layer at an angle to the one underneath it, like a twisted stack of playing cards.

The researchers shot microscopic steel particles at the material to test its impact resistance, and say it shows promise for making flexible body armour.

Matter (2021)

- Biomaterials

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01053-4>



A baboon in Amboseli National Park in Kenya. Baboons with high levels of a stress hormone died earlier than those with low levels. Credit: Fernando A. Campos

Evolution

21 April 2021

Baboons show signs of prolonged stress — and the toll it takes

High levels of a stress hormone and truncated survival are linked for the first time in a free-ranging animal.

Chronic stress cuts the life expectancy of female baboons by up to one-quarter.

Baboons living in the Amboseli basin in Kenya have been studied extensively since 1971, allowing researchers to model environmental and behavioural factors that could affect the animals' survival. To assess the relationship between lifespan and the body's stress response, Fernando Campos at the University of Texas at San Antonio and his colleagues analysed levels of the stress hormone glucocorticoid in faecal samples collected over 19 years from more than 240 female baboons at Amboseli.

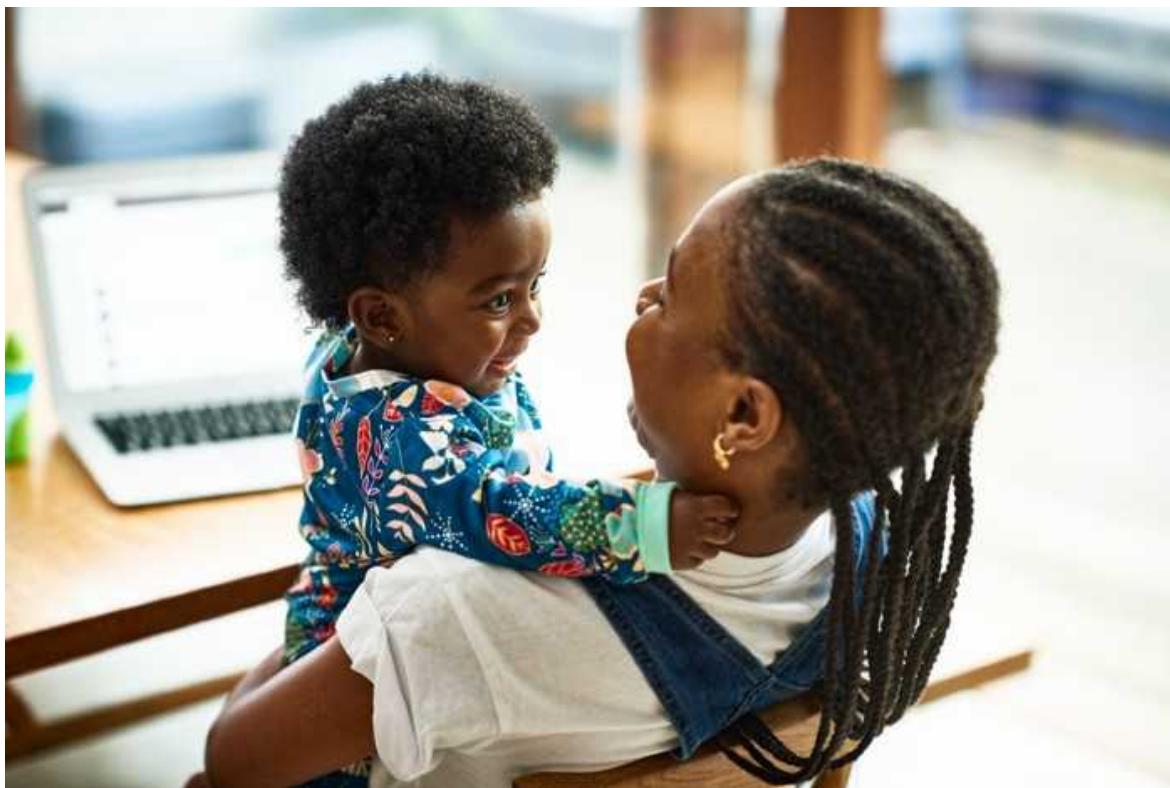
Animals with elevated glucocorticoid levels over a sustained period had shorter lives than those with lower levels, the researchers found. A baboon with glucocorticoid levels in the top 90% for her age throughout adulthood would, on average, die 5.4 years earlier than another that had levels in the lowest 10% for her age, according to the team's models.

Many scientists have suspected that animals with elevated glucocorticoid levels — which are often associated with lifelong stress — have shorter lives, but this is the first time the effect has been shown in animals outside the laboratory.

[*Sci. Adv. \(2021\)*](#)

- [Evolution](#)

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01052-5>



Changing-table surprise: stool samples reveal that babies' microbiomes harbour hundreds of types of antibiotic-resistance gene. Credit: Getty

Microbiology

21 April 2021

Babies' gut bacteria teem with antibiotic-resistance genes

A young child's 'resistome' is shaped by the presence of older siblings and other environmental factors.

Hundreds of genes for antibiotic resistance lurk in the bacteria in babies' guts — and antibiotic consumption during late pregnancy influences such genes in the infant's microbiome.

The human gut is home to trillions of bacteria, including many with genes that help them to fight off antibiotics. To investigate how resistance arises among a person's resident microbes, Søren Sørensen at the University of Copenhagen and his colleagues analysed stool samples from 662 one-year-olds in Denmark.

The team found 409 types of gene that granted resistance to 34 classes of antibiotics. Of those gene types, 167 provided resistance to multiple antibiotics, including critically important drugs that the infants had not taken.

The mixture of resistance genes was influenced by environmental factors, such as whether the children grew up in rural or urban areas. Using antibiotics during pregnancy 40 days or less before childbirth also played a part.

Higher levels of antibiotic-resistance genes were associated with a particular mixture of gut microbes that the team had previously linked to an elevated risk of asthma in later life.

[Cell Host Microbe \(2021\)](#)

- [Microbiology](#)

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01051-6>

News in Focus

- **[Africa's vaccine revolution, fetal-tissue studies and infection protection](#)** [28 April 2021]
News Round-Up • The latest science news, in brief.
- **[India's massive COVID surge puzzles scientists](#)** [21 April 2021]
News • The virus is spreading faster than ever before in India despite previous high infection rates in megacities, which should have conferred some protection.
- **[Lift off! First flight on Mars launches new way to explore worlds](#)** [19 April 2021]
News • NASA's Ingenuity helicopter successfully hovered for 40 seconds in Mars's thin atmosphere.
- **[COVID vaccines and kids: five questions as trials begin](#)** [21 April 2021]
News Explainer • As the first clinical trials in young children start, here's what scientists want to know.
- **[‘The damage is total’: fire rips through historic South African library and plant collection](#)** [19 April 2021]
News • University of Cape Town faces losing ‘irreplaceable’ historical material on anthropology, ecology and politics.
- **[US pledges to dramatically slash greenhouse emissions over next decade](#)** [22 April 2021]
News • Researchers say the commitment to roughly halve emissions is a good start to slowing climate change, but fear it still isn't enough.
- **[Will COVID force public health to confront America’s epic inequality?](#)** [28 April 2021]
News Feature • In California’s San Joaquin Valley, some researchers are turning political to address the social determinants of health.
- **[These cellular clocks help explain why elephants are bigger than mice](#)** [27 April 2021]
News Feature • Biologists are uncovering how tiny timekeepers in our cells might govern body size, lifespan and ageing.

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

NEWS ROUND-UP

28 April 2021

Africa's vaccine revolution, fetal-tissue studies and infection protection

The latest science news, in brief.



A shipment of Oxford–AstraZeneca COVID-19 vaccines lands at Abidjan, Ivory Coast. The African Union wants most vaccines to be produced on the continent by 2040.Credit: Sia Kambou/AFP/Getty

Africa plans a vaccine-manufacturing revolution

Africa's leaders are on a path to ramp up vaccine manufacturing and boost regulatory bodies for medicines. On 13 April, they pledged to increase the share of vaccines used in Africa that are manufactured on the continent from 1% to 60% by 2040. This will involve [building factories and bolstering capacity](#) in research and development.

The COVID-19 pandemic has left Africa woefully short of vaccines, according to John Nkengasong, director of the Africa Centres for Disease Control and Prevention (Africa CDC), based in Addis Ababa. The ambitious move represents an important step in boosting Africa's capacity in public health, he added.

Nkengasong was speaking at a 2-day vaccine summit on 12–13 April, co-organized by Africa CDC and the African Union, and attended by 40,000 delegates. Also taking part were heads of state and leaders from research and business.

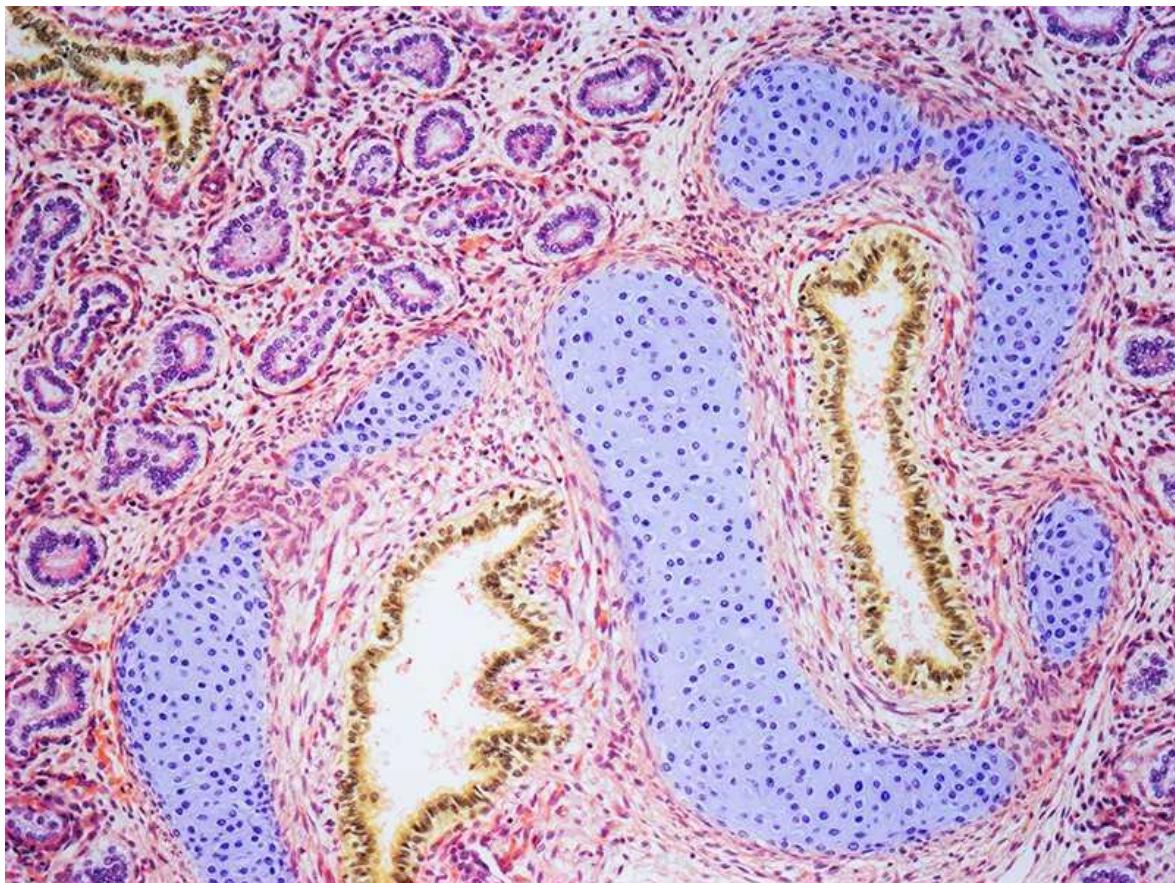
“We have been humbled, all of us, by this pandemic,” said Abdoulaye Diouf Sarr, Senegal’s minister of health and welfare, at the summit. The 1% figure “boggles the mind”, added virologist Salim Abdool Karim, formerly a science adviser to South Africa’s government.

Africa, a continent of 1.2 billion people, has only around 10 vaccine manufacturers. Some 99% of its vaccines are imported; most come from international procurement mechanisms, such as one organized through the United Nations children’s charity UNICEF.

The pandemic has also led to calls to accelerate efforts to establish an African Medicines Agency (AMA) — similar to the European Medicines Agency — which would provide national regulators with guidance on new medicines. The project is being led by the African Union and the Africa CDC. The AMA will cost US\$100 million to establish.

As of February 2021, only eight countries had ratified an international treaty needed to launch the AMA ([B. M. Ncube et al. J. Pharm. Policy Pract. 14, 29; 2021](#)). Fifteen countries will need to ratify it to bring it into existence; Tedros Adhanom Ghebreyesus, director-general of the World Health

Organization, urged countries to quickly enact relevant laws in their national parliaments.



Research on fetal tissue, such as the lung tissue shown here, has drawn controversy for years. Credit: Steve Gschmeissner/Science Photo Library

US lifts Trump-era limits on studies of fetal tissues

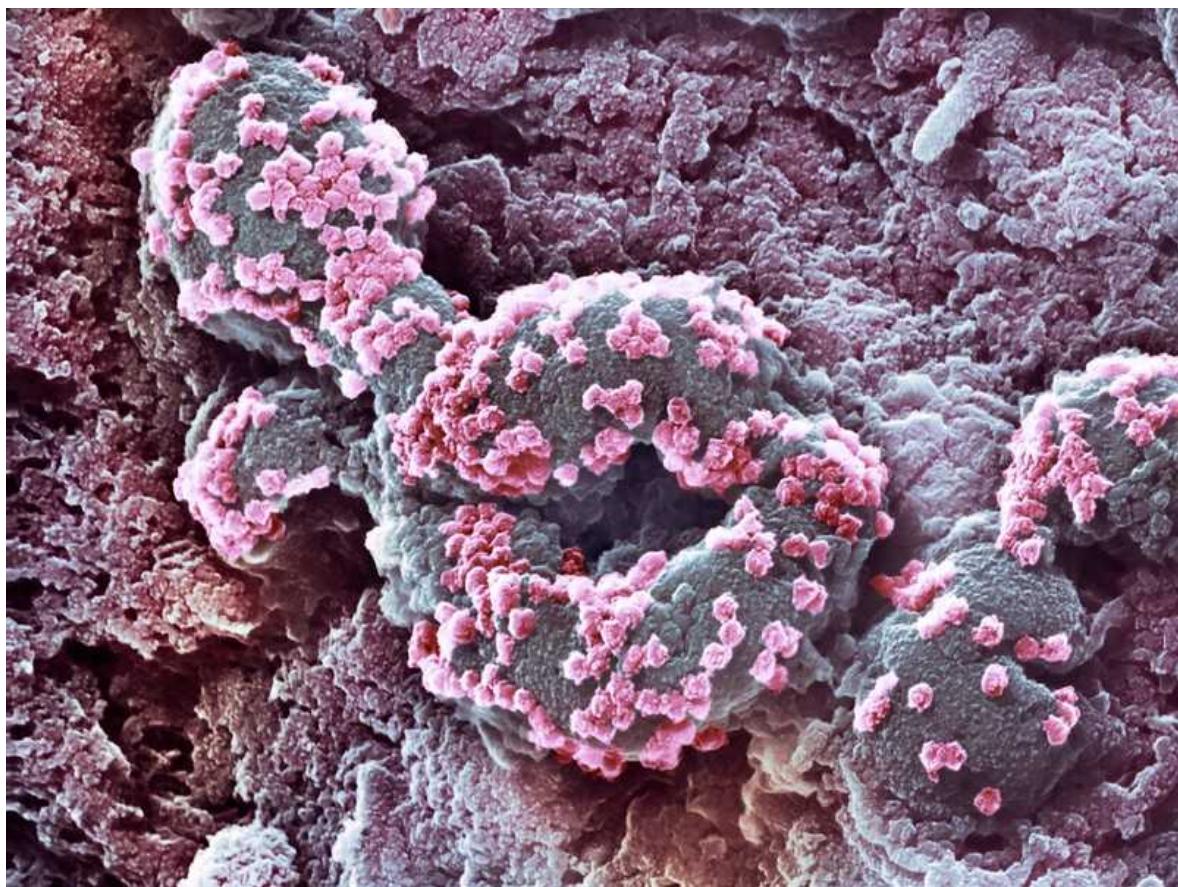
The United States is [reversing restrictions on fetal-tissue research](#) set by former president Donald Trump's administration. Government scientists will be able to resume studies using the biological material, and an extra ethics review of grant proposals submitted by academic researchers has been cancelled.

“That’s good news,” says Lawrence Goldstein, a neuroscientist at the University of California, San Diego, who was a member of the Trump-era board that carried out the ethics reviews. Researchers use fetal tissue

(pictured), obtained from elective abortions, to study a range of subjects, from infectious diseases to human development. The US National Institutes of Health (NIH) announced the reversal on 16 April.

The Trump administration announced limits on fetal-tissue studies in 2019, following pressure from anti-abortion groups.

Kristin Matthews, a science-policy scholar at Rice University's Baker Institute for Public Policy in Houston, Texas, is pleased with the reversal. "There's not a lot of groups that do fetal-tissue research, but what they do is pretty vital," she says, "and so it'll be good to be able to continue that research."



A cell infected with SARS-CoV-2 particles (pink; artificially coloured). Credit: Science Photo Library

Previous infection could cut COVID duration

Recent infection by viruses related to SARS-CoV-2 could reduce the duration of COVID-19 symptoms, according to an analysis of antibodies from 2,000 health-care workers.

Antibodies against the SARS-CoV-2 spike protein can be a powerful defence against COVID-19. But some people have rare antibodies against other coronaviruses that pre-date the pandemic and can bind to other SARS-CoV-2 proteins. To search for a possible link between such antibodies and protection from COVID-19, Scott Hensley at the University of Pennsylvania in Philadelphia and his colleagues studied antibody levels and infection status in about 2,000 local volunteers during two COVID-19 surges ([S. Gouma *et al.* Preprint at medRxiv https://doi.org/f7zp; 2021](#)).

The team found that people with the rare, pre-pandemic antibodies that work against SARS-CoV-2 were not protected from contracting the virus and developing COVID-19. But high concentrations of other antibodies elicited by two betacoronaviruses — a category that also includes SARS-CoV-2 — were associated with a quicker recovery from COVID-19.

The authors speculate that this protection is provided by immune-system players called T cells that were generated in response to previous betacoronavirus infection.

Nature **592**, 665 (2021)

doi: <https://doi.org/10.1038/d41586-021-01085-w>

Jobs from Nature Careers

- - - [All jobs](#)
 -
 - [Research Associate / Postdoc](#)

[Technische Universität Dresden \(TU Dresden\)](#)

Dresden, Germany

JOB POST

- **PhD position in evolutionary ecology of insect immune priming**

Université Bourgogne Franche-Comté (UBFC)

DIJON, France

JOB POST

- **PhD contract Doctoral contract in sociology-demography : Analysis of differential socio-demographic behaviors by sex among 50-74 year olds**

Université Bourgogne Franche-Comté (UBFC)

DIJON, France

JOB POST

- **PhD - Dark ecological networks and land management: methodological and strategic contributions for reducing ecological impacts of artificial light**

Université Bourgogne Franche-Comté (UBFC)

BESANCON, France

JOB POST

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

NEWS

21 April 2021

India's massive COVID surge puzzles scientists

The virus is spreading faster than ever before in India despite previous high infection rates in megacities, which should have conferred some protection.

- [Smriti Mallapaty](#)

1. Smriti Mallapaty

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)



Ahead of a lockdown imposed on 14 April, migrant workers queue at a railway station to depart the city of Mumbai, India. Credit: Getty

The pandemic is sweeping through India at a pace that has staggered scientists. Daily case numbers have exploded since early March: the government reported 273,810 new infections nationally on 18 April. High numbers in India have also helped drive global cases to a daily high of 854,855 in the past week, almost breaking a record set in January.

Just months earlier, antibody data had suggested that many people in cities such as Delhi and Chennai had already been infected, leading some researchers to conclude that the worst of the pandemic [was over in the country.](#)

Researchers in India are now trying to pinpoint what is behind the unprecedented surge, which could be due to an unfortunate confluence of factors, including the emergence of particularly infectious variants, a rise in unrestricted social interactions, and low vaccine coverage. Untangling the causes could be helpful to governments trying to suppress or prevent similar surges around the world.

European countries such as France and Germany are also currently experiencing large outbreaks relative to their size, and nations including Brazil and the United States are reporting high infection rates at around 70,000 a day. But India's daily totals are now some of the highest ever recorded for any country, and are not far off a peak of 300,000 cases seen in the United States on 2 January.

‘Ripple in a bathtub’

COVID-19 case numbers started to drop in India last September, after a high of around 100,000 daily infections. But they began to rise again in March and the current peak is more than double the previous one (see ‘Surging cases of COVID-19’).

“The second wave has made the last one look like a ripple in a bathtub,” says Zarir Udwadia, a clinician-researcher in pulmonary medicine at P D Hinduja Hospital & Medical Research Centre in Mumbai, who spoke to

Nature during a break from working in the intensive-care unit. He describes a “nightmarish” situation at hospitals, where beds and treatments are in extremely short supply.

Shahid Jameel, a virologist at Ashoka University in Sonipat, agrees that the intensity of the current wave is startling. “I was expecting fresh waves of infection, but I would not have dreamt that it would be this strong,” he says.



India's COVID-vaccine woes — by the numbers

Studies that tested for SARS-CoV-2 antibodies — an indicator of past infection — in December and January estimated that more than 50% of the population in some areas of India’s large cities had already been exposed to the virus, which should have conferred some immunity, says Manoj Murhekar, an epidemiologist at the National Institute of Epidemiology in Chennai, who led the work. The studies also suggested that, nationally, some 271 million people had been infected¹ — about one-fifth of India’s population of 1.4 billion.

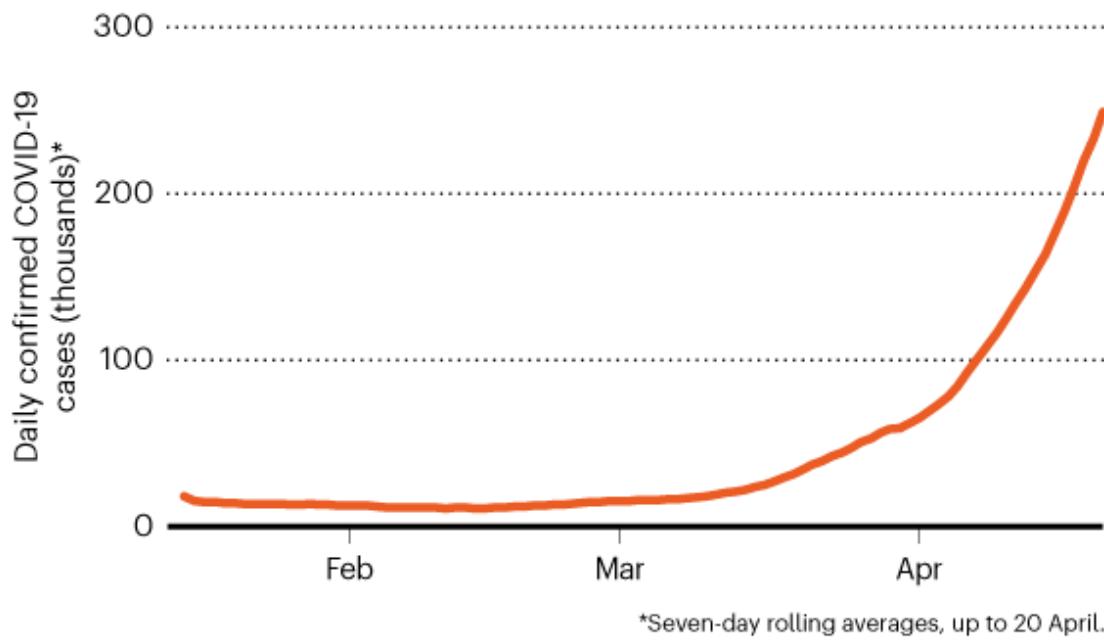
These figures made some researchers optimistic that the next stage of the pandemic would be less severe, says Ramanan Laxminarayan, an epidemiologist in Princeton University, New Jersey, who is based in New Delhi. But the latest eruption of COVID-19 is forcing them to rethink.

One explanation might be that the first wave primarily hit the urban poor. Antibody studies might not have been representative of the entire population and potentially overestimated exposure in other groups, he says.

The antibody data did not reflect the uneven spread of the virus, agrees Gagandeep Kang, a virologist at the Christian Medical College in Vellore, India. “The virus may be getting into populations that were previously able to protect themselves,” she says. That could include wealthier urban communities, in which people isolated during the first wave but had started mingling by the second.

SURGING CASES OF COVID-19

New daily cases in India have been rapidly on the rise since March and have now far surpassed a peak of around 100,000 new cases per day last September.



©nature

Source: Our World in Data

Fast-moving variants?

But some researchers say that the speed and scale of the current outbreak suggest a new ingredient: emerging variants of the virus.

Udwadia has anecdotally observed that entire households are now getting infected — unlike in the first wave of COVID-19, when single individuals would test positive. He attributes this to the presence of more-infectious variants. “If one person in the family has it, I can guarantee that everyone in the family has it,” he says.

Genomic surveillance data show that the variant B.1.1.7, which was first identified in the United Kingdom, has become the dominant form of the virus in the Indian state of Punjab.



[Has COVID peaked? Maybe, but it's too soon to be sure](#)

And a new and potentially concerning variant first identified in India late last year, known as B.1.617, has become dominant in the state of Maharashtra. B.1.617 has drawn attention because it contains two mutations that have been linked to increased transmissibility and an ability to evade immune protection. It has now been detected in 20 other countries. Laboratories in India are trying to culture it to test how fast it replicates, and whether blood from vaccinated individuals can block infection, says Jameel.

The situation in India looks similar to that late last year in Brazil, he adds, where a resurgence of COVID-19 in the city of Manaus coincided with the spread of a highly transmissible variant known as P.1, which might have been able to evade immunity conferred by infections with earlier strains.

But others say that the existing sequencing data are not sufficient to make such claims. “As the numbers of sequences available are low, relative to the number of cases in India, we do need to be cautious,” says David Robertson, a virologist at the University of Glasgow, UK.

Mixing, moving and travelling

Some say that emerging variants account for only a small part of India’s surge in infections. In many regions that are experiencing outbreaks, they don’t make up the majority of genomes sequenced, says Anurag Agrawal, director of the CSIR Institute of Genomics and Integrative Biology in New Delhi.

Srinath Reddy, an epidemiologist and head of the Public Health Foundation of India in New Delhi, argues that people letting their guards down is a bigger driver. “The pandemic resurfaced in a fully open society where people were mixing and moving and travelling,” he says.



[India will supply coronavirus vaccines to the world — will its people benefit?](#)

With cases declining after last September's peak, "there was a public narrative that India had conquered COVID-19", says Laxminarayan. In recent months, large crowds have gathered indoors and outdoors for political rallies, religious celebrations and weddings.

The nationwide [vaccination campaign](#), which kicked off in January, might even have contributed to an uptick in cases, if it caused people to ease public-health measures. "The arrival of the vaccine put everyone into a relaxed mood," says Laxminarayan.

More than 120 million doses have been administered, mostly of an Indian-produced version of the Oxford–AstraZeneca vaccine called Covishield. But that's less than 10% of India's population, so there is still a long way to go. In particular, India needs to ramp up vaccinations in the hardest-hit regions, says Kang.

Some people might have become infected while getting vaccines, says Udwadia, because crowds often share clinic waiting areas with ill people who are waiting to be seen.

Nature **592**, 667-668 (2021)

doi: <https://doi.org/10.1038/d41586-021-01059-y>

References

1. 1.

Murhekar, M. V. *et al.* Preprint at SSRN
<https://dx.doi.org/10.2139/ssrn.3797589> (2021).

[Jobs from Nature Careers](#)

- - - [All jobs](#)
 -

- **Research Associate / Postdoc**

Technische Universität Dresden (TU Dresden)

Dresden, Germany

JOB POST

- **PhD position in evolutionary ecology of insect immune priming**

Université Bourgogne Franche-Comté (UBFC)

DIJON, France

JOB POST

- **PhD contract Doctoral contract in sociology-demography : Analysis of differential socio-demographic behaviors by sex among 50-74 year olds**

Université Bourgogne Franche-Comté (UBFC)

DIJON, France

JOB POST

- **PhD - Dark ecological networks and land management: methodological and strategic contributions for reducing ecological impacts of artificial light**

Université Bourgogne Franche-Comté (UBFC)

BESANCON, France

JOB POST

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01059-y>.

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

NEWS

19 April 2021

- Update [19 April 2021](#)

Lift off! First flight on Mars launches new way to explore worlds

NASA's Ingenuity helicopter successfully hovered for 40 seconds in Mars's thin atmosphere.

- [Alexandra Witze](#)
 1. Alexandra Witze
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)



NASA's Mars helicopter Ingenuity took this shot of the Martian surface during its first flight, on 19 April. Credit: NASA/JPL-Caltech

NASA has pulled off the first powered flight on another world. Ingenuity, the robot rotorcraft that is part of the agency's Perseverance mission, lifted off from the surface of Mars on 19 April, in a 39.1-second flight that is a landmark in interplanetary aviation.

"We can now say that human beings have flown a rotorcraft on another planet," says MiMi Aung, the project's lead engineer at the Jet Propulsion Laboratory (JPL) in Pasadena, California.

Ingenuity's short test flight is the off-Earth equivalent of the Wright brothers piloting their aeroplane above the coastal dunes at Kitty Hawk, North Carolina, in 1903. In tribute, the helicopter carries a postage-stamp-sized piece of muslin fabric from the Wright brothers' plane. "Each world gets only one first flight," says Aung.

The flight came after a one-week delay, because software issues kept the helicopter from transitioning into flight mode two days ahead of a planned flight attempt on 11 April. Today, at 12.34 a.m. US Pacific time, Ingenuity successfully spun its counter-rotating carbon-fibre blades at more than 2,400 revolutions per minute to give it the lift it needed to rise 3 metres into the air. The US\$85-million drone hovered there, and then, in a planned manoeuvre, turned 96 degrees and descended safely back to the Martian surface. “This is just the first great flight,” says Aung.

This video, captured by the Perseverance rover, shows Ingenuity’s 40-second flight. Credit: NASA/JPL-Caltech/ASU/MSSS

Four further flights, each lasting for up to 90 seconds, are planned in the next 2 weeks. The next one is tentatively scheduled for 22 April. In it, Ingenuity will aim to rise 5 metres above the surface, fly laterally for about 2 metres, then fly 2 metres back and land at the same place it took off. Eventually, the helicopter may fly faster and farther, travelling up to 300 metres from its take-off point. Each successive flight will push Ingenuity’s capabilities to see how well the drone fares in Mars’s thin atmosphere, which is just 1% as dense as Earth’s. “We will be pushing the envelope,” says Aung — probably to the point that Ingenuity will ultimately crash, by design.

Space agencies have sent drifting aircraft to other planets before; for example, the Soviet Union’s Vega 1 and Vega 2 missions sent balloons into Venus’s atmosphere in 1985. But Ingenuity’s flight is the first controlled flight on another planet. “I’m just thrilled with the way it has turned out,” says John Grunsfeld, a former astronaut who approved the Ingenuity programme when he served as NASA’s associate administrator for science.

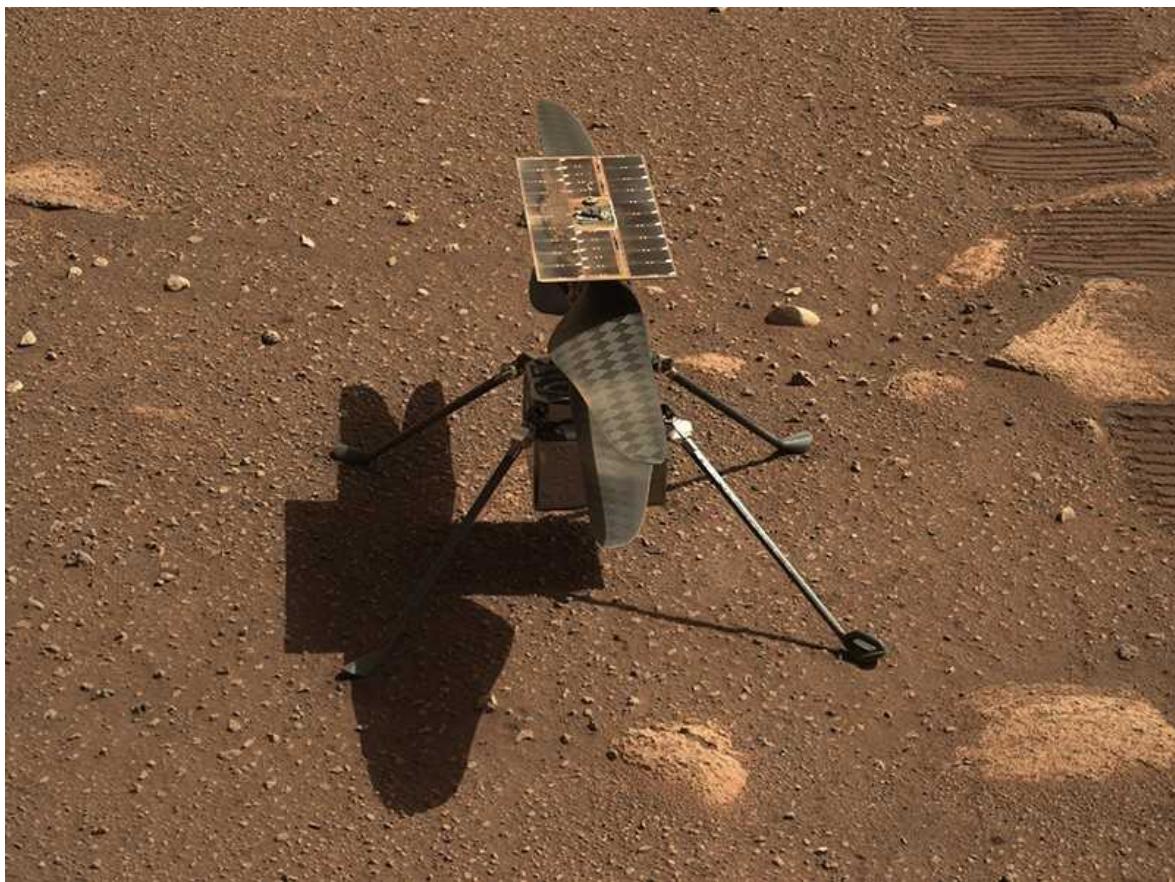


The Perseverance rover used one of its cameras to take a selfie with the Ingenuity helicopter (behind and to the left) on 6 April. Credit: NASA/JPL-Caltech/MSSS

Ingenuity's purpose is to test whether helicopters could be used to explore other worlds. As it flies across the terrain, the craft snaps black-and-white images of the surface below, and colour images looking towards the horizon. Future helicopters could help rovers, or even astronauts, to make their way across the surface, by scouting for interesting areas ahead and relaying images of what the landscape looks like.

Big rotorcraft could also get into areas that are inaccessible to rovers rolling across the ground, says Anubhav Datta, an aerospace engineer at the University of Maryland in College Park who has been working on Mars helicopter concepts for decades. "If we are serious about human missions to Mars, we should be serious about sending large helicopters to truly explore what awaits there," he says. "The most interesting places we want to explore are not on flat land but up the slopes, on the cliffs, down the craters and into the caves." Cameras and other instruments aboard helicopters could capture information about such places.

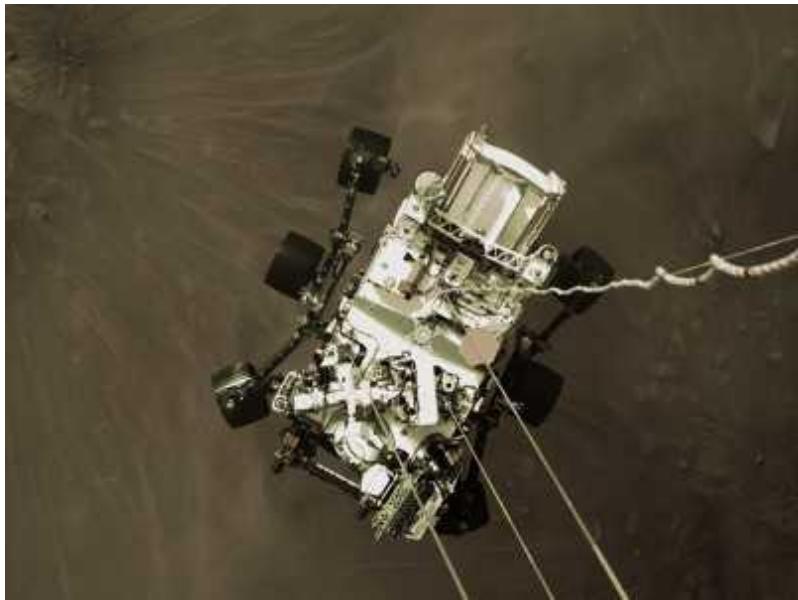
NASA is [already building a car-sized octocopter](#) named Dragonfly that it plans to send to Saturn's moon Titan. Set to launch in 2027, the copter would explore Titan's atmosphere, which is four times denser than Earth's and is rich in primordial organic compounds. That's a very different environment from the one that Ingenuity is experiencing on Mars. But the early flight lessons from Ingenuity will inform Dragonfly's design. "We're looking forward to learning from the Ingenuity team's experience flying in an extraterrestrial sky," says Elizabeth Turtle, a planetary scientist at the Johns Hopkins University Applied Physics Laboratory in Laurel, Maryland, who is Dragonfly's principal investigator.



A close-up of Ingenuity sitting on the Martian surface, ready for its first flight test. Credit: NASA/JPL-Caltech/ASU

Ingenuity [arrived in Mars's Jezero Crater in February](#), nestled under the belly of the Perseverance rover. From its landing site, Perseverance drove to a flat 'airfield' in the crater that is relatively free of rocks, and deposited Ingenuity there. The rover then rolled to a slight rise 65 metres away, a

vantage point from which it watched and videotaped Ingenuity's first take-off and flight.



[Mars video reveals Perseverance rover's daring touchdown](#)

The biggest challenge in designing Ingenuity was making it small and light enough to be carried under Perseverance's belly, while still being capable of flight, says Aung. The helicopter ended up weighing just 1.8 kilograms. Engineers tested it on Earth in a special chamber at JPL from which nearly all the air had been sucked out, to simulate the thin Martian atmosphere.

Compared with a similar-sized helicopter on Earth, Ingenuity has larger blades that spin much faster, to lift it into the thin Martian air. Datta says that he will be anxiously awaiting information on how much power the helicopter takes to hover; this knowledge will help engineers to better understand the aerodynamics on Mars.



The Perseverance rover dropped Ingenuity (shown here with its legs deployed) from its belly on 30 March. Credit: NASA/JPL-Caltech

Another researcher, William Farrell at NASA's Goddard Space Flight Center in Greenbelt, Maryland, is crossing his fingers that Ingenuity will help scientists to gain a better idea of the electrical properties of the Martian atmosphere. To do this, it would need to fly — or at least spin its blades — near dusk on Mars. Farrell and his colleagues have calculated that the moving helicopter blades could become electrically charged through contact with the dust in the surrounding air¹, much as helicopter blades on Earth can build up charge in sand storms. That could cause a faint blue-purplish glow along the blades, best visible in the dim light of dusk. Farrell has asked the Ingenuity team if it could rotate the blades during dusk at some point — and if that happens, he will be watching closely.



A month on Mars: what NASA's Perseverance rover has found so far

The thin Martian atmosphere means that winds there are not particularly strong. Ingenuity can handle winds of a little over 10 metres per second while flying, and stronger winds when it's sitting on the ground. It is powered by solar panels to keep it warm during the freezing Martian nights, when temperatures can sink to -90°C at Jezero Crater.

Ingenuity is designed to last for just 30 Martian days, which end on 4 May. After that, team scientists will turn their attention back to the rover on which it travelled to Mars. Ingenuity will rest in perpetuity in Jezero Crater as Perseverance trundles off on its main mission to collect rock samples for eventual return to Earth.

Nature **592**, 668-669 (2021)

doi: <https://doi.org/10.1038/d41586-021-00909-z>

Updates & Corrections

- **Update 19 April 2021:** This story was updated on 19 April to include new information about the next flights Ingenuity will take.

References

1. 1.

Farrell, W. M., McLain, J. L., Marshall, J. R. & Wang, A. *Planet. Sci. J.* **2**, 46 (2021).

Jobs from Nature Careers

- - - [All jobs](#)
 - - [Research Associate / Postdoc](#)
 - [Technische Universität Dresden \(TU Dresden\)](#)
[Dresden, Germany](#)
[JOB POST](#)
 - [PhD position in evolutionary ecology of insect immune priming](#)
[Université Bourgogne Franche-Comté \(UBFC\)](#)
[DIJON, France](#)
[JOB POST](#)
 - [PhD contract Doctoral contract in sociology-demography : Analysis of differential socio-demographic behaviors by sex among 50-74 year olds](#)

Université Bourgogne Franche-Comté (UBFC)

DIJON, France

JOB POST

- **PhD - Dark ecological networks and land management: methodological and strategic contributions for reducing ecological impacts of artificial light**

Université Bourgogne Franche-Comté (UBFC)

BESANCON, France

JOB POST

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-00909-z>

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

NEWS EXPLAINER

21 April 2021

COVID vaccines and kids: five questions as trials begin

As the first clinical trials in young children start, here's what scientists want to know.

- [Ewen Callaway](#)

1. Ewen Callaway

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)



Students at the Petri primary school in Dortmund, Germany.Credit: Ina Fassbender/AFP via Getty

Parents are clamouring to enrol their kids in the first COVID-19 vaccine trials in young children. “Somebody told me they called and called and called until they were allowed to participate,” says Kawsar Talaat, an infectious-disease physician and vaccine scientist at the Johns Hopkins Bloomberg School of Public Health in Baltimore, Maryland. She is part of a trial that began testing the Pfizer–BioNTech vaccine in children under 12 in late March.

As such trials get under way — Moderna began a similar study of its vaccine last month — scientists are seeking answers to important questions about how safe and effective the vaccines are in kids.

Nature looks at how the trials will account for differences in children’s immune systems and susceptibility to COVID-19, compared with those of adults, as well as the added safety precautions that surround medical research in kids. “Children are not little adults,” stresses Talaat.

Do we even need to vaccinate children?

Children rarely develop severe forms of COVID-19, and deaths from the disease are rarer still. But on rare occasions — one estimate¹ puts it at around one case in 1,000, although it could be even lower — kids who’ve experienced even mild infections can later develop a sometimes deadly condition called multi-system inflammatory syndrome in children (MIS-C). “I’m tired of seeing sick kids. I want to see them protected,” says James Conway, a paediatric infectious-disease specialist and vaccine researcher at the University of Wisconsin–Madison.

Evidence is building that [vaccines might block transmission of SARS-CoV-2](#), so vaccinating children could have beneficial knock-on effects in the wider community. “If we really want to get back to normalcy, [we really need to achieve herd immunity](#) across all the groups that potentially contribute to transmission,” Conway adds.

Children, particularly younger kids, probably aren’t super-spreaders of SARS-CoV-2, as they are for viruses including influenza. But [the emergence of faster-spreading variants](#), along with rising adult vaccination rates in some

countries, means that children and adolescents might soon be contributing more to spread. “COVID transmission is now hottest in younger people. The virus will find ways to survive and spread unless we close off the pathways,” says Talaat.

How will the trials in kids work?

In some ways, vaccine trials in children under 12 years old will be a replay of early trials in adults. The first recipients — who will be on the older end of the spectrum, although trials will eventually include children as young as six months — will receive a range of doses to find one that triggers a strong immune response without too many side effects. “Some of them are too much, some of them are too little. You’re looking for that sweet spot,” says Conway. “It’s the Goldilocks effect.”

Once an ideal dose is identified, several thousand participants will be randomized to receive either two doses of vaccine or of a placebo injection. Researchers will then follow the children for months and even years, to study the safety and effectiveness of the vaccines.

In adult trials, participating individuals provide informed consent. When participants are children, however, their legal guardian must agree to their involvement. But researchers are also obliged to obtain assent from any child participants old enough to understand the trial, says Beate Kampmann, a paediatric infectious-disease specialist and director of the Vaccine Centre at the London School of Hygiene & Tropical Medicine. “Our kids are savvy, they understand. They’ve heard about this all year,” says Talaat, who generally seeks assent from kids aged five and over — and sometimes younger, depending on the child.

Will children and adults respond differently to COVID-19 vaccines?

Children’s immune systems are brimming with cells that haven’t seen pathogens, so they tend to produce a strong immune response to vaccines, says Donna Farber, an immunologist at Columbia University in New York

City. “It’s those first few years of life, where you’re learning about pathogens.”

Early trial results have shown that 12–15-year-olds who received two standard doses of the Pfizer–BioNTech vaccine developed substantially higher levels of virus-blocking antibodies than did 16–25-year-olds in earlier trials. Farber wonders whether children who are even younger will get the same immune response from a lower dose.

Children’s potent immune responses mean that they are more likely than adults to develop a fever after vaccination, says Talaat, so researchers will need to strike a balance between triggering a strong immune response and minimizing the side effects that come with it. However, this might not be such a problem, because children seem to be less bothered by fevers than are adults, says Farber.

As COVID-19 vaccines are tested in ever-younger children, researchers will want to make sure they’re not interfering with immunity generated by routine childhood vaccinations, says Kampmann. The Pfizer–BioNTech trial plans eventually to enrol children under five who might still be due to receive boosters of polio vaccines and jabs against measles, mumps and rubella, as well as other immunizations, but Talaat says children will need to be up to date on their vaccination schedules to participate. Studies of how a COVID-19 vaccine should best be integrated into a child’s immunization schedule will need to come later, she adds.

How will scientists know if vaccines work in children?

We know that vaccines prevent COVID-19 in adults because the clinical trials were designed to show this. They involved tens of thousands of people randomly assigned to receive either the vaccine or a placebo, and showed compelling differences in the rates of disease between the two groups.

In the paediatric trials, which will involve only a few thousand children, there might be too few symptomatic infections to measure efficacy in the same way, says Talaat. It makes more sense, she says, to look at immune

markers after vaccination. “If we see the paediatric immune responses are the same or better than we saw in adults, we can make inferences that the vaccine will be effective.” Both the Moderna and the Pfizer–BioNTech trials list such markers as their primary measures of success.

Conway would like to see good evidence that the vaccines can actually prevent COVID-19 in kids. The Pfizer–BioNTech trial in adolescents recorded 18 cases in the placebo group, and none in those who got the vaccine, so it’s not inconceivable that trials in younger kids will also show such efficacy, says Talaat, but it will depend on community infection rates.

However, if the primary aim of childhood vaccination is to stop transmission, trials should show this, says Christiane Eberhardt, a physician-scientist in clinical vaccinology at the Geneva University Hospitals in Switzerland. This would ideally involve frequent swabs in kids — unlikely to be popular — and in unvaccinated family members. Instead, the Moderna and Pfizer–BioNTech trials intend to look at blood markers of asymptomatic infection, which Eberhardt sees as acceptable under the circumstances. “That’s the closest you can get.”

How will researchers know if the vaccines are safe in young children?

Safety is paramount in clinical trials involving children, and researchers are aware that COVID-19 vaccine trials in kids will get extra scrutiny. “Anything that smears vaccines in general, and makes people question the safety of vaccines in kids, is a step backwards from a public-health standpoint,” says Conway.

Early adult trials paid close attention to the possibility that people who received the vaccine could develop ‘enhanced disease’ if they later became infected. The trials found no evidence for this, but Conway says that paediatric trials should look for immune responses that might exacerbate disease, as well as signs that participants are developing immune reactions similar to those seen in MIS-C.

It is not yet clear how [concerns over very rare blood clots](#) potentially linked to the Oxford–AstraZeneca and Johnson & Johnson vaccines will affect paediatric trials. The University of Oxford, UK, has paused a small trial in kids aged 6–17 that began in February. Johnson & Johnson announced at the start of April that it was set to begin including adolescents in an ongoing trial of its vaccine, but has since paused all its trials as the clots are investigated.

Nature **592**, 670-671 (2021)

doi: <https://doi.org/10.1038/d41586-021-01061-4>

References

1. 1.

Hoang, A. *et al.* *EClinicalMedicine* **24**, 100433 (2020).

Jobs from Nature Careers

- - - [All jobs](#)
 - - [Research Associate / Postdoc](#)
 - [Technische Universität Dresden \(TU Dresden\)](#)
 - [Dresden, Germany](#)
 - [JOB POST](#)
 - [PhD position in evolutionary ecology of insect immune priming](#)
 - [Université Bourgogne Franche-Comté \(UBFC\)](#)

DIJON, France

JOB POST

- **PhD contract Doctoral contract in sociology-demography : Analysis of differential socio-demographic behaviors by sex among 50-74 year olds**

Université Bourgogne Franche-Comté (UBFC)

DIJON, France

JOB POST

- **PhD - Dark ecological networks and land management: methodological and strategic contributions for reducing ecological impacts of artificial light**

Université Bourgogne Franche-Comté (UBFC)

BESANCON, France

JOB POST

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01061-4>

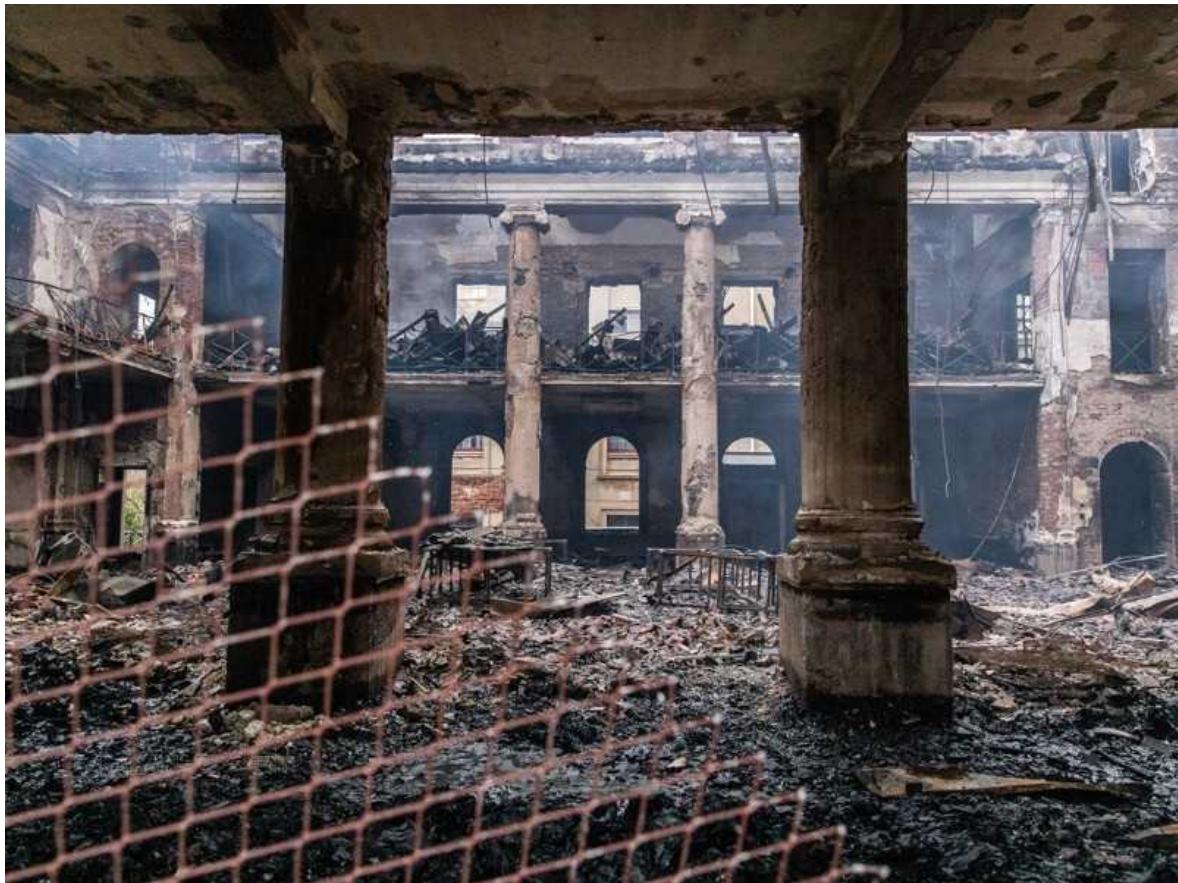
NEWS
19 April 2021

‘The damage is total’: fire rips through historic South African library and plant collection

University of Cape Town faces losing ‘irreplaceable’ historical material on anthropology, ecology and politics.

- [Linda Nordling](#)
 1. Linda Nordling
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)



The gutted interior of the University of Cape Town library's Jagger Reading Room, which was built in the 1930s. Credit: Ashraf Hendricks/GroundUp

Forest fires raging in South Africa's Table Mountain National Park have reached the University of Cape Town (UCT), and gutted the reading room of its main library, which houses irreplaceable documents and records from the country's past.

Amid apocalyptic scenes on 18 April, fire tore through part of the 200-year-old university's campus on the slopes of Table Mountain, torching the library building and destroying a plant research unit.

[Researchers have set up an online page](#) asking for anyone with photos or digital scans of the library's collections to upload them.

“This archive is special for all sorts of reasons, and for me it’s because it includes collections which provide a record of the ordinary lives of ordinary people in the area — from working-class children to Black students

attending night school,” says Sarah Emily Duff, a historian of South Africa based at Colby College in Waterville, Maine. “We lose that texture of everyday life and struggles with a catastrophe like this,” she adds.

Table Mountain, on the continent’s southwestern tip, often has fires in summer and autumn, but the university has been spared catastrophic damage in previous wildfires, says Howard Phillips, a UCT historian, now emeritus. “I am not aware of any other natural fire disaster having hit UCT so devastatingly,” he told *Nature*.

Unseasonable weather

The immediate cause of the wildfires is unknown. One theory suggests that an abandoned cooking fire on the mountain took hold, fuelled by unseasonably hot, dry weather. Cape Town police also confirmed the arrest of a suspected arsonist later that day, although it is not clear whether the individual is thought to have started the main fire, or a smaller blaze on the mountain.

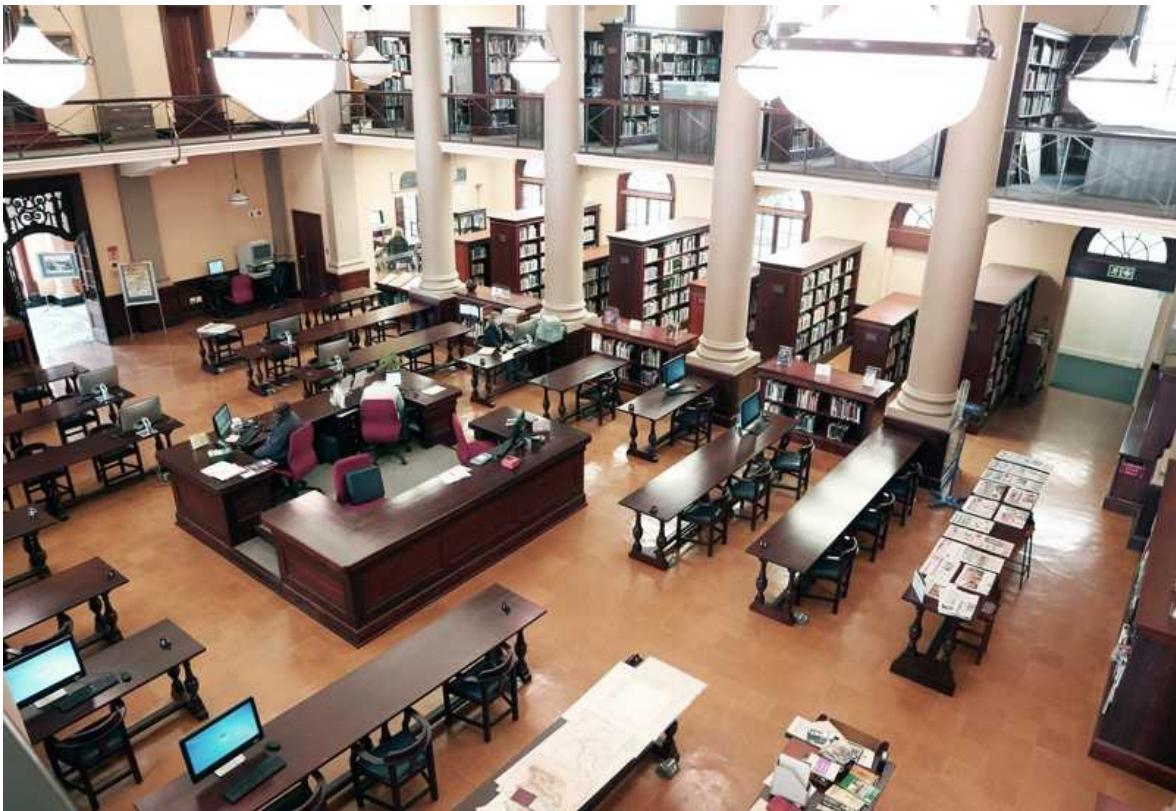
“There are several factors which drove the fire — wind, dry vegetation and extremely hot weather,” said Jermaine Carelse, a spokesperson for Cape Town’s fire services.

Around 150 firefighters arrived at the university and surrounding areas on the morning of the fire and fought the blaze until late evening. Overnight, the winds blew the fires away from the campus, but around 500 firefighters, including volunteers, were still battling the blaze the following day, when it was threatening residential areas in the city centre.

As thick smoke engulfed the campus, hundreds of students were evacuated from their residences. Some were put up in hotels that are largely empty because of the coronavirus pandemic.



Around 150 firefighters battled the blaze all day on Sunday, 18 April.Credit: Nic Bothma/EPA-EFE/Shutterstock



The Jagger Reading Room, photographed in 2016 after a five-year restoration project. Credit: Theresa Schoeman/UCT Libraries

Total destruction

The library complex houses UCT's special collections, which hold irreplaceable artefacts, including annotated watercolours of plants and animals dated from 1881, painted by Indigenous inhabitants of the Cape. It also holds maps, manuscripts and government records from the Cape's mottled past — including its colonial and military history.

Automated shutters that had been installed to protect the library's rarest items were activated by the fire, but it is not yet known whether they succeeded in stopping its spread from the reading room to the special collections nearby.

"Some of our valuable collections have been lost," UCT libraries director Ujala Satgoor said in a statement. "However a full assessment can only be

done once the building has been declared safe and we can enter the building.”

The university’s botany building was also seriously damaged. In the Plant Conservation Unit, where researchers track changes in climate by studying fossilized pollen and comparing historical photos with present-day images, the “damage is total”, says the unit’s leader Timm Hoffman, a historical ecologist.

“We are locked out of the campus, but we’ve seen photographs. There’s nothing left; the entire unit has been destroyed,” Hoffman says. Most of the historical photos and slides lost in the fire will have been digitized, but some were still undergoing that process.

The unit’s fossil-pollen laboratory might be completely destroyed, says UCT palaeoecologist Lindsey Gillson. But the sediment cores from which the lost samples were taken were housed in a different, unharmed, part of the botany building.

The herbarium, which houses specimens from many of the Cape’s unique plants, is understood to have been spared — although it might have sustained water damage.

“We’ve lost computers and microscopes and the samples in the lab. Those are replaceable. I’m just happy nobody was in the lab,” Gillson says. “We’ve lost our home, but we haven’t lost our community.”

Nature **592**, 672 (2021)

doi: <https://doi.org/10.1038/d41586-021-01045-4>

Jobs from Nature Careers

- - - [All jobs](#)
 -

- **Research Associate / Postdoc**

Technische Universität Dresden (TU Dresden)

Dresden, Germany

JOB POST

- **PhD position in evolutionary ecology of insect immune priming**

Université Bourgogne Franche-Comté (UBFC)

DIJON, France

JOB POST

- **PhD contract Doctoral contract in sociology-demography : Analysis of differential socio-demographic behaviors by sex among 50-74 year olds**

Université Bourgogne Franche-Comté (UBFC)

DIJON, France

JOB POST

- **PhD - Dark ecological networks and land management: methodological and strategic contributions for reducing ecological impacts of artificial light**

Université Bourgogne Franche-Comté (UBFC)

BESANCON, France

JOB POST

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01045-4>

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

NEWS

22 April 2021

US pledges to dramatically slash greenhouse emissions over next decade

Researchers say the commitment to roughly halve emissions is a good start to slowing climate change, but fear it still isn't enough.

- [Jeff Tollefson](#)

1. Jeff Tollefson

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)



Joe Biden addresses the audience during the US climate summit on 22 April.Credit: Al Drago/Getty

US President Joe Biden opened a virtual White House climate summit on 22 April by announcing an ambitious new commitment to curb US greenhouse-gas emissions to at least 50% below 2005 levels by the end of this decade. The pledge is the most aggressive made by the country yet, extending and strengthening a commitment made under former president Barack Obama to cut emissions by at least 26% below 2005 levels by 2025.



Can Joe Biden make good on his revolutionary climate agenda?

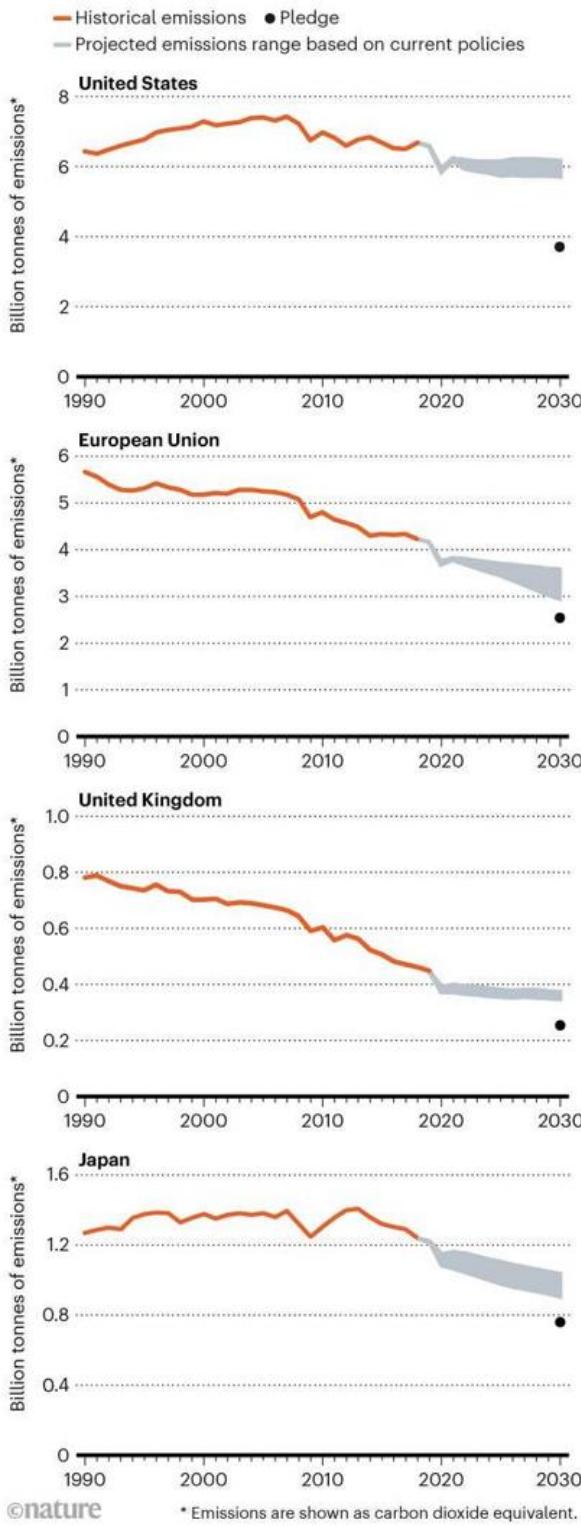
The announcement marks a radical shift from the past four years under former president Donald Trump, who [pulled the United States out of the 2015 Paris climate accord](#) and sought to protect the nation's struggling coal industry.

Calling global warming "the existential crisis of our times", Biden said the United States is resolved to act, and urged fellow world leaders to join the battle. "No nation can solve this crisis on our own," Biden said. All countries, he added, "have to step up".

Forty world leaders are attending the US summit, held over two days, with the goal of building momentum for international action on climate.

NEW COMMITMENTS

The United States held a climate summit this week at which it pledged to cut its greenhouse-gas emissions by at least 50% below 2005 levels by 2030. The European Union and various countries have also made new climate commitments, which will require significant changes in economic and energy policies.



Source: Climate Action Tracker

The new US target is roughly in line with recent commitments by the European Union and a growing cadre of nations that have pledged to halve their emissions by 2030 as a crucial first step on the path to full decarbonization by mid-century. It is also in line with calls from scientists and environmentalists who lobbied the administration for an aggressive commitment during the run-up to this week's summit.



Will the United States make its most dramatic climate pledge yet?

“This level of ambition is what’s necessary to inspire other big emitters to go further and curb their own emissions by 2030,” says Marcene Mitchell, senior vice-president for climate change at the wildlife charity WWF in Washington DC. The next step, says Mitchell, is for Biden to mobilize society and make the commitment a reality.

Although his climate team is still working out the details, Biden has outlined how he will do this in broad strokes. In addition to decarbonizing the electricity sector by 2035 and ramping up the use of electric vehicles to cut transportation emissions, his plan seeks to harness new technologies to curb industrial emissions while protecting forests and natural ecosystems that pull carbon out of the atmosphere. As a down payment, Biden last month put forward a massive US\$2.3-trillion package that would invest heavily in clean-energy technologies, infrastructure and jobs.

Falling short

At the Paris climate meeting in 2015, governments committed to limit global warming to 1.5–2 °C above pre-industrial levels, in an effort to prevent a cascade of potential catastrophic impacts, including increases in extreme weather, the destruction of natural ecosystems and damages to agricultural systems.

Although many governments have begun to bend the emissions curve downwards by adopting climate policies that advance clean-energy development, international commitments have fallen well short of that 2015 goal: on the basis of current policies, the world is currently on track for around a 3 °C rise, according to Climate Action Tracker, an international consortium of scientists and policy specialists that monitors efforts to implement the Paris accord.



[COVID curbed carbon emissions in 2020 — but not by much](#)

The White House climate summit is the first stop this year on the road to a United Nations meeting in November in Glasgow, UK, where governments are expected to formalize their new climate commitments. The world's largest emitter, China, pledged last September to halt its growth in carbon emissions by 2030 and achieve carbon-neutrality by 2060. Last December, the European Union announced its goal of curbing emissions to 55% below 1990 levels by 2030.

Climate observers say the US summit has helped to inject new momentum into international negotiations: Japan and the United Kingdom are among a group of nations that have announced new climate goals this week (see ‘New Commitments’).

Climate Action Tracker is still assessing the overall impact of these pledges, but they are probably not enough to hit the 1.5 °C goal, models show.

Looking at climate policies that governments have already put in place, the group projects that global emissions will rise to the equivalent of between 51 billion and 55 billion tonnes of carbon dioxide in 2030, roughly twice what they would need to be to keep the world on a path to a rise of no more than 1.5 °C. Nonetheless, scientists say the new commitments will help to bend the curve further. If achieved, the US pledge alone would curb global emissions in 2030 by roughly 2 billion tonnes.

“There’s a long way to go, but I’m more optimistic than I was a few months ago,” says Bill Hare, who heads Climate Analytics, a non-profit consultancy in Berlin that is a member of the Climate Action Tracker consortium.

Nature **592**, 673 (2021)

doi: <https://doi.org/10.1038/d41586-021-01071-2>

Jobs from Nature Careers

- - - [All jobs](#)
 -
 - [Research Associate / Postdoc](#)

[Technische Universität Dresden \(TU Dresden\)](#)

[Dresden, Germany](#)

[JOB POST](#)

- **PhD position in evolutionary ecology of insect immune priming**

Université Bourgogne Franche-Comté (UBFC)

DIJON, France

JOB POST

- **PhD contract Doctoral contract in sociology-demography : Analysis of differential socio-demographic behaviors by sex among 50-74 year olds**

Université Bourgogne Franche-Comté (UBFC)

DIJON, France

JOB POST

- **PhD - Dark ecological networks and land management: methodological and strategic contributions for reducing ecological impacts of artificial light**

Université Bourgogne Franche-Comté (UBFC)

BESANCON, France

JOB POST

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

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-00943-x>

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

NEWS FEATURE

27 April 2021

These cellular clocks help explain why elephants are bigger than mice

Biologists are uncovering how tiny timekeepers in our cells might govern body size, lifespan and ageing.

- [Michael Marshall](#) 0

1. Michael Marshall

1. Michael Marshall is a science writer based in Devon, UK.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)



Illustration by Alberto Seveso

In her laboratory in Barcelona, Spain, Miki Ebisuya has built a clock without cogs, springs or numbers. This clock doesn't tick. It is made of genes and proteins, and it keeps time in a layer of cells that Ebisuya's team has grown in its lab. This biological clock is tiny, but it could help to explain some of the most conspicuous differences between animal species.

Animal cells bustle with activity, and the pace varies between species. In all observed instances, mouse cells run faster than human cells, which tick faster than whale cells. These differences affect how big an animal gets, how its parts are arranged and perhaps even how long it will live. But biologists

have long wondered what cellular timekeepers control these speeds, and why they vary.

A wave of research is starting to yield answers for one of the many clocks that control the workings of cells. There is a clock in early embryos that beats out a regular rhythm by activating and deactivating genes. This ‘segmentation clock’ creates repeating body segments such as the vertebrae in our spines. This is the timepiece that Ebisuya has made in her lab.

“I’m interested in biological time,” says Ebisuya, a developmental biologist at the European Molecular Biology Laboratory Barcelona. “But lifespan or gestation period, they are too long for me to study.” The swift speed of the segmentation clock makes it an ideal model system, she says.

Biologists have been studying the segmentation clock since the 1990s, and they know that it runs about twice as fast in mouse embryos as it does in human embryos. The speed at which an embryo develops, or at which different parts of it develop, has an important influence on the adult body. Ebisuya and others want to understand how differences in developmental pace give rise to organisms with such different bodies and behaviours.

In the past three years, answers have begun to emerge. This is mostly because biologists can now grow the tissue that generates the segmentation clock *in vitro*, from human stem cells, and observe its activity in detail.

“What’s truly exciting here is that you can watch it in human [tissue],” says stem-cell biologist Helen Blau at Stanford University in California. “It’s a major advance.”

The findings are already overturning some long-held assumptions about how different animals develop. So far, there is no sign of a master gene controlling the speed of the segmentation clock. Instead, its speed seems to be controlled by the differing rates at which proteins are broken down. Scientists had assumed the speed was mostly constant for each protein across animals, so the discovery might require them to revise some molecular-biology textbooks.

These differences in cellular speed might even help to explain unique features of human development, such as our oversized brains, protracted childhoods and long lives, relative to many other species.

If results from studies of the segmentation clock are true, this tiny, fleeting timepiece could help to reveal the existence of deeper, biochemical principles that shape all our lives.

Haeckel and heterochrony

Speed matters when it comes to building species. Evolution didn't give giraffes long necks by adding extra bones; they have the same number of vertebrae as their stubby-necked okapi relatives. Rather, neck vertebrae in giraffes grow over longer periods of time, which allows them to reach bigger sizes.



The labs growing human embryos for longer than ever before

This variation in the speed at which different body parts develop is called heterochrony, a concept described by Ernst Haeckel, a German zoologist noted for his work on embryo development. Modern developmental biologists regard heterochrony as a key concept that helps to explain a core mystery: at the earliest stages of development, all vertebrate embryos look

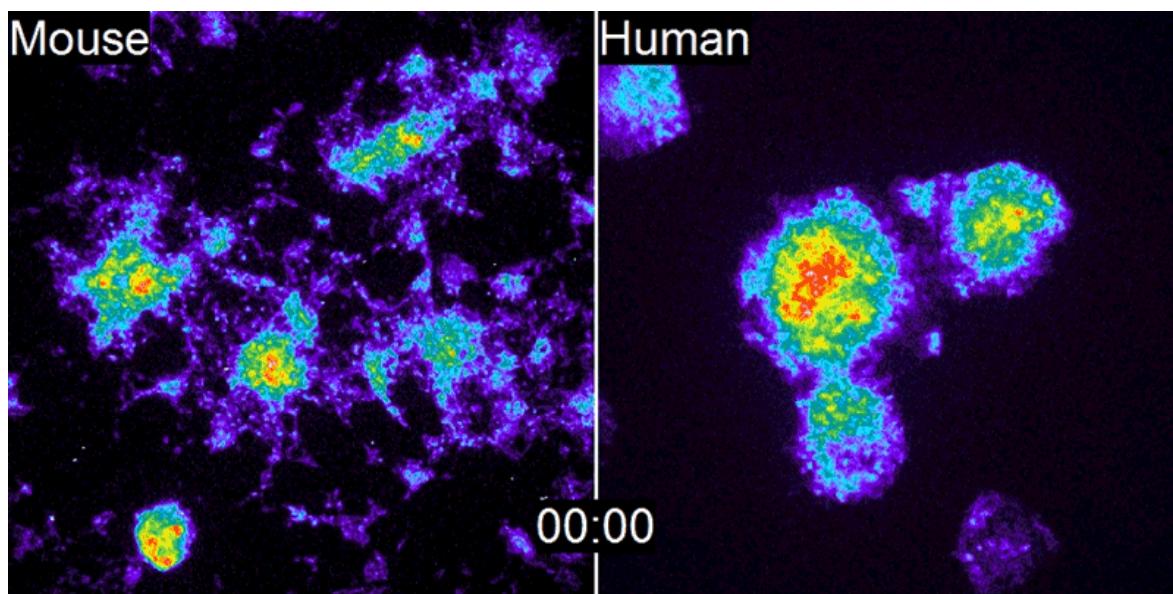
alike, yet as the embryos develop, they become easily recognizable. How do the cells of a human embryo develop into a human baby, and not into an infant chimpanzee or juvenile fish?

A big part of the answer is that the speed at which the parts of the body develop makes a big difference to what the final animal looks like. But what controls the speed of development?

Like Haeckel, modern biologists have found vertebrae and other repeating body segments useful as case studies in how the speed of development shapes animals. Decades ago, they began to investigate how body segments arise during embryogenesis.

As an embryo develops, one of its compartments splits itself into repeating segments known as somites, which run from head to tail. Each somite gives rise to a single vertebra and its associated tissue.

In 1976, two researchers proposed that cells in this compartment might each contain an oscillating mechanism of some kind, which turns itself on and off on a repeating cycle, controlling the production of somites¹. “That remained as a curiosity for some time,” says Olivier Pourquié, a developmental biologist at Harvard Medical School in Boston, Massachusetts. “And then, in the late 1990s, we identified a gene that showed a rhythmic behaviour in the tissue that’s going to form the somite.”



A gene called *Hes7* turns itself on and off twice as fast in cultured mouse stem cells (left) as it does in human cells (right). Credit: M. Matsuda *et al./Science*

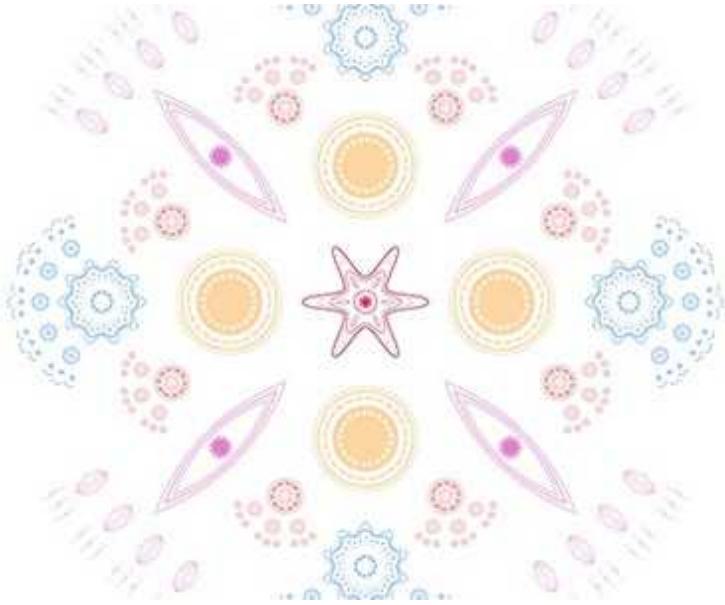
Pourquié's team studied developing chick embryos and found that a gene called *c-hairy1* flicked on and off every 90 minutes² — the time it takes for a somite to form. Waves of *c-hairy1* expression moved along the embryo from tail to head, oscillating in synchrony with the development of somites. Similar segmentation clocks have since been found in mice and other species.

Ever since, Pourquié and other biologists have been trying to take the segmentation clock apart and understand how it works, building a long list of genes and proteins that help the clock to keep time. One key gene is *Hes7*, the mammalian equivalent of the bird gene *c-hairy1*. *Hes7* can repeatedly turn itself on and off, as can several other genes involved in the clock. That makes it “a key pacemaker for the segmentation clock”, says Ryoichiro Kageyama, a developmental biologist at Kyoto University in Japan who has studied the gene for almost two decades.

But it is still unclear why *Hes7* turns on and off at different speeds in different species, and thus how the speed of the segmentation clock is ultimately controlled. A series of studies over the past three years point to an answer.

Unpicking the clock

In 2019 and 2020, several labs showed that they can recreate the human segmentation clock *in vitro*, by culturing stem cells so that they develop into somite-forming tissue^{3–5}. This was the first hard evidence that humans have a segmentation clock — although this was widely expected. More importantly, creating the clock *in vitro* meant that it could be studied in human tissue for the first time, and allowed much more fine-grained analysis of its mechanism.



The trickiest family tree in biology

These studies were made possible by advances in the culturing of stem cells to persuade them to grow into specific tissues, says Pierre Vanderhaeghen, a developmental neurobiologist at Leuven University in Belgium. In human embryos, the segmentation clock is active only between about the third and fourth weeks of development. “That’s even before women know they’re pregnant,” says Pourquié. “So we know nothing about it. This *in vitro* system provides a proxy for us to study.”

In 2018, Kageyama’s team showed that it could take mouse embryonic stem cells and grow them into somite-forming tissue, complete with oscillating *Hes7* gene expression⁶. In 2019 and 2020, three independent groups, led by Pourquié, Ebisuya and stem-cell biologist James Thomson at the Morgridge Institute for Research in Madison, Wisconsin, showed that the same trick could be achieved with human stem cells^{3–5}.

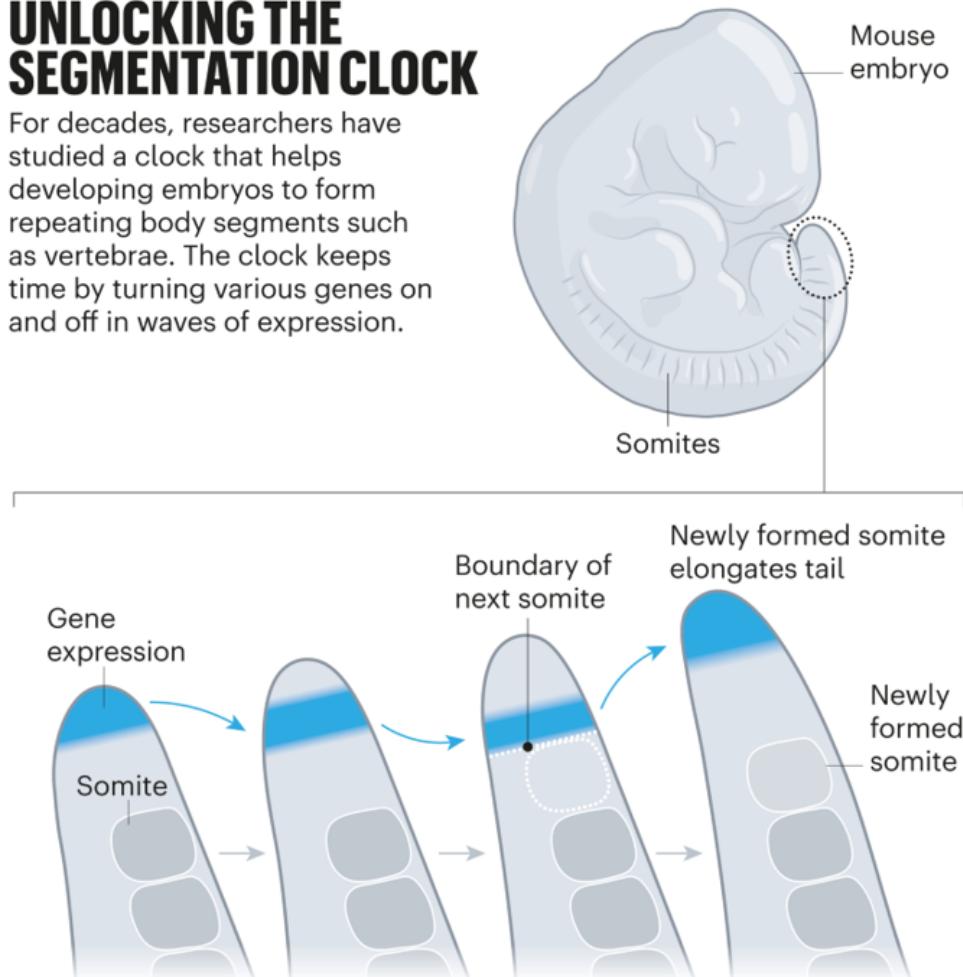
These studies revealed many similarities between the segmentation clock of humans and those of other animals. Analogues of the same genes and proteins are involved in mice and humans, for instance.

But there was one striking difference. The human segmentation clock is slow. Each oscillation takes 5–6 hours, twice as long as the 2–3 hours it takes in mouse embryos: a clear example of heterochrony (see ‘Unlocking

the segmentation clock'). But why is the human segmentation clock so slow, and what is controlling it?

UNLOCKING THE SEGMENTATION CLOCK

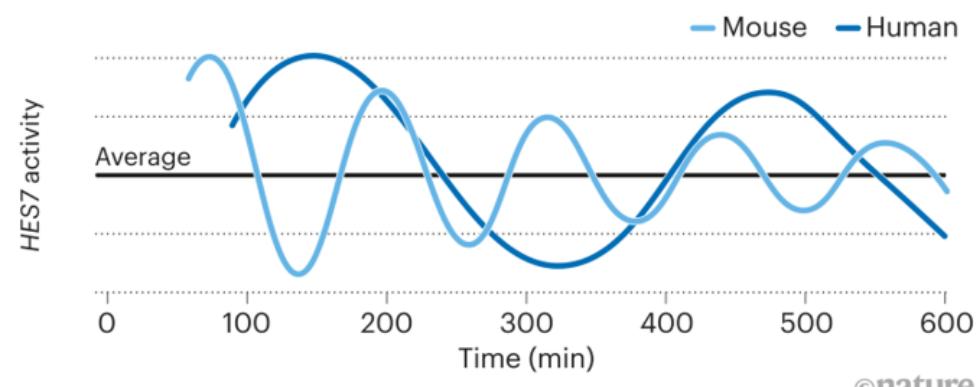
For decades, researchers have studied a clock that helps developing embryos to form repeating body segments such as vertebrae. The clock keeps time by turning various genes on and off in waves of expression.



In the developing embryo, the activity of a group of genes oscillates over time, dictating where the next segment — called a somite — will form.

Human versus mouse

Scientists have recreated the human segmentation clock *in vitro*, and followed the activity of a key gene called *Hes7*. When compared with that of the mouse, the human clock ‘ticks’ half as fast.



©nature

Source: Ref. 7. Graphic: Nik Spencer/*Nature*

Running slow

Two papers published together in *Science* last September^{7,8} offered a possible answer.

Ebisuya's team focused on the *Hes7* gene, which she calls "the core of the segmentation clock". To check whether the human and mouse versions of *Hes7* were controlling the cells' different speeds, they placed human *Hes7* into mouse cells and mouse *Hes7* into human cells, then watched to see whether the human cells started oscillating at mouse speed and vice versa. But the speed of the oscillations hardly changed at all⁷. Something else was influencing *Hes7*.

To explore this, the team considered how *Hes7* actually works. When the gene is active, it produces the *Hes7* protein, and when enough of the protein builds up, this deactivates the gene. Then, once the *Hes7* proteins have been broken down, the gene can reactivate. In this way *Hes7* keeps turning itself on and off.

Ebisuya's team wondered whether the *Hes7* protein might be broken down more slowly in human cells than in mouse cells, and whether this would account for the slower oscillations in *Hes7* activity — and thus the slower segmentation clock. In further experiments, they found that the *Hes7* protein and its RNA template were indeed degraded much more slowly in human cells⁷.

It's not clear exactly why this is. *Hes7* proteins are degraded by structures called proteasomes, after first being tagged for destruction. "But we don't know which part of this degradation process is slower," says Ebisuya.

What is intriguing is that this slow degradation of human proteins is not limited to *Hes7*, or even the segmentation clock. This was borne out by the second study⁸, which was led by James Briscoe, a developmental biologist at the Francis Crick Institute in London. His team studied the differentiation of motor neurons in the spinal cords of mouse and human embryos (see

‘Speedy genes’). This takes place in a different part of the embryo from the formation of vertebrae, and does not involve the segmentation clock. Yet the process is still slow in humans, taking about 2 weeks, compared with 3–4 days in mice.

SPEEDY GENES

A study from 2020 found that as human and mouse stem cells differentiate into motor neurons, they express the same genes in the same sequence but the process is twice as fast in mice — a possible reason why the two creatures develop at different rates.



Source: Ref. 8

The team found that the human proteins took about twice as long to break down as the mouse proteins did — which seemed to be determining the speed at which motor neurons develop. This is strikingly similar to what Ebisuya found, says Briscoe. “We found exactly the same time difference in the spinal cord that she was finding in segmentation.”

The mouse and human proteins are close to identical, adds Briscoe's team member Teresa Rayon. That means it is unlikely that the mouse proteins are inherently less stable. "We suspect it's something to do with how proteins are degraded."

Nevertheless, Blau and Vanderhaeghen are cautious about the idea that the rate of protein degradation is the key to the segmentation clock's variable speed. "I don't know that we know that," says Blau. The studies have ruled out some explanations for the difference in speed, she says, but they don't yet prove that the rate of protein degradation is responsible.

Reaction speed

Meanwhile, in unpublished experiments, Ebisuya is exploring whether all human proteins are degraded more slowly than mouse proteins. "We think there's a general trend that degradation rates become slower in human cells", she says, but it might not apply to all proteins. Her team also has evidence that as well as being degraded at more leisurely rates, proteins are produced more slowly in human cells than in other species. Vanderhaeghen says that some other component of the cell, such as metabolic cycles or mitochondrial processes, might also be running faster or slower in different species.

The researchers are all uncertain why biochemical reactions would be systematically slower in human cells — both how the difference arises mechanistically, and why it arose in evolutionary history. For Pourquié, the mechanism might yet lie in the genome. "There has to be some kind of genetic trick that arose during evolution that leads to this slowing down of all these processes in human cells," he says. But it's not clear why this would be evolutionarily advantageous.



CRISPR's hopeful monsters: gene-editing storms evo-devo labs

The relative sluggishness of human cells could be a product of their composition or complexity, says J. Kim Dale, a developmental biologist at the University of Dundee, UK. For example, the degradation machinery might find itself struggling to keep up with demand, slowing reactions down. “With the information that we have, it would suggest it’s the cell environment,” she says.

The *in vitro* segmentation clock studies could well resolve this question, but also suggest a broader mystery: do human cells run slower than those of other species, not just during specific periods of development, but throughout our lives? If so, that could help to explain why our lifespan is extended compared with that of other species.

It is too early to be sure, but a January study⁹ suggests that these variations in biochemical reaction speeds run deep in biology. A team led by Sina Ghaemmaghami, who studies proteomics at the University of Rochester in New York, compared how rapidly proteins were created and destroyed in the skin cells of 12 mammals, ranging from golden hamsters that barely live 4 years, to humans, to bowhead whales that can live 200 years.

“I thought there was no way we were going to see much difference,” says Ghaemmaghami. Proteomics textbooks often argue that the half-life of a

protein is inherent in its structure, he says, so these highly conserved proteins — which vary little between species — should last about as long in all animals. But in fact the team found a strong inverse correlation with lifespan: longer-lived species had slower turnover of proteins.

Are longer-lived animals simply running slower at the biochemical level? “That’s the million-dollar question,” says Ghaemmaghami. “Is it that the slow turnover is in some way causing long lifespan, or is it that these organisms have long lifespans for a completely independent reason and then can adjust their turnover rate? It’s really hard to know.”

For now, Ghaemmaghami’s working hypothesis is that the slowness is a consequence of long lifespan. He points out that making and breaking proteins rapidly is a good thing, because it ensures the cells are using high-quality proteins — but all that activity releases harmful waste products that can damage the cell. “If you’re a long-lived organism, you can’t just rapidly turn your proteins over, because you damage everything else,” he says. Instead, his team thinks that long-lived animals reduce the overall turnover and pinpoint only damaged proteins for degradation.

Ebisuya thinks that the speed of chemical reactions could be key to these differences, but she wants to work out the mechanism before she is ready to generalize. “I’m still not sure whether similar mechanisms can explain other biological processes,” she says.

It might take time, but the clocks Ebisuya and others have built promise to reveal much more about how animals tick.

Nature **592**, 682-684 (2021)

doi: <https://doi.org/10.1038/d41586-021-01086-9>

References

1. 1.

Cooke, J. & Zeeman, E. C. *J. Theor. Biol.* **58**, 455–476 (1976).

2. 2.

Palmeirim, I., Henrique, D., Ish-Horowicz, D. & Pourquié, O. *Cell* **91**, 639–648 (1997).

3. 3.

Diaz-Cuadros, M. *et al.* *Nature* **580**, 113–118 (2020).

4. 4.

Matsuda, M. *et al.* *Nature* **580**, 124–129 (2020).

5. 5.

Chu, L.-F. *et al.* *Cell Rep.* **28**, 2247–2255 (2019).

6. 6.

Matsumiya, M., Tomita, T., Yoshioka-Kobayashi, K., Isomura, A. & Kageyama, R. *Development* **145**, dev156836 (2018).

7. 7.

Matsuda, M. *et al.* *Science* **369**, 1450–1455 (2020).

8. 8.

Rayon, T. *et al.* *Science* **369**, eaba7667 (2020).

9. 9.

Swovick, K. *et al.* *Mol. Cell Proteom.* **20**, 100041 (2021).

Jobs from Nature Careers

- - - [All jobs](#)

o

- **Research Associate / Postdoc**

Technische Universität Dresden (TU Dresden)

Dresden, Germany

JOB POST

- **PhD position in evolutionary ecology of insect immune priming**

Université Bourgogne Franche-Comté (UBFC)

DIJON, France

JOB POST

- **PhD contract Doctoral contract in sociology-demography : Analysis of differential socio-demographic behaviors by sex among 50-74 year olds**

Université Bourgogne Franche-Comté (UBFC)

DIJON, France

JOB POST

- **PhD - Dark ecological networks and land management: methodological and strategic contributions for reducing ecological impacts of artificial light**

Université Bourgogne Franche-Comté (UBFC)

BESANCON, France

JOB POST

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01086-9>

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

Opinion

- **Restore autonomy for Pakistan's Higher Education Commission** [21 April 2021]
Correspondence •
- **Pakistan: anger mounts over threat to higher education**
[27 April 2021]
Correspondence •
- **China's wildlife protection: add annual reviews and oversight** [27 April 2021]
Correspondence •
- **Peru scandal: gaps in oversight of COVID vaccine trial** [27 April 2021]
Correspondence •

CORRESPONDENCE

21 April 2021

Restore autonomy for Pakistan's Higher Education Commission

- [Farasat Zaman](#)⁰ &
- [Amir Rashid](#)¹

1. [Farasat Zaman](#)

1. Karolinska Institute, Stockholm, Sweden.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Amir Rashid

1. KTH Royal Institute of Technology, Stockholm, Sweden.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

The Pakistani government has cut the contract of the chair of the country's autonomous Higher Education Commission (HEC) from four to two years. Because the chair of the funding and regulatory body — Tariq Banuri, an economist at the University of Utah in Salt Lake City — had already served two years, his appointment was effectively terminated. This unexplained 'sacking' is widely seen by researchers, the public and educationists as an attack on the sector's independence.

These moves fly in the face of the government's policy of fostering meritocracy and transparency in public-sector organizations. Key to that is independence in regulatory policymaking for universities, given their pivotal role in developing a knowledge-based economy, and their international visibility ([P. S. Aithal and S. Aithal *Int. J. Appl. Eng. Mgmt Lett.* 3\(2\), 36–52; 2019](#)).

Prime Minister Imran Khan, a long-standing champion of universities and research, has urged members of the Pakistani diaspora to contribute their research skills to nation-building. Unless HEC autonomy is restored, the nation will become less attractive to Pakistani scientists working abroad and to international researchers, along with the funding that they bring.

Nature **592**, 685 (2021)

doi: <https://doi.org/10.1038/d41586-021-01073-0>

Jobs from Nature Careers

- - - [All jobs](#)
 - - [Research Associate / Postdoc](#)
 - [Technische Universität Dresden \(TU Dresden\)](#)
 - [Dresden, Germany](#)
 - [JOB POST](#)
 - [PhD position in evolutionary ecology of insect immune priming](#)
 - [Université Bourgogne Franche-Comté \(UBFC\)](#)

DIJON, France

JOB POST

- **PhD contract Doctoral contract in sociology-demography : Analysis of differential socio-demographic behaviors by sex among 50-74 year olds**

Université Bourgogne Franche-Comté (UBFC)

DIJON, France

JOB POST

- **PhD - Dark ecological networks and land management: methodological and strategic contributions for reducing ecological impacts of artificial light**

Université Bourgogne Franche-Comté (UBFC)

BESANCON, France

JOB POST

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01073-0>

CORRESPONDENCE

27 April 2021

Pakistan: anger mounts over threat to higher education

- [Aziz Khan](#) 0

1. [Aziz Khan](#)

1. Stanford University, California, USA.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

On 8 April, the Pakistani government passed an ordinance — without any justifiable reason — that will lead to the country's Higher Education Commission (HEC) losing its independence. More than 2,200 academics have now signed a petition calling for action against such political interference (go.nature.com/2rnf4v), which could endanger the future of higher education in Pakistan.

The HEC was established in 2002 as an independent agency to revamp the country's higher education, initially backed by generous funding that has since dwindled. It has been the butt of criticism over the years ([Nature 467, 378–379; 2010](#); [Nature 560, 419; 2018](#)). The commission's success should instead be gauged by its role in promoting a culture of responsible research and education in Pakistan and beyond.

The nation must fiercely protect the autonomy and funding of the HEC, subject to accountability. Experts in science, academic policy and governance can help to rigorously review its policies, programmes and progress every year.

Pakistan's future is interwoven with fair and independent higher education and science. It could learn from the investment and progress in research made by its neighbours in China and India (*Nature* **590**, 184; 2021). The government will need to upgrade its scientific infrastructure and enterprise to broaden its development and make it sustainable.

Nature **592**, 685 (2021)

doi: <https://doi.org/10.1038/d41586-021-01119-3>

Jobs from Nature Careers

- - - [All jobs](#)
 - - [Research Associate / Postdoc](#)
 - [Technische Universität Dresden \(TU Dresden\)](#)
[Dresden, Germany](#)
[JOB POST](#)
 - [PhD position in evolutionary ecology of insect immune priming](#)
[Université Bourgogne Franche-Comté \(UBFC\)](#)
[DIJON, France](#)
[JOB POST](#)
 - [PhD contract Doctoral contract in sociology-demography : Analysis of differential socio-](#)

demographic behaviors by sex among 50-74 year olds

Université Bourgogne Franche-Comté (UBFC)

DIJON, France

JOB POST

- PhD - Dark ecological networks and land management: methodological and strategic contributions for reducing ecological impacts of artificial light

Université Bourgogne Franche-Comté (UBFC)

BESANCON, France

JOB POST

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01119-3>

CORRESPONDENCE

27 April 2021

China's wildlife protection: add annual reviews and oversight

- [Hai-Tao Shi](#)⁰,
- [Jian Wang](#)¹,
- [Huai-Qing Chen](#)² &
- [James F. Parham](#)³

1. Hai-Tao Shi

1. Hainan Normal University, Haikou, Hainan, China.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. [Jian Wang](#)

1. Honghe University, Mengzi, China.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Huai-Qing Chen

1. Peking University, Beijing, China.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. James F. Parham

1. California State University, Fullerton, California, USA.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Now that China has finally updated its List of Wildlife under Special State Protection, a more nimble and responsive approach is needed to aid conservation. The list should be reviewed every year, as well as subjected to the planned five-yearly updates. Species can quickly become endangered in times of rapid development.

The latest additions are the first in more than 30 years (see go.nature.com/2q7sfga). During that time, China has changed profoundly, but the list of protected species has not kept pace. This lag has been disastrous for some animals that were not given the protection they needed.

At least 33 species became extinct in China and many more are critically endangered ([Y. Xie & W. Sung *Integr. Zool.* 2, 26–35; 2007](#); [Z. Jiang *et al. Biodivers. Sci.* 24, 500–551; 2016](#)).

An independent government committee should be created to oversee amendments. When making decisions, it could refer to appendices of the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES) and the ‘red lists’ of threatened species curated by the Chinese Academy of Sciences and the International Union for Conservation of Nature (IUCN). These steps would build on the more forceful approach to managing wildlife that China has taken since the start of the COVID-19 pandemic.

Nature **592**, 685 (2021)

doi: <https://doi.org/10.1038/d41586-021-01111-x>

[Jobs from Nature Careers](#)

- - - [All jobs](#)
 - - [**Research Associate / Postdoc**](#)

[Technische Universität Dresden \(TU Dresden\)](#)
[Dresden, Germany](#)
[JOB POST](#)

 - [**PhD position in evolutionary ecology of insect immune priming**](#)

[Université Bourgogne Franche-Comté \(UBFC\)](#)
[DIJON, France](#)
[JOB POST](#)
 - [**PhD contract Doctoral contract in sociology-demography : Analysis of differential socio-demographic behaviors by sex among 50-74 year olds**](#)

[Université Bourgogne Franche-Comté \(UBFC\)](#)
[DIJON, France](#)
[JOB POST](#)
 - [**PhD - Dark ecological networks and land management: methodological and strategic**](#)

contributions for reducing ecological impacts of artificial light

Université Bourgogne Franche-Comté (UBFC)

BESANCON, France

JOB POST

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01111-x>

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

CORRESPONDENCE

27 April 2021

Peru scandal: gaps in oversight of COVID vaccine trial

- [Carlos F. Cáceres](#)⁰,
- [Alfonso Zavaleta](#)¹ &
- [Gustavo F. González](#)²

1. [Carlos F. Cáceres](#)

1. Cayetano Heredia University, Lima, Peru.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Alfonso Zavaleta

1. Cayetano Heredia University, Lima, Peru.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Gustavo F. González

1. Cayetano Heredia University, Lima, Peru.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

We wish to clarify points in your report on Peruvian COVID vaccine trials that could reflect negatively on Cayetano Heredia University in Lima (UPCH) (*Nature* **592**, 174–175; 2021).

The UPCH conducts hundreds of studies every year and is committed to high-quality, ethical research. However, its Sinopharm trial diverged from global standards by including an off-trial vaccine lot “to protect the research-team and trial-related personnel”. This unusual provision was negotiated between governments in the context of countries ordering millions of doses of vaccines that are still not available.

Following international recommendations (see go.nature.com/3tcahuu), the Peruvian government’s National Institute of Health (INS) appointed a national research-ethics committee to provide prompt review and oversight of COVID vaccine trials. This meant that the study protocol was not reviewed or approved by the UPCH’s research-ethics committee.

The protocol was approved by the INS and the national ethics committee. These bodies organized oversight visits during the trial, but did not ascertain whether or how the extra vaccines were used. Had this happened, the whole process could have been called into question.

The UPCH has started a high-level investigation and appointed a new principal investigator for the study. It will ensure that such scenarios cannot reoccur.

Nature **592**, 685 (2021)

doi: <https://doi.org/10.1038/d41586-021-01120-w>

Jobs from Nature Careers

- - - [All jobs](#)
 -

- **Research Associate / Postdoc**

Technische Universität Dresden (TU Dresden)

Dresden, Germany

JOB POST

- **PhD position in evolutionary ecology of insect immune priming**

Université Bourgogne Franche-Comté (UBFC)

DIJON, France

JOB POST

- **PhD contract Doctoral contract in sociology-demography : Analysis of differential socio-demographic behaviors by sex among 50-74 year olds**

Université Bourgogne Franche-Comté (UBFC)

DIJON, France

JOB POST

- **PhD - Dark ecological networks and land management: methodological and strategic contributions for reducing ecological impacts of artificial light**

Université Bourgogne Franche-Comté (UBFC)

BESANCON, France

JOB POST

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01120-w>

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

Work

- **Minority representation in US science workforce sees few gains** [23 April 2021]

Career News • Proportion of Black and Hispanic researchers has changed little in the past few years.

- **'Less pipetting and more thinking': scientists carry on through the pandemic** [22 April 2021]

Career News • Researchers worldwide have turned to analysing data, planning experiments and writing and reviewing more papers as their lab time has been disrupted, a global survey finds.

- **On the trail of travelling RNA** [26 April 2021]

Where I Work • Molecular biologist Juan Pablo Tosar wants to understand how cells talk to each other.

CAREER NEWS

23 April 2021

Minority representation in US science workforce sees few gains

Proportion of Black and Hispanic researchers has changed little in the past few years.

- [Chris Woolston](#) 0

1. Chris Woolston

1. Chris Woolston is a freelance journalist in Billings, Montana.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)



Black researchers are making small gains in representation in the life sciences. Credit: Getty

Members of minority ethnic groups have made only modest inroads into US science, technology, engineering and mathematics (STEM) positions in recent years, according to an analysis of nearly 20 million people.

[The analysis](#) was conducted by the Pew Research Center, a non-profit organization in Washington DC, and used data collected by the Integrated Public Use Microdata Series American Community Survey. It found that Black people accounted for 9% of the STEM workforce in 2019. That's the same proportion as in 2016, suggesting a lack of progress at a time when many companies and universities had pledged to promote diversity. Over the same period, the proportion of Hispanic workers in STEM jobs rose from 7% to 8%. Black and Hispanic people make up 12% and 17% of the US workforce, respectively.



Collection: Diversity and scientific careers

The data reveal some modest gains for under-represented minorities. For example, Black employees accounted for 6% of all jobs in the life sciences in 2019, up from 4% in 2016. In that same period, the proportion of Hispanic employees in the life sciences increased from 7% to 8%. They also accounted for 8% of all jobs in the mathematics field in 2019, up from 6% in 2016. Over that period, the percentage of Black people in maths-related jobs held steady at 9%.

Asian people, who made up 6% of the working population in the United States in 2019, held 19% of positions in life sciences and 18% of positions in physical sciences. Asian people accounted for 13% of the STEM workforce overall, the same proportion as in 2016.

The report also found that the percentage of women in STEM fields has remained stable. In 2019, women accounted for 47% of the overall workforce and 50% of the STEM workforce.

Links to education

The continued lack of racial diversity is a bit surprising, says Marcus Lambert, a higher-education researcher at SUNY Downstate Health Sciences

University in New York City. He notes that the trend persists despite some gains in the enrolment of members of minority ethnic groups in university science programmes. “The STEM workforce is connected to the educational pathways,” he says. “If people are switching majors and dropping out of STEM programmes, we’re not doing our jobs.”

Lambert was a co-author of a 2020 study that surveyed more than 1,200 biomedical postdoctoral researchers in the United States about their career plans¹. It found that 24% of male postdocs from minority ethnic groups had decided to get out of research or leave science altogether, even though they had already gone far in their training. For comparison, only 14% of white postdocs were planning to quit research or leave science.

Lambert’s survey found that even some highly productive postdocs from minority ethnic groups who had authored a number of high-quality publications didn’t feel completely comfortable in their jobs. “It’s not that they don’t love science,” he says. “Whether it’s politics or negative stereotypes, they don’t feel welcome or completely satisfied.”



How I embrace diversity in my lab

The Pew report found that representation of people from minority ethnic groups varied greatly depending on the job in question. For example, within the life sciences, the proportion of Black people employed as technicians in

the food or agriculture industry (12%) was six times the proportion who worked as conservation scientists and foresters (2%). In the physical sciences, Hispanic representation was more than twice as high in chemical technicians (12%) as in astronomers and physicists (5%).

Many US universities have pledged to increase diversity in hiring, but progress has been notably slow, says Melanie Okoro, a water-quality researcher and chief executive and principal scientist at Eco-Alpha Environmental and Engineering Services, a consulting company in Sacramento, California. Okoro is also a board member of MS PHD's, an initiative sponsored by NASA and the US National Science Foundation to increase minority participation in Earth sciences. Part of the problem at the university level, she says, is that predominantly white institutions have been reluctant to partner with minority institutions, such as Howard University in Washington DC, to identify qualified applicants.

Okoro thinks that private companies such as hers could help drive progress towards more equitable representation in STEM fields. "As someone who owns a company, I'm very intentional in who I hire and the culture that my company can reflect," she says. In the big picture, she says, "significant strides are being made to support a diverse and inclusive workforce, but sustainable changes will take time".

Nature **592**, 805-806 (2021)

doi: <https://doi.org/10.1038/d41586-021-01089-6>

References

1. 1.

Lambert, W. M. *et al.* *eLife* **9**, e48774 (2020).

Jobs from Nature Careers

- -

- [All jobs](#)
 -
- [Research Associate / Postdoc](#)

[Technische Universität Dresden \(TU Dresden\)](#)
[Dresden, Germany](#)

[JOB POST](#)
- [PhD position in evolutionary ecology of insect immune priming](#)

[Université Bourgogne Franche-Comté \(UBFC\)](#)
[DIJON, France](#)

[JOB POST](#)
- [PhD contract Doctoral contract in sociology-demography : Analysis of differential socio-demographic behaviors by sex among 50-74 year olds](#)

[Université Bourgogne Franche-Comté \(UBFC\)](#)
[DIJON, France](#)

[JOB POST](#)
- [PhD - Dark ecological networks and land management: methodological and strategic contributions for reducing ecological impacts of artificial light](#)

Université Bourgogne Franche-Comté (UBFC)

BESANCON, France

JOB POST

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01089-6>

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

CAREER NEWS

22 April 2021

‘Less pipetting and more thinking’: scientists carry on through the pandemic

Researchers worldwide have turned to analysing data, planning experiments and writing and reviewing more papers as their lab time has been disrupted, a global survey finds.

- [Chris Woolston](#) 0

1. Chris Woolston

1. Chris Woolston is a freelance writer in Billings, Montana.

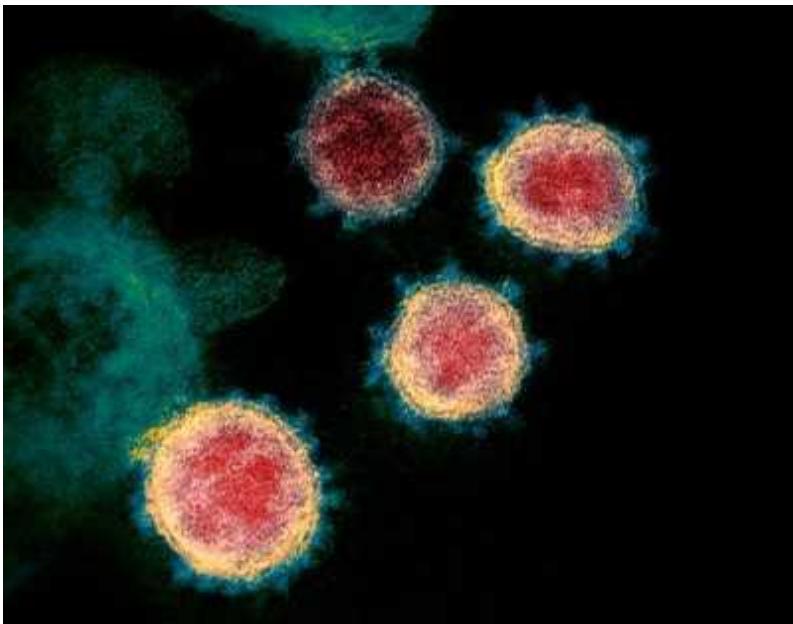
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)



The pandemic has presented obstacles and upsides for researchers globally. Credit: Getty

Nearly all respondents (97%) to a global survey of 2,000 researchers reported that the COVID-19 pandemic has affected their work — and half reported ‘significant’ impact — but most are staying productive despite the disruptions. Those are among the key findings of a [study of users of ResearchGate](#), the Berlin-based scientific social-network and paper-sharing site. The survey was conducted over an 18-hour period on 19 February. The survey report does not disclose where respondents live or their disciplines or career stages.



[Coronavirus and COVID-19: Keep up to date](#)

Two-thirds of respondents are continuing to work on data analysis and experimental planning, even if the experiments themselves have to wait until lockdowns are fully lifted. In the comment section of the survey, a researcher whose institution was shuttered at the time reported that they had found other ways to work. “The positive is that I have had the opportunity to study a large quantity of research papers, write two review articles, and complete one research paper.”

Another researcher noted the extra obstacles to doing research during the pandemic: “A lot of my time is spent devising ways to accomplish tasks that would have been easier in the lab or face-to-face.”

The pandemic has transformed the research workplace. In the survey, more than eight out of ten respondents said they work from home. Of those, about one-third reported that they weren’t working exclusively at home, suggesting they occasionally ventured to the office or the lab.

Keeping busy

The ResearchGate survey underscores the fact that scientists can stay productive even during catastrophe, says Alisa Wolberg, a haematology

researcher at the University of North Carolina (UNC) at Chapel Hill. Wolberg and one of her PhD students, Dre’Von Dobson, co-authored an article published last September¹ that discusses some of the pandemic’s ‘silver linings’ for scientific research.

Wolberg notes that researchers have had more time to ponder the big picture of their work — or, as she puts it, there’s been “less pipetting and more thinking”. “I don’t want to undersell how tragic this experience has been,” she says. “But scientists love to find problems and solve them. We’re going to stay busy with the other things that our jobs entail.”

The survey’s findings don’t precisely correspond with the results of other studies that found that the pandemic has significantly disrupted not just many people’s scientific output, but also their careers. In a [Nature survey of postdocs last year](#), for example, 70% of postdoctoral researchers in South America reported that their careers have suffered during the pandemic.

In the ResearchGate survey, 40% of respondents said they had spent more or much more time searching for and reading scientific literature. Slightly more than half have spent more or much more time writing, submitting and peer-reviewing papers. “With so many scientists at home, papers are being written and submitted en masse,” wrote one respondent. “I’ve had to review four times more papers than usual during this time.”

The pandemic has slowed some of Dobson’s experimental work. On the upside, he’s been able to build up his network of colleagues and mentors through video conferencing. For example, his group at the UNC shares its weekly lab meetings with a group led by James Luyendyk at Michigan State University in East Lansing. “I feel like I’ve gained another mentor,” Dobson says. “If I’m giving a presentation, I can reach out to Jim to see what he thinks.” For 62% of respondents, the amount of time they spend collaborating with others either stayed the same or increased.

All of that connectivity has a downside, however, especially when it comes to teaching duties: 40% of respondents have spent more or much more time teaching during the pandemic, and the comment section reflected widespread frustration with the demands of online instruction. “Online teaching duties require far more preparation,” wrote one respondent. “Work–

life balance has been destroyed, as e-mail availability requires 24–7 responsibility.”

Dobson acknowledges that the pandemic has complicated his research and training — but, like many survey respondents, he’s moving ahead. “The bad things outweigh the good, but there are things that we can all take with us from this experience as we transition to some form of normalcy.”

Nature **592**, 806 (2021)

doi: <https://doi.org/10.1038/d41586-021-01075-y>

References

1. 1.

Dre’Von, A. *et al. Res. Pract. Thromb. Haemost.* **4**, 1083–1086 (2020).

Jobs from Nature Careers

- - - [All jobs](#)
 - - [**Research Associate / Postdoc**](#)

[Technische Universität Dresden \(TU Dresden\)](#)
[Dresden, Germany](#)
[JOB POST](#)

▪ [**PhD position in evolutionary ecology of insect immune priming**](#)

[Université Bourgogne Franche-Comté \(UBFC\)](#)

DIJON, France

JOB POST

- **PhD contract Doctoral contract in sociology-demography : Analysis of differential socio-demographic behaviors by sex among 50-74 year olds**

Université Bourgogne Franche-Comté (UBFC)

DIJON, France

JOB POST

- **PhD - Dark ecological networks and land management: methodological and strategic contributions for reducing ecological impacts of artificial light**

Université Bourgogne Franche-Comté (UBFC)

BESANCON, France

JOB POST

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01075-y>.

WHERE I WORK

26 April 2021

On the trail of travelling RNA

Molecular biologist Juan Pablo Tosar wants to understand how cells talk to each other.

- [Sara Reardon](#) ⁰

1. Sara Reardon

1. Sara Reardon is a freelance science writer in Bozeman, Montana.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Juan Pablo Tosar separating extracellular RNA-containing complexes by size-exclusion chromatography in his lab

Juan Pablo Tosar is a molecular biologist at the Pasteur Institute and the University of the Republic of Uruguay in Montevideo, Uruguay. Pablo Albarenga for *Nature*

When you see someone wearing gloves, you tend to think they're protecting themselves from something. In my laboratory at the Pasteur Institute in Montevideo, Uruguay, we wear gloves to protect our samples from ourselves. We study RNA, and some of the enzymes in our skin can break those molecules down.

I am a molecular biologist, and I aim to understand how cells talk to each other using their RNA. Cells sometimes release RNAs that travel from one cell to another, where they can affect gene expression. We're also studying how these travelling RNAs in human blood or urine samples can be used to

diagnose disease. We can look at all the extracellular RNA in a sample, and determine whether any of the molecules contain abnormalities that might point to the presence of conditions such as cancer.

In this picture from February, I'm using a syringe to inject a sample containing extracellular RNA into a vertical column of porous resin beads. The column separates out the contents of a sample by size: small molecules get stuck in the pores, and bigger stuff drops straight through and comes out the end faster. A protein carrying a molecule of RNA is large and comes out of the column quickly. When the sample is treated with enzymes to break down the RNA, the protein is freed and, because it is now smaller, it comes out of the column later. We study the difference between these descent speeds to see which proteins are interacting with RNA.

Uruguay is a small country, and the resources for research are not great. We often need to get creative — not just in our science, but also in how we manage our resources. If a reviewer asks for an experiment that requires special reagents that would take us months to get, I might need to think of other ways of showing the same thing. Finding alternatives is part of this job. It's a way to approach the most complex questions.

Nature **592**, 808 (2021)

doi: <https://doi.org/10.1038/d41586-021-01088-7>

Jobs from Nature Careers

- - - [All jobs](#)
 - - [Research Associate / Postdoc](#)

[Technische Universität Dresden \(TU Dresden\)](#)

[Dresden, Germany](#)

JOB POST

- [PhD position in evolutionary ecology of insect immune priming](#)

[Université Bourgogne Franche-Comté \(UBFC\)](#)

[DIJON, France](#)

JOB POST

- [PhD contract Doctoral contract in sociology-demography : Analysis of differential socio-demographic behaviors by sex among 50-74 year olds](#)

[Université Bourgogne Franche-Comté \(UBFC\)](#)

[DIJON, France](#)

JOB POST

- [PhD - Dark ecological networks and land management: methodological and strategic contributions for reducing ecological impacts of artificial light](#)

[Université Bourgogne Franche-Comté \(UBFC\)](#)

[BESANCON, France](#)

JOB POST

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

Research

- **[Highly reactive form of magnesium stabilized by bulky ligands](#)** [28 April 2021]

News & Views • Magnesium atoms typically lose two electrons to form chemical compounds. A reactive complex has finally been made in which magnesium keeps all of its electrons, and which can be thought of as a soluble form of the metal.

- **[Life in a carbon dioxide world](#)** [21 April 2021]

News & Views • Microorganisms living in hydrothermal vents that emit carbon dioxide gas provide a striking example of metabolic finesse. This pathway sheds light on microbial ecology in extreme environments and offers clues to early life on Earth.

- **[The path to destruction for D-type cyclin proteins](#)** [14 April 2021]

News & Views • The molecular mechanism governing the destruction of key cell-cycle proteins, D-type cyclins, has been elucidated. This mechanism might underlie the lack of response of some human tumours to an inhibitor treatment.

- **[Ancient continental blocks soldered from below](#)** [28 April 2021]

News & Views • A study of melting in the mantle under northern Canada more than one billion years ago shows that the oldest blocks of continent not only break apart but can also be repaired by the gluing action of major melting episodes.

- **[The central role of DNA damage in the ageing process](#)** [28 April 2021]

Review Article • This Review examines the evidence showing that DNA damage is associated with ageing phenotypes, suggesting that it may have a central role as the cause of ageing.

- **[X-ray quasi-periodic eruptions from two previously quiescent galaxies](#)** [28 April 2021]

Article • X-ray quasi-periodic eruptions are detected from two previously inactive galaxies, with observations suggesting that the very-high-amplitude X-ray bursts may arise from an orbiting compact object.

- **[Transition from an atomic to a molecular Bose–Einstein condensate](#)** [28 April 2021]

Article • A Bose-Einstein condensate of molecules is produced by pairing atoms in an atomic condensate; this transition is the bosonic analog of the Bardeen-Cooper-Schrieffer superfluid to BEC crossover in Fermi gases.

- **Direct observation of chemical short-range order in a medium-entropy alloy** [28 April 2021]
Article • Direct experimental evidence of chemical short-range atomic-scale ordering (CSRO) in a VCoNi medium-entropy alloy is provided via diffraction and electron microscopy, analysed from specific crystallographic directions.
- **Strongly reducing magnesium(0) complexes** [28 April 2021]
Article • Strongly reducing β -diketiminate complexes containing magnesium in its zero oxidation state are reported, among which is a compound with a linear triatomic Mg–Mg–Mg core.
- **Direct assessment of the acidity of individual surface hydroxyls** [28 April 2021]
Article • Non-contact atomic force microscopy measurements are used to probe the hydrogen bond strength of individual surface hydroxyl groups and determine their acidity with atomic precision.
- **Accelerated global glacier mass loss in the early twenty-first century** [28 April 2021]
Article • Analysis of satellite stereo imagery uncovers two decades of mass change for all of Earth's glaciers, revealing accelerated glacier shrinkage and regionally contrasting changes consistent with decadal climate variability.
- **Plume-driven reactivation of deep continental lithospheric mantle** [28 April 2021]
Article • Upwelling of mantle plumes is proposed as a mechanism for craton healing after substantial disruption of their roots, enabling them to return to their original lithospheric thickness.
- **Towards complete and error-free genome assemblies of all vertebrate species** [28 April 2021]
Article • The Vertebrate Genome Project has used an optimized pipeline to generate high-quality genome assemblies for sixteen species (representing all major vertebrate classes), which have led to new biological insights.
- **Universal nomenclature for oxytocin–vasotocin ligand and receptor families** [28 April 2021]
Article • A revised, universal nomenclature for the vertebrate genes that encode the oxytocin and vasopressin–vasotocin ligands and receptors will improve our understanding of gene evolution and facilitate the translation of findings across species.
- **Platypus and echidna genomes reveal mammalian biology and evolution** [06 January 2021]

Article • New reference genomes of the two extant monotreme lineages (platypus and echidna) reveal the ancestral and lineage-specific genomic changes that shape both monotreme and mammalian evolution.

- **REV-ERB in GABAergic neurons controls diurnal hepatic insulin sensitivity** [24 March 2021]

Article • REV-ERB in GABAergic neurons orchestrates the rhythmic sensitivity of hepatic glucose production to insulin-mediated suppression that peaks at wakening, with implications in the extended dawn phenomenon.

- **A biosensor for the direct visualization of auxin** [07 April 2021]

Article • A genetically encoded sensor for the quantitative visualization of auxin distribution in plants enables real-time monitoring of its uptake and clearance by individual cells and within cellular compartments.

- **Structural and biochemical mechanisms of NLRP1 inhibition by DPP9** [17 March 2021]

Article • The 2:1 complex between the inflammatory mediator NLRP1 and the dipeptidyl peptidase DPP9 functions to sequester an inflammasome-forming fragment of NLRP1 and enhance NLRP1 autoinhibition.

- **DPP9 sequesters the C terminus of NLRP1 to repress inflammasome activation** [17 March 2021]

Article • Structures of NLRP1–DPP9 alone and with a small-molecule inhibitor of DPP9 reveal the mechanisms through which NLRP1 is regulated, providing insights into the role of this complex in inflammasome regulation.

- **High CO₂ levels drive the TCA cycle backwards towards autotrophy** [21 April 2021]

Article • In the deltaproteobacterium *Hippea maritima*, the tricarboxylic acid (TCA) cycle can be reversed by high partial pressures of CO₂ for the autotrophic fixation of carbon.

- **CRL4AMBRA1 is a master regulator of D-type cyclins** [14 April 2021]

Article • Biochemical and genetics studies identify CRL4AMBRA1 as the ubiquitin ligase that has a key role in regulating the stability of D-type cyclins during cell-cycle progression.

- **The AMBRA1 E3 ligase adaptor regulates the stability of cyclin D** [14 April 2021]

Article • AMBRA1 is the main regulator of the degradation of D-type cyclins, and loss of AMBRA1 promotes cell proliferation and tumour growth, and reduces the sensitivity of cancer cells to inhibition of CDK4 and CDK6.

- **AMBRA1 regulates cyclin D to guard S-phase entry and genomic integrity** [14 April 2021]
Article • AMBRA1-mediated degradation of cyclin D through CRL4–DDB1 regulates cell proliferation and prevents replication stress in neurodevelopment and cancer.
- **Quantifying forest change in the European Union** [28 April 2021]
Matters Arising •
- **Concerns about reported harvests in European forests** [28 April 2021]
Matters Arising •
- **Reply to Wernick, I. K. et al.; Palahí, M. et al.** [28 April 2021]
Matters Arising •

NEWS AND VIEWS

28 April 2021

Highly reactive form of magnesium stabilized by bulky ligands

Magnesium atoms typically lose two electrons to form chemical compounds. A reactive complex has finally been made in which magnesium keeps all of its electrons, and which can be thought of as a soluble form of the metal.

- [Cameron Jones](#) [ORCID: http://orcid.org/0000-0002-7269-1045](http://orcid.org/0000-0002-7269-1045) ⁰

1. [Cameron Jones](#)

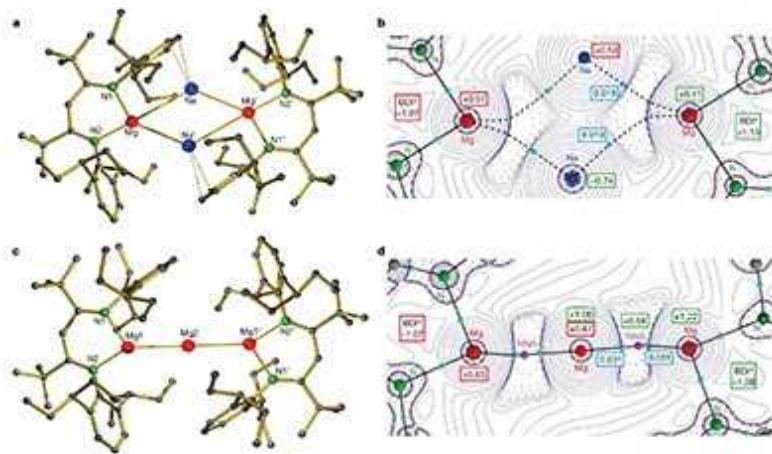
1. Cameron Jones is at the School of Chemistry, Monash University, Melbourne, Victoria 3800, Australia.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

In the past few decades, innovative strategies have been developed to stabilize chemical compounds that contain main-group elements in unusually low oxidation states. (Main-group elements are those on the left and right sides of the periodic table, surrounding the transition metals.) Many of these compounds were previously thought to be impossible to obtain as stable materials. This is especially the case for magnesium, a reactive metal whose compounds almost all contain the element in the +2 oxidation state (equivalent to the atoms losing two electrons and having a charge of +2). [Writing in Nature](#), Rösch *et al.*¹ report that magnesium can form compounds in which it keeps all of its electrons, and thereby exists in the zero oxidation state that is normally characteristic of the pure element. These magnesium(0) compounds are stable at room temperature, but are

highly reactive, with the metal atoms readily surrendering electrons in chemical reactions that are potentially useful for organic synthesis.



[Read the paper: Strongly reducing magnesium\(0\) complexes](#)

It is well established that transition metals can exist in a range of oxidation states in their compounds, including the zero oxidation state. This characteristic is central to the ubiquitous application of such compounds as industrial catalysts. By contrast, the highly reactive *s*-block metals (groups 1 and 2 of the periodic table, excluding hydrogen) were not thought to be capable of forming similar, stable low-oxidation-state compounds, because they tend to lose all of their outermost electrons during compound formation.

This idea was overturned in 2007, when the first compounds of magnesium — a group 2 metal — were prepared in the +1 oxidation state². These compounds consist of two magnesium atoms connected by a chemical bond, with each magnesium atom bound to a bulky organic ligand; the ligand acts as a protective coating that prevents the magnesium ions from reverting to the +2 oxidation state. Other examples of compounds containing group-2 metals in low oxidation states have since been reported, such as beryllium in the zero³ and +1 oxidation states⁴, and calcium in the +1 state⁵, but they typically require special electron-accepting ligands to stabilize the compound concerned.

Such advances in the chemistry of *s*-block metals have inspired chemists to pursue one of the outstanding goals in the field: the preparation of stable magnesium(0) compounds⁶. The logical approach would be to react an extremely bulky magnesium(ii) precursor compound with a powerful reducing agent, such as sodium or potassium metal. The idea is that the reducing agent would deliver two electrons to the magnesium atom, thus forming a magnesium(0) compound.

In pursuit of this goal, researchers from the same laboratory as Rösch *et al.* previously developed an extraordinarily bulky ligand (abbreviated as BDI*), which they used to prepare a magnesium(ii) precursor compound⁷. However, when they reacted the precursor with potassium metal, only one electron was donated to the magnesium atoms, resulting in the formation of the magnesium(i) compound $(BDI^*)Mg-Mg(BDI^*)$. Rösch and colleagues now report that, when they change the reducing reagent to a finely divided (powdered) form of sodium metal, two electrons are donated to the precursor, yielding the remarkable magnesium(0) compound $\{[(BDI^*)Mg^-][Na^+]\}_2$ (Fig. 1).

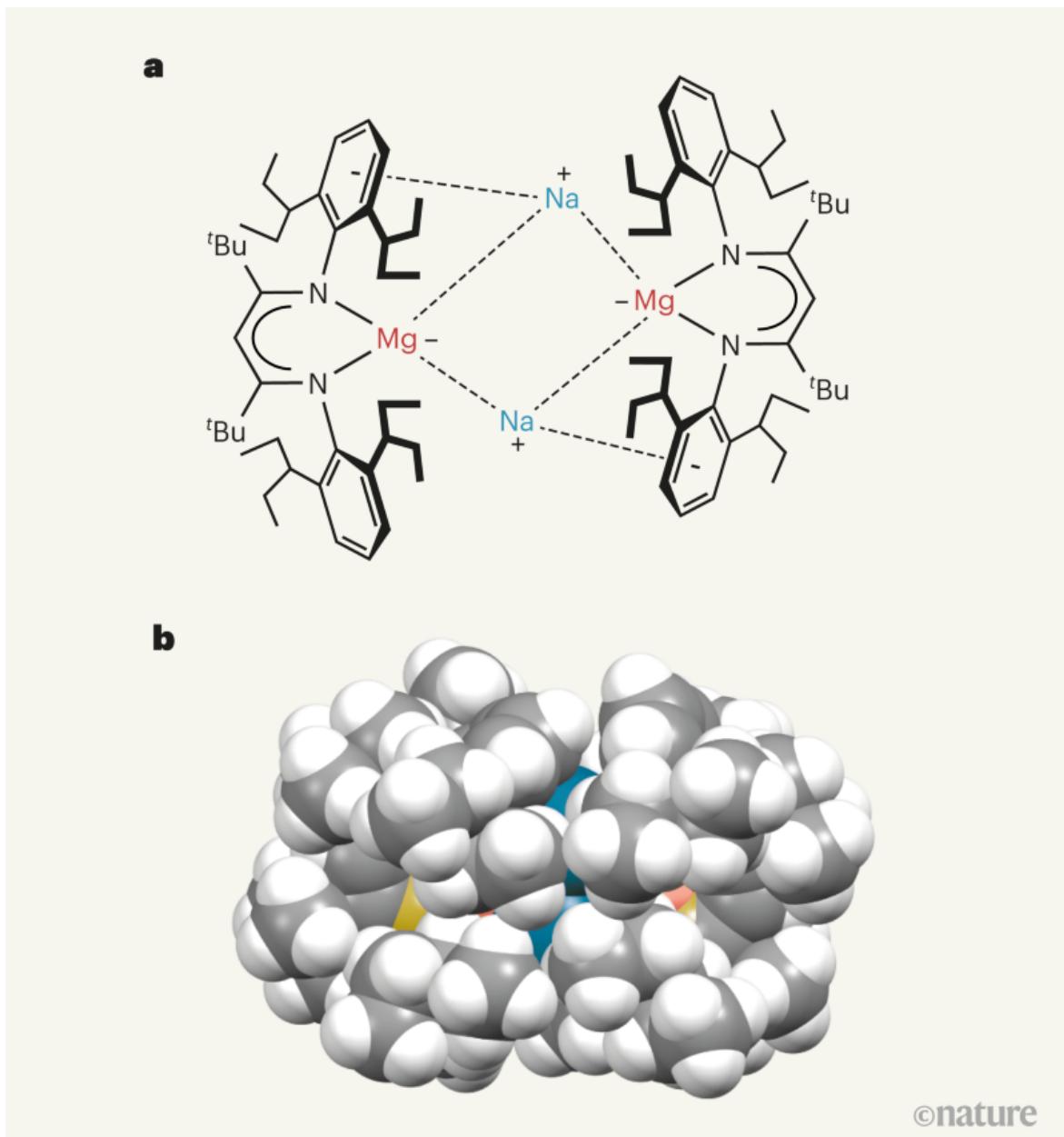
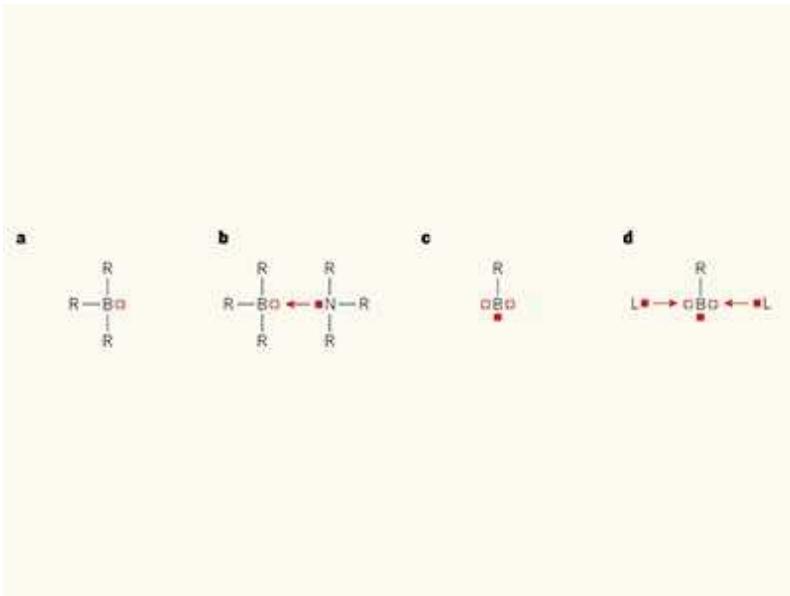


Figure 1 | A complex containing magnesium in the zero oxidation state. Almost all compounds of the reactive metal magnesium contain the metal in its +2 oxidation state. **a**, Rösch *et al.*¹ report the first examples of compounds with magnesium (Mg) in the zero oxidation state, one example of which is shown here. The magnesium(0) atoms are enshrouded by extremely bulky organic ligands, which protect the magnesium centres from decomposition processes. Sodium (Na) ions are also present in the molecule. Broken lines indicate non-covalent bonding interactions between metals and between metals and benzene rings in the molecule. ^tBu represents a *t*-butyl

group, $\text{C}(\text{CH}_3)_3$. Bonds in bold project upwards from the plane of the figure; semicircular lines represent delocalized electrons. **b**, This space-filling model of the molecule shows how the metal atoms are almost completely enclosed by the ligands. Carbon atoms, grey; hydrogen, white; nitrogen, yellow; magnesium, red; sodium, blue.

It is clear that the enormous size of the BDI* ligand is required to stabilize the magnesium(0) compound, preventing the solid form of the compound from decomposing at room temperature. Even so, the compound partially decomposes at room temperature when in solution. This process yields another unprecedented type of compound known as a magnesium cluster, in which three magnesium atoms are bonded together, $(\text{BDI}^*)\text{Mg}-\text{Mg}-\text{Mg}(\text{BDI}^*)$. Astonishingly, this contains magnesium in both the 0 and +1 oxidation states. Rösch *et al.* reasonably posit that the formation of this compound could shed light on the unknown mechanism of formation of Grignard reagents — an important class of magnesium(ii) compound that has been widely used in organic chemistry for more than 120 years.

The size and 3D shape of the BDI* ligand hits the sweet spot when it comes to stabilizing the magnesium(0) compound. The authors report that the compound consists of a central core of magnesium and sodium atoms — $[\text{Mg}_2\text{Na}_2]^{2+}$ — arranged in a ring and enveloped by two BDI* ligands (Fig. 1). Computational analyses reveal that the magnesium atoms have the same number of electrons as magnesium metal, which means that the compound could be viewed as a soluble form of the metal. There is, however, some sharing of electrons between the magnesium and sodium atoms. This doesn't detract from the assignment of the oxidation state of the magnesium atoms as zero, and the observation of a magnesium–sodium ‘bond’ in the compound is itself another first.



Peculiar boron startles again

Given that the magnesium atoms are in the zero oxidation state, the compound should display a level of reactivity similar to that of the elemental metal. In fact, Rösch and co-workers' preliminary experiments show that it is even more reactive than that. For example, it can readily activate (break or weaken) very strong bonds, such as hydrogen–hydrogen and carbon–fluorine bonds, at room temperature. Many other compounds that contain main-group elements in low oxidation states can do the same⁸. A true demonstration of the exceptional reducing ability of the magnesium(0) compound would be the activation of even more staunchly inert molecules, such as dinitrogen (N_2). This seems achievable, given the recent demonstration that dinitrogen can be activated by a transiently formed calcium(i) compound⁹.

A more surprising aspect of the reactivity of Rösch and colleagues' compound is that its magnesium(0) atoms can transfer electrons to its sodium atoms, reducing them back to sodium metal. This seems counter-intuitive, because the reverse process — the reduction of magnesium(ii) to magnesium(0) by sodium metal — was used to make the magnesium(0) compound in the first place. The authors' experimental evidence backs up the observation that the sodium atoms are reduced, but more work is required to examine the processes by which this operates.

Rösch and co-workers' stable magnesium(0) compound is a landmark in the chemistry of the *s*-block elements. It will fundamentally change chemists' views about what can be synthesized using these elements. Moreover, it will help to advance our understanding of — and raise questions about — the unusual 'non-classical' bonding in low-oxidation-state main-group compounds. The development of highly reducing magnesium(0) compounds might also pave the way to their use in chemical reactions that, at present, cannot normally be carried out with *s*-block metals. The future is surely bright for magnesium now that it has hit zero.

Nature **592**, 687–688 (2021)

doi: <https://doi.org/10.1038/d41586-021-01014-x>

References

1. 1.

Rösch, B. *et al.* Nature **592**, 717–721 (2021).

2. 2.

Green, S. P., Jones, C. & Stasch, A. Science **318**, 1754–1757 (2007).

3. 3.

Arrowsmith, M. *et al.* Nature Chem. **8**, 890–894 (2016).

4. 4.

Wang, G. *et al.* J. Am. Chem. Soc. **142**, 4560–4564 (2020).

5. 5.

Kriek, S., Görls, H., Yu, L., Reiher, M. & Westerhausen M. J. Am. Chem. Soc. **131**, 2977–2985 (2009).

6. 6.

Jones, C. *Commun. Chem.* **3**, 159 (2020).

7. 7.

Rösch, B. *et al. Chem. Commun.* **56**, 11402–11405 (2020).

8. 8.

Weetman, C. & Inoue, S. *ChemCatChem* **10**, 4213–4228 (2018).

9. 9.

Rösch, B. *et al. Science* **371**, 1125–1128 (2021).

Jobs from Nature Careers

- - - [All jobs](#)
 - - [Research Associate / Postdoc](#)
 - [Technische Universität Dresden \(TU Dresden\)](#)
[Dresden, Germany](#)
[JOB POST](#)
 - [PhD position in evolutionary ecology of insect immune priming](#)
 - [Université Bourgogne Franche-Comté \(UBFC\)](#)
[DIJON, France](#)
[JOB POST](#)

- **PhD contract Doctoral contract in sociology-demography : Analysis of differential socio-demographic behaviors by sex among 50-74 year olds**

Université Bourgogne Franche-Comté (UBFC)

DIJON, France

JOB POST

- **PhD - Dark ecological networks and land management: methodological and strategic contributions for reducing ecological impacts of artificial light**

Université Bourgogne Franche-Comté (UBFC)

BESANCON, France

JOB POST

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01014-x>

NEWS AND VIEWS

21 April 2021

Life in a carbon dioxide world

Microorganisms living in hydrothermal vents that emit carbon dioxide gas provide a striking example of metabolic finesse. This pathway sheds light on microbial ecology in extreme environments and offers clues to early life on Earth.

- [Martina Preiner](#) ⁰ &
- [William F. Martin](#) ¹

1. [Martina Preiner](#)

1. Martina Preiner is at the Institute of Molecular Evolution, Heinrich Heine University Düsseldorf, Düsseldorf 40225, Germany.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. [William F. Martin](#)

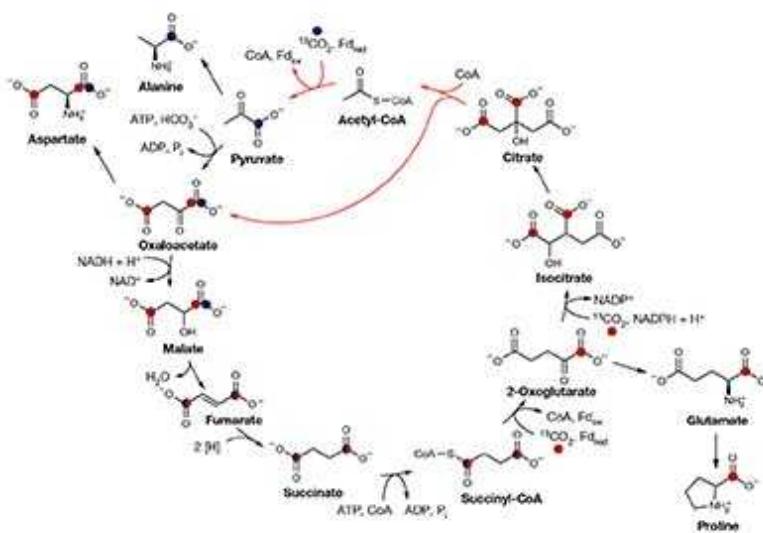
1. William F. Martin is at the Institute of Molecular Evolution, Heinrich Heine University Düsseldorf, Düsseldorf 40225, Germany.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Few chemicals have hit the headlines so widely that everyone knows their formula, but carbon dioxide is an exception. It is so crucial for understanding climate change that we recognize its shorthand name of CO₂

as a threat to our future. For most microbes, however, CO₂ looks more like a feast than a threat. Microbes have tools at their disposal — CO₂-fixation pathways — that enable them to incorporate CO₂ into their cell mass. These pathways are essential for life because all ecosystems on Earth ultimately depend on cells that make organic material from CO₂. [Writing in Nature](#), Steffens *et al.*¹ uncover key details about an ingenious pathway that enables bacteria to thrive in a hydrothermal environment surrounded by gases consisting mainly of CO₂.



[Read the paper: High CO₂ levels drive the TCA cycle backwards towards autotrophy](#)

Steffens and colleagues studied *Hippea maritima* bacteria. These microorganisms shun oxygen, love temperatures near 60°C, and obtain energy from the reaction of hydrogen gas (H₂) with sulfur to make hydrogen sulfide (H₂S). As with all life forms, they need a carbon source to grow. And, like many, they can choose this source depending on what is available in their environment. If a rich diet of protein is on offer, *H. maritima* incorporate this as a building block into their metabolic pathways for growth.

But if *H. maritima* grow in the presence of CO₂ concentrations of 40% (1,000 times higher than atmospheric CO₂ levels), they do some ‘chemical engineering’, using a pathway called the reversed oxidative tricarboxylic acid cycle. That might sound complicated, but it is connected to something familiar — human nutrition. After the food we eat is broken down in the gut, our cells convert the sugars, fats and proteins contained in the food into energy and CO₂ using a pathway called the tricarboxylic acid (TCA) cycle. This is also called the Krebs cycle, after the scientist who discovered it². The TCA cycle is used by nearly all life forms, but it can run backwards in some bacteria³: this change of direction, to give the reversed oxidative TCA cycle (Fig. 1), invests energy that converts CO₂ into amino acids, sugars and lipids.

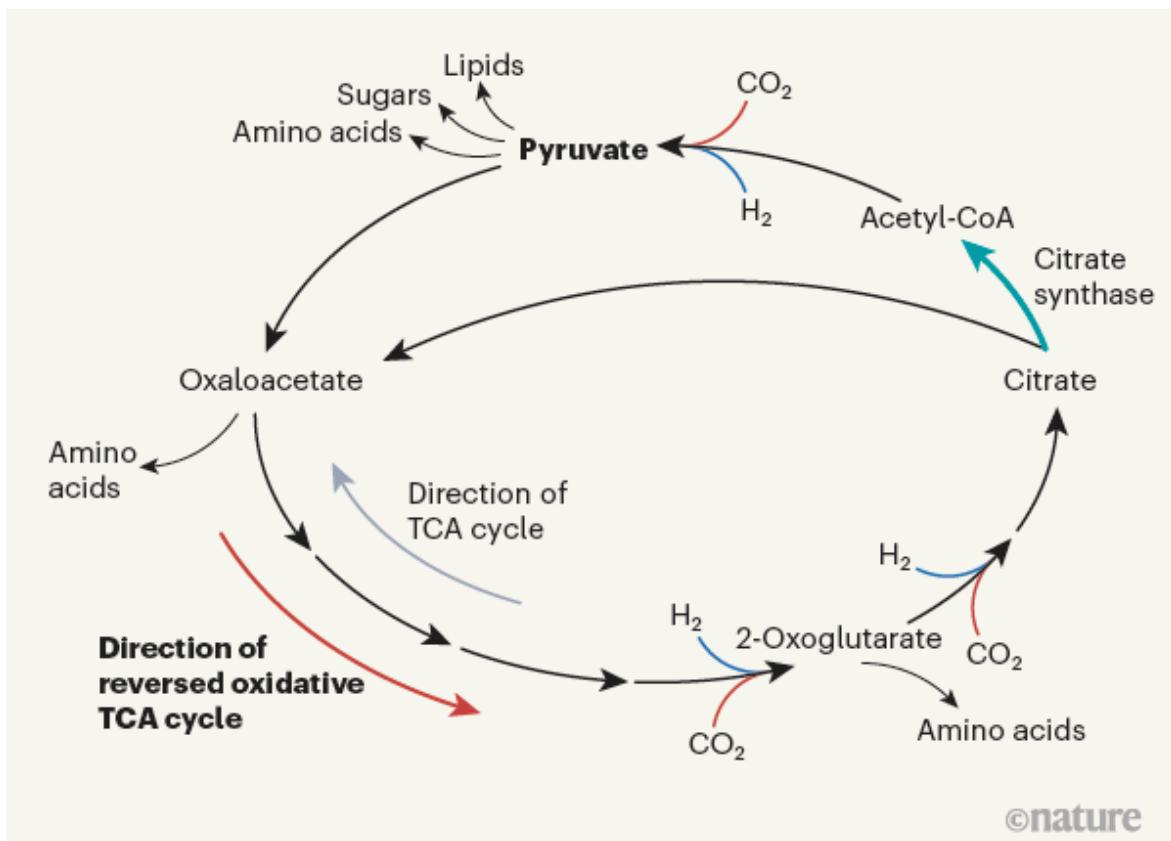


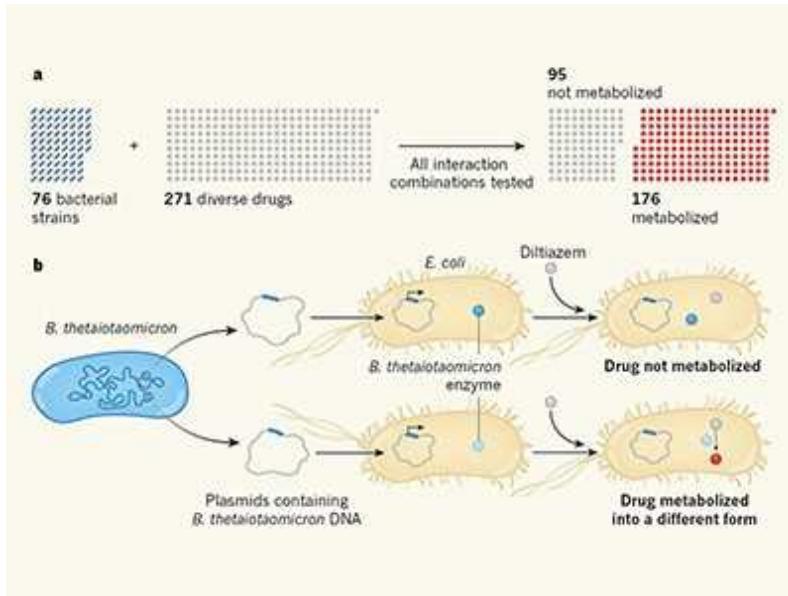
Figure 1 | The reversed oxidative tricarboxylic acid (TCA) cycle. Almost all life forms use the TCA cycle to convert molecules such as amino acids, sugars and lipids into energy and carbon dioxide by means of a pathway that involves molecules such as pyruvate, oxaloacetate, 2-oxoglutarate, citrate

and acetyl coenzyme A (acetyl-CoA). Some bacteria can run this cycle in the reverse direction (it's then called the reversed oxidative TCA cycle), incorporating CO₂ and hydrogen (H₂) to form molecules such as amino acids, sugars and lipids. Steffens *et al.*¹ used approaches such as tracking labelled carbon atoms to reveal the mechanism that enables the bacterium *Hippea maritima* to run the TCA cycle backwards. The authors report that a high level of the enzyme citrate synthase is key to pathway reversal.

What enables the TCA cycle to run in reverse under specific growth conditions has been a mystery, until now. Steffens *et al.* show that *H. maritima*'s secret trick is to adjust levels of a crucial enzyme in an unexpected way, so as to be ready to assimilate high concentrations of CO₂ before they are encountered. This generates an elegant harmony between the microbe's environment and its metabolism.

H. maritima uses the reversed oxidative TCA cycle when high levels of CO₂ are present, and this is where the technical brilliance of Steffens and colleagues' investigation becomes evident. The authors fed the bacteria amino acids and CO₂ labelled with the ¹³C isotope of carbon. Both of these food sources were channelled into the reversed oxidative TCA cycle.

Tracking ¹³C accumulation in the intermediate molecules of this pathway in growing cells enabled the authors to uncover which carbon source the cells used down which route of the pathway. It also enabled them to determine how many full 'turns' of the reversed oxidative TCA cycle occur as carbon is assimilated.



Microbes make metabolic mischief by targeting drugs

This revealed that *H. maritima* preferentially uses CO₂ as its carbon source, but only when CO₂ is in abundant supply. To enable the TCA cycle to run backwards in response to high levels of CO₂, the cells harbour huge amounts of the enzyme citrate synthase. A high level of citrate synthase makes it easier to generate acetyl coenzyme A (acetyl-CoA) molecules, which exit the reversed oxidative TCA cycle by forming pyruvate, which is converted to lipids, sugars and amino acids (Fig. 1). This in turn, invites CO₂ to enter the cycle. In this way, high environmental CO₂ levels push the cycle in the direction of converting CO₂ to acetyl-CoA.

This would cause a logjam at the acetyl-CoA stage of the cycle were it not for the high levels of CO₂. The main connection between the reversed oxidative TCA cycle and other metabolic pathways is the molecule pyruvate, which is made by a reaction involving CO₂ and acetyl-CoA. That reaction, like the two other reactions that incorporate CO₂ in this cycle, is reversible and can run in either direction. A high CO₂ concentration — typically expressed as high partial pressure relative to the total pressure of all the gases present — pushes all three of these reactions forwards. The whole pathway is thereby pushed in the direction of pyruvate production, as long as there is no bottleneck at the reaction catalysed by citrate synthase.

High amounts of that enzyme avert this potential bottleneck, and keep cells poised to exploit high levels of CO₂ if the environment provides them.

Ecosystems with high CO₂ harbour many environments in which resident microbes have genes that encode enzymes of the reversed oxidative TCA cycle, as metagenomics analysis (genome sequencing of microbial communities) has indicated⁴. However, the presence of genes alone cannot reveal in which direction cells are using a pathway because the environment can dictate the flow of substrates, as this exquisitely detailed example of *H. maritima* underscores.



Energy from thin air

Hippea maritima is not the only known example of a bacterium with reversible metabolism. Another example is the bacterium *Thermacetogenium phaeum*, which grows under conditions similar to those that support *H. maritima* (high CO₂ and an absence of oxygen), but in industrial cellulose-processing reactors⁵. If the environment offers abundant H₂ and CO₂, *T. phaeum* grows using these to make the molecule acetate. However, if those gases become scarce and acetate is abundant, the microbe's main metabolic reaction runs backwards⁵, and it survives on the conversion of acetate to H₂ and CO₂. How it achieves this is unknown. Looking at the genes that microbes use in a given environment can reveal

important clues to the secrets of life in microbial communities⁶. But to really understand the chemical reactions that support microbial life, there is no substitute for studies such as those by Steffens and colleagues, which show us, carbon atom by carbon atom, what cells are doing with the substrates that their environment presents.

Individual microbes, such as *H. maritima*, and even whole ecosystems, can thrive from the energy supplied by the reaction of H₂ with CO₂. This not only offers examples of fascinating microbial ecology, but also provides a window into the ancient past, by presenting strategies for growth in conditions thought to be similar to those that the first microbes on Earth encountered³. Those pioneering microbes had to be able to survive on a diet of CO₂ because it was the carbon source that the early Earth had available⁷.

Yet CO₂ is only half of the story. To convert CO₂ into organic compounds, microbes need a source of energy and electrons. For the first ecosystems on Earth, and for *H. maritima* today, the source of chemical energy and electrons for CO₂ fixation is H₂. For four billion years, microbes have been living from the energy provided by the vast amounts of H₂ that Earth's crust constantly generates⁸. Given the effort that *H. maritima* invests in making pyruvate from H₂ and CO₂, it seems almost unimaginable that the very first biochemical pathways could have got going before there were enzymes to assist the carbon-fixing reactions. Yet, surprisingly, H₂ and CO₂ can form pyruvate overnight without any enzyme involvement if simple metal catalysts, of the kinds found in the oceanic crust, are provided⁹. This suggests that early metabolism on Earth was built around the naturally occurring chemistry between CO₂ and H₂ in mineral-rich environments¹⁰.

The chemical reactions that underpin the lifestyle of *H. maritima* thus hark back to a time when the first cells lived in a world of carbon dioxide. By investigating cells that still inhabit such realms today, we can discover some clues about the life and times of the most ancient microbial ancestors.

Nature **592**, 688-689 (2021)

doi: <https://doi.org/10.1038/d41586-021-00977-1>

References

1. 1.
Steffens, L. *et al.* *Nature* **592**, 784–788 (2021).
2. 2.
Krebs, H. A. & Martin, A. *Hans Krebs, Reminiscences and Reflections* (Clarendon, 1981).
3. 3.
Fuchs, G. *Annu. Rev. Microbiol.* **65**, 631–658 (2011).
4. 4.
Sharon, I. & Banfield, J. F. *Science* **342**, 1057–1058 (2013).
5. 5.
Keller, A., Schink, B. & Müller, N. *Front. Microbiol.* **10**, 2785 (2019).
6. 6.
Orsi, W. D. *Nature Rev. Microbiol.* **16**, 671–683 (2018).
7. 7.
Sossi, P. A. *et al.* *Sci. Adv.* **6**, eabd1387 (2020).
8. 8.
Sleep, N. H., Bird, D. K. & Pope, E. C. *Phil. Trans. R. Soc. B* **366**, 2857–2869 (2011).
9. 9.
Preiner, M. *et al.* *Nature Ecol. Evol.* **4**, 534–542 (2020).

10. 10.

Baross, J. A. *Nature* **564**, 42–43 (2018).

Jobs from Nature Careers

•

◦

- [All jobs](#)

◦

- [Research Associate / Postdoc](#)

[Technische Universität Dresden \(TU Dresden\)](#)

[Dresden, Germany](#)

[JOB POST](#)

- [PhD position in evolutionary ecology of insect immune priming](#)

[Université Bourgogne Franche-Comté \(UBFC\)](#)

[DIJON, France](#)

[JOB POST](#)

- [PhD contract Doctoral contract in sociology-demography : Analysis of differential socio-demographic behaviors by sex among 50-74 year olds](#)

[Université Bourgogne Franche-Comté \(UBFC\)](#)

[DIJON, France](#)

JOB POST

- **PhD - Dark ecological networks and land management: methodological and strategic contributions for reducing ecological impacts of artificial light**

Université Bourgogne Franche-Comté (UBFC)

BESANCON, France

JOB POST

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-00977-1>

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

NEWS AND VIEWS

14 April 2021

The path to destruction for D-type cyclin proteins

The molecular mechanism governing the destruction of key cell-cycle proteins, D-type cyclins, has been elucidated. This mechanism might underlie the lack of response of some human tumours to an inhibitor treatment.

- [Charupong Saengboonmee](#)⁰ &
- [Piotr Sicinski](#)¹

1. Charupong Saengboonmee

1. Charupong Saengboonmee is in the Department of Cancer Biology, Dana-Farber Cancer Institute, and in the Department of Genetics, Blavatnik Institute, Harvard Medical School, Boston, Massachusetts 02215, USA; and in the Department of Biochemistry, Faculty of Medicine and Cholangiocarcinoma Research Institute, Khon Kaen University, Khon Kaen, Thailand.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Piotr Sicinski

1. Piotr Sicinski is in the Department of Cancer Biology, Dana-Farber Cancer Institute, and in the Department of Genetics, Blavatnik Institute, Harvard Medical School, Boston, Massachusetts 02215, USA.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Proteins called D-type cyclins (cyclin D1, D2 and D3) are key components of the core cell-cycle engine, which drives cellular division. Writing in *Nature*, [Simoneschi et al.](#)¹, [Chaikovsky et al.](#)² and [Maiani et al.](#)³ provide the long-sought answer to how D-type cyclins are normally degraded.

During cell division, D-type cyclins bind to and activate their enzymatic partners, termed cyclin-dependent kinase 4 (CDK4) and cyclin-dependent kinase 6 (CDK6). These cyclin–CDK4/6 kinases add phosphate groups to the tumour-suppressor proteins RB1, RBL1 and RBL2, and this consequently propels cell division (Fig. 1). The uncontrolled activation of cyclin D–CDK4/6 kinases is a driving force in the development of many types of cancer⁴.

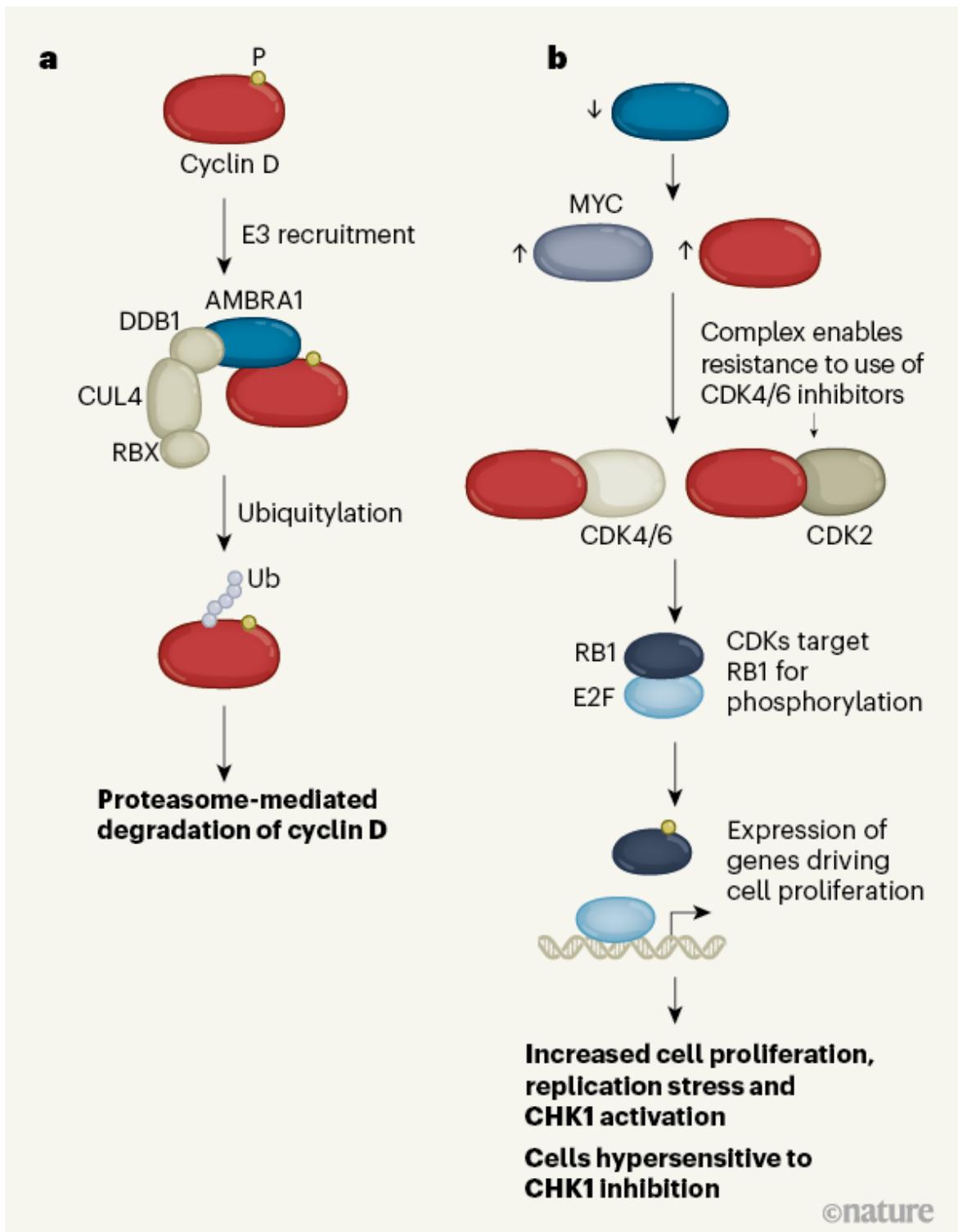
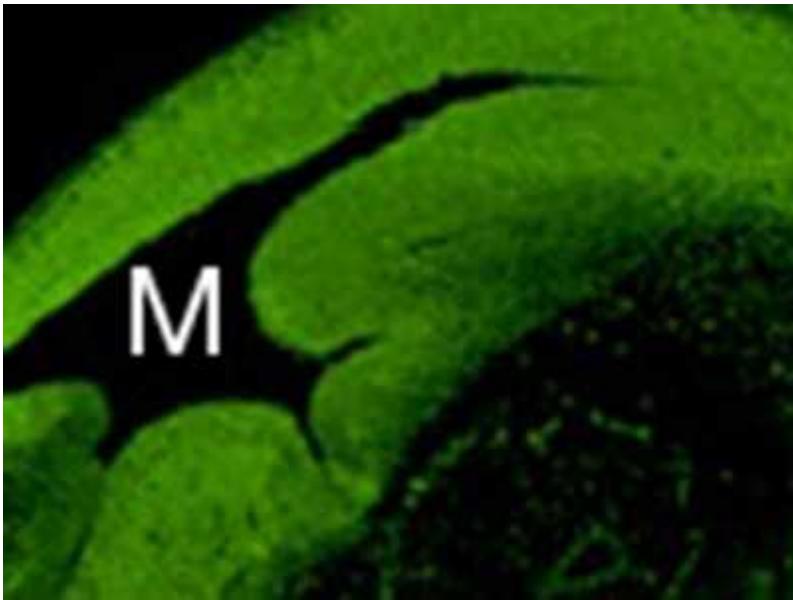


Figure 1 | The destruction of cyclin D protein. Simoneschi *et al.*¹, Chaikovsky *et al.*² and Maiani *et al.*³ reveal how this key cell-cycle regulator is degraded in mammalian cells. **a**, Adding a phosphate group (P)

targets cyclin D for degradation mediated by an E3 protein complex. This complex is called CRL4, and it includes the proteins cullin 4 (CUL4), RBX and DDB1. The protein AMBRA1 targets cyclin D to the E3. The E3 attaches chains of ubiquitin proteins (Ub) to cyclin D, marking it for degradation by a protein complex called the proteasome. **b**, AMBRA1 depletion leads to a rise in the levels of cyclin D and MYC proteins. Cyclin D binds to its usual partners, the enzymes CDK4 and CDK6, which phosphorylate the protein RB1. Phosphorylated RB1 releases E2F transcription factors to drive the expression of genes required for cell-cycle progression. In AMBRA1-depleted cells, cyclin D also forms complexes with the enzyme CDK2. These complexes enable cancer cells to resist treatment by CDK4/6 inhibitors. Cell proliferation driven by high cyclin D results in DNA damage, replication stress and activation of the enzyme CHK1. AMBRA1-depleted cells are hypersensitive to treatment with CHK1 inhibitors.

Interest in cyclin D–CDK4/6 biology has grown lately as small-molecule inhibitors of CDK4/6 kinases have entered the clinic. Remarkable results from clinical trials of these compounds demonstrate their ability to extend the survival times of people with breast cancer. The CDK4/6 inhibitors palbociclib, ribociclib and abemaciclib are approved for the treatment of advanced breast cancers. Moreover, these drugs are being tested in several hundred clinical trials for many different types of cancer^{4,5}.

D-type cyclins have been intensively studied since their discovery in the 1990s, but how they are degraded during the cell cycle has remained controversial. The phosphorylation of the carboxy-terminal region of D-cyclins triggers the destruction of these proteins by a degradation pathway called the ubiquitin–proteasome system⁶. In this system, a cascade of activity by ubiquitin-activating (E1), ubiquitin-conjugating (E2) and ubiquitin-ligase (E3) enzymes attaches chains of several molecules of a small protein, ubiquitin, to target proteins, by a process called ubiquitylation. These ubiquitylated proteins are then destined for degradation in a protein complex called the proteasome.

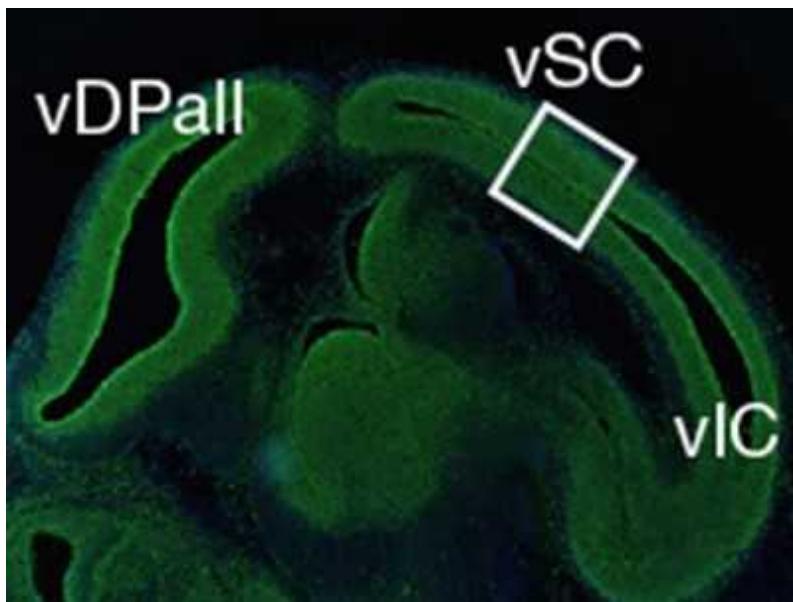


[Read the paper: CRL4AMBRA1 is a master regulator of D-type cyclins](#)

The largest family of E3s are cullin–RING ligases (CRLs). CRLs consist of a cullin protein, a RING protein (which recruits an E2), an adaptor protein and one of the many different substrate-receptor proteins responsible for recruiting the target protein to the E3 complex^{7–9}. Several substrate receptors for an E3 called CRL1 have been implicated in the degradation of cyclin D1, whereas others were postulated to target cyclins D2 and D3 for proteasomal destruction. In addition, cyclin D1 was shown to be ubiquitylated by the anaphase-promoting complex, an E3 complex that targets several cell-cycle proteins⁹. In contrast to those models, other research¹⁰ demonstrated that the level and stability of cyclin D1 was unaffected by depletion of these proteins, indicating that some other E3 regulates the degradation of cyclin D1.

The papers by Simoneschi, Chaikovsky, Maiani and their respective colleagues report that three D-type cyclins are ubiquitylated and targeted for proteasomal degradation by the E3 termed CRL4, which uses the protein AMBRA1 as its substrate receptor. It was already known that AMBRA1 has a key role in regulating autophagy, a process which cells degrade damaged organelles or protein aggregates^{11,12}. AMBRA1 had also been identified as a substrate receptor of E3s, including CRL4^{7,13}. Through a beautiful series of experiments, using techniques from fields including cell biology, molecular

biology and developmental genetics to provide insights in both mice and humans, Simoneschi, Chaikovsky and Maiani *et al.* demonstrate that AMBRA1 depletion in normal cells and cancer cells, as well as in developing mouse embryos, leads to a rise in the level of D-type cyclins. This results in greater phosphorylation of RB1 and more cell proliferation than occurs in cells that have normal amounts of AMBRA1.

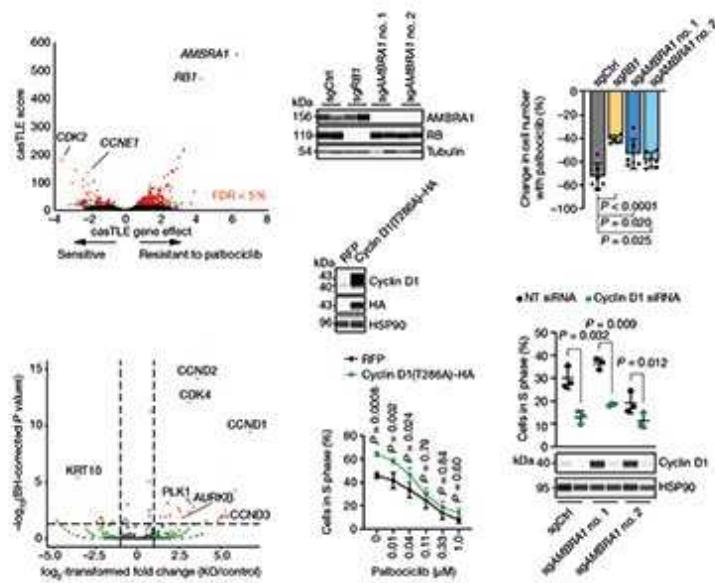


[Read the paper: AMBRA1 regulates cyclin D to guard S-phase entry and genomic integrity](#)

Maiani and colleagues also show that AMBRA1 depletion elevates the levels of the transcription-factor protein N-MYC. This group previously showed¹⁴ that AMBRA1 regulates the stability and activity of a related transcription factor, c-MYC. MYC-family proteins can upregulate the expression of D-type and E-type cyclin proteins¹⁵, thereby accelerating cell-cycle progression.

These observations suggest that AMBRA1 might act as a tumour-suppressor protein. Indeed, mice with only one functional copy of the gene encoding AMBRA1 have a higher incidence of lung, liver and kidney tumours than do normal mice with two functional copies of the gene¹⁴. These new studies provide compelling evidence to support this idea.

The authors demonstrate that the *AMBRA1* gene is mutated in human cancers. As expected, given the ability of *AMBRA1* to promote the degradation of cyclin D1, the authors report that the level of *AMBRA1* in human tumours was inversely correlated with the level of cyclin D1. Moreover, low levels of *AMBRA1* in tumours are associated with poor prognosis for people with cancer. The experimental inactivation of *AMBRA1* in either human tumour cell lines or mouse cells engineered to have cancer-promoting mutations increased the tumour-forming potential of the cells, as assessed after injection into mice. Furthermore, the genetic ablation of *AMBRA1* boosted tumour formation in a mouse model of lung cancer driven by a mutant version of the *Kras* gene, and these *AMBRA1*-deficient tumours had a higher than normal level of D-cyclins. Collectively, these studies indicate that *AMBRA1* normally restrains cell proliferation, largely by stopping D-cyclins from reaching high levels.



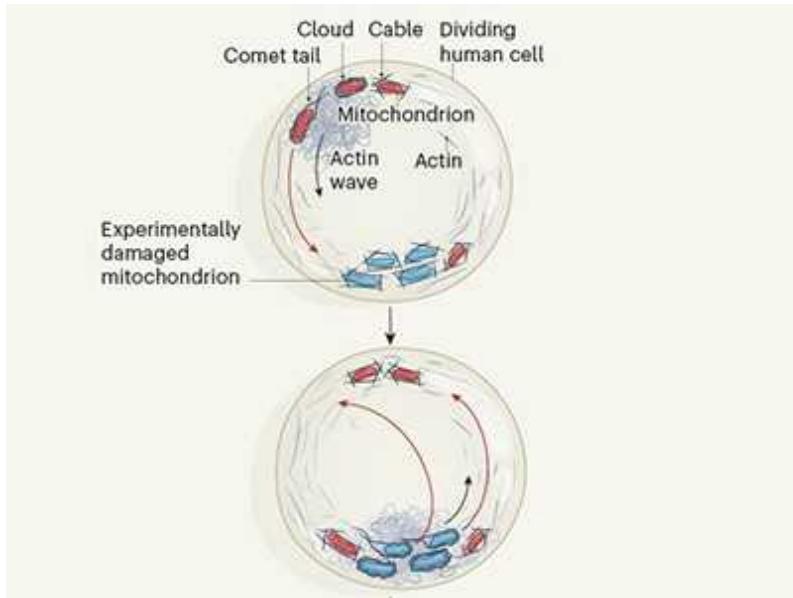
[Read the paper: The *AMBRA1* E3 ligase adaptor regulates the stability of cyclin D](#)

In addition, Chaikovsky, Simoneschi and their respective colleagues demonstrate that the loss of *AMBRA1*, and the concomitant increase in D-cyclins, results in decreased sensitivity of human tumour cells to CDK4/6 inhibitors. Intriguingly, these authors report that, in *AMBRA1*-depleted cells, rather than mainly partnering with CDK4/6, cyclin D1 also forms a

catalytically active complex with the cyclin-dependent kinase enzyme CDK2, and that these complexes are insensitive to CDK4/6 inhibitors.

Maiani and colleagues also show that loss of AMBRA1, and the resulting rise of D-cyclins (and possibly also of other proteins, such as c-MYC), triggers DNA damage and replication stress, which is accompanied by the activation of a kinase enzyme called CHK1. Importantly, Maiani *et al.* report that AMBRA1-depleted cancer cells were hypersensitive to treatment with CHK1 inhibitors, which suggests a potential therapeutic opportunity to target AMBRA1-deficient tumours.

These exciting results raise several crucial questions. For example, does a decrease in AMBRA1 levels underlie existing and acquired resistance of human tumours to CDK4/6 inhibitors? And is the rise in D-cyclins that occurs with AMBRA1 depletion the sole factor responsible for the development of resistance to CDK4/6 inhibitors? Analyses of clinical trials in people with breast cancer^{16–19} reveal no correlation between the presence of extra copies of the gene encoding cyclin D1, or the level of cyclin-D1 messenger RNA or protein in tumours, and a patient's response to CDK4/6 inhibitors. Indeed, Chaikovsky *et al.* found that human cancer cells driven to have higher than normal expression of D-cyclins did not fully recapitulate the characteristics of inhibitor resistance observed on AMBRA1 depletion. Perhaps other AMBRA1-regulated proteins, such as c-MYC, which can upregulate cyclin E protein and activate cyclin E–CDK2 complexes, contribute to treatment resistance.



Mitochondria are mixed during cell division

The observation by Chaikovsky *et al.* and Simoneschi *et al.* of the formation of CDK4/6-inhibitor-resistant cyclin D–CDK2 complexes in AMBRA1-depleted cells is intriguing. Such ‘atypical’ complexes were previously shown to underlie acquired resistance to CDK4/6 inhibition²⁰. It is tempting to speculate that depletion of AMBRA1 somehow promotes the formation of these cyclin D–CDK2 complexes — and that this, along with an elevation of cyclin D levels, results in resistance to CDK4/6 inhibitors. A particularly exciting possibility, arising from Maiani and colleagues’ work, is that CHK1 inhibitors can be used to treat CDK4/6-inhibitor-resistant tumours that have low levels of AMBRA1.

Further studies of the role of AMBRA1 in human cancer are necessary. Is the tumour-suppressive function of AMBRA1 mediated mainly by cyclin D1 or c-MYC, or is it also mediated by other targets? Tumour cells that no longer produce RB1 do not require D-cyclins for cell-cycle progression⁴, so AMBRA1 loss would be expected to be observed in tumours that produce RB1 if D-cyclins are the main target of AMBRA1. It also remains to be determined whether AMBRA1 loss in human tumours is mutually exclusive with mutations affecting the C termini of D-cyclins, which would be expected to render D-cyclins resistant to AMBRA1-mediated degradation. Another unresolved issue is why, as Chaikovsky and colleagues report, low expression of AMBRA1 correlates with high levels of cyclin D1 and

associates with poor survival of lung tumours that have specific types of genetic alteration, such as *Kras* mutations. This effect was not observed for lung tumours that instead have mutant versions of the gene encoding EGFR protein or have wild-type *Kras*.

Regardless of the answers to these questions, the impressive studies by Chaikovsky *et al.*, Maiani *et al.* and Simoneschi *et al.* improve our understanding of mechanisms governing cell-cycle progression.

Nature **592**, 690-691 (2021)

doi: <https://doi.org/10.1038/d41586-021-00889-0>

References

1. 1.

Simoneschi, D. *et al.* Nature **592**, 789–793 (2021).

2. 2.

Chaikovsky, A. C. *et al.* Nature **592**, 794–798 (2021).

3. 3.

Maiani, E. *et al.* Nature **592**, 799–804 (2021).

4. 4.

Otto, T. & Sicinski, P. Nature Rev. Cancer **17**, 93–115 (2017).

5. 5.

Sherr, C. J., Beach, D. & Shapiro, G. I. Cancer Discov. **6**, 353–367 (2016).

6. 6.

Lin, D. I. *et al.* Mol. Cell **24**, 355–366 (2006).

7. 7.

Chen, S.-H. *et al.* *EMBO J.* **37**, e97508 (2018).

8. 8.

Petroski, M. D. & Deshaies, R. J. *Nature Rev. Mol. Cell Biol.* **6**, 9–20 (2005).

9. 9.

Qie, S. & Diehl, J. A. *J. Mol. Med.* **94**, 1313–1326 (2016).

10. 10.

Kanie, T. *et al.* *Mol. Cell Biol.* **32**, 590–605 (2012).

11. 11.

Cianfanelli, V., Nazio, F. & Cecconi, F. *Mol. Cell Oncol.* **2**, e970059 (2015).

12. 12.

Fimia, G. M. *et al.* *Nature* **447**, 1121–1125 (2007).

13. 13.

Jin, J., Arias, E. E., Chen, J., Harper, J. W. & Walter, J. C. *Mol. Cell* **23**, 709–721 (2006).

14. 14.

Cianfanelli, V. *et al.* *Nature Cell Biol.* **17**, 20–30 (2015).

15. 15.

Bretones, G., Delgado, M. D. & Leon, J. *Biochim. Biophys. Acta* **1849**, 506–516 (2015).

16. 16.

Finn, R. S. *et al. Lancet Oncol.* **16**, 25–35 (2015).

17. 17.

Finn, R. S. *et al. Clin. Cancer Res.* **26**, 110–121 (2020).

18. 18.

Li, Z. *et al. Cancer Cell* **34**, 893–905 (2018).

19. 19.

Turner, N. C. *et al. J. Clin. Oncol.* **37**, 1169–1178 (2019).

20. 20.

Herrera-Abreu, M. T. *et al. Cancer Res.* **76**, 2301–2313 (2016).

Competing Financial Interests

P.S. has been a consultant at Novartis, Genovis, Guidepoint, The Planning Shop, ORIC Pharmaceuticals, Syros, Cedilla Therapeutics and Exo Therapeutics. His laboratory has received research funding from Novartis.

Jobs from Nature Careers

- - - [All jobs](#)
 - - [Research Associate / Postdoc](#)
 - [Technische Universität Dresden \(TU Dresden\)](#)
 - [Dresden, Germany](#)

JOB POST

- **PhD position in evolutionary ecology of insect immune priming**

Université Bourgogne Franche-Comté (UBFC)

DIJON, France

JOB POST

- **PhD contract Doctoral contract in sociology-demography : Analysis of differential socio-demographic behaviors by sex among 50-74 year olds**

Université Bourgogne Franche-Comté (UBFC)

DIJON, France

JOB POST

- **PhD - Dark ecological networks and land management: methodological and strategic contributions for reducing ecological impacts of artificial light**

Université Bourgogne Franche-Comté (UBFC)

BESANCON, France

JOB POST

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

NEWS AND VIEWS

28 April 2021

Ancient continental blocks soldered from below

A study of melting in the mantle under northern Canada more than one billion years ago shows that the oldest blocks of continent not only break apart but can also be repaired by the gluing action of major melting episodes.

- [Stephen Foley](#) ⁰ &
- [Craig O'Neill](#) ¹

1. [Stephen Foley](#)

1. Stephen Foley is in the Department of Earth and Environmental Sciences, Macquarie University, Sydney, New South Wales 2109, Australia.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

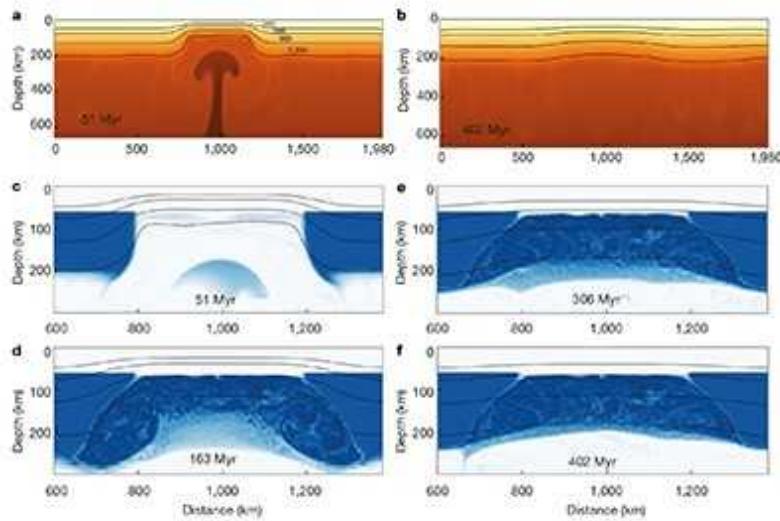
2. [Craig O'Neill](#)

1. Craig O'Neill is in the Department of Earth and Environmental Sciences, Macquarie University, Sydney, New South Wales 2109, Australia.

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

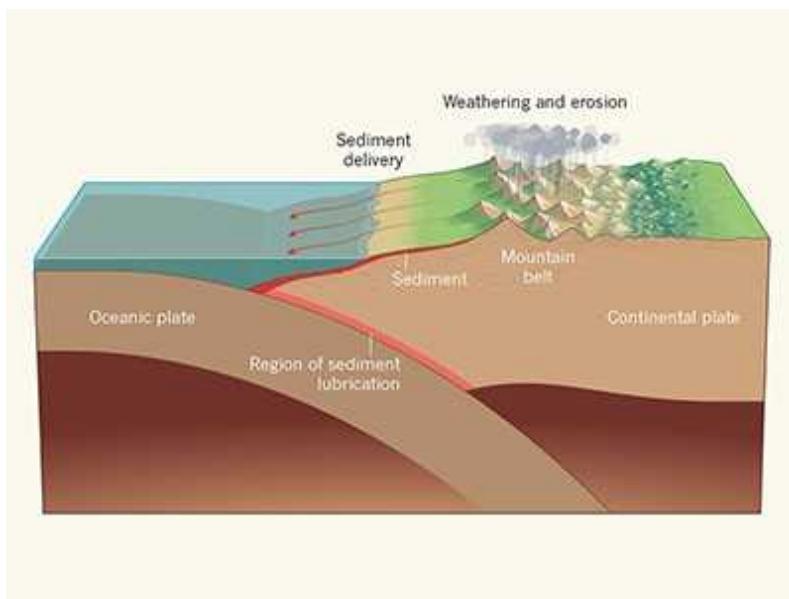
Billions of years of plate tectonics have destroyed much of the evidence about the nature of the earliest continents. The parts of these continents that remain have survived because their composition makes them buoyant and strong, and they float like driftwood on the convecting mantle that has slowly churned beneath them many times over. These ancient continental blocks, known as cratons, were originally thought to be indefinitely stable owing to their strength and buoyancy. However, in the past few years, many studies have cast doubt on this viewpoint and emphasized the break-up and destruction of cratons. [Writing in Nature](#), Liu *et al.*¹ report that cratons not only break apart but can also be fused back together again.



[Read the paper: Plume-driven recratonization of deep continental lithospheric mantle](#)

Mantle plumes are columns of hot, buoyant material that rise upwards through the mantle. On modern Earth, plumes are thought to be responsible for rifting (the splitting apart of tectonic plates) and continental break-up², and for the erosion of cratons³. A large geological structure called the Mackenzie dyke swarm in a craton of northern Canada constitutes the remnants of an enormous plume. This structure comprises a massive collection of vertical slabs of solidified molten rock spanning thousands of kilometres and dating back 1.27 billion years⁴.

High-pressure melting experiments on rocks with similar compositions to those of the Mackenzie dyke swarm have shown that the magma that formed this structure originated from depths of less than 100 kilometres below the planet's surface⁵. Therefore, this magma must have erupted through a thin plate. However, geophysical images of the area today reveal a thick plate (about 200 km thick)⁶ that is typical of cratons, showing that the plate must have thickened after the development of the dyke swarm. Radiometric dating of mantle samples from the area indicates that the plate had thickened by 600 million years ago⁷.

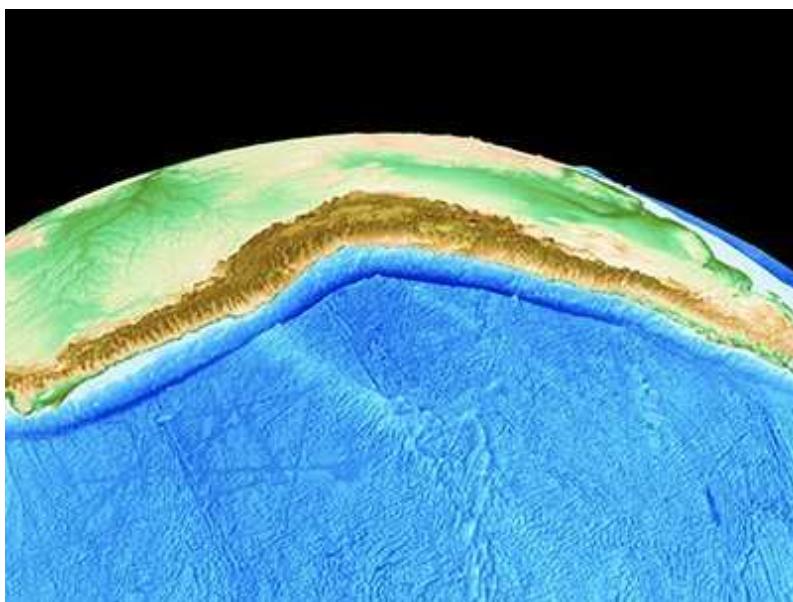


Role of major erosion events in Earth's dynamics

Liu and colleagues combined geodynamic simulations with pressure–temperature calculations for the origin of mantle samples to suggest a mechanism for the thickening of the plate immediately after the dyke swarm formed. They demonstrate in their simulations that the residue from the melting of the plume responsible for the dyke swarm could have helped to repair the damaged craton, essentially filling thin parts of the plate and cementing it from below. This finding indicates that cratons do not simply exist passively for billions of years, but undergo a more complex history than was previously assumed.

Over the past few years, the indestructibility of cratons has been questioned. It has been found that they can be slowly destroyed by chiselling from below

by magmas³, and can crumble from the sides as a result of small-scale mantle convection⁸ or the effects of subduction⁹ (the process in which one plate dives beneath another). Consequently, in some regions, such as northern China, the cratonic plate has thinned over vast areas¹⁰. Cratons had previously been thought to resist the effects of hot mantle material from below, spreading the heat to the sides, much as a frying pan does to the flame beneath. In this scenario, the heat would result in blooms of magmatic activity around the margins of the craton, but it would not affect the base of the craton much.



Lost tectonic history recovered from Earth's deep mantle

However, Liu and colleagues show that mantle melts can repair and thicken cratonic plates in areas where they are thin, even if these areas are disconnected from the main heat source. This finding is consistent with the upward flow of hot mantle material along channels in the base of such plates¹¹. Therefore, this material, which melts as it nears the planet's surface, might preferentially seal the cratons along existing gaps or crevasses in the underside of the continental blocks. If confirmed, this sealing mechanism will change the outlook of many researchers, who have viewed mantle plumes as causing the break-up of continents rather than as a means to stabilize them.

The resealing of continental blocks and thickening of plates also have ramifications for mineral exploration. Many types of mineral deposit are often spatially associated with huge structures known as trans-lithospheric faults that reach from Earth's surface down to the base of the plates, cutting through both the crust and the upper mantle. These faults provide pathways for fluids and melts, which can form mineral deposits in the upper crust when they stop or encounter rocks that contrast, in their chemical properties, with those from which the fluids or melts originated¹². Clearly, trans-lithospheric faults can serve as pathways only if they are open at the bottom. Therefore, when the resealing process occurs, it will close a time window after which no further mineral deposits can form in these large structures.

Why some mantle plumes break up continents and others stitch them together remains unknown. Differences in regional tectonic states and local forces might play a part in determining whether a plume is destructive or constructive (Fig. 1). For instance, if a plume is located below a continent that is prevented from breaking up because of the relative movements of the surrounding plates, the resealing effect becomes more probable than it would otherwise be.

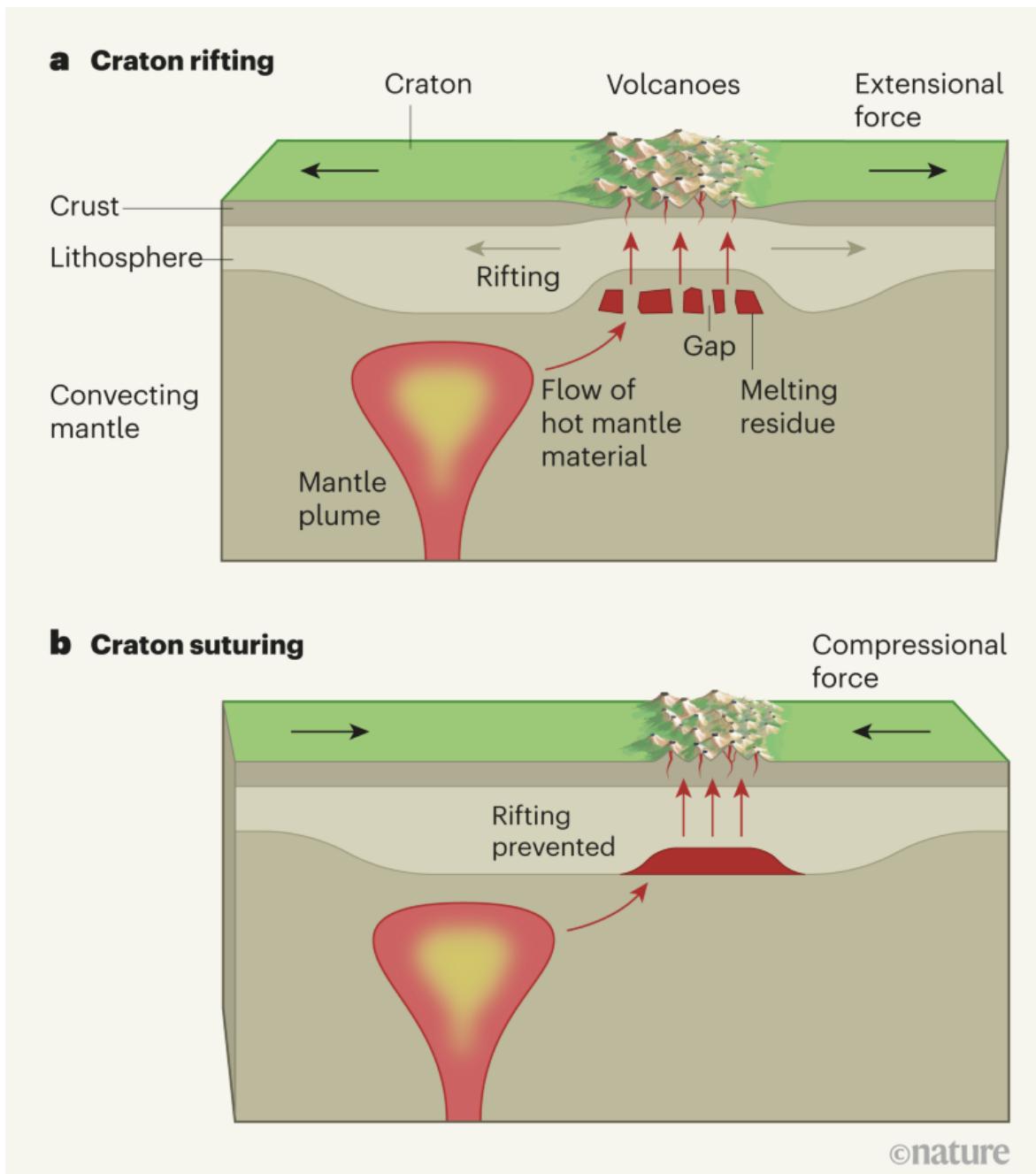


Figure 1 | Rifting and suturing effects of mantle plumes. Liu *et al.*¹ explore how ancient blocks of continent called cratons are influenced by mantle plumes — columns of warm, buoyant material that rise upwards through Earth's convecting mantle. This hot mantle material flows into gaps in the underside of the lithosphere (the rigid layer beneath Earth's crust). It then melts as it approaches the planet's surface, producing abundant volcanism. **a**, If regional forces related to the large-scale movement of

tectonic plates are extensional, the craton can undergo rifting (stretching and thinning of the crust and lithosphere), and the residue from the melting of the hot mantle material cannot plug the gaps in the underside of the lithosphere. **b**, However, if these tectonic forces are compressional, craton rifting is prevented, and the authors show that the melting residue can plug the gaps and thicken the base of the lithosphere.

Another open question is how the frequency of occurrence of mantle plumes, their melting behaviour and their ability to cement continental blocks together have changed over time¹³. Many cratons are made up of smaller blocks that were amalgamated during the Archaean eon more than 2.5 billion years ago. Compared with today, the mantle was then hotter, plumes are thought to have been more frequent, and more of these smaller blocks covered the planet's surface. In those times, the suturing mechanism described by Liu and colleagues might have been common, allowing these smaller blocks to weld together and survive the violence of early Earth.

Nature **592**, 692-693 (2021)

doi: <https://doi.org/10.1038/d41586-021-01087-8>

References

1. 1.

Liu, J. *et al.* Nature **592**, 732–736 (2021).

2. 2.

Hill, R. I. Earth Planet. Sci. Lett. **104**, 398–416 (1991).

3. 3.

Foley, S. F. Nature Geosci. **1**, 503–510 (2008).

4. 4.

Gibson, I. L., Sinha, M. N. & Fahrig, W. F. in *Mafic Dyke Swarms* Spec. Pap. 34 (eds Halls, H. C. & Fahrig, W. F.) 109–121 (Geol. Assoc. Canada, 1987).

5. 5.

Falloon, T. J., Green, D. H., Hatton, C. J. & Harris, K. L. *J. Petrol.* **29**, 1257–1282 (1988).

6. 6.

Schaeffer, A. J. & Lebedev, S. *Earth Planet. Sci. Lett.* **402**, 26–41 (2014).

7. 7.

Armstrong, J. P., Fitzgerald, C. E., Kjarsgaard, B. A., Heaman, L. & Tappe, S. *Int. Kimberlite Conf. Ext. Abstr.* 10IKC-170 (2012).

8. 8.

Muirhead, J. D. *et al. Nature* **582**, 67–72 (2020).

9. 9.

Dave, R. & Li, A. *Geology* **44**, 883–886 (2016).

10. 10.

Zhu, R. X., Chen, L., Wu, F. Y. & Liu, J. L. *Science China Earth Sci.* **54**, 789–797 (2011).

11. 11.

Sleep, N. H., Ebinger, C. J. & Kendall, J.-M. in *The Early Earth: Physical, Chemical and Biological Development* Spec. Publ. 199 (eds Fowler, C. M. R., Ebinger, C. J. & Hawkesworth, C. J.) 135–150 (Geol. Soc. Lond., 2002).

12. 12.

Begg, G. C. *et al. Econ. Geol.* **105**, 1057–1070 (2010).

13. 13.

Fiorentini, M. L. *et al. Sci. Rep.* **10**, 19729 (2020).

Jobs from Nature Careers

- - - [All jobs](#)
 - - [Research Associate / Postdoc](#)
 - [Technische Universität Dresden \(TU Dresden\)](#)
[Dresden, Germany](#)
[JOB POST](#)
 - [PhD position in evolutionary ecology of insect immune priming](#)
[Université Bourgogne Franche-Comté \(UBFC\)](#)
[DIJON, France](#)
[JOB POST](#)
 - [PhD contract Doctoral contract in sociology-demography : Analysis of differential socio-demographic behaviors by sex among 50-74 year olds](#)

Université Bourgogne Franche-Comté (UBFC)

DIJON, France

JOB POST

- **PhD - Dark ecological networks and land management: methodological and strategic contributions for reducing ecological impacts of artificial light**

Université Bourgogne Franche-Comté (UBFC)

BESANCON, France

JOB POST

This article was downloaded by **calibre** from <https://www.nature.com/articles/d41586-021-01087-8>

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

- Review
- [Published: 28 April 2021](#)

The central role of DNA damage in the ageing process

- [Björn Schumacher](#) ORCID: [orcid.org/0000-0001-6097-5238^{1,2}](https://orcid.org/0000-0001-6097-5238),
- [Joris Pothof³](#),
- [Jan Vijg](#) ORCID: [orcid.org/0000-0002-8457-9595^{4,5}](https://orcid.org/0000-0002-8457-9595) &
- [Jan H. J. Hoeijmakers](#) ORCID: [orcid.org/0000-0003-3526-7795^{1,2,3,6}](https://orcid.org/0000-0003-3526-7795)

[Nature](#) volume 592, pages695–703(2021)[Cite this article](#)

- 5513 Accesses
- 1 Citations
- 122 Altmetric
- [Metrics details](#)

Subjects

- [Ageing](#)
- [DNA damage and repair](#)
- [Genomic instability](#)
- [Senescence](#)

Abstract

Ageing is a complex, multifaceted process leading to widespread functional decline that affects every organ and tissue, but it remains unknown whether

ageing has a unifying causal mechanism or is grounded in multiple sources. Phenotypically, the ageing process is associated with a wide variety of features at the molecular, cellular and physiological level—for example, genomic and epigenomic alterations, loss of proteostasis, declining overall cellular and subcellular function and deregulation of signalling systems. However, the relative importance, mechanistic interrelationships and hierarchical order of these features of ageing have not been clarified. Here we synthesize accumulating evidence that DNA damage affects most, if not all, aspects of the ageing phenotype, making it a potentially unifying cause of ageing. Targeting DNA damage and its mechanistic links with the ageing phenotype will provide a logical rationale for developing unified interventions to counteract age-related dysfunction and disease.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

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: DNA damage is the driver of ageing.

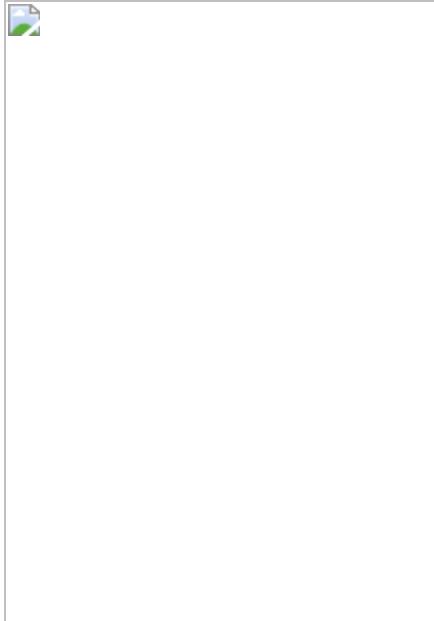
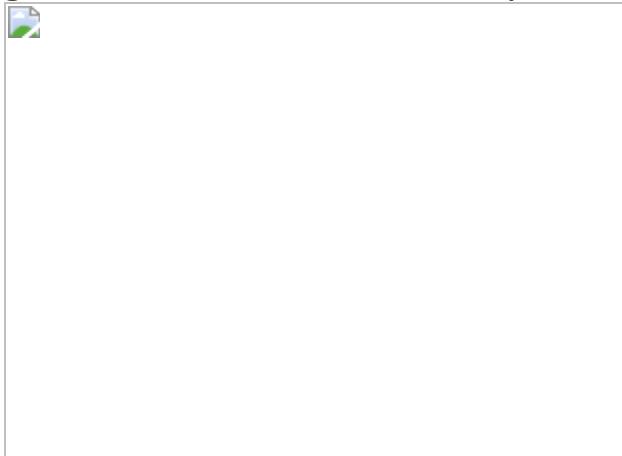


Fig. 2: Molecular, cellular and systemic consequences of DNA damage.



References

1. 1.

Schumacher, B. *The Mystery of Human Aging: Surprising Insights from a Science That's Still Young* (Algora, 2017).

2. 2.

Charlesworth, B. Fisher, Medawar, Hamilton and the evolution of aging. *Genetics* **156**, 927–931 (2000).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

3. 3.

Williams, G. C. Pleiotropy, natural selection, and the evolution of senescence. *Evolution* **11**, 398–411 (1957).

[Google Scholar](#)

4. 4.

López-Otín, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. The hallmarks of aging. *Cell* **153**, 1194–1217 (2013).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

5. 5.

Lindahl, T. Instability and decay of the primary structure of DNA. *Nature* **362**, 709–715 (1993).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

6. 6.

Niedernhofer, L. J. et al. Nuclear genomic instability and aging. *Annu. Rev. Biochem.* **87**, 295–322 (2018).

[CAS](#) [PubMed](#) [Google Scholar](#)

7. 7.

Lubberts, S., Meijer, C., Demaria, M. & Gietema, J. A. Early ageing after cytotoxic treatment for testicular cancer and cellular senescence: time to act. *Crit. Rev. Oncol. Hematol.* **151**, 102963 (2020).

[PubMed](#) [Google Scholar](#)

8. 8.

Gonzalo, S., Kreienkamp, R. & Askjaer, P. Hutchinson–Gilford progeria syndrome: a premature aging disease caused by LMNA gene mutations. *Ageing Res. Rev.* **33**, 18–29 (2017).

[CAS](#) [PubMed](#) [Google Scholar](#)

9. 9.

Vijg, J. *Aging of the Genome* (Oxford Univ. Press, 2007).

10. 10.

Lundblad, V. & Szostak, J. W. A mutant with a defect in telomere elongation leads to senescence in yeast. *Cell* **57**, 633–643 (1989).

[CAS](#) [PubMed](#) [Google Scholar](#)

11. 11.

de Lange, T. Shelterin-mediated telomere protection. *Annu. Rev. Genet.* **52**, 223–247 (2018).

[PubMed](#) [Google Scholar](#)

12. 12.

Fumagalli, M. et al. Telomeric DNA damage is irreparable and causes persistent DNA-damage-response activation. *Nat. Cell Biol.* **14**, 355–365 (2012).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

13. 13.

Abdallah, P. et al. A two-step model for senescence triggered by a single critically short telomere. *Nat. Cell Biol.* **11**, 988–993 (2009).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

14. 14.

Shay, J. W. Role of telomeres and telomerase in aging and cancer. *Cancer Discov.* **6**, 584–593 (2016).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

15. 15.

Martínez, P. & Blasco, M. A. Telomere-driven diseases and telomere-targeting therapies. *J. Cell Biol.* **216**, 875–887 (2017).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

16. 16.

Jaskelioff, M. et al. Telomerase reactivation reverses tissue degeneration in aged telomerase-deficient mice. *Nature* **469**, 102–106 (2011).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

17. 17.

Demanelis, K. et al. Determinants of telomere length across human tissues. *Science* **369**, eaaz6876 (2020).

[PubMed](#) [Google Scholar](#)

18. 18.

Robin, J. D. et al. Telomere position effect: regulation of gene expression with progressive telomere shortening over long distances. *Genes Dev.* **28**, 2464–2476 (2014).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

19. 19.

Hauer, M. H. & Gasser, S. M. Chromatin and nucleosome dynamics in DNA damage and repair. *Genes Dev.* **31**, 2204–2221 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

20. 20.

O’Sullivan, R. J., Kubicek, S., Schreiber, S. L. & Karlseder, J. Reduced histone biosynthesis and chromatin changes arising from a damage signal at telomeres. *Nat. Struct. Mol. Biol.* **17**, 1218–1225 (2010).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

21. 21.

Hu, Z. et al. Nucleosome loss leads to global transcriptional up-regulation and genomic instability during yeast aging. *Genes Dev.* **28**, 396–408 (2014).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

22. 22.

Lu, A. T. et al. DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging* **11**, 303–327 (2019).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

23. 23.

Mortusewicz, O., Schermelleh, L., Walter, J., Cardoso, M. C. & Leonhardt, H. Recruitment of DNA methyltransferase I to DNA repair sites. *Proc. Natl Acad. Sci. USA* **102**, 8905–8909 (2005).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

24. 24.

Wang, S., Meyer, D. H. D. H. & Schumacher, B. H3K4me2 regulates the recovery of protein biosynthesis and homeostasis following DNA damage. *Nat. Struct. Mol. Biol.* **27**, 1165–1177 (2020). **S. Wang and colleagues revealed that, after the repair of DNA lesions, the epigenetic mark H3K4me2 is deposited to restore protein homeostasis and thus antagonize DNA damage-driven ageing.**

[CAS](#) [Article](#) [PubMed](#) [Google Scholar](#)

25. 25.

Ito, T., Teo, Y. V., Evans, S. A., Neretti, N. & Sedivy, J. M. Regulation of cellular senescence by polycomb chromatin modifiers through distinct DNA damage- and histone methylation-dependent pathways. *Cell Rep.* **22**, 3480–3492 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

26. 26.

Sedelnikova, O. A. et al. Senescent human cells and ageing mice accumulate DNA lesions with unrepairable double-strand breaks. *Nat. Cell Biol.* **6**, 168–170 (2004).

[CAS](#) [PubMed](#) [Google Scholar](#)

27. 27.

Rodier, F. et al. DNA-SCARS: distinct nuclear structures that sustain damage-induced senescence growth arrest and inflammatory cytokine secretion. *J. Cell Sci.* **124**, 68–81 (2011).

[CAS](#) [PubMed](#) [Google Scholar](#)

28. 28.

Russo, G. et al. DNA damage and repair modify DNA methylation and chromatin domain of the targeted locus: mechanism of allele methylation polymorphism. *Sci. Rep.* **6**, 33222 (2016).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

29. 29.

Fang, E. F. et al. Defective mitophagy in XPA via PARP-1 hyperactivation and NAD⁺/SIRT1 reduction. *Cell* **157**, 882–896 (2014).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

30. 30.

Oberdoerffer, P. et al. SIRT1 redistribution on chromatin promotes genomic stability but alters gene expression during aging. *Cell* **135**, 907–918 (2008).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

31. 31.

Bahar, R. et al. Increased cell-to-cell variation in gene expression in ageing mouse heart. *Nature* **441**, 1011–1014 (2006). **R. Bahar and colleagues demonstrated that cell-to-cell variation in gene expression increases in aged single cardiomyocytes, suggesting that genome damage could trigger stochastic variations in transcript levels, thus compromising cellular function during ageing.**

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

32. 32.

Takada, K. & Becker, L. E. Cockayne's syndrome: report of two autopsy cases associated with neurofibrillary tangles. *Clin. Neuropathol.* **5**, 64–68 (1986).

[CAS](#) [PubMed](#) [Google Scholar](#)

33. 33.

Lopes, A. F. C. et al. A *C. elegans* model for neurodegeneration in Cockayne syndrome. *Nucleic Acids Res.* **48**, 10973–10985 (2020).

[CAS](#) [Article](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

34. 34.

Bucholtz, N. & Demuth, I. DNA-repair in mild cognitive impairment and Alzheimer's disease. *DNA Repair* **12**, 811–816 (2013).

[CAS](#) [PubMed](#) [Google Scholar](#)

35. 35.

Weissman, L. et al. Defective DNA base excision repair in brain from individuals with Alzheimer's disease and amnestic mild cognitive impairment. *Nucleic Acids Res.* **35**, 5545–5555 (2007).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

36. 36.

Leandro, G. S., Lobo, R. R., Oliveira, D. V. N. P., Moriguti, J. C. & Sakamoto-Hojo, E. T. Lymphocytes of patients with Alzheimer's disease display different DNA damage repair kinetics and expression profiles of DNA repair and stress response genes. *Int. J. Mol. Sci.* **14**, 12380–12400 (2013).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

37. 37.

Sepe, S., Payan-Gomez, C., Milanese, C., Hoeijmakers, J. H. & Mastroberardino, P. G. Nucleotide excision repair in chronic neurodegenerative diseases. *DNA Repair* **12**, 568–577 (2013).

[CAS](#) [PubMed](#) [Google Scholar](#)

38. 38.

Obulesu, M. & Rao, D. M. DNA damage and impairment of DNA repair in Alzheimer's disease. *Int. J. Neurosci.* **120**, 397–403 (2010).

[CAS](#) [PubMed](#) [Google Scholar](#)

39. 39.

Sepe, S. et al. Inefficient DNA repair is an aging-related modifier of Parkinson's disease. *Cell Rep.* **15**, 1866–1875 (2016).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

40. 40.

Jones, L., Houlden, H. & Tabrizi, S. J. DNA repair in the trinucleotide repeat disorders. *Lancet Neurol.* **16**, 88–96 (2017).

[CAS](#) [PubMed](#) [Google Scholar](#)

41. 41.

Gao, R. et al. Mutant huntingtin impairs PNKP and ATXN3, disrupting DNA repair and transcription. *eLife* **8**, e42988 (2019).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

42. 42.

Lodato, M. A. et al. Aging and neurodegeneration are associated with increased mutations in single human neurons. *Science* **359**, 555–559 (2018). **Using single-cell sequencing, this study revealed increasing somatic mutations in neurons during human ageing, indicating that even post-mitotic cell types accrue mutations that might trigger age-associated functional decline and degeneration.**

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

43. 43.

Vermeij, W. P. et al. Restricted diet delays accelerated ageing and genomic stress in DNA-repair-deficient mice. *Nature* **537**, 427–431 (2016). **W. P. Vermeij and colleagues showed that premature ageing in DNA repair-deficient mice could be alleviated by dietary restriction, suggesting that DNA repair is capable of reducing DNA damage infliction and promoting genome stability.**

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

44. 44.

Wei, Y. N. et al. Transcript and protein expression decoupling reveals RNA binding proteins and miRNAs as potential modulators of human aging. *Genome Biol.* **16**, 41 (2015).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

45. 45.

Kelmer Sacramento, E. et al. Reduced proteasome activity in the aging brain results in ribosome stoichiometry loss and aggregation. *Mol. Syst. Biol.* **16**, e9596 (2020).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

46. 46.

Edifizi, D. et al. Multilayered reprogramming in response to persistent DNA damage in *C. elegans*. *Cell Rep.* **20**, 2026–2043 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

47. 47.

Rodier, F. et al. Persistent DNA damage signalling triggers senescence-associated inflammatory cytokine secretion. *Nat. Cell Biol.* **11**, 973–979 (2009). **This study established that senescent cells release cytokines and thus elicit non-cell-autonomous effects such as inflammatory responses.**

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

48. 48.

Williams, A. B. et al. Restoration of proteostasis in the endoplasmic reticulum reverses an inflammation-like response to cytoplasmic DNA in *Caenorhabditis elegans*. *Genetics* **212**, 1259–1278 (2019).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

49. 49.

Harman, D. Aging: a theory based on free radical and radiation chemistry. *J. Gerontol.* **11**, 298–300 (1956).

[CAS](#) [PubMed](#) [Google Scholar](#)

50. 50.

Kauppila, T. E. S., Kauppila, J. H. K. & Larsson, N. G. Mammalian mitochondria and aging: an update. *Cell Metab.* **25**, 57–71 (2017).

[CAS](#) [PubMed](#) [Google Scholar](#)

51. 51.

Trifunovic, A. et al. Premature ageing in mice expressing defective mitochondrial DNA polymerase. *Nature* **429**, 417–423 (2004).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

52. 52.

Kujoth, G. C. et al. Mitochondrial DNA mutations, oxidative stress, and apoptosis in mammalian aging. *Science* **309**, 481–484 (2005).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

53. 53.

Bua, E. et al. Mitochondrial DNA-deletion mutations accumulate intracellularly to detrimental levels in aged human skeletal muscle fibers. *Am. J. Hum. Genet.* **79**, 469–480 (2006).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

54. 54.

Wanagat, J., Cao, Z., Pathare, P. & Aiken, J. M. Mitochondrial DNA deletion mutations colocalize with segmental electron transport system abnormalities, muscle fiber atrophy, fiber splitting, and oxidative damage in sarcopenia. *FASEB J.* **15**, 322–332 (2001).

[CAS](#) [PubMed](#) [Google Scholar](#)

55. 55.

Kraytsberg, Y. et al. Mitochondrial DNA deletions are abundant and cause functional impairment in aged human substantia nigra neurons. *Nat. Genet.* **38**, 518–520 (2006).

[CAS](#) [PubMed](#) [Google Scholar](#)

56. 56.

Taylor, R. W. et al. Mitochondrial DNA mutations in human colonic crypt stem cells. *J. Clin. Invest.* **112**, 1351–1360 (2003).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

57. 57.

Vermulst, M. et al. Mitochondrial point mutations do not limit the natural lifespan of mice. *Nat. Genet.* **39**, 540–543 (2007).

[CAS](#) [PubMed](#) [Google Scholar](#)

58. 58.

O’Hara, R. et al. Quantitative mitochondrial DNA copy number determination using droplet digital PCR with single-cell resolution. *Genome Res.* **29**, 1878–1888 (2019).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

59. 59.

Ameur, A. et al. Ultra-deep sequencing of mouse mitochondrial DNA: mutational patterns and their origins. *PLoS Genet.* **7**, e1002028 (2011).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

60. 60.

Hoch, N. C. et al. XRCC1 mutation is associated with PARP1 hyperactivation and cerebellar ataxia. *Nature* **541**, 87–91 (2017).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

61. 61.

Hayflick, L. & Moorhead, P. S. The serial cultivation of human diploid cell strains. *Exp. Cell Res.* **25**, 585–621 (1961).

[CAS](#) [PubMed](#) [Google Scholar](#)

62. 62.

Bodnar, A. G. et al. Extension of life-span by introduction of telomerase into normal human cells. *Science* **279**, 349–352 (1998).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

63. 63.

Demaria, M. et al. An essential role for senescent cells in optimal wound healing through secretion of PDGF-AA. *Dev. Cell* **31**, 722–733 (2014).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

64. 64.

Krtolica, A., Parrinello, S., Lockett, S., Desprez, P.-Y. & Campisi, J. Senescent fibroblasts promote epithelial cell growth and tumorigenesis: a link between cancer and aging. *Proc. Natl Acad. Sci. USA* **98**, 12072–12077 (2001).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

65. 65.

Baker, D. J. et al. Clearance of p16Ink4a-positive senescent cells delays ageing-associated disorders. *Nature* **479**, 232–236 (2011). **D. J. Baker and colleagues revealed that the elimination of senescent cells could delay ageing and provided the conceptual framework for eliminating senescence cells for therapeutic interventions aimed at extending healthspan.**

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

66. 66.

Childs, B. G. et al. Senescent intimal foam cells are deleterious at all stages of atherosclerosis. *Science* **354**, 472–477 (2016).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

67. 67.

Jeon, O. H. et al. Local clearance of senescent cells attenuates the development of post-traumatic osteoarthritis and creates a pro-regenerative environment. *Nat. Med.* **23**, 775–781 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

68. 68.

Kotsantis, P., Petermann, E. & Boulton, S. J. Mechanisms of oncogene-induced replication stress: jigsaw falling into place. *Cancer Discov.* **8**, 537–555 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

69. 69.

Gorgoulis, V. et al. Cellular senescence: defining a path forward. *Cell* **179**, 813–827 (2019).

[CAS](#) [PubMed](#) [Google Scholar](#)

70. 70.

Andriani, G. A. et al. Whole chromosome instability induces senescence and promotes SASP. *Sci. Rep.* **6**, 35218 (2016).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

71. 71.

Wiley, C. D. et al. Mitochondrial dysfunction induces senescence with a distinct secretory phenotype. *Cell Metab.* **23**, 303–314 (2016).

[CAS](#) [PubMed](#) [Google Scholar](#)

72. 72.

McNeely, T., Leone, M., Yanai, H. & Beerman, I. DNA damage in aging, the stem cell perspective. *Hum. Genet.* **139**, 309–331 (2020).

[PubMed](#) [Google Scholar](#)

73. 73.

Matsumura, H. et al Hair follicle aging is driven by transepidermal elimination of stem cells via COL17A1 proteolysis. *Science* **351**, aad4395 (2016).

[PubMed](#) [Google Scholar](#)

74. 74.

Alyodawi, K. et al. Compression of morbidity in a progeroid mouse model through the attenuation of myostatin/activin signalling. *J. Cachexia Sarcopenia Muscle* **10**, 662–686 (2019).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

75. 75.

Geiger, H., de Haan, G. & Florian, M. C. The ageing haematopoietic stem cell compartment. *Nat. Rev. Immunol.* **13**, 376–389 (2013).

[CAS](#) [PubMed](#) [Google Scholar](#)

76. 76.

Beerman, I., Seita, J., Inlay, M. A., Weissman, I. L. & Rossi, D. J. Quiescent hematopoietic stem cells accumulate DNA damage during aging that is repaired upon entry into cell cycle. *Cell Stem Cell* **15**, 37–50 (2014).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

77. 77.

Flach, J. et al. Replication stress is a potent driver of functional decline in ageing haematopoietic stem cells. *Nature* **512**, 198–202 (2014).

This study established that increasing DNA replication stress in ageing haematopoietic stem cells leads to their functional decline.

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

78. 78.

Mohrin, M. et al. Hematopoietic stem cell quiescence promotes error-prone DNA repair and mutagenesis. *Cell Stem Cell* **7**, 174–185 (2010).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

79. 79.

Rossi, D. J. et al. Deficiencies in DNA damage repair limit the function of haematopoietic stem cells with age. *Nature* **447**, 725–729 (2007).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

80. 80.

Lee-Six, H. et al. Population dynamics of normal human blood inferred from somatic mutations. *Nature* **561**, 473–478 (2018).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

81. 81.

Zink, F. et al. Clonal hematopoiesis, with and without candidate driver mutations, is common in the elderly. *Blood* **130**, 742–752 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

82. 82.

Zhang, L. & Vijg, J. Somatic mutagenesis in mammals and its implications for human disease and aging. *Annu. Rev. Genet.* **52**, 397–419 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

83. 83.

Franco, I. et al. Somatic mutagenesis in satellite cells associates with human skeletal muscle aging. *Nat. Commun.* **9**, 800 (2018).

[ADS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

84. 84.

Ju, Z. et al. Telomere dysfunction induces environmental alterations limiting hematopoietic stem cell function and engraftment. *Nat. Med.* **13**, 742–747 (2007).

[CAS](#) [PubMed](#) [Google Scholar](#)

85. 85.

Mostoslavsky, R. et al. Genomic instability and aging-like phenotype in the absence of mammalian SIRT6. *Cell* **124**, 315–329 (2006).

[CAS](#) [PubMed](#) [Google Scholar](#)

86. 86.

Liu, L. et al. Impaired Notch signaling leads to a decrease in p53 activity and mitotic catastrophe in aged muscle stem cells. *Cell Stem Cell* **23**, 544–556.e4 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

87. 87.

Ou, H.-L., Kim, C. S., Uszkoreit, S., Wickström, S. A. & Schumacher, B. Somatic niche cells regulate the CEP-1/p53-mediated DNA damage response in primordial germ cells. *Dev. Cell* **50**, 167–183.e8 (2019).

[CAS](#) [PubMed](#) [Google Scholar](#)

88. 88.

Kenyon, C., Chang, J., Gensch, E., Rudner, A. & Tabtiang, R. A *C. elegans* mutant that lives twice as long as wild type. *Nature* **366**, 461–464 (1993).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

89. 89.

Saxton, R. A. & Sabatini, D. M. mTOR signaling in growth, metabolism, and disease. *Cell* **168**, 960–976 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

90. 90.

Prasad, R. & Katiyar, S. K. Crosstalk among UV-induced inflammatory mediators, DNA damage and epigenetic regulators facilitates suppression of the immune system. *Photochem. Photobiol.* **93**, 930–936 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

91. 91.

Ermolaeva, M. A. et al. DNA damage in germ cells induces an innate immune response that triggers systemic stress resistance. *Nature* **501**, 416–420 (2013).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

92. 92.

Schumacher, B. et al. Delayed and accelerated aging share common longevity assurance mechanisms. *PLoS Genet.* **4**, e1000161 (2008).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

93. 93.

Niedernhofer, L. J. et al. A new progeroid syndrome reveals that genotoxic stress suppresses the somatotroph axis. *Nature* **444**, 1038–1043 (2006). **This study linked unrepaired DNA damage to genetic regulators of longevity by showing that ILS is attenuated in progeroid mice deficient in DNA repair.**

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

94. 94.

Garinis, G. A. et al. Persistent transcription-blocking DNA lesions trigger somatic growth attenuation associated with longevity. *Nat. Cell Biol.* **11**, 604–615 (2009).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

95. 95.

Mueller, M. M. et al. DAF-16/FOXO and EGL-27/GATA promote developmental growth in response to persistent somatic DNA damage. *Nat. Cell Biol.* **16**, 1168–1179 (2014). **Using *C. elegans*, M. M. Mueller and colleagues demonstrated that the longevity regulator DAF-16 responds to DNA damage and elevates the tolerance of the organism tolerance to persistent DNA lesions.**

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

96. 96.

McCay, C. M., Maynard, L. A., Sperling, G. & Barnes, L. L. Retarded growth, life span, ultimate body size and age changes in the albino rat after feeding diets restricted in calories. *J. Nutr.* **18**, 1–13 (1939).

[CAS](#) [Google Scholar](#)

97. 97.

López-Otín, C., Galluzzi, L., Freije, J. M. P., Madeo, F. & Kroemer, G. Metabolic control of longevity. *Cell* **166**, 802–821 (2016).

[PubMed](#) [Google Scholar](#)

98. 98.

Matsuoka, S. et al. ATM and ATR substrate analysis reveals extensive protein networks responsive to DNA damage. *Science* **316**, 1160–1166 (2007).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

99. 99.

Dominick, G., Bowman, J., Li, X., Miller, R. A. & Garcia, G. G. mTOR regulates the expression of DNA damage response enzymes in long-lived Snell dwarf, GHRKO, and PAPPA-KO mice. *Aging Cell* **16**, 52–60 (2017).

[CAS](#) [PubMed](#) [Google Scholar](#)

100. 100.

Alves-Fernandes, D. K. & Jasiulionis, M. G. The role of SIRT1 on DNA damage response and epigenetic alterations in cancer. *Int. J. Mol. Sci.* **20**, 1–13 (2019).

[Google Scholar](#)

101. 101.

Wu, C. L. et al. Role of AMPK in UVB-induced DNA damage repair and growth control. *Oncogene* **32**, 2682–2689 (2013).

[CAS](#) [PubMed](#) [Google Scholar](#)

102. 102.

Ma, Y., Vassetzky, Y. & Dokudovskaya, S. mTORC1 pathway in DNA damage response. *Biochim. Biophys. Acta* **1865**, 1293–1311 (2018).

[CAS](#) [Google Scholar](#)

103. 103.

Adamowicz, M., Vermezovic, J. & d'Adda di Fagagna, F. NOTCH1 inhibits activation of ATM by impairing the formation of an ATM-FOXO3a-KAT5/Tip60 complex. *Cell Rep.* **16**, 2068–2076 (2016).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

104. 104.

Tian, X. et al. SIRT6 is responsible for more efficient DNA double-strand break repair in long-lived species. *Cell* **177**, 622–638.e22 (2019).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

105. 105.

Kanfi, Y. et al. The sirtuin SIRT6 regulates lifespan in male mice. *Nature* **483**, 218–221 (2012). **This study established that in male mice transgenic expression of SIRT6 leads to reduced IGF-1 signalling and lifespan extension.**

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

106. 106.

Shaposhnikov, M., Proshkina, E., Shilova, L., Zhavoronkov, A. & Moskalev, A. Lifespan and stress resistance in *Drosophila* with overexpressed DNA repair genes. *Sci. Rep.* **5**, 15299 (2015).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

107. 107.

Zhang, L. et al. Single-cell whole-genome sequencing reveals the functional landscape of somatic mutations in B lymphocytes across the human lifespan. *Proc. Natl Acad. Sci. USA* **116**, 9014–9019 (2019).

Using single cell sequencing, L. Zhang and colleagues show increased somatic mutations during normal human ageing occurring in genes and regulatory regions indicating their role in the age-dependent functional decline.

[CAS](#) [PubMed](#) [Google Scholar](#)

108. 108.

Brazhnik, K. et al. Single-cell analysis reveals different age-related somatic mutation profiles between stem and differentiated cells in human liver. *Sci. Adv.* **6**, eaax2659 (2020).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

109. 109.

Franco, I. et al. Whole genome DNA sequencing provides an atlas of somatic mutagenesis in healthy human cells and identifies a tumor-prone cell type. *Genome Biol.* **20**, 285 (2019).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

110. 110.

Yizhak, K. et al. RNA sequence analysis reveals macroscopic somatic clonal expansion across normal tissues. *Science* **364**, eaaw0726 (2019). **This study used RNA sequencing to reveal the clonal expansion of somatic mutations in different human tissues.**

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

111. 111.

Johnson, S. C., Dong, X., Vijg, J. & Suh, Y. Genetic evidence for common pathways in human age-related diseases. *Aging Cell* **14**, 809–817 (2015).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

112. 112.

Day, F. R. et al. Large-scale genomic analyses link reproductive aging to hypothalamic signaling, breast cancer susceptibility and BRCA1-mediated DNA repair. *Nat. Genet.* **47**, 1294–1303 (2015).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

113. 113.

Laven, J. S. E., Visser, J. A., Uitterlinden, A. G., Vermeij, W. P. & Hoeijmakers, J. H. J. Menopause: Genome stability as new paradigm. *Maturitas* **92**, 15–23 (2016).

[CAS](#) [PubMed](#) [Google Scholar](#)

114. 114.

Henshaw, P. S., Riley, E. F. & Stapleton, G. E. The biologic effects of pile radiations. *Radiology* **49**, 349–360 (1947).

[CAS](#) [PubMed](#) [Google Scholar](#)

115. 115.

Alexander, P. The role of DNA lesions in the processes leading to aging in mice. *Symp. Soc. Exp. Biol.* **21**, 29–50 (1967).

[CAS](#) [PubMed](#) [Google Scholar](#)

116. 116.

Jans, J. et al. Powerful skin cancer protection by a CPD-photolyase transgene. *Curr. Biol.* **15**, 105–115 (2005).

[CAS](#) [PubMed](#) [Google Scholar](#)

117. 117.

Hart, R. W. & Setlow, R. B. Correlation between deoxyribonucleic acid excision-repair and life-span in a number of mammalian species. *Proc. Natl Acad. Sci. USA* **71**, 2169–2173 (1974).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

118. 118.

Gensler, H. L. & Bernstein, H. DNA damage as the primary cause of aging. *Q. Rev. Biol.* **56**, 279–303 (1981).

[CAS](#) [PubMed](#) [Google Scholar](#)

119. 119.

de Boer, J. et al. Premature aging in mice deficient in DNA repair and transcription. *Science* **296**, 1276–1279 (2002). **J. de Boer and colleagues established that genetic defects in DNA repair accelerate the mammalian ageing process.**

[ADS](#) [PubMed](#) [Google Scholar](#)

120. 120.

Barzilai, A., Schumacher, B. & Shiloh, Y. Genome instability: linking ageing and brain degeneration. *Mech. Ageing Dev.* **161** (Pt A), 4–18 (2017).

[CAS](#) [PubMed](#) [Google Scholar](#)

121. 121.

Traube, F. R. et al. Isotope-dilution mass spectrometry for exact quantification of noncanonical DNA nucleosides. *Nat. Protoc.* **14**, 283–312 (2019).

[CAS](#) [PubMed](#) [Google Scholar](#)

122. 122.

Mori, T. et al. High levels of oxidatively generated DNA damage 8,5'-cyclo-2'-deoxyadenosine accumulate in the brain tissues of xeroderma pigmentosum group A gene-knockout mice. *DNA Repair* **80**, 52–58 (2019).

[CAS](#) [PubMed](#) [Google Scholar](#)

123. 123.

Van Houten, B., Cheng, S. & Chen, Y. Measuring gene-specific nucleotide excision repair in human cells using quantitative amplification of long targets from nanogram quantities of DNA. *Mutat. Res.* **460**, 81–94 (2000).

[PubMed](#) [Google Scholar](#)

124. 124.

Nakazawa, Y. et al. Ubiquitination of DNA damage-stalled RNAPII promotes transcription-coupled repair. *Cell* **180**, 1228–1244.e24 (2020).

[CAS](#) [PubMed](#) [Google Scholar](#)

125. 125.

Mingard, C., Wu, J., McKeague, M. & Sturla, S. J. Next-generation DNA damage sequencing. *Chem. Soc. Rev.* **49**, 7354–7377 (2020).

[CAS](#) [PubMed](#) [Google Scholar](#)

[Download references](#)

Acknowledgements

The research was supported by the Deutsche Forschungsgemeinschaft (SCHU 2494/3-1, SCHU 2494/7-1, SCHU 2494/10-1, SCHU 2494/11-1, KFO 286, KFO 329, GRK 2407 to B.S., and CECAD EXC 2030–390661388, SFB 829 to B.S. and J.H.J.H.), Deutsche Krebshilfe (70112899), H2020-MSCA-ITN-2018 (Healthage and ADDRESS ITNs) and John Templeton Foundation Grant (61734) to B.S., NIH grants (PO1 AG017242 to J.V. and J.H., and U19 AG056278, U01 ES029519, P01AG047200, U01HL145560, P30AG038072 to J.V.), European Research Council Advanced Grants DamAge and Dam2Age, ONCODE (Dutch Cancer Society), Memorabel and Chembridge (ZonMW) and BBoL (NWO-ENW) to J.P. and J.H.

Author information

Affiliations

1. Institute for Genome Stability in Ageing and Disease, Medical Faculty, University of Cologne, Cologne, Germany

Björn Schumacher & Jan H. J. Hoeijmakers

2. Cologne Excellence Cluster for Cellular Stress Responses in Aging-Associated Diseases (CECAD), Center for Molecular Medicine Cologne (CMMC), University of Cologne, Cologne, Germany

Björn Schumacher & Jan H. J. Hoeijmakers

3. Department of Molecular Genetics, Erasmus University Medical Center, Rotterdam, The Netherlands

Joris Pothof & Jan H. J. Hoeijmakers

4. Department of Genetics, Albert Einstein College of Medicine, Bronx, New York, NY, USA

Jan Vijg

5. Center for Single-Cell Omics, School of Public Health, Shanghai Jiao Tong University School of Medicine, Shanghai, China

Jan Vijg

6. Princess Máxima Center for Pediatric Oncology, Oncode Institute, Utrecht, The Netherlands

Jan H. J. Hoeijmakers

Authors

1. Björn Schumacher

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Joris Pothof

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Jan Vijg

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Jan H. J. Hoeijmakers

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Contributions

All authors contributed to the writing of the article.

Corresponding author

Correspondence to [Björn Schumacher](#).

Ethics declarations

Competing interests

J.V. is a co-founder of SingulOmics Corp. The other authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Jan van Deursen, John Sedivy 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.

Rights and permissions

[Reprints and Permissions](#)

About this article



Check for
updates

Cite this article

Schumacher, B., Pothof, J., Vijg, J. *et al.* The central role of DNA damage in the ageing process. *Nature* **592**, 695–703 (2021).
<https://doi.org/10.1038/s41586-021-03307-7>

[Download citation](#)

- Received: 24 March 2020
- Accepted: 28 January 2021
- Published: 28 April 2021
- Issue Date: 29 April 2021
- DOI: <https://doi.org/10.1038/s41586-021-03307-7>

Further reading

- [**From DNA damage to mutations: All roads lead to aging**](#)
 - Jan Vijg

Ageing Research Reviews (2021)

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

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

X-ray quasi-periodic eruptions from two previously quiescent galaxies

[Download PDF](#)

- Article
- Open Access
- [Published: 28 April 2021](#)

X-ray quasi-periodic eruptions from two previously quiescent galaxies

- [R. Arcodia](#) [ORCID: orcid.org/0000-0003-4054-7978¹](#),
- [A. Merloni](#) [ORCID: orcid.org/0000-0002-0761-0130¹](#),
- [K. Nandra](#) [ORCID: orcid.org/0000-0002-7150-9192¹](#),
- [J. Buchner¹](#),
- [M. Salvato](#) [ORCID: orcid.org/0000-0001-7116-9303¹](#),
- [D. Pasham²](#),
- [R. Remillard](#) [ORCID: orcid.org/0000-0003-4815-0481²](#),
- [J. Comparat¹](#),
- [G. Lamer³](#),
- [G. Ponti](#) [ORCID: orcid.org/0000-0003-0293-3608^{1,4}](#),
- [A. Malyali](#) [ORCID: orcid.org/0000-0002-8851-4019¹](#),
- [J. Wolf¹](#),
- [Z. Arzoumanian⁵](#),
- [D. Bogensberger](#) [ORCID: orcid.org/0000-0002-5924-4822¹](#),
- [D. A. H. Buckley⁶](#),
- [K. Gendreau](#) [ORCID: orcid.org/0000-0001-7115-2819⁵](#),
- [M. Gromadzki](#) [ORCID: orcid.org/0000-0002-1650-1518⁷](#),
- [E. Kara²](#),
- [M. Krumpe³](#),
- [C. Markwardt⁵](#),
- [M. E. Ramos-Caja](#) [ORCID: orcid.org/0000-0002-9117-3251¹](#),
- [A. Rau](#) [ORCID: orcid.org/0000-0001-5990-6243¹](#),
- [M. Schramm⁸](#) &
- [A. Schwone](#) [ORCID: orcid.org/0000-0003-3441-9355³](#)

[Nature](#) volume **592**, pages 704–707(2021) [Cite this article](#)

- 1725 Accesses
- 58 Altmetric
- [Metrics details](#)

Subjects

- [Compact astrophysical objects](#)
- [High-energy astrophysics](#)

Abstract

Quasi-periodic eruptions (QPEs) are very-high-amplitude bursts of X-ray radiation recurring every few hours and originating near the central supermassive black holes of galactic nuclei^{1,2}. It is currently unknown what triggers these events, how long they last and how they are connected to the physical properties of the inner accretion flows. Previously, only two such sources were known, found either serendipitously or in archival data^{1,2}, with emission lines in their optical spectra classifying their nuclei as hosting an actively accreting supermassive black hole^{3,4}. Here we report observations of QPEs in two further galaxies, obtained with a blind and systematic search of half of the X-ray sky. The optical spectra of these galaxies show no signature of black hole activity, indicating that a pre-existing accretion flow that is typical of active galactic nuclei is not required to trigger these events. Indeed, the periods, amplitudes and profiles of the QPEs reported here are inconsistent with current models that invoke radiation-pressure-driven instabilities in the accretion disk^{5,6,7,8,9}. Instead, QPEs might be driven by an orbiting compact object. Furthermore, their observed properties require the mass of the secondary object to be much smaller than that of the main body¹⁰, and future X-ray observations may constrain possible changes in their period owing to orbital evolution. This model could make QPEs a viable candidate for the electromagnetic counterparts of so-called extreme-mass-ratio inspirals^{11,12,13}, with considerable implications for multi-messenger astrophysics and cosmology^{14,15}.

[Download PDF](#)

Main

Given its large collecting area and blind survey strategy, the eROSITA instrument on the Spectrum-Roentgen-Gamma (SRG) space observatory¹⁶ is capable of systematic searches for X-ray sources that are variable on timescales of hours to months (see [Methods](#) for more details). This applies to QPEs, which thus far have only been

detected in X-rays^{1,2}. The first QPE observed by eROSITA, hereafter eRO-QPE1, showed high-amplitude X-ray variability within just a few hours. It showed a strong X-ray signal in two eROSITA survey scans that were preceded, separated and followed by scans in which the signal was much fainter (Fig. 1a). Similar to the two previously known QPE sources—GSN 069¹ and RX J1301.9+2747²—the X-ray spectrum is very soft with most of the counts originating from below approximately 1.5–2 keV and consistent with a thermal black-body emission. As with the light curve, the spectrum shows oscillations from a faint to a bright phase (Fig. 1b). We identify eRO-QPE1 as originating within the nucleus of the galaxy 2MASS 02314715-1020112, for which we measured a spectroscopic redshift of $z = 0.0505$ (see [Methods](#) section ‘The host galaxies of the QPEs’). The related eROSITA quiescence (1σ upper limit) and peak intrinsic 0.5–2-keV luminosities are $<2.1 \times 10^{41} \text{ erg s}^{-1}$ and approximately $9.4 \times 10^{42} \text{ erg s}^{-1}$, respectively, if the X-ray spectra are modelled with a standard accretion disk model (see [Methods](#) section ‘X-ray spectral analysis’).

Fig. 1: The first eROSITA QPE.

 **figure1**

a, eROSITA light curve in the 0.2–0.6-keV and 0.6–2.3-keV energy bands (circles and squares, respectively), with red and orange highlighting faint and bright observations, respectively. The start of the light curve $t_{\text{eRO},0}$ is approximately MJD 58864.843 (MJD, modified Julian date), observed during the first eROSITA all-sky survey (eRASS1). **b**, eROSITA X-ray spectra of the bright and faint states in orange and red, as in **a**. **c**, Background-subtracted XMM-Newton X-ray light curves with 500-s bins for the European Photon Imaging Camera (EPIC) instruments: EPIC-pn (dark grey),

MOS1 (green) and MOS2 (red) in the energy band shown in the legend. The beginning of both observations was contaminated by flares in the background and excluded; the dark grey solid line and contours show the underlying ≤ 1 -keV EPIC-pn light curve to give a zeroth-order extrapolation of the rate, excluding the presence of obvious soft X-ray eruptions. $t_{\text{XMM},0}$ corresponds to the start of the cleaned MOS2 exposure in the first observation, approximately MJD 59057.805. XMM-Newton optical and UV fluxes are shown in the lower panels (units of $\text{erg cm}^{-2} \text{s}^{-1}$, where F_λ is the spectral flux density and λ is the wavelength in angstroms), with non-detections shown as upper limits. **d**, Background-subtracted NICER-XTI light curve. The mean (and dispersion-on) rise-to-decay duration is approximately 7.6 h (~ 1.0 h) and the peak-to-peak separation is approximately 18.5 h (~ 2.7 h). All uncertainties are 1σ , shown as error bars or shaded regions.

[Full size image](#)

Two follow-up observations triggered with the XMM-Newton X-ray telescope confirmed the remarkable bursting nature of the source (Fig. [1c, d](#)). The first observation (hereafter eRO-QPE1-XMM1) found the source in a faint state for approximately 30 ks, followed by a sequence of three consecutive asymmetric bursts, possibly partially overlapping (Fig. [1c](#)), which is behaviour that has not been previously observed in QPEs^{[1,2](#)}. In terms of intrinsic 0.5–2-keV luminosity, after an initial quiescent phase at about $2.3 \times 10^{40} \text{ erg s}^{-1}$ the first burst was characterized by a fast rise and slower decay lasting around 30 ks and peaking at approximately $3.3 \times 10^{42} \text{ erg s}^{-1}$; it was then followed by a second fainter burst (peak at approximately $7.9 \times 10^{41} \text{ erg s}^{-1}$) and by a third, which was the brightest (peak at approximately $2.0 \times 10^{43} \text{ erg s}^{-1}$) but was only caught during its rise. The second XMM-Newton observation (hereafter eRO-QPE1-XMM2) showed an eruption very similar to the first seen in eRO-QPE1-XMM1 in terms of amplitude and luminosity, although lasting for >40 ks, that is, for almost as much as the three in eRO-QPE1-XMM1 combined (Fig. [1c](#)). To better characterize the physics and to determine the duty cycle of these eruptions, we started an intense monitoring campaign with the NICER X-ray instrument aboard the International Space Station (ISS), which revealed 15 eruptions in about 11 days (Fig. [1d](#)).

The second eROSITA QPE we have detected, hereafter eRO-QPE2, showed similar variability patterns and X-ray spectra as eRO-QPE1 during the X-ray sky survey (Fig. [2a, b](#)). We associated it with the galaxy 2MASX J02344872-4419325 and determined a spectroscopic redshift of $z = 0.0175$ (see [Methods](#) section ‘The host galaxies of the QPEs’). The related intrinsic 0.5–2-keV luminosities of the quiescent (1σ upper limit) and peak phases are $<4.0 \times 10^{40} \text{ erg s}^{-1}$ and approximately $1.0 \times 10^{42} \text{ erg s}^{-1}$, respectively. A follow-up observation with XMM-Newton revealed nine eruptions in a single day, oscillating between approximately $1.2 \times 10^{41} \text{ erg s}^{-1}$ and $1.2 \times 10^{42} \text{ erg s}^{-1}$.

in the 0.5–2-keV band (Fig. 2c). In neither eRO-QPE1 nor eRO-QPE2 is there evidence of simultaneous optical/UV variability (see Figs. 1c, 2c), in agreement with the behaviour of GSN 069¹.

Fig. 2: The second eROSITA QPE.



a, b, As in Fig. 1a, b, for eRO-QPE2. The start of the eROSITA light curve is approximately MJD 59023.191. **c**, As in Fig. 1c, for the XMM-Newton observation of eRO-QPE2. $t_{\text{XMM},1}$ corresponds to the start of the cleaned MOS1 exposure, approximately MJD 59067.846. The mean (and related dispersion) of the rise-to-decay duration is about 27 min (~ 3 min), with a peak-to-peak separation of approximately 2.4 h (~ 5 min). All uncertainties are 1σ , shown as error bars or shaded regions.

[Full size image](#)

eRO-QPE1 shows a distribution of QPE rise-to-decay durations with a mean (dispersion) of approximately 7.6 h (~ 1.0 h) and a distribution of peak-to-peak separations of about 18.5 h (~ 2.7 h), as derived from the NICER light curve (Fig. 1d). The duty cycle (mean duration over mean separation) is approximately 41%. Conversely, eRO-QPE2 shows much narrower and more frequent eruptions (see Fig.

[2c](#)): the mean (dispersion) of the rise-to-decay duration is approximately 27 min (~ 3 min), with a peak-to-peak separation of approximately 2.4 h (~ 5 min) and a duty cycle of around 19%. Interestingly, compared to the two previously known QPEs^{[1,2](#)}, eRO-QPE1 and eRO-QPE2 extend the parameter space of QPE widths and recurrence times towards longer and shorter timescales, respectively. We also note that eRO-QPE1 is the most luminous and the most distant QPE discovered so far, and the most extreme in terms of timescales. The duration and recurrence times of the bursts in eRO-QPE1 are approximately an order of magnitude longer than in eRO-QPE2. This could simply be an effect of the timescales scaling with black hole mass^{[17](#)}. We estimated the total stellar mass of the two host galaxies and that of eRO-QPE1 is 4–8 times higher than that of eRO-QPE2. Assuming a standard scaling of the black hole mass with stellar mass (see [Methods](#) section ‘The host galaxies of the QPEs’), this is broadly in agreement with their different X-ray timing properties. Furthermore, peak soft X-ray luminosities of approximately 2×10^{43} erg s $^{-1}$ and 10^{42} erg s $^{-1}$ for eRO-QPE1 and eRO-QPE2, respectively, exclude a stellar-mass black hole origin, and their X-ray positions, within uncertainties, suggest a nuclear origin (Extended Data Figs. [1a](#), [2a](#)).

The optical counterparts of eRO-QPE1 and eRO-QPE2 are local low-mass galaxies with no canonical active galactic nuclei (AGN)-like broad emission lines in the optical nor any infrared photometric excess indicating the presence of hot dust (the so-called torus)^{[18](#)}. In this sense they are similar to GSN 069^{[3](#)} and RX J1301.9+2747^{[4](#)}, although their optical spectra show narrow emission lines with clear AGN-driven ionization^{[3,4](#)}. Instead, the optical counterpart of eRO-QPE1 is easily classified as a passive galaxy from the absence of emission lines (Extended Data Fig. [1b](#)), and in eRO-QPE2 the strong narrow emission lines that are observed classify it as a star-forming galaxy (Extended Data Fig. [2b](#) and Methods section ‘The host galaxies of the QPEs’). This in turn suggests that the two galaxies reported here have not been active for approximately at least the last 10^3 – 10^4 years, assuming narrow-line region light-travel timescales^{[19](#)}. Although the number of known QPEs is too low to reach firm statistical conclusions, our blind search is inherently designed to sample the population of the QPEs’ host galaxies without bias, as opposed to serendipitous or archival discoveries, which rely on the source being previously active and known^{[1,2](#)}. These results hint that perhaps the parent population of QPE hosts consists of more passive galaxies than active galaxies. The X-ray spectra of the QPEs in quiescence are consistent with an almost featureless accretion disk model^{[1,2](#)} (see [Methods](#) section ‘X-ray spectral analysis’), although the inactive nature of the host galaxies of our sources argues against a pre-existing AGN-like accretion system.

A few scenarios to explain the QPEs have been suggested^{[1,10](#)}, some based on the presumed active nature of the host black holes of the QPEs. These include so-called limit-cycle radiation-pressure accretion instabilities (see [Methods](#) section ‘On

accretion flow instabilities’), proposed for GSN 069¹ based on the similarities between its observed properties and two extremely variable stellar-mass black holes, namely GRS 1915+105^{20,21} and IGR J17091-3624²². However, the observed properties of the two QPEs reported here, as well as those of RX J1301.9+2747², are inconsistent with the theoretical predictions of this scenario^{5,6,7,8,9}. In particular, the faster rise and slower decay of eRO-QPE1 would imply a thicker flow in the cold and stable phase than in the hot and unstable phase, contrary to theory⁶. Moreover, the theory predicts that once the period, the duty cycle and the luminosity amplitude are known, only specific values of black hole mass M_{BH} and viscosity parameter α are allowed⁸: for eRO-QPE1 (eRO-QPE2) one solution is found for $M_{\text{BH}} \approx 4 \times 10^6 M_{\odot}$ (M_{\odot} , solar mass) and $\alpha \approx 5$ ($M_{\text{BH}} \approx 3 \times 10^6 M_{\odot}$ and $\alpha \approx 3$), therefore for the expected masses^{1,2} an unphysically high viscosity parameter would be required. Alternatively, more reasonable values of $\alpha \approx 0.1$ and 0.01 would yield very small $M_{\text{BH}} \approx 2.4 \times 10^3 M_{\odot}$ and $M_{\text{BH}} \approx 60 M_{\odot}$ ($M_{\text{BH}} \approx 4.3 \times 10^3 M_{\odot}$ and $M_{\text{BH}} \approx 30 M_{\odot}$) for eRO-QPE1 (eRO-QPE2). Even in this latter scenario and pushing α as high as approximately 0.2 , the resulting thermal timescales for eRO-QPE1 (eRO-QPE2) are $\tau_{\text{th}} \approx 20$ s (35 s) at $20r_g$ ($r_g = GM_{\text{BH}}/c^2$, where G is the gravitational constant and c the speed of light in vacuum), which is orders of magnitude smaller than the observed QPE timescales (more details in Methods section ‘On accretion flow instabilities’).

Extreme or sinusoidal quasi-periodic variability as seen in QPEs is also typically associated with compact object binaries, a scenario which would not require the galactic nuclei to be previously active, as our present evidence suggests. Drawing a simplistic scenario, we assumed the mass of the main body to be in the range of approximately $10^4 M_{\odot} - 10^7 M_{\odot}$ for both eRO-QPE1 and eRO-QPE2 and computed the expected period decrease of a compact binary due to emission of gravitational waves. We inferred that a supermassive black hole binary with a mass ratio of the order of unity²³ is unlikely given the properties of the observed optical, ultraviolet (UV), infrared and X-ray emission in QPEs and the lack of evident periodicity and/or strong period decrease therein. If QPEs are triggered by the presence of a secondary orbiting body, our data suggest its mass (M_2) to be much smaller than the main body. This is in agreement with at least one proposed scenario for the origin of GSN 069, for which the average luminosity in a QPE cycle can be reproduced by a periodic mass-inflow rate from a white dwarf orbiting the black hole with a highly eccentric orbit¹⁰. Our current data for eRO-QPE1 only exclude M_2 larger than approximately $10^6 M_{\odot}$ ($\sim 10^4 M_{\odot}$) for zero (high, ~ 0.9) eccentricity (as a function of the mass of the main body, Extended Data Fig. 7a); instead, for eRO-QPE2 we can already argue that only an orbiting M_2 lower than approximately $10^4 M_{\odot}$ ($\sim 10 M_{\odot}$) is allowed for zero (~ 0.9)

eccentricity (Extended Data Fig. 7b). More details are reported in Methods section ‘On the presence of an orbiting body’.

Future X-ray observations on longer temporal baselines (months or years) will help to constrain or rule out this scenario and to monitor the possible orbital evolution of the system. This picture is also reminiscent of a suggested formation channel of extreme-mass-ratio inspirals^{24,25} and it could make QPEs their electromagnetic messenger^{13,26}. Regardless of their origin, the QPEs seen so far seem to be found in relatively low-mass supermassive black holes ($\sim 10^5 M_\odot - 10^7 M_\odot$) and finding more will help us to understand how black holes are activated in low-mass galaxies, which—although so far a poorly explored mass range in their co-evolution history^{27,28}—is crucial for synergies with future Laser Interferometer Space Antenna (LISA) gravitational wave signals²⁹.

Fig. 3: Phase-folded light curves.



a, b, Median light curve profile (with related 16th and 84th percentile contours) for eRO-QPE1 (**a**) and eRO-QPE2 (**b**), folded at the eruption peaks (see [Methods](#)).

[Full size image](#)

Methods

Blind search for QPEs with eROSITA

eROSITA¹⁶ is the main instrument aboard the Spectrum-Roentgen-Gamma (SRG) mission (R. Sunyaev et al., manuscript in preparation), which was launched on 13 July 2019. On 13 December 2019 it started the first of eight all-sky surveys (eRASS1–eRASS8), each completed in six months, observing in the 0.2–8-keV band. In each survey, as the field of view moves every point of the sky is observed for \sim 40 s every \sim 4 h with the number of times (typically six) varying with the location in the sky, increasing towards high ecliptic latitudes. Our search for QPE candidates starts with a systematic screening of all eROSITA light curves, produced for each detected source on a weekly basis by the eROSITA Science Analysis Software (eSASS; H. Brunner et al., manuscript in preparation). Light curves are binned to yield one data point for each 4-h revolution (called an ‘eROday’). A light curve generated by the eSASS pipeline will trigger a ‘QPE alert’ if it shows two or more high-count states with (at least) one low-count state in between (see Figs. 1a, 2a as examples) in any of its standard energy bands (0.2–0.6 keV, 0.6–2.3 keV, 2.3–5.0 keV). As thresholds, we fixed a relative count-rate ratio (including uncertainties) of 5 if both high and low states are detected, or 3 if the low-count state is consistent with the background. Since neither the survey scans nor QPEs are strictly periodic, every eRASS can be treated as an independent sky to find new candidates. This produces a census of X-ray sources varying on hour-long timescales for each eRASS, albeit only for the specific intermittent pattern described above. Unsurprisingly, the vast majority of the automatically generated alerts are produced by Galactic sources (mainly flaring coronally active stars), but we can filter them out by finding the multi-wavelength counterpart associated with every X-ray source (M. Salvato et al., manuscript in preparation). Good QPE candidates are then selected by screening the handful of alerts with a secure or possible extra-galactic counterpart. Thanks to this process, we identified the two best eROSITA QPE candidates that were worth immediate follow-up, promptly obtained with both XMM-Newton and, in one case, NICER. Given the success of our initial search over the first nine months of the survey, we are confident that we can detect up to around 3–4 good eROSITA QPE candidates every year. Therefore, by the end of the last eROSITA all-sky survey in December 2023 this search may provide a sample of up to approximately 10–15 QPEs.

The two eROSITA QPEs

The first QPE, here named eRO-QPE1, is eRASSU J023147.2-102010, located at the astrometrically corrected X-ray position of $\text{RA}_{\text{J}2000}$, $\text{dec.}_{\text{J}2000} = (02:31:47.26, -10:20:10.31)$, with a total 1σ positional uncertainty of \sim 2.1''. It was observed ten

times between 16 and 18 January 2020 during eRASS1 with 339 s of total exposure. Using the Bayesian cross-matching algorithm NWAY³⁰, we associated eRO-QPE1 with the galaxy 2MASS 02314715-1020112 at RA_{J2000}, dec._{J2000} = (02:31:47.15, −10:20:11.22). The second QPE, here named eRO-QPE2, is eRASSU J023448.9-441931, located at the astrometrically corrected X-ray position of RA_{J2000}, dec._{J2000} = (02:34:48.97, −44:19:31.65), with a total positional uncertainty of ∼3.2''. It was observed 11 times between 23 and 24 June 2020 during eRASS2. It was associated via the same method³⁰ with 2MASX J02344872-4419325, a galaxy at RA_{J2000}, dec._{J2000} = (02:34:48.69, −44:19:32.72). Both galaxies are in the DESI Legacy Imaging Surveys³¹ DR8 footprint (Extended Data Figs. 1a, 2a). X-ray XMM-Newton positions were corrected with the ‘eposcorr’ task cross-correlating the sources in the X-ray image with external optical and infrared catalogues. The counterpart of the QPE itself was excluded from the cross-correlation to obtain a more unbiased estimate of the possible offset from the nucleus. The XMM-Newton X-ray positions are consistent with the nuclei of these galaxies. We took optical spectra of both galaxies with the Southern African Large Telescope (SALT)³² and measured spectroscopic redshifts of 0.0505 and 0.0175 for eRO-QPE1 and eRO-QPE2, respectively (Extended Data Figs. 1b, 2b). More details are shown in Methods sections ‘Data reduction’ and ‘The host galaxies of the QPEs’.

Previous X-ray activity

eRO-QPE1 has not previously been detected in X-rays, although upper limits can be obtained from the XMM-Newton upper limits server for ROSAT³³, both from the survey and a pointed observation (taken in 1991 and 1992, with ∼270 s and ∼5,300 s, respectively), and the XMM-Newton Slew Survey³⁴ (taken in 2004, 2007, 2008 and 2017, all between ∼3 s and ∼8 s of exposure). The ROSAT pointed observation puts a stringent upper limit at $\leq 3.8 \times 10^{-14}$ cgs (cgs, erg s^{−1} cm^{−2}) in the 0.2–2.0-keV band. However, given the very short exposures compared with the timescales of eRO-QPE1, we cannot rule out that QPEs were already ongoing and that all previous missions caught eRO-QPE1 in a faint state. As with eRO-QPE1, eRO-QPE2 has not been previously detected in X-rays. Upper limits were again computed for ROSAT (taken in 1990, ∼480 s of exposure) and the XMM-Newton Slew survey (taken in 2004, 2008, 2012 and 2013, all between ∼4 s and ∼8 s). The most stringent upper limit, at $\leq 8.8 \times 10^{-14}$ cgs in the 0.2–2.0-keV band, comes from ROSAT. It is slightly below the flux observed by XMM-Newton in quiescence in the same band (Extended Data Fig. 4), perhaps indicating that the QPE behaviour only started more recently. For both QPE sources however, the ROSAT and Slew exposures are much shorter than the evolving timescales (the QPE quasi-period and its dispersion), hence they do not provide meaningful constraints on the start of the QPE behaviour.

Data reduction

In this section we report details of the processing of the complete dataset. We show a summary of the observations in Extended Data Table 1.

eROSITA

Members of the German eROSITA consortium (eROSITA-DE) have full and immediate access to survey data at Galactic longitudes $180^\circ < l < 360^\circ$. These data were processed using eSASS v946 (H. Brunner et al., manuscript in preparation). For eRO-QPE1 (eRO-QPE2), photons were extracted choosing a circular aperture of radius $80''$ ($67''$), and background counts were extracted from an annulus (off-centre circle) of inner and outer radii $178''$ ($382''$) and $996''$, respectively, excluding all the other sources detected within the area. eRO-QPE1 was detected with a detection likelihood of 440 and a total number of 119 counts in the 0.2–5.0-keV band. eRO-QPE2 was detected with a detection likelihood of 125 and a total number of 48 counts in the 0.2–5.0-keV band.

XMM-Newton

XMM-Newton data from EPIC MOS1, MOS2³⁵ and EPIC-pn³⁶ cameras and the Optical Monitor (OM)³⁷ were processed using standard tools (SAS v. 18.0.0 and HEAsoft v. 6.25) and procedures. Event files from EPIC cameras were filtered for flaring particle background. Source (background) regions were extracted within a circle of $38''$ and $34''$ in eRO-QPE1 and eRO-QPE2, respectively, centred on the source (in a source-free region). eRO-QPE1 was consecutively observed three times with the U filter, then seven times with UVW1 and nine (eight) times with the UVM2 in the first (second) XMM-Newton observation, each exposure \sim 4,400-s long. The source was detected only in the U and UVW1 with mean magnitudes \sim 19.9 and \sim 20.3 in both XMM-Newton observations (OM light curves in Fig. 1c). eRO-QPE2 was consecutively observed twice with the U filter, then ten times with UVW1, six with UVM2 and three with UVW2; all exposures were 4,400 s. It was almost always detected in all filters with mean magnitudes of \sim 17.4, \sim 17.5, \sim 18.0 and \sim 18.1, for U, UVW1, UVM2 and UVW2 filters, respectively (OM light curves in Fig. 2c). eRO-QPE2 was flagged as extended in the U, UVW1 and UVM2 filters, and therefore the reported absolute magnitudes include at least some contamination from the host galaxy.

NICER

NICER's X-ray Timing Instrument (XTI)^{38,39} onboard the ISS observed eRO-QPE1 between 17 August 2020 and 31 August 2020. Beginning late on 19 August, high-cadence observations were performed during almost every ISS orbit, which is roughly 93 min. All the data were processed using the standard NICER Data Analysis Software (NICERDAS) task ‘nicerl2’. Good time intervals (GTIs) were chosen with standard defaults, yielding \sim 186 ks of exposure time. We further divided the GTIs into intervals of 128 s, and on this basis we extracted the spectra and applied the ‘3C50’ model (R.R. et al., submitted) to determine the background spectra. The light curve for eRO-QPE1 in soft X-rays (Fig. 1d) was determined by integrating the background-subtracted spectrum for each 128-s GTI over the range 0.3–1.0 keV. More detailed spectral analyses of these data will be discussed in a follow-up paper.

SALT

Optical spectra of eRO-QPE1 and eRO-QPE2 were obtained using the Robert Stobie Spectrograph (RSS)⁴⁰ on the Southern African Large Telescope (SALT)³² in September 2020 on the nights of the 24th and the 8th, respectively. The PG900 VPH grating was used to obtain pairs of exposures (900 s and 500 s, respectively) at different grating angles, allowing for a total wavelength coverage of 3,500–7,400 Å. The spectra were reduced using the PySALT package, a PyRAF-based software package for SALT data reductions⁴¹, which includes gain and amplifier cross-talk corrections, bias subtraction, amplifier mosaicing, and cosmetic corrections. The individual spectra were then extracted using standard Image Reduction and Analysis Facility (IRAF) procedures, wavelength calibration (with a calibration lamp exposure taken immediately after the science spectra), background subtraction and extraction of one-dimensional spectra. We could only obtain relative flux calibrations, from observing spectrophotometric standards in twilight, owing to the SALT design, which has a time-varying, asymmetric and underfilled entrance pupil⁴².

X-ray spectral analysis

In this work, X-ray spectral analysis was performed using v3.4.2 of the Bayesian X-ray Analysis software (BXA)⁴³, which connects a nested sampling algorithm (UltraNest⁴⁴; J.B., manuscript in preparation) with a fitting environment. For the latter, we used XSPEC v12.10.1⁴⁵ with its Python-oriented interface pyXSPEC. eROSITA source-plus-background spectra were fitted including a model component for the background, which was determined via a principal component analysis from a large sample of eROSITA background spectra⁴⁶ (J.B. et al., manuscript in preparation). XMM-Newton EPIC-pn spectra were instead fitted using wstat, namely the XSPEC implementation of the Cash statistic⁴⁷, given the good count statistics in both source and background spectra. We quote, unless otherwise stated, median values with the related 16th and 84th percentiles and upper limits at 1σ . Results are also reported in Extended Data Tables 2, 3.

eRO-QPE1

For eRO-QPE1, both eROSITA and XMM-Newton EPIC-pn spectra were fitted with a simple absorbed black body (using the models tbabs⁴⁸ and zbbbody) or accretion disk (tbabs and diskbb⁴⁹), with absorption frozen at the Galactic equivalent hydrogen column density (N_{H}) of $N_{\text{H}} \approx 2.23 \times 10^{20} \text{ cm}^{-2}$, as reported by the HI4PI Collaboration⁵⁰. For eROSITA, we jointly extracted and analysed spectra of the faint states (red points in Fig. 1a) and, separately, of the two bright states observed in eRASS1 (orange points in Fig. 1a). In the eROSITA bright states the temperature, in terms of $k_{\text{B}}T$ (k_{B} , Boltzmann's constant; T , temperature) in eV, is $\langle 138 \rangle - \langle 131 \rangle^{\langle 146 \rangle} \langle \text{rm} \{ \text{eV} \} \rangle$ and $\langle \langle 180 \rangle - \langle 168 \rangle^{\langle 195 \rangle} \langle \text{rm} \{ \text{eV} \} \rangle \rangle$, using zbbbody and diskbb as source models, respectively. The related unabsorbed rest-frame 0.5–2.0-keV fluxes are $\langle 1.6 \rangle - \langle 1.4 \rangle^{\langle 1.8 \rangle} \langle 10 \rangle^{\langle -12 \rangle} \langle \text{rm} \{ \text{cgs} \} \rangle$ and $\langle 1.5 \rangle - \langle 1.4 \rangle^{\langle 1.7 \rangle} \langle 10 \rangle^{\langle -12 \rangle} \langle \text{rm} \{ \text{cgs} \} \rangle$, respectively. The eROSITA spectrum of the faint states combined is consistent with background, with the temperature and unabsorbed rest-frame 0.5–2.0-keV flux constrained to be ≤ 124 eV (≤ 160 eV) and $\leq 3.5 \times 10^{-14}$ cgs ($\leq 3.4 \times 10^{-14}$ cgs) for zbbbody (diskbb). We also analysed the observations of eRO-QPE1 obtained six months later during eRASS2, which triggered our QPE search again: two bright states were observed separated by several faint ones, with fluxes consistent with eRASS1.

We performed time-resolved X-ray spectral analysis on XMM-Newton data, extracting a spectrum in each 500-s time bin of the EPIC-pn light curve, with the exception of the quiescence spectrum, which was extracted and analysed combining all the related time bins of both observations (that is, before $t \approx 26,500$ s in eRO-QPE1-XMM1 and before $t \approx 35,788$ s in eRO-QPE1-XMM2, with times as defined in Fig. 1c). Fit results obtained using XMM-Newton EPIC-pn spectra with diskbb as the source model component are shown in Extended Data Fig. 3. Furthermore, we show for visualization three EPIC-pn spectra and related best-fit models (Extended Data Fig. 5a) corresponding to the quiescence phase and the peak of both XMM-Newton observations. A more thorough X-ray spectral analysis with

other models and additional components for the bright phase will be presented in future work.

eRO-QPE2

For eRO-QPE2, eROSITA's faint and bright phases were also separately combined and analysed (Fig. 2a, b). The faint state as observed by eROSITA is consistent with background. The temperature and normalization of the source cannot be constrained, thus we only quote an upper limit for the unabsorbed rest-frame 0.5–2.0-keV flux of $\leq 1.9 \times 10^{-14}$ cgs ($\leq 5.7 \times 10^{-14}$ cgs) using zbbbody (diskbb). The spectrum of the eROSITA bright states constrains the temperature to $\langle 164 \rangle - \langle 149 \rangle^{+182} \langle \text{rm}\{\text{eV}\} \rangle$ and at $\langle 209 \rangle - \langle 185 \rangle^{+241} \langle \text{rm}\{\text{eV}\} \rangle$, using zbbbody and diskbb as source models, respectively. The related unabsorbed rest-frame 0.5–2.0-keV fluxes are $\langle 1.4 \rangle - \langle 1.2 \rangle^{+1.8} \langle 10 \rangle^{-12} \langle \text{rm}\{\text{cgs}\} \rangle$ and $\langle 1.5 \rangle - \langle 1.2 \rangle^{+1.8} \langle 10 \rangle^{-12} \langle \text{rm}\{\text{cgs}\} \rangle$, respectively. The triggering eROSITA observation was obtained during eRASS2, although a single bright state (thus not satisfying our trigger criterion) was also detected in eRASS1 with the same flux level. For eRO-QPE2, in addition to the Galactic column density ($N_{\text{H}} \approx 1.66 \times 10^{20} \text{ cm}^{-2}$)⁵⁰ we included an absorption component at the redshift of the host galaxy (that is, with the models tbabs, ztbabs, and zbbbody or diskbb). This excess absorption was inferred to be present on the basis of the XMM-Newton spectrum (see below).

For XMM-Newton, we performed time-resolved X-ray spectral analysis for each 150-s time bin of the EPIC-pn light curve. The absorption, in addition to the Galactic value, was first fitted in the XMM-Newton quiescence spectrum, which was extracted combining all the low states in the XMM-Newton light curve (Fig. 2c, Extended Data Fig. 4). The fit yielded $\langle N \rangle - \langle \text{rm}\{H\} \rangle = \langle 0.35 \rangle - \langle 0.30 \rangle^{+0.40} \langle 10 \rangle^{+22} \langle \text{cm}^{-2} \rangle$. In all other observations, including all eROSITA spectra and the rises, peaks and decays in the XMM-Newton light curve, the additional N_{H} was left free to vary between the 10th and 90th percentile of the fitted posterior distribution of the quiescent spectrum. Under the

assumption that absorption did not vary throughout the observation, this ensures that no spurious effects are imprinted on the fit temperature and normalizations owing to degeneracies with N_{H} ; at the same time, in this way parameters are marginalized over a reasonable range in N_{H} . Freezing the value instead would artificially narrow the uncertainties on the temperature and normalizations. Fit results obtained with diskbb as the source model are shown in Extended Data Fig. 4. Furthermore, we show for visualization the EPIC-pn spectra and best-fit models of the quiescence and peak phases (Extended Data Fig. 5b). Similar results are obtained using zbbbody as the source model.

Timing analysis

In Fig. 3 we show the median (with related 16th and 84th percentile contours) light-curve profiles obtained by folding the light curve at the eruption peaks. First, a random representative burst is selected and cross-correlated with the whole light curve. The peaks of this cross-correlation identify the times when the phase is zero. Data are then folded at these phase-zero times to obtain a template median profile, which is then used to repeat the same operation and yield Fig. 3. A phase bin of ~ 0.1 corresponds to $\sim 6,600$ s and ~ 820 s for eRO-QPE1 and eRO-QPE2, respectively. Moreover, XMM-Newton and NICER light curve profiles were fitted with UltraNest⁴⁴. Motivated by the strong asymmetry in eRO-QPE1 (Figs. 1c, d, 3a), we adopted a model with Gaussian rise and an exponential decay, a generic model often adopted for transients⁵¹. eRO-QPE2, on the other hand, can be fitted with a simple Gaussian profile (Fig. 3b), possibly owing to the much shorter timescales. Here we simply highlight the most evident results of timing analysis; a more in depth study of the variability properties of QPEs is deferred to future work. Here the modelling allows us to determine mean values for the duration and recurrence time of the QPEs, which were used for comparison with models of accretion instabilities (see [Methods](#) section ‘On accretion flow instabilities’) and compact object binaries (see [Methods](#) section ‘On the presence of an orbiting body’). The mean rise-to-decay duration for eRO-QPE1, as observed from the NICER light curve (Fig. 1d), is ~ 7.6 h (dispersion of ~ 1.0 h), and the mean peak-to-peak separation is ~ 18.5 h (dispersion of ~ 2.7 h). The related duty cycle (here

computed simply as mean duration over mean separation) is \sim 41%. Conversely, eRO-QPE2 shows much narrower and more frequent eruptions (see Fig. 2c): the mean rise-to-decay duration is \sim 27 min (dispersion of \sim 3 min), with a mean peak-to-peak separation of \sim 2.4 h (dispersion of \sim 5 min) and a duty-cycle of \sim 19%.

The host galaxies of the QPEs

Very little was known on both galaxies from published multi-wavelength catalogues, except for WISE infrared monitoring, indicating W1 – W2 emission consistent with zero, which is typical of inactive galactic nuclei, for the last few years. Most of our knowledge is based on optical spectra taken with SALT after the X-ray observation. The optical counterpart of eRO-QPE1 is classified as a passive galaxy from the absence of emission lines (Extended Data Fig. 1b), whereas eRO-QPE2 shows very strong and narrow [O ii], H β , [O iii], H α , [N ii] and [S ii] in emission (Extended Data Fig. 2b). The high [O ii]/[O iii] value and that H β is as strong as [O iii] are already strongly indicative that star-forming processes are the dominant ionization mechanism⁵². We computed the flux ratios $\log([O\text{ iii}]/H\beta) = -0.05$, $\log([O\text{ ii}]/H\beta) = 0.44$ and $\log([N\text{ ii}]/H\alpha) = -0.68$, as well as the [O ii] equivalent width (EW) $\log \text{EW}_{[\text{O II}]} = 2.56$ and $D_{4,000} = 1.26$, where $D_{4,000}$ is the ratio of the continuum level after and before the 4,000 Å break⁵³. Using standard line diagnostic plots⁵⁴ we can confirm that indeed eRO-QPE2 can be classified as star-forming. Spectroscopic classification of future QPEs will be crucial to confirm whether their host galaxies are indeed preferentially inactive, as our pilot study suggests, or not. A first census in a statistically significant sample may bring new insights, as has been the case for other transients such as tidal disruption events (TDEs)^{55,56,57,58}.

A preliminary analysis of the properties of the host galaxies of the QPEs was performed by fitting the optical spectra (Extended Data Figs. 1b, 2b) with Firefly^{59,60}. We first re-normalized the flux of the optical spectra using the most recent *g*-band and *r*-band archival magnitudes, because SALT spectra are not calibrated to absolute values⁴². For eRO-QPE1, *gri*-band photometry ($g = 18.7 \pm 0.06$ mag, $r = 18.0 \pm 0.05$ mag, $i = 17.8 \pm 0.05$ mag)

was taken on 30 July 2020 with the Rapid Response Robotic Telescope at Fan Mountain Observatory, indicating that the source did not change substantially with respect to archival photometry⁶¹. The total stellar masses inferred with Firefly from the optical spectra are $\langle M \rangle_{\text{last}} \approx \{3.8\}_{-1.9}^{+0.4} \times 10^9 M_\odot$ and $\{1.01\}_{-0.50}^{+0.01} \times 10^9 M_\odot$ for eRO-QPE1 and eRO-QPE2, respectively. Systematic errors and degeneracy due to the use of different stellar population models⁶² would push M_* to higher values for eRO-QPE1 ($\sim 4.8 \times 10^9 M_\odot$) and lower values for eRO-QPE2 ($\sim 0.6 \times 10^9 M_\odot$), enhancing their relative difference. Firefly also yields an estimate of the age of the stellar population and the star-formation rate (SFR), although for medium and low signal-to-noise ratios these estimates are more prone to biases⁵⁹. For eRO-QPE2, the mean signal-to-noise ratio (~ 23) is high enough to yield a fairly reliable $\langle \text{SFR} \rangle \approx \{0.078\}_{-0.066}^{+0.001} M_\odot \text{ yr}^{-1}$, which is also consistent within uncertainties with the SFR that can be estimated from the [O ii] and H α luminosities^{63,64}. For eRO-QPE1, the mean signal-to-noise ratio (~ 8) is lower and no reliable estimate of the SFR was obtained. We therefore inferred an upper limit of $\sim 0.01 M_\odot \text{ yr}^{-1}$ from the absence of narrow emission lines^{63,64}. We report in Extended Data Fig. 6 the M_* –SFR plane with the two QPs reported here, together with normal galaxies, and hosts of known TDEs⁵⁷ and changing-look AGN (CLAGN)⁶⁵, all taken below $z < 0.1$ and within the Sloan Digital Sky Survey MPA-JHU DR7 catalogue⁶⁶. Evidence is mounting that both TDEs^{57,67} and CLAGN⁶⁵ might be over-represented in galaxies in the so-called ‘green valley’, perhaps indicating that they are activated in specific periods of galaxy co-evolution with their central black holes. For QPs, a statistically meaningful sample still needs to be built before reaching any conclusion.

We have estimated that the host galaxy of eRO-QPE1 is more massive than that of eRO-QPE2. We here refrain from quoting absolute values for black hole masses using their scaling relations with the host galaxy properties, because our stellar masses are lower than those used to calibrate them⁶⁸. However, it is worth mentioning that the relative ratio of ~ 4 – 8 in stellar masses between the galaxies of eRO-QPE1 and eRO-QPE2 would propagate to a black hole mass ratio of the order of approximately 10 (ref.

[68](#)). This is in line with the X-ray timing properties in eRO-QPE1 and eRO-QPE2, because their peak-to-peak separation and rise-to-decay duration scale roughly by the same amount. Finally, X-ray emission from eRO-QPE1 and eRO-QPE2 was observed to be positionally consistent with the galaxy nucleus for both objects (Extended Data Figs. [1a](#), [2a](#); Methods section ‘The two eROSITA QPEs’). If a future QPE is found in a more nearby galaxy we can aim to constrain more precisely the X-ray position with respect to the galactic nucleus. This will allow us to determine conclusively whether or not these phenomena are nuclear.

On accretion flow instabilities

Accretion disks⁶⁹ with an accretion rate such that radiation pressure dominates in the inner flow are thought to be subject to thermal–viscous instabilities⁷⁰. The net result of these instabilities is that the luminosity is predicted to oscillate^{5,6,7,8} with timescales and amplitude proportional to the black hole mass and bolometric luminosity^{71,72}. The predicted light curves profiles show first a stable and slow rise in luminosity, as both temperature and surface density increase while matter is slowly accumulated. Thereafter a sharp luminosity burst is produced by a runaway increase (decrease) in temperature (surface density) propagating outwards within the unstable region. Finally, the inner flow, devoid of the matter supply, cools down rapidly and cycles back to the initial stable state with low temperature and density. Both heating and cooling fronts propagate following thermal timescales⁶, where $\tau_{\text{th}} \approx \alpha^{-1}(GM_{\text{BH}}/R^3)^{-1/2}$ (where R is the distance from the black hole). These so-called limit-cycle or ‘heartbeat’ instabilities have been successfully applied to a few accreting sources across all mass scales, for instance to the stellar-mass black holes GRS 1915+105^{20,21,73}, IGR J17091-3624²² and 4XMM J111816.0-324910⁷⁴ and to supermassive black holes in a statistical fashion^{71,72}. The similarity of their timing properties with those of QPEs in GSN 069 is remarkable and naturally led to the proposed connection with limit-cycle instabilities for that object. In particular, the symmetry of the eruptions in GSN 069 was compared to the fast heating and cooling phases of the instability¹, both following similar τ_{th} under the assumption of invariant α across the two phases⁷⁵. The lack of a slow rise before the eruptions in QPEs, predicted by the instability models, could be due to our limited coverage of the full disk temperature profile in the soft X-ray band.

With the observation of the two QPEs we report here we can now argue against at least this type of accretion disk instability as the origin of QPEs. Specifically, the strong asymmetric nature of the eruptions in eRO-QPE1, which show a faster rise and a much slower decay (Fig. 3a), argues against this interpretation. Qualitatively, our data would suggest that QPEs are not related to τ_{th} , because α is not expected to change between the hot and cold

phases in AGN⁷⁵. Moreover, if instead it is the front propagation timescales, following $\tau_{\text{front}} \approx (H/R)^{-1}\tau_{\text{th}}$ (where H is the vertical scale height of the disk), or viscous timescales, following $\tau_{\text{visc}} \approx (H/R)^{-2}\tau_{\text{th}}$, that regulates the rise (decay) in the cold (hot) phase, it would imply a thicker flow in the cold and stable phase than in the hot and unstable phase. This runs contrary to the theoretical expectation that unstable flows should be thicker⁶. The limit-cycle oscillation theory further predicts that once the period, duty cycle and luminosity amplitude are known and a viscosity prescription for the accretion flow is adopted, there are only specific values of M_{BH} and α that unstable sources are allowed to follow⁸. Here we adopt for eRO-QPE1 (eRO-QPE2) a peak-to-peak period $T_{\text{pp}} = 18.5$ h (2.4 h), an amplitude $A \approx 294$ (11) and a duty cycle $D = 41\%$ (19%). The amplitude A is the ratio of the disk luminosity (computed within the 0.001–100-keV range) for peak and quiescence, taken as proxy of the maximum and minimum bolometric luminosity, and D is here defined as the ratio of the flare duration (rise-to-decay T_{rd}) and the period T_{pp} . We begin by adopting a standard viscosity prescription, with the average stress between the annuli proportional to the gas plus radiation pressure⁶⁹ P_{tot} . The allowed M_{BH} and α values for eRO-QPE1 (eRO-QPE2) are $M_{\text{BH}} \approx 4 \times 10^6 M_{\odot}$ and $\alpha \approx 5$ ($M_{\text{BH}} \approx 3 \times 10^6 M_{\odot}$ and $\alpha \approx 3$), therefore an unphysically high viscosity parameter would be required. Considering alternative viscosity prescriptions^{5,8}, for eRO-QPE1 (eRO-QPE2) a more reasonable $\alpha \approx 0.1$ or 0.01 would correspond to allowed $M_{\text{BH}} \approx 2.4 \times 10^3 M_{\odot}$ or $M_{\text{BH}} \approx 60 M_{\odot}$ ($M_{\text{BH}} \approx 4.3 \times 10^3 M_{\odot}$ or $M_{\text{BH}} \approx 30 M_{\odot}$), respectively. The above combinations are either unphysical or very unlikely. Adopting $\alpha \approx 0.2$ and alternative viscosity prescriptions, eRO-QPE1 (eRO-QPE2) would yield an intermediate-mass black hole (IMBH) at $M_{\text{BH}} \approx 0.9 \times 10^4 M_{\odot}$ ($M_{\text{BH}} \approx 1.6 \times 10^4 M_{\odot}$) accreting at ~ 0.1 (~ 0.3) of the Eddington limit in quiescence and at ~ 30 (~ 3) times the Eddington limit at the peak. However, this IMBH scenario would not account for the opposite asymmetry shown by eRO-QPE1 compared to the theoretical predictions, nor would the resulting thermal timescales be self-consistent for either of the two: for eRO-QPE1 (eRO-QPE2) $\tau_{\text{th}} \approx 20$ s (35 s) at $20r_g$ adopting $M_{\text{BH}} \approx 0.9 \times 10^4 M_{\odot}$ ($M_{\text{BH}} \approx$

$1.6 \times 10^4 M_\odot$), which is orders of magnitude smaller than the observed QPE durations, and the rise-to-peak times would be only reconciled with τ_{th} at $\sim 780r_g$ ($\sim 250r_g$). Analogous results can be obtained using the observed properties of RX J1301.9+2747², adopting $T_{\text{pp}} \approx 20$ ks (or the second period $T_{\text{pp}} \approx 13$ ks), $D = 6\%$ (9%) and $A \approx 9.4$, the latter obtained taking the ratio of the quoted quiescence and peak 0.3–2.0-keV flux as proxy for a bolometric luminosity ratio: adopting² $\alpha \approx 0.15$ the allowed black hole mass is $\sim 2.2 \times 10^4 M_\odot$ ($\sim 1.5 \times 10^4 M_\odot$), much lower than the quoted^{2,76} $\sim (0.8 - 2.8) \times 10^6 M_\odot$.

When a given source is in a ‘sweet-spot’ regime in mass accretion rate, some more recent modified viscosity prescriptions of accretion disks predict the presence of a narrow unstable zone placed within an inner inefficient advection-dominated flow and an outer standard geometrically thin and stable flow⁹. This model would reduce the propagation timescales by a factor of approximately dR/R , where dR is the radial extent of the unstable zone at a distance R from the black hole, which may help reconcile the model with the dramatic and fast variability observed in CLAGN⁷⁷. This would not, however, change the inconsistency with the asymmetric shape of the QPEs we report here, nor was it successful in modelling all the observables in GSN 069⁹. In summary, our data for both of the QPEs reported here are inconsistent with published models for radiation pressure instability^{5,6,7,8,9}. The role of more complex or exotic phenomenology⁹ should be further explored.

We also note that a fast-rise exponential decay profile—such as the one in eRO-QPE1—can be naturally produced by ionization instability models, which are used for some bursting stellar-mass accreting compact objects⁷⁸. To our knowledge there is no evidence so far of such instabilities taking place in AGN⁷⁹. In addition, the predicted timescales are many orders of magnitude longer than QPEs for both AGN^{79,80,81} and IMBH masses⁸².

Finally, we discuss disk warping and tearing induced by Lense–Thirring precession^{83,84}, which has been recently qualitatively compared also to QPE sources⁸⁵. In this work we presented evidence of QPEs being observed in

previously inactive galaxies, therefore the accretion flow in these systems should be young. Moreover, a key element of disk warping and tearing due to Lense–Thirring precession is that mass needs to flow in from large inclination with respect to the black hole spin. Both conditions are satisfied if the accretion flow is formed, for instance, by a fully stripped TDE. However, in this case the warped inner flow would be damped very fast⁸⁶, which would be in contrast with QPEs lasting at least months¹ (Figs. 1, 2) or even years². A more quantitative comparison is beyond the reach of this work and of current disk warping and tearing simulations, but this is a promising scenario worth exploring in the future.

On the presence of an orbiting body

Periodic variability is also often associated with binary systems of compact objects²³ and the connection with the quasi-periodic nuclear emission observed in QPEs of interest for future work. We here assume the main body to be a supermassive black hole ranging between approximately $10^4 M_\odot$ and $10^7 M_\odot$ and we first consider the presence of a second orbiting supermassive black hole with a similar mass. There are several reasons which, when combined, disfavour such a scenario. First, simulations show that the accretion flow of such objects is composed by a circum-binary disk with two inner small mini-disks^{87,88,89}, which are thought to produce a quasi-sinusoidal modulated emission^{90,91}. This signature can be detected in transient surveys^{92,93} or in single sources⁹⁴, with a well known extreme case being OJ 287^{95,96}. However, so far there is no evidence of such variability in optical and UV data¹ of QPEs (Fig. 1c, 2c), in particular in eRO-QPE1, which was covered in the *g* and *r* bands by the Zwicky Transient Facility DR3⁹⁷ until the end of 2019. Nor can this prediction be reconciled with the dramatic non-sinusoidal eruptions observed in X-rays, even in the case of binary self-lensing⁹⁸ which can produce sharper bursts, albeit achromatic, and therefore in contrast with the energy dependence of QPEs^{1,2}. Moreover, we do not observe peculiar single- or double-peaked emission lines^{99,100,101} and this cannot be reconciled by enhanced obscuration¹⁰², because infrared photometry in QPEs is not AGN-like (WISE has observed stable $W_1 - W_2 \approx 0$ emission for the past 6–7 yr) and

X-rays do not indicate the presence of strong absorption. Second, supermassive black hole binaries are expected to form mostly via galactic mergers^{103,104}, but the host galaxies of the two QPEs we report here look unperturbed (Extended Data Figs. 1a, 2a). Perhaps most importantly, a binary of supermassive black holes observed with a periodicity of the order of hours, such as the four observed QPEs, would show a large period derivative due to gravitational wave emission, and would be relatively close to merger. To have (at least) four such objects very close to merger within $z \approx 0.02\text{--}0.05$ is very unlikely¹⁰⁵ and would imply that they are much more common in the local Universe than observations suggest^{92,93}.

Under the simplified assumption that the orbital evolution is dominated by gravitational waves emission, Extended Data Fig. 7 shows the allowed parameter space in terms of the derivative of the period, $\langle \dot{P} \rangle$, and M_2 for a range of $M_{\text{BH},1} \approx 10^4 M_\odot\text{--}10^7 M_\odot$ and zero- or high-eccentricity orbit ($e_O \approx 0.9$), given the rest-frame period of both eRO-QPE1 and eRO-QPE2. We have additionally imposed $M_2 \leq M_{\text{BH},1}$. For both sources we can draw a tentative line at the minimum period derivative that, if present, we would have measured already within the available observations: somewhat conservatively, we adopt a period decrease of one cycle over the 15 observed by NICER for eRO-QPE1 and the nine observed by XMM-Newton for eRO-QPE2 (Figs. 1d, 2c). Our constraint on $\langle \dot{P} \rangle$ is not very stringent for eRO-QPE1 and only high M_2 and eccentricities are disfavoured; instead, for eRO-QPE2 only an orbiting IMBH, or smaller, is allowed for zero eccentricity, whereas only a stellar-mass compact object is allowed for high eccentricity ($e_O \approx 0.9$). Future observations of eRO-QPE1 and eRO-QPE2 in the next months may lead to tighter constraints on the mass and eccentricity of the putative orbiting body.

The preliminary conclusion of our analysis is that, if QPEs are driven by the presence of an orbiting body around a central black hole, it is more likely that this is a compact object with a mass considerably smaller than the $\sim 10^4 M_\odot\text{--}10^7 M_\odot$ assumed for the main body. This scenario could make QPEs a viable candidate for the electromagnetic counterparts of the so-called extreme-mass-ratio inspirals (EMRIs)^{11,12,13}, with considerable implications for multi-messenger astrophysics and cosmology^{14,15}.

Interestingly, it has been recently suggested for GSN 069 that a stellar-mass compact object orbiting around a supermassive black hole could be the origin of QPEs: a white dwarf of $\sim 0.2M_{\odot}$ on a highly eccentric orbit ($e_O \approx 0.94$) could reproduce the mass inflow rate needed to produce the observed X-ray luminosity averaged over a QPE cycle¹⁰. This is reminiscent of a suggested, albeit still observationally elusive, EMRI-formation channel^{13,24,25,26}. For GSN 069, a possible explanation of the QPE-free X-ray bright and decaying phase could be given by an accretion flow expanding and intercepting the body at a later time¹; or if the orbiting body was originally a massive star and the stripped envelope produced the TDE-like behaviour of the past decade¹ while the remaining core started interacting with the newly born or expanded accretion flow only at a later stage, which would also explain the relatively small mass required by the white dwarf calculations¹⁰. For the other QPEs that did not show evidence of a past X-ray bright and decaying phase, this scenario is not necessary and the interaction with a second stellar-mass (or more massive) compact object could qualitatively reproduce the periodic behaviour (Extended Data Fig. 7b). Future X-ray observations of the known QPEs would help in further constraining the possible orbital evolution. It should be pointed out that these calculations so far only match the average observed QPE luminosity with the mass inflow rate required to produce it¹⁰, but details on the exact physical mechanism that would produce these X-ray bursts are still unknown (see [Methods](#) section ‘On accretion flow instabilities’).

Finally, we address the lack of UV and optical variability in the scenario of an orbiting body. The X-ray plateau at minimum shows an almost featureless accretion disk thermal spectrum^{1,2} (Extended Data Fig. 5), which could have been built up during the first orbiting cycles. This accretion flow should be unusually small due to the lack of a broad line region^{3,4} (Extended Data Figs. 1b, 2b), which would respond in light-days and that, if present, should have been observed in the SALT spectra taken months after the X-ray QPEs. The lack of strong UV and optical variability might be then due to the fact that the accretion disk is not large enough to even emit strong enough UV–optical radiation to emerge above the galaxy emission, which we can assume to be most of the observed $L \approx 4.0 \times 10^{41} \text{ erg s}^{-1}$ ($L \approx 4.3 \times 10^{41} \text{ erg s}^{-1}$) in the OM-UVW1 filter at

2,910 Å for eRO-QPE1 (eRO-QPE2). Using a simplified but physically motivated accretion disk model¹⁰⁶ for a spin-zero black hole accreting at ~ 0.1 of the Eddington limit, we computed the distance at which the bulk of 2,910-Å radiation would be emitted, namely $\sim 1,100r_g$ and $\sim 500r_g$ for masses of $10^5 M_\odot$ and $10^6 M_\odot$, respectively. This would shift to even larger radii for increasing accretion rate (for example, $\sim 1,850r_g$ and $\sim 860r_g$ at ~ 0.5 of the Eddington limit), whereas even for high spinning sources (dimensionless spin parameter ~ 0.9) the peak of OM-UVW1 flux would still come from $\sim 775r_g$ and $\sim 360r_g$. Furthermore, the predicted OM-UVW1 disk luminosity would be at least one or two orders of magnitude lower than the observed $L \approx 4.0 \times 10^{41}$ erg s $^{-1}$ in the most luminous scenario. Therefore, even a UV–optical eruption 100 times brighter than the plateau would be barely detectable above the galaxy component.

Predicted numbers

Detailed self-consistency calculations on the predicted rate of such EMRI events, as compared to QPE rates, are required but are beyond the scope of this paper. Instead, we can provide here a rough model-independent estimate of the expected number of observed QPEs, regardless of their origin. Convolving the black hole mass function¹⁰⁷ between approximately $10^{4.5} M_\odot$ and $10^{6.5} M_\odot$ up to $z \approx 0.03$ with the eROSITA sensitivity yields approximately 100 sources, a number that is then reduced with some educated guesses on some unknowns: during what fraction of their X-ray bright phase such sources undergo QPE behaviour (the biggest unknown; for example, >20% for GSN 069); how many such sources are obscured and missed ($\sim 2/3$); how many times we detect ongoing QPEs given the eROSITA sampling (depends on the period and the burst duration; for example, $\sim 20\%$ for GSN 069). This results in a (extremely uncertain) number of the order of unity per eRASS scan in the eROSITA-DE hemisphere; however, this is in agreement with our pilot study of the first few months of eROSITA operations. Thus, the low observed numbers do not necessarily imply that these events are a rare phenomenon intrinsically and they can in fact be a fairly common product of the co-evolution of black holes with their host galaxies²⁸. With a statistically meaningful

sample of QPEs, inverting this calculation may allow us to constrain the black hole mass function in a poorly known mass regime²⁷.

Data availability

With the exception of eROSITA proprietary data, the data used in this work are public and available from the corresponding data archives (XMM-Newton, <http://nxsa.esac.esa.int/nxsa-web/#search>; NICER, https://heasarc.gsfc.nasa.gov/docs/nicer/nicer_archive.html) or they will be soon, after the remaining proprietary period expires. Most of data may be available from the corresponding author on reasonable request.

References

1. 1.

Miniutti, G. et al. Nine-hour X-ray quasi-periodic eruptions from a low-mass black hole galactic nucleus. *Nature* **573**, 381–384 (2019).

[ADS](#) [CAS](#) [PubMed](#) [Article](#) [PubMed Central](#) [Google Scholar](#)

2. 2.

Giustini, M., Miniutti, G. & Saxton, R. X-ray quasi-periodic eruptions from the galactic nucleus of RX J1301.9+2747. *Astron. Astrophys.* **636**, L2 (2020).

[ADS](#) [Article](#) [Google Scholar](#)

3. 3.

Miniutti, G. et al. A high Eddington-ratio, true Seyfert 2 galaxy candidate: implications for broad-line region models. *Mon. Not. R. Astron. Soc.* **433**, 1764–1777 (2013).

[ADS](#) [Article](#) [Google Scholar](#)

4. 4.

Sun, L., Shu, X. W. & Wang, T. R. X. J1301.9+2747: a highly variable Seyfert galaxy with extremely soft X-ray emission. *Astrophys. J.* **768**, 167 (2013).

[ADS](#) [Article](#) [Google Scholar](#)

5. 5.

Merloni, A. & Nayakshin, S. On the limit-cycle instability in magnetized accretion discs. *Mon. Not. R. Astron. Soc.* **372**, 728–734 (2006).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

6. 6.

Janiuk, A., Czerny, B. & Siemiginowska, A. Radiation pressure instability driven variability in the accreting black holes. *Astrophys. J.* **576**, 908–922 (2002).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

7. 7.

Janiuk, A. & Czerny, B. On different types of instabilities in black hole accretion discs: implications for X-ray binaries and active galactic nuclei. *Mon. Not. R. Astron. Soc.* **414**, 2186–2194 (2011).

[ADS](#) [Article](#) [Google Scholar](#)

8. 8.

Grzędzielski, M., Janiuk, A., Czerny, B. & Wu, Q. Modified viscosity in accretion disks: application to Galactic black hole binaries, intermediate mass black holes, and active galactic nuclei. *Astron. Astrophys.* **603**, A110 (2017).

[ADS](#) [Article](#) [CAS](#) [Google Scholar](#)

9. 9.

Sniegowska, M., Czerny, B., Bon, E. & Bon, N. Possible mechanism for multiple changing-look phenomena in active galactic nuclei. *Astron. Astrophys.* **641**, A167 (2020).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

10. 10.

King, A. GSN 069 – a tidal disruption near miss. *Mon. Not. R. Astron. Soc.* **493**, L120–L123 (2020).

[ADS](#) [Article](#) [Google Scholar](#)

11. 11.

Amaro-Seoane, P. et al. Intermediate and extreme-mass-ratio inspirals —astrophysics, science applications and detection using LISA. *Class. Quantum Gravity* **24**, R113–R169 (2007).

[MathSciNet](#) [MATH](#) [Article](#) [Google Scholar](#)

12. 12.

Babak, S. et al. Science with the space-based interferometer LISA. V. Extreme mass-ratio inspirals. *Phys. Rev. D* **95**, 103012 (2017).

[ADS](#) [Article](#) [Google Scholar](#)

13. 13.

Wang, Y. Y., Wang, F. Y., Zou, Y. C. & Dai, Z. G. A bright electromagnetic counterpart to extreme mass ratio inspirals. *Astrophys. J. Lett.* **886**, 22 (2019).

[ADS](#) [Article](#) [CAS](#) [Google Scholar](#)

14. 14.

Bloom, J. S. et al. Astro2010 decadal survey whitepaper: coordinated science in the gravitational and electromagnetic skies. Preprint at <https://arxiv.org/abs/0902.1527> (2009).

15. 15.

Tamanini, N. Late time cosmology with LISA: probing the cosmic expansion with massive black hole binary mergers as standard sirens. *J. Phys. Conf. Ser.* **840**, 012029 (2017).

[Article](#) [CAS](#) [Google Scholar](#)

16. 16.

Predehl, P. et al. The eROSITA X-ray telescope on SRG. *Astron. Astrophys.* **647**, A1 (2021).

[Article](#) [Google Scholar](#)

17. 17.

McHardy, I. M., Koerding, E., Knigge, C., Uttley, P. & Fender, R. P. Active galactic nuclei as scaled-up Galactic black holes. *Nature* **444**, 730–732 (2006).

[ADS](#) [CAS](#) [PubMed](#) [Article](#) [PubMed Central](#) [Google Scholar](#)

18. 18.

Netzer, H. Revisiting the unified model of active galactic nuclei. *Annu. Rev. Astron. Astrophys.* **53**, 365–408 (2015).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

19. 19.

Chen, J. et al. The spatial extension of extended narrow line regions in MaNGA AGN. *Mon. Not. R. Astron. Soc.* **489**, 855–867 (2019).

[ADS](#) [Article](#) [Google Scholar](#)

20. 20.

Taam, R. E., Chen, X. & Swank, J. H. Rapid bursts from GRS 1915+105 with RXTE. *Astrophys. J.* **485**, L83–L86 (1997).

[ADS](#) [Article](#) [Google Scholar](#)

21. 21.

Belloni, T., Klein-Wolt, M., Méndez, M., van der Klis, M. & van Paradijs, J. A model-independent analysis of the variability of GRS 1915+105. *Astron. Astrophys.* **355**, 271–290 (2000).

[ADS](#) [Article](#) [Google Scholar](#)

22. 22.

Altamirano, D. et al. The faint ‘heartbeats’ of IGR J17091–3624: an exceptional black hole candidate. *Astrophys. J. Lett.* **742**, 17 (2011).

[ADS](#) [Article](#) [Google Scholar](#)

23. 23.

De Rosa, A. et al. The quest for dual and binary supermassive black holes: a multi-messenger view. *New Astron. Rev.* **86**, 101525 (2019).

[Article](#) [Google Scholar](#)

24. 24.

Sesana, A., Vecchio, A., Eracleous, M. & Sigurdsson, S. Observing white dwarfs orbiting massive black holes in the gravitational wave

and electro-magnetic window. *Mon. Not. R. Astron. Soc.* **391**, 718–726 (2008).

[ADS](#) [Article](#) [Google Scholar](#)

25. 25.

Zalamea, I., Menou, K. & Beloborodov, A. M. White dwarfs stripped by massive black holes: sources of coincident gravitational and electromagnetic radiation. *Mon. Not. R. Astron. Soc. Lett.* **409**, L25–L29 (2010).

[ADS](#) [Article](#) [Google Scholar](#)

26. 26.

Eracleous, M. et al. An arena for multi-messenger astrophysics: inspiral and tidal disruption of white dwarfs by massive black holes. *Bull. Am. Astron. Soc.* **51**, 10 (2019).

[Google Scholar](#)

27. 27.

Kelly, B. V. & Merloni, A. Mass functions of supermassive black holes across cosmic time. *Adv. Astron.* **2012**, 970858 (2012).

[ADS](#) [Article](#) [Google Scholar](#)

28. 28.

Heckman, T. M. & Best, P. N. The coevolution of galaxies and supermassive black holes: insights from surveys of the contemporary Universe. *Annu. Rev. Astron. Astrophys.* **52**, 589–660 (2014).

[ADS](#) [Article](#) [Google Scholar](#)

29. 29.

Amaro-Seoane, P. et al. Laser Interferometer Space Antenna. Preprint at <https://arxiv.org/abs/1702.00786> (2017).

30. 30.

Salvato, M. et al. Finding counterparts for all-sky X-ray surveys with NWAY: a Bayesian algorithm for cross-matching multiple catalogues. *Mon. Not. R. Astron. Soc.* **473**, 4937–4955 (2018).

[ADS](#) [Article](#) [Google Scholar](#)

31. 31.

Dey, A. et al. Overview of the DESI Legacy Imaging Surveys. *Astron. J.* **157**, 168 (2019).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

32. 32.

Buckley, D. A. H., Swart, G. P. & Meiring, J. G. Completion and commissioning of the Southern African Large Telescope. In *Proc. SPIE 6267: Ground-based and Airborne Telescopes* (ed. Stepp L. M.) 62670Z (SPIE, 2006).

33. 33.

Trümper, J. The ROSAT mission. *Adv. Space Res.* **2**, 241–249 (1982).

[ADS](#) [Article](#) [Google Scholar](#)

34. 34.

Saxton, R. D. et al. The first XMM-Newton slew survey catalogue: XMMSL1. *Astron. Astrophys.* **480**, 611–622 (2008).

[ADS](#) [Article](#) [Google Scholar](#)

35. 35.

Turner, M. J. L. et al. The European Photon Imaging Camera on XMM-Newton: the MOS cameras. *Astron. Astrophys.* **365**, L27–L35 (2001).

[ADS](#) [Article](#) [Google Scholar](#)

36. 36.

Strüder, L. et al. The European Photon Imaging Camera on XMM-Newton: the pn-CCD camera. *Astron. Astrophys.* **365**, L18–L26 (2001).

[ADS](#) [Article](#) [Google Scholar](#)

37. 37.

Mason, K. O. et al. The XMM-Newton optical/UV monitor telescope. *Astron. Astrophys.* **365**, L36–L44 (2001).

[ADS](#) [Article](#) [Google Scholar](#)

38. 38.

Gendreau, K. C., Arzoumanian, Z. & Okajima, T. The Neutron star Interior Composition ExploreR (NICER): an Explorer mission of opportunity for soft X-ray timing spectroscopy. In *Proc. SPIE 8443: Space Telescopes and Instrumentation 2012: Ultraviolet to Gamma Ray* (eds Takahashi, T. et al.) 844313 (SPIE, 2012).

39. 39.

Arzoumanian, Z. et al. The neutron star interior composition explorer (NICER): mission definition. In *Proc. SPIE 9144: Space Telescopes and Instrumentation 2014: Ultraviolet to Gamma Ray* (eds Takahashi, T. et al.) 914420 (SPIE, 2014).

40. 40.

Burgh, E. B. et al. Prime Focus Imaging Spectrograph for the Southern African Large Telescope: optical design. In *Proc. SPIE 4841: Instrument Design and Performance for Optical/Infrared Ground-based Telescopes* (eds Iye, M. & Moorwood, A. F. M.) 1463–1471 (SPIE, 2003).

41. 41.

Crawford, S. M. et al. PySALT: the SALT science pipeline. In *Proc. SPIE 7737: Observatory Operations: Strategies, Processes, and Systems III* (eds Silva, D. R. et al.) 773725 (SPIE, 2010).

42. 42.

Buckley, D. A. H. et al. A comparison between SALT/SAAO observations and kilonova models for AT 2017gfo: the first electromagnetic counterpart of a gravitational wave transient – GW170817. *Mon. Not. R. Astron. Soc. Lett.* **474**, L71–L75 (2018).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

43. 43.

Buchner, J. et al. X-ray spectral modelling of the AGN obscuring region in the CDFS: Bayesian model selection and catalogue. *Astron. Astrophys.* **564**, A125 (2014).

[Article](#) [CAS](#) [Google Scholar](#)

44. 44.

Buchner, J. Collaborative nested sampling: Big Data versus complex physical models. *Publ. Astron. Soc. Pacif.* **131**, 108005 (2019).

[ADS](#) [Article](#) [Google Scholar](#)

45. 45.

Arnaud, K. A. XSPEC: the first ten years. In *Astronomical Data Analysis Software and Systems V* (eds Jacoby, G. H. & Barnes, J.) http://iraf.noao.edu/ADASS/adass_proc/adass_95/arnaudk/arnaudk.html (1996).

46. 46.

Simmonds, C., Buchner, J., Salvato, M., Hsu, L.-T. & Bauer, F. E. X. Z. Deriving redshifts from X-ray spectra of obscured AGN. *Astron. Astrophys.* **618**, A66 (2018).

[Article](#) [CAS](#) [Google Scholar](#)

47. 47.

Cash, W. Parameter estimation in astronomy through application of the likelihood ratio. *Astrophys. J.* **228**, 939–947 (1979).

[ADS](#) [Article](#) [Google Scholar](#)

48. 48.

Wilms, J., Allen, A. & McCray, R. On the absorption of X-rays in the interstellar medium. *Astrophys. J.* **542**, 914–924 (2000).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

49. 49.

Mitsuda, K. et al. Energy spectra of low-mass binary X-ray sources observed from TENMA. *Publ. Astron. Soc. Jpn.* **36**, 741–759 (1984).

[ADS](#) [CAS](#) [Google Scholar](#)

50. 50.

HI4PI Collaboration. HI4PI: a full-sky H I survey based on EBHIS and GASS. *Astron. Astrophys.* **594**, A116 (2016).

[Article](#) [CAS](#) [Google Scholar](#)

51. 51.

Van Velzen, S. The first tidal disruption flare in ZTF: from photometric selection to multi-wavelength characterization. *Astrophys. J.* **872**, 198 (2019).

[ADS](#) [Article](#) [CAS](#) [Google Scholar](#)

52. 52.

Silverman, J. D. et al. Ongoing and co-evolving star formation in zCOSMOS galaxies hosting active galactic nuclei. *Astrophys. J.* **696**, 396–410 (2009).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

53. 53.

Zabludoff, A. I. et al. The environment of ‘E+A’ galaxies. *Astrophys. J.* **466**, 104–113 (1996).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

54. 54.

Comparat, J. Investigating emission-line galaxy surveys with the Sloan Digital Sky Survey infrastructure. *Mon. Not. R. Astron. Soc.* **428**, 1498–1517 (2013).

[ADS](#) [Article](#) [Google Scholar](#)

55. 55.

Arcavi, I. et al. A continuum of H- to He-rich tidal disruption candidates with a preference for E+A galaxies. *Astrophys. J.* **793**, 38 (2014).

[ADS](#) [Article](#) [CAS](#) [Google Scholar](#)

56. 56.

French, K. D., Arcavi, I. & Zabludoff, A. Tidal disruption events prefer unusual host galaxies. *Astrophys. J. Lett.* **818**, 21 (2016).

[ADS](#) [Article](#) [CAS](#) [Google Scholar](#)

57. 57.

Law-Smith, J., Ramirez-Ruiz, E., Ellison, S. L. & Foley, R. J. Tidal disruption event host galaxies in the context of the local galaxy population. *Astrophys. J.* **850**, 22 (2017).

[ADS](#) [Article](#) [CAS](#) [Google Scholar](#)

58. 58.

Van Velzen, S. et al. Seventeen tidal disruption events from the first half of ZTF survey observations: entering a new era of population studies. *Astrophys. J.* **908**, 4 (2021).

[ADS](#) [Article](#) [CAS](#) [Google Scholar](#)

59. 59.

Wilkinson, D. M., Maraston, C., Goddard, D., Thomas, D. & Parikh, T. FIREFLY (Fitting IterativEly For Likelihood analYsis): a full spectral fitting code. *Mon. Not. R. Astron. Soc.* **472**, 4297–4326 (2017).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

60. 60.

Maraston, C. et al. Stellar population models based on the SDSS-IV MaStar library of stellar spectra—I. Intermediate-age/old models. *Mon. Not. R. Astron. Soc.* **496**, 2962–2997 (2020).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

61. 61.

Abbott, T. M. C. et al. The Dark Energy Survey: data release 1. *Astrophys. J. Suppl. Ser.* **239**, 18 (2018).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

62. 62.

Maraston, C. & Stroemback, G. Stellar population models at high spectral resolution. *Mon. Not. R. Astron. Soc.* **418**, 2785–2811 (2011).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

63. 63.

Kennicutt, R. C., Jr. Star formation in galaxies along the Hubble sequence. *Annu. Rev. Astron. Astrophys.* **36**, 189–231 (1998).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

64. 64.

Kewley, L. J., Geller, M. J. & Jansen, R. A. [O ii] as a star formation rate indicator. *Astron. J.* **127**, 2002–2030 (2004).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

65. 65.

Dodd, S. A., Law-Smith, J. A. P., Auchettl, K., Ramirez-Ruiz, E. & Foley, R. J. The landscape of galaxies harboring changing-look active galactic nuclei in the local Universe. *Astrophys. J. Lett.* **907**, 21 (2021).

[ADS](#) [Article](#) [CAS](#) [Google Scholar](#)

66. 66.

Brinchmann, J. et al. The physical properties of star-forming galaxies in the low-redshift Universe. *Mon. Not. R. Astron. Soc.* **351**, 1151–1179 (2004).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

67. 67.

Hammerstein, E. et al. Tidal disruption event hosts are green and centrally concentrated: signatures of a post-merger system. *Astrophys. J. Lett.* **908**, 20 (2021).

[ADS](#) [Article](#) [Google Scholar](#)

68. 68.

Reines, A. E. & Volonteri, M. Relations between central black hole mass and total galaxy stellar mass in the local Universe. *Astrophys. J.* **813**, 82 (2015); correction **879**, 133 (2019).

[ADS](#) [Article](#) [Google Scholar](#)

69. 69.

Shakura, N. I. & Sunyaev, R. A. Black holes in binary systems. *Observational appearance*. *Astron. Astrophys.* **24**, 337–355 (1973).

[Google Scholar](#)

70. 70.

Lightman, A. P. & Eardley, D. M. Black holes in binary systems: instability of disk accretion. *Astrophys. J.* **187**, L1–L3 (1974).

[ADS](#) [Article](#) [Google Scholar](#)

71. 71.

Czerny, B., Siemiginowska, A., Janiuk, A., Nikiel-Wroczyński, B. & Stawarz, Ł. Accretion disk model of short-timescale intermittent activity in young radio sources. *Astrophys. J.* **698**, 840–851 (2009).

[ADS](#) [Article](#) [Google Scholar](#)

72. 72.

Wu, Q. et al. The Universal ‘heartbeat’ oscillations in black hole systems across the mass-scale. *Astrophys. J.* **833**, 79 (2016).

[ADS](#) [Article](#) [Google Scholar](#)

73. 73.

Neilsen, J., Remillard, R. A. & Lee, J. C. The physics of the ‘heartbeat’ state of GRS 1915+105. *Astrophys. J.* **737**, 69 (2011).

[ADS](#) [Article](#) [CAS](#) [Google Scholar](#)

74. 74.

Motta, S. et al. The slow heartbeats of an ultraluminous X-ray source in NGC 3621. *Astrophys. J.* **898**, 174 (2020).

[ADS](#) [Article](#) [Google Scholar](#)

75. 75.

Janiuk, A., Czerny, B., Siemiginowska, A. & Szczerba, R. On the turbulent α -disks and the intermittent activity in active galactic nuclei. *Astrophys. J.* **602**, 595–602 (2004).

[ADS](#) [Article](#) [Google Scholar](#)

76. 76.

Shu, X. W. et al. Central engine and host galaxy of RXJ 1301.9+2747: a multiwavelength view of a low-mass black hole active galactic

nuclei with ultra-soft X-ray emission. *Astrophys. J.* **837**, 3 (2017).

[ADS](#) [Article](#) [Google Scholar](#)

77. 77.

MacLeod, C. L. et al. A systematic search for changing-look quasars in SDSS. *Mon. Not. R. Astron. Soc.* **457**, 389–404 (2016).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

78. 78.

Hameury, J. M. A review of the disc instability model for dwarf novae, soft X-ray transients and related objects. *Adv. Space Res.* **66**, 1004–1024 (2020).

[ADS](#) [Article](#) [Google Scholar](#)

79. 79.

Hameury, J. M., Viallet, M. & Lasota, J.-P. The thermal–viscous disk instability model in the AGN context. *Astron. Astrophys.* **496**, 413–421 (2009).

[ADS](#) [Article](#) [Google Scholar](#)

80. 80.

Lin, D. N. C. & Shields, G. A. Accretion disks and periodic outbursts of active galactic nuclei. *Astrophys. J.* **305**, 28 (1986).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

81. 81.

Yan, Z. & Xie, F.-G. A decades-long fast-rise-exponential-decay flare in low-luminosity AGN NGC 7213. *Mon. Not. R. Astron. Soc.* **475**, 1190–1197 (2018).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

82. 82.

Lasota, J.-P. et al. The origin of variability of the intermediate-mass black-hole ULX system HLX-1 in ESO 243-49. *Astrophys. J.* **735**, 89 (2011).

[ADS](#) [Article](#) [Google Scholar](#)

83. 83.

Nixon, C., King, A., Price, D. & Frank, J. Tearing up the disk: how black holes accrete. *Astrophys. J. Lett.* **757**, 24 (2012).

[ADS](#) [Article](#) [Google Scholar](#)

84. 84.

Raj, A., Nixon, C. & Dogan, S. Disk tearing: numerical investigation of warped disk instability. *Astrophys. J.* **909**, 81 (2021).

[ADS](#) [Article](#) [Google Scholar](#)

85. 85.

Raj, A. & Nixon, C. Disk tearing: implications for black hole accretion and AGN variability. *Astrophys. J.* **909**, 82 (2021).

[ADS](#) [Article](#) [Google Scholar](#)

86. 86.

Franchini, A., Lodato, G. & Facchini, S. Lense–Thirring precession around supermassive black holes during tidal disruption events. *Mon. Not. R. Astron. Soc.* **455**, 1946–1956 (2016).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

87. 87.

Farris, B. D., Duffell, P., MacFadyen, A. I. & Haiman, Z. Binary black hole accretion from a circumbinary disk: gas dynamics inside the central cavity. *Astrophys. J.* **783**, 134 (2014).

[ADS](#) [Article](#) [Google Scholar](#)

88. 88.

Bowen, D. B. et al. Quasi-periodic behavior of mini-disks in binary black holes approaching merger. *Astrophys. J. Lett.* **853**, 17 (2018).

[ADS](#) [Article](#) [CAS](#) [Google Scholar](#)

89. 89.

Tang, Y., Haiman, Z. & MacFadyen, A. The late inspiral of supermassive black hole binaries with circumbinary gas discs in the LISA band. *Mon. Not. R. Astron. Soc.* **476**, 2249–2257 (2018).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

90. 90.

Farris, B. D., Duffell, P., MacFadyen, A. I. & Haiman, Z. Characteristic signatures in the thermal emission from accreting binary black holes. *Mon. Not. R. Astron. Soc.* **446**, L36–L40 (2015).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

91. 91.

D’Orazio, D. J., Haiman, Z. & Schiminovich, D. Relativistic boost as the cause of periodicity in a massive black-hole binary candidate. *Nature* **525**, 351–353 (2015).

[ADS](#) [PubMed](#) [Article](#) [CAS](#) [PubMed Central](#) [Google Scholar](#)

92. 92.

Graham, M. J. A systematic search for close supermassive black hole binaries in the Catalina Real-time Transient Survey. *Mon. Not. R. Astron. Soc.* **453**, 1562–1576 (2015).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

93. 93.

Charisi, M. et al. A population of short-period variable quasars from PTF as supermassive black hole binary candidates. *Mon. Not. R. Astron. Soc.* **463**, 2145–2171 (2016).

[ADS](#) [Article](#) [Google Scholar](#)

94. 94.

Graham, M. J. et al. A possible close supermassive black-hole binary in a quasar with optical periodicity. *Nature* **518**, 74–76 (2015).

[ADS](#) [CAS](#) [PubMed](#) [Article](#) [PubMed Central](#) [Google Scholar](#)

95. 95.

Sillanpää, A., Haarala, S., Valtonen, M. J., Sundelius, B. & Byrd, G. G. OJ 287: binary pair of supermassive black holes. *Astrophys. J.* **325**, 628–634 (1988).

[ADS](#) [Article](#) [Google Scholar](#)

96. 96.

Valtonen, M. et al. A massive binary black-hole system in OJ 287 and a test of general relativity. *Nature* **452**, 851–853 (2008).

[ADS](#) [CAS](#) [PubMed](#) [Article](#) [PubMed Central](#) [Google Scholar](#)

97. 97.

Graham, M. J. et al. The Zwicky Transient Facility: science objectives. *Publ. Astron. Soc. Pacif.* **131**, 078001 (2019).

[ADS](#) [Article](#) [Google Scholar](#)

98. 98.

D’Orazio, D. J. & Di Stefano, R. Periodic self-lensing from accreting massive black hole binaries. *Mon. Not. R. Astron. Soc.* **474**, 2975–2986 (2018).

[ADS](#) [Article](#) [Google Scholar](#)

99. 99.

Wang, J.-M. et al. Active galactic nuclei with double-peaked narrow lines: are they dual active galactic nuclei? *Astrophys. J.* **705**, L76–L80 (2009).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

100. 100.

Smith, K. L. et al. A search for binary active galactic nuclei: double-peaked [O iii] AGNs in the Sloan Digital Sky Survey. *Astrophys. J.* **716**, 866–877 (2010).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

101. 101.

Eracleous, M., Boroson, T. A., Halpern, J. P. & Liu, J. A large systematic search for close supermassive binary and rapidly recoiling black holes. *Astrophys. J. Suppl. Ser.* **201**, 23 (2012).

[ADS](#) [Article](#) [CAS](#) [Google Scholar](#)

102. 102.

Pfeifle, R. W. et al. Buried black hole growth in IR-selected mergers: new results from Chandra. *Astrophys. J.* **875**, 117 (2019).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

103. 103.

Khan, F. M., Fiacconi, D., Mayer, L., Berczik, P. & Just, A. Swift coalescence of supermassive black holes in cosmological mergers of massive galaxies. *Astrophys. J.* **828**, 73 (2016).

[ADS](#) [Article](#) [Google Scholar](#)

104. 104.

Kelley, L. Z., Blecha, L. & Hernquist, L. Massive black hole binary mergers in dynamical galactic environments. *Mon. Not. R. Astron. Soc.* **464**, 3131–3157 (2017).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

105. 105.

Payne, A. et al. ASASSN-14ko is a periodic nuclear transient in ESO 253-G003. Preprint at <https://arxiv.org/abs/2009.03321> (2020).

106. 106.

Arcodia, R., Merloni, A., Nandra, K. & Ponti, G. Testing the disk–corona interplay in radiatively-efficient broad-line AGN. *Astron. Astrophys.* **628**, A135 (2019).

[ADS](#) [Article](#) [Google Scholar](#)

107. 107.

Merloni, A. & Heintz, S. A synthesis model for AGN evolution: supermassive black holes growth and feedback modes. *Mon. Not. R. Astron. Soc.* **388**, 1011–1030 (2008).

[ADS](#) [CAS](#) [Google Scholar](#)

[Download references](#)

Acknowledgements

We are grateful to the XMM-Newton Project Scientist N. Schartel for the DDT observation of eRO-QPE2; we also thank the XMM-Newton Science Operations team for performing it along with the two approved ToOs on eRO-QPE1 with a relatively short notice before the end of the visibility window. R.A. thanks M. Gilfanov, K. Dennerl and S. Carpano for discussions. G.P. acknowledges funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement no. 865637). This work is based on data from eROSITA, the primary instrument aboard SRG, a joint Russian–German science mission supported by the Russian Space Agency (Roskosmos), in the interests of the Russian Academy of Sciences represented by its Space Research Institute (IKI), and the Deutsches Zentrum für Luft- und Raumfahrt (DLR). The SRG spacecraft was built by Lavochkin Association (NPOL) and its subcontractors, and is operated by NPOL with support from the Max Planck Institute for Extraterrestrial Physics (MPE). The development and construction of the eROSITA X-ray instrument was led by MPE, with contributions from the Dr. Karl Remeis Observatory Bamberg & ECAP (FAU Erlangen-Nuernberg), the University of Hamburg Observatory, the Leibniz Institute for Astrophysics Potsdam, and the Institute for Astronomy and Astrophysics of the University of Tübingen, with the support of DLR and the Max Planck Society. The Argelander Institute for Astronomy of the University of Bonn and the Ludwig Maximilians Universität München also participated in the science preparation for eROSITA. The eROSITA data shown here were processed using the eSASS/NRTA software system developed by the German eROSITA consortium. Some of the observations were obtained using the Southern African Large Telescope (SALT) as part of the Large Science Programme on transient (2018-2-LSP-001; PI: D.A.H.B.). Polish participation in SALT is funded by grant no. MNiSW DIR/WK/2016/07. D.A.H.B. is supported by the National Research Foundation (NRF) of South Africa. M.G. is supported by the Polish NCN MAESTRO grant

2014/14/A/ST9/00121. M.K. is supported by DFG grant KR 3338/4-1. This work is partly supported by the Optical and Near-infrared Astronomy Inter-University Cooperation Program and the Grants-in-Aid of the Ministry of Education of Japan.

Funding

Open access funding provided by Max Planck Society.

Author information

Affiliations

1. Max-Planck-Institut für Extraterrestrische Physik, Garching, Germany
R. Arcodia, A. Merloni, K. Nandra, J. Buchner, M. Salvato, J. Comparat, G. Ponti, A. Malyali, J. Wolf, D. Bogensberger, M. E. Ramos-Ceja & A. Rau
2. MIT Kavli Institute for Astrophysics and Space Research, Cambridge, MA, USA
D. Pasham, R. Remillard & E. Kara
3. Leibniz-Institut für Astrophysik Potsdam (AIP), Potsdam, Germany
G. Lamer, M. Krumpe & A. Schwope
4. INAF—Osservatorio Astronomico di Brera, Merate, Italy
G. Ponti
5. Astrophysics Science Division, NASA Goddard Space Flight Center, Greenbelt, MD, USA
Z. Arzoumanian, K. Gendreau & C. Markwardt
6. South African Astronomical Observatory, Cape Town, South Africa
D. A. H. Buckley
7. Astronomical Observatory, University of Warsaw, Warsaw, Poland

M. Gromadzki

8. Graduate School of Science and Engineering, Saitama University,
Saitama, Japan

M. Schramm

Authors

1. R. Arcodia

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. A. Merloni

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. K. Nandra

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. J. Buchner

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. M. Salvato

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. D. Pasham

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. R. Remillard

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. J. Comparat

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. G. Lamer

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. G. Ponti

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

11. A. Malyali

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

12. J. Wolf

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

13. Z. Arzoumanian

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

14. D. Bogensberger

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

15. D. A. H. Buckley

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

16. K. Gendreau

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

17. M. Gromadzki

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

18. E. Kara

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

19. M. Krumpe

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

20. C. Markwardt

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

21. M. E. Ramos-Ceja

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

22. A. Rau

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

23. M. Schramm

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

24. A. Schwope

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Contributions

R.A. wrote the article, developed the method to search for QPE candidates, was PI of XMM-Newton and NICER follow-up observations, and performed most of the data analysis. A. Merloni is PS/PI of eROSITA and followed and improved the project, the strategy of XMM-Newton and NICER follow-up observing campaigns and the writing of the article. K.N. followed and improved the project, the strategy of XMM-Newton and NICER follow-up proposals and the writing of the article. J.B. developed part of the code used for spectral and timing analysis and helped R.A. with XMM-Newton data analysis and with the development of the method to search for QPE candidates. M. Salvato led the counterpart association and significantly contributed to the interpretation and analysis of the photometry and optical spectroscopy. D.P. and R.R. performed most of the NICER data analysis. J.C. performed some of the analysis of optical spectroscopy and improved its interpretation. G.L. performed the astrometry of eROSITA and XMM-Newton data. G.P. significantly improved the writing of the article and encouraged the pursuit of this project. A. Malyali and J.W. directly contributed to the analysis of XMM-Newton light curves, optical images and counterpart association. Z.A., K.G. and C.M. scheduled, monitored and processed NICER observations. D.A.H.B. instigated the SALT observations and M.G. processed related data. E.K. and M.K. improved the scientific interpretation and presentation of the manuscript. M.E.R.-C. contributed to development of the method to search for QPE candidates. A.R. contributed to XMM-Newton observing proposals and with A.S. led the triggering of SALT observations from the eROSITA-DE Consortium. M. Schramm performed and processed photometry data from the Rapid Response

Robotic Telescope. All authors contributed to improve the projects from different sides throughout the process.

Corresponding author

Correspondence to [R. Arcodia](#).

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 eRO-QPE1 position and identification.](#)

a, Legacy DR8 image cut-out around the optical counterpart of eRO-QPE1. Red and green circles represent the astrometry-corrected eROSITA and XMM-Newton EPIC-pn positions, respectively, with 1σ positional uncertainties. The EPIC-pn position was corrected excluding the target (blue cross) to ensure an unbiased estimate of the possible positional offset. Image reproduced from Legacy Surveys / D. Lang (Perimeter Institute) under a CC-BY-4.0 licence. **b**, SALT spectra of eRO-QPE1 shown in black and blue with related 1σ errors as shaded regions. The cyan spectrum

represents a re-normalized sky spectrum to guide the eye for the residual sky feature around 5,577 Å.

Extended Data Fig. 2 eRO-QPE2 position and identification.

As in Extended Data Fig. 1, for eRO-QPE2. Green pixels in **a** are artefacts or missing data in the optical image. Image in **a** reproduced from Legacy Surveys / D. Lang (Perimeter Institute) under a CC-BY-4.0 licence.

Extended Data Fig. 3 eRO-QPE1 spectral fit results.

XMM-Newton EPIC-pn light-curve (top panel) and time-resolved spectroscopy fit results for spectra extracted in the 500-s time bins (bottom panels) of the two XMM-Newton observations of eRO-QPE1 (left, XMM1; right, XMM2) using an accretion disk model (diskbb): in particular, the evolution of the peak accretion disk temperature (T_{in} , $k = k_{\text{B}}$) and the normalization (norm), which is proportional to the inner radius once distance and inclination are known. The time evolution of the 0.5–2.0-keV flux ($F_{0.5-2.0\text{keV}}$) and luminosity ($L_{0.5-2.0\text{keV}}$) is also shown in the bottom panel. The quiescence level is fitted by combining the first part of both XMM-Newton observations. It is shown with a dashed line because, due to low counts, the fit is more uncertain (see Extended Data Fig. 5a). Median fit values and fluxes of the high and low eROSITA states are reported with orange and red arrows pointing left (upper limits (UL) are denoted with diagonal arrows). 1σ uncertainties on the fit results are shown with shaded regions around the median.

Extended Data Fig. 4 eRO-QPE2 spectral fit results.

As in Extended Data Fig. 3, for eRO-QPE2. Here the eROSITA upper limit (UL) of the low state is reported at 3σ .

Extended Data Fig. 5 eRO-QPE1 and eRO-QPE2 spectra.

a, XMM-Newton EPIC-pn source plus background (bkg) spectra for eRO-QPE1. Red, orange and green data correspond to quiescence and to the peak

of the second and first XMM-Newton observations, respectively, with error bars showing 1σ uncertainties. The related solid lines show the unabsorbed source model obtained with diskbb, just for visualization. The grey line represents the background spectrum alone. The plateau is shown with a dotted line because, due to low counts, the fit is more uncertain. **b**, As in **a**, for eRO-QPE2. Here green data represent one of the peaks and the additional dashed lines indicate the absorbed source model.

Extended Data Fig. 6 The properties of the host galaxies of the QPEs.

Stellar mass M_* and star-formation rate (SFR) for eRO-QPE1 (blue) and eRO-QPE2 (red), with related 1σ uncertainties; for eRO-QPE1, SFR is largely unconstrained (see [Methods](#) section ‘The host galaxies of the QPEs’). For a comparison, normal galaxies⁶⁶, TDEs⁵⁷ and CLAGN⁶⁵, all below $z < 0.1$, are also shown.

Extended Data Fig. 7 Constraints on a secondary orbiting body.

a, Allowed parameter space in terms of the derivative of the period \dot{P} and secondary mass M_2 for a range of primary mass $M_{\text{BH},1} \approx 10^4 M_\odot - 10^7 M_\odot$ and zero (solid lines) or high orbital eccentricity ($e_O \approx 0.9$, dotted lines), in which we can reproduce the rest-frame period of eRO-QPE1. We have additionally imposed $M_2 \leq M_{\text{BH},1}$. We have drawn an approximate threshold at the minimum value of \dot{P} which we would have measured within the available observations, corresponding to a period decrease of one QPE cycle over the 15 observed by NICER (Fig. [1d](#)). The excluded region is shaded in red. **b**, As in **a**, for eRO-QPE2, and adopting as tentative minimum \dot{P} a period decrease of one cycle over the nine observed with XMM-Newton (Fig. [2c](#)).

Extended Data Table 1 Summary of the observations performed

[Full size table](#)

Extended Data Table 2 Summary of spectral fit results for eRO-QPE1

[Full size table](#)

Extended Data Table 3 Summary of spectral fit results for eRO-QPE2

[Full size table](#)

Supplementary information

[Peer Review File](#)

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

[Reprints and Permissions](#)

About this article



Check for
updates

Cite this article

Arcodia, R., Merloni, A., Nandra, K. *et al.* X-ray quasi-periodic eruptions from two previously quiescent galaxies. *Nature* **592**, 704–707 (2021).
<https://doi.org/10.1038/s41586-021-03394-6>

Download citation

- Received: 23 November 2020
- Accepted: 25 February 2021
- Published: 28 April 2021
- Issue Date: 29 April 2021
- DOI: <https://doi.org/10.1038/s41586-021-03394-6>

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

Download PDF

Associated Content

Massive black holes flaring up time and again

- Tingting Liu

Nature Astronomy News & Views 30 Apr 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03394-6>

- Article
- [Published: 28 April 2021](#)

Transition from an atomic to a molecular Bose–Einstein condensate

- [Zhendong Zhang](#) ORCID: orcid.org/0000-0001-7221-1729¹,
- [Liangchao Chen](#)²,
- [Kai-Xuan Yao](#) ORCID: orcid.org/0000-0002-4071-7060¹ &
- [Cheng Chin](#) ORCID: orcid.org/0000-0003-0278-4630¹

Nature volume **592**, pages 708–711(2021) [Cite this article](#)

- 1689 Accesses
- 133 Altmetric
- [Metrics details](#)

Subjects

- [Quantum mechanics](#)
- [Ultracold gases](#)

Abstract

Molecular quantum gases (that is, ultracold and dense molecular gases) have many potential applications, including quantum control of chemical reactions, precision measurements, quantum simulation and quantum information processing^{1,2,3}. For molecules, to reach the quantum regime

usually requires efficient cooling at high densities, which is frequently hindered by fast inelastic collisions that heat and deplete the population of molecules^{4,5}. Here we report the preparation of two-dimensional Bose–Einstein condensates (BECs) of spinning molecules by inducing pairing interactions in an atomic condensate near a *g*-wave Feshbach resonance⁶. The trap geometry and the low temperature of the molecules help to reduce inelastic loss, ensuring thermal equilibrium. From the equation-of-state measurement, we determine the molecular scattering length to be + 220(± 30) Bohr radii (95% confidence interval). We also investigate the unpairing dynamics in the strong coupling regime and find that near the Feshbach resonance the dynamical timescale is consistent with the unitarity limit. Our work demonstrates the long-sought transition between atomic and molecular condensates, the bosonic analogue of the crossover from a BEC to a Bardeen–Cooper–Schrieffer (BCS) superfluid in a Fermi gas^{7,8,9}. In addition, our experiment may shed light on condensed pairs with orbital angular momentum, where a novel anisotropic superfluid with non-zero surface current is predicted^{10,11}, such as the A phase of ${}^3\text{He}$.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

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: Production of *g*-wave molecular condensate.

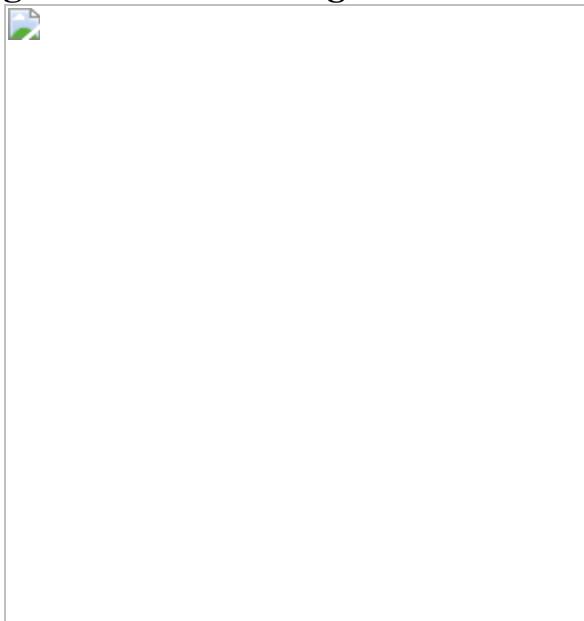


Fig. 2: Equation of state of molecular gases.

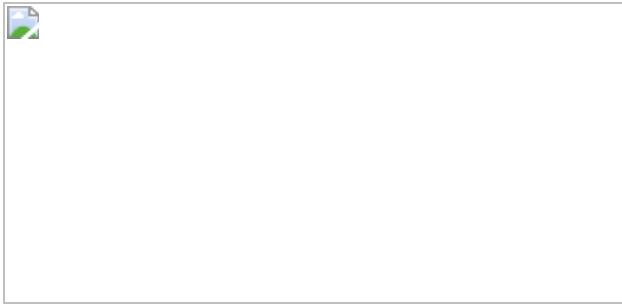


Fig. 3: Stability of g -wave molecular condensate.

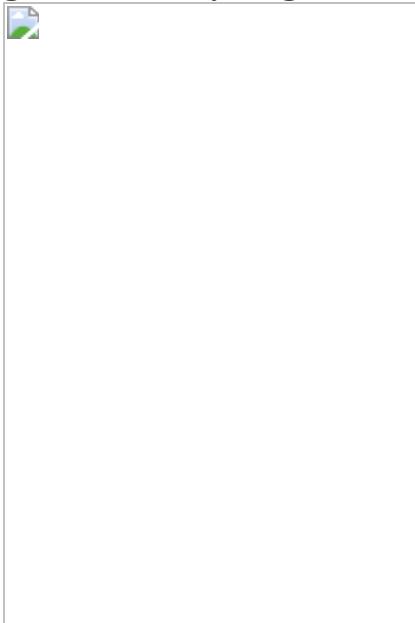
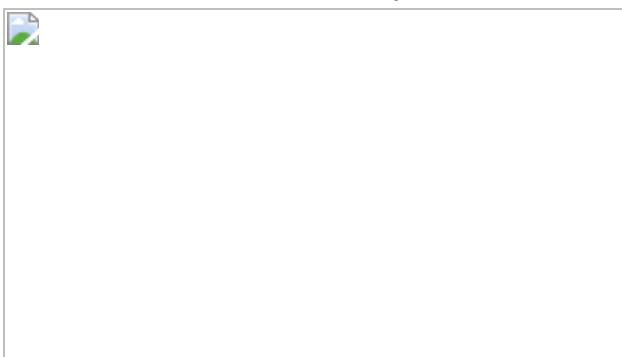


Fig. 4: Unpairing dynamics in a molecular condensate near the g -wave Feshbach resonance at $B_0 = 19.874$ G.



Data availability

The data that support the plots within this paper and other findings of this study are available from the corresponding author upon reasonable request. [Source data](#) are provided with this paper.

Code availability

The codes for the analysis of data shown within this paper are available from the corresponding author upon reasonable request.

References

1. 1.

Bohn, J. L., Rey, A. M. & Ye, J. Cold molecules: progress in quantum engineering of chemistry and quantum matter. *Science* **357**, 1002–1010 (2017).

[ADS](#) [MathSciNet](#) [CAS](#) [Article](#) [Google Scholar](#)

2. 2.

Carr, L. D., DeMille, D., Krems, R. V. & Ye, J. Cold and ultracold molecules: science, technology and applications. *New J. Phys.* **11**, 055049 (2009).

[ADS](#) [Article](#) [Google Scholar](#)

3. 3.

Quéméner, G. & Julienne, P. S. Ultracold molecules under control! *Chem. Rev.* **112**, 4949–5011 (2012).

[Article](#) [Google Scholar](#)

4. 4.

Mayle, M., Quéméner, G., Ruzic, B. P. & Bohn, J. L. Scattering of ultracold molecules in the highly resonant regime. *Phys. Rev. A* **87**, 012709 (2013).

[ADS](#) [Article](#) [Google Scholar](#)

5. 5.

Idziaszek, Z. & Julienne, P. S. Universal rate constants for reactive collisions of ultracold molecules. *Phys. Rev. Lett.* **104**, 113202 (2010).

[ADS](#) [Article](#) [Google Scholar](#)

6. 6.

Köhler, T., Góral, K. & Julienne, P. S. Production of cold molecules via magnetically tunable Feshbach resonances. *Rev. Mod. Phys.* **78**, 1311–1361 (2006).

[ADS](#) [Article](#) [Google Scholar](#)

7. 7.

Romans, M. W. J., Duine, R. A., Sachdev, S. & Stoof, H. T. C. Quantum phase transition in an atomic Bose gas with a Feshbach resonance. *Phys. Rev. Lett.* **93**, 020405 (2004).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

8. 8.

Radzhovskiy, L., Park, J. & Weichman, P. B. Superfluid transitions in bosonic atom-molecule mixtures near a Feshbach resonance. *Phys. Rev. Lett.* **92**, 160402 (2004).

[ADS](#) [Article](#) [Google Scholar](#)

9. 9.

Duine, R. A. & Stoof, H. T. C. Atom–molecule coherence in Bose gases. *Phys. Rep.* **396**, 115–195 (2004).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

10. 10.

Anderson, P. W. & Morel, P. Generalized Bardeen–Cooper–Schrieffer states and the proposed low-temperature phase of liquid He³. *Phys. Rev.* **123**, 1911 (1961).

11. 11.

Ho, T.-L. The Bose-Einstein condensate of g-wave molecules and its intrinsic angular momentum. Preprint at <https://arxiv.org/abs/2101.05431> (2021).

12. 12.

Chen, Q., Stajic, J., Tan, S. & Levin, K. BCS-BEC crossover: from high temperature superconductors to ultracold superfluids. *Phys. Rep.* **412**, 1–88 (2005).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

13. 13.

Giorgini, S., Pitaevskii, L. P. & Stringari, S. Theory of ultracold atomic Fermi gases. *Rev. Mod. Phys.* **80**, 1215–1274 (2008).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

14. 14.

Timmermans, E., Furuya, K., Milonni, P. W. & Kerman, A. K. Prospect of creating a composite Fermi-Bose superfluid. *Phys. Lett. A* **285**, 228–233 (2001).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

15. 15.

De Marco, L. et al. A degenerate Fermi gas of polar molecules. *Science* **363**, 853–856 (2019).

[ADS](#) [Article](#) [Google Scholar](#)

16. 16.

Chin, C., Grimm, R., Julienne, P. & Tiesinga, E. Feshbach resonances in ultracold gases. *Rev. Mod. Phys.* **82**, 1225–1286 (2010).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

17. 17.

Krüger, P., Hadzibabic, Z. & Dalibard, J. Critical point of an interacting two-dimensional atomic Bose gas. *Phys. Rev. Lett.* **99**, 040402 (2007).

[ADS](#) [Article](#) [Google Scholar](#)

18. 18.

Tung, S., Lamporesi, G., Lobser, D., Xia, L. & Cornell, E. A. Observation of the presuperfluid regime in a two-dimensional Bose gas. *Phys. Rev. Lett.* **105**, 230408 (2010).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

19. 19.

Hung, C.-L., Zhang, X., Gemelke, N. & Chin, C. Observation of scale invariance and universality in two-dimensional Bose gases. *Nature* **470**, 236–239 (2011).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

20. 20.

Clark, L. W., Gaj, A., Feng, L. & Chin, C. Collective emission of matter-wave jets from driven Bose–Einstein condensates. *Nature* **551**, 356–359 (2017).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

21. 21.

Herbig, J. et al. Preparation of a pure molecular quantum gas. *Science* **301**, 1510–1513 (2003).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

22. 22.

Mark, M., Meinert, F., Lauber, K. & Nägerl, H.-C. Mott-insulator-aided detection of ultra-narrow Feshbach resonances. *SciPost Phys.* **5**, 055 (2018).

[ADS](#) [Article](#) [Google Scholar](#)

23. 23.

Mark, M. et al. “Stückelberg interferometry” with ultracold molecules. *Phys. Rev. Lett.* **99**, 113201 (2007).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

24. 24.

Ho, T.-L. & Zhou, Q. Obtaining the phase diagram and thermodynamic quantities of bulk systems from the densities of trapped gases. *Nat. Phys.* **6**, 131–134 (2010).

[CAS](#) [Article](#) [Google Scholar](#)

25. 25.

Petrov, D. S. & Shlyapnikov, G. V. Interatomic collisions in a tightly confined Bose gas. *Phys. Rev. A* **64**, 012706 (2001).

[ADS](#) [Article](#) [Google Scholar](#)

26. 26.

Prokof'ev, N. & Svistunov, B. Two-dimensional weakly interacting Bose gas in the fluctuation region. *Phys. Rev. A* **66**, 043608 (2002).

[ADS](#) [Article](#) [Google Scholar](#)

27. 27.

Prokof'ev, N., Ruebenacker, O. & Svistunov, B. Critical point of a weakly interacting two-dimensional Bose gas. *Phys. Rev. Lett.* **87**, 270402 (2001).

[Article](#) [Google Scholar](#)

28. 28.

Hadzibabic, Z. & Dalibard, J. Two-dimensional Bose fluids: an atomic physics perspective. *Riv. Nuovo Cim.* **34**, 389–434 (2011).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

29. 29.

Ferlaino, F. et al. Collisions of ultracold trapped cesium Feshbach molecules. *Laser Phys.* **20**, 23–31 (2010).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

30. 30.

Chin, C. et al. Observation of Feshbach-like resonances in collisions between ultracold molecules. *Phys. Rev. Lett.* **94**, 123201 (2005).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

31. 31.

Idziaszek, Z., Jachymski, K. & Julienne, P. S. Reactive collisions in confined geometries. *New J. Phys.* **17**, 035007 (2015).

[ADS](#) [Article](#) [Google Scholar](#)

32. 32.

Micheli, A. et al. Universal rates for reactive ultracold polar molecules in reduced dimensions. *Phys. Rev. Lett.* **105**, 073202 (2010).

[ADS](#) [Article](#) [Google Scholar](#)

33. 33.

de Miranda, M. H. G. et al. Controlling the quantum stereodynamics of ultracold bimolecular reactions. *Nat. Phys.* **7**, 502–507 (2011).

[Article](#) [Google Scholar](#)

34. 34.

Son, H., Park, J. J., Ketterle, W. & Jamison, A. O. Collisional cooling of ultracold molecules. *Nature* **580**, 197–200 (2020).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

35. 35.

Ho, T.-L. Universal thermodynamics of degenerate quantum gases in the unitarity limit. *Phys. Rev. Lett.* **92**, 090402 (2004).

[ADS](#) [Article](#) [Google Scholar](#)

36. 36.

Eigen, C. et al. Universal prethermal dynamics of Bose gases quenched to unitarity. *Nature* **563**, 221–224 (2018).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

37. 37.

Hung, C.-L. In Situ Probing of Two-Dimensional Quantum Gases. <http://pi.lib.uchicago.edu/1001/cat/bib/8855526> Thesis, Univ. Chicago (2011).

38. 38.

Fano, U. Effects of configuration interaction on intensities and phase shifts. *Phys. Rev.* **124**, 1866–1878 (1961).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

[Download references](#)

Acknowledgements

We thank P. Julienne for discussions and K. Patel for carefully reading the manuscript. This work is supported by National Science Foundation (NSF) grant number PHY-1511696, the Army Research Office Multidisciplinary Research Initiative under grant W911NF-14-1-0003 and the University of Chicago Materials Research Science and Engineering Center, which is funded by the NSF under grant number DMR-1420709.

Author information

Affiliations

1. James Franck Institute, Enrico Fermi Institute and Department of Physics, University of Chicago, Chicago, IL, USA

Zhendong Zhang, Kai-Xuan Yao & Cheng Chin

2. State Key Laboratory of Quantum Optics and Quantum Optics Devices, Institute of Opto-Electronics, Shanxi University, Taiyuan, China

Liangchao Chen

Authors

1. Zhendong Zhang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Liangchao Chen

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Kai-Xuan Yao

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Cheng Chin

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Contributions

L.C. and Z.Z. designed and performed the experiments. Z.Z. analysed the data. All authors contributed to discussions on the experiment and preparation of the manuscript. C.C. supervised the project.

Corresponding author

Correspondence to [Cheng Chin](#).

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 Calibration of magnetic anti-trap potential from the atomic density distribution.

a, Fit of the in situ atomic density profile for determination of the magnetic anti-trap frequencies ω_x and ω_y using equation (2). The top and right panels show line cuts of the 2D atomic density in the x and y directions, crossing at the centre of the anti-trap. We choose the region within the red dashed circle for fit and extraction of the equation of state. **b**, Equation of state of atomic BEC shown in **a**. Each data point represents averaged density within a bin size $\delta\mu/h = 0.25$ Hz and error bars represent one standard deviation. The black solid line is a linear fit to the data, while the black dashed line is an extrapolation of the fit towards the origin. Data values and error bars are estimated as in Fig. 2 from 20 measurements.

Extended Data Fig. 2 Calibration of the optical potential barrier projected by a DMD from the density response measurement of atomic BEC.

a, Images of in situ atomic column density with different central barrier heights determined by different fractions of micromirrors f_{DMD} that are turned on in the DMD. **b**, Example measurements of the proportionality $p(x, y)$ for six pixels at different locations. The solid lines are linear fits to the linear part of the data points, the slope of which gives $p(x, y)$. Data values and error bars are estimated as in Fig. 2 from 9–11 measurements. **c**,

Spatial dependence of the proportionality $p(x, y)$. The upper and right panels are line cuts in the x and y directions crossing the peak value. Data values are determined from the fits and the errorbars represent 95% confidence interval.

Extended Data Fig. 3 Fast equilibration of molecules with atoms during the ramp across the Feshbach resonance.

a, Dynamics of the number of molecules during the magnetic field ramp across the Feshbach resonance at 19.87 G with different ramp speeds of 161 mG ms⁻¹ (red), 80 mG ms⁻¹ (blue) and 54 mG ms⁻¹ (green). **b**, In situ images of molecules during the magnetic field ramp at 80 mG ms⁻¹. Data values and error bars are estimated as in Fig. 2 from 5–7 measurements.

Extended Data Fig. 4 Azimuthally averaged density profiles.

These profiles correspond to the atomic (left) and molecular (right) clouds shown in Fig. 2a. The atomic density profile is flat-topped, whereas the molecular density profile has a dip in the middle.

Extended Data Fig. 5 Dynamics of molecular density profiles in the 2D box trap with magnetic anti-trap potential.

The azimuthally averaged molecular density profiles are shown as a function of the hold time after the formation of molecules. The dips in the middle result from the magnetic anti-trap potential and persist during the first 15 ms after formation of the molecules.

Supplementary information

Peer Review File

Source data

Source Data Fig. 2

Source Data Fig. 3

Source Data Fig. 4

Rights and permissions

Reprints and Permissions

About this article



Check for
updates

Cite this article

Zhang, Z., Chen, L., Yao, KX. *et al.* Transition from an atomic to a molecular Bose–Einstein condensate. *Nature* **592**, 708–711 (2021). <https://doi.org/10.1038/s41586-021-03443-0>

Download citation

- Received: 06 July 2020
- Accepted: 11 March 2021
- Published: 28 April 2021
- Issue Date: 29 April 2021
- DOI: <https://doi.org/10.1038/s41586-021-03443-0>

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03443-0>

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

- Article
- [Published: 28 April 2021](#)

Direct observation of chemical short-range order in a medium-entropy alloy

- [Xuefei Chen](#) ORCID: [orcid.org/0000-0002-6014-0940^{1,2}](https://orcid.org/0000-0002-6014-0940) na1,
- [Qi Wang](#) ORCID: [orcid.org/0000-0003-3418-8809³](https://orcid.org/0000-0003-3418-8809) na1,
- [Zhiying Cheng](#) ORCID: [orcid.org/0000-0002-9880-4124⁴](https://orcid.org/0000-0002-9880-4124) na1,
- [Mingliu Zhu](#) ORCID: [orcid.org/0000-0003-3342-9594¹](https://orcid.org/0000-0003-3342-9594),
- [Hao Zhou⁵](#),
- [Ping Jiang](#) ORCID: [orcid.org/0000-0002-0503-140X¹](https://orcid.org/0000-0002-0503-140X),
- [Lingling Zhou](#) ORCID: [orcid.org/0000-0002-2274-6749^{1,2}](https://orcid.org/0000-0002-2274-6749),
- [Qiqi Xue](#) ORCID: [orcid.org/0000-0003-3440-4866^{1,2}](https://orcid.org/0000-0003-3440-4866),
- [Fuping Yuan](#) ORCID: [orcid.org/0000-0003-3689-6753^{1,2}](https://orcid.org/0000-0003-3689-6753),
- [Jing Zhu](#) ORCID: [orcid.org/0000-0002-2175-9476⁴](https://orcid.org/0000-0002-2175-9476),
- [Xiaolei Wu](#) ORCID: [orcid.org/0000-0001-8270-7621^{1,2}](https://orcid.org/0000-0001-8270-7621) &
- [En Ma](#) ORCID: [orcid.org/0000-0002-7468-4340⁶](https://orcid.org/0000-0002-7468-4340)

[Nature](#) volume 592, pages 712–716(2021) [Cite this article](#)

- 2885 Accesses
- 65 Altmetric
- [Metrics details](#)

Subjects

- [Mechanical properties](#)

- [Metals and alloys](#)

Abstract

Complex concentrated solutions of multiple principal elements are being widely investigated as high- or medium-entropy alloys (HEAs or MEAs)[1,2,3,4,5,6,7,8,9,10,11](#), often assuming that these materials have the high configurational entropy of an ideal solution. However, enthalpic interactions among constituent elements are also expected at normal temperatures, resulting in various degrees of local chemical order[12,13,14,15,16,17,18,19,20,21,22](#). Of the local chemical orders that can develop, chemical short-range order (CSRO) is arguably the most difficult to decipher and firm evidence of CSRO in these materials has been missing thus far[16,22](#). Here we discover that, using an appropriate zone axis, micro/nanobeam diffraction, together with atomic-resolution imaging and chemical mapping via transmission electron microscopy, can explicitly reveal CSRO in a face-centred-cubic VCoNi concentrated solution. Our complementary suite of tools provides concrete information about the degree/extent of CSRO, atomic packing configuration and preferential occupancy of neighbouring lattice planes/sites by chemical species. Modelling of the CSRO order parameters and pair correlations over the nearest atomic shells indicates that the CSRO originates from the nearest-neighbour preference towards unlike (V–Co and V–Ni) pairs and avoidance of V–V pairs. Our findings offer a way of identifying CSRO in concentrated solution alloys. We also use atomic strain mapping to demonstrate the dislocation interactions enhanced by the CSROs, clarifying the effects of these CSROs on plasticity mechanisms and mechanical properties upon deformation.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

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: TEM microstructure of VCoNi MEA.

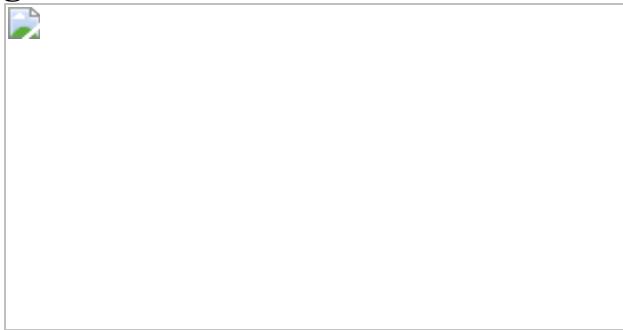


Fig. 2: Evidence of CSRO in the fcc VCoNi.

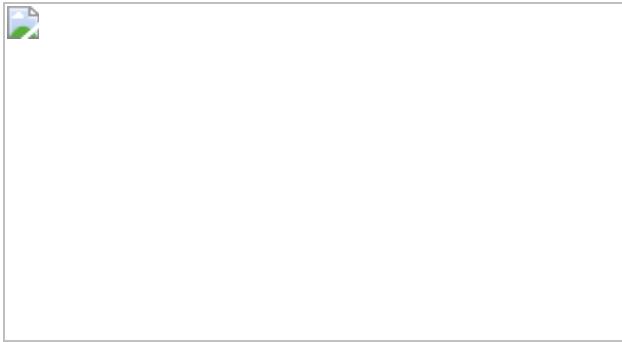


Fig. 3: Chemical mapping indicating element-specific enrichment on alternating atomic planes.

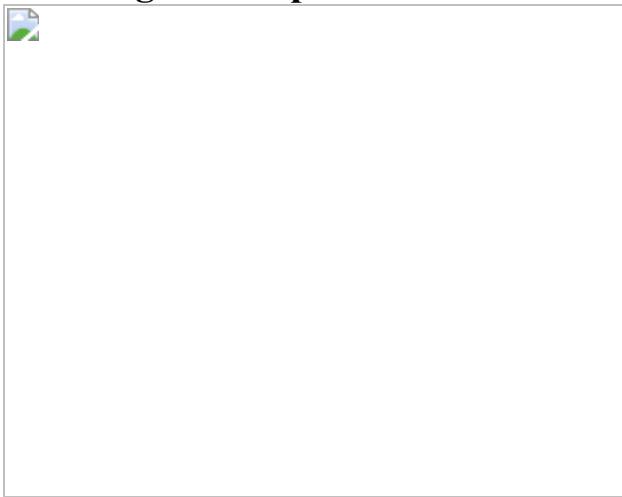


Fig. 4: Interaction between CSRO regions and dislocations.

Data availability

The data generated during and/or analysed during the current study are available from the corresponding author upon reasonable request.

References

1. 1.

Ye, Y. F., Wang, Q., Lu, J., Liu, C. T. & Yang, Y. High-entropy alloy: challenges and prospects. *Mater. Today* **19**, 349–362 (2016).

[CAS](#) [Article](#) [Google Scholar](#)

2. 2.

Miracle, D. B. & Senkov, O. N. A critical review of high entropy alloys and related concepts. *Acta Mater.* **122**, 448–511 (2017).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

3. 3.

Wei, S., He, F. & Tasan, C. C. Metastability in high-entropy alloys: a review. *J. Mater. Res.* **33**, 2924–2937 (2018).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

4. 4.

George, E. P., Raabe, D. & Ritchie, R. O. High-entropy alloys. *Nat. Rev. Mater.* **4**, 515–534 (2019).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

5. 5.

Ma, E. & Wu, X. L. Tailoring heterogeneities in high-entropy alloys to promote strength–ductility synergy. *Nat. Commun.* **10**, 5623 (2019).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

6. 6.

George, E. P., Curtin, W. A. & Tasan, C. C. High entropy alloys: a focused review of mechanical properties and deformation mechanisms. *Acta Mater.* **188**, 435–474 (2020).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

7. 7.

Li, Z., Pradeep, K. G., Deng, Y., Raabe, D. & Tasan, C. C. Metastable high-entropy dual-phase alloys overcome the strength–ductility trade-off. *Nature* **534**, 227–230 (2016).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

8. 8.

Yang, T. et al. Multicomponent intermetallic nanoparticles and superb mechanical behaviors of complex alloys. *Science* **362**, 933–937 (2018).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

9. 9.

Lei, Z. et al. Enhanced strength and ductility in a high-entropy alloy via ordered oxygen complexes. *Nature* **563**, 546 (2018).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

10. 10.

Gludovatz, B. et al. Exceptional damage-tolerance of a medium-entropy alloy CrCoNi at cryogenic temperatures. *Nat. Commun.* **7**, 10602 (2016).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

11. 11.

Liu, C. T. Atomic ordering and structural transformation in the V-Co-Ni ternary alloys. *Metall. Trans.* **4**, 1743–1753 (1973).

[CAS](#) [Article](#) [Google Scholar](#)

12. 12.

Singh, P., Smirnov, A. V. & Johnson, D. D. Atomic short-range order and incipient long-range order in high-entropy alloys. *Phys. Rev. B* **91**, 224204 (2015).

[ADS](#) [Article](#) [Google Scholar](#)

13. 13.

Fernández-Caballero, A. et al. Short-range order in high entropy alloys: theoretical formulation and application to Mo-Nb-Ta-V-W system. *J. Phase Equilibria Diffus.* **38**, 391–403 (2017).

[Article](#) [Google Scholar](#)

14. 14.

Jian, W. R. et al. Effects of lattice distortion and chemical short-range order on the mechanisms of deformation in medium entropy alloy CoCrNi. *Acta Mater.* **199**, 352–369 (2020).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

15. 15.

Ma, Y. et al. Chemical short-range orders and the induced structural transition in high-entropy alloys. *Scr. Mater.* **144**, 64–68 (2018).

[CAS](#) [Article](#) [Google Scholar](#)

16. 16.

Yin, B. L., Yoshida, S. H., Tsuji, N. & Curtin, W. A. Yield strength and misfit volumes of NiCoCr and implications for short-range-order. *Nat. Commun.* **11**, 2507 (2020).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

17. 17.

Zhang, F. X. et al. Local structure and short-range order in a NiCoCr solid solution alloy. *Phys. Rev. Lett.* **118**, 205501 (2017).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

18. 18.

Oh, H. S. et al. Engineering atomic-level complexity in high-entropy and complex concentrated alloys. *Nat. Commun.* **10**, 2090 (2019).

[ADS](#) [Article](#) [Google Scholar](#)

19. 19.

Ding, J., Yu, Q., Asta, M. & Ritchie, R. O. Tunable stacking fault energies by tailoring local chemical order in CrCoNi medium-entropy alloys. *Proc. Natl Acad. Sci. USA* **115**, 8919–8924 (2018).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

20. 20.

Li, Q. J. et al. Strengthening in multi-principal element alloys with local-chemical-order roughened dislocation pathways. *Nat. Commun.* **10**, 3563 (2019).

[ADS](#) [Article](#) [Google Scholar](#)

21. 21.

Ding, Q. et al. Tuning element distribution, structure and properties by composition in high-entropy alloys. *Nature* **574**, 223–227 (2019).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

22. 22.

Zhang, R. et al. Short-range order and its impact on the CrCoNi medium-entropy alloy. *Nature* **581**, 283–287 (2020).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

23. 23.

Sohn, S. S. et al. Ultrastrong medium-entropy single-phase alloys designed via severe lattice distortion. *Adv. Mater.* **31**, 1807142 (2019).

[Article](#) [Google Scholar](#)

24. 24.

Sohn, S. S. et al. High-rate superplasticity in an equiatomic medium-entropy VCoNi alloy enabled through dynamic recrystallization of a duplex microstructure of ordered phase. *Acta Mater.* **194**, 106–117 (2020).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

25. 25.

Williams, D. B. & Carter, C. B. *Transmission Electron Microscopy, A Textbook for Materials Science* (Springer, 2009).

26. 26.

Kuwano, N. et al. In situ TEM observation of long range ordering via short range order in Cu₃Pt. *Bull. Mater. Sci.* **22**, 697–700 (1999).

[CAS](#) [Article](#) [Google Scholar](#)

27. 27.

Van Tendeloo, G. & Amelinckx, S. The origin of diffuse intensity in electron diffraction patterns. *Phase Transit.* **67**, 101–135 (1998).

[Article](#) [Google Scholar](#)

28. 28.

Van Tendeloo, G. & Amelinckx, S. On a simple method to determine the origin of diffuse scattering in electron diffraction patterns. *Scr. Metall.* **20**, 335–339 (1986).

[Article](#) [Google Scholar](#)

29. 29.

Tibshirani, R. Regression shrinkage and selection via the lasso. *J. R. Stat. Soc.* **58**, 267–288 (1996).

[MathSciNet](#) [MATH](#) [Google Scholar](#)

30. 30.

Zunger, A. First-principles statistical mechanics of semiconductor alloys and intermetallic compounds. In *Statics and Dynamics of Alloy Phase Transformations* (eds Turchi, P. E. A. & Gonis, A.) 361–419 (Springer, 1994).

31. 31.

Tamm, A., Aabloo, A., Klintenberg, M., Stocks, M. & Caro, A. Atomic-scale properties of Ni-based FCC ternary and quaternary alloys. *Acta Mater.* **99**, 307–312 (2015).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

32. 32.

Zunger, A., Wei, S., Ferreira, L. G. & Bernard, J. E. Special quasirandom structures. *Phys. Rev. Lett.* **65**, 353–356 (1990).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

33. 33.

Hýtch, M. Quantitative measurement of displacement and strain fields from HREM micrographs. *Ultramicroscopy* **74**, 131–146 (1998).

[Article](#) [Google Scholar](#)

34. 34.

Warren, B. E. *X-Ray Diffraction* (Dover Publications, 1990).

[Download references](#)

Acknowledgements

X.W., F.Y. and P.J. were supported by the National Key Research and Development Program of the Ministry of Science and Technology of China (grant numbers 2019YFA0209900 and 2017YFA0204402), the Basic Science Center Program (grant number 11988102), the Natural Science Foundation of China (grant numbers 11972350 and 11890680), and the Chinese Academy of Sciences (grant number XDB22040503). Z.C. and J.Z. were supported by the Basic Science Center Program (grant number 51788104), the National Key Research and Development Program of the Ministry of Science and Technology of China (2016YFB0700402). E.M. thanks Xi'an Jiaotong University for supporting his work at the Center for Alloy Innovation and Design.

Author information

Author notes

1. These authors contributed equally: Xuefei Chen, Qi Wang, Zhiying Cheng

Affiliations

1. State Key Laboratory of Nonlinear Mechanics, Institute of Mechanics, Chinese Academy of Sciences, Beijing, China

Xuefei Chen, Mingliu Zhu, Ping Jiang, Lingling Zhou, Qiqi Xue, Fuping Yuan & Xiaolei Wu

2. School of Engineering Science, University of Chinese Academy of Sciences, Beijing, China

Xuefei Chen, Lingling Zhou, Qiqi Xue, Fuping Yuan & Xiaolei Wu

3. Science and Technology on Surface Physics and Chemistry Laboratory, Jiangyou, China

Qi Wang

4. National Center of Electron Microscopy in Beijing, School of Materials Science and Engineering, Tsinghua University, Beijing, China

Zhiying Cheng & Jing Zhu

5. Nano and Heterogeneous Materials Center, School of Materials Science and Engineering, Nanjing University of Science and Technology, Nanjing, China

Hao Zhou

6. Center for Alloy Innovation and Design (CAID), State Key Laboratory for Mechanical Behavior of Materials, Xi'an Jiaotong University, Xi'an, China

En Ma

Authors

1. Xuefei Chen

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Qi Wang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Zhiying Cheng

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Mingliu Zhu

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Hao Zhou

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Ping Jiang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Lingling Zhou

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Qiqi Xue

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. Fuping Yuan

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. Jing Zhu

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

11. Xiaolei Wu

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

12. En Ma

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Contributions

X.W. and E.M. conceived the ideas and supervised the project together with J.Z. X.C. and H.Z. performed the TEM, FFT and geometric phase analysis work. Q.W. conducted the pair correlation analysis, DFT and Monte Carlo simulations. Z.C. designed and carried out the STEM experiments. P.J. prepared the materials, samples and heat treatments. P.J., L.Z., Q.X., M.Z. and F.Y. conducted tensile testing, electron backscatter diffraction observations, and mechanical behaviour analysis. All authors participated in the discussions. X.W. and E.M. wrote the paper.

Corresponding authors

Correspondence to [Jing Zhu](#) or [Xiaolei Wu](#) or [En Ma](#).

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 Evidence of CSRO in the fcc phase of VCoNi MEA before tensile deformation.

All results are similar to those after tensile straining shown in Fig. 2. **a-1**, Micro-area EDP with the [112] zone axis. Arrows point to the arrays of superlattice reflections at $\langle\langle 1\bar{2}\bar{2}\bar{3}11\rangle\rangle$ positions. **a-2**, Energy-filtered dark-field TEM image taken using extra reflections. Inset, a close-up view of the area in the dashed-line enclosed square, highlighting an area with CSROs. **a-3**, Statistics showing the size distribution of CSROs, observed in the dark-field TEM images and inverse FFT images. **b-1**, FFT pattern of the fcc phase with the [112] zone axis. The yellow circle highlights a diffuse reflection at $\langle\langle 1\bar{2}\bar{2}\bar{3}11\rangle\rangle$ positions. **b-2**, Inverse FFT image showing the CSROs (circled) that are superimposed on the fcc lattice image. **b-3**, Maps of V, V-Co and V-Ni, respectively, showing two CSROs in two arrays by EDS mapping from the HAADF image with the [112] zone axis in an aberration-corrected TEM. All dashed lines mark the $\langle\langle \bar{3}11\rangle\rangle$ planes intersecting the $\langle\langle 111\rangle\rangle$ plane in plan view: yellow: V-enriched, blue: Co-/Ni-enriched. All scale bars are 0.5 nm.

Extended Data Fig. 2 EDS mapping of the VCoNi alloy with the [112] zone axis.

These are additional maps not included in the main text, showing the distribution of Co, Ni and Co+Ni. The dashed lines in each panel mark (Co,Ni)-enriched {311} planes (blue dashed in Fig. 3), which alternate with V-enriched ones.

Extended Data Fig. 3 Illustration of projecting a (111) plane along [112] and [110] beam directions.

V–Co(Ni) as the nearest neighbour is assumed in this idealized model to be the prevailing CSRO. The numbers 1, 3 and 4 indicate the 1st, 3rd and 4th nearest neighbours, respectively, around a centre atom C. **a**, Plan view of a close-packed (111) plane, projected along the [112] beam direction. The distance between the nearest points in the [110] direction is r^* (compare with Fig. 3d). **b**, Plan view of the same (111) plane, observed along the [110] beam direction. Unlike for the [112] beam direction, when projected along the [110] beam direction (for example, horizontal dashed lines) the centre atom will be directly superimposed onto two unlike 1st neighbours. This mixed column, when compared with the case in **a** (no overlapping of unlike species in the column) blurs the difference (and hence the contrast) from the neighbouring columns. [112] is therefore the preferred beam direction to see the CSRO of interest.

Extended Data Fig. 4 Schematic of element occupancy that exemplifies the CSRO taking fcc lattice sites.

a, Two-dimensional lattice structure of the CSRO, deduced from experimental evidence (the alternating pattern of {311} planes in Fig. 2h and of atomic columns in the {111} plane in Fig. 3d). Note that the red (blue) spheres are meant to represent V (Co,Ni)-enriched atom positions, respectively (i.e., red is not yet V only, but still contains some Co and Ni). The boxed region shows the minimum-sized configuration of the CSRO. **b**, The 3D configuration of the CSRO is based on the motif (left) deduced from observations under both the [112] and [110] z.a. (to be explained in a

future publication), and embedded in the fcc matrix (right). Grey spheres indicate random atoms (V,Co,Ni) without chemical order in the fcc lattice. **c**, Simulated diffraction pattern for the sub-nanometre CSRO configuration embedded in the fcc lattice in **b**, with the [112] zone axis, showing the extra reflections at the $\frac{1}{2}\bar{3}11$ positions (purple diffuse disks).

Extended Data Fig. 5 Strain-field analysis around the CSRO before tensile deformation.

a, Geometric phase analysis strain mapping, superimposed on the lattice image taken with the [110] zone axis. The yellow striped areas with positive strain correspond to the CSROs; two red atoms (V_1 and V_2) are displayed to represent the V-enriched columns. **b**, Strain distribution between the two V atoms in **a**. We note that the spacing between the strain peaks of neighbouring V atoms is 0.28 nm. This figure further illustrates the elastic strains observed in Fig. 4. The atomic radii of V, Co, and Ni are 1.35 Å, 1.26 Å and 1.24 Å, respectively. The larger V atoms and the smaller Co/Ni atoms induce tensile and compressive strain, respectively, in the normal direction of close-packed $\{1\bar{1}1\}$ planes. The yellow striped bands correspond to the CSROs, with tensile (positive) strain induced by the V-enriched columns in the $(1\bar{1}1)$ plane. Two V atoms (red) are placed in the figure to mark such columns. The strain distribution between these neighbouring columns (V_1 and V_2) is shown in **b**. The spacing between the two strain peaks is 0.28 nm (the average value is 0.3 nm), quite close to the measured spacing between two atomic columns (0.26 nm) based on TEM lattice image. This corroborates that the yellow regions of positive strain are due to the V-enriched columns associated with the CSRO. Nearby regions (blue) experience compressive strain (negative).

Supplementary information

Supplementary Information

This file contains Supplementary Notes 1-3 and Supplementary Figs 1-4, which present details of i) definition of pair correlation coefficient, ii) use

of model structures to demonstrate the useful features of pair correlation coefficients, and iii) cluster-expansion-guided Monte Carlo simulations.

[Peer Review File](#)

Rights and permissions

[Reprints and Permissions](#)

About this article



Check for
updates

Cite this article

Chen, X., Wang, Q., Cheng, Z. *et al.* Direct observation of chemical short-range order in a medium-entropy alloy. *Nature* **592**, 712–716 (2021).
<https://doi.org/10.1038/s41586-021-03428-z>

[Download citation](#)

- Received: 27 September 2020
- Accepted: 08 March 2021
- Published: 28 April 2021
- Issue Date: 29 April 2021
- DOI: <https://doi.org/10.1038/s41586-021-03428-z>

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03428-z>

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

- Article
- [Published: 28 April 2021](#)

Strongly reducing magnesium(0) complexes

- [B. Rösch¹](#),
- [T. X. Gentner¹](#),
- [J. Eyselein¹](#),
- [J. Langer](#) [ORCID: orcid.org/0000-0002-2895-4979¹](#),
- [H. Elsen¹](#) &
- [S. Harder](#) [ORCID: orcid.org/0000-0002-3997-1440¹](#)

[Nature](#) volume **592**, pages 717–721(2021) [Cite this article](#)

- 2792 Accesses
- 98 Altmetric
- [Metrics details](#)

Subjects

- [Chemical bonding](#)
- [Ligands](#)
- [Organometallic chemistry](#).
- [Chemical synthesis](#)

Abstract

A complex of a metal in its zero oxidation state can be considered a stabilized, but highly reactive, form of a single metal atom. Such complexes

are common for the more noble transition metals. Although rare examples are known for electronegative late-main-group *p*-block metals or semimetals^{1,2,3,4,5,6}, it is a challenge to isolate early-main-group *s*-block metals in their zero oxidation state^{7,8,9,10,11}. This is directly related to their very low electronegativity and strong tendency to oxidize. Here we present examples of zero-oxidation-state magnesium (that is, magnesium(0)) complexes that are stabilized by superbметally, monoanionic, β -diketiminate ligands. Whereas the reactivity of an organomagnesium compound is typically defined by the nucleophilicity of its organic groups and the electrophilicity of Mg²⁺ cations, the Mg⁰ complexes reported here feature electron-rich Mg centres that are nucleophilic and strongly reducing. The latter property is exemplified by the ability to reduce Na⁺ to Na⁰. We also present a complex with a linear Mg₃ core that formally could be described as a Mg^I–Mg⁰–Mg^I unit. Such multinuclear mixed-valence Mg_{*n*} clusters are discussed as fleeting intermediates during the early stages of Grignard reagent formation. Their remarkably strong reducing power implies a rich reactivity and application as specialized reducing agents.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

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: Low-valent complexes of main group metals and semimetals.

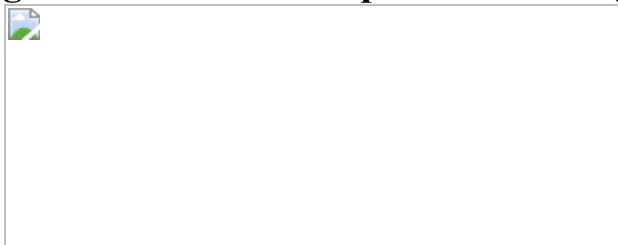


Fig. 2: Synthesis and reactivity of β -diketiminate Mg^0 complexes.

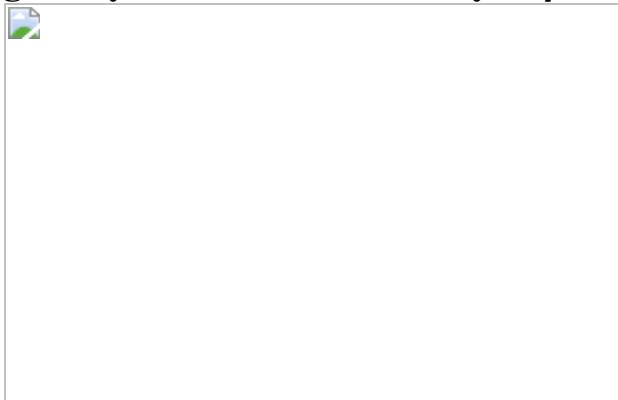


Fig. 3: Molecular structures and Laplacian distribution of the electron density for the Mg^0 complexes.

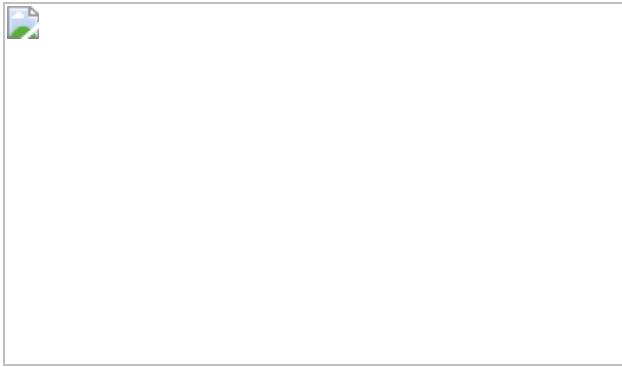
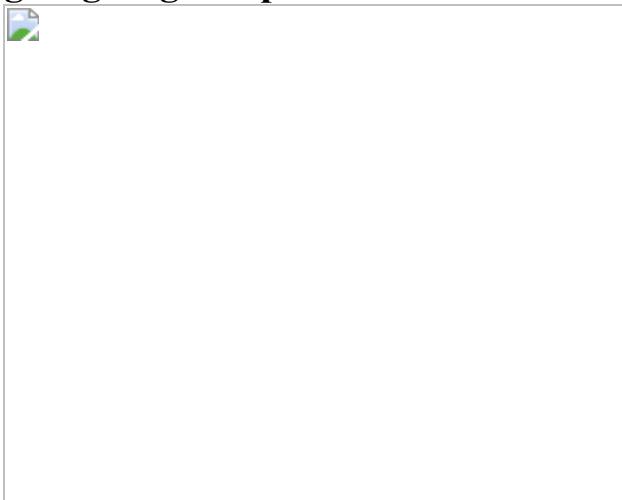


Fig. 4: Calculated energy profiles for stepwise loss of Mg⁰ in trinuclear Mg–Mg–Mg complexes.



Data availability

X-ray data are available free of charge from the Cambridge Crystallographic Data Centre under references CCDC 2045616 (**1**), 2045617 (**2**), 2045618 (**3**), 2045619 (**4**) and 2045620 (**5**). Spectroscopic data that support the findings of this study as well as complementary crystallographic and computational details are included in Supplementary Information. Raw data are available from the corresponding author on reasonable request.

References

1. 1.

Wang, Y. et al. A stable silicon(0) compound with a Si=Si double bond. *Science* **321**, 1069–1071 (2008).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

2. 2.

Sidiropoulos, A., Jones, C., Stasch, A., Klein, S. & Frenking, G. N-Heterocyclic carbene stabilized digermanium(0). *Angew. Chem. Int. Edn* **48**, 9701–9704 (2009).

[CAS](#) [Google Scholar](#)

3. 3.

Braunschweig, H. et al. Ambient-temperature isolation of a compound with a boron-boron triple bond. *Science* **336**, 1420–1422 (2012).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

4. 4.

Jones, C., Sidiropoulos, A., Holzmann, N., Frenking, G. & Stasch, A. An N-heterocyclic carbene adduct of diatomic tin, :Sn=Sn:. *Chem. Commun.* **48**, 9855–9857 (2012).

[CAS](#) [Google Scholar](#)

5. 5.

Mondal, K. C. et al. A stable singlet biradicaloid siladicarbene: (L:)₂Si. *Angew. Chem. Int. Edn* **52**, 2963–2967 (2013).

[CAS](#) [Google Scholar](#)

6. 6.

Li, Y. et al. Acyclic germylones: congeners of allenes with a central germanium atom. *J. Am. Chem. Soc.* **135**, 12422–12428 (2013).

[CAS](#) [PubMed](#) [Google Scholar](#)

7. 7.

Glaunsinger, W. S. et al. Structure and molecular motions in alkaline earth hexammines. *Nature* **271**, 414–417 (1978).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

8. 8.

Wu, X. et al. Observation of alkaline earth complexes $M(CO)_8$ ($M = Ca, Sr, or Ba$) that mimic transition metals. *Science* **361**, 912–916 (2018).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

9. 9.

Wang, Q. et al. Transition-metal chemistry of alkaline-earth elements: the trisbenzene complexes $M(Bz)_3$ ($M = Sr, Ba$). *Angew. Chem. Int. Edn* **58**, 17365–17374 (2019).

[CAS](#) [Google Scholar](#)

10. 10.

Arrowsmith, M. et al. Neutral zero-valent s-block complexes with strong multiple bonding. *Nat. Chem.* **8**, 890–894 (2016).

[CAS](#) [Google Scholar](#)

11. 11.

Couchman, S. A., Holzmann, N., Frenking, G., Wilson, D. J. D. & Dutton, J. L. Beryllium chemistry the safe way: a theoretical evaluation of low oxidation state beryllium compounds. *Dalton Trans.* **42**, 11375–11384 (2013).

[CAS](#) [PubMed](#) [Google Scholar](#)

12. 12.

Green, S. P., Jones, C. & Stasch, A. Stable magnesium(I) compounds with Mg–Mg bonds. *Science* **318**, 1754–1757 (2007).

[ADS](#) [CAS](#) [Google Scholar](#)

13. 13.

Jones, C. Dimeric magnesium(I) β -diketiminates: a new class of quasi-universal reducing agent. *Nat. Rev. Chem.* **1**, 0059 (2017).

[CAS](#) [Google Scholar](#)

14. 14.

Jones, C. Open questions in low oxidation state group 2 chemistry. *Commun. Chem.* **3**, 159 (2020).

[Google Scholar](#)

15. 15.

Gentner, T. X. et al. Low valent magnesium chemistry with a super bulky β -diketiminate ligand. *Angew. Chem. Int. Edn* **58**, 607–611 (2019).

[CAS](#) [Google Scholar](#)

16. 16.

Jones, D. D. L., Douair, I., Maron, L. & Jones, C. Photochemically activated dimagnesium(I) compounds: reagents for the reduction and selective C–H bond activation of inert arenes. *Angew. Chem. Int. Edn* **60**, 7087–7092 (2021).

17. 17.

Rösch, B. et al. Mg-Mg bond polarization induced by a superbulky β -diketiminate ligand. *Chem. Commun.* **56**, 11402–11405 (2020).

[Google Scholar](#)

18. 18.

Hicks, J., Juckel, M., Paparo, A., Dange, D. & Jones, C. Multigram syntheses of magnesium(I) compounds using alkali metal halide supported alkali metals as dispersible reducing agents. *Organometallics* **37**, 4810–4813 (2018).

[CAS](#) [Google Scholar](#)

19. 19.

Cui, C. et al. Synthesis and structure of a monomeric aluminum(I) compound [$\{\text{HC}(\text{CMeNAr})_2\}\text{Al}$] ($\text{Ar}=2,6\text{-}i\text{Pr}_2\text{C}_6\text{H}_3$): a stable aluminum analogue of a carbene. *Angew. Chem. Int. Edn* **39**, 4274–4276 (2000).

[CAS](#) [Google Scholar](#)

20. 20.

Hicks, J., Vasko, P., Goicoechea, J. M. & Aldridge, S. Synthesis, structure and reaction chemistry of a nucleophilic alumanyl anion. *Nature* **557**, 92–95 (2018); correction **560**, E24 (2018).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

21. 21.

Dye, J. L., Ceraso, J. M., Lok Tak, M., Barnett, B. L. & Tehan, F. J. Crystalline salt of the sodium anion (Na^-). *J. Am. Chem. Soc.* **96**, 608–609 (1974).

[CAS](#) [Google Scholar](#)

22. 22.

Camp, C. & Arnold, J. On the non-innocence of “NacNacs”: ligand-based reactivity in β -diketiminate supported coordination compounds. *Dalton Trans.* **45**, 14462–14498 (2016).

[CAS](#) [PubMed](#) [Google Scholar](#)

23. 23.

Slater, J. C. Atomic radii in crystals. *J. Chem. Phys.* **41**, 3199–3204 (1964).

[ADS](#) [CAS](#) [Google Scholar](#)

24. 24.

Lambert, C. & von Ragué Schleyer, P. Are polar organometallic compounds “carbanions”? The gegenion effect on structure and energies of alkali-metal compounds. *Angew. Chem. Int. Edn Engl.* **33**, 1129–1140 (1994).

[Google Scholar](#)

25. 25.

Cao, W. L., Gatti, C., MacDougall, P. J. & Bader, R. F. W. On the presence of non-nuclear attractors in the charge distributions of Li and Na clusters. *Chem. Phys. Lett.* **141**, 380–385 (1987).

[ADS](#) [CAS](#) [Google Scholar](#)

26. 26.

Gentner, T. X. & Mulvey, R. E. Alkali metal mediation: diversity of applications in main group organometallic chemistry. *Angew. Chem. Int. Edn* **59**, 2–18 (2020).

[Google Scholar](#)

27. 27.

Arrowsmith, M. et al. Mononuclear three-coordinate magnesium complexes of a highly sterically encumbered β -diketiminate ligand. *Inorg. Chem.* **53**, 10543–10552 (2014).

[CAS](#) [PubMed](#) [Google Scholar](#)

28. 28.

Bakewell, C., White, A. J. P. & Crimmin, M. R. Addition of carbon-fluorine bonds to a Mg(I)-Mg(I) bond: an equivalent of Grignard formation in solution. *J. Am. Chem. Soc.* **138**, 12763–12766 (2016).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

29. 29.

Hicks, J., Underhill, E. J., Kefalidis, C. E., Maron, L. & Jones, C. A mixed-valence tri-zinc complex, [LZnZnZnL] (L=bulky amide), bearing a linear chain of two-coordinate zinc atoms. *Angew. Chem. Int. Edn* **54**, 10000–10004 (2015).

[CAS](#) [Google Scholar](#)

30. 30.

Bakewell, C., Ward, B. J., White, A. J. P. & Crimmin, M. R. A combined experimental and computational study on the reaction of fluoroarenes with Mg-Mg, Mg-Zn, Mg-Al and Al-Zn bonds. *Chem. Sci.* **9**, 2348–2356 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

31. 31.

Platts, J. A., Overgaard, J., Jones, C., Iversen, B. B. & Stasch, A. First experimental characterization of a non-nuclear attractor in a dimeric magnesium(I) compound. *J. Phys. Chem. A* **115**, 194–200 (2011).

[CAS](#) [PubMed](#) [Google Scholar](#)

32. 32.

Garst, G. F. & Ungvary, F. in *Grignard Reagents: New Developments* (ed. Richey, H.) 185–276 (Wiley, 2000).

33. 33.

Tjurina, L. A. et al. Synthesis of cluster alkyl and aryl Grignard reagents in solution. *Organometallics* **23**, 1349–1351 (2004).

[CAS](#) [Google Scholar](#)

34. 34.

Imizu, Y. & Klabunde, K. J. Metal cluster vs. atom reactivities: magnesium cluster Grignard reagents. *Inorg. Chem.* **23**, 3602–3605 (1984).

[CAS](#) [Google Scholar](#)

35. 35.

Köhn, A., Weigend, F. & Ahlrichs, R. Theoretical study on clusters of magnesium. *Phys. Chem. Chem. Phys.* **3**, 711–719 (2001).

[Google Scholar](#)

36. 36.

Eriksson, L. A. Accurate density functional theory study of cationic magnesium clusters and Mg⁺-rare gas interactions. *J. Chem. Phys.* **103**, 1050–1056 (1995).

[ADS](#) [CAS](#) [Google Scholar](#)

37. 37.

Kruczyński, T., Henke, F., Neumaier, M., Bowen, K. H. & Schnöckel, H. Many Mg-Mg bonds form the core of the $Mg_{16}Cp^*_8Br_4K$ cluster anion: the key to a reassessment of the Grignard reagent (GR) formation process? *Chem. Sci.* **7**, 1543–1547 (2016).

[PubMed](#) [Google Scholar](#)

38. 38.

Velazquez, A., Fernández, I., Frenking, G. & Merino, G. Multimetallocenes. A theoretical study. *Organometallics* **26**, 4731–4736 (2007).

[CAS](#) [Google Scholar](#)

39. 39.

Chase, M. W. Jr *NIST-JANAF Thermochemical Tables* Part 2, 4th edn (Monograph No. 9, J. Phys. Chem. Ref. Data, American Institution of Physics, 1998).

40. 40.

Clegg, W. et al. Synthesis and structures of [(trimethylsilyl)methyl]sodium and -potassium with bi- and tridentate N-donor ligands. *Eur. J. Inorg. Chem.* 721–726 (2011).

41. 41.

Rigaku Oxford Diffraction. CrysAlisPro Software system, version 1.171.40.67a (Rigaku Corporation, 2018).

42. 42.

Dolomanov, O. V., Bourhis, L. J., Gildea, R. J., Howard, J. A. K. & Puschmann, H. OLEX2: a complete structure solution, refinement and analysis program. *J. Appl. Crystallogr.* **42**, 339–341 (2009).

[CAS](#) [Google Scholar](#)

43. 43.

Sheldrick, G. M. SHELXT — Integrated space-group and crystal-structure determination. *Acta Crystallogr. A* **71**, 3–8 (2015).

[MATH](#) [Google Scholar](#)

44. 44.

Sheldrick, G. M. Crystal structure refinement with SHELXL. *Acta Crystallogr. C* **71**, 3–8 (2015).

[MATH](#) [Google Scholar](#)

45. 45.

Frisch, M. J. et al. Gaussian 16 Rev. A.03 (Gaussian, Inc., 2016).

46. 46.

Becke, A. D. A new mixing of Hartree–Fock and local density-functional theories. *J. Chem. Phys.* **98**, 1372–1377 (1993).

[ADS](#) [CAS](#) [Google Scholar](#)

47. 47.

Perdew, J. P. *Electronic Structure of Solids* (Akademie, 1991).

48. 48.

Weigend, F. & Ahlrichs, R. Balanced basis sets of split valence, triple zeta valence and quadruple zeta valence quality for H to Rn: Design and assessment of accuracy. *Phys. Chem. Chem. Phys.* **7**, 3297–3305 (2005).

[CAS](#) [PubMed](#) [Google Scholar](#)

49. 49.

Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* **132**, 154104 (2010).

[ADS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

50. 50.

Reed, A. E., Weinstock, R. B. & Weinhold, F. Natural population analysis. *J. Chem. Phys.* **83**, 735–746 (1985).

[ADS](#) [CAS](#) [Google Scholar](#)

51. 51.

Bader, R. F. W. A quantum theory of molecular structure and its applications. *Chem. Rev.* **91**, 893–928 (1991).

[CAS](#) [Google Scholar](#)

52. 52.

Keith, T. A. AIMAll Version 17.01.25 (TK Gristmill Software, 2017).

[Download references](#)

Acknowledgements

We thank A. Roth (FAU) and the Kolbe Microanalytical Laboratory (Mülheim/Ruhr) for elemental analysis. This work was generously supported by the University of Erlangen-Nürnberg.

Author information

Affiliations

1. Department of Chemistry and Pharmacy, Friedrich-Alexander University Erlangen-Nürnberg, Erlangen, Germany
B. Rösch, T. X. Gentner, J. Eyselein, J. Langer, H. Elsen & S. Harder

Authors

1. B. Rösch
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. T. X. Gentner
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. J. Eyselein
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. J. Langer
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. H. Elsen
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. S. Harder
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Contributions

B.R.: conceptualization, investigation, validation, formal analysis (that is, analyses of raw data from spectroscopy, X-ray diffraction or computational methods), writing of original draft and visualization. T.X.G.: investigation, validation and formal analysis. J.E.: formal analysis. J.L.: formal analysis. H.E.: formal analysis. S.H.: conceptualization, writing of original draft, review and editing, visualization, supervision and project administration.

Corresponding author

Correspondence to [S. Harder](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Jason Dutton 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.

Supplementary information

[Supplementary Information](#)

This file contains: Spectroscopic characterization (including Supplementary Figs. 1 to 41); Crystal structure determination (including Supplementary Figs. 42 to 46 and Supplementary Tables 1 to 5); Comparison of chemical shifts and bond distances for (BDI*)Mg complexes (including Supplementary Table 6); Computational details (including Supplementary

Figs. 47 to 57 and Supplementary Tables 7-15); Arguments in favor for the correct metal assignment in $\{[(\text{BDI}^*)\text{Mg}^-]\text{[Na}^+\}\}_2$ (**1**); and Supplementary References.

Peer Review File

Rights and permissions

Reprints and Permissions

About this article



Cite this article

Rösch, B., Gentner, T.X., Eyselein, J. *et al.* Strongly reducing magnesium(0) complexes. *Nature* **592**, 717–721 (2021).
<https://doi.org/10.1038/s41586-021-03401-w>

Download citation

- Received: 04 December 2020
- Accepted: 26 February 2021
- Published: 28 April 2021
- Issue Date: 29 April 2021
- DOI: <https://doi.org/10.1038/s41586-021-03401-w>

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

Associated Content

[Highly reactive form of magnesium stabilized by bulky ligands](#)

- Cameron Jones

Nature News & Views 28 Apr 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03401-w>

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

- Article
- [Published: 28 April 2021](#)

Direct assessment of the acidity of individual surface hydroxyls

- [Margareta Wagner ORCID: orcid.org/0000-0001-9414-1696^{1,2}](#),
- [Bernd Meyer ORCID: orcid.org/0000-0002-3481-8009^{3,4}](#),
- [Martin Setvin^{1,5}](#),
- [Michael Schmid¹](#) &
- [Ulrike Diebold ORCID: orcid.org/0000-0003-0319-5256¹](#)

[Nature](#) volume **592**, pages 722–725(2021) [Cite this article](#)

- 1156 Accesses
- 65 Altmetric
- [Metrics details](#)

Subjects

- [Chemical physics](#)
- [Scanning probe microscopy](#)

Abstract

The state of deprotonation/protonation of surfaces has far-ranging implications in chemistry, from acid–base catalysis¹ and the electrocatalytic and photocatalytic splitting of water², to the behaviour of minerals³ and biochemistry⁴. An entity’s acidity is described by its proton affinity and its acid dissociation constant pK_a (the negative logarithm of the equilibrium

constant of the proton transfer reaction in solution). The acidity of individual sites is difficult to assess for solids, compared with molecules. For mineral surfaces, the acidity is estimated by semi-empirical concepts, such as bond-order valence sums⁵, and increasingly modelled with first-principles molecular dynamics simulations^{6,7}. At present, such predictions cannot be tested—experimental measures, such as the point of zero charge⁸, integrate over the whole surface or, in some cases, individual crystal facets⁹. Here we assess the acidity of individual hydroxyl groups on In₂O₃(111)—a model oxide with four different types of surface oxygen atom. We probe the strength of their hydrogen bonds with the tip of a non-contact atomic force microscope and find quantitative agreement with density functional theory calculations. By relating the results to known proton affinities of gas-phase molecules, we determine the proton affinity of the different surface sites of In₂O₃ with atomic precision. Measurements on hydroxylated titanium dioxide and zirconium oxide extend our method to other oxides.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

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 clean and hydroxylated $\text{In}_2\text{O}_3(111)$ surface.

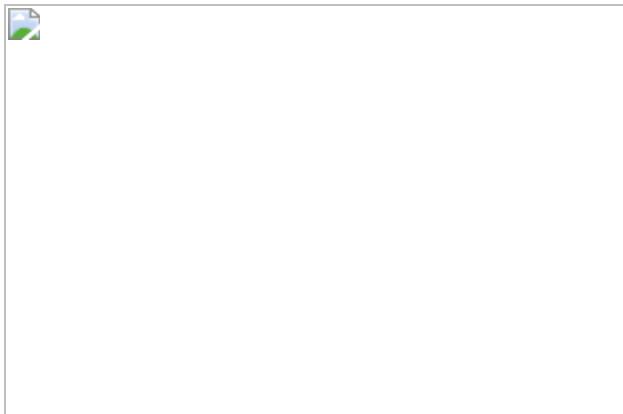


Fig. 2: Probing individual surface hydroxyls with the AFM tip.

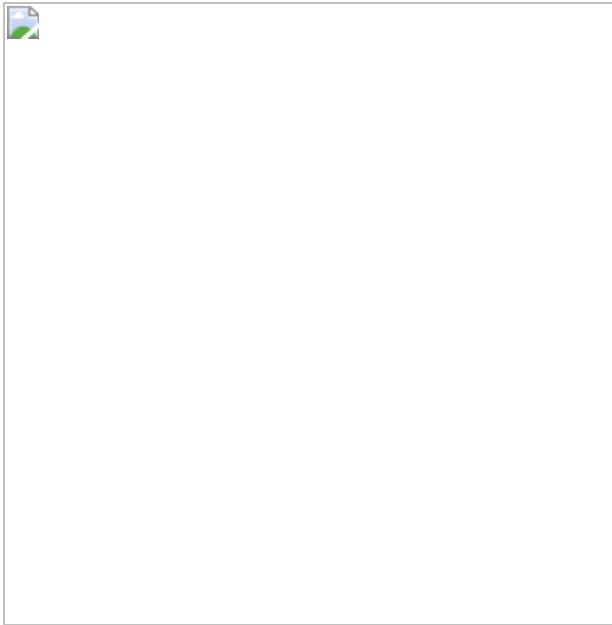


Fig. 3: Evaluation of theoretical force curves of hydroxyls on $\text{In}_2\text{O}_3(111)$ and $\text{TiO}_2(110)$.

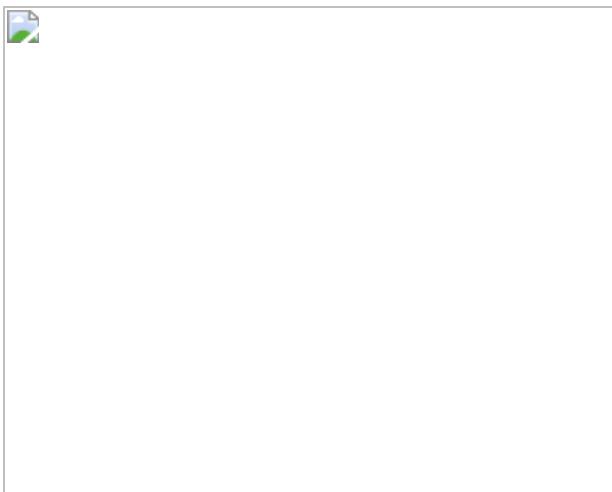
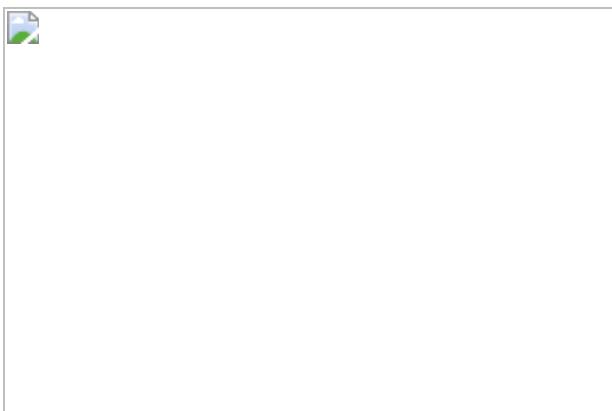


Fig. 4: Relating AFM-measured force minima to PA using a calibration line.



Data availability

The datasets generated and analysed during the current study are available from the corresponding author on reasonable request.

References

1. 1.

Kramer, G. J., van Santen, R. A., Emeis, C. A. & Nowak, A. K. Understanding the acid behaviour of zeolites from theory and experiment. *Nature* **363**, 529–531 (1993).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

2. 2.

She, Z. W. et al. Combining theory and experiment in electrocatalysis: insights into materials design. *Science* **355**, eaad4998 (2017).

[Article](#) [Google Scholar](#)

3. 3.

Brown, G. E. Jr & Calas, G. Mineral–aqueous solution interfaces and their impact on the environment. *Geochem. Perspect.* **1**, 483–742 (2012).

[Article](#) [Google Scholar](#)

4. 4.

Silverman, R. B. *Organic Chemistry of Enzyme-Catalyzed Reactions* (Academic Press, 2002).

5. 5.

Hiemstra, T., Venema, P. & Riemsdijk, W. Intrinsic proton affinity of reactive surface groups of metal (hydr)oxides: the bond valence principle. *J. Colloid Interface Sci.* **184**, 680–692 (1996).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

6. 6.

Cheng, J. & Sprik, M. Acidity of the aqueous rutile TiO₂(110) surface from density functional theory based molecular dynamics. *J. Chem. Theory Comput.* **6**, 880–889 (2010).

[CAS](#) [Article](#) [Google Scholar](#)

7. 7.

Gittus, O. R. von Rudorff, G. F., Rosso, K. M. & Blumberger, J. Acidity constants of the hematite–liquid water interface from ab initio molecular dynamics. *J. Phys. Chem. Lett.* **9**, 5574–5582 (2018).

[CAS](#) [Article](#) [Google Scholar](#)

8. 8.

Kosmulski, M. Compilation of PZC and IEP of sparingly soluble metal oxides and hydroxides from literature. *Adv. Colloid Interface Sci.* **152**, 14–25 (2009).

[CAS](#) [Article](#) [Google Scholar](#)

9. 9.

Bullard, J. W. & Cima, M. J. Orientation dependence of the isoelectric point of TiO₂ (rutile) surfaces. *Langmuir* **22**, 10264–10271 (2006).

[CAS](#) [Article](#) [Google Scholar](#)

10. 10.

Giessibl, F. J. The qPlus sensor, a powerful core for the atomic force microscope. *Rev. Sci. Instrum.* **90**, 011101 (2019).

[ADS](#) [Article](#) [Google Scholar](#)

11. 11.

Gross, L., Mohn, F., Moll, N., Liljeroth, P. & Meyer, G. The chemical structure of a molecule resolved by atomic force microscopy. *Science* **325**, 1110–1114 (2009).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

12. 12.

Setvín, M. et al. Polarity compensation mechanisms on the perovskite surface $\text{KTaO}_3(001)$. *Science* **359**, 572–575 (2018).

[ADS](#) [Article](#) [Google Scholar](#)

13. 13.

Peng, J. et al. The effect of hydration number on the interfacial transport of sodium ions. *Nature* **557**, 701–705 (2018); correction **563**, E18 (2018).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

14. 14.

Lantz, M. A. et al. Quantitative measurement of short-range chemical bonding forces. *Science* **291**, 2580–2583 (2001).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

15. 15.

Sugimoto, Y. et al. Chemical identification of individual surface atoms by atomic force microscopy. *Nature* **446**, 64–67 (2007).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

16. 16.

Gross, L. et al. Measuring the charge state of an adatom with noncontact atomic force microscopy. *Science* **324**, 1428–1431 (2009).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

17. 17.

Onoda, J., Ondráček, M., Jelínek, P. & Sugimoto, Y. Electronegativity determination of individual surface atoms by atomic force microscopy. *Nat. Commun.* **8**, 15155 (2017).

[ADS](#) [Article](#) [Google Scholar](#)

18. 18.

Hunter, E. P. L. & Lias, S. G. Evaluated gas phase basicities and proton affinities of molecules: an update. *J. Phys. Chem. Ref. Data* **27**, 413–656 (1998).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

19. 19.

Gorte, R. J. & White, D. Interactions of chemical species with acid sites in zeolites. *Top. Catal.* **4**, 57–69 (1997).

[CAS](#) [Article](#) [Google Scholar](#)

20. 20.

Larrazábal, G. O., Shinagawa, T., Martín, A. J. & Pérez-Ramírez, J. Microfabricated electrodes unravel the role of interfaces in multicomponent copper-based CO₂ reduction catalysts. *Nat. Commun* **9**, 1477 (2018).

[ADS](#) [Article](#) [Google Scholar](#)

21. 21.

Martin, O. et al. Indium oxide as a superior catalyst for methanol synthesis by CO₂ hydrogenation. *Angew. Chem. Int. Ed.* **128**, 6369–6373 (2016).

[Article](#) [Google Scholar](#)

22. 22.

Hagleitner, D. R. et al. Bulk and surface characterization of In₂O₃(001) single crystals. *Phys. Rev. B* **85**, 115441 (2012).

[ADS](#) [Article](#) [Google Scholar](#)

23. 23.

Wagner, M. et al. Reducing the In₂O₃(111) surface results in ordered indium adatoms. *Adv. Mater. Interfaces* **1**, 1400289 (2014).

[Article](#) [Google Scholar](#)

24. 24.

Capdevila-Cortada, M., Vilé, G., Teschner, D., Pérez-Ramírez, J. & López, N. Reactivity descriptors for ceria in catalysis. *Appl. Catal. B* **197**, 299–312 (2016).

[CAS](#) [Article](#) [Google Scholar](#)

25. 25.

Wagner, M. et al. Resolving the structure of a well-ordered hydroxyl overlayer on In₂O₃(111): nanomanipulation and theory. *ACS Nano* **11**, 11531–11541 (2017).

[CAS](#) [Article](#) [Google Scholar](#)

26. 26.

Yurtsever, A. et al. Understanding image contrast formation in TiO₂ with force spectroscopy. *Phys. Rev. B* **85**, 125416 (2012).

[ADS](#) [Article](#) [Google Scholar](#)

27. 27.

Stetsovich, O. et al. Atomic species identification at the (101) anatase surface by simultaneous scanning tunnelling and atomic force microscopy. *Nat. Commun.* **6**, 7265 (2015).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

28. 28.

Sugimoto, Y. et al. Quantum degeneracy in atomic point contacts revealed by chemical force and conductance. *Phys. Rev. Lett.* **111**, 106803–106805 (2013).

[ADS](#) [Article](#) [Google Scholar](#)

29. 29.

Kowalski, P. M., Meyer, B. & Marx, D. Composition, structure, and stability of the rutile TiO₂(110) surface: oxygen depletion, hydroxylation, hydrogen migration, and water adsorption. *Phys. Rev. B* **79**, 115410 (2009).

[ADS](#) [Article](#) [Google Scholar](#)

30. 30.

Lackner, P. et al. Water adsorption at zirconia: from the ZrO₂(111)/Pt₃Zr(0001) model system to powder samples. *J. Mater.*

Chem. A **6**, 17587–17601 (2018).

[CAS Article](#) [Google Scholar](#)

31. 31.

Huber, F. & Giessibl, F. J. Low noise current preamplifier for qPlus sensor deflection signal detection in atomic force microscopy at room and low temperatures. *Rev. Sci. Instrum.* **88**, 073702 (2017).

[ADS Article](#) [Google Scholar](#)

32. 32.

Štubian, M., Bobek, J., Setvín, M., Diebold, U. & Schmid, M. Fast low-noise transimpedance amplifier for scanning tunneling microscopy and beyond. *Rev. Sci. Instrum.* **91**, 074701 (2020).

[ADS Article](#) [Google Scholar](#)

33. 33.

Sader, J. E. & Jarvis, S. P. Accurate formulas for interaction force and energy in frequency modulation force spectroscopy. *Appl. Phys. Lett.* **84**, 1801–1803 (2004).

[ADS CAS Article](#) [Google Scholar](#)

34. 34.

Giannozzi, P. et al. Quantum ESPRESSO: a modular and open-source software project for quantum simulations of materials. *J. Phys. Condens. Matter* **21**, 395502 (2009).

[Article](#) [Google Scholar](#)

35. 35.

Perdew, J., Burke, K. & Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **77**, 3865–3868 (1996).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

36. 36.

Vanderbilt, D. Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Phys. Rev. B* **41**, 7892–7895 (1990).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

37. 37.

Wagner, M. et al. Well-ordered In adatoms at the $\text{In}_2\text{O}_3(111)$ surface created by Fe deposition. *Phys. Rev. Lett.* **117**, 206101 (2016).

[ADS](#) [Article](#) [Google Scholar](#)

38. 38.

Grimme, S., Antony, J., Ehrlich, S. & Krieg, H. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H–Pu. *J. Chem. Phys.* **132**, 154104 (2010).

[ADS](#) [Article](#) [Google Scholar](#)

39. 39.

Sokolović, I. et al. Resolving the adsorption of molecular O_2 on the rutile $\text{TiO}_2(110)$ surface by noncontact atomic force microscopy. *Proc. Natl Acad. Sci. USA* **117**, 14827–14837 (2020).

[Article](#) [Google Scholar](#)

40. 40.

Hapala, P. et al. Mechanism of high-resolution STM/AFM imaging with functionalized tips. *Phys. Rev. B* **90**, 085421 (2014).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

41. 41.

Linstrom, P. J. & Mallard, W. G. (eds) *NIST Chemistry WebBook: NIST Standard Reference Database Number 69* (National Institute of Standards and Technology, 2018); <https://webbook.nist.gov/chemistry>.

[Download references](#)

Acknowledgements

This work was supported by the Austrian Science Fund (FWF), project V 773-N (Elise-Richter-Stelle, M.W.) and Z 250-N27 (Wittgenstein Prize, U.D.), as well as the German Research Foundation (DFG), Research Unit FOR 1878 (funCOS, B.M.). M.W. and U.D. also acknowledge funding under the Horizon 2020 Research and Innovation Programme under the grant agreement number 810626. M. Setvin acknowledges the support of GAUK Primus/20/SCI/009. Computational resources were provided by LRZ Garching (project pn98fa)and RRZ Erlangen.

Author information

Affiliations

1. Institute of Applied Physics, TU Wien, Vienna, Austria

Margareta Wagner, Martin Setvin, Michael Schmid & Ulrike Diebold

2. Central European Institute of Technology (CEITEC), Brno University of Technology, Brno, Czech Republic

Margareta Wagner

3. Interdisciplinary Center for Molecular Materials, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

Bernd Meyer

4. Computer Chemistry Center, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany

Bernd Meyer

5. Department of Surface and Plasma Science, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic

Martin Setvin

Authors

1. Margareta Wagner

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Bernd Meyer

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Martin Setvin

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Michael Schmid

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Ulrike Diebold

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Contributions

M.W. and M. Setvin conducted the experiments. M.W., M. Setvin and M. Schmid analysed the AFM data. B.M. performed the calculations and derived the scaling relationships. M.W., B.M. and U.D. wrote the manuscript, which was reviewed and edited by all authors. U.D. oversaw the project.

Corresponding author

Correspondence to [Ulrike Diebold](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Leo Gross 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 The clean In₂O₃\(111\) surface imaged with various AFM tip terminations.**](#)

a, b, Structural models of the $\text{In}_2\text{O}_3(111)$ surface, with emphasis on the surface O (red) atoms (**a**) and In (blue, green) atoms (**b**). Also indicated is the unit cell. **c**, AFM image with an O-terminated tip. The intentional functionalization was performed on a reduced rutile $\text{TiO}_2(110)$ surface with adsorbed O_2 molecules by applying voltage pulses (approximately +3 V) above the molecular oxygen species. Previous experience has shown that such a tip functionalization provides excellent resolution of the oxygen sublattice in the repulsive regime³⁹. Such tips provide negligible attractive interaction with the anion lattice and they are rigid; the resulting images are therefore not distorted by bending the tip apex⁴⁰ and clearly show the O sublattice, $\text{O}(\alpha)$ to $\text{O}(\delta)$. **d**, AFM image taken after the tip was gently pushed into the hydroxylated $\text{In}_2\text{O}_3(111)$ surface to induce an OH termination. The more flexible tip termination leads to crests in the images. The dark, round features correspond to In, which are indicated by the green and blue dots (attractive interaction between negatively charged (tip) OH and In anions). The bright maxima correspond to $\text{O}(3c)$.

Extended Data Fig. 2 Distance-dependent STM/AFM images of the clean $\text{In}_2\text{O}_3(111)$ surface.

a, b, O-terminated tip. **a**, Constant-height STM image (tunnelling current, I_T). **b**, Constant-height AFM images (frequency shift, Δf). From top to bottom, the tip–sample distance was progressively reduced in steps of about 50 pm. **c, d**, OH-terminated tip. **c**, Constant-current STM image. **d**, Same as **c** but the distance was reduced in steps of about 40 pm.

Extended Data Fig. 3 Adsorption sites of OH groups on $\text{In}_2\text{O}_3(111)$.

a–f, Constant-height nc-AFM images of three dissociated water molecules in adjoining unit cells, taken with decreasing tip height. The height difference Δz between the subsequent images is 10 to 15 pm. Each dissociated water molecule gives rise to two OH groups adsorbed next to each other, O_WH and O_SH (where the ‘W’ and ‘S’ indicate the origin of the oxygen atom, that is, the water molecule or the surface, respectively)²⁵. In

a, with the tip farthest away, both OH groups are imaged as dark features. As the tip comes closer to the OH groups (**b**, **c**), the O_WH (sticking farther away from the surface than the O_SH) turn bright (onset of repulsive interaction, but overall attractive forces). In **d**, the O_SH also start to be imaged as bright features. Approaching the surface even more, **e**, **f** reveal the O(3c) lattice atoms of the surface; this information was used to determine the adsorption sites of the OH groups experimentally. **g**, Profiles of the frequency shift across the O_SH and O_WH pair as indicated by the arrows in **a–f**. **h**, Same image as **f** with the O(3c) lattice superimposed (O(α), white; O(β), red; O(γ), orange; O(δ), yellow). The site of the O_SH is identified as an O(β) (circle filled black). **i**, Atomic model of the surface including the adsorption sites of the O_SH and O_WH (black circle, filled white), which bridges two In(5c) atoms nearby. The adsorption sites agree with previous STM and DFT results in ref. [25](#).

Extended Data Fig. 4 The hydroxylated In₂O₃(111) surface with increasing water coverages in STM and AFM.

The STM (top) and AFM images (bottom) were acquired at different regions of the surface. **a1**, Oxygen-terminated tip. **b1–d1**, OH-terminated tips. **a**, **a1**, Clean In₂O₃(111) surface. The contrast in STM (empty states) is dominated by the high density of states of the In(5c) and the lower density of states at the In(6c), which gives dark triangles. In AFM, the contrast is dominated by the topmost atoms of the surface, that is, the 12 O(3c) per unit cell. **b**, **b1**, Single dissociated water molecules, O_WH and O_SH. **c**, **c1**, Two and three dissociated water molecules per unit cell. **d**, **d1**, Saturation with three dissociated water molecules per unit cell in symmetry-equivalent sites, giving rise to a ‘propeller-like feature’ consisting of three brighter (O_WH) and darker (O_SH at O(β) site) at equivalent positions. For DFT calculations and structural relaxations see ref. [25](#).

Extended Data Fig. 5 Manipulation of O_SH groups by voltage pulses.

a, STM image before (**a1**) and during (**a2**) the manipulation. Five pulses (two times +2.8 V and three times +2.7 V, 20 pA, marked with crosses) were applied during STM imaging in the centre of individual propeller-like structures. **b, c**, AFM image of the same surface area before the manipulation (**b**; with crosses marking the position of the pulses) and after the manipulation (**c**). **d, e**, Cartoons identifying the various species before (**d**) and after (**e**) the manipulation. In each manipulation, one H per propeller was removed, leaving behind a denuded O(β), indicated in **e** and visible as a smaller, bright dot in the AFM image in **c**. Two H have desorbed and three H have re-adsorbed, visible as very dark features in AFM in **c** and indicated in yellow in **e**.

Extended Data Fig. 6 Data evaluation of $F(z)$ curves.

a, Frequency-shift curves acquired on different OH groups and the $\text{In}_2\text{O}_3(111)$ background measured in-between the OH groups (mostly in regions A and C, see Fig. 1). **b**, The averaged curves from **a** (solid lines) and after background subtraction (dashed lines). **c**, Background-corrected $F(z)$ curves obtained from **b** using Sader's formula³³ ($f_R = 77.7$ kHz, $k = 5,400 \text{ N m}^{-1}$, $A = 100 \text{ pm}$). **d**, The $F(z)$ curves of Fig. 2a showing the whole z range.

Extended Data Fig. 7 Transferability and reproducibility of the method.

$F(z)$ curves on $\text{O}_\text{S}\text{H}/\text{TiO}_2(110)$ and OH/zirconium oxide/ $\text{Pt}_3\text{Zr}(0001)$ obtained with the OH-terminated indium oxide tip. **a**, $\text{TiO}_2(110)$: experimental data. The tip was prepared on hydroxylated $\text{In}_2\text{O}_3(111)$, and $F(z)$ curves were taken on the $\text{O}_\text{W}\text{H}$ and $\text{O}_\text{S}\text{H}(\beta)$ (labelled as such) to ascertain the tip termination ($f_R = 69.6$ kHz, $k = 5,400 \text{ N m}^{-1}$, $A = 60 \text{ pm}$). For reference, the curves from the main text (Fig. 2) are also plotted on the left; for better visibility they are shifted horizontally by -2 \AA . This tip was used to obtain the $F(z)$ curve on $\text{O}_\text{S}\text{H}/\text{TiO}_2(110)$ (brown, shifted to the right). Approaching the $\text{O}_\text{S}\text{H}$ on TiO_2 even closer results in picking up the H. The hydroxylated $\text{TiO}_2(110)$ surface was prepared following a well

established procedure²⁶, which results in one type of hydroxyl at a bridging O(2c) atom. **b**, Calculated short-range $F(z)$ curves on bridging hydroxyls on TiO₂(110) for two OH coverages (1/4 and 1/8 monolayer) obtained with the OH-terminated InO_x tip. **c**, Zirconium oxide: the standard tip was prepared on hydroxylated In₂O₃(111), and $F(z)$ curves were taken on the O_WH and O_SH(β) (labelled as such) to ascertain the tip termination. For reference, the curves from the main text (Fig. 2) are also plotted on the left; for better visibility they are shifted horizontally by -2 \AA . This tip was used to take $F(z)$ curves on hydroxylated zirconium oxide (purple, shifted to the right). **d**, Zirconium oxide: the same type of measurement, but with an unknown tip termination that gives more shallow minima for O_WH and O_SH on hydroxylated In₂O₃(111). Note that both tips measure the same relative positions in the force minima of the strongly bound H on In₂O₃(111) and zirconium oxide, that is, the force on OH/zirconium oxide lies between O_WH and O_SH(β) ($f_R = 77.2 \text{ kHz}$, $k = 5,400 \text{ N m}^{-1}$, $A = 240 \text{ pm}$ (In₂O₃), 250 pm (zirconium oxide)). The ultrathin zirconium oxide layer was prepared following the method of ref. ³⁰. The surface was exposed to 2 langmuir (1 L = $1.33 \times 10^{-6} \text{ mbar s}$) of water at 320 K.

Extended Data Fig. 8 Reactivity of the hydroxylated surface.

a, Atom-resolved PDOS of the four symmetry-inequivalent three-fold coordinated O surface sites on the fully hydroxylated In₂O₃(111) surface with three O_WH and three O_SH(β) per unit cell. The VBM is at 0 eV. **b**, Calculated adsorption energies $\langle\{ E_{\text{ads}} \} \rangle$ for the remaining three unprotonated O(3c) sites (with respect to the H₂ gas phase molecule). The adsorption of water and the formation of the hydroxylated surface structure slightly modifies the reactivity of the unprotonated O(α), O(γ) and O(δ) sites. The saturation of the O(β) sites by protons leads to a strong downward shift of the O(β) 2p states. The VBM is now formed by the O(δ) 2p states, followed by the O 2p states from the O(α) and O(γ) sites. The O(δ), which are the second-most-reactive O species on the uncovered surface, are thus expected to be the most reactive sites on the hydroxylated In₂O₃(111) surface. This is confirmed by the calculated H adsorption energies (**b**). Whereas for O(γ) the H adsorption energies are rather similar

on the clean and on the hydroxylated surface, the local relaxations upon water adsorption make the O(α) sites slightly more and the O(δ) slightly less reactive (see **b** and inset in Fig. [1b](#)). As for the clean surface, the pronounced peak in the PDOS of the three-fold coordinated O_S surface atoms at the VBM also leads to an upward shift of their *p*-band centre. The shift of the *p*-band centre has been taken as an indirect measure of surface reactivity^{[24](#)} and confirms the expected trend in the PA of the different surface sites.

Extended Data Fig. 9 Tip model for the AFM calculations.

a, Relaxed structure of an In₈O₁₂ cluster cut out of an In₂O₃(111) slab. The cut is centred around the high-symmetry B site of the surface and its depth is four In₂O₃(111) trilayers. The In₂O₃(111) trilayers are all equivalent but they are shifted such that the high-symmetry sites with a three-fold symmetry axis (see Fig. [1a](#)) form a stacking sequence of BCAB. The cluster consists of an In atom at the B site together with three O atoms, and six-membered In₃O₃ rings around the A and C sites. The cluster is stoichiometric and stable upon geometry optimization, it has a three-fold symmetry axis and its HOMO–LUMO gap is 1.62 eV (PBE). **b**, Cluster after dissociative adsorption of six water molecules to saturate the broken In–O bonds. The OH groups were added to the six In atoms in the two central planes, and the protons were placed on top of the two-fold bridging O atoms. This increased the HOMO–LUMO gap to 2.57 eV. In the next step, the upper In(OH)₃ cap was removed (dashed circle). This was done to eliminate the top three-fold coordinated In atom and to reduce the height of the tip model. This saves some computer time in the AFM calculations, as the thickness of the vacuum region can be reduced. Removing the cap increases the HOMO–LUMO gap to 2.68 eV. At this point, the cluster still has a three-fold symmetry axis. **c**, Structure of the final tip. One additional water molecule is added to the cluster (**b**). The OH group is adsorbed at the In apex atom of the tip and the proton is added to one of the remaining O atoms. This breaks the three-fold symmetry of the tip. Careful test calculations showed that the specific choice for the adsorption site of the proton is not important. All tips gave basically the same *F*(*z*) curves. In the end, the proton was split into three parts, and three pseudo-hydrogen atoms

with nuclear charge of +1/3 were added to the three O atoms closest to the tip apex indicated by arrows. The HOMO–LUMO gap of the final cluster calculated with the PBE functional is 2.64 eV.

Extended Data Fig. 10 Illustration of the background subtraction for the calculated $F(z)$ curves.

Grey dashed lines are calculated $F(z)$ curves at the high-symmetry A and C sites on the hydroxylated and the water-free $\text{In}_2\text{O}_3(111)$ surface, each for two different azimuthal tip orientations. The average of the eight curves is shown as a grey solid line. In the background-correction procedure, this curve is subtracted from all calculated $F(z)$ curves above surface OH groups. Black and red graphs show the data for the $\text{O}_\text{W}\text{H}$ and $\text{O}_\text{S}\text{H}(\beta)$ groups before background subtraction (dashed, calculated curves for two tip orientations; solid, average of the two dashed curves).

Extended Data Fig. 11 Correlations between force and energy minima, PA and $\text{p}K_\text{a}$, using probe molecules.

a, List of probe molecules with their experimental PA and $\text{p}K_\text{a}$ values. As our aim is to assign PAs to surface O species, only oxygen-based acids and alcohols were selected. The experimental PAs (for the corresponding base, that is, the $\text{X}-\text{O}^-$ anion) are taken from the National Institute of Standards and Technology (NIST) database⁴¹. The number of listed NIST entries is given in parenthesis. In case that data from several experiments are available, averaged values and uncertainties are given. **b**, Correlation between force and energy minima. The linear fit to the calculated minima of the $E(z)$ and $F(z)$ curves for our probe molecules (red dashed line) shows an almost perfect linear correlation of the data points. The value of the energy minimum in the $E(z)$ curves (that is, $F(z) = 0$) is the natural measure for the strength of the hydrogen bond that forms between the OH groups of the tip and the molecules. This linear relationship allows us to focus on establishing a correlation between PA or acidity and the force, and not the energy minimum. This is more convenient, because the force minimum is the natural measure from the AFM experiments. Although the energy minimum can be obtained by integrating the $F(z)$ curves, it is not trivial to

eliminate ambiguities stemming from a proper choice for the zero point of energy, which determines the integration constant. Furthermore, the force minimum is encountered at a larger OH–tip distance than the energy minimum. The force minimum is therefore less influenced by other interactions between the tip and side groups of the molecule or neighbouring atoms on the surface. **c**, Correlation between the PA and acidity constant pK_a of the selected probe molecules. The PA also governs the pK_a in wet, solution-based chemical processes: an OH group with a strongly bound proton (high PA of the O atom) is a weak acid, and an OH group with a weakly bound proton (low PA of the O atom) is a strong acid. A linear fit (red dashed line) to the experimental data in **a** shows the expected trend, that is, strong acids have a low PA, and weak acids bind their proton more strongly. However, in addition to the PA, the pK_a also includes the Gibbs free energy of solvation of the acid (XOH), the conjugated base (XO^-) and the H^+ (see the thermodynamic cycle in Methods). These species have different solubilities, which leads to a large scatter of the data points. **d**, Correlation between calculated AFM force minima and the experimental acidity constants pK_a of the probe molecules. As expected from the discussion in **c**, the force minima show a larger scatter around the red dashed regression line than when plotted with respect to the PA (see Fig. 4). This is because the AFM measurements are done in vacuum and do not include information about solvation free energies. Thus, the prediction of absolute pK_a values for our OH groups on the $In_2O_3(111)$ surface based on the AFM measurements alone is not possible. Still, there is a clear trend that strong acids form strong H bonds with the AFM tip (deep force minimum, low PA), whereas weak acids form weak H bonds (shallow force minimum, high PA). If we assume that the solvation energies of the structurally rather similar OH groups do not differ too much, then the deviation from the regression line would be similar for all of them, which would allow us to predict at least relative pK_a changes between the OH groups by using the slope of the linear regression. The AFM-measured difference in the force minima of 81 pN for the $O_SH(\beta)$ and $O_SH(\gamma)$ sites (see Fig. 2) then translates to a difference in acidity of 5.5 pK_a units, which is reasonable.

Supplementary information

AFM tip approaching an O

Supplementary Video 1 $s\text{H}(\beta)$ hydroxyl. Sequence of configurations from the DFT calculation of a force-distance curve between an $\text{O}_\text{S}\text{H}(\beta)$ surface hydroxyl and an OH-terminated indium oxide tip. When the tip approaches the surface, a hydrogen bond forms between the O atom at the tip apex and the proton of the surface hydroxyl.

AFM tip approaching an O

Supplementary Video 2 $w\text{H}$ hydroxyl. Sequence of configurations from the DFT calculation of a force-distance curve between an $\text{O}_\text{W}\text{H}$ surface hydroxyl and an OH-terminated indium oxide tip. When the tip approaches the surface, a hydrogen bond forms between the O atom at the tip apex and the proton of the surface hydroxyl.

Rights and permissions

Reprints and Permissions

About this article



Check for
updates

Cite this article

Wagner, M., Meyer, B., Setvin, M. *et al.* Direct assessment of the acidity of individual surface hydroxyls. *Nature* **592**, 722–725 (2021).
<https://doi.org/10.1038/s41586-021-03432-3>

Download citation

- Received: 02 January 2020
- Accepted: 08 March 2021
- Published: 28 April 2021
- Issue Date: 29 April 2021
- DOI: <https://doi.org/10.1038/s41586-021-03432-3>

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03432-3>

- Article
- [Published: 28 April 2021](#)

Accelerated global glacier mass loss in the early twenty-first century

- [Romain Hugonet](#) ORCID: orcid.org/0000-0002-0955-1306^{1,2,3},
- [Robert McNabb](#) ORCID: orcid.org/0000-0003-0016-493X^{4,5},
- [Etienne Berthier](#) ORCID: orcid.org/0000-0001-5978-9155¹,
- [Brian Menounos](#)^{6,7},
- [Christopher Nuth](#)^{5,8},
- [Luc Girod](#) ORCID: orcid.org/0000-0003-3627-5885⁵,
- [Daniel Farinotti](#) ORCID: orcid.org/0000-0003-3417-4570^{2,3},
- [Matthias Huss](#)^{2,3,9},
- [Ines Dussaillant](#) ORCID: orcid.org/0000-0003-0617-7731^{1,10},
- [Fanny Brun](#)¹¹ &
- [Andreas Kääb](#) ORCID: orcid.org/0000-0002-6017-6564⁵

[Nature](#) volume **592**, pages 726–731(2021) [Cite this article](#)

- 3899 Accesses
- 1998 Altmetric
- [Metrics details](#)

Subjects

- [Climate-change impacts](#)
- [Cryospheric science](#)
- [Hydrology](#)

Abstract

Glaciers distinct from the Greenland and Antarctic ice sheets are shrinking rapidly, altering regional hydrology¹, raising global sea level² and elevating natural hazards³. Yet, owing to the scarcity of constrained mass loss observations, glacier evolution during the satellite era is known only partially, as a geographic and temporal patchwork^{4,5}. Here we reveal the accelerated, albeit contrasting, patterns of glacier mass loss during the early twenty-first century. Using largely untapped satellite archives, we chart surface elevation changes at a high spatiotemporal resolution over all of Earth's glaciers. We extensively validate our estimates against independent, high-precision measurements and present a globally complete and consistent estimate of glacier mass change. We show that during 2000–2019, glaciers lost a mass of 267 ± 16 gigatonnes per year, equivalent to 21 ± 3 per cent of the observed sea-level rise⁶. We identify a mass loss acceleration of 48 ± 16 gigatonnes per year per decade, explaining 6 to 19 per cent of the observed acceleration of sea-level rise. Particularly, thinning rates of glaciers outside ice sheet peripheries doubled over the past two decades. Glaciers currently lose more mass, and at similar or larger acceleration rates, than the Greenland or Antarctic ice sheets taken separately^{7,8,9}. By uncovering the patterns of mass change in many regions, we find contrasting glacier fluctuations that agree with the decadal variability in precipitation and temperature. These include a North Atlantic anomaly of decelerated mass loss, a strongly accelerated loss from northwestern American glaciers, and the apparent end of the Karakoram anomaly of mass gain¹⁰. We anticipate our highly resolved estimates to advance the understanding of drivers that govern the distribution of glacier change, and to extend our capabilities of predicting these changes at all scales. Predictions robustly benchmarked against observations are critically needed to design adaptive policies for the local- and regional-scale management of water resources and cryospheric risks, as well as for the global-scale mitigation of sea-level rise.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

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: Regional glacier mass changes and their temporal evolution from 2000 to 2019.

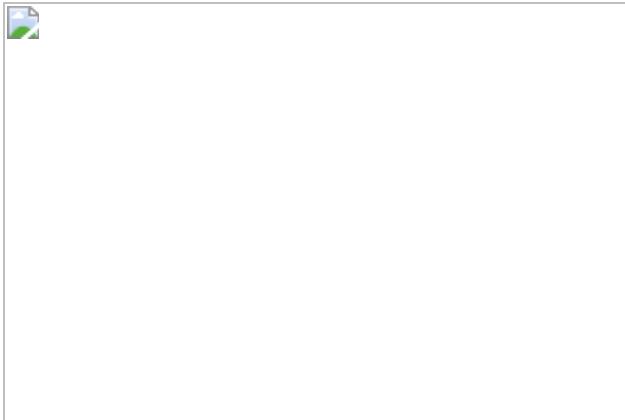


Fig. 2: Spatial distribution of glacier elevation change between 2000 and 2019.

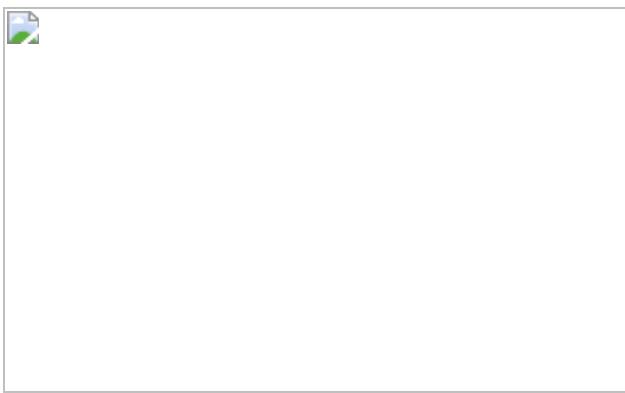


Fig. 3: Comparison to previous global estimates.

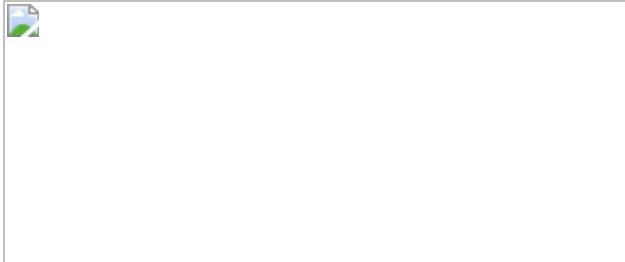
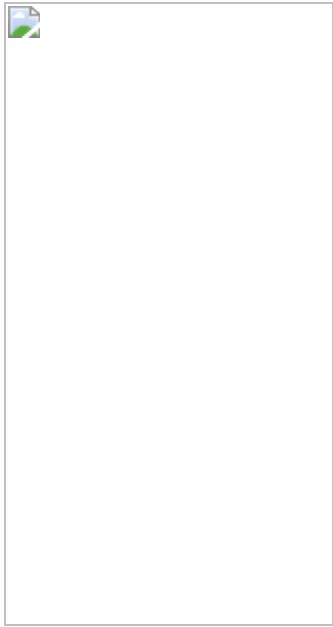


Fig. 4: Decadal patterns of glacier thinning are consistent with decadal variations in precipitation and temperature.



Data availability

Global, regional, tile and per-glacier elevation and mass change time series, elevation change maps for 5-, 10- and 20-year periods at 100 m resolution, and tables in this article are publicly available at <https://doi.org/10.6096/13>. [Source data](#) are provided with this paper.

Code availability

The code developed for the global processing and analysis of all data, and to generate figures and tables in this article, is publicly available at https://github.com/rhugonnet/ww_tv0l_study. Code concomitantly developed for processing ASTER data is available as the Python package pymmaster at <https://github.com/luc-girod/MMASTER-workflows> (with supporting documentation at <https://mmaster-workflows.readthedocs.io>) and for processing DEM time series as the Python package pyddem at <https://github.com/iamdonovan/pyddem> (with supporting documentation at <https://pyddem.readthedocs.io>).

References

1. 1.

Pritchard, H. D. Asia's shrinking glaciers protect large populations from drought stress. *Nature* **569**, 649–654 (2019).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

2. 2.

WCRP Global Sea Level Budget Group. Global sea-level budget 1993–present. *Earth Syst. Sci. Data* **10**, 1551–1590 (2018).

[ADS](#) [Google Scholar](#)

3. 3.

Stoffel, M. & Huggel, C. Effects of climate change on mass movements in mountain environments. *Prog. Phys. Geogr.* **36**, 421–439 (2012).

[Google Scholar](#)

4. 4.

IPCC. *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate* (eds Pörtner, H. O. et al.) (IPCC, 2019).

5. 5.

Gardner, A. et al. A reconciled estimate of glacier contributions to sea level rise: 2003 to 2009. *Science* **340**, 852–857 (2013).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

6. 6.

Nerem, R. S. et al. Climate-change-driven accelerated sea-level rise detected in the altimeter era. *Proc. Natl Acad. Sci. USA* **115**, 2022–2025 (2018).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

7. 7.

IMBIE Team. Mass balance of the Greenland Ice Sheet from 1992 to 2018. *Nature* **579**, 233–239 (2020).

[ADS](#) [Google Scholar](#)

8. 8.

IMBIE team. Mass balance of the Antarctic Ice Sheet from 1992 to 2017. *Nature* **558**, 219–222 (2018).

[ADS](#) [Google Scholar](#)

9. 9.

Smith, B. et al. Pervasive ice sheet mass loss reflects competing ocean and atmosphere processes. *Science* **368**, 1239–1242 (2020).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

10. 10.

Kääb, A., Berthier, E., Nuth, C., Gardelle, J. & Arnaud, Y. Contrasting patterns of early twenty-first-century glacier mass change in the Himalayas. *Nature* **488**, 495–498 (2012).

[ADS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

11. 11.

Kulp, S. A. & Strauss, B. H. New elevation data triple estimates of global vulnerability to sea-level rise and coastal flooding. *Nat. Commun.* **10**, 4844 (2019); author correction **10**, 5752 (2019).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

12. 12.

Immerzeel, W. W. et al. Importance and vulnerability of the world's water towers. *Nature* **577**, 364–369 (2020).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

13. 13.

Marzeion, B., Cogley, J. G., Richter, K. & Parkes, D. Attribution of global glacier mass loss to anthropogenic and natural causes. *Science* **345**, 919–921 (2014).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

14. 14.

Huss, M. & Hock, R. Global-scale hydrological response to future glacier mass loss. *Nat. Clim. Chang.* **8**, 135–140 (2018).

[ADS](#) [Google Scholar](#)

15. 15.

IPCC. *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part B: Regional Aspects* (Cambridge University Press, 2014).

16. 16.

Cauvy-Fraunié, S. & Dangles, O. A global synthesis of biodiversity responses to glacier retreat. *Nat. Ecol. Evol.* **3**, 1675–1685 (2019).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

17. 17.

World Glacier Monitoring Service (WGMS). *Fluctuations of Glaciers Database* https://wgms.ch/data_databaseversions/ (2019).

18. 18.

Bamber, J. L., Westaway, R. M., Marzeion, B. & Wouters, B. The land ice contribution to sea level during the satellite era. *Environ. Res. Lett.* **13**, 063008 (2018); corrigendum **13**, 099502 (2018).

[ADS](#) [Google Scholar](#)

19. 19.

Wouters, B., Gardner, A. S. & Moholdt, G. Global glacier mass loss during the GRACE satellite mission (2002–2016). *Front. Earth Sci.* **7**, 96 (2019).

[ADS](#) [Google Scholar](#)

20. 20.

Ciraci, E., Velicogna, I. & Swenson, S. Continuity of the mass loss of the world's glaciers and ice caps from the GRACE and GRACE Follow-On missions. *Geophys. Res. Lett.* **47**, 226 (2020).

[Google Scholar](#)

21. 21.

Zemp, M. et al. Global glacier mass changes and their contributions to sea-level rise from 1961 to 2016. *Nature* **568**, 382–386 (2019).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

22. 22.

RGI Consortium. *Randolph Glacier Inventory – A Dataset of Global Glacier Outlines*. Technical Report https://www.glims.org/RGI/00_rgi60_TechnicalNote.pdf (Global Land Ice Measurements from Space, 2017).

23. 23.

Huss, M. Density assumptions for converting geodetic glacier volume change to mass change. *Cryosphere* **7**, 877–887 (2013).

[ADS](#) [Google Scholar](#)

24. 24.

Ablain, M. et al. Uncertainty in satellite estimates of global mean sea-level changes, trend and acceleration. *Earth Syst. Sci. Data* **11**, 1189–1202 (2019).

[ADS](#) [Google Scholar](#)

25. 25.

Velicogna, I. et al. Continuity of ice sheet mass loss in Greenland and Antarctica from the GRACE and GRACE Follow-On missions. *Geophys. Res. Lett.* **47**, L11501 (2020).

[Google Scholar](#)

26. 26.

Larsen, C. F. et al. Surface melt dominates Alaska glacier mass balance. *Geophys. Res. Lett.* **42**, 5902–5908 (2015).

[ADS](#) [Google Scholar](#)

27. 27.

Blazquez, A. et al. Exploring the uncertainty in GRACE estimates of the mass redistributions at the Earth surface: implications for the global water and sea level budgets. *Geophys. J. Int.* **215**, 415–430 (2018).

[ADS](#) [Google Scholar](#)

28. 28.

Shean, D. E. et al. A systematic, regional assessment of High Mountain Asia glacier mass balance. *Front. Earth Sci.* **7**, 363 (2020).

[ADS](#) [Google Scholar](#)

29. 29.

Braun, M. H. et al. Constraining glacier elevation and mass changes in South America. *Nat. Clim. Chang.* (2019).

30. 30.

Dehecq, A. et al. Elevation changes inferred from TanDEM-X data over the Mont-Blanc area: impact of the X-band interferometric bias. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **9**, 3870–3882 (2016).

[ADS](#) [Google Scholar](#)

31. 31.

Sandberg Sørensen, L. et al. 25 years of elevation changes of the Greenland Ice Sheet from ERS, Envisat, and CryoSat-2 radar altimetry. *Earth Planet. Sci. Lett.* **495**, 234–241 (2018).

[ADS](#) [Google Scholar](#)

32. 32.

Bevis, M. et al. Accelerating changes in ice mass within Greenland, and the ice sheet's sensitivity to atmospheric forcing. *Proc. Natl Acad. Sci. USA* **116**, 1934–1939 (2019).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

33. 33.

Garreaud, R. D. et al. The Central Chile Mega Drought (2010–2018): a climate dynamics perspective. *Int. J. Climatol.* **40**, 421–439 (2020).

[Google Scholar](#)

34. 34.

Raper, S. C. B. & Braithwaite, R. J. Low sea level rise projections from mountain glaciers and icecaps under global warming. *Nature* **439**, 311–313 (2006).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

35. 35.

Parkes, D. & Marzeion, B. Twentieth-century contribution to sea-level rise from uncharted glaciers. *Nature* **563**, 551–554 (2018).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

36. 36.

Becker, J. J. et al. Global bathymetry and elevation data at 30 arc seconds resolution: SRTM30_PLUS. *Mar. Geod.* **32**, 355–371 (2009).

[Google Scholar](#)

37. 37.

Tielidze, L. G. & Wheate, R. D. The Greater Caucasus glacier inventory (Russia, Georgia and Azerbaijan). *Cryosphere* **12**, 81–94 (2018).

[ADS](#) [Google Scholar](#)

38. 38.

Dunse, T. et al. Glacier-surge mechanisms promoted by a hydro-thermodynamic feedback to summer melt. *Cryosphere* **9**, 197–215 (2015).

[ADS](#) [Google Scholar](#)

39. 39.

McMillan, M. et al. Rapid dynamic activation of a marine-based Arctic ice cap: ice cap dynamic activation. *Geophys. Res. Lett.* **41**, 8902–8909 (2014).

[ADS](#) [Google Scholar](#)

40. 40.

Nuth, C. et al. Dynamic vulnerability revealed in the collapse of an Arctic tidewater glacier. *Sci. Rep.* **9**, 5541 (2019).

[ADS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

41. 41.

Howat, I. M., Negrete, A. & Smith, B. E. The Greenland Ice Mapping Project (GIMP) land classification and surface elevation data sets. *Cryosphere* **8**, 1509–1518 (2014).

[ADS](#) [Google Scholar](#)

42. 42.

Fretwell, P. et al. Bedmap2: improved ice bed, surface and thickness datasets for Antarctica. *Cryosphere* **7**, 375–393 (2013).

[ADS](#) [Google Scholar](#)

43. 43.

NASA/METI/AIST/Japan Space systems & U.S./Japan ASTER Science Team. *ASTER Level 1A Data Set – Reconstructed, Unprocessed Instrument Data*. 2001, NASA EOSDIS Land Processes DAAC, 2001); https://doi.org/10.5067/ASTER/AST_L1A.003.

44. 44.

Porter, C. et al. *ArcticDEM* (Harvard Dataverse, 2018);
[https://dataverse.harvard.edu/dataset.xhtml?
persistentId=doi:10.7910/DVN/OHHUKH.](https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/OHHUKH)

45. 45.

Howat, I. M., Porter, C., Smith, B. E., Noh, M.-J. & Morin, P. The reference elevation model of Antarctica. *Cryosphere* **13**, 665–674 (2019).

[ADS](#) [Google Scholar](#)

46. 46.

Rizzoli, P. et al. Generation and performance assessment of the global TanDEM-X digital elevation model. *ISPRS J. Photogramm. Remote Sens.* **132**, 119–139 (2017).

[ADS](#) [Google Scholar](#)

47. 47.

Vassilaki, D. I. & Stamos, A. A. TanDEM-X DEM: comparative performance review employing LIDAR data and DSMs. *ISPRS J. Photogramm. Remote Sens.* **160**, 33–50 (2020).

[Google Scholar](#)

48. 48.

Nuth, C. & Kääb, A. Co-registration and bias corrections of satellite elevation data sets for quantifying glacier thickness change. *Cryosphere* **5**, 271–290 (2011).

[ADS](#) [Google Scholar](#)

49. 49.

Rupnik, E., Daakir, M. & Pierrot Deseilligny, M. MicMac – a free, open-source solution for photogrammetry. *Open Geospat. Data Softw. Stand.* **2**, 14 (2017).

[Google Scholar](#)

50. 50.

Girod, L., Nuth, C., Kääb, A., McNabb, R. & Galland, O. MMASTER: improved ASTER DEMs for elevation change monitoring. *Remote Sens.* **9**, 704 (2017).

[Google Scholar](#)

51. 51.

Wales, D. J. & Doye, J. P. K. Global optimization by basin-hopping and the lowest energy structures of Lennard–Jones clusters containing up to 110 atoms. *J. Phys. Chem. A* **101**, 5111–5116 (1997).

[CAS](#) [Google Scholar](#)

52. 52.

Noh, M.-J. & Howat, I. M. The surface extraction from TIN based Search-space Minimization (SETSM) algorithm. *ISPRS J. Photogramm. Remote Sens.* **129**, 55–76 (2017).

[ADS](#) [Google Scholar](#)

53. 53.

Dussaillant, I. et al. Two decades of glacier mass loss along the Andes. *Nat. Geosci.* **12**, 802–808 (2019); author correction **13**, 711 (2020).

[ADS](#) [CAS](#) [Google Scholar](#)

54. 54.

Brun, F., Berthier, E., Wagnon, P., Kääb, A. & Treichler, D. A spatially resolved estimate of High Mountain Asia glacier mass balances, 2000–2016. *Nat. Geosci.* **10**, 668–673 (2017); author correction **11**, 543 (2018).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

55. 55.

Toutin, T. Three-dimensional topographic mapping with ASTER stereo data in rugged topography. *IEEE Trans. Geosci. Remote Sens.* **40**, 2241–2247 (2002).

[ADS](#) [Google Scholar](#)

56. 56.

Lacroix, P. Landslides triggered by the Gorkha earthquake in the Langtang valley, volumes and initiation processes. *Earth Planets Space* **68**, 1–10 (2016).

[Google Scholar](#)

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).

[ADS](#) [Google Scholar](#)

58. 58.

Höhle, J. & Höhle, M. Accuracy assessment of digital elevation models by means of robust statistical methods. *ISPRS J. Photogramm. Remote Sens.* **64**, 398–406 (2009).

[ADS](#) [Google Scholar](#)

59. 59.

Williams, C. K. I. & Rasmussen, C. E. *Gaussian Processes for Machine Learning* Vol. 2 (MIT Press, 2006).

60. 60.

Schiefer, E., Menounos, B. & Wheate, R. Recent volume loss of British Columbian glaciers, Canada. *Geophys. Res. Lett.* (2007).

61. 61.

Nuimura, T., Fujita, K., Yamaguchi, S. & Sharma, R. R. Elevation changes of glaciers revealed by multitemporal digital elevation models calibrated by GPS survey in the Khumbu region, Nepal Himalaya, 1992–2008. *J. Glaciol.* **58**, 648–656 (2012).

[ADS](#) [Google Scholar](#)

62. 62.

Willis, M. J., Melkonian, A. K., Pritchard, M. E. & Rivera, A. Ice loss from the Southern Patagonian Ice Field, South America, between 2000 and 2012. *Geophys. Res. Lett.* **39**, L17501 (2012).

[ADS](#) [Google Scholar](#)

63. 63.

Pedregosa, F. et al. Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

[MathSciNet](#) [MATH](#) [Google Scholar](#)

64. 64.

Zwally, H. J., Schutz, R., Hancock, D. & Dimarzio, J. *GLAS/ICESat L2 Global Land Surface Altimetry Data (HDF5)*, Version 34 (NASA Snow and Ice Data Center, 2014); <https://nsidc.org/data/GLAH14>.

65. 65.

Alexandrov, O., McMichael, S. & Beyer., R. A. *IceBridge DMS L3 Ames Stereo Pipeline Photogrammetric DEM, Version 1* (accessed 1 June 2019); <https://nsidc.org/data/IODEM3/versions/1>.

66. 66.

Larsen, C. *IceBridge UAF Lidar Scanner LIB Geolocated Surface Elevation Triplets, Version 1* (accessed 20 February 2020); <https://nsidc.org/data/ILAKS1B/versions/1>.

67. 67.

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).

[ADS](#) [Google Scholar](#)

68. 68.

Harding, D. J. ICESat waveform measurements of within-footprint topographic relief and vegetation vertical structure. *Geophys. Res. Lett.* **32**, L21S10 (2005).

[Google Scholar](#)

69. 69.

Gardelle, J., Berthier, E. & Arnaud, Y. Impact of resolution and radar penetration on glacier elevation changes computed from DEM differencing. *J. Glaciol.* **58**, 419–422 (2012).

[ADS](#) [Google Scholar](#)

70. 70.

McNabb, R., Nuth, C., Kääb, A. & Girod, L. Sensitivity of glacier volume change estimation to DEM void interpolation. *Cryosphere* **13**, 895–910 (2019).

[ADS](#) [Google Scholar](#)

71. 71.

Cressie, N. A. C. *Statistics for Spatial Data* Vol. **4**, 613–617 (Wiley, 1993).

72. 72.

Rolstad, C., Haug, T. & Denby, B. Spatially integrated geodetic glacier mass balance and its uncertainty based on geostatistical analysis: application to the western Svartisen ice cap, Norway. *J. Glaciol.* **55**, 666–680 (2009).

[ADS](#) [Google Scholar](#)

73. 73.

Dehecq, A. et al. Automated processing of declassified KH-9 Hexagon satellite images for global elevation change analysis since the 1970s. *Front. Earth Sci.* **8**, 566802 (2020).

[Google Scholar](#)

74. 74.

Menounos, B. et al. Heterogeneous changes in western North American glaciers linked to decadal variability in zonal wind strength. *Geophys. Res. Lett.* **46**, 200–209 (2018).

[ADS](#) [Google Scholar](#)

75. 75.

Howat, I. M., Smith, B. E., Joughin, I. & Scambos, T. A. Rates of southeast Greenland ice volume loss from combined ICESat and ASTER observations. *Geophys. Res. Lett.* **35**, L17505 (2008).

[ADS](#) [Google Scholar](#)

76. 76.

Wang, D. & Kääb, A. Modeling glacier elevation change from DEM time series. *Remote Sens.* **7**, 10117–10142 (2015).

[ADS](#) [Google Scholar](#)

77. 77.

Cogley, J. G. & Adams, W. P. Mass balance of glaciers other than the ice sheets. *J. Glaciol.* **44**, 315–325 (1998).

[ADS](#) [CAS](#) [Google Scholar](#)

78. 78.

Journel, A. G. & Huijbregts, C. J. *Mining Geostatistics* Vol. 600 (Academic Press, 1978).

79. 79.

Webster, R. & Oliver, M. A. *Geostatistics for Environmental Scientists* (John Wiley & Sons, 2007).

80. 80.

Gräler, B., Pebesma, E. & Heuvelink, G. Spatio-temporal interpolation using gstat. *R J.* **8**, 204 (2016).

[Google Scholar](#)

81. 81.

Mälicke, M. & Schneider, H. D. *Scikit-GStat 0.2.6: A Scipy Flavored Geostatistical Analysis Toolbox Written in Python* (2019);
<https://zenodo.org/record/3531816#.YFsJ737Le00>.

82. 82.

Dussaillant, I., Berthier, E. & Brun, F. Geodetic mass balance of the Northern Patagonian Icefield from 2000 to 2012 using two independent methods. *Front. Earth Sci.* **6**, 8 (2018).

[ADS](#) [Google Scholar](#)

83. 83.

Berthier, E., Scambos, T. A. & Shuman, C. A. Mass loss of Larsen B tributary glaciers (Antarctic Peninsula) unabated since 2002. *Geophys. Res. Lett.* **39**, L13501 (2012).

[ADS](#) [Google Scholar](#)

84. 84.

Granshaw, F. D. & Fountain, A. G. Glacier change (1958–1998) in the North Cascades National Park Complex, Washington, USA. *J. Glaciol.* **52**, 251–256 (2006).

[ADS](#) [CAS](#) [Google Scholar](#)

85. 85.

Pfeffer, W. et al. The Randolph Glacier Inventory: a globally complete inventory of glaciers. *J. Glaciol.* **60**, 537–552 (2014).

[ADS](#) [Google Scholar](#)

86. 86.

Rastner, P. et al. The first complete inventory of the local glaciers and ice caps on Greenland. *Cryosphere* **6**, 1483–1495 (2012).

[ADS](#) [Google Scholar](#)

87. 87.

Bolch, T., Menounos, B. & Wheate, R. Landsat-based inventory of glaciers in western Canada, 1985–2005. *Remote Sens. Environ.* **114**, 127–137 (2010).

[ADS](#) [Google Scholar](#)

88. 88.

Pelto, B. M., Menounos, B. & Marshall, S. J. Multi-year evaluation of airborne geodetic surveys to estimate seasonal mass balance, Columbia and Rocky Mountains, Canada. *Cryosphere* **13**, 1709–1727 (2019).

[ADS](#) [Google Scholar](#)

89. 89.

Wagnon, P. et al. Seasonal and annual mass balances of Mera and Pokalde glaciers (Nepal Himalaya) since 2007. *Cryosphere* **7**, 1769–1786 (2013).

[ADS](#) [Google Scholar](#)

90. 90.

Berthier, E., Schiefer, E., Clarke, G. K. C., Menounos, B. & Rémy, F. Contribution of Alaskan glaciers to sea-level rise derived from satellite imagery. *Nat. Geosci.* **3**, 92–95 (2010).

[ADS](#) [CAS](#) [Google Scholar](#)

91. 91.

Berthier, E., Cabot, V., Vincent, C. & Six, D. Decadal region-wide and glacier-wide mass balances derived from multi-temporal ASTER

Satellite Digital Elevation Models. Validation over the Mont-Blanc area. *Front. Earth Sci.* **4**, 63 (2016).

[ADS](#) [Google Scholar](#)

92. 92.

Glacier Monitoring Switzerland. *Swiss Glacier Volume Change, Release 2018* (2018);
https://doi.glamos.ch/data/volumechange/volumechange_2018_r2018.html.

93. 93.

Bauder, A., Funk, M. & Huss, M. Ice-volume changes of selected glaciers in the Swiss Alps since the end of the 19th century. *Ann. Glaciol.* **46**, 145–149 (2007).

[ADS](#) [Google Scholar](#)

94. 94.

Davaze, L., Rabatel, A., Dufour, A., Hugonnet, R. & Arnaud, Y. Region-wide annual glacier surface mass balance for the European Alps from 2000 to 2016. *Front. Earth Sci.* **8**, 149 (2020).

[ADS](#) [Google Scholar](#)

95. 95.

Schuler, T. V. et al. Reconciling Svalbard Glacier mass balance. *Front. Earth Sci.* **8**, 523646 (2020).

[Google Scholar](#)

96. 96.

Aðalgeirsdóttir, G. et al. Glacier Changes in Iceland From ~1890 to 2019. *Front. Earth Sci.* **8**, 520 (2020).

[ADS](#) [Google Scholar](#)

97. 97.

Hersbach, H. & Dee, D. ERA5 reanalysis is in production. *ECMWF Newslett.* **147**, 5–6 (2016).

[Google Scholar](#)

98. 98.

Skliris, N., Zika, J. D., Nurser, G., Josey, S. A. & Marsh, R. Global water cycle amplifying at less than the Clausius–Clapeyron rate. *Sci. Rep.* **6**, 38752 (2016).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

99. 99.

Sakakibara, D., Sugiyama, S., Sawagaki, T., Marinsek, S. & Skvarca, P. Rapid retreat, acceleration and thinning of Glaciar Upsala, Southern Patagonia Icefield, initiated in 2008. *Ann. Glaciol.* **54**, 131–138 (2013).

[ADS](#) [Google Scholar](#)

100. 100.

Farr, T. G. et al. The Shuttle Radar Topography Mission. *Rev. Geophys.* **45**, RG2004 (2007).

[ADS](#) [Google Scholar](#)

[Download references](#)

Acknowledgements

We thank C. Porter for discussions on ArcticDEM and REMA DEMs, B. Meyssignac for comments on sea-level rise and A. Dehecq for input on the presentation of the manuscript. The GLIMS initiative (in particular J. Kargel and B. Raup) allowed the population of a vast archive of ASTER stereo images over glaciers. Hakai Institute and the University of Northern British Columbia provided computational resources for processing ASTER stereo imagery. SPOT6/7 data were obtained from GEOSUD (ANR-10-EQPX-20, programme ‘Investissements d’Avenir’). ArcticDEM DEMs were provided by the Polar Geospatial Center under NSF-OPP awards 1043681, 1559691 and 1542736 and REMA DEMs were provided by the Byrd Polar and Climate Research Center and the Polar Geospatial Center under NSF-OPP awards 1543501, 1810976, 1542736, 1559691, 1043681, 1541332, 0753663, 1548562, 1238993 and NASA award NNX10AN61G. Computer time was provided through a Blue Waters Innovation Initiative. DEMs were produced using data from DigitalGlobe, Inc. R.H. acknowledges a fellowship from the University of Toulouse. E.B. acknowledges support from the French Space Agency (CNES) through ISIS and TOSCA programmes. R.M., C.N., L.G. and A.K. acknowledge support by ESA through Glaciers_cci and EE10 (4000109873/14/I-NB, 4000127593/19/I-NS, 4000127656/19/NL/FF/gp), and by the European Research Council under the European Union’s Seventh Framework Programme (FP/2007-2013)/ERC grant agreement number 320816. B.M. acknowledges funding from the National Sciences and Engineering Research Council of Canada, the Canada Research Chairs Program, the Tula Foundation and Global Water Futures. R.H., D.F. and M.H. acknowledge funding from the Swiss National Science Foundation, grant number 184634.

Author information

Affiliations

1. LEGOS, Université de Toulouse, CNES, CNRS, IRD, UPS, Toulouse, France

Romain Hugonnet, Etienne Berthier & Ines Dussaillant

2. Laboratory of Hydraulics, Hydrology and Glaciology (VAW), ETH
Zürich, Zürich, Switzerland

Romain Hugonet, Daniel Farinotti & Matthias Huss

3. Swiss Federal Institute for Forest, Snow and Landscape Research
(WSL), Birmensdorf, Switzerland

Romain Hugonet, Daniel Farinotti & Matthias Huss

4. School of Geography and Environmental Sciences, Ulster University,
Coleraine, UK

Robert McNabb

5. Department of Geosciences, University of Oslo, Oslo, Norway

Robert McNabb, Christopher Nuth, Luc Girod & Andreas Kääb

6. Geography Earth and Environmental Sciences, University of Northern
British Columbia, Prince George, British Columbia, Canada

Brian Menounos

7. Hakai Institute, Campbell River, British Columbia, Canada

Brian Menounos

8. The Norwegian Defense Research Establishment, Kjeller, Norway

Christopher Nuth

9. Department of Geosciences, University of Fribourg, Fribourg,
Switzerland

Matthias Huss

10. Department of Geography, University of Zurich, Zurich, Switzerland

Ines Dussaillant

11. IGE, Université Grenoble Alpes, CNRS, IRD, Grenoble INP,
Grenoble, France

Fanny Brun

Authors

1. Romain Hugonnet

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Robert McNabb

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Etienne Berthier

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Brian Menounos

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Christopher Nuth

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Luc Girod

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Daniel Farinotti

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Matthias Huss

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. Ines Dussaillant

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. Fanny Brun

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

11. Andreas Kääb

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Contributions

E.B. and R.H. designed the study with contributions from D.F., M.H. and B.M. L.G., C.N., R.M. and A.K. developed ASTER bias-correction methods. R.H. and R.M. developed glacier elevation GP methods. R.H. implemented spatial statistics methods with inputs from F.B. B.M. assembled and analysed ERA5 data. R.H. performed the processing and analysis of all data with main inputs from E.B., as well as R.M., B.M., D.F., M.H., I.D. and F.B. All authors interpreted the results. R.H. led the writing of the paper and all other co-authors contributed.

Corresponding author

Correspondence to [Romain Hugonet](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Beata Csatho, Thomas Frederikse, Michael Willis 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 Flow chart of the methodology.](#)

Flow diagram describing the processing steps from satellite imagery to global glacier mass change time series. Processing steps correspond to sections in Methods.

[Extended Data Fig. 2 Spatial and temporal coverage of ASTER, ArcticDEM and REMA DEMs.](#)

a–c, Spatial distribution of DEMs as a strip count for ArcticDEM strips above 50° N (**a**), ASTER DEM strips (**b**) and REMA strips below 50° S (**c**), shown on top of a world hillshade³⁶. 67,986 ArcticDEM and 9,369 REMA strips are counted before co-registration to TanDEM-X. This later reduces their number to 40,391 and 3,456, respectively, owing to the limited stable terrain in polar regions. **d**, Temporal distribution of the strip count as a bi-monthly histogram from January 2000 to December 2019. We note that

ArcticDEM and REMA strip footprints ($15\text{ km} \times 50\text{ km}$) are generally much smaller than ASTER DEM strip footprints ($180\text{ km} \times 60\text{ km}$).

Extended Data Fig. 3 Elevation time series estimation.

a–e, Empirical and modelled elevation measurement error (**a**) and temporal covariance of glacier elevation (**b**) estimated globally. These are used to condition the filtering (**c, d**) and elevation time series estimation (**e**) of elevation observations, illustrated here for a $100\text{ m} \times 100\text{ m}$ pixel on the ablation area of Upsala, where a strong nonlinear elevation loss occurred⁹⁹. **a**, Squared measurement error, estimated by the squared NMAD of elevation differences to TanDEM-X on stable terrain as a function of terrain slope and of quality of stereo-correlation. We express the quality of stereo-correlation as a percentage ranging from 0% for poor correlations to 100% for good correlations. **b**, Variance between pairwise glacier elevations in time, or temporal variogram. The empirical temporal variogram is derived from the aggregated median of variances binned by time lags of 0.25 yr. Here, pixels were selected on glacierized terrain showing a linear trend of elevation change (estimated from weighted least squares) between -1.5 and -1.0 m yr^{-1} . The median of the linear trend at these locations (-1.2 m yr^{-1}) was directly used to derive the linear model (orange), which has a quadratic variance. The other models are calibrated so that their sum (dashed black line) matches the empirical variogram. **c**, Spatial and temporal filtering by conditioning a maximum linear elevation change rate from the neighbouring TanDEM-X elevations (see Supplementary Information for further details). **d**, Filtering by successive GP regression fits for credible intervals of size 20σ , 12σ , 9σ , 6σ and 4σ . **e**, Elevation time series of final GP regression after the removal of outliers.

Extended Data Fig. 4 Validation of elevation time series and uncertainties to ICESat and IceBridge.

a–d, ICESat⁶⁴ and IceBridge^{65,66} measurements compared to our surface elevation time series over glacierized terrain in the Saint-Elias Mountains, Alaska (**a–c**) and at the global scale (**d**). **b**, Absolute z-scores (white to purple) are shown on top of the 2000–2019 surface elevation change. z-

scores correspond to elevation differences to ICESat (dashed outlines) or IceBridge (solid outlines), standardized by our time series uncertainty. **c**, Time series for a 100 m × 100 m pixel extracted on the tongue of Agassiz Glacier with neighbouring ICESat and IceBridge elevation differences for demonstration purposes. **d**, Summary of global validation statistics for categories of time, season, region, elevation, observation time lag and total elevation change, with density distributions of measurements for ICESat (light grey) and IceBridge (dark grey). Mean elevation differences, subject to snow-cover biases, are shown only by region (summer mean) and by two-month seasonal component (difference to the annual mean) for each hemisphere.

[Extended Data Fig. 5 Uncertainty analysis of volume changes and validation using high-resolution DEMs.](#)

a–h, Spatial correlation of elevations between the GP time series and ICESat with the time lag to the closest ASTER, ArcticDEM or REMA observation (**a**, **b**), propagation of correlations into specific-volume change uncertainties (**c**), validation of volume change estimates and uncertainties to high-resolution volume changes extracted over the same 588 glaciers and periods (**d–f**) and contribution from all uncertainty sources to the 2000–2019 specific-mass change estimates (**g**, **h**). **a**, An empirical spatial variogram is shown and fitted with a sum of spherical models at correlation lengths of 0.15, 2, 5, 20, 50, 200 and 500 km for elevation differences sampled at 720 days (2 years) from the closest observation. **b**, Spatially correlated variances as a function of the time lag to the closest observation. The model for the variance used during uncertainty propagation is shown in plain lines (sum of quadratic and squared sinusoidal functions optimized by least squares). **c**, Propagation of elevation change uncertainties to volume change uncertainties with varying glacier area. As this computation is specific to the time lag of each pixel to the closest observation, for each glacier, at each time step, **c** refers to an example. The spatial correlations are computed for a time lag to the closest observation, representing the average of our study, of 0–1 yr for 50% of observations, 1–2 yr for 20% of observations, 2–3 yr for 20% of observations and 3–4 yr for 10% of observations. We assume a mean pixel-wise uncertainty of 10 m and simplify by considering only the first step of integration over a continuous

glacierized area (equation (5)). This assumption leads to slightly larger contributions from short-range correlations than with further propagation to the second propagation step between discontinuous glaciers (equation (6)). Uncertainties are largely dominated by short- to long-range spatial correlations. **d**, Comparison of specific-volume changes per glacier with 1σ uncertainties. The mean of differences in estimates over all glaciers does not statistically differ from zero. **e, f**, Theoretical and empirical 1σ uncertainties, and their evolution with glacier size. The theoretical uncertainty is the mean of per-glacier uncertainties derived from spatially integrated variograms and the empirical uncertainty is the NMAD of the difference between high-resolution and GP estimates. **g, h**, Propagation of uncertainty sources to specific-mass changes for each RGI 6.0 region, and all glaciers with and without the Greenland Periphery and the Antarctic and Subantarctic, which are magnified in **h**. Uncertainties are largely dominated by the volume-to-mass conversion uncertainties globally, and by uncertainties in glacier outlines for regions with a relevant share of small glaciers.

Extended Data Fig. 6 Two decades of elevation change over various regions.

a–h, Elevation change of glaciers between 2000 and 2019 in Coropuna, Peru (**a**), Pamir Mountains (**b**), Iceland (**c**), Karakoram Mountains (**d**), European Alps (**e**), Southern Alps, New Zealand (**f**), West Greenland (note the rotated orientation of map) (**g**) and Svalbard (**h**). Except for Svalbard, glacier outlines displayed are from the RGI 6.0. In the background is shown a hillshade derived from several sources^{36,46,100}. In Svalbard, outlines have been updated to include the massive surges of Austfonna Basin 3^{38,39} in the northeast and Nathorstbreen in the southwest⁴⁰, indicated by blue arrows.

Extended Data Fig. 7 Global evolution of 5-year thinning rates.

a–d, Mean elevation change rates aggregated by tiles of $1^\circ \times 1^\circ$ for the periods 2000–2004 (**a**), 2005–2009 (**b**), 2010–2014 (**c**) and 2015–2019 (**d**). The tile area is inversely scaled to the squared 95% confidence interval of the mean elevation change in the tile, and tiles are coloured with mean

elevation change rates, on top of a world hillshade³⁶. The minimum tile area is 10% for a 95% confidence interval larger than 2 m yr^{-1} and tiles are displayed at full size for a 95% confidence interval smaller than 0.5 m yr^{-1} . Region labelling refers to that of Fig. 2. The acceleration of thinning brings the Karakoram anomaly to its apparent end.

Extended Data Table 1 Regional rates of glacier elevation and mass change from 2000 to 2019

[Full size table](#)

Extended Data Table 2 Regional data coverage of elevation time series from 2000 to 2019

[Full size table](#)

Extended Data Table 3 Regional rates of land- and marine-terminating glaciers in maritime regions

[Full size table](#)

Supplementary information

[Supplementary Information](#)

This file contains the Supplementary Methods, Supplementary Discussion, Supplementary Figures 1–9 and Supplementary Tables 1–3.

[Supplementary Table 4](#)

Comparison to IPCC SROCC Table 2A.1 with estimates from this study for periods 2006–2015 and 2000–2019 and recent regional studies (blue). An additional decimal is shown for mass balance rates in Gt yr^{-1} and mm SLE yr^{-1} . Regions are shown in SROCC ordering. SROCC estimates were combined using the most suitable studies (methods ‘xxx’ and ‘x’).

[Peer Review File](#)

Source data

[**Source Data Fig. 1**](#)

[**Source Data Fig. 2**](#)

[**Source Data Fig. 3**](#)

[**Source Data Fig. 4**](#)

Rights and permissions

[Reprints and Permissions](#)

About this article



Check for
updates

Cite this article

Hugonnet, R., McNabb, R., Berthier, E. *et al.* Accelerated global glacier mass loss in the early twenty-first century. *Nature* **592**, 726–731 (2021). <https://doi.org/10.1038/s41586-021-03436-z>

[Download citation](#)

- Received: 03 July 2020
- Accepted: 09 March 2021
- Published: 28 April 2021
- Issue Date: 29 April 2021
- DOI: <https://doi.org/10.1038/s41586-021-03436-z>

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03436-z>

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

- Article
- [Published: 28 April 2021](#)

Plume-driven recratonization of deep continental lithospheric mantle

- [Jingao Liu](#) [ORCID: orcid.org/0000-0002-8800-5898^{1,2}](#),
- [D. Graham Pearson](#) [ORCID: orcid.org/0000-0001-5974-4192²](#),
- [Lawrence Hongliang Wang³](#),
- [Kathy A. Mather⁴](#),
- [Bruce A. Kjarsgaard⁵](#),
- [Andrew J. Schaeffer](#) [ORCID: orcid.org/0000-0002-1286-5883⁶](#),
- [Gordon J. Irvine⁴](#),
- [Maya G. Kopylova⁷](#) &
- [John P. Armstrong⁸](#)

[Nature](#) volume 592, pages732–736(2021)[Cite this article](#)

- 1635 Accesses
- 21 Altmetric
- [Metrics details](#)

Subjects

- [Geochemistry](#)
- [Geodynamics](#)
- [Seismology](#)
- [Volcanology](#)

Abstract

Cratons are Earth's ancient continental land masses that remain stable for billions of years. The mantle roots of cratons are renowned as being long-lived, stable features of Earth's continents, but there is also evidence of their disruption in the recent^{1,2,3,4,5,6} and more distant^{7,8,9} past. Despite periods of lithospheric thinning during the Proterozoic and Phanerozoic eons, the lithosphere beneath many cratons seems to always 'heal', returning to a thickness of 150 to 200 kilometres^{10,11,12}; similar lithospheric thicknesses are thought to have existed since Archaean times^{3,13,14,15}. Although numerous studies have focused on the mechanism for lithospheric destruction^{2,5,13,16,17,18,19}, the mechanisms that re-cratonize the lithosphere beneath cratons and thus sustain them are not well understood. Here we study kimberlite-borne mantle xenoliths and seismology across a transect of the cratonic lithosphere of Arctic Canada, which includes a region affected by the Mackenzie plume event 1.27 billion years ago²⁰. We demonstrate the important role of plume upwelling in the destruction and re-cratonization of roughly 200-kilometre-thick cratonic lithospheric mantle in the northern portion of the Slave craton. Using numerical modelling, we show how new, buoyant melt residues produced by the Mackenzie plume event are captured in a region of thinned lithosphere between two thick cratonic blocks. Our results identify a process by which cratons heal and return to their original lithospheric thickness after substantial disruption of their roots. This process may be widespread in the history of cratons and may contribute to how cratonic mantle becomes a patchwork of mantle peridotites of different age and origin.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

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: Geological map and lithospheric thickness of the Slave craton and surrounding areas, and Os model ages of Slave-craton peridotites.

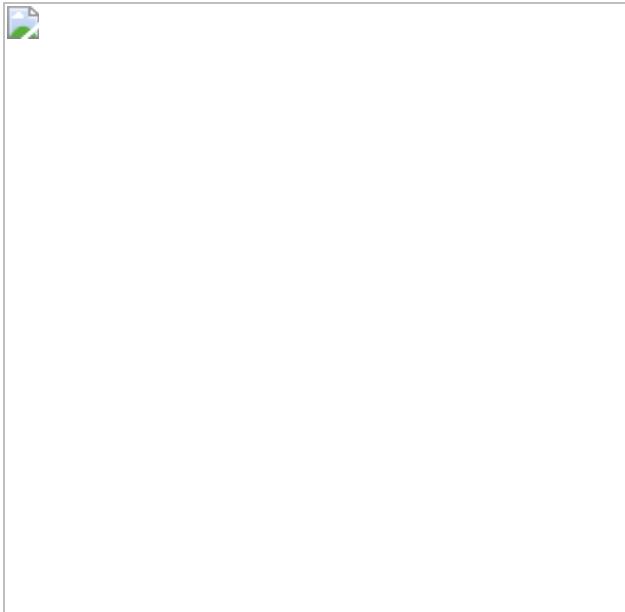


Fig. 2: Whole-rock Yb versus Al₂O₃ content for Slave peridotites.

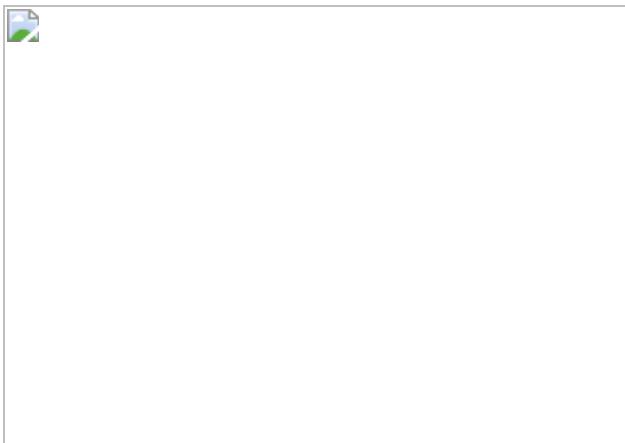
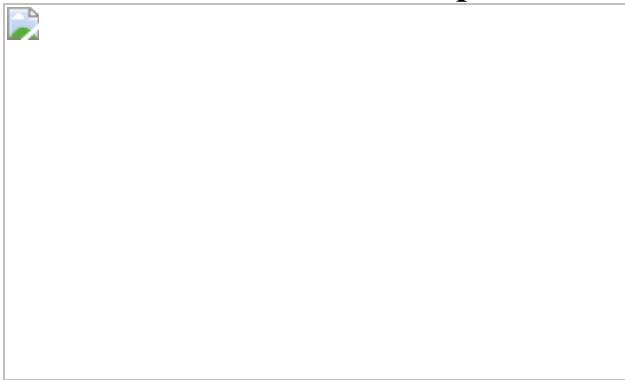


Fig. 3: Numerical modelling of plume residues filling a thin spot between thick cratonic lithospheric roots.



Data availability

The data that support the findings of this study are available at <https://doi.org/10.6084/m9.figshare.13789354>. [Source data](#) are provided with this paper.

Code availability

The code for geodynamic modelling is available on reasonable request to the corresponding author.

References

1. 1.

Menzies, M. A., Fan, W.-M. & Zhang, M. in *Magmatic Processes and Plate Tectonics* (eds Prichard, H. M. et al.) 71–81 (Geological Society, 1993).

2. 2.

Griffin, W. L., Zhang, A. D., O'Reilly, S. Y. & Ryan, C. G. in *Mantle Dynamics and Plate Interactions in East Asia* (eds Flower, M. et al.) 107–126 (American Geophysical Union, 1998).

3. 3.

Griffin, W. L. et al. The origin and evolution of Archean lithospheric mantle. *Precambr. Res.* **127**, 19–41 (2003).

[ADS](#) [CAS](#) [Google Scholar](#)

4. 4.

Carlson, R. W., Irving, A. J., Schulze, D. J. & Hearn, B. C. Timing of Precambrian melt depletion and Phanerozoic refertilization events in the lithospheric mantle of the Wyoming Craton and adjacent Central Plains Orogen. *Lithos* **77**, 453–472 (2004).

[ADS](#) [CAS](#) [Google Scholar](#)

5. 5.

Peace, A., Foulger, G., Schiffer, C. & McCaffrey, K. Evolution of Labrador Sea- Baffin Bay: plate or plume processes? *Geosci. Can.* **44**, 91–102 (2017).

[Google Scholar](#)

6. 6.

Kopylova, M. G., Tso, E., Ma, F., Liu, J. & Pearson, D. G. The Metasomatized Mantle beneath the North Atlantic Craton: insights from Peridotite Xenoliths of the Chidliak Kimberlite Province (NE Canada). *J. Petrol.* **60**, 1991–2024 (2019).

[ADS](#) [CAS](#) [Google Scholar](#)

7. 7.

James, D. E., Fouch, M. J., VanDecar, J. C., van der Lee, S. & Group, K. S. Tectospheric structure beneath southern Africa. *Geophys. Res. Lett.* **28**, 2485–2488 (2001).

[ADS](#) [Google Scholar](#)

8. 8.

Hanson, R. E. et al. Coeval large-scale magmatism in the Kalahari and Laurentian cratons during Rodinia assembly. *Science* **304**, 1126–1129 (2004).

[ADS](#) [CAS](#) [Google Scholar](#)

9. 9.

Ernst, R. E., Wingate, M. T. D., Buchan, K. L. & Li, Z. X. Global record of 1600–700Ma Large Igneous Provinces (LIPs): implications

for the reconstruction of the proposed Nuna (Columbia) and Rodinia supercontinents. *Precambr. Res.* **160**, 159–178 (2008).

[ADS](#) [CAS](#) [Google Scholar](#)

10. 10.

James, D. E. & Fouch, M. J. Formation and evolution of Archaean cratons: insights from southern Africa. *Geol. Soc. Lond. Spec. Publ.* **199**, 1–26 (2002).

[ADS](#) [Google Scholar](#)

11. 11.

Snyder, D. B., Humphreys, E. & Pearson, D. G. Construction and destruction of some North American cratons. *Tectonophysics* **694**, 464–485 (2017).

[ADS](#) [CAS](#) [Google Scholar](#)

12. 12.

Schaeffer, A. J. & Lebedev, S. Imaging the North American continent using waveform inversion of global and USArray data. *Earth Planet. Sci. Lett.* **402**, 26–41 (2014).

[ADS](#) [CAS](#) [Google Scholar](#)

13. 13.

Lee, C.-T. A., Luffi, P. & Chin, E. J. Building and destroying continental mantle. *Annu. Rev. Earth Planet. Sci.* **39**, 59–90 (2011).

[ADS](#) [CAS](#) [Google Scholar](#)

14. 14.

Aulbach, S. Craton nucleation and formation of thick lithospheric roots. *Lithos* **149**, 16–30 (2012).

[ADS](#) [CAS](#) [Google Scholar](#)

15. 15.

Perchuk, A. L., Gerya, T. V., Zakharov, V. S. & Griffin, W. L. Building cratonic keels in Precambrian plate tectonics. *Nature* **586**, 395–401 (2020).

[ADS](#) [CAS](#) [Google Scholar](#)

16. 16.

Foley, S. F. Rejuvenation and erosion of the cratonic lithosphere. *Nat. Geosci.* **1**, 503–510 (2008).

[ADS](#) [CAS](#) [Google Scholar](#)

17. 17.

Wang, H., van Hunen, J. & Pearson, D. G. The thinning of subcontinental lithosphere: the roles of plume impact and metasomatic weakening. *Geochem. Geophys. Geosyst.* **16**, 1156–1171 (2015).

[ADS](#) [Google Scholar](#)

18. 18.

Hu, J. et al. Modification of the Western Gondwana craton by plume–lithosphere interaction. *Nat. Geosci.* **11**, 203–210 (2018); publisher correction **11**, 544 (2018).

[ADS](#) [CAS](#) [Google Scholar](#)

19. 19.

Wu, F.-Y., Yang, J.-H., Xu, Y.-G., Wilde, S. A. & Walker, R. J. Destruction of the North China Craton in the Mesozoic. *Annu. Rev. Earth Planet. Sci.* **47**, 173–195 (2019).

[ADS](#) [CAS](#) [Google Scholar](#)

20. 20.

LeCheminant, A. N. & Heaman, L. M. Mackenzie igneous events, Canada: middle Proterozoic hotspot magmatism associated with ocean opening. *Earth Planet. Sci. Lett.* **96**, 38–48 (1989).

[ADS](#) [CAS](#) [Google Scholar](#)

21. 21.

Shirey, S. et al. in *Deep Mantle Carbon Evolution from the Diamond Record* (eds Orcutt, B. N. et. al) Ch. 5, 89–128 (Cambridge Univ. Press, 2019).

22. 22.

Pearson, D. G. & Wittig, N. in *Treatise of Geochemistry* 2nd edn, Vol. 3, (Holland, H. D. & Turekian, K. K.) Ch. 3.6, 255–292 (Elsevier, 2014).

23. 23.

Smit, K. V., Pearson, D. G., Stachel, T. & Seller, M. Peridotites from Attawapiskat, Canada: Mesoproterozoic reworking of Palaeoarchaean lithospheric mantle beneath the northern Superior Superterrane. *J. Petrol.* **55**, 1829–1863 (2014).

[ADS](#) [CAS](#) [Google Scholar](#)

24. 24.

Bleeker, W. Archaean tectonics: a review, with illustrations from the Slave craton. *Geol. Soc. Spec. Publ.* **199**, 151–181 (2002).

[ADS](#) [Google Scholar](#)

25. 25.

Westerlund, K. J. et al. A subduction wedge origin for Paleoarchean peridotitic diamonds and harzburgites from the Panda kimberlite, Slave craton: evidence from Re–Os isotope systematics. *Contrib. Mineral. Petrol.* **152**, 275 (2006).

[ADS](#) [MathSciNet](#) [CAS](#) [Google Scholar](#)

26. 26.

Hoffman, P. F. United plates of America, the birth of a craton: early Proterozoic assembly and growth of Laurentia. *Annu. Rev. Earth Planet. Sci.* **16**, 543–603 (1988).

[ADS](#) [Google Scholar](#)

27. 27.

Liu, J. et al. Diamondiferous Paleoproterozoic mantle roots beneath Arctic Canada: a study of mantle xenoliths from Parry Peninsula and Central Victoria Island. *Geochim. Cosmochim. Acta* **239**, 284–311 (2018).

[ADS](#) [CAS](#) [Google Scholar](#)

28. 28.

Schmidberger, S. S. et al. Lu–Hf, in-situ Sr and Pb isotope and trace element systematics for mantle eclogites from the Diavik diamond mine: evidence for Paleoproterozoic subduction beneath the Slave craton, Canada. *Earth Planet. Sci. Lett.* **254**, 55–68 (2007).

[ADS](#) [CAS](#) [Google Scholar](#)

29. 29.

Ootes, L. et al. Pyroxenitic magma conduits (ca. 1.86 Ga) in Wopmay orogen and Slave craton: petrogenetic constraints from whole rock and mineral chemistry. *Lithos* **354–355**, 105220 (2020).

[Google Scholar](#)

30. 30.

McKenzie, D. & Bickle, M. J. The volume and composition of melt generated by extension of the lithosphere. *J. Petrol.* **29**, 625–679 (1988).

[ADS](#) [CAS](#) [Google Scholar](#)

31. 31.

Ernst, R. E. & Baragar, W. R. A. Evidence from magnetic fabric for the flow pattern of magma in the Mackenzie giant radiating dyke swarm. *Nature* **356**, 511–513 (1992).

[ADS](#) [Google Scholar](#)

32. 32.

Mather, K. A., Pearson, D. G., McKenzie, D., Kjarsgaard, B. A. & Priestley, K. Constraints on the depth and thermal history of cratonic lithosphere from peridotite xenoliths, xenocrysts and seismology. *Lithos* **125**, 729–742 (2011).

[ADS](#) [CAS](#) [Google Scholar](#)

33. 33.

Rudnick, R. L. & Walker, R. J. Interpreting ages from Re-Os isotopes in peridotites. *Lithos* **112**, 1083–1095 (2009).

[ADS](#) [Google Scholar](#)

34. 34.

Aulbach, S. et al. Mantle formation and evolution, Slave Craton: constraints from HSE abundances and Re–Os isotope systematics of sulfide inclusions in mantle xenocrysts. *Chem. Geol.* **208**, 61–88 (2004).

[ADS](#) [CAS](#) [Google Scholar](#)

35. 35.

Pearson, D. G., Canil, D. & Shirey, S. B. in *Treatise of Geochemistry* 2nd edn, Vol. 3 (Holland, H. D. & Turekian, K. K.) Ch. 3.5, 169–253 (Elsevier, 2014).

36. 36.

Day, J. M. D., Pearson, D. G. & Hulbert, L. J. Rhenium-osmium isotope and platinum-group element constraints on the origin and evolution of the 1.27 Ga Muskox layered intrusion. *J. Petrol.* **49**, 1255–1295 (2008).

[ADS](#) [CAS](#) [Google Scholar](#)

37. 37.

Thompson, R. N. & Gibson, S. A. Subcontinental mantle plumes, hotspots and pre-existing thinspots. *J. Geol. Soc. Lond.* **148**, 973–977 (1991).

[Google Scholar](#)

38. 38.

Stachel, T., Viljoen, K. S., Brey, G. & Harris, J. W. Metasomatic processes in lherzolitic and harzburgitic domains of diamondiferous lithospheric mantle: REE in garnets from xenoliths and inclusions in diamonds. *Earth Planet. Sci. Lett.* **159**, 1–12 (1998).

[ADS](#) [CAS](#) [Google Scholar](#)

39. 39.

Wittig, N. et al. Origin of cratonic lithospheric mantle roots: a geochemical study of peridotites from the North Atlantic Craton, West Greenland. *Earth Planet. Sci. Lett.* **274**, 24–33 (2008).

[ADS](#) [CAS](#) [Google Scholar](#)

40. 40.

Sleep, N. H. Lateral flow and ponding of starting plume material. *J. Geophys. Res. Solid Earth* **102**, 10001–10012 (1997).

[Google Scholar](#)

41. 41.

Stachel, T. et al. The Victor Mine (Superior Craton, Canada): Neoproterozoic lherzolitic diamonds from a thermally-modified cratonic root. *Mineral. Petrol.* **112**, 325–336 (2018).

[ADS](#) [CAS](#) [Google Scholar](#)

42. 42.

Hanson, R. E., Martin, M. W., Bowring, S. A. & Munyanyiwa, H. U-Pb zircon age for the Umkondo dolerites, eastern Zimbabwe: 1.1 Ga large igneous province in southern Africa–East Antarctica and possible Rodinia correlations. *Geology* **26**, 1143–1146 (1998).

[ADS](#) [CAS](#) [Google Scholar](#)

43. 43.

Carlson, R. W., Pearson, D. G. & James, D. E. Physical, chemical, and chronological characteristics of continental mantle. *Rev. Geophys.* **43**, RG1001 (2005).

[ADS](#) [Google Scholar](#)

44. 44.

Liu, J. et al. Age and evolution of the deep continental root beneath the central Rae craton, northern Canada. *Precambr. Res.* **272**, 168–184 (2016).

[ADS](#) [CAS](#) [Google Scholar](#)

45. 45.

Griffin, W. L., O'Reilly, S. Y., Afonso, J. C. & Begg, G. C. The composition and evolution of lithospheric mantle: a re-evaluation and its tectonic implications. *J. Petrol.* **50**, 1185–1204 (2009).

[ADS](#) [CAS](#) [Google Scholar](#)

46. 46.

Kjarsgaard, B. A. in *Decade of North American Geology* Vol. P1 (eds Eckstrand, O. R. et al.) 557–566 (Geological Survey of Canada, 1995).

47. 47.

Steinberger, B. & Becker, T. W. A comparison of lithospheric thickness models. *Tectonophysics* **746**, 325–338 (2018).

[ADS](#) [Google Scholar](#)

48. 48.

Graham, I. et al. Exploration history and geology of the Diavik kimberlites, Lac de Gras, Northwest Territories, Canada. In *Proc. 7th International Kimberlite Conference* Vol. 1 (eds Gurney, J. J. et al.) 262–279 (Red Roof Design, 1999).

49. 49.

Heaman, L. M., Creaser, R. A. & Cookenboo, H. O. Extreme high-field-strength element enrichment in Jericho eclogite xenoliths: a

cryptic record of Paleoproterozoic subduction, partial melting and metasomatism beneath the Slave craton, Canada. *Geology* **30**, 507–510 (2002).

[ADS](#) [CAS](#) [Google Scholar](#)

50. 50.

Armstrong, J. P., Fitzgerald, C. E., Kjarsgaard, B. A., Heaman, L. & Tappe, S. Kimberlites of the Coronation Gulf field, northern Slave craton, Nunavut Canada. In *10th International Kimberlite Conference, Extended Abstracts* abstr. 10IKC-170 (2012); <https://doi.org/10.29173/ikc3723>.

51. 51.

Heaman, L. M., Kjarsgaard, B. A. & Creaser, R. A. The timing of kimberlite magmatism in North America: implications for global kimberlite genesis and diamond exploration. *Lithos* **71**, 153–184 (2003).

[ADS](#) [CAS](#) [Google Scholar](#)

52. 52.

Kopylova, M. G. & Russell, J. K. Chemical stratification of cratonic lithosphere: constraints from the Northern Slave craton, Canada. *Earth Planet. Sci. Lett.* **181**, 71–87 (2000).

[ADS](#) [CAS](#) [Google Scholar](#)

53. 53.

Boyd, F. R. & Mertzman, S. A. in *Magmatic Processes: Physicochemical Principles* (ed. Mysen, B. O.) 13–24 (Geochemical Society, 1987).

54. 54.

Ottley, C. J., Pearson, D. G. & Irvine, G. J. in *Plasma Source Mass Spectrometry: Applications and Emerging Technologies* (eds Holland, J. G. & Tanner, S. D.) 221–230 (Royal Society of Chemistry, 2003).

55. 55.

Harris, G. A. et al. Mantle composition, age and geotherm beneath the Darby kimberlite field, west central Rae Craton. *Mineral. Petrol.* **112**, 57–70 (2018).

[ADS](#) [CAS](#) [Google Scholar](#)

56. 56.

Pearson, D. G. & Woodland, S. J. Solvent extraction/anion exchange separation and determination of PGEs (Os, Ir, Pt, Pd, Ru) and Re-Os isotopes in geological samples by isotope dilution ICP-MS. *Chem. Geol.* **165**, 87–107 (2000).

[ADS](#) [CAS](#) [Google Scholar](#)

57. 57.

Liu, J. & Pearson, D. G. Rapid, precise and accurate Os isotope ratio measurements of nanogram to sub-nanogram amounts using multiple Faraday collectors and amplifiers equipped with $1012\ \Omega$ resistors by N-TIMS. *Chem. Geol.* **363**, 301–311 (2014).

[ADS](#) [CAS](#) [Google Scholar](#)

58. 58.

Luguet, A., Nowell, G. M. & Pearson, D. G. Os-184/Os-188 and Os-186/Os-188 measurements by negative thermal ionisation mass spectrometry (N-TIMS): effects of interfering element and mass fractionation corrections on data accuracy and precision. *Chem. Geol.* **248**, 342–362 (2008).

[ADS](#) [CAS](#) [Google Scholar](#)

59. 59.

Grütter, H. S. Pyroxene xenocryst geotherms: techniques and application. *Lithos* **112** (Suppl. 2) 1167–1178 (2009).

[ADS](#) [Google Scholar](#)

60. 60.

Lebedev, S., Nolet, G., Meier, T. & Van Der Hilst, R. D. Automated multimode inversion of surface and S waveforms. *Geophys. J. Int.* **162**, 951–964 (2005).

[ADS](#) [Google Scholar](#)

61. 61.

Lebedev, S. & Van Der Hilst, R. D. Global upper-mantle tomography with the automated multimode inversion of surface and S-wave forms. *Geophys. J. Int.* **173**, 505–518 (2008).

[ADS](#) [Google Scholar](#)

62. 62.

Schaeffer, A. J. & Lebedev, S. Global shear speed structure of the upper mantle and transition zone. *Geophys. J. Int.* **194**, 417–449 (2013).

[ADS](#) [Google Scholar](#)

63. 63.

Moresi, L. N. & Solomatov, V. S. Numerical investigation of 2D convection with extremely large viscosity variations. *Phys. Fluids* **7**, 2154–2162 (1995).

[ADS](#) [MATH](#) [Google Scholar](#)

64. 64.

Zhong, S., Zuber, M. T., Moresi, L. & Gurnis, M. Role of temperature-dependent viscosity and surface plates in spherical shell models of mantle convection. *J. Geophys. Res. Solid Earth* **105**, 11063–11082 (2000).

[Google Scholar](#)

65. 65.

van Hunen, J., Zhong, S., Shapiro, N. M. & Ritzwoller, M. H. New evidence for dislocation creep from 3-D geodynamic modeling of the Pacific upper mantle structure. *Earth Planet. Sci. Lett.* **238**, 146–155 (2005).

[ADS](#) [Google Scholar](#)

66. 66.

Christensen, U. R. & Yuen, D. A. Layered convection induced by phase transitions. *J. Geophys. Res. Solid Earth* **90**, 10291–10300 (1985).

[Google Scholar](#)

67. 67.

King, S. D. et al. A community benchmark for 2-D Cartesian compressible convection in the Earth's mantle. *Geophys. J. Int.* **180**, 73–87 (2010).

[ADS](#) [Google Scholar](#)

68. 68.

Katz, R. F., Spiegelman, M. & Langmuir, C. H. A new parameterization of hydrous mantle melting. *Geochem. Geophys. Geosyst.* **4**, 1073 (2003).

[ADS](#) [Google Scholar](#)

69. 69.

Schutt, D. L. & Lesher, C. E. Effects of melt depletion on the density and seismic velocity of garnet and spinel lherzolite. *J. Geophys. Res. Solid Earth* **111**, B05401 (2006).

[ADS](#) [Google Scholar](#)

70. 70.

Kennedy, C. S. & Kennedy, G. C. The equilibrium boundary between graphite and diamond. *J. Geophys. Res.* **81**, 2467–2470 (1976).

[ADS](#) [CAS](#) [Google Scholar](#)

71. 71.

Day, H. W. A revised diamond-graphite transition curve. *Am. Mineral.* **97**, 52–62 (2012).

[ADS](#) [CAS](#) [Google Scholar](#)

72. 72.

McDonough, W. F. & Sun, S. S. The Composition of the Earth. *Chem. Geol.* **120**, 223–253 (1995).

[ADS](#) [CAS](#) [Google Scholar](#)

73. 73.

Heaman, L. M., LeCheminant, A. N. & Rainbird, R. H. Nature and timing of Franklin igneous events, Canada: implications for a Late Proterozoic mantle plume and the break-up of Laurentia. *Earth Planet. Sci. Lett.* **109**, 117–131 (1992).

[ADS](#) [CAS](#) [Google Scholar](#)

74. 74.

Mackinder, A., Cousens, B. L., Ernst, R. E. & Chamberlain, K. R. Geochemical, isotopic, and U–Pb zircon study of the central and southern portions of the 780 Ma Gunbarrel Large Igneous Province in western Laurentia. *Can. J. Earth Sci.* **56**, 738–755 (2019).

[ADS](#) [CAS](#) [Google Scholar](#)

75. 75.

Wingate, M., Pirajno, F. & Morris, P. Warakurna large igneous province: a new Mesoproterozoic large igneous province in west-central Australia. *Geology* **32**, 105–108 (2004).

[ADS](#) [Google Scholar](#)

76. 76.

Fishwick, S. & Reading, A. M. Anomalous lithosphere beneath the Proterozoic of western and central Australia: a record of continental collision and intraplate deformation? *Precambr. Res.* **166**, 111–121 (2008).

[CAS](#) [Google Scholar](#)

77. 77.

Rudnick, R. & Nyblade, A. A. The thickness and heat production of Archean lithosphere, constraints from xenolith thermobarometry and surface heat flow. *Geochim. Soc. Spec. Publ.* **6**, 3–12 (1999).

[Google Scholar](#)

78. 78.

Pidgeon, R. T. & Cook, T. J. F. 1214 ± 5 Ma dyke from the Darling Range, southwestern Yilgarn Craton, Western Australia. *Aust. J. Earth Sci.* **50**, 769–773 (2003).

[ADS](#) [Google Scholar](#)

79. 79.

Wang, X.-C., Li, Z.-X., Li, J., Pisarevsky, S. A. & Wingate, M. T. D. Genesis of the 1.21 Ga Marnda Moorn large igneous province by plume–lithosphere interaction. *Precambr. Res.* **241**, 85–103 (2014).

[ADS](#) [CAS](#) [Google Scholar](#)

80. 80.

Geissler, W., Sodoudi, F. & Kind, R. Thickness of the Central and Eastern European lithosphere as seen by S receiver functions. *Geophys. J. Int.* **181**, 604–634 (2010).

[ADS](#) [Google Scholar](#)

81. 81.

Puchkov, V. et al. The ca. 1380 Ma Mashak igneous event of the Southern Urals. *Lithos* **174**, 109–124 (2013).

[ADS](#) [CAS](#) [Google Scholar](#)

82. 82.

Stark, J. C. et al. 1.39 Ga mafic dyke swarm in southwestern Yilgarn Craton marks Nuna to Rodinia transition in the West Australian Craton. *Precambr. Res.* **316**, 291–304 (2018).

[ADS](#) [CAS](#) [Google Scholar](#)

83. 83.

Youbi, N. et al. The 1750Ma Magmatic Event of the West African Craton (Anti-Atlas, Morocco). *Precambr. Res.* **236**, 106–123 (2013).

[ADS](#) [CAS](#) [Google Scholar](#)

84. 84.

Jessell, M., Begg, G. & Miller, M. The geophysical signatures of the West African Craton. *Precambr. Res.* **274**, 3–24 (2016).

[ADS](#) [CAS](#) [Google Scholar](#)

85. 85.

Feng, M., van der Lee, S. & Assumpção, M. Upper mantle structure of South America from joint inversion of waveforms and fundamental mode group velocities of Rayleigh waves. *J. Geophys. Res. Solid Earth* **112**, B04312 (2007).

[ADS](#) [Google Scholar](#)

86. 86.

Reis, N. J. et al. Avanavero mafic magmatism, a late Paleoproterozoic LIP in the Guiana Shield, Amazonian Craton: U–Pb ID-TIMS baddeleyite, geochemical and paleomagnetic evidence. *Lithos* **174**, 175–195 (2013).

[ADS](#) [CAS](#) [Google Scholar](#)

87. 87.

French, J. E., Heaman, L. M., Chacko, T. & Srivastava, R. K. 1891–1883Ma Southern Bastar–Cuddapah mafic igneous events, India: a newly recognized large igneous province. *Precambr. Res.* **160**, 308–322 (2008).

[ADS](#) [CAS](#) [Google Scholar](#)

88. 88.

Maurya, S. et al. Imaging the lithospheric structure beneath the Indian continent. *J. Geophys. Res. Solid Earth* **121**, 7450–7468 (2016).

[ADS](#) [Google Scholar](#)

89. 89.

Shellnutt, J. G., Hari, K. R., Liao, A. C. Y., Denyszyn, S. W. & Vishwakarma, N. A. 1.88 Ga giant radiating mafic dyke swarm across southern India and Western Australia. *Precambr. Res.* **308**, 58–74 (2018).

[ADS](#) [CAS](#) [Google Scholar](#)

[Download references](#)

Acknowledgements

We thank J. Li, Y. Sun, Y. Xu, Y. Luo, R. Cai, K. Zong, S. Woodland, G. Nowell, S. Jackson and C. Sarkar for help with analytical matters. This research was financially supported by the National Natural Science Foundation of China (41822301, 41790451, 41730214), China “1000 Youth Talents Program” and the 111 project (B18048) to J.L., by the Geomapping for Energy & Minerals program (Diamond project) of the Geological Survey of Canada (B.A.K. and D.G.P.) and the Canada Excellence Research Chairs program to D.G.P., and by the National Key R&D Program of China (2019YFA0708400, 2020YFA0714800 and 2019YFC0605403). This is GSC contribution number 20200737 and CUGB petrogeochemical contribution number PGC20150068 (RIG-no. 9).

Author information

Affiliations

1. State Key Laboratory of Geological Processes and Mineral Resources, China University of Geosciences (Beijing), Beijing, China

Jingao Liu

2. Department of Earth and Atmospheric Sciences, University of Alberta,
Edmonton, Alberta, Canada

Jingao Liu & D. Graham Pearson

3. Department of Environmental Analyses, Institute of Energy
Technology, Oslo, Norway

Lawrence Hongliang Wang

4. Department of Earth Sciences, Durham University, Durham, UK

Kathy A. Mather & Gordon J. Irvine

5. Geological Survey of Canada, Ottawa, Ontario, Canada

Bruce A. Kjarsgaard

6. Geological Survey of Canada, Sidney, British Columbia, Canada

Andrew J. Schaeffer

7. Department of Earth and Ocean Sciences, University of British
Columbia, Vancouver, British Columbia, Canada

Maya G. Kopylova

8. Lucara Diamond Corp, Vancouver, British Columbia, Canada

John P. Armstrong

Authors

1. Jingao Liu

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. D. Graham Pearson

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Lawrence Hongliang Wang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Kathy A. Mather

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Bruce A. Kjarsgaard

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Andrew J. Schaeffer

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Gordon J. Irvine

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Maya G. Kopylova

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. John P. Armstrong

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Contributions

D.G.P., J.L. and B.A.K. conceived and designed the study. J.L., D.G.P. and B.A.K. wrote the manuscript and contributed to data interpretation. J.L., K.A.M., G.J.I. and B.A.K. conducted the analyses and data reduction. L.H.W performed numerical modelling. A.J.S. conducted the seismic modelling. All authors contributed to interpreting the data and writing the paper.

Corresponding author

Correspondence to [Jingao Liu](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Stephen Foley 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 Characteristics of seismic velocity in the Slave craton and surrounding areas.](#)

a, Map of seismic velocity from SL2013NA¹² at 150 km depth, shown as the percentage variation from the reference velocity of 4.39 km s⁻¹ (colour

scale). Because this location is in the middle of the stable cratonic shield, the velocities are dominated by positive perturbations, indicative of the colder and stronger cratonic mantle lithosphere. **b**, Map of the maximum lateral gradient in seismic velocity from SL2013NA¹² at 150 km depth, shown as the percentage variation in velocity per 100 km laterally (colour scale). Because this location is in the middle of the stable cratonic shield, the gradients in velocity are lower than at the craton margins or in actively deforming regions. The highest gradients are associated with the northern boundary of the stable cratonic mantle lithosphere north of Victoria Island. A small but non-zero gradient in velocity is observed trending north–northwest from Jericho–Artemisia through western Victoria Island, largely coincident with the main strike of the Mackenzie dyke swarm. Artemisia lies in a region with shorter-scale lateral variations, which suggests that it could have a slightly more fertile composition (as shown by the slightly lower median whole-rock Mg content of 90.4 versus 91.0; Supplementary Table 1) than other Slave localities. White lines in **a** and **b** denote the boundaries of the Archaean and Palaeoproterozoic tectonic domains (as in Fig. 1); kimberlite pipes are indicated as in Fig. 1a (stars).

Extended Data Fig. 2 Box-and-whisker plot of anhydrous whole-rock Al₂O₃ content of Slave-craton peridotites.

Data are provided in Supplementary Table 1. Global on-craton peridotite data are from the literature²². PM, primitive mantle. [Source data](#)

Extended Data Fig. 3 Palaeogeotherms.

Calculated clinopyroxene thermobarometry pressure–temperature (PT) data from xenoliths and till concentrates of the Slave craton (Supplementary Table 6) are fitted to define a mantle geotherm (solid line, with shading representing the 2σ error envelope) using the FITPLOT (parameters are shown in Extended Data Table 2) method³². The left panel shows the cases for Diavik ($n = 65$) and Jericho ($n = 39$); the right panel shows those of Parry Peninsula ($n = 362$) and Central Victoria Island ($n = 196$) (data from ref. ²⁷). Despite no fresh pyroxene minerals to allow a pressure–temperature calculation in Artemisia, application of the Ni-in-garnet thermometer

(Supplementary Table 5) defines the sampled range of lithospheric mantle depths when extrapolated to the palaeogeotherms from other Slave localities in light of the diamondiferous feature in these kimberlites; therefore, the lithosphere thickness beneath Artemisia may be assumed to be similar to that beneath the nearby Jericho. The diamond and graphite transitions from ref. 70 and ref. 71, respectively, are plotted for reference.

[Source data](#)

[Extended Data Fig. 4 Chondrite-normalized whole-rock rare-earth element patterns for Slave peridotites.](#)

Data are provided in Supplementary Table 1. All the Slave peridotites show enrichments of light to middle rare-earth elements relative to heavy rare-earth elements. The data for CI chondrites are from ref. 72. [Source data](#)

[Extended Data Fig. 5 Whole-rock Yb versus Lu content for Slave peridotites.](#)

Curves show the trajectories of residual mantle after polybaric fractional melting, beginning at 2 GPa (spinel facies; orange), 3 GPa (shallow garnet facies; red) and 7 GPa (deep garnet facies; blue). The partial melting calculations followed those in ref. 39, supplementary material C. Unlike all other reported cratonic peridotites, the very high Lu and Yb (heavy rare-earth element) abundances of Artemisia peridotites clearly indicate initial melting in the presence of high modal garnet, that is, ultradeep plume melting. The previously published data from the North Atlantic, Rae and Kaapvaal cratons³⁹ are plotted for comparison. PM, primitive mantle.

[Source data](#)

[Extended Data Fig. 6 Modelling with different depletion buoyancy and strengthening at around 400 Myr.](#)

a–d, Temperature and depletion field at around 400 Myr for models with different depletion buoyancy and strengthening (**a**, B1; **b**, B2; **c**, B3; **d**, V1). The depletion field is a combination of the compositional field of the cratonic root and melt depletion. Melt-induced depletion buoyancy of the

plume residue has an important role and depends on α_d : the smaller α_d is, the more buoyant the plume residue becomes; $\Delta\eta$ is the viscosity strengthening factor due to melt depletion. **e, f**, The composition field of the cratonic root at around 400 Myr for models R (**e**) and B3 (**f**). The sequence shows that, within the parameters of the model, part of the cratonic root can be eroded by plume flow and then become involved in the formation of new lithosphere at the thin spot. Vertical and horizontal axes are in kilometres.

Extended Data Fig. 7 Growth scenarios of cratonic lithosphere when a plume hits.

a, Modelling an example with normal lithosphere at two side boundaries. **b**, Modelling an example with plume not under the thin spot between two cratons. The parameters for plume residue are $\alpha_d = -0.03\%$, $\Delta\eta = 3$ (**a**) and $\alpha_d = -0.02\%$, $\Delta\eta = 1$ (**b**). These two examples show that the proposed mechanism can lead to the growth of cratonic lithosphere in more general situations. Vertical and horizontal axes are in kilometres.

Extended Data Fig. 8 Thickness evolution of the thin spot between two cratons as defined by the 1,300 °C isotherm.

Without any viscosity strengthening, the lithosphere thickness may grow to approximately 200 km with high depletion buoyancy in models R and B3. With slight viscosity strengthening, intermediate buoyancy may also fill and stabilize the thin spot to around 200 km.

Extended Data Fig. 9 Relationship between whole-rock Yb content and $^{187}\text{Os}/^{188}\text{Os}$ calculated at the time of kimberlite eruption or TRD eruption.

Data are provided in Supplementary Table 1. PUM, primitive upper mantle. [Source data](#)

Extended Data Table 1 Summary of major Proterozoic large igneous provinces and their impact on cratons

[Full size table](#)

Extended Data Table 2 FITPLOT parameters used for constructing palaeogeotherms

[Full size table](#)

Supplementary information

[Peer Review File](#)

[Supplementary Video 1](#)

Temperature and viscosity evolution of the mantle hit by a plume. The video shows the evolution of temperature and viscosity field of the mantle through time when a hot plume hits the lithospheric thin-spot between two cratons.

[Supplementary Video 2](#)

Temperature and melt depletion evolution of the mantle hit by a plume. The video shows the evolution of temperature and melt depletion field of the mantle through time when a hot plume hits the lithospheric thin-spot between two cratons.

[Supplementary Table 1](#)

Whole rock major, trace, and Re-Os data of the peridotite xenoliths from Arctic Canada.

[Supplementary Table 2](#)

Whole-rock major element contents (wt%) for reference materials determined using XRF technique.

[Supplementary Table 3](#)

Whole-rock trace element contents (ppm) for reference materials.

Supplementary Table 4

Mineral major element contents (wt%) of the Slave peridotites measured in this study.

Supplementary Table 5

Garnet in situ trace element concentrations (ppb) determined by LA-ICP-MS.

Supplementary Table 6

P-T data used for constructing palaeogeotherms.

Source data

Source Data Fig. 1

Source Data Fig. 2

Source Data Extended Data Fig. 2

Source Data Extended Data Fig. 3

Source Data Extended Data Fig. 4

Source Data Extended Data Fig. 5

Source Data Extended Data Fig. 9

Rights and permissions

Reprints and Permissions

About this article



Cite this article

Liu, J., Pearson, D.G., Wang, L.H. *et al.* Plume-driven retronization of deep continental lithospheric mantle. *Nature* **592**, 732–736 (2021).
<https://doi.org/10.1038/s41586-021-03395-5>

[Download citation](#)

- Received: 31 March 2020
- Accepted: 25 February 2021
- Published: 28 April 2021
- Issue Date: 29 April 2021
- DOI: <https://doi.org/10.1038/s41586-021-03395-5>

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

Associated Content

Ancient continental blocks soldered from below

- Stephen Foley
- Craig O'Neill

Nature News & Views 28 Apr 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03395-5>

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

Towards complete and error-free genome assemblies of all vertebrate species
[Download PDF](#)

- Article
- Open Access
- [Published: 28 April 2021](#)

Towards complete and error-free genome assemblies of all vertebrate species

- [Arang Rhee](#) ORCID: [orcid.org/0000-0002-9809-8127^{1 na1}](https://orcid.org/0000-0002-9809-8127),
- [Shane A. McCarthy^{2,3 na1}](#),
- [Olivier Fedrigo](#) ORCID: [orcid.org/0000-0002-6450-7551^{4 na1}](https://orcid.org/0000-0002-6450-7551),
- [Joana Damas⁵](#),
- [Giulio Formenti](#) ORCID: [orcid.org/0000-0002-7554-5991^{4,6}](https://orcid.org/0000-0002-7554-5991),
- [Sergey Koren](#) ORCID: [orcid.org/0000-0002-1472-8962¹](https://orcid.org/0000-0002-1472-8962),
- [Marcela Uliano-Silva](#) ORCID: [orcid.org/0000-0001-6723-4715^{7,8}](https://orcid.org/0000-0001-6723-4715),
- [William Chow³](#),
- [Arkarachai Fungtammasan⁹](#),
- [Juwan Kim¹⁰](#),
- [Chul Lee¹⁰](#),
- [Byung June Ko¹¹](#),
- [Mark Chaisson¹²](#),
- [Gregory L. Gedman⁶](#),
- [Lindsey J. Cantin⁶](#),
- [Francoise Thibaud-Nissen](#) ORCID: [orcid.org/0000-0003-4957-7807¹³](https://orcid.org/0000-0003-4957-7807),
- [Leanne Haggerty¹⁴](#),
- [Iliana Bista^{2,3}](#),
- [Michelle Smith³](#),
- [Bettina Haase⁴](#),
- [Jacquelyn Mountcastle](#) ORCID: [orcid.org/0000-0003-1078-4905⁴](https://orcid.org/0000-0003-1078-4905),
- [Sylke Winkler](#) ORCID: [orcid.org/0000-0002-0915-3316^{15,16}](https://orcid.org/0000-0002-0915-3316),
- [Sadie Paez^{4,6}](#),
- [Jason Howard¹⁷](#),

- [Sonja C. Vernes](#) ORCID: orcid.org/0000-0003-0305-4584^{18,19,20},
- [Tanya M. Lama](#)²¹,
- [Frank Grutzner](#) ORCID: orcid.org/0000-0002-3088-7314²²,
- [Wesley C. Warren](#)²³,
- [Christopher N. Balakrishnan](#) ORCID: orcid.org/0000-0002-0788-0659²⁴,
- [Dave Burt](#) ORCID: orcid.org/0000-0002-9991-1028²⁵,
- [Julia M. George](#) ORCID: orcid.org/0000-0001-6194-6914²⁶,
- [Matthew T. Biegler](#)⁶,
- [David Iorns](#)²⁷,
- [Andrew Digby](#) ORCID: orcid.org/0000-0002-1870-8811²⁸,
- [Daryl Eason](#)²⁸,
- [Bruce Robertson](#) ORCID: orcid.org/0000-0002-5348-2731²⁹,
- [Taylor Edwards](#) ORCID: orcid.org/0000-0002-7235-6175³⁰,
- [Mark Wilkinson](#)³¹,
- [George Turner](#)³²,
- [Axel Meyer](#) ORCID: orcid.org/0000-0002-0888-8193³³,
- [Andreas F. Kautt](#) ORCID: orcid.org/0000-0001-7792-0735^{33,34},
- [Paolo Franchini](#) ORCID: orcid.org/0000-0002-8184-1463³³,
- [H. William Detrich III](#) ORCID: orcid.org/0000-0002-0783-4505³⁵,
- [Hannes Svardal](#)^{36,37},
- [Maximilian Wagner](#)³⁸,
- [Gavin J. P. Naylor](#)³⁹,
- [Martin Pippel](#) ORCID: orcid.org/0000-0002-8134-5929^{15,40},
- [Milan Malinsky](#) ORCID: orcid.org/0000-0002-1462-6317^{3,41},
- [Mark Mooney](#)⁴²,
- [Maria Simbirsky](#)⁹,
- [Brett T. Hannigan](#)⁹,
- [Trevor Pesout](#)⁴³,
- [Marlys Houck](#)⁴⁴,
- [Ann Misuraca](#)⁴⁴,
- [Sarah B. Kingan](#) ORCID: orcid.org/0000-0002-4900-0189⁴⁵,
- [Richard Hall](#) ORCID: orcid.org/0000-0001-6490-8227⁴⁵,
- [Zev Kronenberg](#)⁴⁵,
- [Ivan Sović](#)^{45,46},
- [Christopher Dunn](#) ORCID: orcid.org/0000-0002-0601-3254⁴⁵,
- [Zemin Ning](#)³,
- [Alex Hastie](#)⁴⁷,
- [Joyce Lee](#) ORCID: orcid.org/0000-0002-3492-1102⁴⁷,
- [Siddarth Selvaraj](#)⁴⁸,

- [Richard E. Green](#)^{43,49},
- [Nicholas H. Putnam](#) ORCID: orcid.org/0000-0002-1315-782X⁵⁰,
- [Ivo Gut](#) ORCID: orcid.org/0000-0001-7219-632X^{51,52},
- [Jay Ghurye](#)^{49,53},
- [Erik Garrison](#)⁴³,
- [Ying Sims](#)³,
- [Joanna Collins](#)³,
- [Sarah Pelan](#)³,
- [James Torrance](#)³,
- [Alan Tracey](#) ORCID: orcid.org/0000-0002-4805-9058³,
- [Jonathan Wood](#)³,
- [Robel E. Dagnew](#)¹²,
- [Dengfeng Guan](#) ORCID: orcid.org/0000-0002-6376-3940^{2,54},
- [Sarah E. London](#) ORCID: orcid.org/0000-0002-7839-2644⁵⁵,
- [David F. Clayton](#) ORCID: orcid.org/0000-0002-6395-3488⁵⁶,
- [Claudio V. Mello](#) ORCID: orcid.org/0000-0002-9826-8421⁵⁷,
- [Samantha R. Friedrich](#) ORCID: orcid.org/0000-0003-0570-6080⁵⁷,
- [Peter V. Lovell](#)⁵⁷,
- [Ekaterina Osipova](#) ORCID: orcid.org/0000-0002-6769-7223^{15,40,58},
- [Farooq O. Al-Ajli](#) ORCID: orcid.org/0000-0002-4692-7106^{59,60,61},
- [Simona Secomandi](#) ORCID: orcid.org/0000-0001-8597-6034⁶²,
- [Heebal Kim](#) ORCID: orcid.org/0000-0003-3064-1303^{10,11,63},
- [Constantina Theofanopoulou](#) ORCID: orcid.org/0000-0003-2014-7563⁶,
- [Michael Hiller](#)^{64,65,66},
- [Yang Zhou](#) ORCID: orcid.org/0000-0003-1247-5049⁶⁷,
- [Robert S. Harris](#)⁶⁸,
- [Kateryna D. Makova](#)^{68,69,70},
- [Paul Medvedev](#)^{69,70,71,72},
- [Jinna Hoffman](#)¹³,
- [Patrick Masterson](#)¹³,
- [Karen Clark](#)¹³,
- [Fergal Martin](#) ORCID: orcid.org/0000-0002-1672-050X¹⁴,
- [Kevin Howe](#)¹⁴,
- [Paul Flicek](#) ORCID: orcid.org/0000-0002-3897-7955¹⁴,
- [Brian P. Walenz](#) ORCID: orcid.org/0000-0001-8431-1428¹,
- [Woori Kwak](#)^{63,73},
- [Hiram Clawson](#)⁴³,
- [Mark Diekhans](#) ORCID: orcid.org/0000-0002-0430-0989⁴³,
- [Luis Nassar](#)⁴³,

- [Benedict Paten](#) ORCID: [orcid.org/0000-0001-8863-3539⁴³](https://orcid.org/0000-0001-8863-3539),
- [Robert H. S. Kraus](#)^{33,74},
- [Andrew J. Crawford](#) ORCID: [orcid.org/0000-0003-3153-6898⁷⁵](https://orcid.org/0000-0003-3153-6898),
- [M. Thomas P. Gilbert](#) ORCID: [orcid.org/0000-0002-5805-7195^{76,77}](https://orcid.org/0000-0002-5805-7195),
- [Guojie Zhang](#) ORCID: [orcid.org/0000-0001-6860-1521^{78,79,80,81}](https://orcid.org/0000-0001-6860-1521),
- [Byrappa Venkatesh](#) ORCID: [orcid.org/0000-0003-3620-0277⁸²](https://orcid.org/0000-0003-3620-0277),
- [Robert W. Murphy](#)⁸³,
- [Klaus-Peter Koepfli](#)⁸⁴,
- [Beth Shapiro](#) ORCID: [orcid.org/0000-0002-2733-7776^{85,86}](https://orcid.org/0000-0002-2733-7776),
- [Warren E. Johnson](#) ORCID: [orcid.org/0000-0002-5954-186X^{84,87,88}](https://orcid.org/0000-0002-5954-186X),
- [Federica Di Palma](#)⁸⁹,
- [Tomas Marques-Bonet](#) ORCID: [orcid.org/0000-0002-5597-3075^{90,91,92,93}](https://orcid.org/0000-0002-5597-3075),
- [Emma C. Teeling](#) ORCID: [orcid.org/0000-0002-3309-1346⁹⁴](https://orcid.org/0000-0002-3309-1346),
- [Tandy Warnow](#)⁹⁵,
- [Jennifer Marshall Graves](#)⁹⁶,
- [Oliver A. Ryder](#) ORCID: [orcid.org/0000-0003-2427-763X^{44,97}](https://orcid.org/0000-0003-2427-763X),
- [David Haussler](#) ORCID: [orcid.org/0000-0003-1533-4575^{43,85}](https://orcid.org/0000-0003-1533-4575),
- [Stephen J. O'Brien](#)^{98,99},
- [Jonas Korlach](#) ORCID: [orcid.org/0000-0003-3047-4250⁴⁵](https://orcid.org/0000-0003-3047-4250),
- [Harris A. Lewin](#) ORCID: [orcid.org/0000-0002-1043-7287^{5,100,101}](https://orcid.org/0000-0002-1043-7287),
- [Kerstin Howe](#) ORCID: [orcid.org/0000-0003-2237-513X^{3 na2}](https://orcid.org/0000-0003-2237-513X),
- [Eugene W. Myers](#) ORCID: [orcid.org/0000-0002-6580-7839^{15,40,102 na2}](https://orcid.org/0000-0002-6580-7839),
- [Richard Durbin](#) ORCID: [orcid.org/0000-0002-9130-1006^{2,3 na2}](https://orcid.org/0000-0002-9130-1006),
- [Adam M. Phillippy](#) ORCID: [orcid.org/0000-0003-2983-8934^{1 na2}](https://orcid.org/0000-0003-2983-8934) &
- [Erich D. Jarvis](#) ORCID: [orcid.org/0000-0001-8931-5049^{4,6,86 na2}](https://orcid.org/0000-0001-8931-5049)

[Nature](#) volume 592, pages737–746(2021)Cite this article

- 39k Accesses
- 1 Citations
- 433 Altmetric
- [Metrics details](#)

Subjects

- [Evolutionary genetics](#)
- [Genome assembly algorithms](#)
- [Molecular evolution](#)

- [Research data](#)

Abstract

High-quality and complete reference genome assemblies are fundamental for the application of genomics to biology, disease, and biodiversity conservation. However, such assemblies are available for only a few non-microbial species^{1,2,3,4}. To address this issue, the international Genome 10K (G10K) consortium^{5,6} has worked over a five-year period to evaluate and develop cost-effective methods for assembling highly accurate and nearly complete reference genomes. Here we present lessons learned from generating assemblies for 16 species that represent six major vertebrate lineages. We confirm that long-read sequencing technologies are essential for maximizing genome quality, and that unresolved complex repeats and haplotype heterozygosity are major sources of assembly error when not handled correctly. Our assemblies correct substantial errors, add missing sequence in some of the best historical reference genomes, and reveal biological discoveries. These include the identification of many false gene duplications, increases in gene sizes, chromosome rearrangements that are specific to lineages, a repeated independent chromosome breakpoint in bat genomes, and a canonical GC-rich pattern in protein-coding genes and their regulatory regions. Adopting these lessons, we have embarked on the Vertebrate Genomes Project (VGP), an international effort to generate high-quality, complete reference genomes for all of the roughly 70,000 extant vertebrate species and to help to enable a new era of discovery across the life sciences.

[Download PDF](#)

Main

Chromosome-level reference genomes underpin the study of functional, comparative, and population genomics within and across species. The first high-quality genome assemblies of human¹ and other model species (for example, *Caenorhabditis elegans*², mouse³, and zebrafish⁴) were put together using 500–1,000-base pair (bp) Sanger sequencing reads of thousands of hierarchically organized clones with 200–300-kilobase (kb) inserts, and chromosome genetic maps. This approach required tremendous manual effort, software engineering, and cost, in decade-long projects. Whole-genome shotgun approaches simplified the logistics (for example, in human⁷ and *Drosophila*⁸), and later next-generation sequencing with shorter (30–150-bp) sequencing reads and short insert sizes (for example, 1 kb) ushered in more affordable and scalable genome sequencing⁹. However, the shorter reads resulted in lower-quality assemblies, fragmented into thousands of pieces, where many genes were missing, truncated, or incorrectly assembled, resulting in annotation and other errors¹⁰. Such

errors can require months of manual effort to correct individual genes and years to correct an entire assembly. Genomic heterozygosity posed additional problems, because homologous haplotypes in a diploid or polyploid genome are forced together into a single consensus by standard assemblers, sometimes creating false gene duplications^{11,12,13,14}.

To address these problems, the G10K consortium^{5,6} initiated the Vertebrate Genomes Project (VGP; <https://vertebratogenomesproject.org>) with the ultimate aim of producing at least one high-quality, near error-free and gapless, chromosome-level, haplotype-phased, and annotated reference genome assembly for each of the 71,657 extant named vertebrate species and using these genomes to address fundamental questions in biology, disease, and biodiversity conservation. Towards this end, having learned the lessons of having too many variables that make conclusions more difficult to reach in the G10K from the G10K Assemblathon 2 effort¹⁵, we first evaluated multiple genome sequencing and assembly approaches extensively on one species, the Anna's hummingbird (*Calypte anna*). We then deployed the best-performing method across sixteen species representing six major vertebrate classes, with a wide diversity of genomic characteristics. Drawing on the principles learned, we improved these methods further, discovered parameters and approaches that work better for species with different genomic characteristics, and made biological discoveries that had not been possible with the previous assemblies.

Complete, accurate assemblies require long reads

We chose a female Anna's hummingbird because it has a relatively small genome (about 1 Gb), is heterogametic (has both Z and W sex chromosomes), and has an annotated reference of the same individual built from short reads¹⁶. We obtained 12 new sequencing data types, including both short and long reads (80 bp to 100 kb), and long-range linking information (40 kb to more than 100 Mb), generated using eight technologies (Supplementary Table 1). We benchmarked all technologies and assembly algorithms (Supplementary Table 2) in isolation and in many combinations (Supplementary Table 3). To our knowledge, this was the first systematic analysis of many sequence technologies, assembly algorithms, and assembly parameters applied on the same individual. We found that primary contiguous sequences (contigs) (pseudo-haplotype; Supplementary Note 1) assembled from Pacific Biosciences continuous long reads (CLR) or Oxford Nanopore long reads (ONT) were approximately 30- to 300-fold longer than those assembled from Illumina short reads (SR), regardless of data type combination or assembly algorithm used (Fig. 1a, Supplementary Table 3). The highest contig NG50s for short-read-only assemblies were about 0.025 to 0.169 Mb, whereas for long reads they were about 4.6 to 7.66 Mb (Fig. 1a); contig NG50 is an assembly metric based on a weighted median of the lengths of its gapless sequences relative to the estimated genome size. After fixing a

function in the PacBio FALCON software¹⁷ that caused artificial breaks in contigs between stretches of highly homozygous and heterozygous haplotype sequences (Supplementary Note 1, Supplementary Table 2), contig NG50 nearly tripled to 12.77 Mb (Fig. 1a). These findings are consistent with theoretical predictions¹⁸ and demonstrate that, given current sequencing technology and assembly algorithms, it is not possible to achieve high contig continuity with short reads alone, as it is typically impossible to bridge through repeats that are longer than the read length.

Fig. 1: Comparative analyses of Anna's hummingbird genome assemblies with various data types.

 figure1

a, Contig NG50 values of the primary pseudo-haplotype. **b**, Scaffold NG50 values. **c**, Number of joins (gaps). **d**, Number of mis-join errors compared with the curated assembly. The curated assembly has no remaining conflicts with the raw data and thus no known mis-joins. *Same as CLR + linked + Opt. + Hi-C, but with contigs generated with an updated FALCON¹⁷ version and earlier Hi-C Salsa version (v2.0 versus v2.2; Supplementary Table 2) for less aggressive contig joining. **e, f**, Hi-C interaction heat maps before and after manual curation, which identified 34 chromosomes. Grid lines indicate scaffold boundaries. Red arrow, example mis-join that was corrected during curation. **g**, Karyotype of the identified chromosomes ($n = 36 + ZW$), consistent with previous findings⁷⁰. **h**, Correlation between estimated chromosome sizes (in Mb) based on karyotype images in **g** and assembled scaffolds in Supplementary Table 4 (bCalAna1) on a log–log scale. v1.0, VGP assembly v1.0 pipeline; linked, 10X Genomics linked reads; Hi-C, Hi-C proximity ligation; 1D, 2D, Oxford Nanopore long reads; NRGene, NRGene paired-end Illumina reads; SR, paired-end Illumina short reads.

[Full size image](#)

Iterative assembly pipeline

Scaffolds generated with all three scaffolding technologies (that is, 10X Genomics linked reads (10XG), Bionano optical maps (Opt.), and Arima Genomics, Dovetail Genomics, or Phase Genomics Hi-C) were approximately 50% to 150% longer than those generated using one or two technologies, regardless of whether we started with short- or long-read-based contigs (Fig. 1b, Extended Data Fig. 1a, Supplementary Table 3). These findings include improvements we made to each approach (Supplementary Note 1, Supplementary Tables 4, 5, Supplementary Fig. 1). Despite similar scaffold continuity, the short-read-only assemblies had from about 18,000 to about 70,000 gaps, whereas the long-read assemblies had substantially fewer (about 400 to about 4,000) gaps (Fig. 1c). Many gaps in the short-read assemblies were in repeat or GC-rich regions. Considering the curated version of this assembly to be more accurate, we also identified roughly 5,000 to 8,000 mis-joins in short-read-based assemblies, whereas long-read-based assemblies had only from 20 to around 700 mis-joins (Fig. 1d). These mis-joins included chimeric joins and inversions. After we curated this assembly for contamination, assembly errors, and Hi-C-based chromosome assignments (Fig. 1e, f), the final hummingbird assembly had 33 scaffolds that closely matched the chromosome karyotype in number (33 of 36 autosomes plus sex chromosomes) and estimated sizes (approximately 2 to 200 Mb; Fig. 1g, h), with only 1 to 30 gaps per autosome (bCalAnn1 in Supplementary Table 6). Of the five autosomes with only one gap each, three (chromosomes 14, 15, and 19) had complete spanning support by at least two technologies (reliable blocks, Extended Data Fig. 1c; bCalAnn1 in Supplementary Table 6), indicating that the chromosome contigs were nearly complete. However, they were missing long arrays of vertebrate telomere repeats within 1 kb of their ends (Extended Data Fig. 1c; bCalAnn1 in Supplementary Tables 6, 7).

Assembly pipeline across vertebrate diversity

Using the formula that gave the highest-quality hummingbird genome, we built an iterative VGP assembly pipeline (v1.0) with haplotype-separated CLR contigs, followed by scaffolding with linked reads, optical maps, and Hi-C, and then gap filling, base call polishing, and finally manual curation (Extended Data Figs. 2a, 3a). We systematically tested our pipeline on 15 additional species spanning all major vertebrate classes: mammals, birds, non-avian reptiles, amphibians, teleost fishes, and a cartilaginous fish (Supplementary Tables 8, 9, Supplementary Note 2). For the zebra finch, we used DNA from the same male as was used to generate the previous reference genome¹⁹, and included a female trio for benchmarking haplotype completeness, where sequenced reads from the parents were used to bin parental haplotype reads from the offspring before assembly²⁰ (Extended Data Figs. 2a, 3b).

We set initial minimum assembly metric goals of: 1 Mb contig NG50; 10 Mb scaffold NG50; assigning 90% of the sequence to chromosomes, structurally validated by at least two independent lines of evidence; Q40 average base quality; and haplotypes assembled as completely and correctly as possible. When these metrics were achieved, most genes were assembled with gapless exon and intron structures¹¹, and fewer than 3% had frame-shift base errors identified in annotation. Q40 is the mathematical inflection point at which genes go from usually containing an error to usually not²¹. Of the curated assemblies (Supplementary Table 10, Supplementary Note 2), 16 of 17 achieved the desired continuity metrics (Extended Data Table 1). Scaffold NG50 was significantly correlated with genome size (Fig. 2a), suggesting that larger genomes tend to have larger chromosomes. On average, 98.3% of the assembled bases had reliable block NG50s ranging from 2.3 to 40.2 Mb; collapsed repeat bases²² with abnormally high CLR read coverage (more than 3 s.d.) ranged from 0.7 to 31.4 Mb per Gb; and the completeness of the genome assemblies ranged from 87.2 to 98.1%, with less than 4.9% falsely duplicated regions, consistent with the false duplication rate we found for the conserved BUSCO vertebrate gene set (Extended Data Table 1, Supplementary Tables 11, 12).

Fig. 2: Impact of repeats and heterozygosity on assembly quality.

 figure2

- a**, Correlation between scaffold NG50 and genome size of the curated assemblies. **b**, Nonlinear correlation between contig NG50 and repeat content, before and after curation. **c**, Correlation between number of gaps per Gb assembled and repeat content. **d**, Correlation between primary assembly size relative to estimated genome size (y axis) and genome heterozygosity (x axis), before and after purging of false duplications. Assembly sizes above 100% indicate the presence of false duplications and those below 100% indicate collapsed repeats. **e**, **f**, Correlations between genome duplication rate using *k*-mers²³ (**e**) and conserved BUSCO vertebrate gene set (**f**), and

genome heterozygosity before and after purging of false duplications. **g, h**, As in **e, f**, but with whole-genome repeat content before and after purging of false duplications. Genome size, heterozygosity, and repeat content were estimated from 31-mer counts using GenomeScope⁷¹, except for the channel bull blenny, as the estimates were unreliable (see [Methods](#)). Repeat content was measured by modelling the *k*-mer multiplicity from sequencing reads. Sequence duplication rates were estimated with Merqury²³ using 21-mers. * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$, of the correlation coefficient: P values and adjusted r^2 from *F*-statistics. $n = 17$ assemblies of 16 species.

[Full size image](#)

Repeats markedly affect continuity

For assemblies generated using our automated pipeline (Extended Data Fig. [3a](#)) before manual curation, all but 2 (the thorny skate and channel bull blenny) of the 17 assemblies exceeded the desired continuity metrics (Supplementary Table [13](#)). In searching for an explanation of these results, we found that contig NG50 decreased exponentially with increasing repeat content, with the thorny skate having the highest repeat content (Fig. [2b](#), Supplementary Table [13](#)). Consequently, after scaffolding and gap filling, we observed a significant positive correlation between repeat content and number of gaps (Fig. [2c](#)). The kākāpō parrot, which had 15% repeat content, had about 325 gaps per Gb, including 2 of 26 chromosomes with no gaps (chromosomes 16 and 18) and no evidence of collapses or low support, suggesting that the chromosomal contigs were complete (bStrHab1 in Supplementary Table [6](#)). By contrast, the thorny skate, with 54% repeat content, had about 1,400 gaps per Gb (Extended Data Table [1](#)); none of its 49 chromosomal-level scaffolds contained fewer than eight gaps, and all had some regions that contained collapses or low support (sAmbRad1 in Supplementary Table [6](#)). Even after curation and other modifications to increase assembly quality (Supplementary Note [2](#)), the number of collapses, their total size, missing bases, and the number of genes in the collapses all correlated with repeat content (Extended Data Fig. [4a–d](#)). The average collapsed length, however, correlated with average CLR read lengths (10–35 kb; Extended Data Fig. [4e](#)). There were no correlations between the number of collapsed bases and heterozygosity or genome size (Extended Data Fig. [4f,g](#)). Depending on species, 77.4 to 99.2% of the collapsed regions consisted of unresolved segmental duplications (Extended Data Fig. [4h](#)). The remainder were high-copy repeats, mostly of previously unknown types (Extended Data Fig. [4i](#)), and of known types such as satellite arrays, simple repeats, long terminal repeats (LTRs), and short and long interspersed nuclear elements (SINEs and LINEs), depending on species (Extended Data Fig. [4j](#)). We found that repeat masking before generating contigs prevented some repeats from making it into the final assembly (Supplementary Note [3](#)). All of the above findings quantitatively demonstrate the

effect that repeat content has on the ability to produce highly continuous and complete assemblies.

Detection and removal of false duplications

During curation, we discovered that one of the most common assembly errors was the introduction of false duplications, which can be misinterpreted as exon, whole-gene, or large segmental duplications. We observed two types of false duplication: 1) heterotype duplications, which occurred in regions of increased sequence divergence between paternal and maternal haplotypes, where separate haplotype contigs were incorrectly placed in the primary assembly (Extended Data Fig. 5a); and 2) homotype duplications, which occurred near contig boundaries or under-collapsed sequences caused by sequencing errors (Extended Data Fig. 5b). False heterotype duplications appeared to occur with higher heterozygosity. For example, during curation of the female zebra finch genome, we found an approximately 1-Mb falsely duplicated heterozygous sequence (Extended Data Fig. 6a). This zebra finch individual had the highest heterozygosity (1.6%) relative to all other genomes (0.1–1.1%). Homotype duplications often occurred at contig boundaries, and were approximately the same length as the sequence reads (Extended Data Fig. 6b,c). We identified and removed false duplications during curation using read coverage, self-, transcript-, optical map- and Hi-C-alignments, and *k*-mer profiles (Extended Data Fig. 6, Supplementary Fig. 2).

Before we purged false duplications, the primary assembly genome size correlated positively with estimated percentage heterozygosity; more heterozygous genomes tended to have assembly sizes bigger than the estimated haploid genome size (Fig. 2d). Similarly, the extra duplication rate in the primary assembly, measured using *k*-mers²³ or conserved vertebrate BUSCO genes²⁴, varied from 0.3% to 30% and trended towards correlation with heterozygosity (Fig. 2e,f, Supplementary Table 13). Apparent false gene duplication rates correlated more strongly with the overall repeat rate in the assemblies (Fig. 2g,h). To remove these false duplications automatically, we initially used Purge_Haplotigs¹³, which removed retained falsely duplicated contigs that were not scaffolded (Extended Data Fig. 5; VGP v1.0–1.5). Later, we developed Purge_Dups¹⁴ to remove both falsely retained contigs and end-to-end duplicated contigs within scaffolds (Extended Data Fig. 5; VGP v1.6), which reduced the amount of manual curation. After we applied these tools, the primary assembly sizes and the *k*-mer and BUSCO gene duplication rates were all reduced, and their correlations with heterozygosity and repeat content were also reduced or eliminated (Fig. 2d-h). These findings indicate that it is essential to properly phase haplotypes and to obtain high consensus sequence accuracy in order to prevent false duplications and associated biologically false conclusions.

Curation is needed for a high-quality reference

Each automated scaffolding method introduced tens to thousands of unique joins and breaks in contigs or scaffolds (Supplementary Table 14). Depending on species, the first scaffolding step with linked reads introduced about 50–900 joins between CLR-generated contigs. Optical maps introduced a further roughly 30–3,500 joins, followed by Hi-C with about 30–700 more joins, and each identified up to several dozen joins that were inconsistent with the previous scaffolding step. Manual curation resulted in an additional 7,262 total interventions for 19 genome assemblies or 236 interventions per Gb of sequence (Supplementary Table 15). When a genome assembly was available for the same or a closely related species, it was used to confirm putative chromosomal breakpoints or rearrangements (Supplementary Table 15). These interventions indicate that even with current state-of-the-art assembly algorithms, curation is essential for completing high-quality reference assemblies and for providing iterative feedback to improve assembly algorithms. A further description of our curation approach and analyses of VGP genomes are presented elsewhere²⁵.

Hi-C scaffolding and cytological mapping

Most large assembled scaffolds of each species spanned entire chromosomes, as shown by the relatively clean Hi-C heat map plots across each scaffold after curation (Extended Data Fig. 7), near perfect correlation between chromosomal scaffold length and karyotypically determined chromosome length (Fig. 1h), and the presence of telomeric repeat motifs on some scaffold ends (Supplementary Table 7). In our VGP zebra finch assembly, all inferred chromosomes were consistent with previously identified linkage groups in the Sanger-based reference, except for chromosomes 1 and 1B (Extended Data Fig. 8a). Their join in the VGP assembly was supported by both single CLR reads and optical maps through the junction. We also corrected nine inversion errors and filled in large gaps at some chromosome ends. In the platypus, we identified 18 structural differences in 13 scaffolds between the VGP assembly and the previous Sanger-based reference anchored to chromosomes using fluorescence in situ hybridization (FISH) physical mapping (Extended Data Fig. 8b, Supplementary Table 16). Of these 18, all were supported with Hi-C, and seven were also supported by both CLR and optical maps in the VGP assembly. Our platypus assembly also filled in many large (approximately 1–30 Mb) gaps and corrected many inversion errors (Extended Data Fig. 8b). Furthermore, we identified seven additional chromosomes (chromosomes 30–36) in the zebra finch, and eight (chromosomes 8, 9, 14, 15, 17, 19, 21, and X4; Extended Data Fig. 8a,b) in the platypus^{26,27}. Relative to the VGP assembly, the earlier short-read Anna's hummingbird assembly was highly fragmented (Extended Data Fig. 8c), despite being scaffolded with seven different Illumina libraries spanning a wide range of insert sizes (0.2–20 kb). The previous climbing

perch assembled chromosomes were even more fragmented and also had large gaps of missing sequence (Extended Data Fig. 8d). On average, $97\% \pm 3\%$ (s.d.) of the assembled bases were assigned to chromosomes (Extended Data Table 1), compared with 76% and 32% in the prior zebra finch and platypus references, respectively. We believe the comparable or higher accuracy of Hi-C relative to genetic linkage or FISH physical mapping is due to the higher sampling rate of Hi-C pairs across the genome. Nonetheless, visual karyotyping is useful for complementary validation of chromosome count and structure²⁸.

Trios help to resolve haplotypes

We were able to assemble the trio-based female zebra finch contigs into separate maternal and paternal chromosome-level scaffolds (Extended Data Fig. 9a) using our VGP trio pipeline (Extended Data Fig. 3b). Compared to the non-trio assembly of the same individual, the trio version had seven- to eightfold fewer false duplications (*k*-mer and BUSCO dups in Supplementary Tables 11, 12), well-preserved haplotype-specific variants (*k*-mer precision/recall 99.99/97.08%), and higher base call accuracy, exceeding Q43 for both haplotypes (Extended Data Table 1). The trio-based assembly was the only assembly with nearly perfect (99.99%) separation of maternal and paternal haplotypes, determined using *k*-mers specific to each²³. We identified haplotype-specific structural variants, including inversions of 4.5 to 12.5 Mb on chromosomes 5, 11, and 13 that were not readily identifiable in the non-trio version (Extended Data Fig. 10a–e). Moving forward, the VGP is prioritising the collection of mother–father–offspring trios where possible, or single parent–offspring duos, to assist with diploid assembly and phasing, as well as the development of improved methods for the assembly of diploid genomes in the absence of parental genomic data, as described in another study²⁹.

Effects of polishing on accuracy

Despite their increased continuity and structural accuracy, CLR-based assemblies required at least two rounds of short-read consensus polishing to reach 99.99% base-level accuracy (one error per 10 kb, Phred³⁰ Q40; Supplementary Table 5). Before polishing, the per-base accuracy was Q30–35 (calculated using *k*-mers). The most common errors were short indels from inaccurate consensus calling during CLR contig formation, which resulted in amino acid frameshift errors. Using our combined approach of long-read and short-read polishing applied on both primary and alternate haplotype sequences together, we polished from 82% to 99.7% of the primary and about 91.3% of the alternate assembly (Supplementary Table 17). Of the remaining unpolished sequence, one haplotype was sometimes reconstructed at substantially lower quality, because most reads aligned to the higher quality haplotype (Extended

Data Fig. 11a). False duplications had similar effects, where the duplicated sequence acted as an attractor during the read mapping. Haplotypes in the more homozygous regions tended to be collapsed by FALCON-Unzip¹⁷. All such cases recruited reads from both haplotypes and thereby caused switch errors, which we confirmed in the trio-based assembly and fixed when excluding read pairs from the other haplotype during polishing (Extended Data Fig. 11b). These findings indicate that both sequence read accuracy and careful haplotype separation are important for producing accurate assemblies.

Sex chromosomes and mitochondrial genomes

Sex chromosomes have been notoriously difficult to assemble, owing to their greater divergence relative to autosomes and high repeat content³¹. We successfully assembled both sex chromosomes (Z, W) for all three avian species, the first W chromosome (to our knowledge) for vocal learning birds (Extended Data Figs. 7, 9b), the X and/or Y chromosome in placental mammals (Canada lynx and two bat species), the X chromosome in the thorny skate, and for the first time, to our knowledge, all ten sex chromosomes (5X and 5Y) in the platypus²⁶ (Extended Data Fig. 9c). The completeness and continuity of the zebra finch Z and W chromosomes were further improved by the trio-based assembly (Extended Data Fig. 9b). However, the sex chromosome assemblies were still more fragmented than the autosomes, probably owing to their lower sequencing depth and high repeat content.

Mitochondrial (MT) genomes, which are expected to be 11–28 kb in size³², were initially found in only six assemblies (Supplementary Table 18). The MT-derived raw reads were present, but they failed to assemble, in part because of minimum read-length cutoffs for the starting contig assembly. Furthermore, if the MT genome was not present during nuclear genome polishing, the raw MT reads were attracted to nuclear MT sequences (NuMTs), incorrectly converting them to the full organelle MT sequence (Extended Data Fig. 11c). To address these issues, we developed a reference-guided MT pipeline and included the MT genome during polishing³³ (Extended Data Fig. 3c; VGP v1.6). With these improvements, we reliably assembled 16 of 17 MT genomes (Supplementary Table 18) and discovered 2 kb of an 83-bp repeat expansion within the control region in the kākāpō (Extended Data Fig. 9d), and *Nad1* and *trnL2* gene duplications in the climbing perch (Extended Data Fig. 9e). These duplications were verified using single-molecule CLR reads that spanned the duplication junctions or even the entire MT genome. Their absence in previous MT references^{34,35} is likely to result from the inability of Sanger or short reads to correctly resolve large duplications. More details on the MT-VGP pipeline and new biological discoveries are reported elsewhere³³.

Improvements to read alignment and annotation

Compared to previous Sanger (zebra finch and platypus) and Illumina (Anna's hummingbird and climbing perch) assemblies, we added about 42–176 Mb of missing sequence and placed 68.5 Mb (zebra finch) to 1.8 Gb (platypus) of previously unplaced sequence within chromosomes. We corrected about 7,800–64,000 mis-joins, and closed 55,177–193,137 gaps per genome (Supplementary Table 19). Consistent with these improvements, both transcriptome RNA sequencing (RNA-seq) data (Fig. 3a) and genome assay for transposase-accessible chromatin using sequencing (ATAC-seq) data (Fig. 3b) aligned with about 5 to 10% greater mapability to our new VGP assemblies compared with the previous assemblies. The NCBI RefSeq and EBI Ensembl annotations revealed: 5,434 to 14,073 more protein-coding transcripts per species, with 94.1 to 97.8% fully supported (Fig. 3c, Supplementary Table 20); only about 100 to 300 partially assembled coding genes, compared with about 1,600 to 5,600 (Fig. 3d); more orthologous coding genes shared with human; and fewer transcripts that required corrections to compensate for premature stop codons or frame-shift indel errors (Extended Data Table 2). The total number of genes annotated went down in the VGP assemblies (Extended Data Table 2), partly because there were fewer false duplications (Supplementary Table 19). Supporting these results, the VGP assemblies had 0 to 13% higher k -mer completeness (95% mean \pm 3.5% s.d. versus $88 \pm 4.3\%$; Extended Data Table 2, Supplementary Table 19; $P = 0.0047$, $n = 4$ prior and 17 VGP assemblies, unpaired t -test).

Fig. 3: Improvements to alignments and annotations in VGP assemblies relative to prior references.

 **figure3**

a, b, Average percentage of RNA-seq transcriptome samples (**a**; $n = 44$, mean \pm s.e.m.) and ATAC-seq genome reads (**b**; $n = 12$) that align to the previous and VGP zebra finch assemblies. Unique reads mapped to only one location in the assembly. Total is the sum of unique and multi-mapped reads. P values are from paired t -test. **c, d**, Total number of coding sequence (CDS) transcripts (full bar) and portion fully supported (inner bar) (**c**) and the number of RefSeq coding genes annotated as partial (**d**) in the previous and VGP assemblies using the same input data. **e–h**, Examples of assembly and associated annotation errors in previous reference assemblies corrected in the new VGP assemblies. See main text for descriptions. **i**, Gene synteny around the *VTR2C* receptor in the platypus shows completely missing genes (*NUDT16*), truncated and duplicated *ARHGAP4*, and many gaps in the earlier Sanger-based assembly compared with the filled in and expanded gene lengths in the new VGP assembly. Assembly accessions are in Supplementary Table [19](#).

[Full size image](#)

An example of a whole-gene heterotype false duplication in the RefSeq annotation of the previous zebra finch reference¹⁹ is the BUSCO gene *SPC25*³⁶, for which each haplotype was correctly placed in the VGP primary and alternate assemblies (Fig. 3e). The *GABRG2* receptor, which shows specialized expression in vocal learning circuits³⁷, had a partial tandem duplication of four of its ten exons, resulting in annotated partial false tandem gene duplications (*GABRG2* and *GABRG2*-like; Fig. 3f). The vitellogenin-2 (*VTG2*) gene, a component of egg yolk in all egg-laying species³⁸, was distributed across 14 contigs in 3 different scaffolds in the previous platypus assembly (Fig. 3g). Two of these scaffolds received two corresponding *VTG2*-like gene annotations, and the third was included as false duplicated intron in *CAPN-13* (red), together causing false amino acid sequences in five exons (blue). The BUSCO *YIPF6* gene, which is associated with inflammatory bowel disease³⁹, was split between two different scaffolds and is thus presumed to be a gene loss in the earlier climbing perch assembly⁴⁰ (Fig. 3h). Each of these genes is now present on long VGP contigs, within validated blocks, with no gaps and no false gene gains or losses (Supplementary Table 21).

Going beyond individual genes, a ten-gene synteny window surrounding the vasotocin receptor 2C gene (*VTR2C*; also known as *AVPR2*), which is involved in blood pressure homeostasis and brain function^{41,42}, was split into 34 contigs on four scaffolds, one of which contained a false haplotype duplication of *ARHGAP4* in the previous platypus assembly⁴³ (Fig. 3i). In our VGP assembly, all eleven genes were in one 37-Mb-long contig within the approximately 50 Mb chromosome 6 scaffold. Furthermore, eight of the eleven genes were remarkably increased in size owing to the addition of previously unknown missing sequences. This chromosomal region was more GC-rich (54%) than the entire chromosome 6 (46%). Thousands of such false gains and losses in previous reference assemblies have been corrected in our VGP assemblies (more details in refs. 27,44), demonstrating that assembly quality has a critical effect on subsequent annotations and functional genomics.

GC-rich regulatory regions of coding genes

We tested whether the higher-quality VGP assemblies enabled new biological discoveries. Notably, beginning about 1.5 kb upstream of protein-coding genes, in 100-bp blocks, there was a steady increase from about 6–20% to about 30–55% of genes having missing sequence in previous references (Fig. 4a); similarly high proportions of genes were missing their subsequent 5' untranslated regions (UTRs) and first exons. This fluctuation in missing sequence was directly proportional to GC content (Fig. 4a). We therefore studied the GC content pattern across all protein-

coding genes in all 16 new VGP assemblies and found a genome-wide signature: a rapid rise in GC content in the roughly 1.5 kb before the transcription start site, in the 5' UTR, and in the first exon, followed by a steady decrease in subsequent exons and returning to near intergenic background levels in the 3' UTR and about 1.5 kb after the transcription termination site (Fig. 4b). The introns had lower GC content, closer to the intergenic background. The intergenic GC content was stable within 30 kb on either side of each gene (Fig. 4b). Mammals, birds, and reptiles had the highest increase (around 20%) in GC content near the start site, followed by the amphibian and skate with medium levels (around 10%). Teleost fishes showed an initial decrease, followed by weaker increase (about 5%) from an already lower GC content (Fig. 4b). Given that the skate represents the sister branch to all other vertebrate lineages sequenced, these findings suggest that teleosts lost at least 5% GC content genome-wide, while maintaining most of the GC content pattern in protein-coding genes. Although it is known that promoter regions can be CpG rich, and GC content can vary between exons and introns^{45,46}, such a systematic pattern, the lineage-specific differences within vertebrates, and the magnitude of these differences had not been previously described, to our knowledge.

Fig. 4: VGP assemblies reveal GC content patterns in protein-coding genes.

 **figure4**

a, Average GC content ($n = 14,000\text{--}18,000$ annotated coding genes; Extended Data Table 2) in VGP assemblies (black) and the percentage of genes with missing sequence in the earlier references (red) based on a Cactus alignment, in 100-bp blocks, 2 kb on either side of all protein-coding genes (left and right), and for UTRs, exons, and introns (middle). **b**, Average GC content (mean \pm s.d. for lineages with more than one species) of the six major vertebrate lineages sequenced, for 30 kb upstream and downstream (in 100-bp blocks, log scale; left and right) and of the UTR, exons, and introns (middle). **c, d**, Left, specialized expression (arrows) shown by in situ hybridization of *DRD1B* in the zebra finch striatum (c) and *ER81* in the arcopallium (d), from Jarvis et al.⁴⁷; the cerebellum was removed from the *ER81* image. Right, ATAC-seq profiles in the GC-rich promoter regions of these genes, showing each

gene's GC content (red is high), the ATAC-seq peaks in striatum (purple) or arcopallium (yellow) neurons, and portions of missing sequence (black) in the previous reference assembly (grey).

[Full size image](#)

We tested whether the newly assembled GC-rich promoter regions contained novel regulatory sequences. Analysing the zebra finch brain, we found that genes with upregulated expression specific to the striatum (for example, *DRD1B*, which encodes a dopamine receptor) had ATAC-seq peaks in the GC-rich promoter and 5' UTR region in striatal neurons, but not in arcopallium neurons (Fig. 4c); conversely, genes (for example, the *ER81* transcription factor) with upregulated expression in the arcopallium (mammalian cortex layer 5 equivalent⁴⁷) had ATAC-seq peaks in the GC-rich region in arcopallium neurons but not striatal neurons (Fig. 4d). These GC-rich regions were missing in the earlier assembly. In addition, the missing region in *DRD1B* led to a false annotation as a two-exon gene⁴⁸, whereas the VGP assembly revealed a single-exon gene (Fig. 4c). These GC-rich promoter regions are candidates for driving cell-type-specific expression. These findings demonstrate the importance of using sequencing chemistry that reads through GC-rich regions, like the CLR method. The earlier hummingbird genome assembly was generated using Illumina TruSeq3 chemistry¹⁶, which was designed to read through GC-rich regions, and yet about 55% of the genes were missing the 100-bp GC-rich region before the start site (Fig. 4a). Another paper contains additional findings on missing regions²⁷.

Chromosomal evolution

We next investigated whether we could gain new insights into chromosome evolution among vertebrates. Given the more than 430 million years (Myr) of evolutionary divergence among the species sampled here, it was difficult to generate whole genome-to-genome alignments across all species. Thus, we focused our initial analyses on 1,147 highly conserved BUSCO vertebrate genes that are shared among our assemblies of all 16 species and the human reference (GRCh38). Human chromosomes mapped with greater orthology to 3.7 ± 1.3 (s.d.) chromosomes on average in other mammals, compared to 5.6 ± 2.2 in amphibians and 9.6 ± 3.3 in teleost fishes (Fig. 5a, Supplementary Table 22). The skate chromosome arrangement was more conserved with tetrapods, mapping to 2.9 ± 1.4 chromosomes on average, compared to 4.8 ± 2.5 in teleost fishes. These findings indicate that, along with a reduction in GC content, the teleost lineage has experienced more massive chromosome rearrangements since divergence from their most recent common ancestor with tetrapods, consistent with a proposed higher rearrangement rate in teleosts⁴⁹.

Fig. 5: Chromosome evolution among bats and other vertebrates.

 figure5

a, Chromosome synteny maps across the species sequenced based on BUSCO gene alignments. Chromosome sizes (bar lengths) are normalized to genome size, to make visualization easier. Genes (lines) are coloured according to the human chromosome to which they belong; those on human chromosome 6 are highlighted in blue and other chromosomes are in lighter shades. The cladogram is from the TimeTree database⁷². **b**, Phylogenetic relationship of the mammalian species sequenced and their inferred chromosome EBR rates (breaks per Myr) on different branches. Red, higher rates than average (0.84); blue, lower than average. **c**, Summary of alignment, gene organization, and functional gene status surrounding a bat interchromosomal EBR involving the homologue of human chromosome 6. End of scaffold (S) or chromosome (Chr.) means that the breakpoint is located at a chromosome arm end; middle means that it is located within a scaffold or chromosome. Scale is relevant for human Chr. 6 only. Actual gene sizes in the non-human species may differ and were drawn to match the annotated human gene sizes for simplicity.

[Full size image](#)

To determine the precise locations of chromosome rearrangements between species, we focused on a shorter evolutionary distance of around 180 Myr among mammals, and added four additional bat species described in our Bat1K study⁵⁰, the human genome reference⁵¹ (GRCh38.p12), and a recently upgraded long-read chicken reference⁵² (galGal6a) as an outgroup. Pairwise whole-genome alignments to the human reference defined homologous synteny blocks and evolutionary breakpoint regions (EBRs) among the species. We found that breakpoint rates (EBRs per Myr) tripled among bats soon after the last mass extinction event (about 66 million years ago (Mya)), a time of rapid bat superfamily divergences⁵³ (about 60 Mya; Fig. 5b). Some rearrangements affected genes. For example, a 1.3-Mb inversion in greater horseshoe bat chromosome 28 (homologous to 29.5 Mb of human chromosome 15; Extended Data Fig. 12a) disrupted *STARD5*, a gene involved in cholesterol homeostasis in liver cells⁵⁴. The rearrangement separated exons 1–5 from exon 6, and disrupted splicing of the transcripts (Extended Data Fig. 12b). Another example was

an EBR that involved fission of an ancestral bat chromosome homologue of human chromosome 6 (boreoeutherian mammal chromosome 5⁵⁵) and was later reused among the different bat lineages in rearrangements that involved the ancestral homologues of human chromosomes 1, 2 and 6 (Fig. 5c, Extended Data Fig. 12c). We also noted a fission in this region in the mouse, rat, and dog genomes⁵⁵. On the basis of the conserved gene order in human and Canada lynx, we inferred that the boreoeutherian ancestral mammal locus corresponding to human 6p22.1 contained 12 genes, including four *ZSCAN* and two *ZKSCAN* transcription factors, and two *GPX* enzyme genes, all associated with sequentially increasing independent gene losses in bats (Fig. 5d). For example, the greater horseshoe bat lost only *ZSCAN12* and *GPX6* to pseudogenization, whereas Kuhl's pipistrelle lost all 12 genes. *ZSCAN* and *ZKSCAN* are involved in cell differentiation, migration and invasion, proliferation, apoptosis, and innate immunity⁵⁶. We speculate that loss of *ZSCAN12* in all six bats could contribute to their immune tolerance to pathogens⁵⁰.

Other biological findings using these VGP assemblies are published elsewhere, and include: 1) more accurate synteny across species, leading to a better understanding of the evolution of and thus a universal nomenclature for the vasotocin (also known as vasopressin) and oxytocin ligand and receptor gene families⁵⁷; 2) greater understanding of the evolution of the carbohydrate 6-O sulfotransferase gene family, which encodes enzymes that modify secreted carbohydrates⁵⁸; 3) the first Bat1K study⁵⁰, which generated a genome-scale phylogeny that better resolves the relationships between bats and other mammals, and which identified changes in bat genes that are involved in immunity and life span, including genes that are relevant to the COVID-19 pandemic⁵⁹; 4) deleterious mutations that have been purged from the last surviving isolated and inbred population of the critically endangered kākāpō⁶⁰; and 5) more complete resolution of the evolution of the complex sex chromosomes in platypus and echidna²⁶. These discoveries were not possible with the previous reference assemblies, and we expect many future discoveries to follow.

Proposed assembly quality metrics

Drawing on the lessons learned from this work, we propose that assembly quality should be summarized using 14 metrics under 6 categories (Table 1; full details in Supplementary Note 4). We summarize the most critical and commonly used metrics using the simple notation $x.y.P.Q.C$, where: $x = \log_{10}[\text{contig NG50}]$, $y = \log_{10}[\text{scaffold NG50}]$, $P = \log_{10}[\text{haplotype phase block NG50}]$, $Q = \text{QV base accuracy}$, and $C = \text{percentage of the assembly assigned to chromosomes}$ (Table 1). Our current minimum VGP standard, for example, is 6.7.P5.Q40.C90. This revises our prior notation^{50,61,62}, which reported log-scaled continuity measured in 'kilobases' rather than 'bases'. The thresholds we chose were based on empirical and quantitative

observations between what is achievable currently and what is aspirational, and the question the assemblies are meant to answer. For example, the short-read paired-end library-based assemblies of the B10K Phase 1 genomes in 2014¹⁶ and the 10XG linked-read assembly of the Anna's hummingbird presented here would be categorized as a 4.5.P7.Q50 assembly, with low continuity but high base accuracy (Table 1). Such a genome would be suitable for use in phylogenomics⁶³ and for population-scale SNP surveys⁶⁴. If, instead, a genome is to be used to study chromosomal evolution, then the VGP-2016 minimum metric 6.7.P5.Q40.C95, with high structural and base accuracies and more than 95% assigned to chromosomes (Table 1), would be necessary. If having GC-rich promoter regions and complete 5' exons in most genes is essential, then long-read approaches that sequence through these regions are necessary. 'Finished' quality (Table 1) is obviously the ideal assembly result, but this level of quality is currently routine only for bacterial and non-vertebrate model organisms with smaller genome sizes that lack large centromeric satellite arrays^{65,66,67} and for organelle genomes, as presented here³³. The possibility of achieving complete, telomere-to-telomere assemblies of vertebrate and other eukaryotic species is foreseeable, given some assembled avian and bat chromosomes with zero gaps in this study, and the recent complete assembly of two human chromosomes^{68,69}.

Table 1 Proposed standards and metrics for defining genome assembly quality
[Full size table](#)

The Vertebrate Genomes Project

Building on this initial set of assembled genomes and the lessons learned, we propose to expand the VGP to deeper taxonomic phases, beginning with phase 1: representatives of approximately 260 vertebrate orders, defined here as lineages separated by 50 million or more years of divergence from each other. Phase 2 will encompass species that represent all approximately 1,000 vertebrate families; phase 3, all roughly 10,000 genera; and phase 4, nearly all 71,657 extant named vertebrate species (Supplementary Note 5, Supplementary Fig. 3). To accomplish such a project within 10 years, we will need to scale up to completing 125 genomes per week, without sacrificing quality. This includes sample permitting, high molecular weight DNA extractions, sequencing, meta-data tracking, and computational infrastructure. We will take advantage of continuing improvements in genome sequencing technology, assembly, and annotation, including advances in PacBio HiFi reads, Oxford Nanopore reads, and replacements for 10XG reads (Supplementary Note 6), while addressing specific scientific questions at increasing levels of phylogenetic refinement. Genomic technology advances quickly, but we believe the principles of our pipeline and the lessons learned will be applicable to future efforts. Areas in which improvement is needed include more accurate and complete haplotype phasing, base-

call accuracy, and resolution of long repetitive regions such as telomeres, centromeres, and sex chromosomes. The VGP is working towards these goals and making all data, protocols, and pipelines openly available (Supplementary Notes 5, 7).

Despite remaining imperfections, our reference genomes are the most complete and highest quality to date for each species sequenced, to our knowledge. When we began to generate genomes beyond the Anna's hummingbird in 2017, only eight vertebrate species in GenBank had genomes that met our target continuity metrics, and none were haplotype phased (Supplementary Table 23). The VGP pipeline introduced here has now been used to complete assemblies of more than 130 species of similar or higher quality (Supplementary Note 5; BioProject PRJNA489243). We encourage the scientific community to use and evaluate the assemblies and associated raw data, and to provide feedback towards improving all processes for complete and error-free assembled genomes of all species.

Methods

Genome assembly naming

For each completed assembly of an individual, we gave that assembly an abbreviated name with the following rules: Lineage/GenusSpecies/Individual#.Assembly#. The first letter, in lowercase, identifies the particular lineage: m, mammals; b, birds; r, reptiles; a, amphibians; f, teleost fish; and s, sharks and other cartilaginous fishes. The next three letters (first in caps) identify the species scientific genus name; the next three letters (first in caps) identifies the specific species name. In the last position is the genome identifier, where integers (1, 2, 3, ...) represent different individuals of the same species, and decimals (1.1, 1.2, 1.3, ...) represent different assemblies of the same individual. For example, the first submission of the curated Anna's hummingbird (*Calyptra anna*) assembly is bCalAnn1.1, and an updated assembly for the same individual is bCalAnn1.2. When the abbreviated lineage or genus and species names for two or more species were identical, we replaced the subsequent letters (fourth, fifth and so on) of the genus or species name until they could be differentiated. We have created abbreviated names for all 71,657 vertebrate species (<http://vgpdb.snu.ac.kr/splist/>; <https://id.tol.sanger.ac.uk/>).

Sample collection

The production of high-quality genome assemblies required us to obtain high-quality cells or tissue that would yield high-molecular-weight (HMW) DNA for long-read sequencing technologies (CLR and ONT) and optical mapping (Bionano). Therefore, we obtained fresh-frozen samples of various tissues (Supplementary Table 8). All samples were obtained according to approved protocols of the respective animal care

and use committees or permits obtained by the respective persons and institutions listed in Supplementary Table 8. Additional details of the samples are on their respective BioSample pages (<https://www.ncbi.nlm.nih.gov/biosample>; accession numbers in Supplementary Table 8). All tissue types tested yielded a sufficient quantity and quality of DNA for sequencing and assembly, but we found that blood worked best for species that have nucleated red blood cells (that is, bird and reptiles), and spleen or cultured cells worked best for mammals, as of to date. Analysis of different tissue types will be presented elsewhere (in preparation).

Isolation of high-molecular-weight DNA

Agarose plug DNA isolation

For tissue, HMW DNA was extracted using the Bionano animal tissue DNA isolation fibrous tissue protocol (cat no. RE-013-10; document number 30071), according to the manufacturer's guidelines. A total of 25–30 mg was fixed in 2% formaldehyde and homogenized using the Qiagen TissueRuptor or manual tissue disruption. For nucleated blood, 27–54 µl was used with an adapted protocol (Bionano, personal communication) of the Bionano Prep Blood and Cell Culture DNA Isolation Kit (cat no. RE-130-10). Lysates were embedded into agarose plugs and treated with Proteinase K and RNase A. Plugs were then purified by drop dialysis with 1× TE. DNA quality was assessed using pulse field gel electrophoresis (PFGE) (Pippin Pulse, SAGE Science, Beverly, MA) or the Femto Pulse instrument (Agilent). PFGE revealed that we isolated ultra-high-molecular-weight DNA between ~100 and ~500 kb long.

Phenol–chloroform gDNA extraction

For some samples, we performed phenol–chloroform extractions for HMW gDNA. Snap-frozen tissue was pulverized into a fine powder with a mortar and pestle in liquid nitrogen. The powdered tissue was lysed overnight at 55 °C in high-salt tissue lysis buffer (400 mM NaCl, 20 mM Tris base (pH 8.0), 30 mM EDTA (pH 8.0), 0.5% SDS, 100 µg/ml Proteinase K), and powdered lung tissue was lysed overnight in Qiagen G2 lysis buffer (cat no. 1014636, Qiagen, Hilden, Germany) containing 100 µg/ml Proteinase K at 55 °C. RNA was removed by incubation in 50 µg/ml RNase A for 1 h at 37 °C. HMW gDNA was purified with two washes of phenol–chloroform-IAA equilibrated to pH 8.0, followed by two washes of chloroform-IAA, and precipitated in ice-cold 100% ethanol. Filamentous HMW gDNA was either spooled with shepherds hooks or collected by centrifugation. HMW gDNA was washed twice with 70% ethanol, dried for 20 min at room temperature and eluted in TE. For the flier cichlid muscle gDNA sample used for PacBio CLR and 10XG libraries, glycogen was precipitated by adding 1/10 (v/v) 0.3 M sodium acetate, pH 6.0 to the extracted genomic DNA, mixing carefully and spinning at room temperature at 10,000g. PFGE revealed

thatDNA molecule length was between 50 and 300 kb—often lower in size than that obtained with the agarose plug but sufficient for long-range sequencing of CLR and linked read data types.

Others

We also used the Qiagen MagAttract HMW DNA kit (cat no. 67563) and the KingFisher Cell and Tissue DNA kit (Thermo Scientific; cat no. 97030196), following the manufacturers' guidelines. These protocols yielded HMW DNA ranging from 30 to 50 kb. The Genomic Tip (Qiagen) kit was also used for tissue-based extraction of HMW DNA.

Libraries and sequencing

PacBio libraries and sequencing

DNA obtained from agarose plugs was sheared down to ~40 kb fragment size with a MegaRuptor device (Diagenode, Belgium) and fragmented using Covaris g-tubes (520079) or by needle shearing. PacBio large insert libraries were prepared with either the SMRTbell Template Prep Kit 1.0-SPv3 (no.100-991-900) or the SMRTbell Express Template Prep Kit v1 (no. 101-357-000). Libraries were size-selected between 12 and 25 kb using Sage BluePippin (Sage Science, USA), depending on the DNA quality and extraction method. These libraries were sequenced on either RSII or Sequel I instruments, at least 60 \times coverage per species using Sequel Binding Kit and Sequencing Plate versions 2.0 and 2.1 with 10-h movie time (Supplementary Table 9).

10X Chromium libraries and sequencing

Unfragmented HMW DNA from the agarose plugs was used to generate linked read libraries on the 10X Genomics Chromium platform (Genome Library Kit & Gel Bead Kit v2 PN-120258, Genome Chip Kit v2 PN-120257, i7 Multiplex Kit PN-120262) following the manufacturer's guidelines. We sequenced the 10X libraries at ~60 \times coverage per species on an Illumina NovaSeq S4 150-bp PE lane.

Bionano libraries and optical map imaging

Unfragmented ultra-HMW DNA from the agarose plugs was labelled using either two different nicking enzymes (BspQI and BssSI) or a direct labelling enzyme (DLE1) following the Bionano Prep Labelling NLRs (document number 30024) and DLS protocols, respectively (document number 30206). Labelled samples were then imaged on a Bionano Irys or on a Bionano Saphyr instrument. For all species, we aimed for at least 100× coverage per label (Supplementary Table 9).

Hi-C libraries and sequencing

Chromatin interaction (Hi-C) libraries were generated using either Arima Genomics, Dovetail Genomics, or Phase libraries on muscle, blood, or other tissue with *in vivo* cross-linking (Supplementary Table 9) and sequenced on Illumina instruments. Arima-HiC preparations were performed by Arima Genomics (<https://arimagenomics.com/>) using the Arima-HiC kit that uses two enzymes (P/N: A510008). The resulting Arima-HiC proximally ligated DNA was then sheared, size-selected around 200–600 bp using SPRI beads, and enriched for biotin-labelled proximity-ligated DNA using streptavidin beads. From these fragments, Illumina-compatible libraries were generated using the KAPA Hyper Prep kit (P/N: KK8504). The resulting libraries were PCR amplified and purified with SPRI beads. The quality of the final libraries was checked with qPCR and Bioanalyzer, and then sequenced on Illumina HiSeq X at ~60× coverage following the manufacturer's protocols. Dovetail-HiC preparations were performed by Dovetail using a single-enzyme (DpnII) proximity ligation approach. Phase-HiC libraries were made by Phase Genomics using a Proximo Hi-C Library single-enzyme reaction.

Quality control

Before we performed any assembly, all genomic data of all data types from each sample were used to screen potential outlier libraries, outlier sequencing runs, or accidental species contamination with Mash⁷³ by measuring sequence similarity (Supplementary Fig. 4). When running

Mash, we used 21-mers to generate sketches with sketch size of 10,000 and compared among each sequencing run, and then differences assessed between sequencing sets.

Genome size, repeat content, and heterozygosity estimations

These estimations were made with k -mer-based methods applied to the Illumina short reads obtained from 10XG linked sequencing libraries. After trimming off barcodes during scaff10x⁷⁴ preprocessing, canonical 31-mer counts were collected using Meryl²³. With the resulting 31-mer histogram, GenomeScope⁷¹ was used to estimate the haploid genome length, repeat content, and heterozygosity. The thorny skate linked read data failed quality control, which we suspect was due to low complexity sequences from the high repeat content (54.1%) of the genome; so k -mers were collected later from Illumina whole-genome sequencing reads instead. The genome size and repeat content of the channel bull blenny were estimated from an alternative method that looks at the mode of long read overlap coverage and WindowMasker⁷⁵, as the estimated genome size from GenomeScope was almost doubling the known haploid genome size (1.29 Gb versus 0.6 Gb) and repeat content (28.0% versus 58.0%), for reasons related to either the quality of the 10X data or species differences.

Benchmarking assembly steps with the Anna's hummingbird

To develop the VGP standard pipeline, we compared various scaffolding, gap filling, and polishing tools. Default options were used unless otherwise noted. Detailed software versions are listed in Supplementary Table 2.

Contigging and scaffolding

FALCON⁷⁶ and FALCON-Unzip¹⁷ (smrtanalysis 3.0.0) were used to generate contigs that used CLR. Canu⁷⁷ 1.5+67 was used to generate the combined PacBio CLR and Oxford Nanopore ONT assembly. To benchmark scaffolding with linked reads, we used scaff10x⁷⁴ 2.0. For the linked read-only assembly, Supernova 2⁷⁸ was used. For the optical maps, two-enzyme hybrid scaffolding was used in the Bionano Solve v3.2.1

software, using BspQI and BssSI initially, as well as DLE1 later when the technology was developed. For benchmarking Hi-C in scaffolding, Salsa 2.2⁷⁹ was used for scaffolding results in Fig. 1a, with Hi-C reads generated from Arima Genomics. Additional comparisons for the Hi-C libraries were performed using assemblies provided by Dovetail Genomics and Phase Genomics (Supplementary Table 3). We used Hi-C from Arima Genomics as it had the smallest number of PCR duplicates and better coverage for short and long interactions at the time of comparison (Supplementary Fig. 1). Assembly statistics from HiRise, Proximo HiC, 3D-DNA⁸⁰ and Arima Hi-C are available in Supplementary Table 3. We concluded that all Hi-C scaffolding algorithms had similar performance. We decided to use Salsa, as HiRise and Proximo HiC were not open access, and 3D-DNA was computationally expensive on the DNAnexus platform. For short read assemblies, other than Supernova and the NRGene assembly, the assembly GCA_000699085.1¹⁶ was used for benchmarking, which was generated with Illumina paired-end, multiple mate-pair libraries and the SoapDeNovo⁸¹ assembler. The NRGene assembly was provided by the company with DeNovo Magic.

Gap filling

We ran PBJelly with support --capturedOnly --spanOnly parameters, to avoid greedy gap closures with no spanning read support. For conservatively filling sequences, we compared different parameters in output stage with --minreads 1 and --minreads 4 in addition to no restrictions. We found that the number of gaps closed was similar to the gaps filled with Arrow⁷⁶ (Supplementary Table 4) and chose not to run PBJelly⁸² for future assemblies.

Short-read polishing

Illumina polishing benchmarking was performed using Longranger⁸³ 2.1.3 and Pilon⁸⁴ 1.21 with --fix bases, local option (Supplementary Table 5). Later, for the VGP pipeline, we used FreeBayes⁸⁵ as Pilon⁸⁴ was not computationally scalable for large genomes with the updated Longranger 2.2.2.

Base-level accuracy estimate

Base-level accuracy was measured using a mapping-based approach and later using the *k*-mer-based approach²³. To determine the number of rounds to polish, we used Illumina paired-end reads from the hummingbird¹⁶.

Mis-joins and missed-joins

The curated hummingbird assembly was mapped to the target assemblies with MashMap2⁸⁶ with --filter_mode one-to-one --pi 95 using 5 kb segments (-s 5000) for CLR assemblies and 1 kb (-s 1000) for SR assemblies to compensate for the shorter contig sizes, as contigs smaller than a segment size will be excluded from the alignment. The number of mis-joins and missed joins were identified using the assembly_comparison.pl used in the ‘Curation’ section below ([Supplementary Methods](#), Supplementary Fig. 5).

VGP standard genome assembly pipeline 1.0 to 1.6

All 17 genomes were assembled with the VGP pipeline (Extended Data Fig. [2a](#)) for benchmark purposes, with some uncurated. The pale spear-nosed bat, greater horseshoe bat, Canada lynx, platypus, male and female zebra finch, kākāpō, Anna’s hummingbird, Goode’s thornscrub tortoise, flier cichlid, and blunt-snouted clingfish assemblies were generated using the VGP pipeline 1.0 to 1.6 and curated for submission to NCBI and EBI public archives. The curated and submitted two-lined caecilian, zig-zag eel, climbing perch, channel bull blenny, eastern happy, and thorny skate assemblies were generated using a similar process developed in parallel ([Supplementary Note 2](#)). Two submitted curated versions of the female zebra finch were made, one using the standard VGP pipeline and the other using the VGP trio pipeline, so that comparative analyses could be performed by others.

Contigging

For PacBio data, contigs were generated from subreads using FALCON⁷⁶ and FALCON-Unzip¹⁷, with one round of Arrow polishing (smrtanalysis 5.1.0.26412). A minimum read length of 2 kb or a cutoff at which reads longer than the cutoff include 50× coverage was used, whichever was longer. For calculating read coverage, we used estimated genome size from <http://www.genomesize.com/> when available, or from the literature (Supplementary Table 11) while waiting for 10XG sequencing to estimate genome size using *k*-mers. FALCON and FALCON-Unzip were run with default parameters, except for computing the overlaps. Raw read overlaps were computed with DALIGNER parameters -k14 -e0.75 -s100 -l2500 -h240 -w8 to better reflect the higher error rate in early PacBio sequel I and II. Pread (presassembled read) overlaps were computed with DALIGNER parameters -k24 -e.90 -s100 -l1000 -h600 intending to collapse haplotypes for the FALCON step to better unzip genomes with high heterozygosity rate. FALCON-Unzip outputs both a pseudo-haplotype and a set of alternate haplotigs that represent the secondary alleles. We refer to these outputs as the primary contig set (c1) and alternate contig set (c2).

Purging false duplications

Heterotype false duplications occurred despite setting FALCON⁷⁶ parameters to resolve up to 10% haplotype divergence. FALCON-Unzip¹⁷ also incorrectly retained some secondary alleles in the primary contig set, which appeared as false duplications. To reduce these false duplications, we ran Purge_Haplots¹³, first during curation (VGP v1.0 pipeline) and then later after contig formation (VGP v1.5 pipeline). To do the former, Purge_Haplots was run on the primary contigs (c1), and identified haplotigs were mapped to the scaffolded primary assembly with MashMap2⁸⁶ for removal. In the latter, identified haplotigs were moved from the primary contigs (c1) to the alternate haplotig set (p2). The remaining primary contigs were referred to as p1; p2 combined with c2 was referred to as q2. Later, in the VGP v1.6 pipeline, we replaced Purge_Haplots with Purge_Dups¹⁴, a new program developed by several of the authors in response to Purge_Haplots not removing partial false duplication at contig boundaries. Purging also removes excessive low-coverage (junk) and high-coverage (repeats) contigs. To calculate the

presence and overall success of purging false duplications, we used a k -mer approach ([Supplementary Methods](#), Supplementary Fig. 6).

Scaffolding with 10XG linked reads

The 10X Genomics linked reads were aligned to the primary contigs (p1), and an adjacency matrix was computed from the barcodes using scaff10x⁷⁴ v2.0–2.1. Two rounds of scaffolding were performed. The first round was run with parameters -matrix 2000 -reads 12 -link 10, and the second round with parameters -matrix 2000 -reads 8 -link 10. A gap of 100 bp (represented with ‘N’s) was inserted between joined contigs. The resulting primary scaffold set was named s1.

Scaffolding with Bionano optical maps

Bionano cmaps were generated using the Bionano Pipeline in non-haplotype assembly mode and used to further scaffold the s1 assembly with Bionano Solve v3.2.1⁸⁷. We began with a one-enzyme nick map (BspQI), followed by a two-enzyme nick map (BspQI and BssSI), and then with a DLE-1 one-enzyme non-nicking approach when the later data type became available (Supplementary Table 9). Scaffold gaps were sized according to the software estimate. The resulting scaffold set was named s2.

Scaffolding with Hi-C reads

Hi-C reads were aligned to the s2 scaffolds using the Arima Genomics mapping pipeline⁸⁸. In brief, both ends of a read pair were mapped independently using BWA-MEM⁸⁹ with the parameter -B8, and filtered when mapping quality was <10. Chimeric reads containing a restriction enzyme site were trimmed from the restriction site onward, leaving only the 5' end. The filtered single-read alignments were then rejoined as paired read alignments. The processed alignments were then used for scaffolding with Salsa2⁷⁹, which analyses the normalized frequency of Hi-C interactions between all pairs of contig ends to determine a likely ordering and orientation of each. We used parameters -m yes -i 5 -p yes to allow Salsa2 to break potentially mis-assembled contigs and perform five iterations of

scaffolding. After feedback from curation, later versions of Salsa were developed, which more conservatively determine the number of iterations (v2.1) and actively break at mis-assemblies (v2.2), and run for the Canada lynx, Goode's thornscrub tortoise, and two-lined caecilian. The restriction enzyme(s) used to generate each library were specified using parameters -e GATC,GANTC for Arima and -e GATC for Dovetail and Phase Genomics Hi-C data. The resulting Hi-C scaffolded assembly was named s3.

Consensus polishing

To polish bases in both haplotypes with minimal alignment bias, we concatenated the alternate haplotig set (c2 in v1.0 or q2 in v1.5–1.6) to the scaffolded primary set (s3) and the assembled mitochondrial genome (mitoVGP in v1.6). We then performed another round of polishing with Arrow (smrtanalysis 5.1.0.26412) using PacBio CLR reads, aligning with pbalign --minAccuracy=0.75 --minLength=50 --minAnchorSize=12 --maxDivergence=30 –concordant --algorithm=blasr --algorithmOptions=--useQuality --maxHits=1 --hitPolicy=random --seed=1 and consensus polishing with variantCaller --skipUnrecognizedContigs haploid -x 5 -q 20 -X120 –v --algorithm=arrow. While this round of polishing resulted in higher QV for all genomes herein considered, we noticed that it was particularly sensitive to the coverage cutoff parameter (-x). This is because Arrow generates a de novo consensus from the mapped reads without explicitly considering the reference sequence. Later, we found that the second round of Arrow polishing sometimes reduced the QV accuracy for some species. Upon investigation, this issue was traced back to option -x 5, which requires at least 5 reads to call consensus. Such low minimum requirements can lead to uneven polishing in low coverage regions. To avoid this behaviour, we suggest to increase the -x close to the half sequence coverage (for example, 30× when 60× was used for assembly) and check QV before moving forward.

For genomes with a combined assembly size larger than 4 Gb, we used Minimap2⁹⁰ with parameters -ax map-pb instead of Blasr⁹¹ to overcome reference index size limitations.

Two more rounds of base-pair polishing were performed with linked reads. The reads were aligned with Longranger align 2.2.2, which incorporates the Lauriat for barcode-aware alignment⁸³. From the alignments, homozygous mismatches (variants) were called with FreeBayes⁸³ v1.2.0 using default options. Consensus was called with bcftools consensus⁹² with -i'QUAL>1 && (GT="AA" || GT = 'Aa') -Hla.

VGP Trio Pipeline v1.0–v1.6

The trio pipeline is similarly designed to the standard pipeline, except for the use of parental data (Extended Data Fig. 3b). When parental genomes are available, the child's CLR reads are binned to maternal and paternal haplotypes, and assembled separately as haplotype-specific contigs (haplotigs) using TrioCanu²⁰. In brief, parental specific marker k -mers were collected using Meryl²³ from the parental Illumina WGS reads of the parents. These markers were filtered and used to bin the child's CLR read. A haplotype was assigned given the markers observed, normalized by the total markers in each haplotype. The subsequent purging, scaffolding, and polishing steps were similarly updated with the use of Purge_Dups¹⁴ (v1.6). We extended binning to linked reads and Hi-C reads, by excluding read pairs that had any parental-specific marker. The binned Hi-C reads were used to scaffold its haplotype assembly, and polished with the binned linked reads from the observation of haplotype switching using the standard polishing approach. During curation, one of the haplotype assemblies with the higher QV and/or contiguity was chosen as the representative haplotype. The heterogametic sex chromosome from the unchosen haplotype was added to the representative assembly. However, while curating several trios, we found that in regions of low divergence between shared parental homogametic sex chromosomes (that is, X or Z), a small fraction of offspring CLR data was mis-assigned to the wrong haplotype. This mis-alignment resulted in a duplicate, low-coverage offspring X or Z assembly in the paternal (for mammals) or maternal (for birds) haplotype, respectively, which required removal during curation. We are working on methods to improve the binning accuracy for resolution of this issue going forward.

For the female zebra finch in particular, contigs were generated before the binning was automated in the Canu assembler as TrioCanu1.7, and therefore a manual binning process was applied as described in the original Trio-binning paper²⁰ ([Supplementary Methods](#)). Contigs were assembled for each haplotype using the binned reads, excluding unclassified reads. The contigs were polished with two rounds of Arrow polishing using the binned reads, and scaffolded following the v1.0 pipeline with no purging. Additional scaffolding rounds with Bionano (s4) and Hi-C were applied. Scaffolds were renamed according to the primary scaffold assembly of the same individual (s5), with sex chromosomes grouped as Z in the paternal assembly and W in the maternal assembly following synteny to the Z chromosome from the curated male zebra finch VGP assembly. Two rounds of SR polishing were applied using linked reads, by mapping on both haplotypes. After haplotype switches were discovered, additional rounds of polishing were applied using binned linked reads ([Supplementary Methods](#)).

Mitochondrial genome assembly

Similar to other recent methods^{93,94}, we developed a reference-guided MT assembly pipeline. MT reads in the raw CLR data were identified by mapping the whole read set to an existing reference sequence of the specific species or of closely related species using Blasr. Filtered mtDNA CLRs were assembled into a single contig using Canu v1.8, polished with Arrow using CLR and then FreeBayes v1.0.2 together with bcftools v1.9 using short reads from the 10XG data (Extended Data Fig. 3c). The overlapping sequences at the ends of the contig were trimmed, and the remaining contig sequence circularized. The mitoVGP pipeline is made available at <https://github.com/VGP/vgp-assembly/tree/master/mitoVGP>. A more detailed protocol description of the assembly pipeline and new discoveries from the MT assemblies are published elsewhere³³.

Curation

The VGP genome assembly pipeline produces high quality assemblies, yet no automated method to date is free from the production of errors, especially during the scaffolding stages. To minimize the impact of the remaining algorithmic shortcomings, we subjected all assemblies to rigorous manual curation. All data generated for a species in this study and other publicly available data (for example, genetic maps, gene sets and genome assemblies of the same or closely related species) were collated, aligned to the primary assembly and analysed in gEVAL⁹⁵ (<https://vgp-geval.sanger.ac.uk/index.html>), visualizing discordances in a feature browser and issue lists. In parallel, Hi-C data were mapped to the primary assembly and visualized using Juicebox⁹⁶ and/or HiGlass⁹⁷. With these data, genome curators identified mis-joins, missed joins and other anomalies, and corrected the primary assembly accordingly. No change was made without unambiguous evidence from available data types; for example, a Hi-C suggested join would not be made unless supported by BioNano maps, long-read data, or gene alignments. When sequencing the heterogametic sex, we identified sex chromosomes based on half coverage,

homology alignments to sex chromosomes in other species, and the presence of sex chromosome-specific genes.

Contamination removal

A succession of searches was used to identify potential contaminants in the generated assemblies.

- 1) A megaBLAST⁹⁸ search against a database of common contaminants (ftp://ftp.ncbi.nlm.nih.gov/pub/kitts/contam_in_euks.fa.gz) requiring $e \leq 1 \times 10^{-4}$, reporting matches with $\geq 98\%$ sequence identity and match length 50–99 bp, $\geq 94\%$ and match length 100–199 bp, or $\geq 90\%$ and match length 200 bp or above.
- 2) A vecscreen (<https://www.ncbi.nlm.nih.gov/tools/vecscren/>) search against a database of adaptor sequences (ftp://ftp.ncbi.nlm.nih.gov/pub/kitts/adaptors_for_screening_euks.fa)
- 3) After soft-masking repeats using Windowmasker⁷⁵, a megaBLAST search against chromosome-level assemblies from RefSeq requiring $e \leq 1 \times 10^{-4}$, match score ≥ 100 , and sequence identity $\geq 98\%$; regions matching highly conserved rDNAs were ignored.

Manual inspection of the results was necessary to differentiate contamination from conservation and/or horizontal gene transfer. Adaptor sequences were masked; other contaminant sequences were removed. Assemblies were also checked for runs of Ns at the ends of scaffolds, created as artefacts of the iterative scaffolding process, and when found they were trimmed.

Organelle genomes

These were detected by a megaBLAST search against a database of known organelle genomes requiring $e \leq 1 \times 10^{-4}$, sequence identity $\geq 90\%$, and match length ≥ 500 ; the databases are available at <ftp://ftp.ncbi.nlm.nih.gov/blast/db/FASTA/mito.nt.gz> and ftp://ftp.ncbi.nlm.nih.gov/refseq/release/plastid/*genomic.fna.gz. Only

scaffolds consisting entirely of organelle sequences were assumed to be organelle genomes, and replaced by the genome from the separate organelle assembly pipeline. Organelle matches embedded in nuclear sequences that were found to be NuMTs were kept.

False duplication removal

Retained false duplications were identified using Purge_Haplotigs¹³ run either after scaffolding and polishing (Anna's hummingbird, kākāpō, male zebra finch, female zebra finch, platypus, pale spear-nosed bat, and greater horseshoe bat) or on the c1 before scaffolding (two-lined caecilian, flier cichlid, Canada lynx, and Goode's thornscrub tortoise). Subsequent manual curation identified additional haplotypic duplications for the listed assemblies and also those that were not treated with Purge_Haplotigs (Eastern happy, climbing perch, zig-zag eel). The evidence used included read coverage, sequence self-comparison, transcript alignments, Bionano map alignments and Hi-C 2D maps, all confirming the superfluous nature of one allele. The identified haplotype duplications were moved from the primary to the alternate assembly.

Chromosome assignment

For a scaffold to be annotated as a chromosome, we used evidence from Hi-C as well as genetic linkage or FISH karyotype mapping when available. For Hi-C evidence, we considered a scaffold as a complete chromosome (albeit with gaps) when there was a clear unbroken diagonal in the Juicebox or HiGlass plots for that scaffold and no other large scaffolds that could be joined to that same scaffold; if present and no unambiguous join was possible, we named it as an unlocalized scaffold for that chromosome.

When we could not find evidence of a complete chromosome, we kept the scaffold number for its name. We named all evidence-validated scaffolds as chromosomes down to the smallest Hi-C box unit resolution allowed with these characteristics. When there was an established chromosome terminology for a given species or set of species, we use the established terminology except when our new assemblies revealed errors in the older assembly, such as scaffold/chromosome fusions, fissions, rearrangements,

and non-chromosome names. For species without an established chromosome terminology, we named the scaffolds as chromosomes numbers 1, 2, 3 …, in descending order of scaffold size. For the sex chromosomes, we used the letters X and Y for mammals and Z and W for birds.

Using comparative genomics to assess assembly structure

In cases where a high-quality chromosome-level genome was available for a closely related species, comparative genome analysis was performed. The polished primary assembly (t3.p) was mapped to the related genome using MashMap2⁸⁶ with --pi 75 -s 300000. The number of chromosomal differences was identified using a custom script available at https://github.com/jdamas13/assembly_comparison. This resulted in the identification of ~60 to ~450 regions for each genome assembly flanking putative misassemblies or lineage-specific genome rearrangements. To identify which were real misassemblies, the identified discrepancies were communicated to the curation team for manual verification (see above).

To identify any possible remaining mis-joins, each curated avian and mammalian assembly was compared with the zebra finch (taeGut2) or human (hg38) genomes, respectively. Pairwise alignments between each of the VGP assemblies and the clade reference were generated with LastZ⁹⁹ (version 1.04) using the following parameters: $C = 0$ $E = 30$ $H = 2000$ $K = 3000$ $L = 2200$ $O = 400$. The pairwise alignments were converted into the UCSC ‘chain’ and ‘net’ formats with axtChain (parameters: -minScore = 1000 -verbose = 0 -linearGap = medium) followed by chainAntiRepeat, chainSort, chainPreNet, chainNet and netSyntenic, all with default parameters¹⁰⁰. Pairwise synteny blocks were defined using maf2synteny¹⁰¹ at 100-, 300-, and 500-kb resolutions. Evolutionary breakpoint regions were detected and classified using an ad hoc statistical approach¹⁰². This analysis identified 2 to 90 genomic regions per assembly that could be flanking misassemblies, lineage-specific chromosome rearrangements, or reference-specific chromosome rearrangements (116 in the human and 26 in the zebra finch). Determining the underlying cause for each of the flagged regions will need further verification. All alignments are

available for visualization at the Evolution Highway comparative chromosome browser (<http://eh-demo.ncsa.illinois.edu/vgp/>).

Annotation

NCBI and Ensembl annotation pipeline used in this study are described in the [Supplementary Methods](#).

Evaluation

Detailed methods for other types of evaluation, including BUSCO runs, mis-join and missed-join identification, reliable blocks, collapsed repeats, telomeres, RNA-seq and ATAC-seq mapping, and false gene duplications are in the [Supplementary Methods](#). No statistical methods were used to predetermine sample size, the experiments were not randomized, and the investigators were not blinded to group during experiments and outcome assessment.

Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

Data availability

All raw data, intermediate and final assemblies are publicly available via GenomeArk (<https://vgp.github.io/genomeark>), archived on NCBI/EBI BioProject under accession [PRJNA489243](#) with annotations, and browsable on the UCSC Genome Browser (<https://hgdownload.soe.ucsc.edu/hubs/VGP/>). The final primary assembly from the automated pipeline before curation is browsable on gEVAL (<https://vgp-geval.sanger.ac.uk>) with all four raw data mappings. The VGP assembly pipeline is available as a stand-alone pipeline (<https://github.com/VGP/vgp-assembly>) as well as a workflow on DNAnexus (<https://platform.dnanexus.com/>). A VGP-specific assembly hub portal in the U.C. Santa Cruz browser is available as a gateway to

access all VGP genome assemblies and annotations (<https://hgdownload.soe.ucsc.edu/hubs/VGP>).

Code availability

All codes used in the VGP Assembly Pipeline and the VGP Trio Pipeline are publicly available at <https://github.com/VGP/vgp-assembly/tree/master/pipeline>.

References

1. 1.

International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).

[ADS](#) [Google Scholar](#)

2. 2.

Sulston, J. et al. The *C. elegans* genome sequencing project: a beginning. *Nature* **356**, 37–41 (1992).

[PubMed](#) [ADS](#) [Google Scholar](#)

3. 3.

Mouse Genome Sequencing Consortium. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520–562 (2002).

[Google Scholar](#)

4. 4.

Howe, K. et al. The zebrafish reference genome sequence and its relationship to the human genome. *Nature* **496**, 498–503 (2013).

[CAS](#) [PubMed](#) [PubMed Central](#) [ADS](#) [Google Scholar](#)

5. 5.

Genome 10K Community of Scientists. Genome 10K: a proposal to obtain whole-genome sequence for 10,000 vertebrate species. *J. Hered.* **100**, 659–674 (2009).

[PubMed Central](#) [Google Scholar](#)

6. 6.

Koepfli, K.-P., Paten, B., the Genome 10K Community of Scientists & O'Brien, S. J. The Genome 10K Project: a way forward. *Annu. Rev. Anim. Biosci.* **3**, 57–111 (2015).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

7. 7.

Venter, J. C. et al. The sequence of the human genome. *Science* **291**, 1304–1351 (2001).

[CAS](#) [PubMed](#) [ADS](#) [Google Scholar](#)

8. 8.

Adams, M. D. et al. The genome sequence of *Drosophila melanogaster*. *Science* **287**, 2185–2195 (2000).

[PubMed](#) [Google Scholar](#)

9. 9.

Shendure, J. & Ji, H. Next-generation DNA sequencing. *Nat. Biotechnol.* **26**, 1135–1145 (2008).

[CAS](#) [PubMed](#) [Google Scholar](#)

10. 10.

Yin, Z.-T. et al. Revisiting avian ‘missing’ genes from de novo assembled transcripts. *BMC Genomics* **20**, 4 (2019).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

11. 11.

Korlach, J. et al. De novo PacBio long-read and phased avian genome assemblies correct and add to reference genes generated with intermediate and short reads. *Gigascience* **6**, 1–16 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

12. 12.

Kelley, D. R. & Salzberg, S. L. Detection and correction of false segmental duplications caused by genome mis-assembly. *Genome Biol.* **11**, R28 (2010).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

13. 13.

Roach, M. J., Schmidt, S. A. & Borneman, A. R. Purge Haplotigs: allelic contig reassignment for third-gen diploid genome assemblies. *BMC Bioinformatics* **19**, 460 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

14. 14.

Guan, D. et al. Identifying and removing haplotypic duplication in primary genome assemblies. *Bioinformatics* **36**, 2896–2898 (2020).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

15. 15.

Bradnam, K. R. et al. Assemblathon 2: evaluating de novo methods of genome assembly in three vertebrate species. *Gigascience* **2**, 10 (2013).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

16. 16.

Zhang, G. et al. Comparative genomics reveals insights into avian genome evolution and adaptation. *Science* **346**, 1311–1320 (2014).

[CAS](#) [PubMed](#) [PubMed Central](#) [ADS](#) [Google Scholar](#)

17. 17.

Chin, C.-S. et al. Phased diploid genome assembly with single-molecule real-time sequencing. *Nat. Methods* **13**, 1050–1054 (2016).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

18. 18.

Bresler, G., Bresler, M. & Tse, D. Optimal assembly for high throughput shotgun sequencing. *BMC Bioinformatics* **14** (Suppl. 5), S18 (2013).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

19. 19.

Warren, W. C. et al. The genome of a songbird. *Nature* **464**, 757–762 (2010).

[CAS](#) [PubMed](#) [PubMed Central](#) [ADS](#) [Google Scholar](#)

20. 20.

Koren, S. et al. De novo assembly of haplotype-resolved genomes with trio binning. *Nat. Biotechnol.* (2018).

21. 21.

Koren, S., Phillippy, A. M., Simpson, J. T., Loman, N. J. & Loose, M. Reply to ‘Errors in long-read assemblies can critically affect protein prediction’. *Nat. Biotechnol.* **37**, 127–128 (2019).

[CAS](#) [PubMed](#) [Google Scholar](#)

22. 22.

Vollger, M. R. et al. Long-read sequence and assembly of segmental duplications. *Nat. Methods* **16**, 88–94 (2019).

[CAS](#) [PubMed](#) [Google Scholar](#)

23. 23.

Rhie, A., Walenz, B. P., Koren, S. & Phillippy, A. M. Merqury: reference-free quality, completeness, and phasing assessment for genome assemblies. *Genome Biol.* **21**, 245 (2020).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

24. 24.

Waterhouse, R. M. et al. BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mol. Biol. Evol.* **35**, 543–548 (2018).

[CAS](#) [PubMed](#) [Google Scholar](#)

25. 25.

Howe, K. et al. Significantly improving the quality of genome assemblies through curation. *Gigascience* **10**, giaa153 (2021).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

26. 26.

Zhou, Y. et al. Platypus and echidna genomes reveal mammalian biology and evolution. *Nature* <https://doi.org/10.1038/s41586-020-03039-0> (2021).

27. 27.

Kim, J. et al. False gene and chromosome losses affected by assembly and sequence errors. Preprint at <https://doi.org/10.1101/2021.04.09.438906> (2021).

28. 28.

Lewin, H. A., Graves, J. A. M., Ryder, O. A., Graphodatsky, A. S. & O'Brien, S. J. Precision nomenclature for the new genomics. *Gigascience* **8**, giz086 (2019).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

29. 29.

Kronenberg, Z. N. et al. Extended haplotype phasing of de novo genome assemblies with FALCON-Phase. *Nat. Commun.* <https://doi.org/10.1038/s41467-020-20536-y> (2021).

30. 30.

Ewing, B., Hillier, L., Wendl, M. C. & Green, P. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* **8**, 175–185 (1998).

[CAS](#) [PubMed](#) [Google Scholar](#)

31. 31.

Tomaszkiewicz, M., Medvedev, P. & Makova, K. D. Y and W chromosome assemblies: approaches and discoveries. *Trends Genet.* **33**, 266–282 (2017).

[CAS](#) [PubMed](#) [Google Scholar](#)

32. 32.

Kolesnikov, A. A. & Gerasimov, E. S. Diversity of mitochondrial genome organization. *Biochem. (Mosc.)* **77**, 1424–1435 (2012).

[CAS](#) [Google Scholar](#)

33. 33.

Formenti, G. et al. Complete vertebrate mitogenomes reveal widespread repeats and gene duplications. *Genome Biol.* (in the press).

34. 34.

Harrison, G. L. A. et al. Four new avian mitochondrial genomes help get to basic evolutionary questions in the late cretaceous. *Mol. Biol. Evol.* **21**, 974–983 (2004).

[CAS](#) [PubMed](#) [Google Scholar](#)

35. 35.

Zhao, H. et al. The complete mitochondrial genome of the *Anabas testudineus* (Perciformes, Anabantidae). *Mitochondrial DNA A DNA Mapp. Seq. Anal.* **27**, 1005–1007 (2016).

[CAS](#) [PubMed](#) [Google Scholar](#)

36. 36.

Suzuki, A. et al. How the kinetochore couples microtubule force and centromere stretch to move chromosomes. *Nat. Cell Biol.* **18**, 382–392 (2016).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

37. 37.

Pfenning, A. R. et al. Convergent transcriptional specializations in the brains of humans and song-learning birds. *Science* **346**, 1256846 (2014).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

38. 38.

Robinson, R. For mammals, loss of yolk and gain of milk went hand in hand. *PLoS Biol.* **6**, e77 (2008).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

39. 39.

Brandl, K. et al. Yip1 domain family, member 6 (Yipf6) mutation induces spontaneous intestinal inflammation in mice. *Proc. Natl Acad. Sci. USA* **109**, 12650–12655 (2012).

[CAS](#) [PubMed](#) [ADS](#) [Google Scholar](#)

40. 40.

Malmstrøm, M. et al. Evolution of the immune system influences speciation rates in teleost fishes. *Nat. Genet.* **48**, 1204–1210 (2016).

[PubMed](#) [Google Scholar](#)

41. 41.

Japundžić-Žigon, N., Lozić, M., Šarenac, O. & Murphy, D. Vasopressin & oxytocin in control of the cardiovascular system: an updated review. *Curr. Neuropharmacol.* **18**, 14–33 (2020).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

42. 42.

Cataldo, I., Azhari, A. & Esposito, G. A review of oxytocin and arginine-vasopressin receptors and their modulation of autism spectrum disorder. *Front. Mol. Neurosci.* **11**, 27 (2018).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

43. 43.

Warren, W. C. et al. Genome analysis of the platypus reveals unique signatures of evolution. *Nature* **453**, 175–183 (2008).

[CAS](#) [PubMed](#) [PubMed Central](#) [ADS](#) [Google Scholar](#)

44. 44.

Ko, B. J. et al. Widespread false gene gains caused by duplication errors in genome assemblies. Preprint at <https://doi.org/10.1101/2021.04.09.438957> (2021).

45. 45.

Lemaire, S. et al. Characterizing the interplay between gene nucleotide composition bias and splicing. *Genome Biol.* **20**, 259 (2019).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

46. 46.

Zhang, L., Kasif, S., Cantor, C. R. & Broude, N. E. GC/AT-content spikes as genomic punctuation marks. *Proc. Natl Acad. Sci. USA* **101**, 16855–16860 (2004).

[CAS](#) [PubMed](#) [ADS](#) [Google Scholar](#)

47. 47.

Jarvis, E. D. et al. Global view of the functional molecular organization of the avian cerebrum: mirror images and functional columns. *J. Comp. Neurol.* **521**, 3614–3665 (2013).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

48. 48.

Kubikova, L., Wada, K. & Jarvis, E. D. Dopamine receptors in a songbird brain. *J. Comp. Neurol.* **518**, 741–769 (2010).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

49. 49.

Sémon, M. & Wolfe, K. H. Rearrangement rate following the whole-genome duplication in teleosts. *Mol. Biol. Evol.* **24**, 860–867 (2007).

[PubMed](#) [Google Scholar](#)

50. 50.

Jebb, D. et al. Six reference-quality genomes reveal evolution of bat adaptations. *Nature* **583**, 578–584 (2020).

[CAS](#) [PubMed](#) [ADS](#) [Google Scholar](#)

51. 51.

Schneider, V. A. et al. Evaluation of GRCh38 and de novo haploid genome assemblies demonstrates the enduring quality of the reference assembly. *Genome Res.* **27**, 849–864 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

52. 52.

Warren, W. C. et al. A new chicken genome assembly provides insight into avian genome structure. *G3 (Bethesda)* **7**, 109–117 (2017).

[CAS](#) [Google Scholar](#)

53. 53.

Meredith, R. W. et al. Impacts of the Cretaceous Terrestrial Revolution and KPg extinction on mammal diversification. *Science* **334**, 521–524 (2011).

[CAS](#) [PubMed](#) [ADS](#) [Google Scholar](#)

54. 54.

Rodriguez-Agudo, D. et al. StarD5: an ER stress protein regulates plasma membrane and intracellular cholesterol homeostasis. *J. Lipid Res.* **60**, 1087–1098 (2019).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

55. 55.

Kim, J. et al. Reconstruction and evolutionary history of eutherian chromosomes. *Proc. Natl Acad. Sci. USA* **114**, E5379–E5388 (2017).

[CAS](#) [PubMed](#) [Google Scholar](#)

56. 56.

Lin, B., Dutta, B. & Fraser, I. D. C. Systematic investigation of multi-TLR sensing identifies regulators of sustained gene activation in macrophages. *Cell Syst.* **5**, 25–37.e3 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

57. 57.

Theofanopoulou, C., Gedman, G. L., Cahill, J. A., Boeckx, C. & Jarvis, E. D. Universal nomenclature for oxytocin-vasotocin ligand and receptor families. *Nature* <https://doi.org/10.1038/s41586-020-03040-7> (2021).

58. 58.

Ocampo Daza, D. & Haitina, T. Reconstruction of the carbohydrate 6-O sulfotransferase gene family evolution in vertebrates reveals novel member, CHST16, lost in amniotes. *Genome Biol. Evol.* **12**, 993–1012 (2020).

[PubMed](#) [Google Scholar](#)

59. 59.

Damas, J. et al. Broad host range of SARS-CoV-2 predicted by comparative and structural analysis of ACE2 in vertebrates. *Proc. Natl Acad. Sci. USA* **117**, 22311–22322 (2020).

[CAS](#) [PubMed](#) [Google Scholar](#)

60. 60.

Dussex, N. et al. Population genomics reveals the impact of long-term small population size in the critically endangered kākāpō. *Cell Genom.* (in the press).

61. 61.

Teeling, E. C. et al. Bat biology, genomes, and the Bat1K project: to generate chromosome-level genomes for all living bat species. *Annu. Rev. Anim. Biosci.* **6**, 23–46 (2018).

[PubMed](#) [Google Scholar](#)

62. 62.

Lewin, H. A. et al. Earth BioGenome Project: sequencing life for the future of life. *Proc. Natl Acad. Sci. USA* **115**, 4325–4333 (2018).

[CAS](#) [PubMed](#) [Google Scholar](#)

63. 63.

Jarvis, E. D. et al. Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science* **346**, 1320–1331 (2014).

[CAS](#) [PubMed](#) [PubMed Central](#) [ADS](#) [Google Scholar](#)

64. 64.

Li, S. et al. Genomic signatures of near-extinction and rebirth of the crested ibis and other endangered bird species. *Genome Biol.* **15**, 557 (2014).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

65. 65.

Koren, S. & Phillippy, A. M. One chromosome, one contig: complete microbial genomes from long-read sequencing and assembly. *Curr. Opin. Microbiol.* **23**, 110–120 (2015).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

66. 66.

Jenjaroenpun, P. et al. Complete genomic and transcriptional landscape analysis using third-generation sequencing: a case study of *Saccharomyces cerevisiae* CEN.PK113-7D. *Nucleic Acids Res.* **46**, e38 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

67. 67.

Tyson, J. R. et al. MinION-based long-read sequencing and assembly extends the *Caenorhabditis elegans* reference genome. *Genome Res.* **28**, 266–274 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

68. 68.

Miga, K. H. et al. Telomere-to-telomere assembly of a complete human X chromosome. *Nature* **585**, 79–84 (2020).

[CAS](#) [PubMed](#) [PubMed Central](#) [ADS](#) [Google Scholar](#)

69. 69.

Logsdon, G. A. et al. The structure, function and evolution of a complete human chromosome 8. *Nature* <https://doi.org/10.1038/s41586-021-03420-7> (2021).

70. 70.

Beçak, M. L., Beçak, W., Roberts, F. L., Shoffner, R. N. & Volpe, P. (eds.) *Chromosome Atlas: Fish, Amphibians, Reptiles, and Birds* Vol. 2 (Springer, 1973).

71. 71.

Vrture, G. W. et al. GenomeScope: fast reference-free genome profiling from short reads. *Bioinformatics* **33**, 2202–2204 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

72. 72.

Kumar, S., Stecher, G., Suleski, M. & Hedges, S. B. TimeTree: a resource for timelines, timetrees, and divergence times. *Mol. Biol. Evol.* **34**, 1812–1819 (2017).

[CAS](#) [PubMed](#) [Google Scholar](#)

73. 73.

Ondov, B. D. et al. Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol.* **17**, 132 (2016).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

74. 74.

Ning, Z. & Harry, E. Scaff10X <https://github.com/wtsi-hpag/Scaff10X>.

75. 75.

Morgulis, A., Gertz, E. M., Schäffer, A. A. & Agarwala, R. WindowMasker: window-based masker for sequenced genomes. *Bioinformatics* **22**, 134–141 (2006).

[CAS](#) [PubMed](#) [Google Scholar](#)

76. 76.

Chin, C.-S. et al. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat. Methods* **10**, 563–569 (2013).

[CAS](#) [PubMed](#) [Google Scholar](#)

77. 77.

Koren, S. et al. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. *Genome Res.* **27**, 722–736 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

78. 78.

Weisenfeld, N. I., Kumar, V., Shah, P., Church, D. M. & Jaffe, D. B. Direct determination of diploid genome sequences. *Genome Res.* **27**, 757–767 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

79. 79.

Ghurye, J. et al. Integrating Hi-C links with assembly graphs for chromosome-scale assembly. *PLoS Comput. Biol.* **15**, e1007273 (2019).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

80. 80.

Lieberman-Aiden, E. et al. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**, 289–293 (2009).

[CAS](#) [PubMed](#) [PubMed Central](#) [ADS](#) [Google Scholar](#)

81. 81.

Luo, R. et al. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience* **1**, 18 (2012).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

82. 82.

English, A. C. et al. Mind the gap: upgrading genomes with Pacific Biosciences RS long-read sequencing technology. *PLoS ONE* **7**, e47768 (2012).

[CAS](#) [PubMed](#) [PubMed Central](#) [ADS](#) [Google Scholar](#)

83. 83.

Bishara, A. et al. Read clouds uncover variation in complex regions of the human genome. *Genome Res.* **25**, 1570–1580 (2015).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

84. 84.

Walker, B. J. et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE* **9**, e112963 (2014).

[PubMed](#) [PubMed Central](#) [ADS](#) [Google Scholar](#)

85. 85.

Garrison, E. & Marth, G. Haplotype-based variant detection from short-read sequencing. Preprint at <http://arxiv.org/abs/1207.3907> (2012).

86. 86.

Jain, C., Koren, S., Dilthey, A., Phillippy, A. M. & Aluru, S. A fast adaptive algorithm for computing whole-genome homology maps. *Bioinformatics* **34**, i748–i756 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

87. 87.

Bionano Genomics, Inc. Bionano Software Downloads.
<https://bionanogenomics.com/support/software-downloads/>.

88. 88.

Arima Genomics, Inc. Arima Genomics Mapping Pipeline.
https://github.com/ArimaGenomics/mapping_pipeline.

89. 89.

Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

90. 90.

Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

91. 91.

Chaisson, M. J. & Tesler, G. Mapping single molecule sequencing reads using basic local alignment with successive refinement (BLASR): application and theory. *BMC Bioinformatics* **13**, 238 (2012).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

92. 92.

Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

93. 93.

Dierckxsens, N., Mardulyn, P. & Smits, G. NOVOPlasty: de novo assembly of organelle genomes from whole genome data. *Nucleic Acids Res.* **45**, e18 (2017).

[PubMed](#) [Google Scholar](#)

94. 94.

Soorni, A., Haak, D., Zaitlin, D. & Bombara, A. Organelle_PBA, a pipeline for assembling chloroplast and mitochondrial genomes from PacBio DNA sequencing data. *BMC Genomics* **18**, 49 (2017).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

95. 95.

Chow, W. et al. gEVAL — a web-based browser for evaluating genome assemblies. *Bioinformatics* **32**, 2508–2510 (2016).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

96. 96.

Durand, N. C. et al. Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. *Cell Syst.* **3**, 99–101 (2016).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

97. 97.

Kerpedjiev, P. et al. HiGlass: web-based visual exploration and analysis of genome interaction maps. *Genome Biol.* **19**, 125 (2018).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

98. 98.

Camacho, C. et al. BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421 (2009).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

99. 99.

Harris, R. S. Improved Pairwise Alignment of Genomic DNA. Thesis, Pennsylvania State Univ. (2007).

100. 100.

Kent, W. J., Baertsch, R., Hinrichs, A., Miller, W. & Haussler, D. Evolution's cauldron: duplication, deletion, and rearrangement in the mouse and human genomes. *Proc. Natl Acad. Sci. USA* **100**, 11484–11489 (2003).

[CAS](#) [PubMed](#) [ADS](#) [Google Scholar](#)

101. 101.

Kolmogorov, M., Raney, B., Paten, B. & Pham, S. Ragout—a reference-assisted assembly tool for bacterial genomes. *Bioinformatics* **30**, i302–i309 (2014).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

102. 102.

Farré, M. et al. Novel insights into chromosome evolution in birds, archosaurs, and reptiles. *Genome Biol. Evol.* **8**, 2442–2451 (2016).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

103. 103.

Guan, D. Asset. <https://github.com/dfguan/asset>.

104. 104.

Tarailo-Graovac, M. & Chen, N. Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr. Protoc. Bioinformatics* **25**, 4.10.1–4.10.14 (2009).

[Google Scholar](#)

105. 105.

Krumsiek, J., Arnold, R. & Rattei, T. Gepard: a rapid and sensitive tool for creating dotplots on genome scale. *Bioinformatics* **23**, 1026–1028 (2007).

[CAS](#) [PubMed](#) [Google Scholar](#)

106. 106.

Harry, E. PretextView. <https://github.com/wtsi-hpag/PretextView>.

107. 107.

Kurtz, S. et al. Versatile and open software for comparing large genomes. *Genome Biol.* **5**, R12 (2004).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

Nattestad, M. Dot. <https://github.com/MariaNattestad/dot>.

[Download references](#)

Acknowledgements

We thank the following persons for feedback and support: R. Johnson, E. Karlsson, K. Lindblad Toh, W. Jun, I. Korf, W. Haerty, G. Etherington, B. Clavijo, and A. Komissarov for discussions in the early stages of the project; R. Fuller for help with the G10K website maintenance, and H. Segal for help with VGP website development; M. Linh Pham for help with initial grant writing; L. Shalmiyev for administrative help; D. Church, G. Kol, K. Baruch, O. Barad, I. Liachko, E. Muzychenco, S. Garg, and M. Kolmogorov for preliminary analyses performed on one or more genomes; K. Oliver, C. Corton and J. Skelton for data generation; E. Harry for technical support in scaff10x and Pretext; C. Mazzoni for coordinating students and training at Leibniz Institute for Zoo and Wildlife Research and Berlin Center for Genomics in Biodiversity Research; and M. Driller, C. Caswara, M. Vafadar, N. Hill, D. De Panis, A. Whibley, B. Maloney, C. Mitchell, G. Gallo, J. Gaige, K. Amoako-Boadu, M. Jose Gomez, M. Montero, D. Ratnikov, S. Brown, S. Zylka, S. Marcus, and T. Carrasco for completing training and testing the VGP pipeline by producing ordinal representative genome assemblies not described in this manuscript. We thank our company partners (listed below), NCBI, EBI, and Amazon AWS, including AWS for sponsoring sequence storage. J. Fekcs and D. Leja created the animal images, and J. Kim modified them to silhouettes. We thank them for their permission to publish. A.R., S.K., B.P.W. and A.M.P. were supported by the Intramural Research Program of the NHGRI, NIH (1ZIAHG200398). A.R. was also supported by the Korea Health Technology R&D Project through KHIDI, funded by the Ministry of Health & Welfare, Republic of Korea (HI17C2098). S.A.M., I.B. and R.D. were supported by Wellcome Trust grant WT207492; W.C., M. Smith, Z.N., Y.S., J.C., S. Pelan, J.T., A.T., J.W. and Kerstin Howe by WT206194; L.H., F.M., Kevin Howe and P. Flliceck by WT108749/Z/15/Z, WT218328/B/19/Z and the European Molecular Biology Laboratory. O.F. and E.D.J. were

supported by Howard Hughes Medical Institute and Rockefeller University start-up funds for this project. J.D. and H.A.L. were supported by the Robert and Rosabel Osborne Endowment. M.U.-S. received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement (750747). F.T.-N., J. Hoffman, P. Masterson and K.C. were supported by the Intramural Research Program of the NLM, NIH. C.L., B.J.K., J. Kim and H.K. were supported by the Marine Biotechnology Program of KIMST, funded by the Ministry of Ocean and Fisheries, Republic of Korea (20180430). M.C. was supported by Sloan Research Fellowship (FG-2020-12932). S.C.V. was funded by a Max Planck Research Group award from the Max Planck Society, and a Human Frontiers Science Program (HFSP) Research grant (RGP0058/2016). T.M.L., W.E.J. and the Canada lynx genome were funded by the Maine Department of Inland Fisheries & Wildlife (F11AF01099), including when W.E.J. held a National Research Council Research Associateship Award at the Walter Reed Army Institute of Research (WRAIR). C.B. was supported by the NSF (1457541 and 1456612). D.B. was funded by The University of Queensland (HFSP - RGP0030/2015). D.I. was supported by Science Exchange Inc. (Palo Alto, CA). H.W.D. was supported by NSF grants (OPP-0132032 ICEFISH 2004 Cruise, PLR-1444167 and OPP-1955368) and the Marine Science Center at Northeastern University (416). G.J.P.N. and the thorny skate genome were funded by Lenfest Ocean Program (30884). M.P. was funded by the German Federal Ministry of Education and Research (01IS18026C). M. Malinsky was supported by an EMBO fellowship (ALTF 456-2016). The following authors' contributions were supported by the NIH: S. Selvaraj (R44HG008118); C.V.M., S.R.F., P.V.L. (R21 DC014432/DC/NIDCD); K.D.M. (R01GM130691); H.C. (5U41HG002371-19); M.D. (U41HG007234); and B.P. (R01HG010485). D.G. was supported by the National Key Research and Development Program of China (2017YFC1201201, 2018YFC0910504 and 2017YFC0907503). F.O.A. was supported by Al-Gannas Qatari Society and The Cultural Village Foundation-Katara, Doha, State of Qatar and Monash University Malaysia. C.T. was supported by The Rockefeller University. M. Hiller was supported by the LOEWE-Centre for Translational Biodiversity Genomics (TBG) funded by the Hessen State Ministry of Higher Education, Research and the Arts (HMWK). H.C. was supported by the NHGRI (5U41HG002371-19).

R.H.S.K. was funded by the Max Planck Society with computational resources at the bwUniCluster and BinAC funded by the Ministry of Science, Research and the Arts Baden-Württemberg and the Universities of the State of Baden-Württemberg, Germany (bwHPC-C5). B.V. was supported by the Biomedical Research Council of A*STAR, Singapore. T.M.-B. was funded by the European Research Council under the European Union's Horizon 2020 research and innovation programme (864203), MINECO/FEDER, UE (BFU2017-86471-P), Unidad de Excelencia María de Maeztu, AEI (CEX2018-000792-M), a Howard Hughes International Early Career award, Obra Social "La Caixa" and Secretaria d'Universitats i Recerca and CERCA Programme del Departament d'Economia i Coneixement de la Generalitat de Catalunya (GRC 2017 SGR 880). E.C.T. was supported by the European Research Council (ERC-2012-StG311000) and an Irish Research Council Laureate Award. M.T.P.G. was supported by an ERC Consolidator Award 681396-Extinction Genomics, and a Danish National Research Foundation Center Grant (DNRF143). T.W. was supported by the NSF (1458652). J. M. Graves was supported by the Australian Research Council (CEO561477). E.W.M. was partially supported by the German Federal Ministry of Education and Research (01IS18026C). Complementary sequencing support for the Anna's hummingbird and several genomes was provided by Pacific Biosciences, Bionano Genomics, Dovetail Genomics, Arima Genomics, Phase Genomics, 10X Genomics, NRGene, Oxford Nanopore Technologies, Illumina, and DNAexus. All other sequencing and assembly were conducted at the Rockefeller University, Sanger Institute, and Max Planck Institute Dresden genome labs. Part of this work used the computational resources of the NIH HPC Biowulf cluster (<https://hpc.nih.gov>). We acknowledge funding from the Wellcome Trust (108749/Z/15/Z) and the European Molecular Biology Laboratory. We thank Le Comité Scientifique Régional du Patrimoine Naturel and Direction de l'Environnement, de l'Aménagement et du Logement, Guyanne for research approvals and export permits.

Author information

Author notes

1. These authors contributed equally: Arang Rie, Shane A. McCarthy, Olivier Fedrigo
2. These authors jointly supervised this work: Kerstin Howe, Eugene W. Myers, Richard Durbin, Adam M. Phillippy, Erich D. Jarvis

Affiliations

1. Genome Informatics Section, Computational and Statistical Genomics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA

Arang Rie, Sergey Koren, Brian P. Walenz & Adam M. Phillippy

2. Department of Genetics, University of Cambridge, Cambridge, UK

Shane A. McCarthy, Iliana Bista, Dengfeng Guan & Richard Durbin

3. Wellcome Sanger Institute, Cambridge, UK

Shane A. McCarthy, William Chow, Iliana Bista, Michelle Smith, Milan Malinsky, Zemin Ning, Ying Sims, Joanna Collins, Sarah Pelan, James Torrance, Alan Tracey, Jonathan Wood, Kerstin Howe & Richard Durbin

4. Vertebrate Genome Lab, The Rockefeller University, New York, NY, USA

Olivier Fedrigo, Giulio Formenti, Bettina Haase, Jacquelyn Mountcastle, Sadie Paez & Erich D. Jarvis

5. The Genome Center, University of California Davis, Davis, CA, USA

Joana Damas & Harris A. Lewin

6. Laboratory of Neurogenetics of Language, The Rockefeller University, New York, NY, USA

Giulio Formenti, Gregory L. Gedman, Lindsey J. Cantin, Sadie Paez, Matthew T. Biegler, Constantina Theofanopoulou & Erich D. Jarvis

7. Leibniz Institute for Zoo and Wildlife Research, Department of Evolutionary Genetics, Berlin, Germany

Marcela Uliano-Silva

8. Berlin Center for Genomics in Biodiversity Research, Berlin, Germany

Marcela Uliano-Silva

9. DNAnexus Inc., Mountain View, CA, USA

Akarachai Fungtammasan, Maria Simbirsky & Brett T. Hannigan

10. Interdisciplinary Program in Bioinformatics, Seoul National University, Seoul, Republic of Korea

Juwan Kim, Chul Lee & Heebal Kim

11. Department of Agricultural Biotechnology and Research Institute of Agriculture and Life Sciences, Seoul National University, Seoul, Republic of Korea

Byung June Ko & Heebal Kim

12. University of Southern California, Los Angeles, CA, USA

Mark Chaisson & Robel E. Dagnaw

13. National Center for Biotechnology Information, National Library of Medicine, NIH, Bethesda, MD, USA

Francoise Thibaud-Nissen, Jinna Hoffman, Patrick Masterson & Karen Clark

14. European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Genome Campus, Hinxton, UK

Leanne Haggerty, Fergal Martin, Kevin Howe & Paul Flicek

15. Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany

Sylke Winkler, Martin Pippel, Ekaterina Osipova & Eugene W. Myers

16. DRESDEN-concept Genome Center, Dresden, Germany

Sylke Winkler

17. Novogene, Durham, NC, USA

Jason Howard

18. Neurogenetics of Vocal Communication Group, Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands

Sonja C. Vernes

19. Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands

Sonja C. Vernes

20. School of Biology, University of St Andrews, St Andrews, UK

Sonja C. Vernes

21. University of Massachusetts Cooperative Fish and Wildlife Research Unit, Amherst, MA, USA

Tanya M. Lama

22. School of Biological Science, The Environment Institute, University of Adelaide, Adelaide, South Australia, Australia

Frank Grutzner

23. Bond Life Sciences Center, University of Missouri, Columbia, MO, USA

Wesley C. Warren

24. Department of Biology, East Carolina University, Greenville, NC,
USA

Christopher N. Balakrishnan

25. UQ Genomics, University of Queensland, Brisbane, Queensland,
Australia

Dave Burt

26. Department of Biological Sciences, Clemson University, Clemson, SC,
USA

Julia M. George

27. The Genetic Rescue Foundation, Wellington, New Zealand

David Iorns

28. Kākāpō Recovery, Department of Conservation, Invercargill, New
Zealand

Andrew Digby & Daryl Eason

29. Department of Zoology, University of Otago, Dunedin, New Zealand

Bruce Robertson

30. University of Arizona Genetics Core, Tucson, AZ, USA

Taylor Edwards

31. Department of Life Sciences, Natural History Museum, London, UK

Mark Wilkinson

32. School of Natural Sciences, Bangor University, Gwynedd, UK

George Turner

33. Department of Biology, University of Konstanz, Konstanz, Germany
Axel Meyer, Andreas F. Kautt, Paolo Franchini & Robert H. S. Kraus
34. Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA
Andreas F. Kautt
35. Department of Marine and Environmental Sciences, Northeastern University Marine Science Center, Nahant, MA, USA
H. William Detrich III
36. Department of Biology, University of Antwerp, Antwerp, Belgium
Hannes Svardal
37. Naturalis Biodiversity Center, Leiden, The Netherlands
Hannes Svardal
38. Institute of Biology, Karl-Franzens University of Graz, Graz, Austria
Maximilian Wagner
39. Florida Museum of Natural History, University of Florida, Gainesville, FL, USA
Gavin J. P. Naylor
40. Center for Systems Biology, Dresden, Germany
Martin Pippel, Ekaterina Osipova & Eugene W. Myers
41. Zoological Institute, University of Basel, Basel, Switzerland
Milan Malinsky

42. Tag bio, San Francisco, CA, USA

Mark Mooney

43. UC Santa Cruz Genomics Institute, University of California, Santa Cruz, CA, USA

Trevor Pesout, Richard E. Green, Erik Garrison, Hiram Clawson, Mark Diekhans, Luis Nassar, Benedict Paten & David Haussler

44. San Diego Zoo Global, Escondido, CA, USA

Marlys Houck, Ann Misuraca & Oliver A. Ryder

45. Pacific Biosciences, Menlo Park, CA, USA

Sarah B. Kingan, Richard Hall, Zev Kronenberg, Ivan Sović, Christopher Dunn & Jonas Korlach

46. Digital BioLogic, Ivanić-Grad, Croatia

Ivan Sović

47. Bionano Genomics, San Diego, CA, USA

Alex Hastie & Joyce Lee

48. Arima Genomics, San Diego, CA, USA

Siddarth Selvaraj

49. Dovetail Genomics, Santa Cruz, CA, USA

Richard E. Green & Jay Ghurye

50. Independent Researcher, Santa Cruz, CA, USA

Nicholas H. Putnam

51. CNAG-CRG, Centre for Genomic Regulation, Barcelona Institute of Science and Technology, Barcelona, Spain

Ivo Gut

52. Universitat Pompeu Fabra, Barcelona, Spain

Ivo Gut

53. Department of Computer Science, University of Maryland College Park, College Park, MD, USA

Jay Ghurye

54. School of Computer Science and Technology, Center for Bioinformatics, Harbin Institute of Technology, Harbin, China

Dengfeng Guan

55. Department of Psychology, Institute for Mind and Biology, University of Chicago, Chicago, IL, USA

Sarah E. London

56. Department of Genetics and Biochemistry, Clemson University, Clemson, SC, USA

David F. Clayton

57. Department of Behavioral Neuroscience, Oregon Health and Science University, Portland, OR, USA

Claudio V. Mello, Samantha R. Friedrich & Peter V. Lovell

58. Max Planck Institute for the Physics of Complex Systems, Dresden, Germany

Ekaterina Osipova

59. Monash University Malaysia Genomics Facility, School of Science, Selangor Darul Ehsan, Malaysia

Farooq O. Al-Ajli

60. Tropical Medicine and Biology Multidisciplinary Platform, Monash University Malaysia, Selangor Darul Ehsan, Malaysia

Farooq O. Al-Ajli

61. Qatar Falcon Genome Project, Doha, Qatar

Farooq O. Al-Ajli

62. Department of Biosciences, University of Milan, Milan, Italy

Simona Secomandi

63. eGnome, Inc., Seoul, Republic of Korea

Heebal Kim & Woori Kwak

64. LOEWE Centre for Translational Biodiversity Genomics, Frankfurt, Germany

Michael Hiller

65. Senckenberg Research Institute, Frankfurt, Germany

Michael Hiller

66. Goethe-University, Faculty of Biosciences, Frankfurt, Germany

Michael Hiller

67. BGI-Shenzhen, Shenzhen, China

Yang Zhou

68. Department of Biology, Pennsylvania State University, University Park, PA, USA

Robert S. Harris & Kateryna D. Makova

69. Center for Medical Genomics, Pennsylvania State University, University Park, PA, USA

Kateryna D. Makova & Paul Medvedev

70. Center for Computational Biology and Bioinformatics, Pennsylvania State University, University Park, PA, USA

Kateryna D. Makova & Paul Medvedev

71. Department of Computer Science and Engineering, Pennsylvania State University, University Park, PA, USA

Paul Medvedev

72. Department of Biochemistry and Molecular Biology, Pennsylvania State University, University Park, PA, USA

Paul Medvedev

73. Hoonygen, Seoul, Korea

Woori Kwak

74. Department of Migration, Max Planck Institute of Animal Behavior, Radolfzell, Germany

Robert H. S. Kraus

75. Department of Biological Sciences, Universidad de los Andes, Bogotá, Colombia

Andrew J. Crawford

76. Center for Evolutionary Hologenomics, The GLOBE Institute,
University of Copenhagen, Copenhagen, Denmark

M. Thomas P. Gilbert

77. University Museum, NTNU, Trondheim, Norway

M. Thomas P. Gilbert

78. China National Genebank, BGI-Shenzhen, Shenzhen, China

Guojie Zhang

79. Villum Center for Biodiversity Genomics, Section for Ecology and
Evolution, Department of Biology, University of Copenhagen,
Copenhagen, Denmark

Guojie Zhang

80. State Key Laboratory of Genetic Resources and Evolution, Kunming
Institute of Zoology, Chinese Academy of Sciences, Kunming, China

Guojie Zhang

81. Center for Excellence in Animal Evolution and Genetics, Chinese
Academy of Sciences, Kunming, China

Guojie Zhang

82. Institute of Molecular and Cell Biology, A*STAR, Biopolis,
Singapore, Singapore

Byrappa Venkatesh

83. Centre for Biodiversity, Royal Ontario Museum, Toronto, Ontario,
Canada

Robert W. Murphy

84. Smithsonian Conservation Biology Institute, Center for Species Survival, National Zoological Park, Washington, DC, USA

Klaus-Peter Koepfli & Warren E. Johnson

85. Department of Ecology and Evolutionary Biology, University of California Santa Cruz, Santa Cruz, CA, USA

Beth Shapiro & David Haussler

86. Howard Hughes Medical Institute, Chevy Chase, MD, USA

Beth Shapiro & Erich D. Jarvis

87. The Walter Reed Biosystematics Unit, Museum Support Center MRC-534, Smithsonian Institution, Suitland, MD, USA

Warren E. Johnson

88. Walter Reed Army Institute of Research, Silver Spring, MD, USA

Warren E. Johnson

89. Department of Biological Sciences, Earlham Institute, University of East Anglia, Norwich, UK

Federica Di Palma

90. Institute of Evolutionary Biology (UPF-CSIC), PRBB, Barcelona, Spain

Tomas Marques-Bonet

91. Catalan Institution of Research and Advanced Studies (ICREA), Barcelona, Spain

Tomas Marques-Bonet

92. Centre for Genomic Regulation (CRG), Barcelona Institute of Science and Technology (BIST), Barcelona, Spain

Tomas Marques-Bonet

93. Institut Català de Paleontologia Miquel Crusafont, Universitat Autònoma de Barcelona, Barcelona, Spain

Tomas Marques-Bonet

94. School of Biology and Environmental Science, University College Dublin, Dublin, Ireland

Emma C. Teeling

95. Department of Computer Science, The University of Illinois at Urbana-Champaign, Urbana, IL, USA

Tandy Warnow

96. School of Life Science, La Trobe University, Melbourne, Victoria, Australia

Jennifer Marshall Graves

97. Department of Evolution, Behavior, and Ecology, University of California San Diego, La Jolla, CA, USA

Oliver A. Ryder

98. Laboratory of Genomics Diversity-Center for Computer Technologies, ITMO University, St. Petersburg, Russian Federation

Stephen J. O'Brien

99. Guy Harvey Oceanographic Center, Halmos College of Natural Sciences and Oceanography, Nova Southeastern University, Fort Lauderdale, FL, USA

Stephen J. O'Brien

100. Department of Evolution and Ecology, University of California Davis, Davis, CA, USA

Harris A. Lewin

101. John Muir Institute for the Environment, University of California Davis, Davis, CA, USA

Harris A. Lewin

102. Faculty of Computer Science, Technical University Dresden, Dresden, Germany

Eugene W. Myers

Authors

1. Arang Rhei

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Shane A. McCarthy

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Olivier Fedrigo

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Joana Damas

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Giulio Formenti

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Sergey Koren

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Marcela Uliano-Silva

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. William Chow

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. Arkarachai Fungtammasan

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. Juwan Kim

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

11. Chul Lee

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

12. Byung June Ko

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

13. Mark Chaisson

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

14. Gregory L. Gedman

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

15. Lindsey J. Cantin

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

16. Francoise Thibaud-Nissen

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

17. Leanne Haggerty

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

18. Iliana Bista

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

19. Michelle Smith

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

20. Bettina Haase

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

21. Jacquelyn Mountcastle
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

22. Sylke Winkler
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

23. Sadye Paez
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

24. Jason Howard
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

25. Sonja C. Vernes
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

26. Tanya M. Lama
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

27. Frank Grutzner
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

28. Wesley C. Warren
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

29. Christopher N. Balakrishnan

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

30. Dave Burt

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

31. Julia M. George

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

32. Matthew T. Biegler

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

33. David Iorns

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

34. Andrew Digby

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

35. Daryl Eason

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

36. Bruce Robertson

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

37. Taylor Edwards

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

38. Mark Wilkinson

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

39. George Turner

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

40. Axel Meyer

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

41. Andreas F. Kautt

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

42. Paolo Franchini

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

43. H. William Detrich III

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

44. Hannes Svardal

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

45. Maximilian Wagner

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

46. Gavin J. P. Naylor

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

47. Martin Pippel

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

48. Milan Malinsky

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

49. Mark Mooney

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

50. Maria Simbirsky

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

51. Brett T. Hannigan

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

52. Trevor Pesout

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

53. Marlys Houck

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

54. Ann Misuraca

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

55. Sarah B. Kingan

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

56. Richard Hall

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

57. Zev Kronenberg

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

58. Ivan Sović

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

59. Christopher Dunn

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

60. Zemin Ning

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

61. Alex Hastie

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

62. Joyce Lee

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

63. Siddarth Selvaraj

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

64. Richard E. Green

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

65. Nicholas H. Putnam

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

66. Ivo Gut

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

67. Jay Ghurye

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

68. Erik Garrison

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

69. Ying Sims

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

70. Joanna Collins

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

71. Sarah Pelan

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

72. James Torrance

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

73. Alan Tracey

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

74. Jonathan Wood

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

75. Robel E. Dagnew

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

76. Dengfeng Guan

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

77. Sarah E. London

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

78. David F. Clayton

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

79. Claudio V. Mello

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

80. Samantha R. Friedrich

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

81. Peter V. Lovell

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

82. Ekaterina Osipova

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

83. Farooq O. Al-Ajli

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

84. Simona Secomandi

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

85. Heebal Kim

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

86. Constantina Theofanopoulou

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

87. Michael Hiller

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

88. Yang Zhou

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

89. Robert S. Harris

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

90. Kateryna D. Makova

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

91. Paul Medvedev

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

92. Jinna Hoffman

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

93. Patrick Masterson

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

94. Karen Clark

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

95. Fergal Martin

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

96. Kevin Howe

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

97. Paul Flicek

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

98. Brian P. Walenz

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

99. Woori Kwak

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

100. Hiram Clawson

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

101. Mark Diekhans

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

102. Luis Nassar

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

103. Benedict Paten

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

104. Robert H. S. Kraus

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

105. Andrew J. Crawford

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

106. M. Thomas P. Gilbert

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

107. Guojie Zhang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

108. Byrappa Venkatesh

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

109. Robert W. Murphy
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

110. Klaus-Peter Koepfli
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

111. Beth Shapiro
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

112. Warren E. Johnson
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

113. Federica Di Palma
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

114. Tomas Marques-Bonet
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

115. Emma C. Teeling
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

116. Tandy Warnow
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

117. Jennifer Marshall Graves
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

118. Oliver A. Ryder
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

119. David Haussler
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

120. Stephen J. O'Brien
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

121. Jonas Korlach
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

122. Harris A. Lewin
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

123. Kerstin Howe
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

124. Eugene W. Myers
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

125. Richard Durbin

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

126. Adam M. Phillippy

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

127. Erich D. Jarvis

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Contributions

Wrote the paper and co-coordinated the study: A.R., E.D.J., A.M.P., R.D., E.W.M., Kerstin Howe, S.A.M., O.F. Coordination with vendors: J. Korlach, S. Selvaraj, R.E.G., A.H., M. Mooney. Collected samples: M.T.P.G., W.E.J., R.W.M., G.Z., B.V., M.T.B., J. Howard, S.C.V., T.M.L., F.G., W.C.W., D.B., J. M. George, M.T.B., D.I., A.D., D.E., B.R., T.E., M. Wilkinson, G.T., A. Meyer, A.F.K., P. Franchini, H.W.D., H.S., M. Wagner, G.J.P.N., R.D., E.D.J., E.C.T., R.H.S.K. Generated genome data: O.F., I.B., M. Smith, B.H., J.M., S.W., C.B., A. Meyer, A.F.K., P. Franchini, I.G., D.F.C., C.V.M. Generated genome assemblies: A.R., S.A.M., S.K., M.P., S.B.K., R.H., J.G., Z.N., J.L., B.P.W., M. Malinsky. Generated/modified software: S.K., A.R., S.B.K., R.H., Z.K., J. Korlach, I.S., C.D., Z.N., A.H., J.L., J.G., E.G., C.V.M., S.R.F., N.H.P. Pipeline development: A.R., S.A.M., G.F., S.K., M.U.-S., A.F., M. Simbirsky, B.T.H., T.P., M.P., E.W.M., R.D., A.M.P. Generated MT assemblies: G.F., J. Korlach. Curation: Kerstin Howe, W.C., Y.S., J.C., S. Pelan, J.T., A.T., J.W., Y.Z., J.D., H.A.L. Sex chromosomes: Y.Z., R.S.H., K.D.M., P. Medvedev, J. M. Graves. Hummingbird karyotype analyses: M. Houck, A. Misuraca, M.P., E.W.M., E.D.J. Annotation: F.T.-N., L.H., J. Hoffman, P. Masterson, K.C., F.M., Kevin Howe, P. Fliecek, D.B. Evaluation analysis: A.R., J.D., M.U.-S., J. Kim, C.L., B.J.K., M.C., G.L.G., L.J.C., F.T.-N., L.H., J. M. George, J.G., R.E.D., D.G., S.E.L., D.F.C., C.V.M., S.R.F., P.V.L., E.O., F.O.A.-A., S.

Secomandi, C.T., M. Hiller, H.K., Kerstin Howe, E.W.M., R.D., A.M.P., E.D.J. Biological findings: J.D., J. Kim, C.L., B.J.K., G.L.G., L.J.C., H.A.L., A.R., E.D.J. Data availability: A.R., S.A.M., W.C., A.F., S. Paez, M. Simbirsky, B.T.H., B.P.W., W.K., H.C., M.D., L.N., B.P., A.M.P., E.D.J. G10K council, founders, and coordination of VGP: T.M.-B., A.J.C., F.D.P., R.D., M.T.P.G., E.D.J., K.-P.K., H.A.L., R.W.M., E.W.M., E.C.T., B.V., G.Z., A.M.P., S. Paez, J. M. Graves, O.A.R., D.H., S.J.O., T.W. and B.S. All authors reviewed the manuscript.

Corresponding authors

Correspondence to [Kerstin Howe](#) or [Eugene W. Myers](#) or [Richard Durbin](#) or [Adam M. Phillippy](#) or [Erich D. Jarvis](#).

Ethics declarations

Competing interests

During the contributing period, B.T.H., M. Simbirsky, A.F. and M. Mooney were employees of DNAnexus Inc. S.B.K., R.H., Z.K., J. Korlach, I.S. and C.D. were full-time employees at Pacific Biosciences, a company developing single-molecule long read sequencing technologies. R.E.G., N.H.P., and J.G. were affiliated with Dovetail Genomics, a company developing genome assembly tools, including Hi-C. I.G. was affiliated with Oxford Nanopore Technologies, a company generating long read sequencing technologies. A.H. and J.L. were employees of Bionano Genomics, a company developing optical maps for genome assembly. S. Selvaraj was an employee of Arima Genomics, a company developing Hi-C data for genome assemblies. R.D. is a scientific advisory board member of Dovetail Inc. P. Flicek is a member of the Scientific Advisory Boards of Fabric Genomics, Inc., and Eagle Genomics, Ltd. H.C. receives royalties from the sale of UCSC Genome Browser source code, LiftOver, GBiB, and GBiC licenses to commercial entities. S.K. has received travel funds to speak at symposia organized by Oxford Nanopore. M.D. and L.N. receive royalties from licensing of UCSC Genome Browser. For W.E.J., the content here is not to be construed as the views of the DA or DOD. All other authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Michael Schatz, Justin Zook 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 Assessment of completeness of the Anna's hummingbird assembly.

a, b, Steps and NG50 continuity values of the VGP assembly pipeline that gave the highest quality assembly for Anna's hummingbird (**a**) and Canada lynx (**b**) in this study. The specific steps are outlined further in Extended Data Fig. [2a](#), and Methods. **c**, Whole-genome alignment of CLR (red), linked reads (green), optical maps (blue), and Hi-C reads (purple) of the Anna's hummingbird, along with telomere motif (TTAGGG and its reverse complement, yellow) and gaps (grey) using Asset software^{[103](#)}. For each data type, the first row shows the mapped coverage, and the second shows the number of counts of low coverage or signs of collapsed repeats. Larger chromosomal scaffolds (1–19) have fewer gaps and low coverage or collapsed regions compared with the micro chromosomes (20–33). Chromosomes 14, 15 and 19 of the Anna's hummingbird were the most structurally reliable scaffolds, having only one gap each with no low-support regions. We defined reliable blocks as those supported by at least two technologies. Reliable blocks excluded regions with structural assembly errors, such as collapsed repeats or unresolved segmental duplications. Low-support regions are those where the reliable blocks row has a peak.

Extended Data Fig. 2 VGP assembly pipeline applied across multiple species.

a, Iterative assembly pipeline of sequence data types (coloured as in **b**) with increasing chromosomal distance. Thin bars, sequence reads; thick black bars, assembled contigs; black bars with space and arcing links, scaffolds; grey bars, gaps placed by previous steps; thick red border, tracking of an example contig in the pipeline. The curation step shows an example of a mis-assembly break identified by sequence coverage (grey, left) and an example of an inversion error (right) detected by the optical map. **b**, Intra-molecule length distribution of the four data types used to generate the assemblies of 16 vertebrate species, weighted by the fraction of bases in each length bin (log scaled). Molecule length above 1 kb was measured from read length for CLR, estimated molecule coverage for linked reads,

raw molecule length for optical maps, and interaction distance for Hi-C reads. For each species, the fragment length distribution of each data type was similar to those for the Anna's hummingbird, with differences primarily influenced by tissue type, preservation method, and collection or storage conditions (unpublished data).

Extended Data Fig. 3 Flow charts of assembly pipelines used to generate high-quality assemblies in this study.

a, Standard VGP assembly pipeline when sequencing data of one individual, that generated the highest quality assemblies: generate primary pseudo-haplotype and alternate haplotype contigs with CLR using FALCON-Unzip¹⁷; generate scaffolds with linked reads using Scaff10x⁷⁴; break mis-joins and further scaffold with optical maps using Solve⁸⁷; generate chromosome-scale scaffolds with Hi-C reads using Salsa2⁷⁹; fill in gaps and polish base-errors with CLR using Arrow (Pacific BioSciences); perform two or more rounds of short-read polishing with linked reads using FreeBayes⁸⁵; and perform expert manual curation to correct potential assembly errors using gEVAL^{25,95} **b**, Standard VGP trio assembly pipeline when DNA is available for a child and parents²⁰. Dashed line indicates that the other haplotype went through the same steps before curation. In addition to the curated assemblies of both haplotypes, a representative haplotype with both sex chromosomes is submitted. **c**, Mitochondrial assembly pipeline. Figure key applies to **a–c**. Steps newly introduced in v1.5–v1.6 are highlighted in light blue. c, contigs; p, purged false duplications from primary contigs; q, purged alternate contigs; s, scaffolds; t, polished scaffolds. Further details and instructions are available elsewhere³³ and at <https://github.com/VGP/vgp-assembly>.

Extended Data Fig. 4 Relationship between collapses and genomic characteristics.

a, Correlation between the total number of collapses and percentage repeat content estimated in the submitted curated versions of $n = 17$ genomes from 16 species. **b**, Correlation between total number of bases in collapsed regions per Gb and repeat content. **c**, Correlation between total missing bases collapsed per Gb and repeat content. **d**, Correlation between total number of genes (coding and non-coding) in the collapsed regions and repeat content. **e**, Lack of correlation between the average collapsed size and repeat content. **f**, Lack of correlation between the total number of collapses and percentage heterozygosity. **g**, Lack of correlation between the

total number of collapses per Gb and genome size. Genome size, heterozygosity, and repeat content were estimated from 31-mer counts using GenomeScope⁷¹. Reported are adjusted r^2 and P values from F -statistics. **h**, Cumulative collapsed bases per Gb in each collapse and percentage repeat masked. Each circle is coloured by species with its size relative to the length of the collapse as it appears in the assembly. Collapses above the horizontal bar (>90%) are further classified as collapsed high-copy repeats, and those below the horizontal bar are classified as segmental duplications (low-copy repeats). **i**, Major repeat types in collapsed high-copy repeats. Most of the repeats were masked only with WindowMasker⁷⁵, with no annotation available by RepeatMasker¹⁰⁴. **j**, Minor repeat types in collapsed repeats. This is a breakdown of the repeats categorized as ‘Others’ in **i**, owing to the smaller scale. Bar colours in **i** and **j** are as in **h**. Note smaller scale in **j** compared with **i**. Collapsed satellite arrays were almost exclusively found in the platypus, comprising ~2.5 Mb. Collapsed simple repeats were the major source in the thorny skate (~400 kb). There was a higher proportion of LTRs in birds, LINEs and SINEs in mammals, and DNA repeats in the amphibian. Among the genes in the collapses, many were repetitive short non-coding RNAs. P values from F -statistics.

Extended Data Fig. 5 False duplication mechanisms in genome assembly.

a, False heterotype (haplotype) duplications occurs when more divergent sequence reads from each haplotype A (blue) and B (red) (maternal and paternal) form greater divergent paths in the assembly graph (bubbles), while nearly identical homozygous sequences (black) become collapsed. When the assembly graph is properly formed and correctly resolved (green arrow), one of the haplotype-specific paths (red or blue) is chosen for building a ‘primary’ pseudo-haplotype assembly and the other is set apart as an ‘alternate’ assembly. When the graph is not correctly resolved (purple arrow), one of four types of pattern are formed in the contigs and subsequent scaffolds. Depending on the supporting evidence, the scaffolder either keeps these haplotype contigs on separate scaffolds or brings them together on the same scaffold, often separated by gaps: 1. Separate contigs: both contigs are retained in the primary contig set, an error often observed when haplotype-specific sequences are highly diverged. 2. Flanking

contigs: the assembly graph is partially formed, connecting the homozygous sequence of the 5' side to one haplotype (blue) and the 3' side to the other haplotype (red). 3. Partial flanking contigs: only one haplotype (blue) flanks one side of the homozygous sequence. 4. Failed connecting of contigs: all haplotype sequences fail to properly connect to flanking homozygous sequences. **b**, False homotype duplications occur where a sequence from the same genomic locus is duplicated, and are of two types: 1. Overlapping sequences at contig boundaries: in current overlap-layout-consensus assemblers, branching sequences in assembly graphs that are not selected as the primary path have a small overlapping sequence (purple), dovetailing to the primary path where it originated a branch. The size of the duplicated sequence is often the length of a corrected read. Subsequent scaffolding results in tandem duplicated sequences with a gap between. 2. Under-collapsed sequences: sequencing errors in reads (red x) randomly or systematically pile up, forming under-collapsed sequences. Subsequent duplication errors in the scaffolding are similar to the heterotype duplications. Purge_haplotigs¹³ align sequences to themselves to find a smaller sequence that aligns fully to a larger contig or scaffold, and removes heterotype duplication types 1, 3, and 4. Purge_dups¹⁴ additionally uses coverage information to detect heterotype duplication type 2 and homotype duplications. We distinguished the two types of duplications by: 1) haplotype-specific variants in reads aligning at half coverage to each heterotype duplication; 2) differing consensus quality that resulted from read coverage fluctuations when aligning reads to homotype duplications; and 3) *k*-mer copy number anomalies in which homotype duplications were observed in the assembly with more than the expected number of copies.

[Extended Data Fig. 6 False duplication examples fixed during manual curation.](#)

a, An example of a heterotype duplication in the female zebra finch, non-trio assembly. Left, a self-dot plot of this region generated with Gepard¹⁰⁵, with sequences coloured by haplotypes. Gaps, duplicated sequences (green and purple), and haplotype-specific marker densities are indicated at the top. Right, a detailed alignment view of the green haplotype duplication with paternal and maternal markers, self-alignment components, transcripts annotated, contigs, bionano maps, and repeat components displayed in

gEVAL⁹⁵. **b**, Example of a homotype duplication found in the hummingbird assembly. These were caused by an algorithm bug in FALCON, which was later fixed. **c**, Example of a combined duplication involving both heterotype (green) and homotype (orange) duplications. Assembly graph structure is shown on the left for clarity, highlighting the overlapping sites at the contig boundary shaded following the duplication type. Assembly errors including the above false duplications were detected and fixed during the curation process.

Extended Data Fig. 7 Evidence of near-complete chromosome scaffolds in the VGP assemblies.

Shown are Hi-C interaction heat maps for each species after curation, visualized with PretextView¹⁰⁶. A scaffold is considered a putative arm-to-arm chromosome when all Hi-C read pairs in a row and column map to a square (that is, an assembled chromosome) on the diagonal without any other interactions off the diagonal. Those with remaining off-diagonal matches to smaller scaffolds are not linked because of ambiguous order or orientation, and are instead submitted as ‘unlocalized’ belonging to the relevant chromosome. Bands at the top of each heat map show scaffolds identified as X, Z (blue) or Y, W (red) sex chromosomes. The Hi-C map of fAstCal1 is not included as we had no remaining tissue left of the animal used to generate Hi-C reads.

Extended Data Fig. 8 Comparison of chromosomal organization between previous and new VGP assemblies.

a, Zebra finch male compared to a previous reference assembly of the same animal. **b**, Platypus male compared with a previous reference female assembly (so the Y chromosomes are absent in the previous reference). **c**, Hummingbird female compared to a previous reference of the same animal. **d**, Climbing perch compared to a previous reference. Each row represents a VGP-generated chromosome for the target species. Colours depict identity with the reference (see key to the right); more than one colour indicates reorganization in the VGP assembly relative to the reference. The lines within each block depict orientation relative to the reference; a positive

slope is the same orientation as the reference, whereas a negative slope is the inverse orientation. Gaps are white boxes with no lines, in the reference relative to the VGP assembly. A white box for the entire chromosome means a newly identified chromosome in the VGP assembly. Top 20 is the longest 20 scaffolds of the hummingbird and climbing perch assemblies. Accession numbers of the assemblies compared are listed in Supplementary Table 19.

Extended Data Fig. 9 Haplotype-resolved sex chromosomes and mitochondrial genomes.

a, Alignment scatterplot, generated with MUMmer NUCmer¹⁰⁷, visualized with dot¹⁰⁸, of maternal and paternal chromosomes from the female zebra finch trio-based assembly. Blue, same orientation; red, inversion; orange, repeats between haplotypes. The paternal Z chromosome is highly divergent from the maternal W, and thus mostly unaligned. **b**, Alignment scatterplot of assembled Z and W chromosomes across the three bird species, approximated with MashMap2⁸⁶. Segments of 300 kb (green), 500 kb (blue), and 1 Mb (purple) are shaded darker with higher sequence identity, with a minimum of 85%. The smaller size and higher repeat content of the W chromosome are clearly visible. **c**, X and Y chromosome segments of the mammals (platypus, Canada lynx, pale spear-nosed bat, and greater horseshoe bat) showing a higher density of repeats within the mammalian X chromosome than the avian Z chromosome. **d**, VGP kākāpō mitochondrial genome assembly reveals previously missing repetitive sequences (adding 2,232 bp) in the origin of replication region, containing an 83-bp repeat unit. **e**, VGP climbing perch mitochondrial genome assembly showing a duplication of *trnL2* and partial duplication of *Nad1*, which were absent from the prior reference. Orange arrows and red lines, tRNA genes and their alignments; dark grey arrows and grey shading, all other genes and their alignments; black, non-coding regions; green line, conventional starting point of the circular sequence.

Extended Data Fig. 10 Large haplotype inversions with direct evidence in the zebra finch trio assembly.

a, Two inversions (green and red) in chromosome 5 found from the MUMmer NUCmer¹⁰⁷ alignment of the maternal and paternal haplotype assemblies, visualized with dot¹⁰⁸. **b**, Hi-C interaction plot showing that the trio-binned Hi-C data remove most of the interactions from the other haplotype (red arrows), which could be erroneously classified as a mis-assembly if only one haplotype was used as a reference. **c**, An 8.5-Mb inversion found on chromosome 11 and a complicated 8.1-Mb rearrangement on chromosome 13 between maternal and paternal haplotypes. **d**, No mis-assembly signals were detected from the binned Hi-C interaction plots, indicating that the haplotype-specific inversions are real. **e**, Half the PacBio CLR span and Bionano optical maps agree with the inversion breakpoints in chromosome 11, supporting the haplotype-specific inversion.

Extended Data Fig. 11 Polishing artefacts.

a, An example of uneven mapping coverage in the primary and alternate sequence pair of the Anna's hummingbird assembly. In this example, the alternate (alt) sequence was built at higher quality, attracting all linked-reads for polishing. The matching locus in the primary (pri) assembly was left unpolished, resulting in frameshift errors in the *TLK1* gene. **b**, Haplotype-specific markers (red for maternal, blue for paternal) and error markers found in the assembly on the Z chromosome (inherited from the paternal side) of the trio-binned female zebra finch assembly. Each row shows markers before short-read polishing, mapping all reads to both haplotype assemblies, and polishing by mapping paternally binned reads to the paternal assembly. Polishing improves QV, but introduces haplotype switch errors when using reads from both haplotypes as shown in row 2. This can be avoided when using haplotype binned reads for polishing. **c**, Example of over-polishing. The nuclear mitochondria (NuMT) sequence was transformed as a full mitochondria (MT) sequence during long-read polishing owing to the absence of the MT contig, where the NuMT attracted all long reads from the MT. In comparison, the trio-binned assembly had the MT sequence assembled in place, preventing mis-placing of MT reads during read mapping.

Extended Data Fig. 12 Chromosome evolution among the bat species sequenced.

a, Genes surrounding an inversion in the greater horseshoe bat, relative to human chromosome 15 (red highlight). The *STARD5* gene is directly disrupted by this inversion, which separates exons 1–5 from exon 6 in the greater horseshoe bat. **b**, RNA-seq tracks showing the lack of RNA splicing evidence of *STARD5* transcripts in the greater horseshoe bat (bottom) in comparison to the pale spear-nosed bat where the *STARD5* gene is not disrupted (top). **c**, Circos plots of chromosome organization relationships between the each of the analysed bats and segments of the human chromosomes 1, 2, 6 and 10. Red star, breakpoint location in human chromosome 6, depicting the fission of the boreoeutherian chromosome 5 in the bat ancestor; blue star, the region upstream of the breakpoint in the bats; green star, the region downstream of the breakpoint in the bats. The red starred breakpoint was confirmed as reused, as opposed to assembly errors, in chromosomal rearrangements of the pale spear-nosed bat, Egyptian fruit bat, and greater horseshoe bat. There is no evidence of reuse for the velvety free-tailed bat. We could not confirm breakpoint reuse in the greater mouse-eared bat or Kuhl's pipistrelle at the chromosomal scale because they were on small scaffolds that may not be completely assembled.

Extended Data Table 1 Summary metrics of the curated and submitted vertebrate species assemblies

[Full size table](#)

Extended Data Table 2 Annotation summary statistics in previous and newly assembled VGP reference genomes

[Full size table](#)

Supplementary information

Supplementary Information

This file contains Supplementary text, Supplementary Notes 1-7, Supplementary Figures 1-6 and Supplementary references.

Reporting Summary

Supplementary Tables

This file contains Supplementary Tables 1-23.

Peer Review File

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

Reprints and Permissions

About this article



Check for
updates

Cite this article

Rhie, A., McCarthy, S.A., Fedrigo, O. *et al.* Towards complete and error-free genome assemblies of all vertebrate species. *Nature* **592**, 737–746

(2021). <https://doi.org/10.1038/s41586-021-03451-0>

[Download citation](#)

- Received: 22 May 2020
- Accepted: 12 March 2021
- Published: 28 April 2021
- Issue Date: 29 April 2021
- DOI: <https://doi.org/10.1038/s41586-021-03451-0>

Further reading

- [Complete vertebrate mitogenomes reveal widespread repeats and gene duplications](#)
 - Giulio Formenti
 - , Arang Rhie
 - , Jennifer Balacco
 - , Bettina Haase
 - , Jacquelyn Mountcastle
 - , Olivier Fedrigo
 - , Samara Brown
 - , Marco Rosario Capodiferro
 - , Farooq O. Al-Ajli
 - , Roberto Ambrosini
 - , Peter Houde
 - , Sergey Koren
 - , Karen Oliver
 - , Michelle Smith
 - , Jason Skelton
 - , Emma Betteridge
 - , Jale Dolucan
 - , Craig Corton

- , Iliana Bista
- , James Torrance
- , Alan Tracey
- , Jonathan Wood
- , Marcela Ulianó-Silva
- , Kerstin Howe
- , Shane McCarthy
- , Sylke Winkler
- , Woori Kwak
- , Jonas Körlich
- , Arkarachai Fungtammasan
- , Daniel Fordham
- , Vania Costa
- , Simon Mayes
- , Matteo Chiara
- , David S. Horner
- , Eugene Myers
- , Richard Durbin
- , Alessandro Achilli
- , Edward L. Braun
- , Adam M. Phillippy
- & Erich D. Jarvis

Genome Biology (2021)

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Download PDF](#)

Associated Content

Special

Vertebrate Genomes Project

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03451-0>

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

Universal nomenclature for oxytocin–vasotocin ligand and receptor families

[Download PDF](#)

- Article
- Open Access
- [Published: 28 April 2021](#)

Universal nomenclature for oxytocin–vasotocin ligand and receptor families

- [Constantina Theofanopoulou](#) ORCID: orcid.org/0000-0003-2014-7563^{1,2,3},
- [Gregory Gedman](#)¹,
- [James A. Cahill](#)¹,
- [Cedric Boeckx](#) ORCID: orcid.org/0000-0001-8882-9718^{2,3,4} &
- [Erich D. Jarvis](#) ORCID: orcid.org/0000-0001-8931-5049^{1,5}

[Nature](#) volume 592, pages 747–755(2021) [Cite this article](#)

- 2438 Accesses
- 1 Citations
- 141 Altmetric
- [Metrics details](#)

Subjects

- [Comparative genomics](#)
- [Evolutionary genetics](#)
- [Phylogenetics](#)

Abstract

Oxytocin (OXT; hereafter OT) and arginine vasopressin or vasotocin (AVP or VT; hereafter VT) are neurotransmitter ligands that function through specific receptors to control diverse functions^{1,2}. Here we performed genomic analyses on 35 species that span all major vertebrate lineages, including newly generated high-contiguity

assemblies from the Vertebrate Genomes Project^{3,4}. Our findings support the claim⁵ that *OT* (also known as *OXT*) and *VT* (also known as *AVP*) are adjacent paralogous genes that have resulted from a local duplication, which we infer was through DNA transposable elements near the origin of vertebrates and in which *VT* retained more of the parental sequence. We identified six major oxytocin–vasotocin receptors among vertebrates. We propose that all six of these receptors arose from a single receptor that was shared with the common ancestor of invertebrates, through a combination of whole-genome and large segmental duplications. We propose a universal nomenclature based on evolutionary relationships for the genes that encode these receptors, in which the genes are given the same orthologous names across vertebrates and paralogous names relative to each other. This nomenclature avoids confusion due to differential naming in the pre-genomic era and incomplete genome assemblies, furthers our understanding of the evolution of these genes, aids in the translation of findings across species and serves as a model for other gene families.

[Download PDF](#)

Main

OT and *VT* act as hormones or neurotransmitters that—through their respective G-protein-coupled receptors—regulate a wide range of biological functions, including uterine contractions and milk ejection in placental mammals; copulation, bond formation, thermoregulation, nesting behaviour and social vocalizations (for oxytocin) across many vertebrate and some invertebrate groups; and antidiuresis, blood pressure, parental care and reproduction (for vasotocin) in mammals and/or other vertebrates and invertebrates^{1,2} (Supplementary Table 1, Supplementary Note 1). In the pre-genomic era, small differences in amino acids of the *OT* and *VT* hormones in different species or lineages led biochemists to give them and their receptors different names: for example, mesotocin in birds, reptiles and frogs, and isotocin in teleost fish, for the apparent oxytocin complement of mammals; and vasopressin in mammals for the apparent vasotocin complement in other vertebrates⁶.

It has previously been hypothesized that *OT* and *VT* are the product of a local duplication near the origin of vertebrates⁵. However, the evolutionary trajectory of the receptors is under debate^{7,8,9,10}. One recent view⁹ is that the genes that encode the *OT* and *VT* receptors (hereafter, OTR-VTRs) evolved through two rounds of whole-genome duplication in the ancestor of cyclostomes. An alternative view¹⁰ posits that the OTR-VTRs evolved by one round of whole-genome duplication shared by agnathans and gnathostomes, followed by segmental duplications. However, these studies used highly fragmented genome assemblies and inconsistent annotations, and could not conclusively resolve the evolution of the OTR-VTRs. The resulting varied

biochemical-based and evolutionary-based terminologies have led to confusion as regards the orthology and paralogy of these genes, which is emblematic of a wider problem in gene nomenclature.

Here we analysed the genomes of 35 species that span all the major vertebrate lineages as well as an additional 4 outgroup genomes from invertebrate lineages (Supplementary Table 2); these included several species that were sequenced with long-read and long-range scaffolding technologies by the Vertebrate Genomes Project (VGP) (<https://vertebratogenomesproject.org/>), which filled gaps and corrected errors of previous shorter-read assemblies³. On the basis of gene synteny, sequence identity, family tree and other analyses, we propose that *OT* and *VT* are paralogous genes that arose through a local duplication via DNA transposable elements near the origin of vertebrates. We propose that the OTR-VTR genes evolved by a combination of whole-genome duplication and segmental duplication, which led to six receptors near the origin of jawed vertebrates with lineage-specific losses and gains thereafter. With this improved understanding of the relations between the OTR-VTR genes, we propose a universal vertebrate nomenclature based on evolutionary relationships (Table 1).

Table 1 Previous and proposed terminology for genes encoding OT and VT ligands and receptors in vertebrates

[Full size table](#)

Approach

In all genomes, we initially searched for *OT*, *VT* and OTR-VTR genes using pair-wise BLAST and BLAT analyses, and analysed the synteny of these genes from microchromosomal to macrochromosomal scales between and within species. We then assessed congruence between synteny, sequence identity and gene family trees.

Evolution of the VT and OT ligands

On the basis of BLAST searches, sequence identity and manual microsynteny analyses within a ten-gene window (microchromosomal), we found the human *VT* orthologue (that is, *AVP*) in all vertebrates analysed (Fig. 1a, Supplementary Fig. 1, Supplementary Table 3). Only the putative *VT* in hagfish did not have genes in synteny with any other vertebrate (presumably owing to the fragmented assembly), but the gene tree of this putative *VT* formed an immediate node with the lamprey *VT* (Extended Data Fig. 1a)—which suggests it is the *VT* homologue. In jawed vertebrates (after the divergence of the lamprey and hagfish), we found the *OT* orthologue directly adjacent to *VT* except in teleost fish (Supplementary Fig. 1, Supplementary Table 4a). In teleosts, *OT* was translocated nearby on the same chromosome (or to a separate

chromosome in zebrafish; Supplementary Table 4a), which supports more rearrangements in teleosts¹¹. The spotted gar—which represents the divergence of the holosteans, sister to the teleosts—had both *OT* and *VT* together in the translocated *OT* region found in teleosts, indicating that there was first a translocation and then a relocation of *VT* in teleosts near its original location. A previous short-read assembly of the megabat had *OT* as the only gene on a scaffold, indicating a fragmented assembly. The pale spear-nosed bat assembly from the VGP³ and Bat1K project⁴ revealed a local triplication of the *OT* and *VT* genes. We found support for this triplication using single Pacbio long reads and Bionano optical maps that spanned the entire region (Extended Data Fig. 1b,c, Supplementary Note 2a): such duplications are known to be hard to assemble with short reads¹². An *OT* orthologue was not found in lampreys and hagfish, which provides support for a previous result in lamprey¹³. This previous report was inconclusive owing to the fact that the assembly was generated from the sized-down, programmed and rearranged somatic genome, whereas we analysed a long-read germline genome of the sea lamprey¹⁴; the inshore hagfish data are from a short-read germline genome.

Fig. 1: Phylogenetic distribution and local gene duplication.



a, Phylogenetic distribution of *OT*, *VT* and OTR-VTR genes among vertebrates. Filled circles, presence of a gene; empty circles, loss of a gene; no circle, the gene never evolved in that lineage. Phylogenetic tree based on ref. [36](#). *Unresolved relationship for whether hagfishes and lampreys constitute a single phylum or two separate phyla^{[35,37](#)}. **b**, Local chromosomal organization of the *OT* and *VT* region. Representation of the position (in kb), orientation (+ or -) of *OT* and *VT* genes (exons + introns) in human chromosome 12, intron length (scale, 100 bases), GC content and DNA transposable elements with terminal inverted repeats (TE-TIRs) (green).

[Full size image](#)

In three of the four invertebrate species we analysed, we identified a single gene that was structurally similar (3–4 exons)—but not syntenic—to the vertebrate *VT* (Supplementary Table [5](#)), supporting previous findings^{[15,16,17](#)}. The exception was amphioxus, which had three copies of the *VT* gene: two on the same scaffold 23 kb apart from each other, and the other on another scaffold in a paralogous syntenic territory (Supplementary Table [5](#), Supplementary Note [2b](#)). Two of these three genes had previously been noted^{[13,18](#)}, which—together with our data—indicates several lineage-specific duplications in amphioxus.

To test the hypothesis that vertebrate *OT* could be a tandem duplication of *VT*^{[5](#)}, we searched for DNA transposable elements, which are known to drive gene duplications^{[19](#)}. We found transposable elements around *OT* (for example, in human and chimpanzee), but not around *VT* (Fig. [1b](#), Supplementary Tables [6, 7](#)). These transposable elements had terminal inverted repeats, which are known to transpose through a cut-and-paste mechanism that creates an extra copy at the donor site^{[19](#)}. We searched for other features that are encountered in duplicated genes, such as intron shortening and/or an increase in GC content^{[20](#)}: both of the human *OT* introns were shorter than the *VT* introns, with the first *OT* intron also being 13% richer in GC content (77.9% versus 64.6%) (Fig. [1b](#)). These relationships varied among species, with the elephant shark—representing a more basal vertebrate divergence than that of human—showing a large decrease in length of only the first *OT* intron compared to *VT* (3,226 bp versus 1,158 bp) but similar GC content.

We also found that the orientation of the genes was tail-to-head *OT*-to-*VT* (same direction) in nearly all vertebrates (including marsupial mammals)—except for placental and monotreme mammals, in which the orientation was tail-to-tail (*OT* inverted) (Supplementary Table [8](#)). This is indicative of the fact that, after the original *OT* tandem duplication of *VT*, *OT* inversions either occurred independently at the origin of monotremes and placental mammals (as previously suggested^{[13](#)}) or occurred at the origin of mammals with marsupials reverting back to the tail-to-head orientation. We also identified an independent *OT* inversion in the spotted gar

(Supplementary Table 8). The totality of our findings suggest that *OT* is a local tandemly duplicated copy of *VT* that arose after the divergence of jawed vertebrates, which was followed by divergences in introns, GC content, gene orientation, translocations and further duplications in different lineages.

A universal nomenclature for *OT* and *VT*

On the basis of these findings, we propose a universal nomenclature in which oxytocin (that is, *OT*) and vasotocin (that is, *VT*) are used for these genes in all jawed vertebrates, and *VT* is used in all jawless vertebrates and closely related invertebrates. We believe that these genes should be named in this manner because it portrays their evolutionary history, as is standard practice for other genes that are orthologous across species (for example, *FOXP1*) and paralogous within species (for example, *FOXP2*, *FOXP3* and *FOXP4*). According to this practice, the genes encoding these two peptides would be named vasopressin 1 (*AVP1*) and vasopressin 2 (*AVP2*), vasotocin 1 (*VT1*) and vasotocin 2 (*VT2*) or oxytocin 1 (*OT1*) and oxytocin 2 (*OT2*). As we realize that this would be a far-reaching shift from the existing nomenclature, we propose that the common origin of these genes be portrayed through the shared suffix -tocin, and paralogy conveyed through different root words oxy- and vaso-. Vasotocin is a name that is already used by most scientific communities focusing on non-mammalian species (Table 1). Furthermore, the name ‘arginine vasopressin’ (*AVP*) entails that this gene encodes an arginine as the eighth amino acid, which is not the case for all mammals⁶. For non-mammalian species, this means that the peptides currently known as mesotocin, isotocin, glumitocin, valitocin, aspargtocin and neurophysin in different lineages would now be called by one orthologous name (that is, oxytocin) (Table 1).

Six vertebrate OTR–VTRs

Our manual microsynteny analysis within a ten-gene window revealed six paralogous receptors among vertebrates (Fig. 1a, Supplementary Tables 3, 4b–e). Most vertebrate species had four or five of the six receptors, and some had further lineage-specific duplications (Extended Data Fig. 9, Supplementary Note 2c–e). For greater clarity, we present our findings using our proposed nomenclature of the root names for the ligands, with evolution-based suffixes (Table 1) from evidence highlighted in the ‘Evolution of the VT and OT ligands’ section.

On the basis of microsynteny analyses, we found the gene that is commonly defined as the oxytocin receptor (*OXTR*) in mammals or the mesotocin receptor (*MTR*) in birds (henceforth referred to as *OTR*) in a well-conserved syntenic region in nearly all of the vertebrates we examined (Fig. 2a, Supplementary Fig. 2). We found similar results for the gene known as arginine vasopressin receptor 1 (*AVPR1*) or arginine vasopressin

receptor 1A (*AVPR1A*) in mammals, or vasotocin receptor 4 (*VT4*) in some non-mammals (henceforth referred to as vasotocin receptor 1A (*VTR1A*)) (Supplementary Fig. 3). By contrast, the gene known as *AVPR3* or *AVPR1B* in mammals or as vasotocin receptor 3 (*VT3*) in non-mammals (henceforth referred to as vasotocin receptor 1B (*VTR1B*)) was present in all tetrapods, sharks and coelacanths, but was absent in other fish and lampreys (Supplementary Fig. 4). The syntenic territory of *VTR1B* was present in all fish except lampreys, which indicates a gain of *VTR1B* after the divergence from lampreys that was followed by a loss in holosteans and teleosts after their divergence from coelacanths and sharks (Fig. 1a). In addition, teleosts showed rearranged syntenic gene blocks on each side of *OTR*, each side of *VTR1A* and on one side of the (lost) *VTR1B* (Supplementary Figs. 2–4).

Fig. 2: Interspecies and intraspecies synteny analyses.



a, Example of interspecies ten-gene microsynteny for *OTR* across vertebrates. Same colour, orthologous genes. Black boxes, genome rearrangements. *OTRa* in the sea lamprey and zebrafish is orthologous to *OTR* in all other vertebrates. Human *OTR* is currently known as *OXTR*; tropical clawed frog *otr* is currently known as *oxtr*. **b**, Intraspecies 10-Mb macrosynteny among 6 chromosomes (block colours) for all OTR-VTR gene regions in humans whether present (*OTR*, *VTR1A*, *VTR1B* and *VTR2C*) or deleted (*VTR2A* and *VTR2B*). Gene families are listed alphabetically on the left. In the blue column, underlined genes were found within a 10-Mb window of *VTR1A* on chromosome 12. In the pink column, underlined genes were found within a 10-Mb window of *OTR* on chromosome 3; genes in black bold were found within a 10-Mb window of the deleted *VTR2A* on chromosome 12 or (in blue bold) 7, or within the 10-Mb window of the deleted *VTR2B* on chromosome 3. Orange column, all genes listed (in black bold) were found within a 10-Mb window of *VTR1B* on chromosome 1 (orange block). Yellow column, all genes listed (in black bold) were found within a 10-Mb window of *VTR2C* on chromosome X (yellow block). Green column, an alternative syntenic territory of *VTR2B* (green) was also found at a different location of chromosome 3. Genes not in bold are found outside of the strict 10-Mb window, but are on the same chromosome as the respective OTR-VTR gene.

[Full size image](#)

The gene known as vasotocin receptor 1 (*VT1*) in birds, and by several other names in other lineages (Table 1) (henceforth referred to as vasotocin receptor 2A (*VTR2A*)), was found in conserved synteny in reptiles, mammals and some fish, although its syntenic territory was present in all of these lineages (Supplementary Fig. 5)—indicating independent losses (Fig. 1a). The gene known as arginine vasopressin receptor 4 (*AVPR4*) in fish (henceforth referred to as vasotocin receptor 2B (*VTR2B*)) was detected only in fishes and lampreys but its syntenic territory was detected in all vertebrates (Supplementary Fig. 6), which indicates a loss in the tetrapod ancestor (Fig. 1a). The gene commonly known as arginine vasopressin receptor 2 (*AVPR2*) in mammals or as *AVPR2A* in fishes (henceforth referred to as vasotocin receptor2 C (*VTR2C*)) was found in all vertebrates except for lampreys, elephant sharks and birds (Supplementary Fig. 7). In birds, the absent *VTR2C* was part of a larger block of about 20 genes that has been deleted²¹. These findings indicate a gain of *VTR2C* in vertebrates after the divergence of elephant sharks, followed by a loss in birds (Fig. 1a). Again, teleosts showed rearranged syntenies on either side of *VTR2B* and on one side of *VTR2A* and *VTR2C* (Supplementary Figs. 5–7).

In all of the species we assembled to chromosomal resolution, *OTR* and *VTR2B* were syntenic on the same chromosome and separated by 10–30 genes (Extended Data Fig. 2a, Supplementary Table 4b). Similarly, *VTR1A* and *VTR2A* were also on the same chromosome or scaffold and separated by 4–50 genes, except in mammals and fish (Supplementary Table 4c). In mammals, the syntenic genes (including *VTR1A*) on one

side of the deleted *VTR2A* were on chromosome 12 (human nomenclature), whereas those on the other side were on chromosome 7^{8,9} (Extended Data Fig. 2b, Supplementary Table 4c), which indicates a fission that possibly involved the loss of *VTR2A* in mammals. In fish, there were complex patterns of rearrangements and duplications but some species (for example, the three-spined stickleback, gar, coelacanth and elephant shark) still contain *VTR1A* and *VTR2A* on the same chromosome (Supplementary Table 4c), which indicates lineage-specific chromosomal fissions and other rearrangements. In hagfish, we found only two VTR genes. These genes are located on two different scaffolds, one containing *VTR1* and the other containing *VTR2*, and each is equally syntenic for gene families containing the *OTR* and *VTR2B* combination and the *VTR1A* and *VTR2A* combination in other vertebrates (Supplementary Table 4f,g); this indicates an ancestral relationship to both chromosome combinations, possibly via duplication. Higher-quality germline assemblies for hagfish should reveal whether these two scaffolds are really separate or are part of the same chromosome. Finally, *VTR1B* and *VTR2C* were singly found on different chromosomes or scaffolds in all species in which they were present (Supplementary Table 4d,e). We verified these findings with an independent, automated, more-quantitative and longer-range measure by using SynFind²² on alignments in up to 100-gene macrosynteny windows around the receptors, in all major lineage combinations (Extended Data Fig. 3, Supplementary Note 3).

Chromosomal orthology and paralogy of OTR-VTRs

To assess whether the interspecies synteny we identified was due to local segmental synteny within a chromosome or to entire chromosomal-scale orthology, we generated dot plots using SynMap2 for entire chromosomes or scaffolds that contained OTR-VTR genes, focusing on comparisons between species that represent major vertebrate lineages and which were sequenced at chromosomal resolution. By examining basal branches, we found the sea lamprey scaffold that contained the combination of *OTR* and *VTR2B* had the highest number of syntenic hits (30–60 genes) to the chromosome in all other vertebrates that had the combination of *OTR* and *VTR2B* (Fig. 3a). We found a similar result between species in chromosomes containing the combination of *VTR1A* and *VTR2A* (Fig. 3b). Exceptions included some fish, in which another chromosome had a similar number of syntenic hits (Fig. 3a,b)—consistent with an extra chromosome parologue from an additional whole-genome duplication. Mammals were also an exception: here, the orthology of the sea lamprey scaffold containing the *VTR1A*-*VTR2A* combination was split between two chromosomes (for example, human chromosomes 7 and 12) (Fig. 3b), consistent with a fission event. These results indicate that these two gene combinations are syntenic, because each belongs to a chromosome orthologue of vertebrates after the split with lampreys. The second highest gene hits in most species were to the chromosome that contained the other receptor combination (Fig. 3a,b), which indicates that the chromosomes containing

the *VTR1A*-*VTR2A* and *OTR*-*VTR2B* combinations may be paralogues from a whole-genome duplication. The third highest number of syntenic gene hits were to chromosomes that contained *VTR1B* or *VTR2C* (in no particular order) (Fig. [3a, b](#)), which suggests possible paralogous segmental duplications. Similar—but not as strong—results were found for an apparent duplicate sea lamprey scaffold that contained one VTR gene (Extended Data Fig. [4a, b](#), Supplementary Note [1c](#)).

Fig. 3: Analysis with SynMap2 identified syntenic gene hits between sea lamprey scaffolds containing two receptors each and chromosomes of other species.

 **figure3**

Bar graphs were created from dot plots. **a**, Sea lamprey scaffold 27 is most syntenic with chromosomes of other species that contain the *OTR-VTR2B* combination. **b**, Sea lamprey scaffold 10 is most syntenic with chromosomes of other species that contain the *VTR1A-VTR2A* combination. The minimum number of aligned homologous gene pairs to be considered syntenic was 3 at a 20-gene maximum distance in each species. For comparisons including human, the minimum number was set to 2. *Significant differences between chromosomes with the highest number of gene hits within a

species ($P < 0.05$; χ^2 test, two-sided; $n = 199$ genes located on scaffold 27; $n = 246$ genes located on scaffold 10).

[Full size image](#)

When we used the two scaffolds that contained the *VTR1* and *VTR2* hagfish genes as references, we found fewer syntenic gene hits to chromosomes of other species: chromosomes with the *OTR-VTR2B* and *VTR1A-VTR2A* combinations showed the highest number of hits, with no clear preference between them (Extended Data Fig. 4c–f, Supplementary Note 4). These findings further support a deep ancestral paralogy between chromosomes that contain these two receptor combinations, with possible ancestral chromosome representatives in hagfish. Our findings are consistent with previous studies identifying chromosomal paralogues^{14,23,24,25,26}, and further reveal newly identified candidate chromosomal paralogues in species with genomes that—to our knowledge—have not yet been compared (sea lamprey versus medaka, frog versus medaka and so on) (Fig. 3).

Rapid radiation of the VTR1 and VTR2 families

We next assessed paralogues within species to help to determine evolutionary relationships among the receptors. We analysed 10-Mb macrosynteny windows between and within chromosomes of the same species (intraspecies) for all 6 receptors (whether present or deleted). Within species (for example, human), we found paralogous ‘gene families’ in syntenic blocks around all OTR-VTR genes (Fig. 2b, Supplementary Table 9a), which supports the notion that at least parts of these chromosomes are paralogous (human chromosomes 12 and 7, versus chromosome 3). We also found an extra gene-family territory on human chromosome 3 that is syntenic with the *VTR2B* territory (Fig. 2b). However, we did not detect any *VTR1* or *VTR2* genes within a species that shared substantially more synteny than others (Supplementary Table 9a, b). Further, no gene family was present in the territory of all six present or deleted receptors. At a more microscale level, among the exons and introns of OTR-VTR genes (Extended Data Fig. 5) and in flanking microRNAs and long non-coding RNAs (Extended Data Figs. 6, 7, Supplementary Table 10), we also did not find any one gene with more similarity to another that would allow us to make further conclusions about the evolution of gene subfamilies (for example, sea lamprey and human) (Supplementary Notes 5, 6). Most sequences were too similarly divergent to inform paralogy (Supplementary Note 5). Overall, the lack of further macrosynteny and simple sequence-divergence resolution of paralogues within a species—combined with the better resolution on chromosome orthologues and paralogues between species in which these gene regions reside—suggest a rapid radiation of receptor evolution near the origin of vertebrates.

Single chromosome origin of OTR-VTRs

To assess receptor origins, we performed ancestral analyses by mapping regions that contain the OTR-VTR genes against reconstructed chromosomes of the vertebrate or chordate ancestor from four independent studies^{14,23,25,27}. Chromosome fragments containing the OTR-VTR genes in vertebrates all mapped back to the same reconstructed chromosome (Supplementary Table 11). A previous study²⁸ suggested that *VTR2C* maps back to a separate ancestral chromosome: we believe that this is inaccurate because the region that contains the gene we name *VTR2C* is entirely missing from the reconstruction that was used²³, although it is present in other reconstructions^{14,25,27} that were based on higher-quality amphioxus and sea lamprey genome assemblies (Supplementary Table 11). These findings are consistent with the fact that that vast majority of invertebrates that have been examined have only one *VTR* gene^{13,17} (Supplementary Table 5).

Synteny, phylogeny and receptor evolution

We generated phylogenetic trees of the receptor gene family across vertebrates using alignments of both the exon nucleotide (RAxML) and amino acid (TreeFam and TreeBeST5) sequences, and then mapped our revised synteny-based naming onto the tree leaves. BLAST nucleotide comparisons alone, and previous nomenclatures based on these analyses (Table 1), yielded many contradictions with the synteny-defined orthologues (Supplementary Table 12). We believe this is due to BLAST not returning alignments of the entire sequence, which in turn is due to larger lineage divergences between those gene regions (Supplementary Note 5). By contrast, our phylogenetic sequence analyses revealed tree topologies with almost 1:1 consistency to our synteny-defined relationships (Fig. 4).

Fig. 4: OTR-VTR gene family trees.

 **figure4**

a, Tree topology inferred with the phylogenetic maximum likelihood method on an exon nucleotide alignment (MAFFT), with 1,000 non-parametric bootstrap replicates. Bootstrap values are shown as percentages at the branch points (values <50% were considered less informative). The tree is rooted with the three *VTR* genes we found in

amphioxus. The gene names of the current accessions (see Table 1 and Supplementary Tables 4a–e for full list of synonyms) were written over according to our revised synteny-based orthology. Scale bar, phylogenetic distance of 0.78 substitutions. **b**, Tree topology inferred with the phylogenetic TreeFam method on an amino acid alignment generated via the Ensembl ‘gene tree’ tool (gene tree identifier: ENSGT00760000119156). Left, red boxes denote inferred gene duplication node; blue boxes denote inferred speciation events; and turquoise boxes denote ambiguous nodes. Right, green bars denote multiple amino acid alignment made with MUSCLE; white areas denote gaps in the alignment; and dark green bars denote consensus alignments. Gene names are revised according to our synteny-based orthology; Extended Data Fig. 8 shows a tree with the current nomenclature in Ensembl.

[Full size image](#)

The combined phylogeny and synteny analyses supported a single *VTR* gene shared with an invertebrate ancestor (that is, represented in sea squirt). This receptor then duplicated into what we designate the ancestral *VTR1* and *VTR2* on the same chromosome (that is, represented in hagfish) (Fig. 4). Thereafter, the trees suggest that these two genes expanded into three genes in the *VTR1* subfamily (*VTR1A*, *OTR* and *VTR1B*) and three in the *VTR2* subfamily (*VTR2A*, *VTR2B* and *VTR2C*), respectively, with *VTR1A* and *VTR2A* on the same chromosome and the paralogous *OTR* and *VTR2B* on another chromosome. Thereafter, the sister relationship of *VTR1A* and *VTR1B* in both trees suggests that one directly gave rise to the other, after the divergence of jawless fish (based on absence of *VTR1B* in lampreys). Likewise, the sister relationship of *VTR2B* and *VTR2A* in the nucleotide tree is consistent with the synteny findings of one giving rise to the other, and—together—their sister relationship to *VTR2C* is consistent with one of them giving rise to it, after the divergence of sharks (based on absence of *VTR2C* in sharks).

In stark contradiction to the synteny findings, the lamprey *VTR1A* and *OTR* genes each clustered outside of their respective synteny-defined *VTR1* homologues among species and the same occurred for lamprey *VTR2A* and *VTR2B* for the *VTR2* homologues (Fig. 4), which implies lamprey-specific duplications. One possible explanation for these contradictions is that there could be convergence within the lamprey *OTR-VTR* genes (possibly owing to higher GC content²⁹) or that the divergence was so rapid at the origin of vertebrates that the true relationship is not easy to resolve using gene tree inference. Consistent with the latter, the bootstrap support values are low (72–74%) for a more recent gene duplication. Consistent with the former, the lamprey exon sequences of all 4 receptors were among the highest in GC content (60–69%) compared to other species (Supplementary Table 13). The three amphioxus *VTR* sequences in our exonic tree cluster within species at 100% support (Fig. 4a), consistent with lineage-specific duplications (Supplementary Note 2b). There were some differences in local relationships in the exon and amino acid trees

(Fig. 4a, b), but these did not affect the major conclusions here (Supplementary Note 7).

A universal nomenclature for OTR-VTR genes

On the basis of the above findings, we propose a universal nomenclature for the OTR-VTR genes in which their root terms follow the ligand names (oxytocin receptor (OTR) and vasotocin receptor (VTR)) and their enumeration terms (1A, 1B, 2A, 2B and 2C) follow their evolutionary history: the numbers 1 and 2 designate the original duplication, and the letters A, B and C designate the subsequent subfamily duplications. The only exceptions we made were the decisions not to rename *OTR* as *VTR1B* or *VTR1B* as *VTR1C* (as the evolutionary history warrants), because we felt this might be too radical of a departure from the common use. This is further justified in that—although there is crosstalk in OT and VT binding to these receptors—OT is the dominant ligand for OTR³⁰. For the VTR2 genes, we reordered the enumerations according to the inferred chronological order of duplications: *VTR2A* and *VTR2B* for the genes we found in lampreys, and *VTR2C* for the gene that evolved in the ancestor of bony fishes. This universal nomenclature gives a single name to each gene across vertebrates. The gene that is commonly known as arginine vasopressin receptor 1A (*AVRP1A*) in mammals, vasotocin receptor 4 (*VT4*) in birds and vasotocin receptor (*VasR*) in frogs would, in our revised nomenclature, be called vasotocin receptor 1A (*VTR1A*) (Table 1). The gene that is commonly known as oxytocin receptor (*OXTR*) in mammals, vasotocin receptor 3 (*VT3*) or mesotocin receptor (*MTR*) in birds and frogs, and isotocin receptor (*ITR*) in fish would be called oxytocin receptor (*OTR*). Similar changes from multiple names to a single name apply to the other four receptors (Table 1).

Interpretations and evolutionary hypotheses

We considered a model of OTR-VTR evolution in the context of two competing hypotheses of vertebrate genome evolution: one round of whole-genome duplication and segmental duplications (Fig. 5a) versus two rounds of whole-genome duplication (Fig. 5b). For both hypotheses, we propose that the single *VTR* in the vertebrate ancestor had a tandem segmental duplication on the same chromosome at over 550 million years ago³¹ that gave rise to the ancestral *VTR1* and *VTR2* genes. Thereafter, in a one round of whole-genome duplication in a gnathostome ancestor, one copy of each of the *VTR1* and *VTR2* genes gave rise to the *VTR1A*-*VTR2A* combination on one chromosome parologue and the *OTR*-*VTR2B* combination on the other chromosome parologue. From here, the two hypotheses diverge. In hypothesis 1 (Fig. 5a), a segmental translocated duplication of the chromosomal region containing *VTR1A* gives rise to *VTR1B* in the ancestor of jawed vertebrates at over 500 million years ago

and a segmental translocated duplication of the region containing *VTR2B* gives rise to *VTR2C* in the common ancestor of other vertebrates with bony fish at over 450 million years ago. Segmental duplications have been found in other gene families at these evolutionary time points^{32,33}. In hypothesis 2 (Fig. 5b), two rounds of whole-genome duplication before the divergence of gnathostomes from cyclostomes lead to four more receptors in the ancestor of jawed vertebrates at over 500 million years ago. Both hypotheses agree with lineage-specific losses of *VTR2B* in the ancestor of tetrapods, of *VTR2A* independently in mammals and teleost fish, of *VTR1B* in holostean and teleost fish, and of *VTR2C* in birds. However, for hypothesis 2 to be true, complete independent losses of thus-far unidentified *VTR1C* and *VTR2D* genes and associated chromosome segments in an extinct species before its divergence into all other vertebrate lineages would be required. Our results are more parsimoniously explained by hypothesis 1 (one-round of whole-genome duplication)¹⁴ with prior and subsequent segmental duplications (Supplementary Note 8). Such a vertebrate evolutionary scenario is consistent with expectations given a simple random mutational model³⁴ that requires as few as 6 mutational steps, whereas models that invoke two rounds of whole-genome duplication require at least 9 steps in our case (Supplementary Note 8) or 12–18 under previous assumptions³⁴. Our findings of only two receptor genes (*VTR1* and *VTR2*) in inshore hagfish is more consistent with the paraphyly (separate phyla) of lampreys and hagfishes than the monophyly hypothesis³⁵, as it does not require further inference of independent gene losses or gains or incomplete lineage sorting (Fig. 1a). Thus, our findings may have repercussions on a wider and highly debated topic—that of the evolution of vertebrate genomes (Supplementary Note 9).

Fig. 5: Two proposed hypotheses for the evolution of OTR-VTR genes.



a, Hypothesis 1 proposes the receptors evolved by an initial segmental duplication (SD), then a one round of whole-genome duplication (WGD), followed by two segmental duplications in different vertebrate lineages and then by losses (red X) in specific lineages. **b**, Hypothesis 2 proposes the initial segmental duplication was followed by two rounds of whole-genome duplication and specific losses (X), including in all vertebrates (blue X). Lines connecting genes indicate that they are on the same chromosome in most species. Alignments between sets of genes indicate the closest related paralogues.

[Full size image](#)

Our revised understanding of the receptor relationships allows a more holistic view of their functions. We generated a multiprotein coding sequence alignment among the highest-quality assemblies of all six receptors, and found that the seven transmembrane domains and associated polar amino acids that interact with OT or VT have remained highly conserved in sequence or amino acid type, even 550 million years after their common origin (Extended Data Fig. 10, Supplementary Note 10). By contrast, the extracellular OT or VT binding domains and the intracellular G-protein binding domain became highly diversified from one receptor to another, predicting greater diversity in initial ligand binding and subsequent intracellular signalling. Nine amino acid sites distinguish the VTR1 and VTR2 subfamilies (most in or near the transmembrane regions), but only one of these is in the G-protein binding region (Extended Data Fig. 10). All of the receptors use diacylglycerol, inositol triphosphate and Ca^{2+} for second-messenger intracellular signalling—except for VTR2C, which uses cAMP (Supplementary Table 31). The tissues in which the receptors have their highest expression include the brain (except VTR2A), with expression being highest in the adrenal gland. We could not find data available for signalling or expression for VTR2B in fish, but predict it will be similar to members of the VTR2A and VTR2C subfamilies. Finally, our findings that the OTR (represented by lamprey divergence) evolved millions of years before the OT ligand (represented by elephant shark divergence) suggest that the ancestral VT may have originally acted through the OTR before OT evolved. This suggestion is supported by findings that, in some species, OT and VT bind to the OTR at similar efficiencies⁸; a greater response of OTR to OT over VT is found for the first time in teleost fish⁸.

In summary, we believe that our revised evolution-based and universal nomenclature will make translating findings across vertebrates much easier. It will help to inform our understanding of crosstalk between some of the ligands and receptors, our understanding of genome evolution and could serve as a model for a broader universal gene nomenclature.

Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized, and investigators were not blinded to allocation during experiments and outcome assessment.

Overall synteny and BLASTn analyses

To define orthology in the *OT*, *VT* and OTR-VTRs in all vertebrates, we used interspecies synteny analyses at three scales: a manual 10-gene window microsynteny analyses using BLAT and BLAST^{38,39} searches and cross-species genome alignments; a more automated 100-gene macrosynteny window using SynFind and GeVo²²; and automated chromosomal-scale alignments with syntenic dot plots using SynMap2⁴⁰. To define paralogy and further trace the evolutionary history of the genes, we used intraspecies synteny analysis, searching for paralogous genes in 10-Mb windows. Microsynteny was useful for determining orthologous and paralogous relationships between genes in the majority of the vertebrate lineages. Macrosynteny was useful for determining orthologous and paralogous relationships in cases in which the microsynteny was weaker, such as between genes found in lampreys or hagfish with the rest of the vertebrates. Sequence identity was determined using BLASTn to understand relationships between the per cent identity and synteny (Supplementary Table 12). Only results with a bit score >40 and hits with high probability *E* value $<10^{-4}$ were kept. We describe the specific methods for each synteny approach in ‘Microsynteny between species in approximately 10-gene windows’, ‘Macrosynteny between species in approximately 100-gene windows’ and ‘Chromosome-scale macrosynteny between species’.

Microsynteny between species in approximately 10-gene windows

We ran microsynteny analysis by manually scanning annotated alignments for 5 protein-coding genes before and after each focus gene (Supplementary Table 4a–e) in 35 species spanning all major vertebrate lineages (Supplementary Table 2). The candidate genes in each species (accession number or gene identifier in Supplementary Table 4a–e) were first selected by BLAT and BLAST searches using the UCSC genome browser and alignment (<http://genome.ucsc.edu/>)³⁸ and the SynFind tool from the CoGe comparative genomics research platform²². The NCBI and Ensembl⁴¹ (v.95) database genome alignments were used to identify the neighbouring genes. For the neighbouring genes, we kept in our Supplementary Tables the family gene names used in the genome annotation of each species, even though in some cases—we believe erroneously—different family names have been given to the orthologous gene in different species (for example, *FABP1A* in spotted gar and *FABP1B.1* in stickleback; Supplementary Table 4a). For the species that had more lineage-specific duplications, we labelled the gene that shared more synteny with the orthologue in other vertebrate

lineages with ‘a’ (for example, *OTRa*), and labelled the copy with ‘b’ (for example, *OTRb*). We listed the aliases in NCBI and Ensembl for each focus gene in each organism (‘Aliases’ column in Supplementary Table 4a–e) and included the most frequent ones in Table 1. When our target genes appeared to be lost in a species (no initial BLAST hit), we searched the surrounding gene territory to determine whether only the gene of interest or a larger block of genes were deleted, or whether the deletion was due to an incomplete genome assembly or assembly artefact.

For some species with more-fragmented genome assemblies or annotations or greater divergences in NCBI and Ensembl, we analysed other higher-quality assemblies and annotations. This included the VGP zebra finch, Anna’s hummingbird, pale spear-nosed bat and platypus genome assemblies³. For the Japanese lamprey, we included previously published synteny data¹⁰. For the sea lamprey, we used the assembly of the germline genome¹⁴ and analysed it with BLAST, Genome Browser and Gene Search tools (<https://genomes.stowers.org/organism/Petromyzon/marinus>). For amphioxus, we used the BLAST and Gene Brower tools available at <https://genome.jgi.doe.gov/Brafl1/Brafl1.home.html> with the latest version of the amphioxus genome (*Branchiostoma floridae* v.2.0), whereas previously reported data¹³ are based on the first version of the genome (*B. floridae* v.1.0). For the inshore hagfish genome assembly, the contigs were relatively short and not fully annotated, and thus we first BLAT-searched all the *OT*, *VT* and OTR-VTR sequences of all the aforementioned species against the hagfish genome in Ensembl, found two putative OTR-VTRs in two separate contigs in the hagfish assembly, and then used the ‘Region comparison’ tool of Ensembl to map each gene of these contigs against the human, zebrafish and lamprey genomes (Supplementary Table 4f,g). BLAST gave many gene hits in the hagfish genome, but only with short segments aligning to *OT* and *VT* orthologues in other species. Thus, to determine whether they were real *OT* or *VT* orthologues, we used the ‘Gene Tree’ tool of Ensembl that constructs a phylogeny using the entire protein sequence, with the sea lamprey *VT* as reference. For the receptors, we used our data from the SynMap2 dot plots (described in ‘Chromosome-scale macrosynteny between species’) and included in the synteny of the hagfish receptors the gene hits that appear on the chromosomes in which the OTR-VTRs are located in human, chicken, zebrafish and sea lamprey (Supplementary Table 4f,g).

Macrosynteny between species in approximately 100-gene windows

We generated gene sequence alignments between pairs of species using SynFind²² (density, LastZ defaults). This results in a matrix containing syntenic gene-hit values in the reference species relative to query species along with their chromosomal locations. This data matrix was parsed and visualized using a custom R script (https://github.com/ggedman/OT_VT_synteny). First, a 100-gene window centred around a given receptor gene in the reference organism (x axis) was defined using

biomaRt (v.3.10). As we move 5' (left) or 3' (right) from zero (the focus gene) and if synteny exists, the number of gene hits for a given receptor in the query species shows a cumulative increase. This allowed us to identify large stretches of homologous sequences interspersed by divergent sequences.

Chromosome-scale macrosynteny between species

We used SynMap2⁴⁰ to generate syntenic dot plots of chromosome sequence alignments between species that contain OTR-VTRs (Supplementary Tables 15–30). SynMap2 identifies collinear sets of genes or regions of sequence similarity to infer synteny between two sequences, and generates a dot plot of the results. We used the default parameters (as of December 2018), except for ‘Minimum number of aligned pairs’. This parameter defines the minimum number of homologous genes (based on last default parameters) that should be found in a 20-gene distance for these genes to be considered syntenic and to appear on the dot plot. In more closely related lineages, we selected three as a minimum number (for example, between sea lamprey on the one hand, and Japanese medaka, or zebrafish, frog and chicken genomes on the other); for more distantly related species, we used two (for example, between hagfish on the one hand, and sea lamprey, or Japanese medaka, zebrafish, frog, chicken and human genomes on the other). Additionally, because the hagfish contigs were shorter than most other assemblies (making synteny more difficult to identify), we also ran a dot plot with 1 as the minimum number to search for all possible homologous hits, regardless of synteny.

To test for significant differences, we ran a χ^2 test with distinct samples of genes on the difference of the proportions between the first two chromosomes with the highest number of gene hits, using the number of genes in the super-scaffold of the reference species (for example, sea lamprey) as sample size: Supplementary Table 30 provides confidence intervals, degrees of freedom and *P* values. For cases that reached significance, to confirm that the number of hits between two species was independent of the number of protein-coding genes located on the chromosome of the query organism, we applied a gene density-normalization test to rule out the possibility that the chromosomes with most gene hits were owing to them containing the most genes: we did not find such correlations with our macrosynteny analyses.

Macrosynteny within species in approximately 10-Mb windows

We primarily used the human genome, as it is the best assembled genome and therefore subject to generating fewer errors. We listed all genes found in a 10-Mb window from the present OTR-VTRs (for example, *OTR*, *VTR1A*, *VTR1B* and *VTR2C* in mammals) as well as absent ones (for example, *VTR2A* and *VTR2B*, which are absent in mammals). We chose a 10-Mb window because this genomic region size

often captured macrosynteny of >40 genes, allowing within- and between-species macrosynteny analyses described in ‘Macrosynteny between species in approximately 100-gene windows’ to be comparable. We then searched each gene in the HUGO Gene Nomenclature Committee Database (<https://www.genenames.org/>) to classify its gene-family. For the deleted genes, we defined their territories by manually identifying in the human genome the genes around spotted gar *VTR2B* and chicken *VTR2A*; some of these syntenies around the deleted OTR-VTRs had previously been identified^{8,9}, which we confirmed.

Evolutionary history analyses of *OT* and *VT*

We noted annotated DNA transposable elements in the UCSC Genome Browser in close vicinity of the *OT* and *VT* genes (except for the elephant shark genome, which was not annotated for DNA transposable elements), and thus we quantitatively searched for adjacent transposable elements in the human and chimpanzee genomes using RepeatMasker (<http://genome.ucsc.edu/>)³⁸ and obtained information for each specific transposable element using Dfam 2.0¹². We calculated GC content using ENDMEMO (<http://www.endmemo.com/bio/gc.php>). We aligned the introns of human *OT* and *VT* in all possible combinations using DIALIGN⁴² and compared intron lengths using Serial Cloner v.2.6 (http://serialbasics.free.fr/Serial_Cloner.html). For relative *OT* and *VT* orientations, we examined whether they were in the same direction (tail-to-head) or in opposite directions (tail-to-tail) in the annotated positions in each species. In the cases in which *OT* and *VT* were found in opposite directions, we determined which gene was inverted according to the orientation of other genes in the territory. In addition to the genomes used for all other analyses of this study (Supplementary Table 2), we also used the koala (*Phascolarctos cinereus*) (phaCin_unsw_v4.1; GCF_002099425.1) and the grey short-tailed opossum (*Monodelphis domestica*) (MonDom5; GCF_000002295.2) genomes to include orientation data from the marsupial clade in Supplementary Table 8.

Evolutionary history analyses of OTR-VTRs

To assess in which ancestral vertebrate chromosomes the OTR-VTRs originated, we used four ancestral chromosome models from the literature^{14,23,25,27}, in which the reconstructed chromosomes were based on different species and different genome qualities. Specifically, human, mouse, dog, chicken and tetraodon genomes were used in ref. 23; human, chicken, stickleback, pufferfish, sea squirt, amphioxus, sea urchin, fruitfly and sea anemone genomes in ref. 27; human, chicken and sea lamprey (somatic) genomes in ref. 25; and chicken, spotted gar and sea lamprey (germline) genomes in ref. 14. We searched for the presence of annotated OTR-VTRs in four outgroup invertebrate lineages (through literature review, BLAST and BLAT searches)

—namely in sea squirt, roundworm, California sea hare and amphioxus. For the amphioxus genome (*B. floridae* v.2.0), we performed BLAT queries on OTR-VTR FASTA sequences from all species studied using the JGI genome browser (<https://genome.jgi.doe.gov/portal/>).

To test which sea lamprey receptor(s) most probably represents the orthologous ancestral gene(s), we compared the sea lamprey OTR-VTRs in all possible combinations to each other using BLASTn (same parameters). We compared the exons and introns of the identified genes separately to understand the divergence of the paralogous genes, following a previously proposed paradigm⁴³, using the maximum score and per cent identities of the comparisons that were above the threshold (maximum score >40 and *E* value < 10⁻⁴). We performed a similar analysis for *VTR1B* and *VTR2A* in elephant shark and coelacanth, to test whether sequence identity can help to solve ancestry questions. To shed light on the orthology between the inshore hagfish and the sea lamprey OTR-VTRs, we compared their exons and introns as well.

To analyse conserved non-coding RNA synteny around the OTR-VTRs, we looked for them in alignments in all the species studied in Ensembl, in the miRbase (<http://www.mirbase.org/>; release 22), and the miRviewer⁴⁴ database (28 February 2012 update). We aligned (BLASTn) long non-coding RNA regions within species (sea lamprey and human).

Gene tree phylogeny analyses

Exonic nucleotide tree

Exonic sequences from all the OTR-VTRs from representative species that had the most-complete assembled genes were aligned with MAFFT under the E-INS-i parameter set, which is optimized for sequences with multiple conserved domains and long gaps. Any incomplete non-lamprey OTR-VTR of less than 1,000 bp was excluded, as alignments on short sequences often lack power to resolve species relationships, resulting in weakly supported gene trees. Because of the basal phylogenetic position of the lamprey, all lamprey OTR-VTRs (754 bp and longer) were included. From this alignment, we generated a phylogenetic maximum likelihood tree using GTRGAMMA model of RAXML (version 8.2.10)⁴⁵, with 1,000 replicates. We calculated the GC content of all the exonic sequences using <http://www.endmemo.com/bio/gc.php> (Supplementary Table 13).

Protein amino acid tree

A maximum likelihood phylogenetic tree was constructed on one representative amino acid sequence for every gene in every species, using TreeFAM and TreeBeST5 pipeline in the gene tree tool package of Ensembl (https://www.ensembl.org/info/genome/compara/homology_method.html). Thereafter, we manually curated the Ensembl tree (gene tree identifier: ENSGT00760000119156) using the universal nomenclature that we propose here. All the sequences used to generate both trees, sequence alignments and Newick files can be found at <https://github.com/constantinatheo/otvt>.

Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

Data availability

All the data used in this study can be found in Supplementary Tables 2–30, and at <https://github.com/constantinatheo/otvt>. Any other relevant data are available from the corresponding authors upon request.

Code availability

The code used in this study can be found at <https://github.com/constantinatheo/otvt> and https://github.com/ggedman/OT_VT_synteny.

References

1. 1.

Knobloch, H. S. & Grinevich, V. Evolution of oxytocin pathways in the brain of vertebrates. *Front. Behav. Neurosci.* **8**, 31 (2014).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

2. 2.

Meyer-Lindenberg, A., Domes, G., Kirsch, P. & Heinrichs, M. Oxytocin and vasopressin in the human brain: social neuropeptides for translational medicine. *Nat. Rev. Neurosci.* **12**, 524–538 (2011).

[CAS](#) [PubMed](#) [Google Scholar](#)

3. 3.

Rhie, A. et al. Towards complete and error-free genome assemblies of all vertebrate species. *Nature* <https://doi.org/10.1038/s41586-021-03451-0> (2021).

4. 4.

Jebb, D. et al. Six reference-quality bat genomes reveal evolution of bat adaptations. *Nature* **583**, 578–584 (2020).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

5. 5.

Hoyle, C. H. Neuropeptide families and their receptors: evolutionary perspectives. *Brain Res.* **848**, 1–25 (1999).

[CAS](#) [PubMed](#) [Google Scholar](#)

6. 6.

Acher, R. & Chauvet, J. Structure, processing and evolution of the neurohypophysial hormone–neurophysin precursors. *Biochimie* **70**, 1197–1207 (1988).

[CAS](#) [PubMed](#) [Google Scholar](#)

7. 7.

Ocampo Daza, D., Lewicka, M. & Larhammar, D. The oxytocin/vasopressin receptor family has at least five members in the gnathostome lineage, including two distinct V2 subtypes. *Gen. Comp. Endocrinol.* **175**, 135–143 (2012).

[CAS](#) [PubMed](#) [Google Scholar](#)

8. 8.

Yamaguchi, Y. et al. The fifth neurohypophysial hormone receptor is structurally related to the V2-type receptor but functionally similar to V1-type receptors. *Gen. Comp. Endocrinol.* **178**, 519–528 (2012).

[CAS](#) [PubMed](#) [Google Scholar](#)

9. 9.

Lagman, D. et al. The vertebrate ancestral repertoire of visual opsins, transducin alpha subunits and oxytocin/vasopressin receptors was established by duplication of their shared genomic region in the two rounds of early vertebrate genome duplications. *BMC Evol. Biol.* **13**, 238 (2013).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

10. 10.

Mayasich, S. A. & Clarke, B. L. The emergence of the vasopressin and oxytocin hormone receptor gene family lineage: clues from the characterization of vasotocin receptors in the sea lamprey (*Petromyzon marinus*). *Gen. Comp. Endocrinol.* **226**, 88–101 (2016).

[CAS](#) [PubMed](#) [Google Scholar](#)

11. 11.

Jaillon, O. et al. Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature* **431**, 946–957 (2004).

[ADS](#) [PubMed](#) [Google Scholar](#)

12. 12.

Hubley, R. et al. The Dfam database of repetitive DNA families. *Nucleic Acids Res.* **44**, D81–D89 (2016).

[CAS](#) [PubMed](#) [Google Scholar](#)

13. 13.

Gwee, P. C., Tay, B. H., Brenner, S. & Venkatesh, B. Characterization of the neurohypophyseal hormone gene loci in elephant shark and the Japanese lamprey: origin of the vertebrate neurohypophyseal hormone genes. *BMC Evol. Biol.* **9**, 47 (2009).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

14. 14.

Smith, J. J. et al. The sea lamprey germline genome provides insights into programmed genome rearrangement and vertebrate evolution. *Nat. Genet.* **50**, 270–277 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

15. 15.

Kawada, T., Sekiguchi, T., Itoh, Y., Ogasawara, M. & Satake, H. Characterization of a novel vasopressin/oxytocin superfamily peptide and its receptor from an ascidian, *Ciona intestinalis*. *Peptides* **29**, 1672–1678 (2008).

[CAS](#) [PubMed](#) [Google Scholar](#)

16. 16.

Garrison, J. L. et al. Oxytocin/vasopressin-related peptides have an ancient role in reproductive behavior. *Science* **338**, 540–543 (2012).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

17. 17.

Liutkeviciute, Z., Koehbach, J., Eder, T., Gil-Mansilla, E. & Gruber, C. W. Global map of oxytocin/vasopressin-like neuropeptide signalling in insects. *Sci. Rep.* **6**, 39177 (2016).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

18. 18.

Roch, G. J., Tello, J. A. & Sherwood, N. M. At the transition from invertebrates to vertebrates, a novel GnRH-like peptide emerges in amphioxus. *Mol. Biol. Evo.* **31**, 765–778 (2013).

[Google Scholar](#)

19. 19.

Wicker, T. et al. A unified classification system for eukaryotic transposable elements. *Nat. Rev. Genet.* **8**, 973–982 (2007).

[CAS](#) [PubMed](#) [Google Scholar](#)

20. 20.

Rayko, E., Jabbari, K. & Bernardi, G. The evolution of introns in human duplicated genes. *Gene* **365**, 41–47 (2006).

[CAS](#) [PubMed](#) [Google Scholar](#)

21. 21.

Lovell, P. V. et al. Conserved syntenic clusters of protein coding genes are missing in birds. *Genome Biol.* **15**, 565 (2014).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

22. 22.

Lyons, E. & Freeling, M. How to usefully compare homologous plant genes and chromosomes as DNA sequences. *Plant J.* **53**, 661–673 (2008).

[CAS](#) [PubMed](#) [Google Scholar](#)

23. 23.

Nakatani, Y., Takeda, H., Kohara, Y. & Morishita, S. Reconstruction of the vertebrate ancestral genome reveals dynamic genome reorganization in early vertebrates. *Genome Res.* **17**, 1254–1265 (2007).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

24. 24.

Nakatani, Y. & McLysaght, A. Genomes as documents of evolutionary history: a probabilistic macrosynteny model for the reconstruction of ancestral genomes. *Bioinformatics* **33**, i369–i378 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

25. 25.

Smith, J. J. et al. Sequencing of the sea lamprey (*Petromyzon marinus*) genome provides insights into vertebrate evolution. *Nat. Genet.* **45**, 415–421, e1–e2 (2013).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

26. 26.

Uno, Y. et al. Inference of the protokaryotypes of amniotes and tetrapods and the evolutionary processes of microchromosomes from comparative gene mapping.

PLoS ONE **7**, e53027 (2012).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

27. 27.

Putnam, N. H. et al. The amphioxus genome and the evolution of the chordate karyotype. *Nature* **453**, 1064–1071 (2008).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

28. 28.

Yun, S. et al. Prevertebrate local gene duplication facilitated expansion of the neuropeptide GPCR superfamily. *Mol. Biol. Evol.* **32**, 2803–2817 (2015).

[CAS](#) [PubMed](#) [Google Scholar](#)

29. 29.

Zhang, H. et al. Lampreys, the jawless vertebrates, contain only two ParaHox gene clusters. *Proc. Natl Acad. Sci. USA* **114**, 9146–9151 (2017).

[CAS](#) [PubMed](#) [Google Scholar](#)

30. 30.

Song, Z. & Albers, H. E. Cross-talk among oxytocin and arginine-vasopressin receptors: relevance for basic and clinical studies of the brain and periphery. *Front. Neuroendocrinol.* **51**, 14–24 (2018).

[CAS](#) [PubMed](#) [Google Scholar](#)

31. 31.

Kumar, S. & Hedges, S. B. A molecular timescale for vertebrate evolution. *Nature* **392**, 917–920 (1998).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

32. 32.

Buechi, H. B. & Bridgham, J. T. Evolution of specificity in cartilaginous fish glycoprotein hormones and receptors. *Gen. Comp. Endocrinol.* **246**, 309–320 (2017).

[CAS](#) [PubMed](#) [Google Scholar](#)

33. 33.

Venkatesh, B. et al. Elephant shark genome provides unique insights into gnathostome evolution. *Nature* **505**, 174–179 (2014).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

34. 34.

Smith, J. J. & Keinath, M. C. The sea lamprey meiotic map improves resolution of ancient vertebrate genome duplications. *Genome Res.* **25**, 1081–1090 (2015).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

35. 35.

Miyashita, T. et al. Hagfish from the Cretaceous Tethys Sea and a reconciliation of the morphological–molecular conflict in early vertebrate phylogeny. *Proc. Natl Acad. Sci. USA* **116**, 2146–2151 (2019).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

36. 36.

Green, R. E. et al. Three crocodilian genomes reveal ancestral patterns of evolution among archosaurs. *Science*, **346**, 1254449 (2014).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

37. 37.

Heimberg, A. M., Cowper-Sal Lari, R., Sémon, M., Donoghue, P. C. J. & Peterson, K. J. microRNAs reveal the interrelationships of hagfish, lampreys, and gnathostomes and the nature of the ancestral vertebrate. *Proc. Natl Acad. Sci. USA* **107**, 19379–19383 (2010).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

38. 38.

Kent, W. J. BLAT—the BLAST-like alignment tool. *Genome Res.* **12**, 656–664 (2002).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

39. 39.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).

[CAS](#) [PubMed](#) [Google Scholar](#)

40. 40.

Haug-Batzell, A., Stephens, S. A., Davey, S., Scheidegger, C. E. & Lyons, E. SynMap2 and SynMap3D: web-based whole-genome synteny browsers. *Bioinformatics* **33**, 2197–2198 (2017).

[CAS](#) [PubMed](#) [Google Scholar](#)

41. 41.

Zerbino, D. R. et al. Ensembl 2018. *Nucleic Acids Res.* **46**, D754–D761 (2018).

[CAS](#) [Google Scholar](#)

42. 42.

Morgenstern, B., Prohaska, S. J., Pöhler, D. & Stadler, P. F. Multiple sequence alignment with user-defined anchor points. *Algorithms Mol. Biol.* **1**, 6 (2006).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

43. 43.

Xu, G., Guo, C., Shan, H. & Kong, H. Divergence of duplicate genes in exon–intron structure. *Proc. Natl Acad. Sci. USA* **109**, 1187–1192 (2012).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

44. 44.

Kiezun, A. et al. miRviewer: a multispecies microRNA homologous viewer. *BMC Res. Notes* **5**, 92 (2012).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

45. 45.

Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

46. 46.

Yachdav, G. et al. MSAViewer: interactive JavaScript visualization of multiple sequence alignments. *Bioinformatics* **32**, 3501–3503 (2016).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

47. 47.

Gimpl, G. & Fahrenholz, F. The oxytocin receptor system: structure, function, and regulation. *Physiol. Rev.* **81**, 629–683 (2001).

[CAS](#) [PubMed](#) [Google Scholar](#)

[Download references](#)

Acknowledgements

We thank the leadership of the VGP for early access to high-quality assemblies: in particular, S. Vernes for the pale spear-nosed bat of the Bat1K; G. Zhang for the platypus; and the B10K group for the Anna’s hummingbird and zebra finch assemblies. We thank many colleagues—especially D. Larhammar, Y. Nakatani and J. Smith—for useful discussions on the revised gene nomenclature and findings of this study, and also for their feedback on how to assess genome duplications. C.T. was supported by funds from the Rockefeller University and the Generalitat de Catalunya; G.G. from the Rockefeller University and an NSF Graduate Research Fellowships Program (GRFP); J.A.C. from the Rockefeller University; C.B. from the Spanish Ministry of Economy and Competitiveness/FEDER (grant FFI2016-78034-C2-1-P), MEXT/JSPS Grant-in-Aid for Scientific Research on Innovative Areas 4903 (Evolinguistics: JP17H06379; principal investigator K. Okanoya) and Generalitat de Catalunya (Government of Catalonia) – 2017-SGR-341; and E.D.J. from the Howard Hughes Medical Institute and the Rockefeller University.

Author information

Affiliations

1. Laboratory of Neurogenetics of Language, Rockefeller University, New York, NY, USA

Constantina Theofanopoulou, Gregory Gedman, James A. Cahill & Erich D. Jarvis

2. Section of General Linguistics, University of Barcelona, Barcelona, Spain

Constantina Theofanopoulou & Cedric Boeckx

3. University of Barcelona Institute for Complex Systems, Barcelona, Spain

Constantina Theofanopoulou & Cedric Boeckx

4. ICREA, Barcelona, Spain

Cedric Boeckx

5. Howard Hughes Medical Institute, Chevy Chase, MD, USA

Erich D. Jarvis

Authors

1. Constantina Theofanopoulou
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Gregory Gedman
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. James A. Cahill

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Cedric Boeckx

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Erich D. Jarvis

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Contributions

C.T. conceived the idea, developed the evolutionary proposal, performed BLAST, BLAT, synteny, ancestral, phylogenetic (protein), long non-coding RNA and transposable element analyses, reviewed the literature, made all Figures and Tables and wrote the first draft. G.G. performed BLAST, BLAT and synteny analyses. J.A.C. performed phylogenetic exon tree analyses. C.B. cosupervised the study. E.D.J. conceived the idea, cosupervised the study, proposed the revised nomenclature and helped to write the manuscript. All authors contributed to editing the manuscript.

Corresponding authors

Correspondence to [Constantina Theofanopoulou](#) or [Erich D. Jarvis](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Jeremiah Smith 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 Lineage specific *OT-VT* specializations.

a, Protein phylogenetic tree for *VT* in hagfish and lamprey relative to other vertebrates. Maximum likelihood amino acid phylogenetic tree generated via the Ensembl 'Gene tree' tool (gene tree identifier: ENSGT00390000004511) that uses the Gene Orthology/Paralogy prediction method pipeline. The longest available protein of each species was used. The tree is reconciled with a species tree, generated by TreeBeST. Left, red boxes, inferred gene duplication node; blue boxes, inferred speciation events; turquoise boxes, ambiguous nodes. Right, green bars, multiple amino acid alignment made with MUSCLE; white areas, gaps in the alignment; dark green bars, consensus alignments. We curated the tree and renamed genes using the universal nomenclature proposed in this Article. The tree with the current nomenclature used in the annotations of these genomes can be found at

[http://www.ensembl.org/Multi/GeneTree/Image?
collapse=none;db=core;gt=ENSGT00390000004511](http://www.ensembl.org/Multi/GeneTree/Image?collapse=none;db=core;gt=ENSGT00390000004511).

b, Triplication of the pale spear-nosed bat *OT-VT* region. An approximately 10-gene window of synteny between human, megabat and pale spear-nosed bat is shown. In megabat, *OT*, *VT* and their syntenic genes are found in three different scaffolds (three boxes). In the pale spear-nosed bat with a higher-quality assembly, a syntenic triplication of the *OT-VT* region is found. **c**, gEVAL alignment analyses (<https://geval.sanger.ac.uk/index.html>). This panel shows gapless Pacbio-based long-read contigs (dark blue) and gapless Bionano optical maps (yellow), which span through the entire region with

the *OT* and *VT* duplications in the pale spear-nosed bat, without any noticeable assembly errors.

Extended Data Fig. 2 Lost VTR receptors in human, representative of mammals.

a, Genomic territory of the deleted *VTR2B* in the human genome. The genomic territory before the spotted gar *VTR2B* (top) was found in human chromosome 3 (49–51 Mb), 40 Mb after the location of human *OTR* (bottom). The genomic territory before the spotted gar *VTR2B* was also found in human chromosome 3 (3–5 Mb), 5 Mb before the location of human *OTR*. Text colours denote orthologous genes. Solid black region links two different regions on chromosome 3 in the human genome. **b**, Genomic territory of the deleted *VTR2A* in the human genome. The genomic territory before the chicken *VTR2A* (top) was found in human chromosome 7 (100–115 Mb) (bottom). The genomic territory after the chicken *VTR2A* was found in human chromosome 12 (40–43 Mb). The solid back region links two regions from chromosomes 7 and 12 in the human genome.

Extended Data Fig. 3 Macrosynteny SynFind analyses.

a–d, Comparisons between closely related species (human and chimpanzee) for four receptors, showing maximum syntenies found using this method. **e–h**, Comparisons between intermediately related species (human and chicken) for the same four receptors in **a–d**. **i–l**, Comparisons between distantly related species (human and fish). **m–p**, Comparisons between distantly related non-human species. On the *x* axis, 0 represents the query OTR-VTR in the query organism and the numbers represent the number of genes on the 5' (left) and 3' (right) of the query OTR-VTR in the genome. The *y* axis shows the cumulative number of the matched homologous (orthologous or paralogous) syntenic genes in the reference genome for each reference receptor. For example, in **a** the chimpanzee *OTR* region (red line) shows 17 syntenic gene matches within 20 genes 5' (left) of human *OTR*, and 18 matches within 20 genes 3' (right) of human *OTR*. If the reference OTR-VTR does not show any match, then it is 0 on the *y* axis (for example, the chimpanzee *VTR1B* (shown in green in **a**)); if the reference

OTR-VTR matches only the query OTR-VTR, it reaches 1 (for example, chimpanzee *VTR1A* (shown in blue in **d**) was orthologous only to human *VTR2C*). If the reference OTR-VTR is not orthologous to the query OTR-VTR, but does show gene matches in the neighbouring territory, then it indicates a deletion of the receptor in the query species (for example, chicken *VTR1A* (shown in blue in **f**)).

[Extended Data Fig. 4 Additional chromosomal SynMap2 analyses with lamprey and hagfish.](#)

Bar graphs were created from dot plots. **a**, Additional scaffold in sea lamprey (scaffold 49) with a *VTR* gene is most syntenic with chromosomes of other species containing a *VTR1*-*VTR2* combination, with a 3-gene minimum per 20-gene window criterion. **b**, SynMap2 dot plot between sea lamprey scaffold 49 and scaffolds 10 and 27, with a 3-gene minimum per 10-gene window criterion. **c**, The inshore hagfish scaffold FYBX02010521.1, in which the putative *VTR2* is located, is most syntenic with chromosomes of other species containing a *VTR2A* or *VTR2B* sequence, with a 1-gene minimum per 20-gene window criterion. **d**, Same synteny analyses as in **c**, but with a 2-gene minimum per 20-gene window criterion. **e**, The inshore hagfish scaffold FYBX02010841.1, in which the putative *VTR1* is located, is most syntenic with chromosomes of other species containing a *VTR1A* or *OTR* sequence, with a 1-gene minimum per 20-gene window criterion. **f**, Same synteny analyses as in **e**, but with a 2-gene minimum per 20-gene window criterion.

[Extended Data Fig. 5 Interspecies BLASTn comparisons between exons and introns of all sea lamprey OTR-VTRs in multiple combinations.](#)

a, Three-way comparisons of sea lamprey *VTR1A*, *VTR2A* and *VTR2B* exons (boxes) and introns (lines), and two-way comparisons of *OTRa* and *VTR2B*. **b**, Two-way comparisons of sea lamprey exons and introns of *VTR2B* with *VTR2A*, and *OTRa* with *VTR1A*. Maximum scores and per cent identities are shown for the alignments that exceed a threshold (maximum score > 40 and *E* value < 10^{-4}). Sequence length is shown in bp.

Extended Data Fig. 6 Interspecies non-coding RNA paralogous synteny analyses.

a, Long non-coding RNAs around the *OTR* and *VTR1* genes within human. **b**, Long non-coding RNAs around the OTR-VTRs in sea lamprey. Lines connect the long non-coding RNAs that shared identity beyond a threshold (maximum score > 40 and E value $< 10^{-4}$) in the BLASTn comparisons. Maximum score (bit score) and per cent identity are shown for each pair of long non-coding RNAs. Genomic location is in Mb.

Extended Data Fig. 7 Intraspecies BLASTn comparisons between exons and introns of OTR-VTRs.

a, Two-way comparisons of exons (boxes) and introns (lines) of elephant shark *VTR1B* with sea lamprey *VTR1A* and *OTRa*. **b**, Two-way comparisons of exons and introns of coelacanth *VTR2C* with sea lamprey *VTR2A* and *OTRa*. Maximum scores and per cent identities are shown for the alignments that yielded results beyond a threshold (maximum score > 40 and E value $< 10^{-4}$). Sequence length is shown in bp.

Extended Data Fig. 8 Protein phylogenetic tree for OTR-VTRs with the currently used gene nomenclature.

The same amino acid tree as in Fig. 4b, but labelled with the nomenclature used to date. Further variations within large vertebrate groups, such as tetrapods (for example, *VT1* to *VT4* in birds, *AVPR3* in mammals and *AVPR4* in fish), are not shown.

Extended Data Fig. 9 Microsynteny for *VTR2A* across vertebrates and *VTR2A* and *OTRb* within sea lamprey.

An approximately 14-gene window around the *VTR2A* orthologue across species is shown. In the sea lamprey, *OTRb* is our revised nomenclature for PMZ_0045207-RA/PMZ_0032217-RA on scaffold 49 (Supplementary Table 14), and *VTR2A* is our revision for PMZ_0042163-RA on scaffold 10 (Supplementary Table 14). Orthologous genes are filled with the same

colour; genes found in the territory of the sea lamprey are further outlined in black lines. Further discussion is provided in Supplementary Note 2.

Extended Data Fig. 10 MAFFT alignment of the OTR-VTRs of the best-quality assemblies available (human for *OTR*, *VTR1A*, *VTR1B* and *VTR2C*; zebra finch for *VTR2A*; and clingfish for *VTR2B*).

The MAFFT alignment using the FFT-NS-I parameter was visualized with the MSA viewer⁴⁶. The identifiers and protein sequences used, along with the alignment file can be found in <https://github.com/constantinatheo/otvt>. The functional annotation of transmembrane domains (TM) and intracellular loops (IT) and binding domains is based on findings with *OTR*⁴⁷. Amino acids marked with an asterisk are the OT polar-interacting sites to the receptor; amino acids marked with a # are differences between the VTR1 and VTR2 subfamilies. Colour coding of the amino acids is according to Clustal X (blue, hydrophobic; red, positive charge; green, polar; pink, cysteines; orange, glycines; yellow, prolines; cyan, aromatic; <http://www.jalview.org/help/html/colourSchemes/clustal.html>).

Supplementary information

Supplementary Information

This file contains Supplementary Notes 1-10, Supplementary Tables 1 and 31, and Supplementary Figures 1-7.

Reporting Summary

Supplementary Tables

This file contains Supplementary Tables 2-30. (XLSX)

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



Check for
updates

Cite this article

Theofanopoulou, C., Gedman, G., Cahill, J.A. *et al.* Universal nomenclature for oxytocin–vasotocin ligand and receptor families. *Nature* **592**, 747–755 (2021). <https://doi.org/10.1038/s41586-020-03040-7>

Download citation

- Received: 16 October 2019
- Accepted: 29 May 2020
- Published: 28 April 2021
- Issue Date: 29 April 2021

- DOI: <https://doi.org/10.1038/s41586-020-03040-7>

Further reading

- [Towards complete and error-free genome assemblies of all vertebrate species](#)
 - Arang Rhie
 - , Shane A. McCarthy
 - , Olivier Fedrigo
 - , Joana Damas
 - , Giulio Formenti
 - , Sergey Koren
 - , Marcela Uliano-Silva
 - , William Chow
 - , Arkarachai Fungtammasan
 - , Juwan Kim
 - , Chul Lee
 - , Byung June Ko
 - , Mark Chaisson
 - , Gregory L. Gedman
 - , Lindsey J. Cantin
 - , Francoise Thibaud-Nissen
 - , Leanne Haggerty
 - , Iliana Bista
 - , Michelle Smith
 - , Bettina Haase
 - , Jacquelyn Mountcastle
 - , Sylke Winkler
 - , Sadye Paez
 - , Jason Howard
 - , Sonja C. Vernes
 - , Tanya M. Lama
 - , Frank Grutzner
 - , Wesley C. Warren
 - , Christopher N. Balakrishnan

- , Dave Burt
- , Julia M. George
- , Matthew T. Biegler
- , David Iorns
- , Andrew Digby
- , Daryl Eason
- , Bruce Robertson
- , Taylor Edwards
- , Mark Wilkinson
- , George Turner
- , Axel Meyer
- , Andreas F. Kautt
- , Paolo Franchini
- , H. William Detrich
- , Hannes Svardal
- , Maximilian Wagner
- , Gavin J. P. Naylor
- , Martin Pippel
- , Milan Malinsky
- , Mark Mooney
- , Maria Simbirsky
- , Brett T. Hannigan
- , Trevor Pesout
- , Marlys Houck
- , Ann Misuraca
- , Sarah B. Kingan
- , Richard Hall
- , Zev Kronenberg
- , Ivan SoviÄ
- , Christopher Dunn
- , Zemin Ning
- , Alex Hastie
- , Joyce Lee
- , Siddarth Selvaraj
- , Richard E. Green
- , Nicholas H. Putnam
- , Ivo Gut

- , Jay Ghurye
- , Erik Garrison
- , Ying Sims
- , Joanna Collins
- , Sarah Pelan
- , James Torrance
- , Alan Tracey
- , Jonathan Wood
- , Robel E. Dagnew
- , Dengfeng Guan
- , Sarah E. London
- , David F. Clayton
- , Claudio V. Mello
- , Samantha R. Friedrich
- , Peter V. Lovell
- , Ekaterina Osipova
- , Farooq O. Al-Ajli
- , Simona Secomandi
- , Heebal Kim
- , Constantina Theofanopoulou
- , Michael Hiller
- , Yang Zhou
- , Robert S. Harris
- , Kateryna D. Makova
- , Paul Medvedev
- , Jinna Hoffman
- , Patrick Masterson
- , Karen Clark
- , Fergal Martin
- , Kevin Howe
- , Paul Flicek
- , Brian P. Walenz
- , Woori Kwak
- , Hiram Clawson
- , Mark Diekhans
- , Luis Nassar
- , Benedict Paten

- , Robert H. S. Kraus
- , Andrew J. Crawford
- , M. Thomas P. Gilbert
- , Guojie Zhang
- , Byrappa Venkatesh
- , Robert W. Murphy
- , Klaus-Peter Koepfli
- , Beth Shapiro
- , Warren E. Johnson
- , Federica Di Palma
- , Tomas Marques-Bonet
- , Emma C. Teeling
- , Tandy Warnow
- , Jennifer Marshall Graves
- , Oliver A. Ryder
- , David Haussler
- , Stephen J. OâBrien
- , Jonas Korlach
- , Harris A. Lewin
- , Kerstin Howe
- , Eugene W. Myers
- , Richard Durbin
- , Adam M. Phillippy
- & Erich D. Jarvis

Nature (2021)

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Download PDF](#)

Associated Content

Special

Vertebrate Genomes Project

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-020-03040-7>

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

Platypus and echidna genomes reveal mammalian biology and evolution

[Download PDF](#)

- Article
- Open Access
- [Published: 06 January 2021](#)

Platypus and echidna genomes reveal mammalian biology and evolution

- [Yang Zhou](#) ORCID: [orcid.org/0000-0003-1247-5049^{1,2}](https://orcid.org/0000-0003-1247-5049),
^{nal},
- [Linda Shearwin-Whyatt](#) ORCID: [orcid.org/0000-0002-4504-6534³](https://orcid.org/0000-0002-4504-6534),
^{nal},
- [Jing Li](#)⁴,
^{nal},
- [Zhenzhen Song](#)^{1,5},
- [Takashi Hayakawa](#) ORCID: [orcid.org/0000-0002-1037-0350^{6,7}](https://orcid.org/0000-0002-1037-0350),
- [David Stevens](#)³,
- [Jane C. Fenelon](#) ORCID: [orcid.org/0000-0001-8771-5196⁸](https://orcid.org/0000-0001-8771-5196),
- [Emma Peel](#) ORCID: [orcid.org/0000-0002-2335-8983⁹](https://orcid.org/0000-0002-2335-8983),
- [Yuanyuan Cheng](#) ORCID: [orcid.org/0000-0002-1747-9308⁹](https://orcid.org/0000-0002-1747-9308),
- [Filip Pajpach](#)³,
- [Natasha Bradley](#)³,
- [Hikoyu Suzuki](#)¹⁰,
- [Masato Nikaido](#)¹¹,
- [Joana Damas](#)¹²,
- [Tasman Daish](#)³,
- [Tahlia Perry](#)³,
- [Zexian Zhu](#)⁴,
- [Yuncong Geng](#)¹³,
- [Arang Rhie](#)¹⁴,
- [Ying Sims](#)¹⁵,
- [Jonathan Wood](#)¹⁵,
- [Bettina Haase](#)¹⁶,
- [Jacquelyn Mountcastle](#)¹⁶,
- [Olivier Fedrigo](#)¹⁶,
- [Qiye Li](#) ORCID: [orcid.org/0000-0002-5993-0312¹](https://orcid.org/0000-0002-5993-0312),
- [Huanming Yang](#)^{1,17,18,19},

- [Jian Wang](#)^{1,17},
- [Stephen D. Johnston](#) ORCID: orcid.org/0000-0002-0290-5458²⁰,
- [Adam M. Phillippy](#) ORCID: orcid.org/0000-0003-2983-8934¹⁴,
- [Kerstin Howe](#)¹⁵,
- [Erich D. Jarvis](#) ORCID: orcid.org/0000-0001-8931-5049^{21,22},
- [Oliver A. Ryder](#) ORCID: orcid.org/0000-0003-2427-763X²³,
- [Henrik Kaessmann](#) ORCID: orcid.org/0000-0001-7563-839X²⁴,
- [Peter Donnelly](#) ORCID: orcid.org/0000-0001-9495-3408²⁵,
- [Jonas Korlach](#) ORCID: orcid.org/0000-0003-3047-4250²⁶,
- [Harris A. Lewin](#) ORCID: orcid.org/0000-0002-1043-7287^{12,27,28},
- [Jennifer Graves](#)^{29,30,31},
- [Katherine Belov](#)⁹,
- [Marilyn B. Renfree](#) ORCID: orcid.org/0000-0002-4589-0436⁸,
- [Frank Grutzner](#) ORCID: orcid.org/0000-0002-3088-7314³,
- [Qi Zhou](#) ORCID: orcid.org/0000-0002-7419-2047^{4,32,33} &
- [Guojie Zhang](#) ORCID: orcid.org/0000-0001-6860-1521^{1,2,34,35}

[Nature](#) volume 592, pages756–762(2021)[Cite this article](#)

- 42k Accesses
- 1 Citations
- 989 Altmetric
- [Metrics details](#)

Subjects

- [Comparative genomics](#)
- [Evolutionary genetics](#)

Abstract

Egg-laying mammals (monotremes) are the only extant mammalian outgroup to therians (marsupial and eutherian animals) and provide key insights into mammalian evolution^{1,2}. Here we generate and analyse reference genomes of the platypus (*Ornithorhynchus anatinus*) and echidna (*Tachyglossus aculeatus*), which represent the only two extant monotreme lineages. The nearly complete platypus genome assembly has anchored almost the entire genome onto chromosomes, markedly improving the genome continuity and gene annotation. Together with our echidna

sequence, the genomes of the two species allow us to detect the ancestral and lineage-specific genomic changes that shape both monotreme and mammalian evolution. We provide evidence that the monotreme sex chromosome complex originated from an ancestral chromosome ring configuration. The formation of such a unique chromosome complex may have been facilitated by the unusually extensive interactions between the multi-X and multi-Y chromosomes that are shared by the autosomal homologues in humans. Further comparative genomic analyses unravel marked differences between monotremes and therians in haptoglobin genes, lactation genes and chemosensory receptor genes for smell and taste that underlie the ecological adaptation of monotremes.

[Download PDF](#)

Main

The iconic egg-laying monotremes of Australasia represent one of the three major mammalian lineages. The monotreme lineage comprises two extant families, the semi-aquatic Ornithorhynchidae (platypus) and the terrestrial Tachyglossidae (echidna). At present, the single species of platypus has a restricted distribution in Eastern Australia, whereas four echidna species (*T. aculeatus* and three *Zaglossus* spp.) are present in Australia and New Guinea ([Supplementary Information](#)). Platypuses and echidnas feature radical differences in diet (carnivorous compared with insectivorous), neurophysiology (electroreception-oriented compared with olfaction-oriented), as well as specific intraspecific conflict and defence adaptations¹. Owing to their distinct ecological, anatomical and physiological features, monotremes are interesting mammals well-suited for the study of the evolution of ecological adaptation. Of particular interest are their sex chromosomes, which originated independently from those of therian mammals through additions of autosomes onto an ancestral XY pair, resulting in a multiple sex chromosome system that assembles as a chain during meiosis³.

The previous female platypus genome assembly (OANA5) provided many important insights into monotreme biology and mammalian evolution. However, only about 25% of its sequence was assigned to chromosomes². The incomplete platypus assembly without Y chromosome sequences and lack of an echidna genome have limited the interpretation of the evolution of mammals and monotremes. Here we combined PacBio long-read, 10× linked-read, chromatin conformation (Hi-C) and physical map data to produce a highly accurate chromosome-scale assembly of the platypus genome. We also produced a less-continuous assembly for the short-beaked echidna, which enables us to infer the genomic changes that occurred in the ancestral monotremes and other mammals.

Chromosome-scale monotreme genomes

Our new male platypus genome assembly (mOrnAna1) shows a 1,390-fold improvement for the contig N50 and a 49-fold improvement for the scaffold N50 compared with the previous Sanger-based assembly (OANA5) (Fig. [1a](#)). We performed extensive error correction and manual curation to polish and anchor the assembly at the chromosome scale (Extended Data Fig. [1a, b](#)). Ambiguous chromosome assignments were resolved with fluorescence in situ hybridization (FISH) experiments (Extended Data Fig. [1c, d](#)). We also produced a male echidna genome (mTacAcu1) from a variety of short- and long-insert-size libraries, and further scaffolded it using the same methods as in platypus. The resulting mTacAc1 sequence shows better sequence continuity than OANA5, with a scaffold N50 size of 32.51 Mb (Supplementary Table [2](#)).

Fig. 1: Chromosome assembly of monotreme and mammalian genome evolution.

 **figure1**

a, The contig length distribution among the three monotreme assemblies shows a large improvement in the sequence continuity of the platypus assembly, and at least equivalent quality of the echidna assembly. **b**, Mammalian karyotype evolution trajectory. $2n = 60$ ancestral karyotypes were inferred for the common ancestor of mammals. Conserved blocks were colour-coded in accordance with their chromosomal source in the mammalian ancestor. Numbers of estimated rearrangements are shown for each branch. Silhouettes of the human and opossum are from <https://www.flaticon.com/>. Silhouettes of the platypus and Tasmanian devil are created by S. Werning and are reproduced under the Creative Commons Attribution 3.0 Unported licence (<http://creativecommons.org/licenses/by/3.0/>).

[Full size image](#)

To study the origin and evolution of monotreme sex chromosomes, we greatly improved the assembly of the platypus sex chromosomes. We anchored 172 Mb (92% compared to 22% in OANA5) X-borne sequences to chromosomes (Supplementary Tables 4, 6). This includes one 1.6-Mb segment that was previously misassigned to chromosome 14 (Extended Data Fig. 1e). We determined all of the pseudoautosomal regions (PARs) except for X4, on the basis of the different read coverage between sexes and representation of FISH markers (Supplementary Table 3). We also mapped 92% of the platypus Y-borne sequences to the five Y chromosomes using PacBio reads produced using Y-borne bacterial artificial chromosome (BAC) clones⁴ (Supplementary Tables 5, 6). Owing to a lack of echidna linkage markers, we used the platypus X chromosomes as a reference to anchor a similar length (177 Mb, 96%) of X chromosomes and identified 8.6 Mb Y-borne sequences in echidna.

In the final curated platypus genome (mOrnAna1) 98% of the sequence was assigned to the 21 autosomes, 5 X and 5 Y chromosomes (Supplementary Table 7), with putative telomeres and centromeres annotated for half of the chromosomes (Supplementary Table 8). mOrnAna1 fills around 90% of the gaps in OANA5 (Supplementary Table 9), recovering 161 Mb of previously missed genomic sequences, most of which are long interspersed nuclear elements (LINE)/L2 and short interspersed nuclear elements (SINE)/MIR (Supplementary Tables 10, 11). We also removed 68 Mb of redundant sequences in OANA5 (Extended Data Fig. 1f-h). The repeat elements comprising about half of the monotreme genomes are dominated by LINE/L2 elements that are more similar to reptile genomes than therian mammals (which comprise mostly LINE/L1)⁵ (Supplementary Table 12). The highly continuous assembly also substantially improves gene annotation. We identified 20,742 and 22,029 protein-coding genes in mOrnAna1 and mTacAcul, respectively (Supplementary Table 13). Specifically, 19,576 coding exons from 8,303 platypus genes were recovered from the gapped regions of OANA5. Among them, 454 genes were completely missed in OANA5, and 3,961 fragmented genes in OANA5 now have complete open-reading frames. We corrected 2,395 genes that were previously split or misannotated in OANA5 (Extended Data Fig. 1i,j).

Insights into mammalian genome evolution

Our phylogenomic reconstruction shows that monotremes diverged from therians around 187 million years ago, and the two monotremes diverged around 55 million years ago (Extended Data Fig. 2a). This estimate provides a date for the monotreme–therian split that is earlier than previous estimates (about 21 million years ago)², but agrees with recent analyses of few genes and fossil evidence⁶. We also inferred that monotremes had similar genome substitution rates (approximately 2.6×10^{-3} substitutions per site per million years) compared with other mammals

(Supplementary Table 15). About 14 Mb of mammalian specific highly conserved elements were identified by comparison among vertebrates (Methods): around 90% of elements were located in non-coding regions (Extended Data Fig. 2c), and are associated with genes that are enriched in processes such as brain development (Extended Data Fig. 2d,e, [Supplementary Results](#) and Supplementary Tables 18–20).

Next we used chromosome information from human, opossum, Tasmanian devil, platypus, chicken and common wall lizard genomes to reconstruct the mammalian ancestral karyotype (Methods). This analysis reveals 30 mammalian ancestral chromosomes (MACs) ($2n = 60$) at a resolution of 500 kb, covering around 66% of the human genome and approximately 67% of the platypus genome (Fig. 1b and Supplementary Tables 24–26). Of these, 25 MACs were maintained without breaks in a single chromosome of the therian ancestor, and 17 of them have fused with other MACs in therians. Sixteen MACs were still maintained in a single human chromosome, but only MAC28 had not undergone any intrachromosomal rearrangements during therian evolution (Extended Data Fig. 2f,g). We detected at least 918 chromosome breakage events, and confirmed that the X chromosome in humans was derived from the fusion of an original therian X chromosome with an autosomal region after the divergence from marsupials⁷ (Fig. 1b and Extended Data Fig. 2f,g). The five X chromosomes in platypus were derived from different MACs by multiple fusion and translocation events.

We found that gene families associated with the immune response and hair growth were expanded considerably in the mammalian ancestor, perhaps contributing to the evolution of immune adaptation and fur, respectively, in mammals (Supplementary Table 30). We further manually annotated major histocompatibility complex (MHC) genes and other immune genes ([Supplementary Results](#)). As in nonmammalian vertebrates, the monotreme MHC class Ia genes colocalize with antigen-processing genes and MHC class II genes (Extended Data Fig. 3a and Supplementary Table 31). The defensin genes gave rise to unique defensin-like peptides (OavDLP genes) in platypus venom⁸. By contrast, echidna has only one single *OavDLP* pseudogene (Extended Data Fig. 3f-h), suggesting the loss of the key venom gene family in this species.

Monotreme sex chromosome evolution

To elucidate the detailed genomic composition of the monotreme sex chromosomes, we compared regions that share sequences between the sex chromosomes—that is, the PARs—with regions that have become sexually differentiated (SDRs). PAR boundaries show a sharp shift in the female-to-male sequencing coverage ratio as expected (Fig. 2a and Extended Data Fig. 4a). Both monotremes showed generally nonbiased gene expression levels between sexes within PARs, but pronounced female-

biased expression within SDRs, indicating the absence of complete chromosome-wide dosage compensation in monotremes as previously suggested⁹ (Extended Data Fig. 4b).

Fig. 2: Origin and evolution of the sex chromosomes of the platypus.



a, Genomic composition of the platypus sex chromosomes. From the outer to inner rings: the X chromosomes with the PARs (light colours) and SDRs (dark colours) labelled; the assembled Y chromosome fragments within SDRs showing the colour-scaled sequence divergence levels with the homologous X chromosomes; female-to-male (F/M) ratios of short sequencing-read coverage in non-overlapping 5-kb windows; F/M expression ratios (each red dot is one gene) of the adult kidney and the smoothed expression trend; and GC content in non-overlapping 2-kb windows. In addition, we labelled the positions on the X chromosome ring of the gametologue pairs that have suppressed recombination before the divergence of monotremes ('shared', orange triangles) or after the divergence ('independent', blue triangles). **b**, Homology between X and Y chromosomes of platypus. In particular, most of Y5 shows homology with X1 and X2, which suggests an ancestral ring conformation of the platypus sex chromosomes. We also labelled the position of the putative sex-determining gene *AMH*. The platypus silhouette is created by S. Werning and is reproduced under the Creative Commons Attribution 3.0 Unported licence (<http://creativecommons.org/licenses/by/3.0/>).

[Full size image](#)

The short PARs of platypus chromosomes X2–X5 have a significantly higher GC content (one-sided Wilcoxon rank-sum test, $P < 0.01$) than the SDRs or the longer

PARs (Extended Data Fig. 4c), which probably reflects strong GC-biased gene conversion that is caused by a high recombination rate¹⁰. This is similar to the pattern of the short GC-rich human PAR, the recombination rate of which is 17-fold higher than the genome-wide average¹¹. Notably, chicken orthologous sequences of these monotreme PARs are all located on the microchromosomes, which also have a high GC content¹² (one-sided Wilcoxon rank-sum test, $P < 0.01$) (Extended Data Fig. 4c, d). This highly conserved recombination landscape might be partially selected in monotremes for maintaining the sequence polymorphism and balanced dosage of MHC genes, which reside in the PARs of the chromosome X3–Y3 and Y4–X5 pairs in platypus¹³ (Extended Data Fig. 3a). The regional selection for high recombination may also counteract further expansion of SDRs on these sex chromosomes.

Sex chromosomes of both eutherians and birds formed through stepwise suppression of recombination, resulting in a pattern of pairwise sequence divergence between SDRs termed ‘evolutionary strata’^{14,15}. We identified at least seven strata in monotremes, named S0 to S6 from the oldest to the youngest strata (Fig. 2a and Extended Data Fig. 4a), by ranking their levels of pairwise synonymous sequence divergence between the X–Y gametologues and the phylogeny (Extended Data Fig. 5a, b). All but the most recent strata (S5 and S6) are shared by platypus and echidna. However, the PARs that border S5 and S6, as well as the shorter PARs of chromosomes X2 and X5 (Extended Data Fig. 5c, d), formed independently after their divergence. Overall, the distribution of evolutionary strata suggested a time order of incorporating different ancestral autosomes into the sex chromosome chain: it started from the S0 region of X1 containing a sex-determining gene (see below), followed by X2, X3 and X5. X4 and individual regions of X3 and X1 underwent suppression of recombination after the monotreme divergence.

Despite episodes of independent evolution, most sex chromosome regions of the platypus and echidna are homologous (Extended Data Fig. 6a), suggesting that the complex formed in the monotreme ancestor¹⁶. To reconstruct its origin, we projected the platypus sex chromosomes onto their chicken homologues (Supplementary Table 39). This refined homology map (Extended Data Fig. 4d) suggests that both fusions and reciprocal translocations among the ancestral micro- and macrochromosomal fragments gave rise to the monotreme sex chromosome complex. The platypus X chromosomes contain homologous sequences of the entire or partial chicken microchromosomes 11, 16, 17, 25 and 28. These microchromosomes also have orthologues in the spotted gar¹⁷, suggesting that they were ancestral vertebrate microchromosomes, and fused in the ancestral monotreme or mammalian chromosomes. Evidence of reciprocal translocations came from the observation that parts of every two neighbouring sex chromosomes are homologous to two adjacent regions of the same chicken chromosome (Extended Data Fig. 6c, d). For example, platypus chromosomes X1 and X2 are both homologous to parts of chicken

microchromosome 12 and chromosome 13, whereas X2 and X3 are both homologous to chicken chromosome 2.

Notably, X1 at one end of the meiotic chain and Y5 at the other share this alternately overlapping relationship, and both are homologous to chicken microchromosome 28. Indeed, most of the genes on Y5 are not found on its pairing partner X5, but on X1 (Fig. 2b and Supplementary Table 40). Chromosomes X1 and Y5 do not pair at meiosis, but this homology suggests that the origin of the extant monotreme sex chromosome complex involved the opening of the ancestral chromosomal ‘ring’ as degeneration proceeded¹⁸. A conserved vertebrate sex-determining gene, the anti-Mullerian hormone, is located on chromosome Y5 (*AMHY*) and S0 of chromosome X1 (*AMHX*)¹⁴ (Fig. 2b). The ancestral X1–Y5 pairing region that encompasses *AMH* could, therefore, be the site at which homologous recombination was first suppressed. The degeneration of chromosome Y5 then caused the loss of homology with X1 and led to the break of the chromosome ring. Indeed, synonymous substitution rates (dS) between the retained X1–Y5 gametologue pairs are significantly higher (one-sided Wilcoxon ranked-sum test, $P < 0.01$) than those of any other sex chromosome pairs (Extended Data Fig. 6e). A chromosome ring configuration has been reported in plants¹⁹, but not in any animal species. Alternatively, the ancestral ring structure might have evolved after the emergence of the proto-X1–Y5 pair by translocations that involve other autosomes, so that sexually antagonistic alleles could be linked to the sex-determining genes²⁰.

Interactions between sex chromosomes

The platypus sex chromosomes exhibit an unusual association with each other compared to autosomes during and after meiosis²¹. As little is known about their spatial organization in platypus somatic cells, we investigated this using Hi-C data (male liver) and chromosomal FISH with sex-chromosome-specific and autosomal BAC probes (male fibroblasts). Notably, Hi-C data showed that chromosomes Y2 and Y3 undergo frequent interchromosomal interactions, whereas autosomes confine their interactions mostly within chromosomes (Fig. 3a and Extended Data Fig. 7a–d). FISH showed that chromosomes Y2 and Y3 signals overlapped more frequently (5.2- and 7.6-fold) than signals between chromosomes Y2 and X1 or Y2 and an autosome (chromosome 17) ($P = 8.67 \times 10^{-4}$ and 8.57×10^{-5} , respectively) (Fig. 3b, c and Supplementary Table 41). These interactions allow us to predict a zigzag three-dimensional conformation of the sex chromosomes at interphase (Extended Data Fig. 7e). A similar pattern was also present in echidna (Extended Data Fig. 7f). Notably, the high interaction frequency is conserved in human orthologous autosomal regions (Fig. 3a), suggesting functional importance unrelated to the evolution or function of sex chromosomes.

Fig. 3: Interactions between the platypus sex chromosomes.

 figure3

a, Interchromosomal interactions among the platypus sex chromosomes detected by Hi-C data of liver tissue in platypus (top) and human (bottom). The bars between the Hi-C panels show the platypus sex chromosomes and their orthologues in the human genome. Grey, intrachromosomal interactions; red, interchromosomal interactions. Red lines link the regions with significantly high interchromosomal interactions. The interchromosomal interactions seem to be conserved in mammals, as indicated by the homologous chromosomal fragments of the human and platypus sex chromosomes and their Hi-C contact patterns. **b**, FISH with BAC probes to detect sex chromosomes Y2, Y3 or X1 and autosome chromosome 17 (*WSB1*) in interphase platypus fibroblasts. Examples show no interaction between chromosomes Y2 and Y3 (top, $n = 593$, 3 independent experiments) and interaction (bottom, $n = 56$, 3 independent experiments). Scale bars, 10 μm . **c**, The significantly higher frequency of interaction between Y2 and Y3 than that between Y2 and X1, and between Y2 and *WSB1* (chromosome 17). $n = 185, 206, 258$ cells for the three independent replicate experiments of Y2–Y3, $n = 258, 250, 205$ cells for the three independent replicate experiments of Y2–X1, $n = 298, 262, 220$ cells for the three independent replicate experiments of Y2–*WSB1*. Data are mean \pm s.d. *** $P < 0.001$ (Y2–Y3 versus Y2–X1, $P = 0.0004675$; Y2–Y3 versus Y2–*WSB1*, $P = 6.376 \times 10^{-5}$), one-sided Fisher's exact

test. **d**, Putative CTCF-binding-site density plot showing its enrichment among homologous regions in the platypus, human and chicken genomes.

[Full size image](#)

We further examined the distribution of putative binding sites of the CTCF protein, which is usually enriched at the boundaries of topologically associated domains (TADs) and mediates both intra- and interchromosomal interactions²². This revealed considerable enrichment of putative CTCF-binding sites at the TAD boundaries of the platypus genome (Extended Data Fig. 7g), which are more enriched along the interacting sex chromosomes X2 and X4, as well as along their orthologous regions in human and chicken (Fig. 3d and Extended Data Fig. 7h). These results suggest that an ancestral interaction landscape facilitated by local enrichment of CTCF-binding sites could have promoted the reciprocal translocations between spatially adjacent autosomal fragments that gave rise to the sex chromosome complex in the monotremes.

Eco-evolutionary adaptation of diet

Platypuses consume aquatic invertebrates whereas echidnas feed predominantly on social insects. Although the recent ancestor of monotremes had adult teeth, both extant monotremes lack teeth²³. Of eight genes involved in tooth development²⁴, four genes were lost in both monotreme genomes, suggesting that the loss occurred in their recent common ancestors (Extended Data Fig. 8a and Supplementary Table 42), consistent with other toothless or enamel-less eutherians²⁵. Echidnas (but not platypuses) further lost two enamel genes. Analysis of genes involved in stomach function revealed that the considerable loss of digestive genes (reported in platypus²⁶) is shared with echidna and probably occurred in the monotreme ancestor, although *NGN3*—which is essential for stomach and pancreas development—has been maintained in both species (Extended Data Fig. 8b–g and Supplementary Table 43).

Chemosensory systems mediate animal behaviour that is essential for survival and reproduction through the direct interaction with environmental chemical cues²⁷. For example, eutherian mammals have more than 25 copies of bitter taste receptor genes (*TAS2R* genes)^{27,28}, whereas this gene family is considerably smaller in monotremes (Extended Data Fig. 9a) with only 7 in platypus (Supplementary Tables 44, 45). The number is reduced to three in echidna (Fig. 4a and [Supplementary Results](#)). This reduction is also observed in pangolins, which suggests convergent evolution that results from the insectivore diet of both echidnas and pangolins²⁹.

Fig. 4: Genomic features related to biological characteristics of the monotremes.

 **figure4**

a, Differences in numbers of TAS2R, OR and V1R genes between platypus and echidna. **b**, Phylogeny and synteny of the *HP* gene. Regions are not drawn to scale. **c**, Synteny conservation of the region surrounding caseins (CSN genes) and the ancestral teeth genes (*ODAM*, *FDCSP*, *AMTN*, *AMBN* and *ENAM*). Silhouettes of the human, opossum, koala and frog are from <https://www.flaticon.com/>. Silhouettes of the platypus and Tasmanian devil are created by S. Werning and the emu silhouette is created by D. Naish (vectorized by T. M. Keesey); all three silhouettes are reproduced

under the Creative Commons Attribution 3.0 Unported licence (<http://creativecommons.org/licenses/by/3.0/>).

[Full size image](#)

The nasal cavity of the platypus is closed off during diving and the size of the main olfactory bulb of the platypus is much smaller than that of the echidna¹. Consistent with this, the number of olfactory receptors (OR genes) in platypus (299) is much smaller than in echidna (693) (Fig. 4a and Supplementary Table 46). The difference in the large olfactory bulb and OR repertoire in echidna may contribute to the ability to search for odours of underground prey, whereas the platypus relies on electroreception to detect prey in the water. However, the size of the accessory olfactory bulb is larger in the platypus than in the echidna¹. The accessory olfactory bulb receives projections from the vomeronasal organ, and there is a marked expansion of the number of vomeronasal type-1 receptors (V1R genes) in the platypus (262) compared with the echidna (28) (Fig. 4a and Supplementary Table 47). Vomeronasal receptors probably have important roles in courtship, parental care, induction of lactation and milk ejection in monotremes²³. Therefore, the diversification of the olfactory bulb and accessory olfactory bulb systems in monotremes provide an interesting example of the eco-evolutionary trade-off. V1R amplification has been associated with the size of the vomeronasal organ and nocturnal activity³⁰. This is also consistent with the fact that the platypus closes its eyes when diving and therefore relies entirely on other senses underwater and in the burrow.

Haemoglobin degradation in monotremes

The semi-aquatic lifestyle of the platypus is supported by particularly high haemoglobin levels and large numbers of small red blood cells³¹. The haemoglobin–haem detoxification system in mammals provides efficient clearance to minimize oxidative damage³² in which haptoglobin is the haemoglobin chaperone³² and free haem is bound by haemopexin and alpha-1 microglobulin³³.

Both the haemopexin and alpha-1 microglobulin genes are found in the monotreme genomes, whereas the haptoglobin gene is absent (Fig. 4b, Extended Data Fig. 10a, b and Supplementary Table 48), which suggests that monotremes evolved a haemoglobin clearance system that is different from that of other mammals. Haptoglobin evolved in the common ancestor of vertebrates from an immune gene of the MASP family³³ but has neofunctionalized in mammals to bind to haemoglobin with a higher affinity and to bind to the CD163A receptor, which is also absent in monotremes, for clearance in macrophages³⁴. The absence of the haptoglobin gene and CD163A in monotremes suggests that the neofunctionalization of haptoglobin

happened after the divergence of monotremes from therians, not before it as previously thought³⁴, and long after the evolution of enucleated red blood cells in the common ancestor of mammals³⁵. Several nonmammalian vertebrates have lost haptoglobin, including chicken³⁴ (Fig. 4b), in which an alternative, secreted CD163 family member, *PIT54*, is the haemoglobin-binding chaperone³³. Phylogenetic analysis shows that monotremes lack genes that cluster with haptoglobin in the MASP family or a *PIT54* orthologue (Extended Data Fig. 10c–e and Supplementary Table 50). We confirmed the expansion of the CD163 family in platypus² (ten members) and found five in echidna, compared with two and three in humans and mice, respectively (Extended Data Fig. 10e, f). As mammalian CD163A can bind to haemoglobin in the absence of haptoglobin³⁶ and one CD163 family member has become the haemoglobin chaperone in chicken, the CD163 family protein(s) may have evolved this role in monotremes.

Transition from oviparity to viviparity

Monotremes provide the key to understanding how viviparity evolved in mammals. They are not as dependent on egg proteins as egg-laying avian and reptilian species owing to their nutrient acquisition from uterine secretions^{23,37}, and the subsequent reliance of the young on lactation. Whereas reptiles have three functional copies of the major egg protein vitellogenin (*VTG*)³⁸, in monotremes we found only one functional copy (*VTG2*) (Extended Data Fig. 10g and Supplementary Table 52) and a partial sequence for *VTG1*.

Similar to marsupials, monotremes have an extended lactation period and the composition of the milk changes dynamically as the development progresses to match the changing needs of the young³⁷. *SPINT3*, a major milk-specific protein that is present in early lactation of therians with a probable role in the protection of immunoincompetent young in marsupials³⁹, is absent in monotremes. Syntenic analysis confirmed that this region is conserved in platypus but contains two copies of a new protein that contains a Kunitz domain (Extended Data Fig. 10h and Supplementary Table 53). The Kunitz family is a rapidly evolving family, and one of the new members could have a immunoprotective function similar to *SPINT3* in monotremes.

The monotreme genomes contain most of the milk genes that have been identified in therian mammals^{38,40}. Most mammals have three casein genes⁴¹, which encode the most abundant milk proteins secreted throughout lactation (Fig. 4c). In addition to these genes, monotremes have extra caseins that are not found in therian mammals, with unknown functions, an extra copy of *CSN2* (*CSN2B*) (previously reported⁴⁰) and *CSN3* (*CSN3B*) in platypus (described here), which has the classic structure of *CSN3*⁴² (Extended Data Fig. 10i and Supplementary Table 54).

All caseins are members of the secretory calcium-binding phosphoprotein (SCPP) gene family and are thought to have evolved from other SCPP genes, namely the teeth-related gene *ODAM* through its derivatives *FDCSP* and *SCPPPQI*⁴². As reported above (see ‘Eco-evolutionary adaptation of diet’), extant monotremes appear to have lost both *ODAM* and *FDCSP*. Syntenic analysis showed that the additional monotreme casein genes (*CSN2B* and *CSN3B*) are found in the same therian chromosomal region as *ODAM* and *FDCSP* and within the casein locus (Fig. 4c), providing further evidence that caseins evolved from odontogenic genes.

Summary

Complete and accurate reference genomes and annotations are critical for evolutionary and functional analyses. It remains a challenge to produce a highly accurate chromosome-level assembly, particularly for differentiated sex chromosomes. We have produced a high-quality platypus genome using a combination of single-molecule sequencing technology and multiple sources of physical mapping methods to assign most of the sequences to a chromosome-scale assembly. This permits better-resolved analyses of the origin and diversification of the complex sex chromosome system that evolved specifically in monotremes. We delineate ancient and lineage-specific changes in the sensory system, haemoglobin degradation and reproduction that represent some of the most fascinating biology of platypus and echidna. The new genomes of both species will enable further insights into therian innovations and the biology and evolution of these extraordinary egg-laying mammals.

Methods

Data reporting

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Ethics and sample collection

Pmale08, Pmale09 and Emale12 were collected under AEC permits S-49-2006, S-032-2008 and S-2011-146 at Upper Barnard River (New South Wales, Australia) during the breeding season. Emale01 was collected under San Diego Zoo Global IACUC approval 18-024 and vouched at San Diego Natural History Museum.

Sequencing and assembling

Skeletal muscle of Pmale09 was used for PacBio, 10X and BioNano genome sequencing and the liver of Pmale09 was used for Hi-C (Phase Genomics); the liver of Pmale08 was used for Chicago Hi-C (Dovetail Genomics). Heart muscle of Emale01 was used for a variety of library construction and Illumina sequencing analyses. Muscle of Emale12 was used for 10X and BioNano genome sequencing and liver of Emale12 was used for Hi-C (Phase Genomics). Echidna RNA was extracted from brain, cerebellum, kidney, liver, testis and ovary and sequenced using a previously published procedure⁴³. Platypus Y chromosome BAC isolation via hybridization was performed using a previous published procedure⁴ and sequenced with PacBio. The platypus genome was assembled following VGP assembly pipeline v.1.0. The echidna genome was assembled using Platanus⁴⁴ (v.1.2.1) and followed by three steps of scaffolding in the order of 10X, BioNano and Hi-C. Manual curation was performed for both assemblies. Details are available in the [Supplementary Methods](#).

Sex-borne sequence identification

Female and male reads were mapped to the genome using BWA ALN⁴⁵ (v.0.7.12). The read depth of each sex was calculated in 5-kb non-overlapping windows to identify X-borne sequences and 2-kb non-overlapping windows to identify Y-borne sequences, normalized against the median depth. To identify X-borne sequences, we calculated the female-to-male (F/M) depth ratio of regions that were covered by both sexes in each scaffold, requiring a minimum coverage of 80%, and assigned sequences to X-borne if the depth ratio ranged between 1.5 and 2.5. To identify Y-borne sequences, we calculated the F/M depth ratio as well as the F/M coverage ratio and assigned scaffolds to Y-borne if either ratio was within the range of 0.0–0.3. Parameter evaluation details are available in the [Supplementary Methods](#).

Chromosome anchoring

We collected 75 BAC and 179 marker genes (Supplementary Table 3) and ordered them according to their relative order from those papers. Protein sequences of the gene markers were compared to mOrnAna1 using TBLASTN and the best hit was kept, after which the markers were analysed using GeneWise⁴⁶ (v.2.4.1) to obtain the location within a scaffold. BAC-end reads were mapped to the assembly using BWA MEM⁴⁷ (v.0.7.12) and the best hits were kept. We also used the anchored sequences of OANA5 except for the sequence of chromosome 14 to anchor the scaffolds into chromosomes. Scaffolds were orientated and ordered first based on the order of FISH or gene markers then on the order in OANA5. All identified PARs were included in chromosome X. We collected assembled Y contigs from a previously published study⁴ and generated some Y-BAC PacBio sequencing data. Assembled Y contigs were mapped to the platypus assembly using BWA MEM and Y-BAC PacBio reads were

mapped using minimap2 (v.2.13)⁴⁸. As evidence of both Y2 and Y3 were found on scaffold_229_arrow_ctg1 and scaffold_269_arrow_ctg1 and the covered regions overlapped, these two scaffolds were excluded from the chromosome Y classification. Classified Y-borne scaffolds failed to anchor and orient due to the lack of information. We also curated and anchored some echidna X-borne scaffolds to chromosome X based on Mashmap⁴⁹ (v.2.0) one-to-one results with platypus⁵⁰.

Annotation

We identified repetitive elements in both assemblies using the same pipeline, which included homologue-based and de novo prediction. For the homology-based method, we used default repeat library from Repbase (v.21.11)⁵¹ for RepeatMasker (v.4.0.6)⁵², trf (v.4.07)⁵³ and Proteinmasker (v.4.0.6)⁵² to annotate. For the de novo method, we first ran RepeatModeler (v.1.0.8) to construct the consensus sequence library for each monotreme using their genome as input, then aligned the genome against each consensus library to identify repeats using RepeatMasker. Gene annotation was performed by merging the homology, de novo prediction and transcriptome analyses to build a consensus gene set of each species. Protein sequences from human, mouse, opossum, platypus, chicken and green lizard (*Anolis carolinensis*) from Ensembl⁵⁴ (release 87) were aligned to the genome using TBLASTN⁵⁵ (v.2.2.26) ($e < 1 \times 10^{-5}$). Candidate gene regions were refined using GeneWise for more accurate gene models. We randomly selected 1,000 high score homology-based genes to train Augustus⁵⁶ (v.3.0.3) for de novo prediction on a repeat N-masked genome. We also mapped RNA-sequencing reads of the platypus from a previously published study⁵⁷ and echidna to their respective assemblies using HISAT2⁵⁸ (v.2.0.4), and constructed transcripts using stringTie⁵⁹ (v.1.2.3). Results from these three methods were merged into a nonredundant gene set. Possible retrogenes were filtered according to their hit to SwissProt database⁶⁰ (release 2015_12) or Iprscan⁶¹ (v.5.16-55.0). We used the SwissProt database ($e < 1 \times 10^{-5}$) to annotate the function of the genes. Iprscan was used to annotate the GO of genes. Detailed descriptions of the manual annotation, curation and phylogenetic analysis of genes related to imprinting, immune system, reproduction and haemoglobin degradation can be found in the [Supplementary Methods](#).

Gap analysis

We identified gap-filling regions using an alignment-based strategy similar to a previously published study⁶². We considered gaps for which both flanking regions mapped to mOrnAna1 as closed gaps. Only properly closed gaps defined by (1) both flanking regions were aligned but did not overlap and (2) closed gap size were within

100 times the estimated gap size in OANA5 were considered for repeat and gene improvement analysis.

Redundant sequences analysis

We performed two rounds of Mashmap with parameters ‘-f one-to-one -s 2000’ using mOrnAna1 as reference and OANA5 as query. A one-to-one relationship was obtained in the first round of Mashmap. In the second round of Mashmap, those OANA5 sequences that were unmapped in the first round of mapping were used as query. Candidate redundant sequences were obtained from the second round Mashmap result, but excluded regions that were gaps in OANA5. Female and male reads were then mapped to OANA5 and mOrnAna1 using BWA ALN and normalized by the mode depth.

Gene set comparison

We performed LASTZ⁶³ (v.1.04.00) alignment using OANA5 as reference with parameter set ‘--hspthresh=4500 --gap=600,150 --ydrop=15000 --notransition --format=axt’ and a score matrix for the comparison of closely related species to generate a chain file for gene location liftover from OANA5 to mOrnAna1. Gene coordinates in OANA5 were first converted to mOrnAna1 using in-house-generated scripts with the chain file. We searched for overlap between the converted OANA5 gene set and mOrnAna1 gene set. Fragmented genes were defined as multiple converted OANA5 genes that overlapped with a single mOrnAna1 gene. A one-to-one gene pair between the two gene sets was defined as the liftover of the OANA5 gene when it overlapped with only one mOrnAna1 gene. Only one-to-one pairs were used for the comparison of open-reading frame completeness. We defined a gene as having a complete open-reading frame if its first codon is a start codon and the last codon is a stop codon.

Identification of one-to-one orthologues and synteny blocks between the human sequence and sequences of other species

We defined one-to-one orthologues between the human sequence and the sequences of other species by considering both reciprocal best BLASTP hits (RBH) and synteny, taking the human sequence as reference, as previously described⁶⁴. First, we conducted BLASTP for all protein sequences from human and other species including mouse, opossum, platypus, echidna, chicken and green lizard with an *e*-value cut-off of 1×10^{-7} , and combined local alignments with the SOLAR (<http://treesoft.svn.sourceforge.net/viewvc/treesoft/>). Next, we identified RBH orthologues between human and every other species on the basis of the following

parameters: alignment score, alignment rate and identity. From these RBH orthologues, we retained those pairs with conserved synteny across species. Synteny was determined based on their flanking genes. If RBH orthologous gene pairs shared the same flanking genes, we retained the genes for downstream analyses. Finally, we merged pairwise orthologue lists according to the human coordinates. In this way, we produced the final one-to-one orthologue set across species.

We used the human genome as the reference and aligned it with other species using LASTZ with parameter set ‘--hsptthresh=4500 --gap=600,150 --ydrop=15000 –notransition --format=axt’ and a score matrix for the comparison of closely related species. Alignments were converted into ‘chain’ and ‘net’ results with different levels of alignment scores using utilities of the UCSC Genome Browser (<http://genomewiki.ucsc.edu/index.php/>), and the pairwise synteny blocks between genomes of each species and the human genome were extracted according to the net result. Only alignments larger than 10 kb were kept. The synteny blocks were further cleaned of overlapping genes. N50 and the total length of the synteny block inferred from each human–species pair were calculated based on the human coordinates.

Phylogenetic analysis

The phylogenetic tree was constructed using concatenated four-degenerated sites from the 7,946 one-to-one orthologues using RaxML⁶⁵ (v.8.2.4) with parameter set ‘-m GTRCAT -# 100 -p 12345 -x 12345 -f a’ and chicken and green lizard were specified as the outgroup. MCMCTree in PAML⁶⁶ (v.4.7) was used to estimate divergence time of each species with calibration points obtained from a previously published study⁶⁷ using the same data. Points and time range included the most recent common ancestor of human–mouse, 85–94 million years ago; human–opossum 150–167 million years ago; human–platypus, 163–191 million years ago, human–chicken, 297–326 million years ago, anole–chicken, 276–286 million years ago. The seed used for MCMC was 1192664277.

Substitution rate analysis

We first performed pairwise whole-genome LASTZ alignment using 12 mammals (*Macaca mulatta*, *Tupaia belangeri*, *Mus musculus*, *Canis lupus familiaris*, *Myotis lucifugus*, *Bos taurus*, *Sorex araneus*, *Loxodonta africana*, *Dasyurus novemcinctus*, *Monodelphis domestica*, *O. anatinus* and *T. aculeatus*) with the human genome as the reference genome, with the parameter set ‘--step=19 --hsptthresh=2200 -inner=2000 --ydrop=3400 --gappedthresh=10000 --format=axt’ and a score matrix for the comparison of distantly related species. Pairwise alignments were merged using MULTIZ⁶⁸ (v.11.2). The four-degenerated site alignment was extracted based on the human gene set (Ensembl release 87), concatenated and fed to phyloFit in the PHAST

package⁶⁹ (v.1.5) for the calculation of branch lengths (substitution per site). The substitution rate was calculated by dividing the branch length to the mammalian common ancestor to the mammal–reptile divergence time.

Gene family analysis

Gene families across the seven species were generated using orthoMCL⁷⁰ (v.2.0.9) with BLASTP results ($e < 1 \times 10^{-7}$) and was fed to CAFÉ⁷¹ (v.4.2) along with the phylogenetic tree. We first estimated the assembly error by excluding families with more than 100 members. Then the estimated rate was used to infer the family size at every node for each family. The ancestral node gene number of families with more than 100 members among extant species were inferred separately. We extracted genes based on the human gene set for GO enrichment (χ^2 test) of the significantly expanded family (Viterbi $P < 0.05$) for the mammalian ancestor. A false-discovery rate (FDR) adjustment was used for multiple-test corrections in GO enrichment analyses.

Mammalian-specific highly conserved element analysis

We used the same MULTIZ alignment of the substitution rate analysis and identified mammalian-specific highly conserved elements (MSHCEs) using a similar strategy as has previously been described⁷². At least 80% of species and at least one species in eutherians, marsupials and monotremes were required to be present in alignments. Type-I MSHCEs were defined as HCEs to which no outgroup could be aligned; type-II MSHCEs were HCEs that were significantly conversed ($P < 0.01$) in mammals compared to mammals + outgroup calculated using phyloP (Benjamini–Hochberg adjusted). We considered four sets of outgroup combinations: (1) green lizard only; (2) chicken only; (3) two reptiles and one frog; and (4) two reptiles, one frog and one fish, and only kept those that were significantly conserved in all four sets of statistical tests (Benjamini–Hochberg adjusted $P < 0.01$). Only elements ≥ 20 bp were kept for further analysis.

To annotate MSHCEs to possible functional elements, we used the human annotation (Ensembl release 87) as a reference and classified the elements into the coding sequence, 5' and 3' untranslated regions, non-coding RNA, pseudogene, intron, upstream 10-kb region (from start codon), downstream 10-kb region (from the stop codon) and intergenic regions, with the same hierarchical order if the regions overlapped. Genes located within the upstream or downstream 10-kb range of MSHCEs were considered to be MSHCE-associated genes, and ordered by the length of the element. The top-300 MSHCE-associated genes were used in the GO enrichment analysis (χ^2 test, FDR-adjusted) and visualized using REVIGO⁷³.

Mammalian karyotype reconstruction

We used pairwise LASTZ alignments of the opossum, Tasmanian devil, platypus, chicken and common wall lizard (*Podarcis muralis*) genomes to the human genome as input. Echidna was not used here as most of the sequences were not anchored to chromosomes, which would lead to a more fragmented reconstruction. With the net and chain results, conserved segments that were uniquely and universally presented in all six species were obtained using inferCARs⁷⁴ (release 2006-Jun-16). Marsupial and therian ancestral karyotypes were inferred using ANGES⁷⁵ (v.1.01) using the branch-and-bound algorithm, and the resulting continuous ancestral regions (CARs) were further reorganized based on the previously predicted configuration⁷⁶ (Supplementary Tables 22, 23). We replaced the conserved segments of the human, opossum and Tasmanian devil genomes with those of the reconstructed therian ancestral karyotype and reconstructed marsupial ancestral karyotype using ANGES with the same parameters except setting the target reconstruction node to mammalian ancestor. We reorganized CARs on the basis of gene synteny among ingroups and outgroups inferred using MCScanX⁷⁷ (release 08-05-2012), requiring that there is synteny across CARs in at least one ingroup–outgroup pair (Supplementary Tables 22, 23). Pairwise MCScanX was run among the six species with BLASTP ($e < 1 \times 10^{-7}$).

Rearrangement events in each lineage were inferred using GRIMM⁷⁸ (v.2.1) by taking the karyotypes of the most recent ancestor and the child as input. The breakpoint number in each lineage was calculated on the basis of the output of GRIMM using an in-house-generated script, in which one breakpoint was counted in fission, two breakpoints were counted in translocation, and one or two breakpoints were counted in inversion, depending on whether the inversion happened at the end of the chromosome. Calculations were done using resolutions of 500 kb and 300 kb, and using the raw ANGES output and reorganized output, respectively (Supplementary Table 28). Differences in breakpoint rates compared to the average of all branches were tested as previously described⁷⁹.

Gametologue identification

We used BLASTP to compare all Y-borne genes to all X-borne genes ($e < 1 \times 10^{-5}$) and kept the best hit for each Y-borne gene. Candidate gametologue pairs were further confirmed if both of the genes were mapped to the same gene in NCBI or the SwissProt database. Four gametologues (platypus *AMHX* and *FEMICX* from OANA5, and *SDHAY* and *HNRNPKY* from ref. 14) were added as they were missing in mOrnAna1. Translated genes were aligned using PRANK⁸⁰, filtered using Gblock⁸¹, and converted back into the alignment of the coding sequence. dS was calculated using codeml in PAML with ‘runmode=-2’.

Demarcate evolutionary strata

We aligned all platypus Y-borne scaffolds (N-masked) to all platypus X-borne sequences (N-masked), and aligned all echidna Y-borne scaffolds (N-masked) to all echidna X-borne sequences (N-masked), using LASTZ with the parameter set ‘--step=19 --hspthresh=2200 --inner=2000 --ydrop=3400 --gappedthresh=10000 --format=axt’ and a score matrix set for the comparison of distantly related species. On the basis of the net and ‘maf’ results, the identity of each alignment block was calculated in a 2-kb non-overlapped window and the aligned Y-borne sequences were oriented along the X chromosomes. Identity along X chromosomes was colour-coded for visualization.

Expression calculation

RNA-sequencing reads of platypus (SRP102989) and echidna were mapped to the genome using HISTA2. Uniquely mapped reads were used in the calculation and normalization of the reads per kilobase per million reads (RPKM) using DESeq⁸² (v.1.28.0) to generate an expression matrix for each species. For tissues that were available in both sexes, we computed the median RPKM of each X-borne gene, and computed its F/M RPKM ratio (requiring RPKM in both sexes to be ≥ 1) to determine dosage-compensation status. We used the median expression value in each tissue to calculate the tissue specificity index TAU⁸³ for each gene. We defined tissue-specific expression as a gene that shows at least twofold higher expression in tissue with the highest expression than in any other tissue, the highest RPKM > 1 and TAU > 0.8 .

Building genome-wide Hi-C interaction maps

Genome-wide interaction maps at a 100-kb resolution were generated for platypus, echidna and human (SRX641267) with HiC-Pro⁸⁴ (v.2.10.0). For echidna, we only retained scaffolds > 10 kb as the large number of short scaffolds would cause ICE normalization failure. The normalized sex chromosomes submatrix was extracted for quantification and plotting with ggplot2 (v.3.2.1). For human, we used the scaled homologous sequences of platypus for quantification and plotting.

Identification of TADs and CTCF-binding sites

HiC-Pro interaction maps were transformed to h5 format using hicConvertFormat and fed to hicFindTADs with the parameters ‘--outPrefix TAD --numberOfProcessors 32 --correctForMultipleTesting fdr’ to identify TADs with HiCExplorer⁸⁵ (v.3.0). The human CTCF motif⁸⁶ was used as a bait by fimo in MEME⁸⁷ (v.4.12.0) to identify putative CTCF-binding sites. CTCF densities in every 100 kb non-overlapping sliding

window along the platypus sex chromosomes or scaled homologous sequences of echidna, human and chicken were compared.

FISH

BACs were obtained from the Children's Hospital Oakland Research Institute from the platypus BAC library CH236: CH236-775N6 13q2; CH236-97I3 15p1 and CUGI BAC/EST resource centre from the platypus BAC library Oa_Ab: Oa_Bb-155A12 autosomal (*WSB1*); Oa_Bb-145P09 Y2; Oa_Bb-397I21 Y3. The Super_Scaffold_40-specific probe was amplified from platypus genomic DNA. Gene ENSOANT00000009075.3 was amplified using primers GTCTAAAGACAAGTGTACATCTGTGAC and GTGACTTCTCTTGCAGAACACAC. The 3.9-kb product was cloned into pGEM-T Easy (Promega). BAC probes were directly labelled with dUTP Alexa Fluor 594-dUTP, aminoallyl-dUTP-XX-ATTO-488 (Jena Bioscience) using the Nick Translation Kit (Roche Diagnostics) and the Super_Scaffold_40-specific probe labelled with biotin using the Biotin-Nick Translation Mix (1175824919, Roche Diagnostics). The FISH protocol was carried out on cultured fibroblasts from platypus (authenticated by karyotype, not mycoplasma tested) obtained from animals captured at the Upper Barnard River (New South Wales, Australia) during the breeding season (AEC permits S-49-2006, S-032-2008 and S-2011-146) as previously described⁸⁸ with the following exceptions. Slides were denatured at 70 °C for 3 min in 70% formamide in 2× SSC, 1 mg DNA probe was used per slide, pre-annealing of repetitive DNA sequences was done at 37 °C for 30–60 min. Detection of biotin-labelled probes was done using Rhodamine Avidin D (Vector Laboratories, A-2002), goat Biotinylated anti-avidin D (Vector Laboratories, BA-0300) and Rhodamine Avidin D. Slides were blocked in 4 × SSC, 1% BSA fraction V, for 30 min at 37 °C. Rhodamine Avidin D and Biotinylated anti-avidin D and the second Rhodamine Avidin D were diluted in 4 × SSC, 1% BSA fraction V and were incubated on slides for 45 min at 37 °C, after each step washes were done in 4 × SSC, 4 × SSC, 0.1% triton, 4 × SSC at room temperature for 10 min each. Slides were mounted in VECTASHIELD with DAPI (Vector Laboratories, H-1200). Sample size was determined according to ref. ⁸⁹, but was limited by material availability. Images were captured on a Nikon Ti Microscope using NIS-Elements AR 4.20.00 software and processed with ImageJ (v.2.0.0). Fisher's exact test was performed with matrix containing mean of associated and non-associated cells from the three replicates. No blinding nor randomization was performed.

Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

Data availability

The platypus whole-genome shotgun project has been deposited at GenBank (project accessions [PRJNA489114](#) and [PRJNA489115](#)), CNSA (<https://db.cngb.org/cnsa/>) of CNGBdb (accession CNP0000130) and GenomeArk (https://vgp.github.io/genomeark-curated-assembly/Ornithorhynchus_anatinus/). The echidna whole-genome shotgun project has been deposited at GenBank (project accession [PRJNA576333](#)), CNSA of CNGBdb (accession CNP0000697) and GenomeArk at (https://vgp.github.io/genomeark/Tachyglossus_aculeatus/). Echidna RNA-sequencing data have been deposited at GenBank (project accession [PRJNA591380](#)) and CNSA of CNGBdb (accession CNP0000779). Public database used in this study include: NCBI (<https://www.ncbi.nlm.nih.gov/>), Ensembl (release 87) (<http://dec2016.archive.ensembl.org/index.html>), Uniprot (<https://www.uniprot.org/>) and Repbase (<https://www.girinst.org/repbase/>). Accession codes of genes are available in Supplementary Tables [31](#), [33](#), [37](#), [49](#), [51](#).

Code availability

In-house-generated scripts used in this study are shared on GitHub (<https://github.com/ZhangLabSZ/MonotremeGenome>).

References

1. 1.

Ashwell, K. *Neurobiology of Monotremes: Brain Evolution in Our Distant Mammalian Cousins* (CSIRO PUBLISHING, 2013).

2. 2.

Warren, W. C. et al. Genome analysis of the platypus reveals unique signatures of evolution. *Nature* **453**, 175–183 (2008).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

3. 3.

Grützner, F. et al. In the platypus a meiotic chain of ten sex chromosomes shares genes with the bird Z and mammal X chromosomes. *Nature* **432**, 913–917 (2004).

[ADS](#) [PubMed](#) [Google Scholar](#)

4. 4.

Kortschak, R. D., Tsend-Ayush, E. & Grützner, F. Analysis of SINE and LINE repeat content of Y chromosomes in the platypus, *Ornithorhynchus anatinus*. *Reprod. Fertil. Dev.* **21**, 964–975 (2009).

[CAS](#) [PubMed](#) [Google Scholar](#)

5. 5.

Boissinot, S. & Sookdeo, A. The evolution of LINE-1 in vertebrates. *Genome Biol. Evol.* **8**, 3485–3507 (2016).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

6. 6.

Phillips, M. J., Bennett, T. H. & Lee, M. S. Molecules, morphology, and ecology indicate a recent, amphibious ancestry for echidnas. *Proc. Natl Acad. Sci. USA* **106**, 17089–17094 (2009).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

7. 7.

Bellott, D. W. et al. Mammalian Y chromosomes retain widely expressed dosage-sensitive regulators. *Nature* **508**, 494–499 (2014).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

8. 8.

Whittington, C. M. et al. Defensins and the convergent evolution of platypus and reptile venom genes. *Genome Res.* **18**, 986–994 (2008).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

9. 9.

Julien, P. et al. Mechanisms and evolutionary patterns of mammalian and avian dosage compensation. *PLoS Biol.* **10**, e1001328 (2012).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

10. 10.

Rousselle, M., Laverré, A., Figuet, E., Nabholz, B. & Galtier, N. Influence of recombination and GC-biased gene conversion on the adaptive and nonadaptive substitution rate in mammals versus birds. *Mol. Biol. Evol.* **36**, 458–471 (2019).

[CAS](#) [PubMed](#) [Google Scholar](#)

11. 11.

Hinch, A. G., Altemose, N., Noor, N., Donnelly, P. & Myers, S. R. Recombination in the human pseudoautosomal region PAR1. *PLoS Genet.* **10**, e1004503 (2014).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

12. 12.

Burt, D. W. Origin and evolution of avian microchromosomes. *Cytogenet. Genome Res.* **96**, 97–112 (2002).

[CAS](#) [PubMed](#) [Google Scholar](#)

13. 13.

Dohm, J. C., Tsend-Ayush, E., Reinhardt, R., Grützner, F. & Himmelbauer, H. Disruption and pseudoautosomal localization of the major histocompatibility complex in monotremes. *Genome Biol.* **8**, R175 (2007).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

14. 14.

Cortez, D. et al. Origins and functional evolution of Y chromosomes across mammals. *Nature* **508**, 488–493 (2014).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

15. 15.

Zhou, Q. et al. Complex evolutionary trajectories of sex chromosomes across bird taxa. *Science* **346**, 1246338 (2014).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

16. 16.

Veyrunes, F. et al. Bird-like sex chromosomes of platypus imply recent origin of mammal sex chromosomes. *Genome Res.* **18**, 965–973 (2008).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

17. 17.

Braasch, I. et al. The spotted gar genome illuminates vertebrate evolution and facilitates human–teleost comparisons. *Nat. Genet.* **48**, 427–437 (2016).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

18. 18.

Gruetzner, F., Ashley, T., Rowell, D. M. & Marshall Graves, J. A. How did the platypus get its sex chromosome chain? A comparison of meiotic multiples and sex chromosomes in plants and animals. *Chromosoma* **115**, 75–88 (2006).

[PubMed](#) [Google Scholar](#)

19. 19.

Golczyk, H., Massouh, A. & Greiner, S. Translocations of chromosome end-segments and facultative heterochromatin promote meiotic ring formation in evening primroses. *Plant Cell* **26**, 1280–1293 (2014).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

20. 20.

de Waal Malefijt, M. & Charlesworth, B. A model for the evolution of translocation heterozygosity. *Heredity* **43**, 315–331 (1979).

[Google Scholar](#)

21. 21.

Casey, A. E., Daish, T. J., Barbero, J. L. & Grützner, F. Differential cohesin loading marks paired and unpaired regions of platypus sex chromosomes at prophase I. *Sci. Rep.* **7**, 4217 (2017).

[ADS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

22. 22.

Dixon, J. R. et al. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **485**, 376–380 (2012).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

23. 23.

Griffiths, M. *The Biology of Monotremes* (Academic, 1978).

24. 24.

Meredith, R. W., Zhang, G., Gilbert, M. T., Jarvis, E. D. & Springer, M. S. Evidence for a single loss of mineralized teeth in the common avian ancestor. *Science* **346**, 1254390 (2014).

[PubMed](#) [Google Scholar](#)

25. 25.

Springer, M. S. et al. Odontogenic ameloblast-associated (ODAM) is inactivated in toothless/enamelless placental mammals and toothed whales. *BMC Evol. Biol.* **19**, 31 (2019).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

26. 26.

Ordoñez, G. R. et al. Loss of genes implicated in gastric function during platypus evolution. *Genome Biol.* **9**, R81 (2008).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

27. 27.

Hayakawa, T., Suzuki-Hashido, N., Matsui, A. & Go, Y. Frequent expansions of the bitter taste receptor gene repertoire during evolution

of mammals in the Euarchontoglires clade. *Mol. Biol. Evol.* **31**, 2018–2031 (2014).

[CAS](#) [PubMed](#) [Google Scholar](#)

28. 28.

Johnson, R. N. et al. Adaptation and conservation insights from the koala genome. *Nat. Genet.* **50**, 1102–1111 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

29. 29.

Liu, Z. et al. Dietary specialization drives multiple independent losses and gains in the bitter taste gene repertoire of Laurasiatherian mammals. *Front. Zool.* **13**, 28 (2016).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

30. 30.

Hunnicutt, K. E. et al. Comparative genomic analysis of the pheromone receptor class 1 family (V1R) reveals extreme complexity in mouse lemurs (genus, *Microcebus*) and a chromosomal hotspot across mammals. *Genome Biol. Evol.* **12**, 3562–3579 (2020).

[CAS](#) [PubMed](#) [Google Scholar](#)

31. 31.

Johansen, K., Lenfant, C. & Grigg, G. C. Respiratory properties of blood and responses to diving of platypus *Ornithorhynchus anatinus* (Shaw). *Comp. Biochem. Physiol.* **18**, 597–608 (1966).

[CAS](#) [PubMed](#) [Google Scholar](#)

32. 32.

Alayash, A. I. Haptoglobin: old protein with new functions. *Clin. Chim. Acta* **412**, 493–498 (2011).

[CAS](#) [PubMed](#) [Google Scholar](#)

33. 33.

Wicher, K. B. & Fries, E. Haptoglobin, a hemoglobin-binding plasma protein, is present in bony fish and mammals but not in frog and chicken. *Proc. Natl Acad. Sci. USA* **103**, 4168–4173 (2006).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

34. 34.

Redmond, A. K. et al. Haptoglobin is a divergent masp family member that neofunctionalized to recycle hemoglobin via CD163 in mammals. *J. Immunol.* **201**, 2483–2491 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

35. 35.

Huttenlocker, A. K. & Farmer, C. G. Bone microvasculature tracks red blood cell size diminution in Triassic mammal and dinosaur forerunners. *Curr. Biol.* **27**, 48–54 (2017).

[CAS](#) [PubMed](#) [Google Scholar](#)

36. 36.

Schaer, D. J. et al. CD163 is the macrophage scavenger receptor for native and chemically modified hemoglobins in the absence of haptoglobin. *Blood* **107**, 373–380 (2006).

[CAS](#) [PubMed](#) [Google Scholar](#)

37. 37.

Griffiths, M. *Echidnas* (Pergamon, 1968).

38. 38.

Brawand, D., Wahli, W. & Kaessmann, H. Loss of egg yolk genes in mammals and the origin of lactation and placentation. *PLoS Biol.* **6**, e63 (2008).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

39. 39.

Pharo, E. A. et al. The mammary gland-specific marsupial *ELP* and eutherian *CTI* share a common ancestral gene. *BMC Evol. Biol.* **12**, 80 (2012).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

40. 40.

Lefèvre, C. M., Sharp, J. A. & Nicholas, K. R. Characterisation of monotreme caseins reveals lineage-specific expansion of an ancestral casein locus in mammals. *Reprod. Fertil. Dev.* **21**, 1015–1027 (2009).

[PubMed](#) [Google Scholar](#)

41. 41.

Holt, C., Carver, J. A., Ecroyd, H. & Thorn, D. C. Invited review: Caseins and the casein micelle: their biological functions, structures, and behavior in foods. *J. Dairy Sci.* **96**, 6127–6146 (2013).

[CAS](#) [PubMed](#) [Google Scholar](#)

42. 42.

Kawasaki, K., Lafont, A. G. & Sire, J. Y. The evolution of milk casein genes from tooth genes before the origin of mammals. *Mol. Biol. Evol.* **28**, 2053–2061 (2011).

[CAS](#) [PubMed](#) [Google Scholar](#)

43. 43.

Cardoso-Moreira, M. et al. Gene expression across mammalian organ development. *Nature* **571**, 505–509 (2019).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

44. 44.

Kajitani, R. et al. Efficient de novo assembly of highly heterozygous genomes from whole-genome shotgun short reads. *Genome Res.* **24**, 1384–1395 (2014).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

45. 45.

Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

46. 46.

Birney, E., Clamp, M. & Durbin, R. Genewise and genomewise. *Genome Res.* **14**, 988–995 (2004).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

47. 47.

Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

48. 48.

Li, H. Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* **34**, 3094–3100 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

49. 49.

Jain, C., Koren, S., Dilthey, A., Phillippy, A. M. & Aluru, S. A fast adaptive algorithm for computing whole-genome homology maps. *Bioinformatics* **34**, i748–i756 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

50. 50.

Rens, W. et al. The multiple sex chromosomes of platypus and echidna are not completely identical and several share homology with the avian Z. *Genome Biol.* **8**, R243 (2007).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

51. 51.

Bao, W., Kojima, K. K. & Kohany, O. Repbase update, a database of repetitive elements in eukaryotic genomes. *Mob. DNA* **6**, 11 (2015).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

52. 52.

Chen, N. Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr. Protoc. Bioinformatics* **5**, 4.10.1–4.10.14 (2004).

[Google Scholar](#)

53. 53.

Benson, G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* **27**, 573–580 (1999).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

54. 54.

Yates, A. et al. Ensembl 2016. *Nucleic Acids Res.* **44**, D710–D716 (2016).

[CAS](#) [PubMed](#) [Google Scholar](#)

55. 55.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410 (1990).

[CAS](#) [PubMed](#) [Google Scholar](#)

56. 56.

Stanke, M., Schöffmann, O., Morgenstern, B. & Waack, S. Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. *BMC Bioinformatics* **7**, 62 (2006).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

57. 57.

Brawand, D. et al. The evolution of gene expression levels in mammalian organs. *Nature* **478**, 343–348 (2011).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

58. 58.

Kim, D., Langmead, B. & Salzberg, S. L. HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* **12**, 357–360 (2015).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

59. 59.

Pertea, M. et al. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **33**, 290–295 (2015).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

60. 60.

UniProt Consortium. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, D506–D515 (2019).

[Google Scholar](#)

61. 61.

Jones, P. et al. InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240 (2014).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

62. 62.

Bickhart, D. M. et al. Single-molecule sequencing and chromatin conformation capture enable de novo reference assembly of the domestic goat genome. *Nat. Genet.* **49**, 643–650 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

63. 63.

Harris, R. S. *Improved Pairwise Alignment of Genomic DNA*. PhD thesis, Pennsylvania State Univ. (2007).

64. 64.

Zhang, G. et al. Comparative genomics reveals insights into avian genome evolution and adaptation. *Science* **346**, 1311–1320 (2014).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

65. 65.

Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

66. 66.

Yang, Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

67. 67.

Benton, M. J. et al. Constraints on the timescale of animal evolutionary history. *Palaeontol. Electronica* **18**, 1–106 (2015).

[Google Scholar](#)

68. 68.

Blanchette, M. et al. Aligning multiple genomic sequences with the threaded blockset aligner. *Genome Res.* **14**, 708–715 (2004).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

69. 69.

Hubisz, M. J., Pollard, K. S. & Siepel, A. PHAST and RPHAST: phylogenetic analysis with space/time models. *Brief. Bioinform.* **12**, 41–51 (2011).

[CAS](#) [PubMed](#) [Google Scholar](#)

70. 70.

Li, L., Stoeckert, C. J., Jr & Roos, D. S. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* **13**, 2178–2189 (2003).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

71. 71.

Han, M. V., Thomas, G. W., Lugo-Martinez, J. & Hahn, M. W. Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3. *Mol. Biol. Evol.* **30**, 1987–1997 (2013).

[CAS](#) [PubMed](#) [Google Scholar](#)

72. 72.

Seki, R. et al. Functional roles of Aves class-specific *cis*-regulatory elements on macroevolution of bird-specific features. *Nat. Commun.* **8**, 14229 (2017).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

73. 73.

Supek, F., Bošnjak, M., Škunca, N. & Šmuc, T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS ONE* **6**, e21800 (2011).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

74. 74.

Ma, J. et al. Reconstructing contiguous regions of an ancestral genome. *Genome Res.* **16**, 1557–1565 (2006).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

75. 75.

Jones, B. R., Rajaraman, A., Tannier, E. & Chauve, C. ANGES: reconstructing ANcestral GEnomeS maps. *Bioinformatics* **28**, 2388–2390 (2012).

[CAS](#) [PubMed](#) [Google Scholar](#)

76. 76.

Deakin, J. E. et al. Reconstruction of the ancestral marsupial karyotype from comparative gene maps. *BMC Evol. Biol.* **13**, 258 (2013).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

77. 77.

Wang, Y. et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* **40**, e49 (2012).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

78. 78.

Tesler, G. GRIMM: genome rearrangements web server. *Bioinformatics* **18**, 492–493 (2002).

[CAS](#) [PubMed](#) [Google Scholar](#)

79. 79.

Kim, J. et al. Reconstruction and evolutionary history of eutherian chromosomes. *Proc. Natl Acad. Sci. USA* **114**, E5379–E5388 (2017).

[CAS](#) [PubMed](#) [Google Scholar](#)

80. 80.

Löytönoja, A. Phylogeny-aware alignment with PRANK. *Methods Mol. Biol.* **1079**, 155–170 (2014).

[PubMed](#) [Google Scholar](#)

81. 81.

Talavera, G. & Castresana, J. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* **56**, 564–577 (2007).

[CAS](#) [PubMed](#) [Google Scholar](#)

82. 82.

Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

83. 83.

Yanai, I. et al. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics* **21**, 650–659 (2005).

[CAS](#) [Google Scholar](#)

84. 84.

Servant, N. et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol.* **16**, 259 (2015).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

85. 85.

Ramírez, F. et al. High-resolution TADs reveal DNA sequences underlying genome organization in flies. *Nat. Commun.* **9**, 189 (2018).

[ADS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

86. 86.

Jolma, A. et al. DNA-binding specificities of human transcription factors. *Cell* **152**, 327–339 (2013).

[CAS](#) [PubMed](#) [Google Scholar](#)

87. 87.

Bailey, T. L. & Elkan, C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **2**, 28–36 (1994).

[CAS](#) [PubMed](#) [Google Scholar](#)

88. 88.

Tsend-Ayush, E. et al. Higher-order genome organization in platypus and chicken sperm and repositioning of sex chromosomes during mammalian evolution. *Chromosoma* **118**, 53–69 (2009).

[PubMed](#) [Google Scholar](#)

89. 89.

Ling, J. Q. et al. CTCF mediates interchromosomal colocalization between *Igf2/H19* and *Wsb1/Nfl*. *Science* **312**, 269–272 (2006).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

90. 90.

Parra, Z. E. et al. Comparative genomic analysis and evolution of the T cell receptor loci in the opossum *Monodelphis domestica*. *BMC*

Genomics **9**, 111 (2008).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

91. 91.

Van Laere, A. S., Coppieters, W. & Georges, M. Characterization of the bovine pseudoautosomal boundary: documenting the evolutionary history of mammalian sex chromosomes. *Genome Res.* **18**, 1884–1895 (2008).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

[Download references](#)

Acknowledgements

We thank members of BGI-Shenzhen, China National GeneBank and VGP, and P. Baybayan, R. Hall and J. Howard for help carrying out the sequencing of the platypus and echidna genomes, M. Asahara for discussion, and D. Charlesworth for comments. Work was supported by the Strategic Priority Research Program of the Chinese Academy of Sciences (XDB31020000), the National Key R&D Program of China (MOST) grant 2018YFC1406901, International Partnership Program of Chinese Academy of Sciences (152453KYSB20170002), Carlsberg foundation (CF16-0663) and Villum Foundation (25900) to G.Z. Q.Z. is supported by the National Natural Science Foundation of China (31722050, 31671319 and 32061130208), Natural Science Foundation of Zhejiang Province (LD19C190001), European Research Council Starting Grant (grant agreement 677696) and start-up funds from Zhejiang University. F.G., L.S.-W. and T.D. are supported by Australian Research Council (FT160100267, DP170104907 and DP110105396). M.B.R., J.C.F. and S.D.J. are supported by the Australian Research Council (LP160101728). We acknowledge the Kyoto University Research Administration Office for support and Human Genome Center, the Institute of Medical Science, the University of Tokyo for the super-computing resource for supporting T.H.'s research facilities. T.H. was financed by JSPS KAKENHI grant numbers 16K18630 and

19K16241 and the Sasakawa Scientific Research Grant from the Japan Science Society. The echidna RNA-sequencing analysis was supported by H.K.'s grant from the European Research Council (615253, OntoTransEvol). This work was supported by Guangdong Provincial Academician Workstation of BGI Synthetic Genomics No.

2017B090904014 (H.Y.), Robert and Rosabel Osborne Endowment, Howard Hughes Medical Institute (E.D.J.), Rockefeller University start-up funds (E.D.J.), Intramural Research Program of the National Human Genome Research Institute, National Institutes of Health (A.R. and A.M.P.), Korea Health Technology R&D Project through the Korea Health Industry Development Institute HI17C2098 (A.R.). This work used the computational resources of BGI-Shenzhen and the NIH HPC Biowulf cluster (<https://hpc.nih.gov>). Animal icons are from <https://www.flaticon.com/> (made by Freepik) and <http://phylopic.org/>.

Author information

Author notes

1. These authors contributed equally: Yang Zhou, Linda Shearwin-Whyatt, Jing Li

Affiliations

1. BGI-Shenzhen, Shenzhen, China

Yang Zhou, Zhenzhen Song, Qiye Li, Huanming Yang, Jian Wang & Guojie Zhang

2. Villum Center for Biodiversity Genomics, Section for Ecology and Evolution, Department of Biology, University of Copenhagen, Copenhagen, Denmark

Yang Zhou & Guojie Zhang

3. School of Biological Sciences, The Environment Institute, The University of Adelaide, Adelaide, South Australia, Australia

Linda Shearwin-Whyatt, David Stevens, Filip Pajpach, Natasha Bradley, Tasman Daish, Tahlia Perry & Frank Grutzner

4. MOE Laboratory of Biosystems Homeostasis and Protection and Zhejiang Provincial Key Laboratory for Cancer Molecular Cell Biology, Life Sciences Institute, Zhejiang University, Hangzhou, China

Jing Li, Zexian Zhu & Qi Zhou

5. BGI Education Center, University of Chinese Academy of Sciences, Shenzhen, China

Zhenzhen Song

6. Faculty of Environmental Earth Science, Hokkaido University, Sapporo, Japan

Takashi Hayakawa

7. Japan Monkey Centre, Inuyama, Japan

Takashi Hayakawa

8. School of BioSciences, The University of Melbourne, Melbourne, Victoria, Australia

Jane C. Fenelon & Marilyn B. Renfree

9. School of Life and Environmental Sciences, The University of Sydney, Sydney, New South Wales, Australia

Emma Peel, Yuanyuan Cheng & Katherine Belov

10. digzyme Inc, Tokyo, Japan

Hikoyu Suzuki

11. School of Life Science and Technology, Tokyo Institute of Technology, Tokyo, Japan

Masato Nikaido

12. The Genome Center, University of California, Davis, CA, USA

Joana Damas & Harris A. Lewin

13. Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD, USA

Yuncong Geng

14. Genome Informatics Section, Computational and Statistical Genomics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA

Arang Rie & Adam M. Phillippy

15. Tree of Life Programme, Wellcome Sanger Institute, Cambridge, UK

Ying Sims, Jonathan Wood & Kerstin Howe

16. The Vertebrate Genome Lab, The Rockefeller University, New York, NY, USA

Bettina Haase, Jacquelyn Mountcastle & Olivier Fedrigo

17. James D. Watson Institute of Genome Sciences, Hangzhou, China

Huanming Yang & Jian Wang

18. University of the Chinese Academy of Sciences, Beijing, China

Huanming Yang

19. Guangdong Provincial Academician Workstation of BGI Synthetic Genomics, BGI-Shenzhen, Shenzhen, China

Huanming Yang

20. School of Agriculture and Food Sciences, The University of Queensland, Gatton, Queensland, Australia

Stephen D. Johnston

21. Laboratory of Neurogenetics of Language, The Rockefeller University, New York, NY, USA

Erich D. Jarvis

22. Howard Hughes Medical Institute, Chevy Chase, MD, USA

Erich D. Jarvis

23. San Diego Zoo Global, Escondido, CA, USA

Oliver A. Ryder

24. Center for Molecular Biology of Heidelberg University (ZMBH), DKFZ-ZMBH Alliance, Heidelberg, Germany

Henrik Kaessmann

25. Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK

Peter Donnelly

26. Pacific Biosciences, Menlo Park, CA, USA

Jonas Korlach

27. Department of Evolution and Ecology, College of Biological Sciences, University of California, Davis, CA, USA

Harris A. Lewin

28. Department of Reproduction and Population Health, School of Veterinary Medicine, University of California, Davis, CA, USA
Harris A. Lewin
29. Research School of Biology, Australian National University, Canberra, Australian Capital Territory, Australia
Jennifer Graves
30. Institute for Applied Ecology, University of Canberra, Canberra, Australian Capital Territory, Australia
Jennifer Graves
31. School of Life Sciences, La Trobe University, Melbourne, Victoria, Australia
Jennifer Graves
32. Department of Neuroscience and Developmental Biology, University of Vienna, Vienna, Austria
Qi Zhou
33. Center for Reproductive Medicine, The 2nd Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, China
Qi Zhou
34. State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China
Guojie Zhang
35. Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, China
Guojie Zhang

Authors

1. Yang Zhou

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Linda Shearwin-Whyatt

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Jing Li

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Zhenzhen Song

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Takashi Hayakawa

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. David Stevens

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Jane C. Fenelon

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Emma Peel

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. Yuanyuan Cheng

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. Filip Pajpach

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

11. Natasha Bradley

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

12. Hikoyu Suzuki

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

13. Masato Nikaido

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

14. Joana Damas

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

15. Tasman Daish

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

16. Tahlia Perry

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

17. Zexian Zhu

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

18. Yuncong Geng

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

19. Arang Rhie

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

20. Ying Sims

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

21. Jonathan Wood

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

22. Bettina Haase

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

23. Jacquelyn Mountcastle

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

24. Olivier Fedrigo

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

25. Qiye Li

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

26. Huanming Yang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

27. Jian Wang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

28. Stephen D. Johnston

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

29. Adam M. Phillippy

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

30. Kerstin Howe

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

31. Erich D. Jarvis

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

32. Oliver A. Ryder

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

33. Henrik Kaessmann

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

34. Peter Donnelly

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

35. Jonas Korlach

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

36. Harris A. Lewin

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

37. Jennifer Graves

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

38. Katherine Belov

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

39. Marilyn B. Renfree

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

40. Frank Grutzner

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

41. Qi Zhou

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

42. Guojie Zhang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Contributions

G.Z., Q.Z. and F.G. initiated and designed the monotreme genome project with early input from J.G. G.Z., F.G., Q.L., T.D., B.H., J.M., O.F., E.D.J., O.A.R., H.K., P.D., J.K., T.P., H.Y. and J. Wang coordinated the project and were involved in the collection, extraction and sequencing of samples. A.M.P., A.R., G.Z., Y.Z., K.H., Y.S., J. Wood, Q.L., Q.Z., F.G. and L.S.-W. performed genome assembling, evaluation and chromosome assignment. G.Z., Y.Z., Z.S., Y.G. and F.G. performed the annotation and mammalian macro-evolutionary analysis including divergence time estimation, gene family and MSHCE analysis. G.Z., Y.Z., H.A.L., J.D. and Q.Z. performed ancestral karyotype analysis. Q.Z., J.L., Z.Z., F.G., F.P., G.Z., Y.Z., L.S.-W. and Y.G. performed analysis to sex chromosome and FISH validation. K.B., E.P., Y.C., Y.Z., D.S., Z.S., G.Z. and F.G. performed the immune gene analysis. Y.Z., J.C.F., M.B.R., Z.S. and G.Z. performed the tooth-related gene analysis. N.B., F.G. and Y.Z. performed the digestive gene analysis and PCR validation. T.H., H.S., M.N. and F.G. performed the chemosensory gene analysis. L.S.-W., F.G., Y.Z. and Z.S. performed the haemoglobin-degradation gene analysis. M.B.R., J.C.F. and S.J. performed the reproductive gene analysis. G.Z., Q.Z., F.G., Y.Z., L.S.-W., J.L., M.B.R., J.C.F., K.B., E.P., Y.C., D.S., N.B., F.P., T.H., H.A.L., J.D., A.M.P., A.R., K.H., J. Wood, O.F. and J.G. wrote the manuscript.

Corresponding authors

Correspondence to [Frank Grutzner](#) or [Qi Zhou](#) or [Guojie Zhang](#).

Ethics declarations

Competing interests

J.K. is an employee of Pacific Biosciences, a company that develops single-molecule sequencing technologies.

Additional information

Peer review information *Nature* thanks Janine Deakin, Rebecca Johnson and Hugues Roest Crollius 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 Platypus genome assembly and evaluation.](#)

a, b, Hi-C two-dimensional juicebox maps of mOrnAnal before (**a**) and after (**b**) manual assembly curation. The grey lines depict scaffold boundaries. The off-diagonal matches between scaffolds indicate potential missed joins, whereas ‘empty’ areas within scaffold boundaries indicate misjoins. The gEVAL-supported manual assembly curation led to a notably improved arrangement with >96% of the assembly sequence inside chromosome-scale scaffolds. **c, d**, The Super_Scaffold_40 was misassigned to chromosome 15 in OANA5 but FISH on metaphase spreads from platypus fibroblasts map it to chromosome 13. **c**, Co-hybridization of the BAC of chromosome 15 (green, top arrow) and Super_Scaffold_40 probe (red, bottom arrow) showing an absence of co-localization (14 nuclei scored, 2 independent experiments). Inset, interphase example. **d**, Co-

hybridization of the BAC of chromosome 13 (green) and Super_Scaffold_40 probe (red) showing co-localization (arrows, 40 nuclei scored, 5 independent experiments). Scale bars, 10 μ m. **e**, An example of scaffold chromosome misassignment in OANA5. Female-to-male (F/M) depth ratio, normalized female depth and normalized male depth along OANA5 chromosome 14 in 5-kb non-overlapping windows. Depth ratio, normalized female depth and normalized male depth all suggest that OANA5 chromosome 14 should be an X-borne rather than autosomal sequence. **f, g**, Normalized depth distribution of redundant sequences and one-to-one sequences in male (**f**) and female (**g**). Redundant sequences (red) in OANA5 are probably assembly artefacts due to heterozygotes of the sequenced individual of OANA5, and are therefore featured with 0.5 \times normalized depth in OANA5 but 1 \times normalized depth in mOrnAna1 in both male and female. One-to-one sequences in OANA5 (black) have 1 \times normalized depth in both OANA5 and mOrnAna1 in reads that are mapped from both sexes as expected. Each dot represents one mapping region between OANA5 and mOrnAna1 by Mashmap, and the normalized depth values of each dot are calculated as the mean depth across the mapping region in OANA5 and mOrnAna1. The small peak in one-to-one sequence density plot in the male indicates candidate X-linked sequences. **h**, Example redundant sequences Contig40802, Contig44497 and Contig35847 in OANA5 that could be interpreted as false duplications. Dot plot is generated between the target region of mOrnAna1 chromosome 1 and OANA5 Contig1255, Contig40802, Contig44497 and Contig35847 by FlexiDot. Candidate redundant sequences are those mapped to the same region in mOrnAna1 chromosome 1, highlighted by dashed lines in the dot plot and grey in the normalized depth plot. Normalized male and female read depths along each sequence are calculated in 500-bp windows, and plotted along each sequence. Although the normalized depth is always around 1 in the region of mOrnAna1 chromosome 1, normalized depth drops half in Contig40802, Contig44497, Contig35847 and the aligned regions in Contig1255, indicating that Contig40802, Contig44497 and Contig35847 are probably redundant sequences in OANA5. **i, j**, Examples of gene annotation artefacts in OANA5: *CIT* (**i**) and *PBRM1* (**j**) have been fragmented into multiple small artificial genes in OANA5 (purple) but have now been fully recovered in mOrnAna1 (orange). Orthologous human

genes (grey) are also shown to indicate that the mOrnAna1 rather than OANA5 annotation has a similar gene structure to that of the human genes.

Extended Data Fig. 2 Mammalian genome evolution.

a, Phylogenetic tree constructed using fourfold degenerate sites from 7,946 one-to-one orthologues among seven representative species (human, mouse, opossum, platypus, echidna, chicken and green lizard). The fossil time calibration of the nodes marked by circles were obtained from a previously published study⁶⁷. The numbers of gene families that have undergone significant (Viterbi $P < 0.05$) lineage-specific expansions (green) and contractions (red) are marked on each branch. Exact P values are available in Supplementary Table 29. No multiple-testing correction was applied. **b**, Examples of some imprinting gene clusters improved in mOrnAna1 compared to OANA5. The first line of each synteny plot represents mOrnAna1 and the second line represents OANA5. Names of genes that have been found to be imprinted in human and mouse are highlighted in black, and non-imprinting genes in red. Fragmented genes with alignment rate lower than 70% are marked by triangles. The double slash represents the intermediate region longer than 100 kb. **c**, Distribution of MSHCEs on genomic elements. **d**, Enriched GO terms in the top-300 MSHCE-associated genes. P values of enrichment are calculated using a χ^2 test, and FDRs are computed to adjust for multiple testing. GO terms are clustered based on semantic similarity. GO terms related to nervous system development are highlighted in bold. **e**, A case of one MSHCE in *BCL11A* that overlaps with the enhancer signals inferred from H3K27ac ChIP-seq experiments at 8.5 and 12 weeks after conception (p.c.w.). **f**, Evolution highway comparative chromosome browser visualization of reconstructed MACs at a 500-kb resolution. Blocks overlaid on each MAC represent human syntenic fragments. Numbers within blocks indicate the homologous human chromosome. **g**, Evolution highway comparative chromosome browser visualization of the human genome at a 500-kb resolution, with each block overlaid on each human chromosome representing putative chromosome fragments of the ancestral mammalian genome. Numbers within blocks depict the ancestral mammal chromosome numbers.

Silhouettes of the human and opossum are from <https://www.flaticon.com/>. The silhouette of the platypus is created by S. Werning and is reproduced

under a Creative Commons Attribution 3.0 Unported licence (<http://creativecommons.org/licenses/by/3.0/>).

Extended Data Fig. 3 Evolution of immune gene family in monotremes.

a, MHC genes in platypus and echidna are located on two different chromosomes, but the classical class I and II genes involved in antigen presentation are located within a single cluster in each genome. **b**, Phylogenetic relationship of class I genes in representative mammals and chicken. Classical class I genes (red) in monotremes exhibit high similarity, which is rarely observed in other species. Only bootstrap values with >50% support are shown. **c**, **d**, Phylogenetic relationship of MHC class II alpha (**c**) and beta (**d**) genes. Genes with prefix ‘HLA’, ‘Modo’, ‘Phci’, ‘Oran’, ‘Taac’ and ‘Gaga’ indicate genes in human, opossum, koala, platypus, echidna and chicken, respectively. Only bootstrap values with >50% support are shown. **e**, Phylogenetic relationship among putative functional V γ sequences from platypus (yellow), echidna (purple), koala (green), mouse (orange), human (red), sheep (grey), cow (dark red) and chicken (dark yellow). Groups according to a previous study⁹⁰ are displayed around the outside of the tree, with the putative marsupial–monotreme-specific group denoted by a ‘?’ . Only bootstrap values with greater than 50% support are shown. **f**, Synteny conservation of beta-defensin genes in monotremes and loss of functional venom defensins in echidna. Venom defensins (OavDLP genes) and venom-like defensin (DEFB-VL genes) are shown in red. Only putative functional defensins are shown. **g**, Putative OavDLP loss in echidna. OavDLP genes and *DEFB-VL* each contain two exons (indicated by a box and triangle) in platypus. Both exons of platypus *DEFB-VL* can be mapped to echidna chromosome X2. A single platypus OavDLP exon can be mapped to echidna chromosome X2 while the second exons cannot. Grey links indicate platypus–echidna LASTZ alignment. **h**, Phylogenetic relationship of *DEFB-VL* and OavDLP genes suggested that ancestral monotremes had all three OavDLP genes but that echidna has lost the two of them (*OavDLP-B* and *OavDLP-C*). Branch length is not shown. ta, echidna; oa, platypus. Silhouettes of the human, opossum, koala and frog are from <https://www.flaticon.com/>. The silhouette of the platypus is created by S. Werning and is reproduced under a Creative Commons

Attribution 3.0 Unported licence
(<http://creativecommons.org/licenses/by/3.0/>).

Extended Data Fig. 4 Genomic composition of monotreme sex chromosomes.

a, Composition of the echidna sex chromosomes. The circos plot (from outer to inner rings) shows: X chromosomes with PARs shown as light colours and SDRs as dark colours; assembled Y chromosome fragments showing the colour-scaled sequence similarity levels with homologous X chromosomes; normalized F/M ratios of Illumina DNA-sequencing depth in non-overlapping 5-kb windows; F/M expression ratios (each red dot is one gene) of adult kidney and smoothed expression ratio trend; and GC content in non-overlapping 2-kb windows. In addition, Y-linked fragments with a similar level of sequence divergence from the X chromosome indicate a pattern of evolutionary strata. As expected, F/M DNA depth ratio is centred at 1 at PARs, but is around 2 at SDRs. Some PARs show significantly higher GC content than the regions that suppressed recombination between X and Y. **b**, Partial dosage compensation in monotremes. The four point range plots show \log_2 -transformed values of the male-to-female (M/F) expression ratio in the brain, kidney, heart and liver of platypus and echidna. As expected, \log_2 -transformed values of the M/F expression ratio is close to 0 for genes on autosomes (A) and PARs, whereas for genes on SDRs, the expression is female-biased in all tissues, which suggests that monotremes have partial dosage compensation. Whiskers indicate the 25–75th percentiles and circles are the median value. **c**, Some PARs show significantly higher GC content than SDRs. For platypus, some PARs (X2-PAR-S, X3-PAR-S, X4-PAR-L, X5-PAR-S and X5-PAR-L (where -S is the shorter PAR of the chromosome and -L the longer PAR)) show significantly ($P < 0.01$) higher GC content (1-kb non-overlapping windows) than the SDRs of the same chromosome, which are labelled as asterisks in the heat map. We also checked their orthologous sequences in chicken, as a proxy for the ancestral status before the chromosome became a sex chromosome, and found similarly higher GC content in the orthologous region of PARs than those of SDRs in chicken. *** $P < 0.01$ (all $P < 2.2 \times 10^{-16}$), one-sided Wilcoxon rank-sum test. **d**, Atlas of orthologous chicken fragments along

each platypus sex chromosome. The PARs between the platypus X and Y chromosomes are indicated by crosses. We also labelled the position of the putative sex-determining gene *AMH*.

Extended Data Fig. 5 Evolution of PARs after the platypus and echidna divergence.

a, The distribution of pairwise dS values of platypus and echidna sex chromosomes. In both platypus and echidna, gametologue pairs in the X1 S0 region (Fig. 2), which is largely homologous to chicken chromosome 28, have a higher dS value than those of any other sex-linked regions. This suggests that X1 S0 is the oldest evolutionary stratum. Therefore, we also show platypus genes of X2 with an orthologue on chicken chromosome 28 separately from others (X1_S0_chr28). Following the order of dS values of different chromosome regions, we inferred the time order of formation of evolutionary strata, called S0–S6. For platypus, $n = 5, 5, 2, 2, 1, 1, 4, 2$ and 6 XY gametologue pairs are plotted, from left to right. For echidna, $n = 7, 2, 1, 1, 4, 2$ and 1 XY gametologue pairs are plotted, from left to right. Box plots show median, quartiles (boxes) and range (whiskers). **b**, Phylogenetic tree examples of gametologues that evolved in the common ancestor of monotremes (*EF2* in X2) and independently in two monotreme species (*IRF4* in X3). **c**, Alignments of platypus and echidna X chromosomes (PAR, light colours; SDR, dark colours; the top chromosomes are from platypus and the bottom chromosomes are from echidna) were used to infer X2-PAR-S and X5-PAR-L of platypus evolved independently from echidna after their divergence, given their different lengths. This is supported by the Venn diagrams of PAR genes between platypus and echidna, in which most genes are not shared within independently evolved PARs. **d**, Alignments of PAR–SDR boundaries between platypus and echidna. Alignments of genes (± 1 Mb around the boundaries) support independent evolution of X2-PAR-S and X5-PAR-L in platypus and echidna, as most of their genes are not homologous at the PAR–SDR boundaries (blue, PAR genes; red, SDR genes; platypus, top chromosome, echidna, bottom chromosome). We used lines to connect the genes of the two species, whenever they are orthologous to each other. For each X chromosome, we also labelled their repeat information. Six repeat tracks between each X pair are shown, from top to bottom: the overall repeat content of platypus; LINE/L2 elements of

platypus; SINE/MIR elements of platypus; SINE/MIR elements of echidna; LINE/L2 elements of echidna and overall repeat content of echidna. We did not find obvious repeat enrichment at PAR–SDR boundaries, as shown previously in cow⁹¹.

Extended Data Fig. 6 Sex chromosome evolution in monotremes.

a, Mummerplot showing homology between platypus (*x* axis) and echidna (*y* axis) X chromosomes. Blue lines: forward alignment; red lines: reverse alignment. For echidna, X1, X2 and X3 are homologous to platypus X1, X2 and X3, respectively. Echidna X4 is homologous to platypus X5. And for echidna X5, it is not homologous to any platypus sex chromosome, and instead it is homologous to platypus chromosome 12. **b**, Homology between platypus X chromosomes (*x* axis) and human chromosomes (*y* axis). **c**, Homologous relationships between platypus sex chromosomes and chicken. **d**, Alignment between platypus and chicken showing the alternating pairing pattern of the platypus sex chromosome chain. **e**, X/Y pairwise dS comparison between gametologues on X1–Y5 pair ($n = 18$) and other sex chromosome pairs ($n = 10$). Box plots show median, quartiles (boxes) and range (whiskers). *** $P < 0.001$ ($P = 0.0002954$), one-sided Wilcoxon rank-sum test.

Extended Data Fig. 7 Chromatin conformation of monotreme sex chromosomes.

a, Hi-C interactions between platypus sex chromosomes, with chromosome 1 shown as control. **a, b**, There are unexpected interchromosomal interactions (shown in red) between platypus sex chromosomes detected by Hi-C data (**a**), whereas most interactions are within the same chromosomes (shown in red in **b**) for the other chromosomes (**b**). **c**, The Hi-C interchromosomal interactions among platypus sex chromosome (inter_XY, $n = 2,711$ 100-kb windows) is significantly higher than that among autosomes (inter_A, $n = 14,342,930$ 100-kb window). Box plots show median, quartiles (boxes) and range (whiskers). *** $P < 0.0001$ ($P < 2.2 \times 10^{-16}$), one-sided Wilcoxon rank-sum test. **d**, The interaction strength is higher between Y2 and Y3 than the interaction strengths between Y2 and other chromosomes. $n = 1,002, 228, 5,025, 67,313$ and $6,904,867$ 100-kb windows are shown in Y2-Y2, Y2-Y3, Y2-other.sex.chr, Y2-A and A-A, respectively. Box plots show median, quartiles (boxes) and

range (whiskers). *** $P < 0.0001$ ($P < 2.2 \times 10^{-16}$), one-sided Wilcoxon rank-sum test. **e**, Inferred three-dimensional structure of the platypus sex chromosome system. X chromosomes are shown in red and Y chromosomes in blue, with PARs in light colour. Interchromosomal interactions inferred from Hi-C are shown by dashed lines. **f**, Hi-C interactions reveal unexpected interchromosomal interactions between the echidna sex chromosomes. **g**, Putative CTCF-binding sites are enriched at TAD boundaries in platypus and echidna sex chromosomes. For each X chromosome of platypus, we calculated their putative CTCF-binding-site density per 10 kb and plotted them along the ± 500 kb of TAD boundaries. Platypus X4 and echidna X5 are not shown because less than 10 TAD boundaries are detected. **h**, Putative CTCF-binding-site density plot showing its enrichment among the homologous regions of platypus, echidna, human and chicken.

Extended Data Fig. 8 Loss of dietary-related genes in monotremes.

a, Tooth-related gene loss in representative mammals and reptiles. **b–f**, Potential loss of digestion-related genes in both monotremes shown by whole-genome alignment and read mapping. In each panel there are three lines in the synteny plot, representing the orthologous region of the genes in platypus, human and echidna from top to bottom, respectively. Grey links indicate human–platypus and human–echidna LASTZ alignments. Each rectangle or triangle represents an exon. Fragmented genes are marked by dashed lines. Illumina reads of platypus and echidna are aligned to the platypus or human genome (Ensembl release 87) and the flanking region of each gene is visualized by pyGenomicTrack. *GAPDH* region is also plotted as a control. **g**, RT–PCR expression analysis shows expression of *NGN3* in brain, stomach, intestine and pancreas of both platypus and echidna. These results are similar to other mammals. This, together with sequencing results, shows that *NGN3* in monotremes is present and is likely to be functioning normally. *NGN3*, *NGN3* primers; b-actin, β -actin primers; -ve, negative control, no template; gDNA, genomic DNA template; brain, brain cDNA template; stom, stomach cDNA template; int, intestine cDNA template; panc, pancreas cDNA template. Lanes 1 (top), 1, 8 (middle) and 9 (bottom) are a 100-bp DNA ladder: 1,517, 1,200, 1,000, 900, 800, 700, 600, 500/517,

400, 300, 200 and 100 bp. Expected sizes of PCR products for *NGN3* in platypus is 157 bp and for echidna 145 bp, and the PCR product for the β -actin genomic region is 597 bp and cDNA is 348 bp. Silhouettes of human and opossum are from <https://www.flaticon.com/>. The silhouette of the platypus is created by S. Werning and is reproduced under a Creative Commons Attribution 3.0 Unported licence (<http://creativecommons.org/licenses/by/3.0/>).

Extended Data Fig. 9 Taste-receptor evolution and olfactory-receptor organization in monotremes.

a, Maximum-likelihood mammalian-wide gene tree of the bitter taste receptors (TAS2R genes). There are 28 eutherian (Eu), 27 marsupial (Ma) and 7 monotreme-specific (Mo) orthologous gene groups (supported by $\geq 95\%$ bootstrap values), where the nodes of orthologous gene group clades are indicated by white open circles. Bootstrap values of $\geq 70\%$ in the nodes connecting orthologous gene group clades are indicated by asterisks. There are 3 therian (I, II and III), 2 eutherian (I and II), 3 marsupial (I, II and III) and one monotreme-specific clusters in which massive expansion events occurred in the common ancestor of each taxon after the split from its previous ancestors. **b**, Genomic organization of the intact class I olfactory receptor (OR) cluster spanning over 1.2 Mb on platypus chromosome 2 (138,375,798–139,616,970 bp). The vertical lines indicate the 48 intact class I OR genes. The white open box indicates the J element, a presumable *cis*-regulatory element (enhancer) for the mammalian class I OR cluster (chromosome 2: 139,639,465–139,639,907 bp). Silhouettes of human, opossum and koala are from <https://www.flaticon.com/>. Silhouettes of the platypus and Tasmanian devil are created by S. Werning and are reproduced under a Creative Commons Attribution 3.0 Unported licence (<http://creativecommons.org/licenses/by/3.0/>).

Extended Data Fig. 10 Genomic features related to haemoglobin clearance and reproduction in monotremes.

a, b, Confirmation of *HP* absence in monotremes by whole-genome alignment (**a**) and read mapping (**b**). Grey links indicate human–platypus

and human–echidna LASTZ alignments. Illumina reads of platypus and echidna are aligned to the human genome (Ensembl release 87) and coding regions of *HP* are visualized by pyGenomicTrack. Limited coverage is found at the exons of *HP*, suggesting the absence of *HP* in monotremes. **c**, Phylogenetic tree of HP and related proteases across different species using the maximum-likelihood method. Node IDs are in format of ‘species geneID’. Branch length is not shown here. **d**, Gene synteny plot of the *PIT54* region between chicken and platypus. Echidna is not shown in the figure as the flanking orthologues of *PIT54* are on different scaffolds, preventing us from determining the presence of the gene by synteny. **e**, Phylogenetic tree of members of the group B scavenger receptor cysteine-rich family across different species using the neighbour-joining method. Gene IDs are formatted as ‘species geneID’. Branch length is not shown here. **f**, Confirmation of *SCART1* number difference by dot plot and mapping depth of *SCART1* orthologous regions between platypus and echidna. The region of the *SCART1* cluster in platypus is plotted along the *x* axis while the sequence of echidna is plotted along the *y* axis. Lines in dot plot are visualized according to LASTZ alignment between the two species. Normalized male and female read depths along each sequence is calculated in 500-bp windows, and plotted along each sequence. Normalized depth of both sexes, especially those in the shading region, is centred at 1 along both species, confirming the *SCART1* number difference between the two species is true and is not due to assembly issues. **g**, Synteny conservation of vitellogenin genes. Synteny conservation of the region surrounding the vitellogenin (*VTG*) genes *VTG1*, *VTG2* and *VTG3*. Pseudogenes are marked by a dashed outline. Monotremes have pseudogene *VTG1*, functional *VTG2* and no *VTG3*; and there is a pseudogene *VTG2* in koala. Syntenic maps are shown for human (*Homo sapiens*), koala (*Phascolarctos cinereus*), chicken (*Gallus gallus*), platypus (*O. anatinus*) and echidna (*T. aculeatus*). Koala scaffold 1, NW_018343984.1; koala scaffold 2, NW_018344134.1. Gene distances are not to scale. **h**, Synteny conservation of regions containing *SPINT3*. Synteny conservation of the region surrounding serine peptidase inhibitor, Kunitz-type, 3 (*SPINT3*). No copy of *SPINT3* is detected in platypus but many of the other flanking genes in the region are conserved. Other members with a WFDC domain are detected including two Kunitz-domain members that did not align to any known gene (labelled *KDCP1*). Syntenic maps are reported for human (*H. sapiens*), cow (*B. taurus*), grey

short-tailed opossum (*Monodelphis domestica*), koala (*P. cinereus*) and platypus (*O. anatinus*). Koala scaffold 1, NW_018343967.1; koala scaffold 2, NW_018344098.1. Gene distances are not to scale. **i**, Casein 3 (CSN3) protein sequence alignment in monotremes. All three CSN3 proteins identified in the monotremes have the classic five-exon structure of CSN3 with the untranslated exons I and IV (not shown), the signal peptide in exon II, a small exon III coding for 11 residues, a pSER cluster (S**) at the 5' end of exon IV and a relatively large P/Q-rich exon IV. OA, *O. anatinus* (platypus); TA, *T. aculeatus* (short-beaked echidna). Silhouettes of human, opossum and koala are from <https://www.flaticon.com/>. The silhouette of the platypus is created by S. Werning and is reproduced under a Creative Commons Attribution 3.0 Unported licence (<http://creativecommons.org/licenses/by/3.0/>).

Supplementary information

Supplementary Information

Details on the sample collection and analytical methods used in this study. Also includes Supplementary Results with the detailed analyses results, as well as Supplementary Tables 1, 2, 6, 7, 9, 13-16, 18, 21, 24, 25, 34, 38, 41-45 and 50.

Reporting Summary

Supplementary Tables

This file includes Supplementary Tables 3-5, 8, 10-12, 17, 19, 20, 22, 23, 26-33, 35-37, 39, 40, 46-49 and 51-54.

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



Check for
updates

Cite this article

Zhou, Y., Shearwin-Whyatt, L., Li, J. *et al.* Platypus and echidna genomes reveal mammalian biology and evolution. *Nature* **592**, 756–762 (2021). <https://doi.org/10.1038/s41586-020-03039-0>

[Download citation](#)

- Received: 04 December 2019
- Accepted: 30 July 2020
- Published: 06 January 2021
- Issue Date: 29 April 2021
- DOI: <https://doi.org/10.1038/s41586-020-03039-0>

Further reading

- [Genomics of the Parasitic Nematode Ascaris and Its Relatives](#)

- Jianbin Wang

Genes (2021)

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Download PDF](#)

Associated Content

Special

[Vertebrate Genomes Project](#)

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-020-03039-0>

- Article
- [Published: 24 March 2021](#)

REV-ERB in GABAergic neurons controls diurnal hepatic insulin sensitivity

- [Guolian Ding](#)^{1,2,3} na1,
- [Xin Li](#)³ na1,
- [Xinguo Hou](#) [ORCID: orcid.org/0000-0003-2045-1290](#)⁴ na1,
- [Wenjun Zhou](#)³ na1,
- [Yingyun Gong](#) [ORCID: orcid.org/0000-0002-8570-6904](#)^{3,5},
- [Fuqiang Liu](#)⁴,
- [Yanlin He](#) [ORCID: orcid.org/0000-0002-5471-9016](#)^{6,7},
- [Jia Song](#)⁴,
- [Jing Wang](#)⁴,
- [Paul Basil](#)³,
- [Wenbo Li](#) [ORCID: orcid.org/0000-0002-9126-2529](#)³,
- [Sichong Qian](#)³,
- [Pradip Saha](#)^{3,8},
- [Jinbang Wang](#)⁴,
- [Chen Cui](#)⁴,
- [Tingting Yang](#)^{3,6},
- [Kexin Zou](#)^{1,2},
- [Younghun Han](#)⁹,
- [Christopher I. Amos](#)⁹,
- [Yong Xu](#) [ORCID: orcid.org/0000-0002-4908-1572](#)^{6,8},
- [Li Chen](#) [ORCID: orcid.org/0000-0001-7670-8062](#)⁴ &
- [Zheng Sun](#) [ORCID: orcid.org/0000-0002-6858-0633](#)^{3,8}

- 5860 Accesses
- 143 Altmetric
- [Metrics details](#)

Subjects

- [Hypothalamus](#)
- [Type 2 diabetes](#)

Abstract

Systemic insulin sensitivity shows a diurnal rhythm with a peak upon waking^{1,2}. The molecular mechanism that underlies this temporal pattern is unclear. Here we show that the nuclear receptors REV-ERB- α and REV-ERB- β (referred to here as ‘REV-ERB’) in the GABAergic (γ -aminobutyric acid-producing) neurons in the suprachiasmatic nucleus (SCN) (SCN^{GABA} neurons) control the diurnal rhythm of insulin-mediated suppression of hepatic glucose production in mice, without affecting diurnal eating or locomotor behaviours during regular light–dark cycles. REV-ERB regulates the rhythmic expression of genes that are involved in neurotransmission in the SCN, and modulates the oscillatory firing activity of SCN^{GABA} neurons. Chemogenetic stimulation of SCN^{GABA} neurons at waking leads to glucose intolerance, whereas restoration of the temporal pattern of either SCN^{GABA} neuron firing or REV-ERB expression rescues the time-dependent glucose metabolic phenotype caused by REV-ERB depletion. In individuals with diabetes, an increased level of blood glucose after waking is a defining feature of the ‘extended dawn phenomenon’^{3,4}. Patients with type 2 diabetes with the extended dawn phenomenon exhibit a differential temporal pattern of expression of REV-ERB genes compared to patients with type 2 diabetes who do not have the extended dawn phenomenon. These findings provide mechanistic insights into how the central circadian clock regulates the diurnal rhythm of hepatic insulin sensitivity, with

implications for our understanding of the extended dawn phenomenon in type 2 diabetes.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

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: REV-ERB in GABAergic neurons regulates rhythmic hepatic insulin sensitivity.



Fig. 2: REV-ERB regulates the diurnal rhythm of activity of SCN^{GABA} neurons.

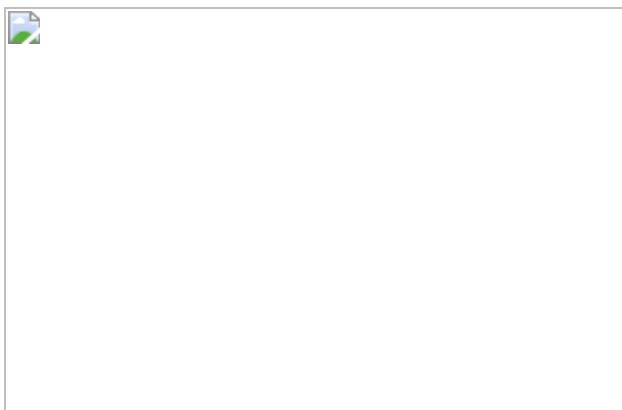
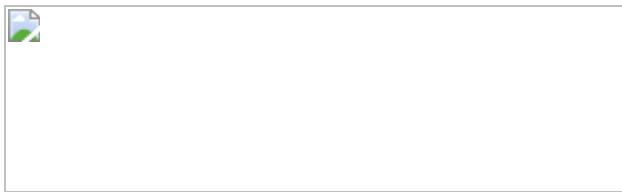


Fig. 3: The rhythmicity of SCN^{GABA} neural activity and REV-ERB expression regulates rhythmic glucose metabolism.



Fig. 4: The dawn phenomenon is associated with altered expression of REV-ERB.



Data availability

The data that support the findings of this study are freely available from the corresponding authors upon request. RNA-seq data are available at the Gene Expression Omnibus (GEO) with the accession code [GSE150840](#). [Source data](#) are provided with this paper.

References

1. 1.

Shi, S. Q., Ansari, T. S., McGuinness, O. P., Wasserman, D. H. & Johnson, C. H. Circadian disruption leads to insulin resistance and obesity. *Curr. Biol.* **23**, 372–381 (2013).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

2. 2.

Coomans, C. P. et al. Detrimental effects of constant light exposure and high-fat diet on circadian energy metabolism and insulin sensitivity. *FASEB J.* **27**, 1721–1732 (2013).

[CAS](#) [PubMed](#) [Google Scholar](#)

3. 3.

O’Neal, T. B. & Luther, E. E. Dawn phenomenon. <https://www.statpearls.com/articlelibrary/viewarticle/20266/> (StatPearls Publishing, 2020).

4. 4.

Monnier, L., Colette, C., Dejager, S. & Owens, D. Magnitude of the dawn phenomenon and its impact on the overall glucose exposure in type 2 diabetes: is this of concern? *Diabetes Care* **36**, 4057–4062 (2013).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

5. 5.

Hastings, M. H., Maywood, E. S. & Brancaccio, M. Generation of circadian rhythms in the suprachiasmatic nucleus. *Nat. Rev. Neurosci.* **19**, 453–469 (2018).

[CAS](#) [PubMed](#) [Google Scholar](#)

6. 6.

Zhang, R., Lahens, N. F., Ballance, H. I., Hughes, M. E. & Hogenesch, J. B. A circadian gene expression atlas in mammals: implications for biology and medicine. *Proc. Natl Acad. Sci. USA* **111**, 16219–16224 (2014).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

7. 7.

Cho, H. et al. Regulation of circadian behaviour and metabolism by REV-ERB- α and REV-ERB- β . *Nature* **485**, 123–127 (2012).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

8. 8.

Zhang, Y. et al. Discrete functions of nuclear receptor Rev-erba couple metabolism to the clock. *Science* **348**, 1488–1492 (2015).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

9. 9.

Doi, M. et al. Circadian regulation of intracellular G-protein signalling mediates intercellular synchrony and rhythmicity in the suprachiasmatic nucleus. *Nat. Commun.* **2**, 327 (2011).

[ADS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

10. 10.

Tu, S. et al. Takusan: a large gene family that regulates synaptic activity. *Neuron* **55**, 69–85 (2007).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

11. 11.

Panda, S. et al. Coordinated transcription of key pathways in the mouse by the circadian clock. *Cell* **109**, 307–320 (2002).

[CAS](#) [PubMed](#) [Google Scholar](#)

12. 12.

Adelmant, G., Bègue, A., Stéhelin, D. & Laudet, V. A functional Rev-erba responsive element located in the human Rev-erba promoter mediates a repressing activity. *Proc. Natl Acad. Sci. USA* **93**, 3553–3558 (1996).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

13. 13.

Carroll, M. F. & Schade, D. S. The dawn phenomenon revisited: implications for diabetes therapy. *Endocr. Pract.* **11**, 55–64 (2005).

[PubMed](#) [Google Scholar](#)

14. 14.

Porcellati, F., Lucidi, P., Bolli, G. B. & Fanelli, C. G. Thirty years of research on the dawn phenomenon: lessons to optimize blood glucose control in diabetes. *Diabetes Care* **36**, 3860–3862 (2013).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

15. 15.

Cuesta, M., Boudreau, P., Cermakian, N. & Boivin, D. B. Rapid resetting of human peripheral clocks by phototherapy during simulated night shift work. *Sci. Rep.* **7**, 16310 (2017).

[ADS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

16. 16.

Akashi, M. et al. Noninvasive method for assessing the human circadian clock using hair follicle cells. *Proc. Natl Acad. Sci. USA* **107**, 15643–15648 (2010).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

17. 17.

la Fleur, S. E., Kalsbeek, A., Wortel, J., Fekkes, M. L. & Buijs, R. M. A daily rhythm in glucose tolerance: a role for the suprachiasmatic nucleus. *Diabetes* **50**, 1237–1243 (2001).

[PubMed](#) [Google Scholar](#)

18. 18.

Coomans, C. P. et al. The suprachiasmatic nucleus controls circadian energy metabolism and hepatic insulin sensitivity. *Diabetes* **62**, 1102–1108 (2013).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

19. 19.

Foppen, E., Tan, A. A., Ackermans, M. T., Fliers, E. & Kalsbeek, A. Suprachiasmatic nucleus neuropeptides and their control of endogenous glucose production. *J. Neuroendocrinol.* **28**, <https://doi.org/10.1111/jne.12365> (2016).

20. 20.

Kalsbeek, A., Yi, C.-X., La Fleur, S. E. & Fliers, E. The hypothalamic clock and its control of glucose homeostasis. *Trends Endocrinol. Metab.* **21**, 402–410 (2010).

[CAS](#) [PubMed](#) [Google Scholar](#)

21. 21.

Bolli, G. B. et al. Demonstration of a dawn phenomenon in normal human volunteers. *Diabetes* **33**, 1150–1153 (1984).

[CAS](#) [PubMed](#) [Google Scholar](#)

22. 22.

Van Cauter, E., Polonsky, K. S. & Scheen, A. J. Roles of circadian rhythmicity and sleep in human glucose regulation. *Endocr. Rev.* **18**, 716–738 (1997).

[PubMed](#) [Google Scholar](#)

23. 23.

Boden, G., Chen, X. & Urbain, J. L. Evidence for a circadian rhythm of insulin sensitivity in patients with NIDDM caused by cyclic changes in hepatic glucose production. *Diabetes* **45**, 1044–1050 (1996).

[CAS](#) [PubMed](#) [Google Scholar](#)

24. 24.

Radziuk, J. & Pye, S. Diurnal rhythm in endogenous glucose production is a major contributor to fasting hyperglycaemia in type 2 diabetes. Suprachiasmatic deficit or limit cycle behaviour? *Diabetologia* **49**, 1619–1628 (2006).

[CAS](#) [PubMed](#) [Google Scholar](#)

25. 25.

Albus, H., Vansteensel, M. J., Michel, S., Block, G. D. & Meijer, J. H. A GABAergic mechanism is necessary for coupling dissociable ventral and dorsal regional oscillators within the circadian clock. *Curr. Biol.* **15**, 886–893 (2005).

[CAS](#) [PubMed](#) [Google Scholar](#)

26. 26.

Choi, H. J. et al. Excitatory actions of GABA in the suprachiasmatic nucleus. *J. Neurosci.* **28**, 5450–5459 (2008).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

27. 27.

Freeman, G. M., Jr, Krock, R. M., Aton, S. J., Thaben, P. & Herzog, E. D. GABA networks destabilize genetic oscillations in the circadian pacemaker. *Neuron* **78**, 799–806 (2013).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

28. 28.

Vong, L. et al. Leptin action on GABAergic neurons prevents obesity and reduces inhibitory tone to POMC neurons. *Neuron* **71**, 142–154 (2011).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

29. 29.

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).

[CAS](#) [PubMed](#) [Google Scholar](#)

30. 30.

Siepka, S. M. & Takahashi, J. S. Methods to record circadian rhythm wheel running activity in mice. *Methods Enzymol.* **393**, 230–239 (2005).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

31. 31.

Atasoy, D., Aponte, Y., Su, H. H. & Sternson, S. M. A FLEX switch targets channelrhodopsin-2 to multiple cell types for imaging and long-range circuit mapping. *J. Neurosci.* **28**, 7025–7030 (2008).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

32. 32.

Sprengel, R. & Hasan, M. T. Tetracycline-controlled genetic switches. *Handb. Exp. Pharmacol.* **178**, 49–72 (2007).

[CAS](#) [Google Scholar](#)

33. 33.

Ochoa, C. D., Alexeyev, M., Pastukh, V., Balczon, R. & Stevens, T. *Pseudomonas aeruginosa* exotoxin Y is a promiscuous cyclase that increases endothelial tau phosphorylation and permeability. *J. Biol. Chem.* **287**, 25407–25418 (2012).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

34. 34.

Hockemeyer, D. et al. A drug-inducible system for direct reprogramming of human somatic cells to pluripotency. *Cell Stem Cell* **3**, 346–353 (2008).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

35. 35.

Roth, B. L. DREADDs for neuroscientists. *Neuron* **89**, 683–694 (2016).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

36. 36.

Krashes, M. J. et al. Rapid, reversible activation of AgRP neurons drives feeding behavior in mice. *J. Clin. Invest.* **121**, 1424–1428 (2011).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

37. 37.

Armbruster, B. N., Li, X., Pausch, M. H., Herlitze, S. & Roth, B. L. Evolving the lock to fit the key to create a family of G protein-coupled receptors potently activated by an inert ligand. *Proc. Natl Acad. Sci. USA* **104**, 5163–5168 (2007).

[ADS](#) [PubMed](#) [Google Scholar](#)

38. 38.

Ren, H. et al. FoxO1 target Gpr17 activates AgRP neurons to regulate food intake. *Cell* **149**, 1314–1326 (2012).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

39. 39.

Liu, T. et al. Fasting activation of AgRP neurons requires NMDA receptors and involves spinogenesis and increased excitatory tone. *Neuron* **73**, 511–522 (2012).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

40. 40.

Fenselau, H. et al. A rapidly acting glutamatergic ARC→PVH satiety circuit postsynaptically regulated by α-MSH. *Nat. Neurosci.* **20**, 42–51 (2017).

[CAS](#) [PubMed](#) [Google Scholar](#)

41. 41.

Itri, J., Michel, S., Waschek, J. A. & Colwell, C. S. Circadian rhythm in inhibitory synaptic transmission in the mouse suprachiasmatic nucleus. *J. Neurophysiol.* **92**, 311–319 (2004).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

42. 42.

Kim, D. et al. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* **14**, R36 (2013).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

43. 43.

Anders, S., Pyl, P. T. & Huber, W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

44. 44.

Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol.* **11**, R106 (2010).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

45. 45.

Danne, T. et al. International consensus on use of continuous glucose monitoring. *Diabetes Care* **40**, 1631–1640 (2017).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

46. 46.

Faraco, G. et al. Dietary salt promotes neurovascular and cognitive dysfunction through a gut-initiated TH17 response. *Nat. Neurosci.* **21**, 240–249 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

47. 47.

Peixoto, R. T., Wang, W., Croney, D. M., Kozorovitskiy, Y. & Sabatini, B. L. Early hyperactivity and precocious maturation of corticostriatal circuits in *Shank3B*^{-/-} mice. *Nat. Neurosci.* **19**, 716–724 (2016).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

48. 48.

Witton, J. et al. Hippocampal circuit dysfunction in the Tc1 mouse model of Down syndrome. *Nat. Neurosci.* **18**, 1291–1298 (2015).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

49. 49.

Xu, P. et al. Estrogen receptor- α in medial amygdala neurons regulates body weight. *J. Clin. Invest.* **125**, 2861–2876 (2015).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

50. 50.

Perusini, J. N. et al. Optogenetic stimulation of dentate gyrus engrams restores memory in Alzheimer's disease mice. *Hippocampus* **27**, 1110–1122 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

51. 51.

Wang, W. et al. Chemogenetic activation of prefrontal cortex rescues synaptic and behavioral deficits in a mouse model of 16p11.2 deletion syndrome. *J. Neurosci.* **38**, 5939–5948 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

[Download references](#)

Acknowledgements

We thank M. Lazar for the *Nr1d1*^{loxP} and *Nr1d2*^{loxP} mice; Q. Tong and Y. Xu for critical reading of the manuscript and technical guidance; H. Liu for technical consultation; X. Wang for statistics consultation; C. Yu and S. S. B. Lee for technical assistance; K. Oka and B. Arenkiel for viral vector production; and C. Ljungberg (U54HD083092 and 1S10OD016167) for some of the histology studies. The Mouse Metabolism and Phenotyping Core at Baylor College of Medicine was supported by R01DK114356 and UM1HG006348. The authors were supported by grants from the National Natural Science Foundation of China (81971458 and 31671222 to G.D.); the American Diabetes Association (ADA1-17-PDF-138 to Y. He, ADA1-19-PDF-012 to W.Z. and NIH P20 GM135002 to Y. He); the US Department of Agriculture (Cris51000-064-01S to Y.X.); the National Key R&D Program of China (2016YFC0901204 and 2018YFC1311801 to L.C. and 2017YFC1001300 to G.D.); and the NIH (R01DK111436, R01HL153320, RF1AG069966 and R01ES027544 to Z.S.). We also thank the John S. Dunn Foundation, the Mrs. Clifford Elder White Graham Endowed Research Fund, the Cardiovascular Research Institute at Baylor

College of Medicine, the Dan L Duncan Comprehensive Cancer Center (P30CA125123), the Texas Medical Center Digestive Diseases Center (P30DK056338), the SPORE program in lymphoma at Baylor College of Medicine (P50 CA126752) and the Gulf Coast Center for Precision Environmental Health (P30ES030285).

Author information

Author notes

1. These authors contributed equally: Guolian Ding, Xin Li, Xinguo Hou, Wenjun Zhou

Affiliations

1. Obstetrics and Gynecology Hospital, Institute of Reproduction and Development, Fudan University, Shanghai, China

Guolian Ding & Kexin Zou

2. Shanghai Key Laboratory of Embryo Original Diseases, Shanghai, China

Guolian Ding & Kexin Zou

3. Department of Medicine, Division of Diabetes, Endocrinology and Metabolism, Baylor College of Medicine, Houston, TX, USA

Guolian Ding, Xin Li, Wenjun Zhou, Yingyun Gong, Paul Basil, Wenbo Li, Sichong Qian, Pradip Saha, Tingting Yang & Zheng Sun

4. Department of Endocrinology, Qilu Hospital of Shandong University, Jinan, China

Xinguo Hou, Fuqiang Liu, Jia Song, Jing Wang, Jinbang Wang, Chen Cui & Li Chen

5. Department of Endocrinology and Metabolism, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China

Yingyun Gong

6. USDA/ARS Children's Nutrition Research Center, Department of Pediatrics, Baylor College of Medicine, Houston, TX, USA

Yanlin He, Tingting Yang & Yong Xu

7. Laboratory of Brain Glycemia and Metabolism Control, Pennington Biomedical Research Center, Baton Rouge, LA, USA

Yanlin He

8. Department of Molecular and Cellular Biology, Baylor College of Medicine, Houston, TX, USA

Pradip Saha, Yong Xu & Zheng Sun

9. Department of Medicine, Section of Epidemiology and Population Sciences, Baylor College of Medicine, Houston, TX, USA

Younghun Han & Christopher I. Amos

Authors

1. Guolian Ding

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Xin Li

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Xinguo Hou

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Wenjun Zhou

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Yingyun Gong

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Fuqiang Liu

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Yanlin He

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Jia Song

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. Jing Wang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. Paul Basil

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

11. Wenbo Li

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

12. Sichong Qian

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

13. Pradip Saha

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

14. Jinbang Wang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

15. Chen Cui

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

16. Tingting Yang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

17. Kexin Zou

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

18. Younghun Han

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

19. Christopher I. Amos

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

20. Yong Xu

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

21. Li Chen

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

22. Zheng Sun

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Contributions

Z.S. conceived the study. G.D. identified the mouse phenotype and oversaw the human study. X.L. performed gene expression analysis, chemogenetic studies, and stereotaxic injections in mice. X.H. recruited study participants and supervised the human study. W.Z. performed histological studies and ChIP. Y.G. coordinated the insulin clamp. W.L. made DNA constructs. S.Q. performed some of the mouse metabolic tests. J.S. performed gene expression analyses in human samples. J.W., F.L., J.W. and C.C. collected human blood samples, made clinical measurements and coordinated clinical studies. Y. He performed patch clamp recording. P.B. performed the initial mouse crossbreeding. G.D., X.L. and Y.G. maintained the mouse lines. P.S. performed CLAMS and insulin clamp analyses. G.D., X.L., W.Z., Y.G., J.S., J.W., Y. He, P.B., T.Y. and Z.S. analysed the data. K.Z., Y. Han and C.I.A. performed statistical analyses. Y.X., X.H. and Z.S. interpreted the data. L.C. and Z.S. obtained funding. G.D. and Z.S. wrote the manuscript with input from other authors.

Corresponding authors

Correspondence to [Li Chen](#) or [Zheng Sun](#).

Ethics declarations

Competing interests

The authors declare no financial or non-financial conflict of interest. No patent was involved in the study.

Additional information

Peer review information *Nature* thanks Erik D. Herzog, Satchidananda Panda 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 Behavioural characterization of knockout mice.](#)

a, RNAscope analysis of *Nrl1d1* (REV-ERB- α) gene expression at ZT6–9 and ZT18–21 in three-month-old wild-type mice. Scale bars, 200 μ m **b**, Representative wheel-running actogram for five-month-old mice in light–dark conditions (LD) or constant darkness (DD). **c**, Phase angle of the light entrainment in the last day of the light–dark conditions ($n = 7$ mice). **d–i**, Representative chi-square periodograms (**d**, **e**, **h**, **i**) and period length (**f**, **g**) for five-month-old mice in light–dark conditions or constant darkness ($n = 7$ mice). **j**, Average wheel-running activity in constant darkness after normalization to the intrinsic period (τ) ($n = 7$ mice). Data are mean \pm s.e.m. * $P < 0.05$ by two-sided t -test. Statistical details are in Supplementary Table 1. [Source data](#)

[Extended Data Fig. 2 Metabolic characterization of knockout mice on a normal chow diet.](#)

a, Daily food intake in three-month-old mice in home cages ($n = 4$ cages across 20 days). **b, c**, Food intake measured by the comprehensive laboratory animal monitoring system (CLAMS) in the 6 h (**b**) or 12 h (**c**) before GTT analyses at four months old ($n = 5$ mice). Box plot centre lines, box limits and whiskers represent the median, quartiles and minimum and maximum values, respectively. **d**, Blood glucose levels in four-month-old mice ($n = 14$ wild-type mice; $n = 10$ knockout mice). **e**, Serum insulin levels ($n = 10$ mice per group). **f**, Blood glucagon levels ($n = 12$ mice per group). **g**, Blood corticosterone levels ($n = 11$ mice per group). **h**, Blood GLP-1 levels ($n = 12$ wild-type mice; $n = 10$ knockout mice). **i**, Blood growth hormone (GH) levels ($n = 11$ wild-type mice; $n = 12$ knockout mice). **j, k**, GTTs in five-month-old mice at ZT6–8 (**j**) or ZT12–14 (**k**), with *vgat-cre* mice serving as the wild-type control ($n = 7$ mice). **l**, Body weight for clamp analyses at five months old ($n = 4$ mice). **m, n**, Blood glucose levels (**m**) and glucose infusion rate (**n**) during clamp analyses ($n = 4$ mice). **o**, Hyperinsulinaemia-mediated suppression of hepatic glucose production in the clamp analyses ($n = 4$ mice). Data are mean \pm s.e.m. * $P < 0.05$ by two-way ANOVA or two-sided *t*-test. Statistical details are in Supplementary Table 1. [Source data](#)

Extended Data Fig. 3 Metabolic characterization of knockout mice on a high-fat diet.

a, Body weight on HFD. HFD started at 10 months ($n = 12$ mice). **b**, GTTs at ZT6–8 after two weeks on HFD ($n = 12$ mice). **c**, GTTs at ZT12–14 after three weeks on HFD ($n = 12$ mice). **d**, ITTs at ZT6–8 after four weeks on HFD ($n = 12$ mice). **e**, ITTs at ZT12–14 after five weeks on HFD ($n = 12$ mice). **f**, Injection of streptozotocin (STZ) six weeks after HFD. **g, h**, Body weight (**g**) and blood glucose levels (**h**) at ZT10 after STZ injection ($n = 12$ mice). **i**, Blood glucose levels at the indicated ZTs two weeks after STZ injection, $n = 12$ mice. Data are mean \pm s.e.m. * $P < 0.05$ by two-way ANOVA or two-sided *t*-test. Statistical details are in Supplementary Table 1. [Source data](#)

Extended Data Fig. 4 Gene expression analysis of different brain regions.

a–e, RT–qPCR analysis of the indicated brain-region-specific marker genes (*Vip* (a), *Pmch* (b), *Crf* (c), *Rfrp* (also known as *Npyf*) (d) and *Pomc1* (also known as *Pomc*) (e)) for brain regions isolated from both wild-type and knockout mice at ZT6 at the age of three months ($n = 12$ mice). Box plot centre lines, box limits and whiskers represent the median, quartiles and minimum and maximum values, respectively. **f–i**, RT–qPCR analysis comparing mRNA expression of *Nr1d1* (f), *Nr1d2* (g), *Bmal1* (h) and *Npas2* (i) in wild-type and knockout mice at ZT6 at the age of three months ($n = 6$ mice). Data are mean \pm s.e.m. * $P < 0.05$ by two-way ANOVA or two-sided *t*-test. Statistical details are in Supplementary Table 1. [Source data](#)

Extended Data Fig. 5 Electrophysiological and molecular characterization of knockout mice.

a–c, Brain slice patch-clamp representative traces for spontaneous firing (a), mEPSCs (b) and mIPSCs (c) at ZT12–14. **d–g**, Temporal pattern of expression of *Rgs16* (d), α7-Takusan (*Gm10406*) (e), *Nr1d1* (f) and *Nr1d2* (g) in the hypothalamus in light–dark conditions from CircaDB (<http://circadb.hogeneschlab.org>). **h–k**, RT–qPCR analysis of the mRNA levels of *Nr1d1* (h), *Nr1d2* (i), *Bmal1* (j) and *Npas2* (k) in the SCN of three-month-old mice ($n = 6$ mice). Primers for *Nr1d1* and *Nr1d2* did not span the floxed exons. **l**, RNAscope of *Rgs16* at the SCN in wild-type and knockout mice at the indicated ZTs. Scale bars, 100 μm. **m**, Quantification of *Rgs16* staining ($n = 5$ wild-type mice at ZT4; $n = 3$ knockout mice at ZT4; $n = 4$ wild-type or knockout mice at ZT16). **n**, In situ hybridization analysis of Takusan *Gm3500* staining. Scale bars, 25 μm. **o**, Quantification of in situ hybridization analysis of Takusan *Gm3500* ($n = 4$ wild-type mice at ZT4; $n = 6$ wild-type mice at ZT16; $n = 3$ knockout mice at ZT4 or ZT16). **p**, Genome browser views of transcription start sites (green arrows) and nearby AGGTCA elements (red arrows) for the indicated genes on GRCm38. **q**, ChIP–qPCR analysis of *Nr1d1* in the hypothalamus of three-month-old wild-type mice at ZT9 and ZT21—the peak and the trough of REV-ERB-α expression, respectively ($n = 4$ samples). Hypothalami from five mice were pooled as one sample. The negative control primers target a gene desert region on chromosome 17. The primer sequences of ChIP–qPCR assays are in Supplementary Table 6. Data are mean \pm s.e.m.

* $P < 0.05$ by two-way ANOVA or two-sided t -test. Statistical details are in Supplementary Table 1. [Source data](#)

Extended Data Fig. 6 Metabolic characterization of mice overexpressing RGS16 or α 7-Takusan in SCN^{GABA} neurons.

a, Validation of the injection with GFP fluorescence signals. Scale bar, 200 μ m. **b–e**, GTTs (**b**, **c**) and ITTs (**d**, **e**) at the indicated ZTs in four-month-old *vgat-cre* mice injected with AAV-FLEX vectors encoding GFP, RGS16 or α 7-Takusan ($n = 7$ mice). **f**, Body weight of three-month-old *vgat-cre* mice at three weeks after AAV injection ($n = 14$ mice). Data are mean \pm s.e.m. * $P < 0.05$ for RGS16 or α 7-Takusan versus the GFP control by two-way ANOVA followed by Holm–Sidak’s test. Statistical details are in Supplementary Table 1. [Source data](#)

Extended Data Fig. 7 Rhythmicity of SCN^{GABA} neuron firing in glucose metabolism.

a, Experimental design for chemogenetic activation of the SCN^{GABA} neurons in wild-type mice with hM3Dq. **b**, Body weight of *vgat-cre* mice injected with AAV expressing hM3Dq or control mCherry ($n = 11$ mice). Mice were injected at the age of two months. **c**, Experimental design for chemogenetic repression of the SCN^{GABA} neurons in wild-type and knockout mice with hM4Di. **d**, Body weight of wild-type and knockout mice injected with AAV expressing hM4Di ($n = 12$ wild-type mice; $n = 14$ knockout mice). Mice were injected at the age of two months. **e**, **f**, GTTs in wild-type or knockout mice injected with AAV expressing hM4Di at the indicated ZTs in the presence of saline (**e**) or CNO (**f**) ($n = 12$ wild-type mice; $n = 14$ knockout mice). Data are mean \pm s.e.m. * $P < 0.05$ by two-sided t -test. [Source data](#)

Extended Data Fig. 8 Rhythmicity of REV-ERB expression in SCN^{GABA} neurons in glucose metabolism.

a, Experimental design for inducible re-expression of REV-ERB- α in the SCN^{GABA} neurons of knockout mice. Mice were injected virus at the age of 2.5 months. **b**, Body weight at the time of euthanasia ($n = 9$ mice). **c**, **d**, GTTs in 4–4.5-month-old mice at ZT6–8 after injection of doxycycline at ZT0 (**c**) or ZT18 (**d**) ($n = 9$ mice). **e**, RT–qPCR analysis of the SCN from knockout mice with inducible re-expression of REV-ERB- α . Doxycycline was injected at ZT0 and the brain was collected at ZT12–14 ($n = 4$ mice). Data are mean \pm s.e.m. * $P < 0.05$ by two-sided t -test. Statistical details are in Supplementary Table 1. [Source data](#)

Extended Data Fig. 9 Assessment of CGM performance.

a, A representative comparison between fingertip glucometer reading and CGM reading for an individual at different times of the day. **b**, Pearson correlation coefficient between CGM and fingertip readings ($n = 16$ DP–; $n = 11$ DP+). **c**, MARD, the average of the absolute error between all CGM values and matched reference values ($n = 16$ DP–; $n = 11$ DP+). Data are mean \pm s.e.m. [Source data](#)

Supplementary information

Supplementary Tables

This file contains Supplementary Tables 1–6. Supplementary Table 1. Statistical tests. Statistical details were shown for results with significant differences, including the names of the statistical methods, p values, and sample sizes for those not included in the main figure legends due to limited space. Supplementary Table 2. Primer sequences for RT-qPCR and ChIP-qPCR. Nucleic acid sequences (from the 5' end to the 3' end) were shown for RT-qPCR primers, ChIP-qPCR primers, and the ISH probe. Supplementary Table 3. Inclusion and exclusion criteria for patient recruitment. Supplementary Table 4. Characteristics of human subjects. Supplementary Table 5. Cardiopulmonary Coupling-Polysomnography (CPC-PSG) of human subjects. Supplementary Table 6. Medication usage in human subjects.

Reporting Summary

Source data

[Source Data Fig. 1](#)

[Source Data Fig. 2](#)

[Source Data Fig. 3](#)

[Source Data Fig. 4](#)

[Source Data Extended Data Fig. 1](#)

[Source Data Extended Data Fig. 2](#)

[Source Data Extended Data Fig. 3](#)

[Source Data Extended Data Fig. 4](#)

[Source Data Extended Data Fig. 5](#)

[Source Data Extended Data Fig. 6](#)

[Source Data Extended Data Fig. 7](#)

[Source Data Extended Data Fig. 8](#)

[Source Data Extended Data Fig. 9](#)

Rights and permissions

[Reprints and Permissions](#)

About this article



Cite this article

Ding, G., Li, X., Hou, X. *et al.* REV-ERB in GABAergic neurons controls diurnal hepatic insulin sensitivity. *Nature* **592**, 763–767 (2021).
<https://doi.org/10.1038/s41586-021-03358-w>

Download citation

- Received: 17 June 2020
- Accepted: 12 February 2021
- Published: 24 March 2021
- Issue Date: 29 April 2021
- DOI: <https://doi.org/10.1038/s41586-021-03358-w>

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

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

A biosensor for the direct visualization of auxin

[Download PDF](#)

- Article
- Open Access
- [Published: 07 April 2021](#)

A biosensor for the direct visualization of auxin

- [Ole Herud-Sikimić](#) [ORCID: orcid.org/0000-0003-3791-9064](#)^{1 na1},
- [Andre C. Stiel](#) [ORCID: orcid.org/0000-0001-8675-6797](#)^{1,2 na1},
- [Martina Kolb](#)^{1 na1},
- [Sooruban Shanmugaratnam](#) [ORCID: orcid.org/0000-0002-2614-6046](#)^{1,3},
- [Kenneth W. Berendzen](#) [ORCID: orcid.org/0000-0001-9870-6164](#)⁴,
- [Christian Feldhaus](#)¹,
- [Birte Höcker](#) [ORCID: orcid.org/0000-0002-8250-9462](#)^{1,3} &
- [Gerd Jürgens](#) [ORCID: orcid.org/0000-0003-4666-8308](#)¹

[Nature](#) volume 592, pages768–772(2021)[Cite this article](#)

- 14k Accesses
- 338 Altmetric
- [Metrics details](#)

Subjects

- [Auxin](#)
- [Cellular imaging](#)
- [Optical imaging](#)
- [Plant physiology](#)

Abstract

One of the most important regulatory small molecules in plants is indole-3-acetic acid, also known as auxin. Its dynamic redistribution has an essential role in almost every

aspect of plant life, ranging from cell shape and division to organogenesis and responses to light and gravity^{1,2}. So far, it has not been possible to directly determine the spatial and temporal distribution of auxin at a cellular resolution. Instead it is inferred from the visualization of irreversible processes that involve the endogenous auxin-response machinery^{3,4,5,6,7}; however, such a system cannot detect transient changes. Here we report a genetically encoded biosensor for the quantitative *in vivo* visualization of auxin distribution. The sensor is based on the *Escherichia coli* tryptophan repressor⁸, the binding pocket of which is engineered to be specific to auxin. Coupling of the auxin-binding moiety with selected fluorescent proteins enables the use of a fluorescence resonance energy transfer signal as a readout. Unlike previous systems, this sensor enables direct monitoring of the rapid uptake and clearance of auxin by individual cells and within cell compartments in planta. By responding to the graded spatial distribution along the root axis and its perturbation by transport inhibitors—as well as the rapid and reversible redistribution of endogenous auxin in response to changes in gravity vectors—our sensor enables real-time monitoring of auxin concentrations at a (sub)cellular resolution and their spatial and temporal changes during the lifespan of a plant.

[Download PDF](#)

Main

The tryptophan-derived metabolite indole-3-acetic acid (IAA) has an important role in plants, triggering a multitude of developmental processes and responses to environmental cues^{1,2}. Much progress has been made towards a mechanistic understanding of the nuclear events that transform auxin perception into transcriptional responses^{9,10,11}. Other studies have investigated the basic machinery involved in the polar and non-vectorial release of auxin from the cell—which occurs through the action of PINFORMED efflux transporters and ABCB transporters—within a tissue context, and have resulted in computer models of how canalized auxin flow mediates developmental or physiological processes^{12,13,14}. By contrast, owing to technical limitations (reviewed in ref. ¹⁵), very little is known about the actual distribution of auxin in tissues at single-cell resolution. At present, plant biologists can use only proxies to visualize auxin distribution, such as the auxin-dependent expression of reporter genes (for example, using the systems DR5::GUS³; DR5::ER-GFP⁴ and DR5::NLS-3xGFP⁵). However, this indirect approach is characterized by latencies and can be affected by modulation of the auxin signalling machinery. More recently, IAA levels have been inferred from auxin-dependent degradation—and thus signal reduction—of fluorescent proteins linked to domain II of an IAA inhibitory protein (for example, DII-VENUS⁶ and R2D2⁷). A limitation of these approaches is

their irreversibility, which precludes the visualization of transient changes in auxin levels.

The ideal sensor for the visualization of auxin dynamics in planta should have the following features: first, physical interaction of the sensor with auxin should elicit a fluorescent signal in a reversible manner, so that changes in auxin concentration can be monitored; second, the sensitivity of the sensor should be sufficiently high to image the dynamic auxin distribution over time; third, the sensor should be targetable to different subcellular compartments—locations that are out of reach for the conventional proxies, which rely on gene expression or protein degradation; and fourth, the sensor should not contain components that are involved in plant metabolism or regulation, such that both interference with auxin responses and regulation of the sensor by the plant are avoided.

With these boundary conditions in mind, we developed a genetically encoded, fully reversible biosensor for *in vivo* imaging of auxin gradients with high spatial and temporal resolution, starting from the bacterial tryptophan repressor (TrpR). IAA resembles tryptophan: both contain an indole ring, the 3-position of which is connected to an amino acid moiety in TRP and a carboxyl group in IAA (Fig. 1a). The dimeric TrpR undergoes a conformational change upon binding TRP^{16,17}, and fluorescent proteins fused to TrpR can relay this change, generating a fluorescence resonance energy transfer (FRET) signal as a convenient readout for *in vivo* measurements¹⁸ (Fig. 1b). Furthermore, TrpR exhibits low affinity towards IAA⁸. This makes TrpR an ideal starting point for developing an auxin-specific, genetically encoded FRET biosensor¹⁹. Our design efforts were aimed at improving the affinity and specificity of IAA binding, while abolishing TRP binding. We assumed a comparable binding mode for the indole ring of both TRP and IAA, and focused our design around the TrpR residues in the vicinity of the amino group of TRP (Fig. 1c), aiming to improve the specificity for the carboxyl group of IAA. This selection was later expanded to include adjacent residues. Altogether, 2,000 variants were generated in successive rounds of mutagenesis and were screened for an increase in FRET signal upon the addition of IAA (Fig. 1d, Extended Data Fig. 1). Improved variants were checked for ligand specificity using a library of substances that are similar to IAA and are reportedly present in *Arabidopsis* (Extended Data Table 1). To confirm improvements in binding affinity, selected TrpR variants were analysed by isothermal titration calorimetry (Extended Data Table 2a). Furthermore, the structures of several variants were elucidated by X-ray crystallography, to guide mutagenesis experiments (Extended Data Table 2b, Supplementary Table 1a–h).

Fig. 1: Summary of the design process.

 **figure1**

a, Chemical structures of TRP and IAA. **b**, Principle of the sensor design. Only in the presence of IAA (red) are the fluorophores (mNeonGreen and Aquamarine) sufficiently close and in the correct orientation for energy transfer (E_{FRET}). N and C represent the N and C termini of the proteins, respectively; L represents the linker; and λ_{ex} and λ_{em} represent the excitation and emission wavelengths, respectively. **c**, Structure of the binding pocket of TrpR with ligand in side view (boxed) (modified from ref. ⁸). Interactions with the side chains of R84, S88 and T44 (second TrpR chain) as well as the backbone carbonyl groups of L41 and L43 (second TrpR chain) are shown explicitly. Further residues that were mutated in this study are indicated with arrows. **d**, Major steps in the design of the sensor (AuxSen), and their cumulative contribution to the change in FRET ratio ($\Delta FRET$) plotted against IAA concentration . Template sensor construct, TrpR–eCFP–Venus (blue squares); engineered binding pocket for IAA, TrpR(M42F/T44L/T81M/N87G/S88Y)–eCFP–Venus (green diamonds); optimized fluorophore combination, TrpR(M42F/T44L/T81M/N87G/S88Y)–mNeonGreen–Aquamarine (purple triangles);

AuxSen, TrpR(M42F/T44L/T81M/N87G/S88Y)–mNeonGreen–Aquamarine with optimised linkers I, II and III (light blue inverted triangles).

[Full size image](#)

Our structural analysis showed that, when binding to TrpR, IAA is flipped by 180° compared with TRP (Fig. 2a), with the carboxyl group of IAA facing the opening of the TrpR binding pocket (Fig. 2b). TRP is anchored by interactions with the surrounding protein residues, whereas IAA binding shows no such stabilization, which is reflected in the poor binding affinity of this interaction (Extended Data Table 2a). In engineering the auxin sensor, we identified variants that stabilize and favour this IAA-binding mode. Primarily, a serine-to-tyrosine mutation at position 88 (S88Y) was found to entirely block TRP binding owing to the bulky side chain of Y88, while simultaneously favouring IAA binding through interaction of the carboxyl group of IAA with the guanidino group of R84 and the hydroxyl group of Y88 (Fig. 2c). The affinity for IAA was improved further by optimizing hydrophobic interactions of its indole ring with the TrpR binding pocket; to this end, the mutations T44L and T81M were incorporated in the final sensor design (Fig. 2d). During the engineering process, we also monitored the binding of indole-3 acetonitrile (IAN)—which could potentially compete with IAA—to TrpR (Extended Data Table 2a). The binding modes of IAN and IAA are markedly similar (Extended Data Fig. 2a, b); however, mutations such as N87G exert discriminating effects through small changes in the positioning of Y88 (Extended Data Fig. 2c). Finally, we identified mutations that have no favourable effect on IAA affinity but improve the FRET readout—probably through changes in the packing, and therefore the orientation, of the attached fluorescent proteins (Extended Data Fig. 2d). We then optimized the fluorophores and the linker combinations (Extended Data Fig. 3) to yield our final sensor, which we term ‘AuxSen’, with the composition mNeonGreen–TrpR–Aquamarine–TrpR, in which TrpR is TrpR(M42F/T44L/T81M/N87G/S88Y) (Figs. 1d, 2d).

Fig. 2: Structure of AuxSen and critical steps in the engineering process.

 **figure2**

a, b, Structure of TrpR bound to the native ligand TRP (purple, Protein Data Bank (PDB) ID: 1ZT9) (**a**) and to the design-target IAA (green) (**b**). IAA is rotated by 180° in the binding pocket compared with TRP. Owing to a lack of stabilization when binding to TrpR, IAA shows conformational freedom; two alternative conformations are shown. **c,** The mutation S88Y sterically precludes the positioning of TRP (transparent purple) in the binding site while favouring the binding of IAA. **d,** Structure of the final AuxSen variant (TrpR(M42F/T44L/T81M/N87G/S88Y)) bound to IAA. The ligand is firmly packed in the enhanced hydrophobic pocket of TrpR and is anchored to R84 as well as Y88, resulting in a high affinity of AuxSen for IAA. All structures are superimposed on the C α of residues 20–60 of both chains. Red dashed lines show polar interactions between ligand and side-chain atoms. The subscript ‘bb’ labels residues that have interactions of backbone atoms with the ligand.

[Full size image](#)

In vitro, the FRET ratio of AuxSen changed by a factor of three upon treatment with 50 μ M IAA, which is within the range of cellular auxin concentrations²⁰. The signal was stable at cytosolic pH, and in the presence of reducing or oxidizing environments and all tested salt ions (Extended Data Fig. 4). The specificity of AuxSen for IAA was assessed using other indole derivatives. Among these, AuxSen showed the highest affinity for IAA; although a response was observed for other compounds, their binding affinities were reduced by around one order of magnitude (Extended Data Fig. 5). Of these indole derivatives, only IAN is present in substantial amount in plants (Extended Data Table 1). However, roots show a growth response to treatment with IAN, and modelling suggests that the IAA receptor SCFTIR¹ could bind IAN²¹. Thus, IAN is probably sequestered and is therefore unlikely to interfere with auxin sensing in the plant.

As a first step to confirm the functionality of AuxSen in vivo, we expressed a nuclear-localized version of the sensor transiently from the viral 35S promoter in cell-suspension protoplasts²², and quantified the FRET response by flow cytometry. The FRET ratio increased with the auxin concentration in the medium over four orders of magnitude, starting at 3 μ M IAA (Fig. 3a, b, Extended Data Figs. 6, 7), with the baseline FRET signal thought to reflect the endogenous level of auxin in the protoplast population. The sensitivity of the sensor could therefore be sufficient to report endogenous auxin levels.

Fig. 3: FRET ratio of AuxSen in response to auxin treatment.

 figure3

a, b, The FRET ratio obtained by flow cytometry in *Arabidopsis* protoplasts. **a**, Dose-response curve, normalized to the minimum FRET ratio (mean \pm s.e.m.; $n = 3$ biologically independent samples). **b**, Baseline fluorescence intensity ('Log_Height' (a.u.)) without exogenous IAA; the relevant area is boxed (Extended Data Figs. [6,7](#)).

c–f, Changes in the FRET ratio (colour bar) in root nuclei incubated in 10 µM IAA recorded for 1 h (**c, d**) or for 10 min (**e, f**). **c, e**, Images (DMSO, control). Scale bar, 100 µm. **d, f**, Quantitative analyses. The thick lines represent the mean (\pm s.e.m.), and the thin lines each represent independent single-seedling measurements ($n = 14$ (experimental), $n = 10$ (control) in **d**, $n = 9$ in **f**). **g, h**, Changes in the FRET ratio (colour bar) in root nuclei following the washout of IAA. **g**, Images obtained after incubation with 10 µM IAA for 1 h (top) or 10 min (bottom). Images were taken before or immediately after IAA treatment, or 10 min after the end of IAA treatment. Scale bar, 100 µm. **h**, Quantitative analysis. The thick lines represent the mean (\pm s.e.m.), and the thin lines each represent independent single-seedling measurements ($n = 5$ (10 min), $n = 3$ (1 h)). **i, j**, Change in the FRET ratio of ER-localized AuxSen in response to 100 µM IAA. **i**, Expression of ER-localized AuxSen (green; first and third rows) and FRET ratio (colour bar; second and fourth rows) of root tissue treated with IAA (top two rows) or DMSO (bottom two rows; control). Scale bar, 50 µm. **j**, Quantitative analysis of cells with high (broken lines; IAA, $n = 4$; DMSO, $n = 4$) or low (solid lines; IAA, $n = 5$; DMSO, $n = 5$) levels of AuxSen accumulation. The thick lines represent the mean (\pm s.d.), the thin lines each represent individual cells. Inset, AuxSen expression in the ER.

[Full size image](#)

We generated approximately 250 transgenic *Arabidopsis* lines bearing a dexamethasone-inducible expression system (*pBay-bar-pRPS5a-mGAL4-VP16-GR_UAS_NLS_AuxSen*) integrated as a single transgene, and selected 10 lines expressing the sensor stably in the fourth generation. The strong, ubiquitously active promoter *RPS5A* drives expression of the yeast transcription factor Gal4p, the nuclear uptake of which is induced by dexamethasone, and the binding of Gal4p to the *UAS* promoter results in the expression of the nuclear-localized auxin sensor. Strongly expressing lines were identified by mNeonGreen fluorescence in the root tips after dexamethasone induction overnight. To examine the response of AuxSen to auxin in planta, we treated seedlings with 10 µM IAA and recorded the FRET signal over time (Fig. 3c–f). After 10 min, the FRET signal in root nuclei had reached a maximum, and then remained constant for another 50 min; by contrast, treatment with the solvent DMSO did not induce any response (Fig. 3c, d). To reveal the speed of nuclear auxin accumulation, we measured the FRET signal over shorter time intervals. The maximum signal was reached within 2 min, with more than 80% achieved after only 1 min (Fig. 3e, f). The uptake of auxin from the extracellular space therefore seems to be highly efficient. We then investigated the reversibility of nuclear accumulation by washing out IAA. After 10 min of incubation of seedlings in 10 µM IAA, which gave the maximum FRET ratio, the medium was changed to DMSO (Fig. 3g, h). The FRET ratio gradually decreased, almost reaching the pre-incubation value 30 min later, which suggests slow IAA efflux from the cells (Fig. 3g, h). We repeated the experiment but extended the IAA incubation time to 1 h. This accelerated IAA efflux

such that the FRET ratio decreased to its pre-incubation level within 10 min of IAA withdrawal (Fig. 3g, h). This increased efflux of IAA correlated with two- to threefold higher expression levels of PIN efflux carriers after 1 h compared to 10 min of IAA incubation (Extended Data Fig. 8). In conclusion, this FRET-based sensor can report the dynamics of transient auxin accumulation. Although traditional reporter systems can detect the response to auxin uptake⁶, the irreversibility of reporter translation or degradation obscures the transient nature of the auxin response. Our data suggest that IAA uptake is a constitutive, fast process, whereas the efflux from the cell occurs on demand, as if auxin regulates its own export.

To explore whether auxin might accumulate in other subcellular compartments—those that cannot be accessed by reporters that are based on gene expression or protein degradation—we targeted AuxSen to the lumen of the endoplasmic reticulum (ER). The localization of PIN and PIN-LIKE putative auxin transporters in the ER membrane has led to speculation that auxin accumulation in the ER could be an ancient mechanism of auxin homeostasis²³. Incubating seedlings in 100 μM IAA for 15 min led to a strong increase in the FRET ratio compared with the DMSO control, suggesting that IAA is transported into the ER (Fig. 3i, j). Different cells within the same seedling root displayed different levels of AuxSen accumulation in the ER (Fig. 3i). Nonetheless, the change in FRET ratio upon exposure to IAA did not differ substantially between cells expressing higher or lower levels of AuxSen (Fig. 3j). In conclusion, this FRET-based auxin sensor can faithfully report auxin concentrations from a subcellular compartment such as the ER, which is inaccessible to traditional auxin-response reporters.

The spatial distribution of endogenous auxin in the seedling root has been inferred from the steady-state expression levels of auxin-response reporters such as *pDR5:GFP* or *p35S:DII-VENUS*, which display a pronounced maximum at the quiescent centre of the root meristem near the root tip⁶. Consistent with this, the FRET ratio of our auxin sensor steadily increased towards the root tip, although there was no prominent maximum (Fig. 4a, b). To assess the contribution of transport to the steady-state distribution of auxin, we incubated the seedlings in brefeldin A, which impairs auxin transport by inhibiting the polar recycling of the auxin efflux transporter PIN1²⁴. FRET ratios were increased in the tip of treated roots compared to untreated controls (Fig. 4a, b). This increase presumably resulted from ongoing IAA synthesis in the root tip while efflux was impaired, which is consistent with mass spectrometry analysis of IAA biosynthesis in cell-sorted GFP-expressing lines²⁰ and with the expression of auxin-biosynthesis genes in the root tip²⁵. We conclude that AuxSen reports the perturbation of endogenous auxin distribution, which highlights its specificity for the detection of auxin.

Fig. 4: FRET ratio of auxin sensor in response to redistribution of endogenous auxin.

 figure4

a, b, Change in the nuclear FRET ratio (colour bar) within the root tip of individual seedlings treated with brefeldin A (BFA) or DMSO. **a,** Images. **b,** Quantitative

analysis after treatment with BFA (10 µM for 10 h, magenta; $n = 10$) or DMSO (control, green; $n = 9$). The thick lines represent the mean (\pm s.e.m.), the thin lines each represent individual seedlings. **c, d**, The response of AuxSen to root gravitropism. **c**, Top, nuclear FRET ratio (colour bar) before and after turning (first and second images), and before and after turning back to the near-vertical position (third and fourth images). Bottom, Cartoons of the seedlings, with arrows indicating the direction of the gravity vector. The signal moves from the left side to the bottom and back to the left side; the colour scale indicates the relative FRET ratio. **d**, Quantified response of the sensor in individual roots. The thick lines represent the mean (\pm s.e.m.), the thin lines each represent individual roots ($n = 10$). Scale bars, 100 µm (**a, c**). **e**, Diagram of root tip. Nuclei within 100 µm above the quiescent centre (asterisk) and at least 10 µm from the midline (red line) were analysed (dark triangles). Green, right/bottom; magenta, left/top.

[Full size image](#)

The redirection of root growth in response to changes in orientation of the gravity vector is a prime example of rapid auxin signalling, which involves redistribution of endogenous auxin within the root tip²⁶. Relocalization of the auxin efflux transporter PIN3 from the (former) basal to the (former) lateral plasma membrane has been detected 2 min after turning the root to a horizontal position²⁷. However, to our knowledge, no concomitant change in auxin distribution has been reported. The earliest change that has been detected, using the auxin-responsive degradation reporter DII-VENUS, was a reduction in fluorescence in the new lower side 30 min after the onset of gravity stimulation²⁸. We equilibrated seedling roots mounted in a near-vertical position for 1 h, then changed their orientation by 135° and monitored the nuclear FRET ratios over time (Fig. 4c, d). Within 1 min, there was a distinct increase in FRET ratio on the new lower side of the root tip, corresponding to more than 80% of the maximum value reached after 5 min. We then returned the roots to the near-vertical position after 30 min, and detected substantial recovery of the pre-stimulation distribution of FRET signals within 1 min (Fig. 4c, d). Gravistimulation therefore elicits a fast and reversible response of the auxin transport system in the root tip. In conclusion, this auxin sensor reports the rapid and reversible redistribution of endogenous auxin accumulation.

Our design approach has yielded a new sensor, AuxSen, for the pervasive plant signalling molecule auxin. Starting from a tryptophan sensor, we optimized the affinity and specificity for the binding of IAA, and improved the signal intensity through choice of the FRET pair and optimization of the linkers. Our results provide a proof-of-principle that this detection system can visualize the dynamic redistribution of auxin as well as subcellular pools of auxin, which cannot be achieved with the auxin reporters that are currently in use. An example of insights into auxin dynamics that

have been made possible by the resolving power of AuxSen is the contrast between the efficient uptake and the slower, conditional efflux of auxin. Furthermore, AuxSen enables changes in auxin distribution to be distinguished from changes in auxin response, which is a prerequisite for investigating the complex regulatory network that underlies the biological effects of this major signalling molecule in plant growth and development.

Methods

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Cloning of in-vitro auxin sensors

The TrpR sensors Trp-CTY and Trp-CTYT were gifts from W. Frommer (Addgene plasmids 13533 and 13534). The fluorophores tested were also donated: Aquamarine by F. Merola (Addgene plasmid 42888), Clover and mRuby2 by K. Beam (Addgene plasmid 49089), mKO1 by K. Thorn (Addgene plasmid 44642) and mCherry by M. Bayer²⁹. eGFP was amplified from pGIIK NLS:3xEGFP³⁰. mNeonGreen, mWasabi and mTFP1 were purchased from Allele Biotechnology and Pharmaceuticals, mKate2 and TagRFP were from Evrogen.

For the initial screening, we first used the Trp-CTY sensor and mutated the residues T44 and S88 individually to all possible amino acids. To this end, we generated primers with 15–16 bp overlaps around the exchanged amino acid codons. We first mutated the amino acids sequence randomly with a degenerate primer and screened 96 clones. Variants not found were then generated by targeted mutagenesis. For example, to generate all T44 variants we first used the primers
CCTGATGCTGnnnCCAGATGAGCGCG and
CGCGCTCATCTGGnnnCAGCATCAGG, to generate the missing T44C variant we then used the primers CCTGATGCTGtgtCCAGATGAGCGCG and
CGCGCTCATCTGGacaCAGCATCAGG. The most promising candidates were introduced by targeted mutagenesis into Trp-CTYT and TrpR without fluorescent proteins, which were analysed by isothermal titration calorimetry (ITC) and structural studies. Trp-CTYT with mutations T44L and S88Y was generated in two steps: we first used the primers CCTGATGCTGctaCCAGATGAGCGCG and
CGCGCTCATCTGGtagCAGCATCAGG to generate Trp-CTYT with T44L and then GATTACGCGTGGATCTAACtac-CTGAAAGCCGCGCCC and
GGGCGCGGCTTCAGtaGTTAGATCCACGCGTAATC to generate Trp-CTYL with the mutations T44L and S88Y. To produce the recombinant proteins, we cloned the TrpR domain into the pET21 expression vector and generated the variants by

targeted mutagenesis. This procedure was repeated with the most promising variants after each round. The primer sequences are available upon request.

The final variant was then codon-optimized for *Arabidopsis* and synthesized by Thermo Fisher Scientific GENEART. To allow an easy exchange of the fluorophores we added restriction enzyme sites at the ends: BamH1 and XhoI around the first fluorophore and ApaI and HindIII around the second. All fluorescent proteins were tested in the fluorophore I–TrpR–fluorophore II–TrpR configuration. Having identified Aquamarine and mNeonGreen as the optimal pair, we introduced all final binding-pocket variants into the backbone by site-directed mutagenesis.

Ligands used for screening and testing

IAA, TRP, IAN, indole-3-carboxaldehyde, indole, indole-3-acetyl alanine, indole-3-acetyl aspartic acid, indole-3-acetamide, indole-3-ethanol, l-kynurenine, 2-oxindole-3-acetic acid, phenylalanine, picloram, tryptamine, $(\text{NH}_4)_2\text{SO}_4$, CaCl_2 , DTT, H_2O_2 , NH_4NO_3 , KNO_3 and yucasin were purchased from Sigma-Aldrich; 4-hydroxyindole-3-carbaldehyde and 5-hydroxyindole-3-carboxylic acid from Santa Cruz Biotechnology; 1-naphthaleneacetic acid, KCl and DMSO from Carl-Roth; 2,4-dichlorophenoxyacetic acid (2,4-D) from Alfa Aesar; indole-3-acetyl glucose from TRC; indole-3-butyric acid from Serva; NaCl from Merck; NPA from Supelco; and BFA from Thermo Fisher Scientific.

Mutagenesis

Trp-CTY variants were generated by site-directed mutagenesis with degenerate or specific oligonucleotides purchased from Sigma-Aldrich. Amplification was carried out using Pfu polymerase (Thermo Fisher Scientific).

More than one thousand oligonucleotides were used; the sequences and resulting vector maps are available upon request. Each variant was sequenced and screened in crude extract of sonicated bacteria for IAA binding; promising candidates were confirmed as purified proteins. The linkers were generated by site-directed mutagenesis, including linkers with 15–16 bp overlap and 3–9 degenerated nucleotides in the middle. To generate linkers shorter than the original ones, parts were deleted, whereas for longer linkers a fixed sequence was inserted into the middle to reduce the risk of generating stop codons by having too many degenerate nucleotides in the sequence.

Protein expression and purification for screening

For protein expression, bacteria were grown in the dark on plates with LB-agar supplemented with ampicillin for 3 days at room temperature. To measure crude extracts, we resuspended the bacteria in 20 mM MOPS pH 7.2, sonicated the suspension with an MS 73 probe (Bandelin) and centrifuged the sample with a tabletop centrifuge (Eppendorf). Protein extraction was performed with His Spin trap columns according to the manufacturer's instructions (GE Healthcare). We resuspended the bacteria of one Petri dish in 2 ml binding buffer and sonicated with a MS 73 probe. Buffer exchange was performed with illustra NAP-25 columns (GE Healthcare). All measurements were performed in 20 mM MOPS (Sigma-Aldrich) with an Infinite F200 plate reader (Tecan).

Ligands used for crystallization and ITC

IAA, IAN and TRP used for crystallization and ITC were dissolved in 50 mM Tris/300 mM NaCl pH 8 buffer containing 1% DMSO if necessary.

Protein purification for crystallography and ITC

After subcloning to pET21b(+), wild-type TrpR and all variants were expressed in *E. coli* BL21(DE3) and purified with a NiIMAC column and a subsequent Superdex-S75 gel-filtration column. All purification steps and measurements were based on the above 50 mM Tris/300 mM NaCl pH 8 buffer.

Crystallization, data collection and processing

Crystals of wild-type TrpR and variants with different ligands were obtained by standard vapour diffusion in sitting drop plates. The crystals were cryoprotected if needed and flash-cooled in liquid nitrogen. Data for single crystals were collected at the synchrotron beamline PXII (Swiss Light Source) at 100 K and 0.5 degree images were recorded on a Pilatus 6 M detector. Only variant TrpR(M42F/T44L/T81M/N87G/S88Y)-IAA was recorded at MX Beamlines BL14.1 at BESSY II (Helmholtz-Zentrum Berlin für Materialien und Energie). Data were indexed, integrated and scaled with the program XDS and converted with XDSConv³¹. Molecular replacement was performed with Phenix using the coordinates of wild-type TrpR (PDB ID: 1WRP³² or 1TRO³³) as search model. Model building was performed using the program Coot³⁴, and refinement was performed using Phenix³⁵. Details on crystallization conditions, data and refinement statistics for all structures are summarized in Extended Data Table 2b and Supplementary Table 1.

ITC

ITC was performed using a VP-ITC (MicroCal). The protein concentration was adjusted to 74 μ M and 730 μ M ligand solutions were prepared using the above buffer containing 1% DMSO. Measurements were performed at 20 °C with a stirring speed of 300 rpm, reference power 15 μ cal s⁻¹ and spacing of 300 s between injections. The data were analysed using the MicroCal program. Binding data were derived from sigmoidal fits based on a one-site binding model from two measurements for each variant. Heat-of-dilution baselines for the ligands alone were subtracted from the experimental data as previously described³⁶. The pH dependence of IAA binding to the variant TrpR(M42F/T44L/T81M/N87G/S88Y) (AuxSen) was measured on a NanoITC LV device with a stir rate of 300 rpm, 15 injections with 2 μ l and 300 s spacing between injections. The 170 μ l cell was overfilled with 400 μ l protein to ensure air-free filling. The protein was added at a concentration of 100 μ M and the ligand IAA at 1 mM at 20 °C. For each pH the data were recorded in the respective buffer: (i) 50 mM Tris pH 8.5, 300 mM NaCl; (ii) 50 mM Tris pH 8.0, 300 mM NaCl; (iii) 50 mM Tris pH 7.5, 300 mM NaCl; (iv) 50 mM Tris pH 7.0, 300 mM NaCl; (v) 50 mM MES pH 6.5, 300 mM NaCl; (vi) 50 mM MES pH 6.0, 300 mM NaCl; (vii) 50 mM MES pH 5.5, 300 mM NaCl; (viii) 50 mM citrate buffer pH 5.0, 300 mM NaCl; and (ix) 50 mM citrate buffer pH 4.5, 300 mM NaCl.

Test of different FRET pairs

We tested pairs of Aquamarine, mCerulean3, mTFP1 and mTurquoise2 with Clover, Ypet and mNeonGreen; Aquamarine was additionally tested with eGFP and mWasabi, and mTFP1 with TagRFP. Furthermore, we tested mNeonGreen, Clover and Ypet with TagRFP and mRuby2. mKO1 was tested with mCherry, mKate2, mNeonGreen and mWasabi. TagRFP was also tested with mTFP1, mWasabi, mKate2 and mCherry.

Constructs for in-vivo AuxSen experiments

For protoplast expression, we cloned the final version of the sensor into pJIT60²² (pJIT60-2xp35S:NLS:AuxSen).

For expression in transgenic plants, we cloned the in-vitro optimized AuxSen in constructs for conditional two-component expression, using the pBay-bar vector (a gift from M. Bayer³⁷). pRPS5a-mGAL4-VP16-GR-UAS_NLS was amplified from a pGII plasmid and inserted into pBay-bar digested with KpnI and BamHI with Gibson Assembly (In-Fusion Cloning, Takara Bio Europe SAS) according to the manufacturer 's instructions. In a second step, AuxSen and ocs terminator were inserted into *pBay-bar pRPS5a-mGAL4-VP16-GR_UAS_NLS* digested with BamHI. To obtain the individual spectra, we replaced *AuxSen* by *mNeonGreen* or *Aquamarine*. These constructs were used for transforming plants and as a template for the unmix matrix in Fiji. To generate the ER-localized auxin sensor SP:AuxSen:HDEL, we removed the

NLS from the nuclear AuxSen construct and inserted the signal peptide of an *Arabidopsis* vacuolar basic chitinase and the HDEL ER retention sequence³⁸ in frame with the coding sequence N-terminally and C-terminally, respectively.

Flow cytometry of protoplasts transiently expressing AuxSen

Protoplasts were prepared from suspension cell cultures and transfected as previously described³⁹, using 12-ml PP tubes and 12 µg of construct pJIT60-2xp35S:NLS:AuxSen per 120 µl of protoplasts (3.5×10^6 per ml) per transfection. On the next day, transfected protoplasts were pooled, filtered through 100-µm nylon mesh and split into 200-µl aliquots. Each IAA stock (in DMSO) was added 1:100 with a timing offset to account for the 5-min measurement cycle, ensuring a 1-h treatment for each sample, performed in triplicate. Cytometric analysis was set up on a Beckman Coulter MoFlo XDP (100 µm CytoNozzle, 30.5/30.0 psi, PBS sheath) to excite mNeonGreen at 488 nm (70 mW, elliptical focus) and capture peak FL1 (534/30) and shoulder FL2 (585/29) emission; Aquamarine at 405 nm excitation (100 mW, spherical focus) and capture peak FL9 (465/30) and shoulder FL10 (529/28) emission. Data were collected and processed using Summit 5.5 (Beckman Coulter). The main light-scattering gate was determined by identifying the population expressing the greatest amount of reporter. The FRET response was the ratio mean of FL10/FL9, with the auxin response moving towards FL10, directly calculated in Summit 5.5. Representative plots were drawn with FCS Express v.6.06.0033 (deNovo).

Plant material and growth conditions

Wild-type *Arabidopsis thaliana* (accession Col-0) plants were used for transformation. Plants were grown on soil at 24 °C, 65% relative humidity under long-day conditions (16-h illumination and 8-h dark period). Seeds were surface-sterilized, stratified for 2 days at 4 °C and grown on half-strength Murashige and Skoog agar plates containing 1% sucrose (0.5MS + S) (Serva). After 1 week plants were transferred to soil.

For imaging, 4-day-old seedlings were transferred to 0.5MS+S 25 µM DEX agar plates, and 16 h later to microscope slides on which they were incubated in 0.5MS+S + IAA or DMSO (control) at the specified concentrations and for the indicated periods of time. To preserve field-of-view and optimal buffer exchange we fixed the cover slip and root with double-sided adhesive tape (Tesa, type 05338, Beiersdorf). To exchange the buffer, we completely emptied the slide on a paper tissue and refilled from the side with the pipette. For BFA treatment, seedlings were transferred to 0.5MS+S 25 µM DEX plates containing either 10 µM BFA or 0.1% (v/v) DMSO (control), and mounted on microscope slides 10 h later.

Imaging

The imaging of seedling roots was performed with an LSM780 confocal laser scanning microscope, running ZEN 2.3 black SP1 as acquisition software (Zeiss) and using a Zeiss LD C-Apochromat 40×/1,1 W Korr for all experiments except the gravitropism experiments, which were recorded with a Zeiss Plan-Apochromat 20×/0.8. Spectral imaging was performed using the QUASAR detection unit on the same system: Aquamarine was excited for FRET ratio measurement at 405 nm using 5 fluorescent channels (419–455 nm, 454–491 nm, 490–526 nm, 525–562 nm and 561–598 nm); subsequently, mNeonGreen was imaged for segmentation of the regions of interest with excitation at 488 nm and detection of 3 fluorescent channels (490–526 nm, 525–562 nm and 561–598 nm). Gravitropism imaging requiring control over the direction of gravity was performed with a custom-made horizontal imaging kit that can be equipped on most inverted microscope stands for wide-field and confocal imaging. The kit consists of two pieces: a holder for the objective with a mirror for changing the direction of the optical axis of the system from a vertical direction to a horizontal direction (components from Thorlabs). The second piece is a rotatable, vertical sample holder that can be mounted into a standard multiwell plate holder (components from Fischertechnik).

We used spectral FRET⁴⁰ to be able to control for influences from sources of autofluorescence, which can be abundant in plant tissues⁴¹. Spectral FRET therefore also has the advantage that the method can be adapted by adjusting the number of acquisition channels if other sources of autofluorescence are present in different plant tissues.

All analyses were performed using a current version of Fiji⁴². First, signals generated were linearly unmixed⁴³ using J. Walter's spectral unmixing plugin (<https://imagej.nih.gov/ij/plugins/spectral-unmixing.html>). The unmixing matrix was generated with mNeonGreen, and Aquamarine as fluorophore controls and a wild-type Col-0 control for background autofluorescence. After this, the images were manually thresholded on all channels to remove unspecific signals and saturated areas, regions of interest (ROIs) for the cell nuclei were automatically generated based on the 488 nm/490–526 nm-channel data using an adaptive threshold plugin (by Q. Tseng, <https://sites.google.com/site/qingzongtseng/adaptivethreshold>) and the 'Watershed' and 'Analyze Particles' functions of ImageJ. We analysed all pixels of the image only for the ER. The FRET ratio was calculated by spectral unmixing of the channels using 'Spectral Unmix' version 1.3 (by J. Walter, <https://imagej.nih.gov/ij/plugins/spectral-unmixing.html>) and a precomputed unmixing matrix (see above) yielding the Aquamarine and mNeonGreen emission for division. Unmixed ROIs were colour-coded using the 'ROI Color Coder' plugin (BAR library, by T. Ferreira, http://imagejdocu.tudor.lu/doku.php?id=macro:roi_color_coder). In general, the colour scales were adapted for each experiment best reflecting the differences. Nuclei consisting of areas that were too small or those that had unrealistic high FRET ratios

(owing to insufficient Aquamarine signal) were omitted (reflected as black in the colour coding).

Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

Data availability

All processed data generated or analysed during this study are included in this Article and its Supplementary Information. Imaging (Herud_et_al_2021_figures_3_and_4_raw_data) and in vitro source data (Herud_et_al_2021_suppl_figures_S1_S3-S5_S7_and_table_S2a_raw_data) are available at Zenodo (<https://doi.org/10.5281/zenodo.4524537>). The data for the FACS measurements in Fig. 3a,b and Extended Data Figs. 6, 7 can be found at <https://flowrepository.org> (FR-FCM-Z3FL). Coordinates and structure factors for all reported X-ray crystallography structures have been deposited in the PDB under accession codes [6EJW](#), [6EJZ](#), [6ENI](#), [6EKP](#), [6ENN](#), [6ELB](#), [6ELF](#) and [6ELG](#). Materials (transgenic lines and plasmids) will be made available via the *Arabidopsis* Biological Resource Center (ABRC) in Columbus (OH, USA) and the Nottingham *Arabidopsis* Stock Centre (NASC), Nottingham (UK).

Code availability

The ImageJ macros used to analyse data (Herud_et_al_2021_ImageJ_analysis_macros) are available from Zenodo (<https://doi.org/10.5281/zenodo.4524537>).

References

1. 1.

Enders, T. A. & Strader, L. C. Auxin activity: Past, present, and future. *Am. J. Bot.* **102**, 180–196 (2015).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

2. 2.

Paque, S. & Weijers, D. Q&A: Auxin: the plant molecule that influences almost anything. *BMC Biol.* **14**, 67 (2016).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

3. 3.

Ulmasov, T., Murfett, J., Hagen, G. & Guilfoyle, T. J. Aux/IAA proteins repress expression of reporter genes containing natural and highly active synthetic auxin response elements. *Plant Cell* **9**, 1963–1971 (1997).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

4. 4.

Friml, J. et al. Efflux-dependent auxin gradients establish the apical-basal axis of *Arabidopsis*. *Nature* **426**, 147–153 (2003).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

5. 5.

Weijers, D. et al. Auxin triggers transient local signaling for cell specification in *Arabidopsis* embryogenesis. *Dev. Cell* **10**, 265–270 (2006).

[CAS](#) [PubMed](#) [Google Scholar](#)

6. 6.

Brunoud, G. et al. A novel sensor to map auxin response and distribution at high spatio-temporal resolution. *Nature* **482**, 103–106 (2012).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

7. 7.

Liao, C. Y. et al. Reporters for sensitive and quantitative measurement of auxin response. *Nat. Methods* **12**, 207–210 (2015).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

8. 8.

Marmorstein, R. Q., Joachimiak, A., Sprinzl, M. & Sigler, P. B. The structural basis for the interaction between l-tryptophan and the *Escherichia coli trp*

aporepressor. *J. Biol. Chem.* **262**, 4922–4927 (1987).

[CAS](#) [PubMed](#) [Google Scholar](#)

9. 9.

Retzer, K., Butt, H., Korbei, B. & Luschnig, C. The far side of auxin signaling: fundamental cellular activities and their contribution to a defined growth response in plants. *Protoplasma* **251**, 731–746 (2014).

[CAS](#) [PubMed](#) [Google Scholar](#)

10. 10.

Leyser, O. Auxin signaling. *Plant Physiol.* **176**, 465–479 (2018).

[CAS](#) [PubMed](#) [Google Scholar](#)

11. 11.

Ma, Q., Grones, P. & Robert, S. Auxin signaling: a big question to be addressed by small molecules. *J. Exp. Bot.* **69**, 313–328 (2018).

[CAS](#) [PubMed](#) [Google Scholar](#)

12. 12.

Adamowski, M. & Friml, J. PIN-dependent auxin transport: action, regulation, and evolution. *Plant Cell* **27**, 20–32 (2015).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

13. 13.

Bennett, T., Hines, G. & Leyser, O. Canalization: what the flux? *Trends Genet.* **30**, 41–48 (2014).

[CAS](#) [PubMed](#) [Google Scholar](#)

14. 14.

Naramoto, S. Polar transport in plants mediated by membrane transporters: focus on mechanisms of polar auxin transport. *Curr. Opin. Plant Biol.* **40**, 8–14 (2017).

[CAS](#) [PubMed](#) [Google Scholar](#)

15. 15.

Pařízková, B., Pernisová, M. & Novák, O. What has been seen cannot be unseen —detecting auxin *in vivo*. *Int. J. Mol. Sci.* **18**, 2736 (2017).

[PubMed Central](#) [Google Scholar](#)

16. 16.

Tsapakos, M. J., Haydock, P. V., Hermodson, M. & Somerville, R. L. Ligand-mediated conformational changes in Trp repressor protein of *Escherichia coli* probed through limited proteolysis and the use of specific antibodies. *J. Biol. Chem.* **260**, 16383–16394 (1985).

[CAS](#) [PubMed](#) [Google Scholar](#)

17. 17.

Zhang, R. G. et al. The crystal structure of *trp* aporepressor at 1.8 Å shows how binding tryptophan enhances DNA affinity. *Nature* **327**, 591–597 (1987).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

18. 18.

Kaper, T. et al. Nanosensor detection of an immunoregulatory tryptophan influx/kynurenone efflux cycle. *PLoS Biol.* **5**, e257 (2007).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

19. 19.

Hamers, D., van Voorst Vader, L., Borst, J. W. & Goedhart, J. Development of FRET biosensors for mammalian and plant systems. *Protoplasma* **251**, 333–347 (2014).

[CAS](#) [PubMed](#) [Google Scholar](#)

20. 20.

Petersson, S. V. et al. An auxin gradient and maximum in the *Arabidopsis* root apex shown by high-resolution cell-specific analysis of IAA distribution and synthesis. *Plant Cell* **21**, 1659–1668 (2009).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

21. 21.

Katz, E. et al. The glucosinolate breakdown product indole-3-carbinol acts as an auxin antagonist in roots of *Arabidopsis thaliana*. *Plant J.* **82**, 547–555 (2015).

[CAS](#) [PubMed](#) [Google Scholar](#)

22. 22.

Herud, O., Weijers, D., Lau, S. & Jürgens, G. Auxin responsiveness of the MONOPTEROS-BODENLOS module in primary root initiation critically depends on the nuclear import kinetics of the Aux/IAA inhibitor BODENLOS. *Plant J.* **85**, 269–277 (2016).

[CAS](#) [PubMed](#) [Google Scholar](#)

23. 23.

Barbez, E. & Kleine-Vehn, J. *Divide Et Impera*—cellular auxin compartmentalization. *Curr. Opin. Plant Biol.* **16**, 78–84 (2013).

[CAS](#) [PubMed](#) [Google Scholar](#)

24. 24.

Geldner, N. et al. The *Arabidopsis* GNOM ARF-GEF mediates endosomal recycling, auxin transport, and auxin-dependent plant growth. *Cell* **112**, 219–230 (2003).

[CAS](#) [PubMed](#) [Google Scholar](#)

25. 25.

Brumos, J. et al. Local auxin biosynthesis is a key regulator of plant development. *Dev. Cell* **47**, 306–318.e5 (2018).

[CAS](#) [PubMed](#) [Google Scholar](#)

26. 26.

Su, S.-H., Gibbs, N. M., Jancewicz, A. L. & Masson, P. H. Molecular mechanisms of root gravitropism. *Curr. Biol.* **27**, R964–R972 (2017).

[CAS](#) [PubMed](#) [Google Scholar](#)

27. 27.

Friml, J., Wiśniewska, J., Benková, E., Mendgen, K. & Palme, K. Lateral relocation of auxin efflux regulator PIN3 mediates tropism in *Arabidopsis*. *Nature* **415**, 806–809 (2002).

[ADS](#) [PubMed](#) [Google Scholar](#)

28. 28.

Band, L. R. et al. Root gravitropism is regulated by a transient lateral auxin gradient controlled by a tipping-point mechanism. *Proc. Natl Acad. Sci. USA* **109**, 4668–4673 (2012).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

29. 29.

Musielak, T. J., Slane, D., Liebig, C. & Bayer, M. A versatile optical clearing protocol for deep tissue imaging of fluorescent proteins in *Arabidopsis thaliana*. *PLoS ONE* **11**, e0161107 (2016).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

30. 30.

Takada, S. & Jürgens, G. Transcriptional regulation of epidermal cell fate in the *Arabidopsis* embryo. *Development* **134**, 1141–1150 (2007).

[CAS](#) [PubMed](#) [Google Scholar](#)

31. 31.

Kabsch, W. XDS. *Acta Crystallogr. D* **66**, 125–132 (2010).

[CAS](#) [Google Scholar](#)

32. 32.

Lawson, C. L. et al. Flexibility of the DNA-binding domains of *trp* repressor. *Proteins* **3**, 18–31 (1988).

[CAS](#) [PubMed](#) [Google Scholar](#)

33. 33.

Otwinowski, Z. et al. Crystal structure of *trp* repressor/operator complex at atomic resolution. *Nature* **335**, 321–329 (1988).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

34. 34.

Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).

[Article](#) [Google Scholar](#)

35. 35.

Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).

[CAS](#) [Google Scholar](#)

36. 36.

Machius, M. et al. Structural and biochemical basis for polyamine binding to the Tp0655 lipoprotein of *Treponema pallidum*: putative role for Tp0655 (TpPotD) as a polyamine receptor. *J. Mol. Biol.* **373**, 681–694 (2007).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

37. 37.

Wang, K. et al. Independent parental contributions initiate zygote polarization in *Arabidopsis thaliana*. Preprint at <https://doi.org/10.1101/2020.12.02.407874> (2020).

38. 38.

Haseloff, J., Siemering, K. R., Prasher, D. C. & Hodge, S. Removal of a cryptic intron and subcellular localization of green fluorescent protein are required to mark transgenic *Arabidopsis* plants brightly. *Proc. Natl Acad. Sci. USA* **94**, 2122–2127 (1997).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

39. 39.

Mehlhorn, D. G., Wallmeroth, N., Berendzen, K. W. & Grefen, C. 2in1 vectors improve in planta BiFC and FRET analyses. *Methods Mol. Biol.* **1691**, 139–158 (2018).

[CAS](#) [PubMed](#) [Google Scholar](#)

40. 40.

Piston, D. W. & Kremers, G. J. Fluorescent protein FRET: the good, the bad and the ugly. *Trends Biochem. Sci.* **32**, 407–414 (2007).

[CAS](#) [PubMed](#) [Google Scholar](#)

41. 41.

Donaldson, L. Autofluorescence in plants. *Molecules* **25**, 2393 (2020).

[CAS](#) [PubMed Central](#) [Google Scholar](#)

42. 42.

Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).

[CAS](#) [Google Scholar](#)

43. 43.

Neher, R. & Neher, E. Optimizing imaging parameters for the separation of multiple labels in a fluorescence image. *J. Microsc.* **213**, 46–62 (2004).

[MathSciNet](#) [CAS](#) [PubMed](#) [Google Scholar](#)

44. 44.

Lambert, T. J. FPbase: a community-editable fluorescent protein database. *Nat. Methods* **16**, 277–278 (2019).

[CAS](#) [PubMed](#) [Google Scholar](#)

45. 45.

Novák, O. et al. Tissue-specific profiling of the *Arabidopsis thaliana* auxin metabolome. *Plant J.* **72**, 523–536 (2012).

[ADS](#) [PubMed](#) [Google Scholar](#)

46. 46.

Tam, Y. Y., Epstein, E. & Normanly, J. Characterization of auxin conjugates in *Arabidopsis*. Low steady-state levels of indole-3-acetyl-aspartate, indole-3-acetyl-glutamate, and indole-3-acetyl-glucose. *Plant Physiol.* **123**, 589–596 (2000).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

47. 47.

Mashiguchi, K. et al. The main auxin biosynthesis pathway in *Arabidopsis*. *Proc. Natl Acad. Sci. USA* **108**, 18512–18517 (2011).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

48. 48.

Normanly, J., Cohen, J. D. & Fink, G. R. *Arabidopsis thaliana* auxotrophs reveal a tryptophan-independent biosynthetic pathway for indole-3-acetic acid. *Proc. Natl Acad. Sci. USA* **90**, 10355–10359 (1993).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

49. 49.

Böttcher, C. et al. The biosynthetic pathway of indole-3-carbaldehyde and indole-3-carboxylic acid derivatives in *Arabidopsis*. *Plant Physiol.* **165**, 841–853 (2014).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

[Download references](#)

Acknowledgements

We thank Addgene for distributing plasmids donated by W. Frommer, F. Merola, K. Beam and K. Thorn; O. Griesbeck, H. Schäfer, A. Holz, A.-C. Barragan-Lopez and C. Brancato for technical assistance; M. Perez Ortega for the initial steps towards the macro script; M. Belkacemi for sequencing of the variants; and M. Bayer for

discussions and suggestions. C.F. thanks R. Palmisano and M. Küppers for help with gravitropism kit design. O.H.-S., M.K. and C.F. thank the staff of the Light Microscopy Facility at MPI for Developmental Biology for their support. G.J. thanks R. Hertel for discussions on auxin-related matters. We further thank the beamline staff at the Swiss Light Source and at BESSY for support. B.H. acknowledges the financial support and the allocation of synchrotron beam time by HZB as well as funding by the Deutsche Forschungsgemeinschaft (grant HO 4022/2-3). This study was supported by the Max Planck Society.

Funding

Open access funding provided by Max Planck Society.

Author information

Author notes

1. These authors contributed equally: Ole Herud-Sikimić, Andre C. Stiel, Martina Kolb

Affiliations

1. Max Planck Institute for Developmental Biology, Tübingen, Germany

Ole Herud-Sikimić, Andre C. Stiel, Martina Kolb, Sooruban Shanmugaratnam, Christian Feldhaus, Birte Höcker & Gerd Jürgens

2. Institute for Biological and Medical Imaging, Helmholtz Zentrum Munich, German Research Center for Environmental Health, Munich, Germany

Andre C. Stiel

3. Department of Biochemistry, University of Bayreuth, Bayreuth, Germany

Sooruban Shanmugaratnam & Birte Höcker

4. Centre for Plant Molecular Biology, University of Tübingen, Tübingen, Germany

Kenneth W. Berendzen

Authors

1. Ole Herud-Sikimić
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Andre C. Stiel
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Martina Kolb
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Sooruban Shanmugaratnam
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Kenneth W. Berendzen
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Christian Feldhaus
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Birte Höcker
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Gerd Jürgens
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Contributions

O.H.-S., A.C.S., B.H. and G.J. conceived the idea and designed the experiments. O.H.-S., A.C.S., M.K., S.S. and K.W.B. performed the experiments. O.H.-S. performed

cloning and mutagenesis of auxin sensor constructs for in vitro analysis, A.C.S. and S.S. carried out biochemistry and structure analysis of auxin sensor variants, and O.H.-S., A.C.S. and S.S. performed ligand-binding studies. M.K. cloned the auxin sensor constructs for in-vivo analysis, generated transgenic lines, analysed AuxSen expression, performed microscopic imaging analyses and determined auxin transporter expression. K.W.B. performed the transfection and flow cytometry analysis of the auxin sensor in protoplasts. C.F. developed the imaging workflow and constructed the microscope adapter for gravitropism-related measurements. O.H.-S., A.C.S., B.H. and G.J. wrote the manuscript with input from all authors.

Corresponding authors

Correspondence to [Birte Höcker](#) or [Gerd Jürgens](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Richard Napier 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 FRET response of several binding-domain variants to increasing substrate concentration.](#)

a, Wild-type TrpR. **b–q**, Engineered TrpR variants. Each mark indicates a single measurement. Circle, IAA; rhombus, TRP; square, IAN.

[Extended Data Fig. 2 Details observed in the crystal structures.](#)

a, Structure of IAA in the binding pocket of TrpR(M42F/T44L/T81I/S88Y). **b**, Structure of IAN bound to the same variant as in **a**. **c**, Overlay of IAA in

TrpR(T44L/T81M/S88Y) (magenta) and TrpR(T44L/T81M/N87G/S88Y) (green). **d**, Structural overview of variants TrpR(T44L/T81M/N87G/S88Y) (green), TrpR(T44L/T81M/S88Y) (magenta), and TrpR(M42F/T44L/T81M/N87G/S88Y) (AuxSen, gold). It is apparent that AuxSen differs from the two intermediate structures regarding the overall arrangement of the helices. **e–g**, The structure of TrpR(S88Y/T44L) and all variant structures based on it show a slight relocation of the backbone of residues 70–90. This is probably due to the fact that all structures based on this variant crystallize in the orthorhombic space group $P2_12_12_1$ (**f**) as opposed to the tetragonal space group $P4_3$ found for TrpR–IAA and TrpR(S88Y)–IAA (**e**). Both geometries have been found in earlier crystal structures of TrpR, for example PDB 1ZT9 (tetragonal) and 2OZ9 (orthorhombic). It seems that the introduction of the T44L mutation strongly favours crystallization in the orthorhombic geometry. In the $P2_12_12_1$ space group, crystals form more extensive crystal contacts. The structure overlay (**g**) shows how several residues are displaced (residues that have symmetry mates within 3 Å are shown in red, and the ligand IAA is shown in green). However, interactions and positioning of the ligands are maintained. Nonetheless, we only compare backbone coordinates between variants of the same space group, to exclude misinterpretations due to crystal contacts.

Extended Data Fig. 3 Parameters tested for potential influence on the change of FRET ratio.

a, Change in FRET ratio upon IAA treatment plotted against the dissociation constant (K_d) of the same variant as determined by ITC. **b**, FRET ratio changes do not correlate with the Förster distance. Blue–yellow pairs are marked in green, yellow–red pairs are in orange. Blue–yellow pairs, in general, show a higher FRET ratio change upon IAA treatment, but a similar range of Förster distances as the yellow–red ones. **c**, FRET ratio changes (in per cent) of several variants tested with two different fluorophore pairs. Variants showing a strong response with one fluorophore pair usually also show a strong response with another pair (correlation coefficient = 0.6). **d–j**, Effects of mutations in linkers. **d**, Structure of the construct. The IAA-binding TrpR variants were cloned as tandem repeats into the construct containing donor and acceptor fluorophores, analogous to ref. ¹⁸. The positional effect of the fluorescent proteins probably stems from slight rearrangements of the overall backbone in both TrpR subunits. Predominantly, this involves helix E of the reading-head motif, which mediates the DNA interaction in the natural function of TrpR. Because helix E is towards the C-terminal end of the chain, it is thought that fluorescent proteins positioned at this end will experience a larger positional relocation and thus show a more dynamic range of the FRET signal. **e–g**, First-round linker mutations. All three linkers were mutated, but no pattern for the optimal linker length could be determined. One linker II variant was chosen for further mutations. **h, i**, Second-round linker

mutations. Linkers I and III were mutated in the variant obtained in the first round, with no changes in the optimized linker II. **j**, Third-round linker mutations. Linker I was further mutated in the variant containing mutations in linkers II and III. The linker length axis indicates the number of amino acid residues.

Extended Data Fig. 4 pH, salt and redox sensitivity of AuxSen.

a, The FRET ratio is slightly affected by changes in the pH, but fully functional in the range of pH values within the plant cell. Red, pH 6.0; blue, pH 6.5; green, pH 7.0; violet, pH 7.5; cyan, pH 8.0. **b**, The FRET ratio is not strongly affected by salts and changes in the redox potential. Black, control; red, 1 mM $(\text{NH}_4)_2\text{SO}_4$; blue, 1 mM CaCl_2 ; green, 10 mM NH_4NO_3 ; violet, 10 mM DTT; cyan, 10 mM H_2O_2 ; yellow, 10 mM KCl; orange squares, 10 mM KNO_3 ; orange diamonds, 10 mM NaCl. **c**, ITC data of AuxSen binding to IAA, measured at different pH values. Data are mean \pm s.d., derived from 3 technical replicates.

Extended Data Fig. 5 Affinities of AuxSen for auxin-related compounds.

a–j, Compounds with weak affinities; **k–v**, compounds with no affinity. The change in FRET ratio (y axis) is plotted against increasing concentrations of the individual compounds in μM (x axis). Each mark indicates a single measurement.

Extended Data Fig. 6 FRET ratio of AuxSen quantified in protoplasts.

Cell-suspension-culture protoplasts were transfected with *pJIT60-2xp35SS:NLS:AuxSen* and the FRET response was captured by flow cytometry as the ratio of the emission between the Aquamarine peak intensity (FL9) and mNeonGreen peak intensity (FL10) as ‘Log_Height’ (a.u.). **a**, Normalized excitation and emission spectra for mNeonGreen and Aquamarine. The 405-nm laser line is shown as a vertical bar and bandpass filters are shown as open boxes, as rendered by FPBase⁴⁴. As FRET occurs, the emission output is shifted from the FL9 to the FL10 bandpass. **b**, Baseline level of AuxSen FRET response, without exogenous IAA. **c**, Maximal response level of AuxSen with 10 mM IAA.

Extended Data Fig. 7 Demonstration of relevant population regions and Aquamarine, mNeonGreen and AuxSen fluorescent emission in cell-suspension-culture protoplasts.

Bivariate plots from left to right are as follows: forward versus side-scatter log area ungated; emission peak to shoulder FL1 (534/30) versus FL2 (585/29); forward versus side-scatter log area with cells of interest marked and gated for **b–d**; emission peak-to-

shoulder FL9 (465/30) versus FL10 (529/28). Arrows indicate ‘gating’, meaning that the following plot is restricted to those data points that fall within that particular window. **a**, Cells only transfected with water. **b**, Cells transfected with *pJIT60-2xp35SS:NLS:Aquamarine*. Cells expressing Aquamarine were used to determine which scattering population produced the fluorescent protein. This gate is followed for mNeonGreen emission. **c**, Cells transfected with *pJIT60-2xp35SS:NLS:mNeonGreen*. Cells expressing mNeonGreen were used to determine which scattering population produced the fluorescent protein. This gate is followed for Aquamarine emission. **d**, Cells transfected with *pJIT60-2xp35SS:NLS:AuxSen*. Cells expressing AuxSen identified by their mNeonGreen emission then restricted to the ‘cells of interest’ scattering population that produced the greatest amount of protein. The FRET-response region was then made to encompass the entire range of possible fluorescence, including the shift in the FL9 to the FL10 bandpasses. See also Supplementary Table 2. Flow cytometry basic gate statistics and FRET-ratios based on (‘FL10-Log_Height’/‘FL9-Log_Height’) ratio are plotted against time for the final gate FRET response.

Extended Data Fig. 8 Quantitative analysis of *PIN* mRNA accumulation in seedlings treated with 10 µM IAA.

Seedlings were kept in MS medium (control; 0 min, purple) or transferred to a solution containing 10 µM IAA. RNA was extracted from 6 seedlings each after 10 min (blue) or 60 min (orange) and subjected to analysis by quantitative PCR with reverse transcription. Data are mean ± s.d. derived from 3 technical replicates, and black dots represent individual values.

Extended Data Table 1 In vivo concentration of indole derivatives as reported in plant extracts. See refs. [21,45,46,47,48,49](#)

[Full size table](#)

Extended Data Table 2 Dissociation constants and overview of crystallographic data for TrpR and selected variants

[Full size table](#)

Supplementary information

Supplementary Tables

This file contains Supplementary Table 1 (Crystallographic data) and Supplementary Table 2 (Flow cytometry statistics).

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



Check for
updates

Cite this article

Herud-Sikimić, O., Stiel, A.C., Kolb, M. *et al.* A biosensor for the direct visualization of auxin. *Nature* **592**, 768–772 (2021). <https://doi.org/10.1038/s41586-021-03425-2>

Download citation

- Received: 02 October 2018
- Accepted: 05 March 2021
- Published: 07 April 2021
- Issue Date: 29 April 2021
- DOI: <https://doi.org/10.1038/s41586-021-03425-2>

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Download PDF](#)

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03425-2>

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

Structural and biochemical mechanisms of NLRP1 inhibition by DPP9

[Download PDF](#)

- Article
- Open Access
- [Published: 17 March 2021](#)

Structural and biochemical mechanisms of NLRP1 inhibition by DPP9

- [Menghang Huang^{1,na1}](#),
- [Xiaoxiao Zhang ORCID: orcid.org/0000-0003-3762-2287^{1,na1}](#),
- [Gee Ann Toh ORCID: orcid.org/0000-0002-7417-8376²](#),
- [Qin Gong³](#),
- [Jia Wang¹](#),
- [Zhifu Han¹](#),
- [Bin Wu ORCID: orcid.org/0000-0002-0883-8006^{3,4}](#),
- [Franklin Zhong ORCID: orcid.org/0000-0002-0516-6021^{2,5}](#) &
- [Jijie Chai ORCID: orcid.org/0000-0001-7591-3873^{1,6,7}](#)

[Nature](#) volume 592, pages773–777(2021) [Cite this article](#)

- 8895 Accesses
- 2 Citations
- 29 Altmetric
- [Metrics details](#)

Subjects

- [Cryoelectron microscopy](#)
- [NOD-like receptors](#)

Abstract

Nucleotide-binding domain, leucine-rich repeat receptors (NLRs) mediate innate immunity by forming inflammasomes. Activation of the NLR protein NLRP1 requires autocleavage within its function-to-find domain (FIIND)^{1,2,3,4,5,6,7}. In resting cells, the dipeptidyl peptidases DPP8 and DPP9 interact with the FIIND of NLRP1 and suppress spontaneous NLRP1 activation^{8,9}; however, the mechanisms through which this occurs remain unknown. Here we present structural and biochemical evidence that full-length rat NLRP1 (rNLRP1) and rat DPP9 (rDPP9) form a 2:1 complex that contains an autoinhibited rNLRP1 molecule and an active UPA–CARD fragment of rNLRP1. The ZU5 domain is required not only for autoinhibition of rNLRP1 but also for assembly of the 2:1 complex. Formation of the complex prevents UPA-mediated higher-order oligomerization of UPA–CARD fragments and strengthens ZU5-mediated NLRP1 autoinhibition. Structure-guided biochemical and functional assays show that both NLRP1 binding and enzymatic activity are required for DPP9 to suppress NLRP1 in human cells. Together, our data reveal the mechanism of DPP9-mediated inhibition of NLRP1 and shed light on the activation of the NLRP1 inflammasome.

[Download PDF](#)

Main

In the mammalian innate immune system, the detection of pathogen-derived or host-derived signals by NLRs induces their oligomerization, forming multiprotein complexes called inflammasomes that mediate inflammatory cell death and cytokine secretion¹⁰. NLRs generally consist of an N-terminal domain, a central nucleotide-binding and oligomerization domain (NOD) and a C-terminal leucine-rich repeat (LRR) domain. NLRP1, like CARD8 (refs. ^{4,7}), contains an unusual domain known as FIIND¹¹ (Fig. ^{1a}), which contains the subdomains ZU5 (found in the tight-junction protein ZO-1 and the netrin receptor UNC5) and UPA (conserved in UNC5, the death-domain-containing protein PIDD and proteins of the ankyrin family). Autoproteolysis between these two subdomains is a prerequisite for the activation of NLRP1 (refs. ^{4,7}) and of CARD8 (ref. ¹²). *Bacillus anthracis* lethal factor is the best-characterized pathogen-derived trigger for the activation of rodent NLRP1 (refs. ^{5,13}). Lethal factor cleaves mouse NLRP1B close to its N terminus and induces proteasomal degradation of the entire N-terminal NOD–LRR–ZU5 fragment via the N-end rule pathway^{1,2,3}. This liberates the active UPA–CARD fragment that rapidly oligomerizes to engage downstream inflammasome effectors such as apoptosis-associated speck-like protein containing a CARD (ASC) and pro-caspase-1 (refs. ^{1,3}). This unique mechanism involving ‘functional degradation’ is conserved in the activation of human NLRP1 (hNLRP1) by the 3C proteases of enteroviruses¹⁴ and in the activation of CARD8 by HIV-1 protease¹⁵.

Fig. 1: Crystal structure of the FIIND of rNLRP1.

 figure1



a, Schematic of domain structures of hNLRP1, rNLRP1 and CARD8. CARD, caspase activation and recruitment domain; PYD, pyrin domain. **b**, Crystal structure of the rNLRP1 FIIND. The ZU5 and UPA subdomains are shown in pink and blue, respectively. The catalytic residues of the FIIND are labelled and shown in stick representation. **c**, A close-up view of the catalytic site of the FIIND. **d**, Mutation of the catalytic residue H942 abolishes autoproteolysis of the rNLRP1 FIIND. Wild-type and H942 mutant rNLRP1 FIIND proteins were purified from insect cells and visualized by SDS-PAGE followed by Coomassie blue staining. See Supplementary Fig. 1 for gel raw data.

[Full size image](#)

DPP8 and DPP9 are related intracellular prolyl peptidases that are implicated in immune regulation and in other viral cellular processes¹⁶. They are endogenous inhibitors of the NLRP1 inflammasome in humans^{8,9,12} and in rodents^{6,8,12,17,18}. Notably, the FIIND of hNLRP1 is necessary and sufficient for interaction with human DPP9 (hDPP9) (ref. 9). Furthermore, inhibitors of class IV DPPs—such as valine

boroproline (VbP)—or knockout of DPP8 and DPP9 specifically activate NLRP1 and/or CARD8 (refs. [6](#),[8](#),[9](#),[12](#)). Like lethal factor, VbP also induces the proteasome-mediated degradation of the NLRP1B N-terminal fragment, but this induction is independent of N-degron recognition. VbP and lethal factor therefore trigger rodent NLRP1 activation through two distinct, independent pathways. At present, the mechanisms that underlie the DPP9-mediated inhibition of NLRP1 remain unknown.

Crystal structure of the autoinhibited FIIND

A crystal structure of the rNLRP1 FIIND of rNLRP1 shows that the purified protein is autocleaved at the predicted position between F968 and S969 (Extended Data Table [1](#), Extended Data Fig. [1a](#)). This is further confirmed by SDS–PAGE analysis of the crystals (Extended Data Fig. [1b](#)). The structure of rNLRP1 FIIND resembles that of the autoinhibited netrin receptor UNC5b^{[19](#)} (Extended Data Fig. [1c](#)). Inter-domain interaction between ZU5 and UPA largely involves the first β -strand (β 13) of UPA, which forms two anti- β sheets with ZU5 (Fig. [1b](#)); this explains how the ZU5 domain can block the release of the active UPA–CARD fragment^{[4](#),[7](#)}.

The catalytically essential FS motif^{[20](#)} is conserved in NLRP1 homologues and in CARD8 (Extended Data Fig. [1d](#)). In the structure of rNLRP1, F968 from this motif points into a hydrophobic pocket (Fig. [1c](#)), as is observed for the corresponding phenylalanine residue in the FIIND-containing proteins NUP98 (ref. [21](#)) and PIDD^{[22](#)}. H942 of rNLRP1, which is highly conserved in NLRP1 proteins and in CARD8 (Extended Data Fig. [1d](#)), is located adjacent to S969 from the FS motif (Fig. [1c](#)). Mutation of H942 of rNLRP1 resulted in a complete loss of autocleavage (Fig. [1d](#)), as was previously observed after mutation of the corresponding residues H270 of CARD8 or H1186 of hNLRP1 (ref. [20](#)). Collectively, these data suggest that H942 of rNLRP1 is a catalytic residue.

Architecture of the 2:1 rNLRP1–rDPP9 complex

Gel-filtration experiments confirmed the formation of a stable complex between full-length rNLRP1 and rDPP9 proteins purified from insect cells (Extended Data Fig. [2a](#))—consistent with the results of previous studies^{[3](#),[8](#),[9](#)}. A similar result was also obtained using hDPP9 and CARD8 (Extended Data Fig. [2b](#)). After purification by gel filtration, the rNLRP1–rDPP9 complex was analysed by cryo-electron microscopy (cryo-EM) (Extended Data Fig. [3](#)). Two-dimensional class averages showed that rDPP9 formed dimers, but only one subunit was bound to NLRP1 in most of the particles (Extended Data Fig. [3b](#)). We used the particles with one rDPP9 subunit bound by rNLRP1 for further cryo-EM analysis. After 3D classification, a subset of 182,116 particles was

used for final 3D reconstruction, generating a map with a global resolution of 3.18 Å (Fig. 2a, Extended Data Fig. 3c–e, Extended Data Table 2).

Fig. 2: Assembly mechanism of the 2:1 rNLRP1–rDPP9 complex.

 [figure2](#)

a, Top, the final cryo-EM density of the rNLRP1–rDPP9 complex at 3.18 Å. Colour codes for domain structures are indicated. Bottom, model of the 2:1 rNLRP1–rDPP9 complex. The three interfaces that mediate the rNLRP1–rDPP9 interaction are shown in coloured boxes. Red, ZU5-binding site; blue, UPA-binding site; yellow, UPA–UPA-binding site. **b**, Detailed interactions between ZU5 and rDPP9 at the ZU5-binding site within the red-framed region in **a**. Hydrogen-bonding interactions are indicated by red dashed lines. **c**, Detailed interactions between UPA and rDPP9 at the UPA-binding site within the blue-framed region in **a**. **d**, 3D reconstruction of the rNLRP1 FIIND–CARD(S969A)–rDPP9 complex. Highlighted within the open ellipse is the vacant UPA-binding site. **e**, The ZU5 subdomain is displaced after interaction of the second rNLRP1 with rDPP9. The preformed rNLRP1 FIIND(S969A)–rDPP9 complex was incubated with fully autoprocessed rNLRP1 FIIND and analysed by gel-filtration experiments. Protein fractions from the gel-filtration assay were visualized by SDS–PAGE followed by Coomassie-blue staining. See Supplementary Fig. 1 for gel raw data.

[Full size image](#)

rDPP9 forms a homodimer (Fig. 2a) that is nearly identical to the dimeric hDPP9 (ref. 23). Several flexible loop regions were not observed in the reported apo-hDPP9 structure, but their equivalents are well defined in the structure of the rNLRP1-bound rDPP9 (Extended Data Fig. 4a). Unexpectedly, the liganded rDPP9 subunit is bound by two rNLRP1 molecules (Fig. 2a), which we hereafter term the 2:1 rNLRP1–rDPP9 complex. The first rNLRP1 molecule contains the complete FIIND, which has a conformation nearly identical to that observed in the crystal structure of the free

FIIND (Extended Data Fig. 4b); the remaining domains of this rNLRP1 molecule are not discernible in the cryo-EM density. In the second rNLRP1 molecule, only the UPA domain is well defined.

Three surfaces mediate the formation of the 2:1 rNLRP1–rDPP9 complex (Fig. 2a). The first surface is mediated by ZU5 of rNLRP1, which binds to one lateral side of the β -propeller domain of rDPP9 (termed the ZU5-binding site of rDPP9) (Fig. 2a,b, red box). The second interface is formed by the deep insertion of an N-terminal loop (N-loop) from the UPA domain of the second rNLRP1 molecule into the rDPP9 substrate-binding channel (termed the UPA-binding site) (Fig. 2a,c, blue box). Notably, the UPA N-loop forms a β -sheet with ZU5 in the autoinhibited FIIND (Fig. 1b). Additionally, the homodimerization interface of the two UPA domains also contributes to formation of the 2:1 complex (termed the UPA dimerization site) (Fig. 2a, yellow box).

Specific interaction between rNLRP1 and rDPP9

The interactions within the ZU5-interacting site consist of extensive contacts between β 10, the α 1– β 6 loop of the rNLRP1 FIIND ZU5 subdomain, and the β -propeller domain of rDPP9 (Fig. 2b). β 10 of FIIND forms three main-chain hydrogen bonds with the N-terminal portion of a long loop of rDPP9, forming an anti- β -sheet-like structure (Fig. 2b). The rDPP9 loop also makes hydrophobic contacts with the middle part of the α 1– β 6 loop of FIIND. The amino acids of this ZU5-interacting site are conserved in DPP8 but not in DPP4 (Extended Data Fig. 4c), which explains why DPP4 fails to inhibit NLRP1B^{8,24}. In addition, the short α -helix in the α 1– β 6 loop of FIIND packs against the other two loops of rDPP9. The substrate-binding groove of rDPP9 is completely blocked by the UPA N-loop (Fig. 2c). Consistent with previous studies^{8,9,25}, N-terminal sequencing indicated that the seven non-structured, N-terminal residues of the UPA N-loop were not cleaved by DPP9 (Extended Data Fig. 5a).

rNLRP1(S969A) forms a 1:1 complex with rDPP9

The catalytically inactive mutant rDPP9(S729A)—in which the serine residue at position 729 is mutated to alanine—formed a stable complex with the rNLRP1 FIIND–CARD fragment, as determined by gel-filtration experiments (Extended Data Fig. 5b). rNLRP1 FIIND in the first bound position adopts a nearly identical conformation to that of free rNLRP1 FIIND in the crystal structure (Extended Data Fig. 4b), which suggests that a non-autocleavable rNLRP1 can interact with rDPP9. Indeed, the auto-cleavage mutant fragment rNLRP1 FIIND–CARD(S969A) retained rDPP9-binding activity (Extended Data Fig. 5c), consistent with previous

observations^{8,12}. Notably, the cryo-EM structure of rDPP9 in complex with rNLRP1 FIIND–CARD(S969A) (Extended Data Fig. 6) is nearly identical to that of the wild-type complex encompassing the first rNLRP1 molecule, except that no clear density was found in the second rNLRP1-binding position (Fig. 2d). This result indicates that rNLRP1 FIIND–CARD(S969A) can interact with the ZU5-binding site but not with the UPA-binding site, establishing that autoproteolysis is required for rNLRP1 to fit into the second position to form the 2:1 complex. This could also explain why autoproteolysis-deficient NLRP1 mutants retain DPP9-binding activity, but to a lesser extent^{8,9}.

The ZU5 domain of rNLRP1 was not observed in the second position (Fig. 2a), because when rNLRP1 docks into this position this domain becomes either flexible or dissociated. To differentiate between these two possibilities, we tested whether an autocleaved rNLRP1 FIIND fits into the second position in a preformed 1:1 rNLRP1 FIIND(S969A)–rDPP9 complex. Gel filtration showed that the autocleaved FIIND did interact with rNLRP1 FIIND(S969A)–rDPP9 (Fig. 2e). In the resulting complex, the ZU5 and UPA subdomains of rNLRP1 became sub-stoichiometric, which indicates that ZU5 dissociates from UPA after rNLRP1 binding in the second position.

NLRP1 blocking by DPP9 enzyme and binding activity

Point mutations at S900 of rNLRP1 and L101 of rDPP9, which were predicted to disrupt the ZU5-binding site, were found to abrogate or markedly compromise the interaction between rNLRP1 and rDPP9 (Fig. 3a,b, Extended Data Fig. 7a). This suggests that the first binding position is required in order for the second rNLRP1 to bind rDPP9. Similarly, mutation of L131 of hDPP9—the residue corresponding to L101 of rDPP9—resulted in no interaction with hNLRP1 at all in 293T cells (Fig. 3c). Notably, the corresponding mutation (L131E) in hDPP9 also resulted in a loss of interaction with CARD8 (Extended Data Fig. 7b). This suggests that the mechanism of ZU5-mediated interaction with DPP9 might be conserved among rNLRP1, hNLRP1 and CARD8.

Fig. 3: Both protease activity and FIIND binding are important for the inhibition of NLRP1 by DPP9.

 **figure3**

a, rNLRP1 FIIND with an N-terminal GST tag (GST–NLRP1 FIIND) was used to pull down non-tagged rDPP9 in vitro. **b**, Lysates from 293T cells transfected with the indicated Flag-tagged rNLRP1 and Myc-tagged rDPP9 proteins were subjected to anti-Flag immunoprecipitation (IP). **c**, As in **b** using Flag–hDPP9 and haemagglutinin

(HA)-tagged hNLRP1. **d**, ASC–GFP DPP8/DPP9 double-knockout 293T cells were transfected with the indicated constructs and treated with and without VbP for 24 h. Bar graphs represent data from biological triplicates. Two-way ANOVA, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$, NS, not significant. **e**, Immortalized N/TERT-1 keratinocytes stably transduced with Tet-ON-3×Flag DPP9 lentiviruses were treated with either doxycycline (DOX) or VbP for 24 h. Lysates and the culture media were analysed by immunoblotting. **f**, Images of N/TERT-1 keratinocytes treated as in **e**. Pyroptotic cells are indicated by yellow arrows. Scale bar, 150 μ m. See Supplementary Figs. [1](#) and [3](#) for gel raw data.

[Source data](#)

[Full size image](#)

To determine whether the binding of NLRP1 is functionally important for its inhibition by DPP9, we imaged hDPP8/hDPP9 double-knockout 293T cells that expressed ASC–GFP and co-expressed hNLRP1 and hDPP9 variants (ref. [9](#)). As anticipated, wild-type hDPP9 efficiently inhibited the hNLRP1-dependent formation of ASC–GFP specks (Fig. [3d](#), Extended Data Fig. [7c](#)). By contrast, the binding-deficient but enzymatically active mutant hDPP9(L131E) (Extended Data Fig. [7d](#)) and the catalytically dead but binding-active mutant hDPP9(S759A) were both defective in doing so (Fig. [3d](#), Extended Data Fig. [7c](#)). These findings indicate that both catalytic activity and binding to the FIIND are important for the hDPP9-mediated inhibition of hNLRP1.

hDPP9(L131E)—but not hDPP9(S759A)—retained sensitivity towards VbP (Fig. [3d](#)), which provides further evidence that the protease activity of hDPP9 has a critical role in hNLRP1 inhibition. Furthermore, simultaneous mutations of these two residues were additive in eliminating the inhibitory effect of hDPP9 on the formation of ASC specks (Fig. [3d](#)). Expression of wild-type hDPP9—but not hDPP9(L131E), hDPP9(S759A) or hDPP9(L131E/S759A)—in hDPP8/hDPP9 double-knockout 293T cells fully inhibited IL-1 β cleavage by caspase-1 (Extended Data Fig. [7e](#)), further confirming that both catalytic activity and FIIND binding are essential for hNLRP1 inhibition.

rNLRP1-bound rDPP9 displayed protease activity, which was inhibited by VbP (Extended Data Fig. [7f](#)). The hDPP9(S759A)–hNLRP1 complex is therefore biochemically similar to VbP-bound hDPP9–hNLRP1. Overexpression of hDPP9(S759A) was expected to outcompete endogenous hDPP9 for hNLRP1 binding to form the hDPP9(S759A)–hNLRP1 complex, which is defective in hNLRP1 inhibition. Consistent with our previous results^{[26](#)}, VbP treatment substantially induced the secretion of mature IL-1 β (molecular mass 17 kDa) and extensive death of human keratinocytes (Fig. [3e,f](#), Extended Data Fig. [8a](#)), indicating that caspase-1 had been activated in these cells. Supporting our prediction, the doxycycline-induced expression of hDPP9(S759A)—but not of the hNLRP1-binding-deficient hDPP9(L131E) or

hDPP9(S759A/L131E) mutants—in these cells caused IL-1 β secretion and pyroptosis (Fig. 3e,f, Extended Data Fig. 8a). Similar observations were also made using 293T cells expressing ASC–GFP and NLRP1 (DPP8 $^+$ DPP9 $^+$) (Extended Data Fig. 8b). These data further support the dual requirement of the catalytic function of hDPP9 and its direct binding to hNLRP1 in order to suppress the activation of hNLRP1 in human cells.

Inhibition of UPA–CARD oligomerization by ZU5

The residues N1032, P1034, P1035 and V1037, at the centre of the UPA–UPA surface (Fig. 4a), are conserved among NLRP1 proteins from different species and CARD8 (Extended Data Fig. 1d). Notably, hNLRP1-2, an hNLRP1 splice isoform that lacks this loop region, displayed impaired cell-killing activity when ectopically expressed in MCF7 cells²⁷, supporting an essential role of this loop in hNLRP1-mediated cell death. Further supporting this conclusion, mutations of P1278 and L1281 of hNLRP1—from the corresponding UPA dimerization interface—substantially reduced VbP-induced ASC speck formation in ASC–GFP-expressing 293T cells (Fig. 4b, Extended Data Fig. 8c,d).

Fig. 4: Mechanism of ZU5-mediated NLRP1 inhibition by DPP9.



a, Cartoon representation of the UPA dimer in the 2:1 rNLRP1–rDPP9 complex. Key residues that mediate the formation of the dimeric UPA are shown in stick representation. **b**, The effect of the mutants hNLRP1(P1278E) and hNLRP1(L1281E)

on the VbP-induced activation of hNLRP1 in 293T cells. Bar graphs represent data from three inductions. Student's *t*-test, **P* < 0.05, ***P* < 0.01. **c**, Negative staining electron microscopy analysis of human wild-type and mutant UPA–CARD fragments. Scale bar, 100 nm. **d**, Alignment of the UPA homodimer from the 2:1 rNLRP1–rDPP9 complex (in cartoon) with the crystal structure of rNLRP1 FIIND (shown as a transparent surface). **e**, Structural alignment of the oligomeric UPA–CARD fragment (shown in surface representation, PDB ID: 6K7V) with the rNLRP1–rDPP9 complex. The two UPA molecules (in blue and cyan) from the complex were used as the template for alignment. **f**, Working model of the DPP9-mediated inhibition of NLRP1 and the pathogen-induced activation of NLRP1. In resting cells, an autoinhibited rNLRP1 interacts with a dimeric rDPP9 via its autoinhibitory ZU5 domain (1). This interaction enables rDPP9 to recruit an autocleaved NLRP1, resulting in dissociation of the N-terminal segment from the C-terminal UPA–CARD fragment and formation of a 2:1 rNLRP1–rDPP9 complex (2). The two UPA–CARD fragments in the complex are sequestered from oligomerization by interaction with the active sites of rDPP9 and interaction with the DPP9-bound autoinhibited rNLRP1 via UPA–UPA dimerization. There may exist a DPP9 substrate that has a role in NLRP1 inhibition, because protease activity is important in order for rDPP9 to suppress NLRP1. Such a substrate could also have a role in NLRP1 activation in the absence of DPP9. Pathogen-induced proteasomal degradation would lead to the release of the active UPA–CARD fragments from the complex (3). The released UPA–CARD fragment then oligomerizes (4) to recruit ASC for the activation of downstream immune signalling (5).

[Source data](#)

[Full size image](#)

UPA forms a ring-like oligomer, which then brings the CARDs into close proximity for the efficient, filament-like polymerization of hNLRP1 UPA–CARD²⁸. The UPA–UPA interaction surface observed in the 2:1 NLRP1–DPP9 complex might be similar to that in the UPA–CARD filament. Indeed, the wild-type UPA–CARD fragment of hNLRP1 has been found to form filamentous structures^{28,29}, but the UPA–UPA interface mutants—UPA–CARD(P1278E) and UPA–CARD(L1281E)—did not (Fig. 4c), which indicates that the UPA–UPA dimer interface is required for higher-order UPA oligomerization and activation of the hNLRP1 inflammasome.

VbP had little effect on the interaction of the rNLRP1 FIIND with rDPP9 (Extended Data Fig. 8e), which is consistent with data from NLRP1B and CARD8 but contrasts with those from hNLRP1 (refs. 8,9). It is of interest to note that the linker region between UPA and CARD in hNLRP1 is longer than that in NLRP1B and CARD8 (Extended Data Figs. 1d, 8f). Deletion analysis suggested a role of the linker region in the sensitivity of rDPP9-mediated hNLRP1 inhibition to VbP (Extended Data Fig. 8f).

The mutation hNLRP1(P1214R), which is associated with autoinflammatory diseases, could perturb the interaction of hNLRP1 with hDPP9 at the UPA-binding site⁹.

In contrast to free UPA^{28,29}, the FIIND was monomeric in solution (Extended Data Fig. 8g), which suggests that the ZU5 domain inhibits UPA dimerization or oligomerization. Supporting this notion, structural comparison revealed that ZU5 in the monomeric FIIND sterically hinders UPA dimerization in the 2:1 complex (Fig. 4d). In its activated state, the UPA–CARD fragment forms helical filaments in which dimeric UPA spirally wraps around the inner CARDs²⁸. Alignment of the dimeric UPA from the 2:1 rNLRP1–rDPP9 complex with that from the UPA–CARD filament²⁸ showed that ZU5 is positioned to block the spiral growth of the dimeric UPA (Fig. 4e). Collectively, these results show that ZU5 negatively regulates NLRP1 activation by directly or indirectly inhibiting the formation of UPA–CARD filaments.

Discussion

Sequestration of the potent UPA–CARD fragment in the 2:1 rNLRP1–rDPP9 complex can block the UPA-mediated formation of functional UPA–CARD filaments (Fig. 4f). This is consistent with the idea that sequestration of active domains of NLRs is a general strategy in the regulation of inflammasomes³⁰. Our data suggested that the ZU5 domain is also important for the inhibition of DPP9-independent UPA–CARD activation (Fig. 4d, Extended Data Fig. 8g). The ZU5 domain therefore seems to be critical for the negative regulation of both DPP9-independent and DPP9-dependent NLRP1 activation. The degradation of N-terminal fragments of NLRP1—induced by either lethal toxin or 3C proteases or by chemical inhibition of DPP9—can similarly disrupt ZU5-mediated interactions and consequently release the autoinhibited and the sequestered UPA–CARD fragments in the 2:1 NLRP1–DPP9 complex (Fig. 4f). Thus, through disruption of the ZU5-dependent interaction between NLRP1 and DPP9, N-terminal degradation of NLRP1 is the unifying mechanism of NLRP1 activation^{1,2,3,14}.

It remains unknown why the protease activity of DPP9 is important for NLRP1 inhibition. A plausible explanation is the existence of a substrate(s) that is required for inhibition, but how the substrate is involved is unclear. The requirement for protease activity raises the possibility that perturbations of this activity by pathogen-derived or host-derived molecules can trigger NLRP1 activation. If this is the case, the 2:1 NLRP1–DPP9 complex could represent a bona fide receptor that senses diverse signals that induce the N-terminal degradation of NLRP1 or perturb the protease activity of DPP9.

Methods

Data reporting

No statistical methods were used to predetermine sample size. The experiments were not randomized and the investigators were not blinded to allocation during experiments and outcome assessment.

Protein expression and purification

The genes encoding full-length rNLRP1 (GenBank ID: HM060632.1) and full-length rDPP9 (NCBI Reference Sequence: NM_001305241.1) were synthesized by Genewiz. The constructs of rNLRP1 (residues 1–1218), rDPP9 (residues 1–862, wild-type and all mutants), rNLRP1 FIIND (residues 822–1122, wild-type and all mutants) and rNLRP1 FIIND–CARD (residues 822–1218) were generated by a standard PCR-based cloning strategy and cloned into pFastBac-1 vector with an N-terminal GST tag or with no tag, and their identities were confirmed by sequencing. All the proteins were expressed using the Bac-to-Bac baculovirus expression system (Invitrogen) in sf21 cells at 28 °C. One litre of cells (2.5×10^6 cells per ml, medium from Expression Systems) was infected with 20 ml baculovirus at 28 °C. After growth at 28 °C for 48 h, the cells were collected, resuspended in the buffer containing 25 mM Tris-HCl pH 8.0 and 150 mM NaCl, and lysed by sonication. The soluble fraction was purified from the cell lysate using Glutathione Sepharose 4B beads (GS4B, GE Healthcare). The proteins were then digested with PreScission protease (GE Healthcare) to remove the GST tag and further purified by gel filtration (SuperoseTM 6 prep grade XK 16/70; GE Healthcare). To prepare the rNLRP1 FIIND for crystallization trials, the purified rNLRP1 FIIND (residues 822–1122) was concentrated to about 8.0 mg ml⁻¹ in buffer containing 100 mM NaCl, 10 mM Tris-HCl pH 8.0. For co-expression of rNLRP1 and rDPP9, one litre of sf21 cells were co-infected with 10 ml recombinant baculovirus of rNLRP1 and rDPP9, and then the rNLRP1–rDPP9 complex was purified using GS4B beads. Similar protocols were used to purify the complex containing full-length GST–CARD8 and hDPP9. For cryo-EM investigation, the purified rNLRP1–rDPP9 complex was concentrated to about 0.3 mg ml⁻¹ in buffer containing 25 mM Tris-HCl pH 8.0, 150 mM NaCl and 3 mM DTT.

Recombinant hNLRP1 UPA–CARD tagged with a removable Snap domain was expressed using bacterial vectors as the form of inclusion bodies. After cellular lysis, the cellular pellet was collected after centrifugation at 30,000g for 30 min at 4 °C. Several additional washes using wash buffer (20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 1% Triton-X and 1 mM DTT) were performed until a pure white pellet was obtained. The pellet was dissolved in 6 M guanidinium, and centrifuged at 30,000g for a second time for 30 min at room temperature to remove contaminants. The denatured soluble proteins were then gradually dialysed against 3, 2, 1.5, 1, 0.8 and 0.6 M guanidinium in dialysis buffer (20 mM Tris-HCl, pH 8.0, 150 mM NaCl, 5 mM β-mercaptoethanol)

in the cold room, and eventually in fresh dialysis buffer without guanidinium. The refolded proteins were centrifuged for a third time at 10,000g for 10 min to remove misfolded aggregates. The soluble refolded fractions were then subjected to biochemical analysis and negative-stain electron microscopy experiments. The Snap tag was removed by 3C proteases, and the product was further purified by reverse Ni-NTA purification.

Gel-filtration assay

The GST–rNLRP1 FIIND(S969A)–rDPP9 complex and rNLRP1 FIIND proteins purified as described in the previous section were subjected to gel filtration (Superose 6, 10/30; GE Healthcare) in buffer containing 10 mM Tris pH 8.0 and 100 mM NaCl. The purified rNLRP1 FIIND was left at 18 °C for two weeks to obtain its fully autoprocessed form. The fully processed rNLRP1 FIIND was then incubated with the purified GST–rNLRP1 FIIND(S969A)–rDPP9 complex at a molar ratio of about 1:1 in 4 °C for 150 min before gel-filtration analysis. Samples from relevant fractions were applied to an SDS–PAGE gel and visualized by Coomassie blue staining. A similar procedure was used to assay the interaction of rDPP9 with other rNLRP1 mutant proteins.

Pull-down assay

Sf21 cells (50 ml; 2.5×10^6 cells per ml, medium from Expression Systems) were infected with 1 ml baculovirus of GST–rNLRP1 FIIND (wild-type or mutants), and the proteins were expressed and purified as described in the section ‘Protein expression and purification’. In brief, the proteins were purified from the cell lysate using 300 µl GS4B resin (GS4B, GE Healthcare), and incubated with an excess of purified wild-type or mutant rDPP9 proteins on ice for 60 min. The resin was washed with 1 ml buffer containing 10 mM Tris pH 8.0, 100 mM NaCl five times, and eluted with 300 µl buffer containing 25 mM Tris pH 8.0, 150 mM NaCl, 15 mM GSH. The eluted samples were analysed by SDS–PAGE and visualized by Coomassie blue staining. A similar procedure was used to assay the interaction between full-length CARD8 and hDPP9.

To test the effect of VbP on the rNLRP1–rDPP9 interaction, 2 mM VbP was added to the purified rDPP9, GST–rNLRP1(S969A)–rDPP9 complex or GST–rNLRP1 FIIND–rDPP9 complex. After 60-min incubation, the samples were individually incubated with wild-type or mutant rNLRP1 FIIND and 100 µl GS4B resin on ice for 60 min. After extensive washing, the proteins bound in the resin were eluted and analysed by SDS–PAGE and visualized by Coomassie blue staining.

Enzymatic activity assay

To measure rDPP9 protease activity, a stock solution of substrate (10 mM Gly-Pro-AMC) was prepared in DMSO. Purified wild-type or mutant rDPP9 was diluted to 1 μ M to a final volume of 100 μ l in buffer containing 10 mM Tris pH 8.0 and 100 mM NaCl. The substrate Gly-Pro-AMC (10 μ l of 10 μ M solution, dissolved in DMSO) was added to the mixture. Substrate cleavage was measured by the liberated AMC fluorescence signal recorded at room temperature in a luminescence spectrometer at excitation and emission wavelengths of 380 nm and 500 nm, respectively, over a period of 30 min.

To measure the protease activity of hDPP9, 293T cells were transfected with hDPP9 and lysed 48 h post transfection in 1 \times Tris-buffered saline (TBS) with 0.25% NP40. Lysate (0.3 μ g) was mixed with 0.1 μ l 100 mM Gly-Pro-AMC in 50 μ l lysis buffer. AMC fluorescence (380 nm excitation; 500 nm emission) was monitored at room temperature for 30 min at 1-min intervals.

Edman degradation by the PPSQ-33A system

The phenylthiohydantoin amino acid was separated in the reversed-phase mode of high-performance liquid chromatography using the differences between the retention times of different amino acids, and the amount of UV absorbance at specific wavelengths was detected. The samples were transferred to the PVDF membrane and five cycles were set. The amino acid sequences of each sample were determined from the chromatograms obtained in each cycle evaluation performed by comparing chromatograms with those in the previous and subsequent cycles and identifying the phenylthiohydantoin amino acids that had the greatest increase in abundance.

Cryo-EM sample preparation and data collection

An aliquot of 3 μ l of purified rNLRP1–rDPP9 or rNLRP1 FIIND–CARD(S969A)–rDPP9 complex was applied to holey carbon grids (Quantifoil Au 1.2/1.3, 300 mesh), which were glow-discharged for 30 s at middle level in Harrick Plasma after 2 min evacuation. The grids were then blotted by filter papers (Ted Pella) for 2.5 s at 8 °C and 100% humidity, then flash-frozen in liquid ethane using FEI Vitrobot Marke IV.

Cryo-EM data for rNLRP1–rDPP9 and rNLRP1 FIIND–CARD(S969A)–rDPP9 were collected on a Titan Krios electron microscope operated at 300 kV, equipped with a Gatan K2 Summit direct electron detector and a Gatan Quantum energy filter (an additional Cs-corrector that was used for rNLRP1 FIIND–CARD(S969A)–rDPP9 data collection). A total of 7,157 and 4,971 micrograph stacks were automatically recorded using AutoEMation in super-resolution mode for rNLRP1–rDPP9 and rNLRP1 FIIND–CARD(S969A)–rDPP9, at a nominal magnification of 130,000 \times and 105,000 \times , respectively. Defocus values varied from -1.0 μ m to -2.0 μ m for both

datasets³¹. Dose rates during the collection of data for rNLRP1–rDPP9 and rNLRP1 FIIND–CARD(S969A)–rDPP9 were 10 and 11 electrons per pixel per second, respectively. For both datasets, the exposure time of 5.6 s was dose-fractionated into 32 sub-frames, leading to a total accumulated dose of approximate 50 electrons per Å² for each stack.

Image processing and 3D reconstruction

The stacks of rNLRP1–rDPP9 and rNLRP1 FIIND–CARD(S969A)–rDPP9 recorded in super-resolution mode were motion-corrected using MotionCor2 and binned twofold, resulting in a physical pixel size of 1.061 Å per pixel and 1.091 Å per pixel, respectively³². Meanwhile, dose weighting for the summed micrographs was performed³³. CTFFIND4 was then used to estimate the contrast transfer function (CTF) parameters³⁴. On the basis of the CTF estimation, 7,033 and 4,667 micrographs were manually selected for rNLRP1–rDPP9 and rNLRP1 FIIND–CARD(S969A)–rDPP9, respectively, and were further processed in Relion 3.1. Approximately 2,000 particles were manually picked and 2D-classified to generate initial templates for autopicking. In the end, 2,700,586 and 1,725,380 particles were automatically picked for rNLRP1–rDPP9 and rNLRP1 FIIND–CARD(S969A)–rDPP9, respectively, using Relion 3.1. After several rounds of reference-free 2D classification, 1,430,734 particles for rNLRP1–rDPP9 and 1,117,656 particles for rNLRP1 FIIND–CARD(S969A)–rDPP9 were subjected to 3D classification, using the initial 3D reference models obtained by ab initio calculation from Relion 3.1. Particles from good 3D classes, with better overall structure features, were selected for 3D refinement. After global 3D refinement and post-processing, the resolution was 3.07 Å with a particle number of 343,648 for rNLRP1–rDPP9, and 3.69 Å with a particle number of 252,425 for rNLRP1 FIIND–CARD(S969A)–rDPP9.

To improve the quality of the density of the NLRP1 section in the rNLRP1–rDPP9 map, the rNLRP1–rDPP9 particles after 3D refinement were then subjected to a further round of focused 3D classification with a local mask generated using Chimera. A previously reported focused 3D classification procedure was adopted to select the 3D class with good density³⁵. Ultimately, a subset of 182,116 particles after focused 3D classification were subjected to a final 3D refinement and yielded a global reconstruction at 3.18 Å after postprocess.

2D classification, 3D classification and 3D autorefinement were all performed using Relion 3.1 (refs. ^{36,37,38}). The resolutions were determined by gold-standard Fourier shell correlation³⁹. Local resolution distribution was evaluated⁴⁰ using Relion 3.1.

Crystallization, data collection and structure determination

Crystallization of rNLRP1 was performed by hanging-drop vapour-diffusion methods, mixing 1 μ l of 8 mg ml $^{-1}$ protein with 1 μ l of reservoir solution at 18 °C. Good-quality crystals of rNLRP1 FIIND were obtained in buffer containing 1.0 M ammonium sulfate, 0.1 M Bis-Tris pH 5.5, 1% w/v polyethylene glycol 3,350. All the crystals were flash-frozen in reservoir buffer to which glycerol (15%) was added as the cryoprotectant to prevent radiation damage. The diffraction dataset was collected at the Shanghai Synchrotron Radiation Facility (SSRF) on the beamline BL19U1 using a CCD detector and was processed using HKL2000 software package. The crystal structure of rNLRP1 FIIND was determined by PHASER_MR with the structure of NUC5b as the search model. The model from the molecular replacement was manually rebuilt to the sequence of rNLRP1 FIIND in the program Coot⁴¹ and subsequently subjected to refinement by the program Refine_Phenix⁴². Data collection, processing, and refinement statistics are summarized in Extended Data Table 1.

Model building and refinement

The EM density map of rNLRP1–rDPP9 was used for model building, as the quality of density for rNLRP1 was sufficient for sequence assignment. The model of hDPP9 (PDB ID: 6EOQ)²³, along with two copies of the rNLRP1 FIIND crystal structure that we determined as described in the previous section, were docked into the EM density map of rNLRP1–rDPP9 in Chimera⁴³. The sequence of hDPP9 was changed to that of rDPP9, the whole model containing two rNLRP1 FIIND molecules and a rDPP9 dimer was then adjusted manually in the program Coot⁴¹, and refined against the EM map by Phenix in real space with secondary structure and geometry restraints⁴². The final model of the rNLRP1–rDPP9 complex was validated using MolProbity and EMRinger in the Phenix package⁴². The model statistics are summarized in Extended Data Table 2.

ASC–GFP transfection in 293T cells and ASC–GFP speck formation assay

293T ASC–GFP and 293T ASC–GFP DPP8/DPP9 double-knockout cells have been previously described⁹. All transfections were carried out using Lipofectamine 2000 (Thermo Fisher). For immunoprecipitation, cells were collected 48 h post transfection. For the ASC–GFP speck assay, cells were fixed 24 h post transfection and counterstained with DAPI or Hoescht before wide-field fluorescence imaging. The number of nuclei per field of view was counted in ImageJ using the following image processing steps: ‘Threshold’ (20–30 to 255); ‘Watershed’; and ‘Analyze Particles’ (200–infinity). ASC specks were counted in ImageJ in the GFP channel using ‘Find Maxima’ (prominence = 20).

Inflammasome activation assays in immortalized keratinocytes

Immortalized human keratinocytes (N/TERT-1) were a gift from H. Reinwald (Harvard University) (Material Transfer Agreement to Skin Research Institute of Singapore). Stably transduced N/TERT-1 cells were induced with doxycycline (1 µg ml⁻¹) for 24 h. Immunoblotting antibodies used were as follows: anti-IL-1β p17 specific (CST, 83186S); GAPDH (Santa Cruz Biotechnology, sc-47724); IL-1β (R&D systems, AF-201); anti-Flag tag (Sigma Aldrich, F3165). All horseradish peroxidase (HRP)-conjugated secondary antibodies were purchased from Jackson Immunoresearch (goat anti-mouse IgG, 115-035-166; goat anti-rabbit IgG, 111-035-144; and donkey anti-goat IgG, 705-005-147).

Reporting summary

Further information on research design is available in the [Nature Research Reporting Summary](#) linked to this paper.

Data availability

The atomic coordinates and structure factors have been deposited in the RCSB Protein Data Bank (PDB) and Electron Microscopy Data Bank (EMDB). The PDB codes of the rNLRP1 FIIND and rNLRP1–rDPP9 structures are [7CRV](#) and [7CRW](#), respectively. The EMDB codes of the rNLRP1–rDPP9 and rNLRP1 FIIND–CARD(S969A)–rDPP9 structures are [EMD-30458](#) and [EMD-30459](#), respectively. For gel source images, see Supplementary Information. All other data or materials can be obtained from the corresponding author upon request. [Source data](#) are provided with this paper.

References

1. 1.

Chui, A. J. et al. N-terminal degradation activates the NLRP1B inflammasome. *Science* **364**, 82–85 (2019).

[CAS](#) [Article](#) [ADS](#) [Google Scholar](#)

2. 2.

Xu, H. et al. The N-end rule ubiquitin ligase UBR2 mediates NLRP1B inflammasome activation by anthrax lethal toxin. *EMBO J.* **38**, e101996 (2019).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

3. 3.

Sandstrom, A. et al. Functional degradation: A mechanism of NLRP1 inflammasome activation by diverse pathogen enzymes. *Science* **364**, eaau1330 (2019).

[CAS Article](#) [Google Scholar](#)

4. 4.

Finger, J. N. et al. Autolytic proteolysis within the function to find domain (FIIND) is required for NLRP1 inflammasome activity. *J. Biol. Chem.* **287**, 25030–25037 (2012).

[CAS Article](#) [Google Scholar](#)

5. 5.

Levinsohn, J. L. et al. Anthrax lethal factor cleavage of Nlrp1 is required for activation of the inflammasome. *PLoS Pathog.* **8**, e1002638 (2012).

[Article](#) [Google Scholar](#)

6. 6.

Okondo, M. C. et al. Inhibition of Dpp8/9 activates the Nlrp1b inflammasome. *Cell Chem. Biol.* **25**, 262–267.e5 (2018).

[CAS Article](#) [Google Scholar](#)

7. 7.

Frew, B. C., Joag, V. R. & Mogridge, J. Proteolytic processing of Nlrp1b is required for inflammasome activity. *PLoS Pathog.* **8**, e1002659 (2012).

[CAS Article](#) [Google Scholar](#)

8. 8.

Griswold, A. R. et al. DPP9's enzymatic activity and not its binding to CARD8 inhibits inflammasome activation. *ACS Chem. Biol.* **14**, 2424–2429 (2019).

[CAS Article](#) [Google Scholar](#)

9. 9.

Zhong, F. L. et al. Human DPP9 represses NLRP1 inflammasome and protects against autoinflammatory diseases via both peptidase activity and FIIND domain binding. *J. Biol. Chem.* **293**, 18864–18878 (2018).

[CAS](#) [Article](#) [Google Scholar](#)

10. 10.

Mitchell, P. S., Sandstrom, A. & Vance, R. E. The NLRP1 inflammasome: new mechanistic insights and unresolved mysteries. *Curr. Opin. Immunol.* **60**, 37–45 (2019).

[CAS](#) [Article](#) [Google Scholar](#)

11. 11.

Pathan, N. et al. TUCAN, an antiapoptotic caspase-associated recruitment domain family protein overexpressed in cancer. *J. Biol. Chem.* **276**, 32220–32229 (2001).

[CAS](#) [Article](#) [Google Scholar](#)

12. 12.

Johnson, D. C. et al. DPP8/DPP9 inhibitor-induced pyroptosis for treatment of acute myeloid leukemia. *Nat. Med.* **24**, 1151–1156 (2018).

[CAS](#) [Article](#) [Google Scholar](#)

13. 13.

Hellmich, K. A. et al. Anthrax lethal factor cleaves mouse Nlrp1b in both toxin-sensitive and toxin-resistant macrophages. *PLoS ONE* **7**, e49741 (2012).

[CAS](#) [Article](#) [ADS](#) [Google Scholar](#)

14. 14.

Robinson, K. S. et al. Enteroviral 3C protease activates the human NLRP1 inflammasome in airway epithelia. *Science* **370**, eaay2002 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

15. 15.

Wang, Q. et al. CARD8 is an inflammasome sensor for HIV-1 protease activity. *Science* <https://doi.org/10.1126/science.abe1707> (2021).

16. 16.

Zhang, H., Chen, Y., Keane, F. M. & Gorrell, M. D. Advances in understanding the expression and function of dipeptidyl peptidase 8 and 9. *Mol. Cancer Res.* **11**, 1487–1496 (2013).

[CAS Article](#) [Google Scholar](#)

17. 17.

de Vasconcelos, N. M. et al. DPP8/DPP9 inhibition elicits canonical Nlrp1b inflammasome hallmarks in murine macrophages. *Life Sci. Alliance* **2**, e201900313 (2019).

[Article](#) [Google Scholar](#)

18. 18.

Gai, K. et al. DPP8/9 inhibitors are universal activators of functional NLRP1 alleles. *Cell Death Dis.* **10**, 587 (2019).

[Article](#) [Google Scholar](#)

19. 19.

Wang, R. et al. Autoinhibition of UNC5b revealed by the cytoplasmic domain structure of the receptor. *Mol. Cell* **33**, 692–703 (2009).

[CAS Article](#) [Google Scholar](#)

20. 20.

D'Osualdo, A. et al. CARD8 and NLRP1 undergo autoproteolytic processing through a ZU5-like domain. *PLoS ONE* **6**, e27396 (2011).

[Article](#) [ADS](#) [Google Scholar](#)

21. 21.

Hodel, A. E. et al. The three-dimensional structure of the autoproteolytic, nuclear pore-targeting domain of the human nucleoporin Nup98. *Mol. Cell* **10**, 347–358

(2002).

[CAS](#) [Article](#) [Google Scholar](#)

22. 22.

Tinel, A. et al. Autoproteolysis of PIDD marks the bifurcation between pro-death caspase-2 and pro-survival NF- κ B pathway. *EMBO J.* **26**, 197–208 (2007).

[CAS](#) [Article](#) [Google Scholar](#)

23. 23.

Ross, B. et al. Structures and mechanism of dipeptidyl peptidases 8 and 9, important players in cellular homeostasis and cancer. *Proc. Natl Acad. Sci. USA* **115**, E1437–E1445 (2018).

[CAS](#) [Article](#) [Google Scholar](#)

24. 24.

Okondo, M. C. et al. DPP8 and DPP9 inhibition induces pro-caspase-1-dependent monocyte and macrophage pyroptosis. *Nat. Chem. Biol.* **13**, 46–53 (2017).

[CAS](#) [Article](#) [Google Scholar](#)

25. 25.

Griswold, A. R. et al. A chemical strategy for protease substrate profiling. *Cell Chem. Biol.* **26**, 901–907.e6 (2019).

[CAS](#) [Article](#) [Google Scholar](#)

26. 26.

Zhong, F. L. et al. Germline NLRP1 mutations cause skin inflammatory and cancer susceptibility syndromes via inflammasome activation. *Cell* **167**, 187–202.e17 (2016).

[CAS](#) [Article](#) [Google Scholar](#)

27. 27.

Hlaing, T. et al. Molecular cloning and characterization of DEFCAP-L and -S, two isoforms of a novel member of the mammalian Ced-4 family of apoptosis

proteins. *J. Biol. Chem.* **276**, 9230–9238 (2001).

[CAS](#) [Article](#) [Google Scholar](#)

28. 28.

Qin, G. et al. Structural basis for distinct inflammasome complex assembly by human NLRP1 and CARD8. *Nat. Commun.* **12**, 188 (2021).

[Article](#) [Google Scholar](#)

29. 29.

Hollingsworth, L. R. et al. Mechanism of filament formation in UPA-promoted CARD8 and NLRP1 inflammasomes. *Nat. Commun.* **12**, 189 (2021).

[Article](#) [Google Scholar](#)

30. 30.

Pedraza-Alva, G., Pérez-Martínez, L., Valdez-Hernández, L., Meza-Sosa, K. F. & Ando-Kuri, M. Negative regulation of the inflammasome: keeping inflammation under control. *Immunol. Rev.* **265**, 231–257 (2015).

[CAS](#) [Article](#) [Google Scholar](#)

31. 31.

Lei, J. & Frank, J. Automated acquisition of cryo-electron micrographs for single particle reconstruction on an FEI Tecnai electron microscope. *J. Struct. Biol.* **150**, 69–80 (2005).

[Article](#) [Google Scholar](#)

32. 32.

Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).

[CAS](#) [Article](#) [Google Scholar](#)

33. 33.

Grant, T. & Grigorieff, N. Measuring the optimal exposure for single particle cryo-EM using a 2.6 Å reconstruction of rotavirus VP6. *eLife* **4**, e06980 (2015).

[Article](#) [Google Scholar](#)

34. 34.

Mindell, J. A. & Grigorieff, N. Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.* **142**, 334–347 (2003).

[Article](#) [Google Scholar](#)

35. 35.

Bai, X. C., Rajendra, E., Yang, G., Shi, Y. & Scheres, S. H. Sampling the conformational space of the catalytic subunit of human γ -secretase. *eLife* **4**, e11182 (2015).

[Article](#) [Google Scholar](#)

36. 36.

Scheres, S. H. A Bayesian view on cryo-EM structure determination. *J. Mol. Biol.* **415**, 406–418 (2012).

[CAS](#) [Article](#) [Google Scholar](#)

37. 37.

Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

38. 38.

Zivanov, J., Nakane, T. & Scheres, S. H. W. Estimation of high-order aberrations and anisotropic magnification from cryo-EM datasets in RELION-3.1. *IUCrJ* **7**, 253–267 (2020).

39. 39.

Rosenthal, P. B. & Henderson, R. Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003).

[CAS](#) [Article](#) [Google Scholar](#)

40. 40.

Kucukelbir, A., Sigworth, F. J. & Tagare, H. D. Quantifying the local resolution of cryo-EM density maps. *Nat. Methods* **11**, 63–65 (2014).

[CAS](#) [Article](#) [Google Scholar](#)

41. 41.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).

[CAS](#) [Article](#) [Google Scholar](#)

42. 42.

Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66**, 213–221 (2010).

[CAS](#) [Article](#) [Google Scholar](#)

43. 43.

Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

[CAS](#) [Article](#) [Google Scholar](#)

[Download references](#)

Acknowledgements

We thank J. Lei and X. Li at Tsinghua University for data collection, and P. Schulze-Lefert at Max Planck Institute for discussion. We acknowledge support with cryo-EM facilities from the Tsinghua University Branch of the China National Center for Protein Sciences (Beijing), and support with computational facilities from the Bio-Computing Platform cluster. We thank the Shanghai Synchrotron Radiation Facility (SSRF) for assistance with data collection, and M. Han in Technology Center for Protein Science Tsinghua University for Edman sequencing analysis. This research was funded by the National Natural Science Foundation of China (31421001 to J.C.), the Alexander von Humboldt Foundation (a Humboldt professorship to J.C.), Max Planck-Gesellschaft (a Max Planck fellowship to J.C.), the Ministry of Health, Singapore, NMRC grant (MOH-000382-00 to W.B.), the Concern Foundation (F.L.Z),

a Nanyang Assistant Professorship (F.L.Z.) and the National Research Foundation fellowship (NRF-NRFF11-2019-0006 to F.L.Z.). We also acknowledge help and advice from members of the Reversade laboratory at Agency for Science, Technology and Research (A*STAR).

Funding

Open access funding provided by Max Planck Society.

Author information

Author notes

1. These authors contributed equally: Menghang Huang, Xiaoxiao Zhang

Affiliations

1. Beijing Advanced Innovation Center for Structural Biology, Tsinghua-Peking Joint Center for Life Sciences, Center for Plant Biology, School of Life Sciences, Tsinghua University, Beijing, China

Menghang Huang, Xiaoxiao Zhang, Jia Wang, Zhifu Han & Jijie Chai

2. Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

Gee Ann Toh & Franklin Zhong

3. School of Biological Sciences, Nanyang Technological University, Singapore, Singapore

Qin Gong & Bin Wu

4. Institute of Structural Biology, Nanyang Technological University, Singapore, Singapore

Bin Wu

5. Skin Research Institute of Singapore (SRIS), Singapore, Singapore

Franklin Zhong

6. Max Planck Institute for Plant Breeding Research, Cologne, Germany

Jijie Chai

7. Institute of Biochemistry, University of Cologne, Cologne, Germany

Jijie Chai

Authors

1. Menghang Huang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Xiaoxiao Zhang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Gee Ann Toh

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Qin Gong

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Jia Wang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Zhifu Han

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Bin Wu

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Franklin Zhong

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. Jijie Chai

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Contributions

J.C., F.Z., M.H. and Z.H. designed the project. J.C. and F.Z. supervised the project. M.H., X.Z., G.A.T., Q.G., J.W. and Z.H. performed the experiments. The data were analysed by all authors. J.C., F.Z. and B.W. wrote the manuscript with input from all authors.

Corresponding authors

Correspondence to [Franklin Zhong](#) or [Jijie Chai](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Zachary Billman, Edward Miao 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 Complete autocleavage of rNLRP1 FIIND in crystals and sequence alignment among rNLRP1, hNLRP1, NLRP1B and CARD8.](#)

a, Electron density around the active site of rNLRP1 FIIND. The autocleavage site is between F968 and S969 of rNLRP1. **b**, SDS-PAGE analysis of the crystals of rNLRP1

FIIND. The crystals were collected and dissolved in SDS–PAGE buffer. Lanes 1–4 represent crystals from four different wells. **c**, Structural alignment of rNLRP1 FIIND and UNC5b. The crystal structure of rNLRP1 FIIND was aligned with that of UNC5b (PDB code: 3G5B). The colour code is indicated on the figure. **d**, Sequence alignment of the FIIND-CARD domains of rNLRP1, hNLRP1, NLRP1B and CARD8. The rDPP9-interacting residues around the ZU5-binding site, the UPA-binding site and residues at the UPA homodimeric interface are highlighted with red, blue and yellow squares, respectively. The two residues from the catalytic FS motif are indicated by purple dots, and the linker region between UPA and CARD is highlighted within the red box. See Supplementary Fig. 1 for gel raw data.

Extended Data Fig. 2 Gel-filtration analysis of the rNLRP1–rDPP9 and the CARD8–hDPP9 interactions.

a, Full-length GST–rNLRP1 and rDPP9 were expressed in insect cells. The complex was purified through GS4B resin. After elution, GST was removed using precision protease and the complex was subjected to gel filtration. Left, the gel-filtration profile of the complex. Right, the protein fractions were visualized by SDS–PAGE followed by Coomassie-blue staining. **b**, The interaction between full-length GST–CARD8 and hDPP9 was assayed as described in **a**. See Supplementary Fig. 1 for gel raw data.

Extended Data Fig. 3 3D reconstruction of the rNLRP1–rDPP9 complex.

a, A representative cryo-EM image of the rNLRP1–rDPP9 complex. **b**, Representative views of 2D class averages of the rNLRP1–rDPP9 complex used for 3D reconstruction. 2D classes with different rDPP9-binding features are shown in different colours. **c**, Flow chart of cryo-EM data processing and 3D reconstruction of the rNLRP1–rDPP9 complex. **d**, The final EM density map of the rNLRP1–rDPP9 complex. The colour code shows the local resolution estimated using Relion. **e**, Fourier shell correlation curves at 0.143 for the final model of the rNLRP1–rDPP9 complex unmasked (red) or masked (black), and at 0.5 for rNLRP1–rDPP9 global model refined against the 3.18 Å rNLRP1–rDPP9 global map (green), the first half map (blue) and the second half map (cyan). The small difference between the blue and cyan curves after the estimated resolution indicates no overfitting of the coordinate refinement.

Extended Data Fig. 4 Structural and sequence alignment.

a, Left, structural alignment between the first rDPP9, second rDPP9, apo-hDPP9 and hDPP9–1G244 (in cartoon representation). The colour code is indicated on the figure. Better-defined loop regions in the rNLRP1-bound rDPP9 subunit are highlighted with open frames. Right, electron density around the stabilized loop regions induced by

rNLRP1. All densities are generated in PyMOL and contoured at 5 (region coloured blue) or 7 (region coloured yellow) sigma. **b**, Alignment of the cryo-EM structure of the rNLRP1 FIIND (with ZU5 shown in pink and UPA in blue) with the crystal structure of the rNLRP1 FIIND (in cyan). **c**, Sequence alignment of rDPP9, hDPP9, hDPP8, hDPP4 and hDPP7. The rNLRP1-interacting residues from the ZU5- and UPA-binding site are marked with red and blue squares, respectively.

Extended Data Fig. 5 Functional assays to analyse the influence of DPP9 protease activity on the interaction between DPP9 and NLRP1.

a, N-terminal sequence analysis of rNLRP1 FIIND by Edman degradation. Shown on the left and the right are N-terminal sequences of UPA from rNLRP1 FIIND and rNLRP1 FIIND–rDPP9, respectively. **b**, Catalytic activity of rDPP9 is dispensable for the binding of rDPP9 to rNLRP1. The catalytic mutant rDPP9(S729A) with GST fused at the N terminus was co-expressed with rNLRP1 FIIND–CARD in insect cells. The complex was purified through GS4B resin. After elution, GST was removed using precision protease and the complex was subjected to gel filtration. Left, the gel-filtration profile of the complex. Right, the protein fractions were visualized by SDS-PAGE followed by Coomassie-blue staining. **c**, Gel-filtration analysis of the autoprocessing-deficient S969A mutant of rNLRP1 FIIND–CARD in complex with rDPP9. The two proteins were co-expressed in insect cells and purified using affinity chromatography. The purified complex was further cleaned using gel filtration. Left, the gel-filtration profile of the complex. Right, the fractions of the two proteins were visualized by SDS-PAGE followed by Coomassie-blue staining. See Supplementary Fig. 1 for gel raw data.

Extended Data Fig. 6 3D reconstruction of the rNLRP1 FIIND–CARD(S969A)–rDPP9 complex.

a, A representative cryo-EM image of the rNLRP1 FIIND–CARD(S969A)–rDPP9 complex. **b**, Representative 2D class averages of the rNLRP1 FIIND–CARD(S969A)–rDPP9 complex. **c**, Flow chart of cryo-EM data processing and 3D reconstruction of the rNLRP1 FIIND–CARD(S969A)–rDPP9 complex. **d**, The final EM density map of the rNLRP1 FIIND–CARD(S969A)–rDPP9 complex. The colour code shows the local resolution estimated using Relion. **e**, Fourier shell correlation curve (at 0.143) of the final reconstruction of the rNLRP1 FIIND–CARD(S969A)–rDPP9 complex.

Extended Data Fig. 7 DPP9 interacts with FIIND from rNLRP1 and CARD8.

a, Mutagenesis analysis of the ZU5-binding site of rDPP9. N-terminally GST-fused wild-type or mutant rNLRP1 FIIND and rDPP9 were individually purified from insect cells. Wild-type or mutant rNLRP1 FIIND was used to pull down rDPP9 (including rDPP9(L101E)) with GS4B resin. After extensive washing, proteins bound to the GS4B resin were visualized by SDS-PAGE followed by Coomassie-blue staining. **b**, The interaction of hDPP9 with CARD8. N-terminally GST-fused CARD8 was used to pull down non-tagged wild-type or mutant hDPP9 in vitro. **c**, Top, representative images of DPP8/DPP9 double-knockout 293T cells transfected with hNLRP1 and wild-type hDPP9 or hDPP9 mutants. Bottom, expression of wild-type and DPP9 mutants in DPP8/DPP9 double-knockout 293T cells. The soluble lysate was blotted for DPP9, NLRP1 or GAPDH. **d**, Mutation of L101 of rDPP9 (left) or its equivalent L131 of hDPP9 (right) has no effect on the protease activity. Each experiment was repeated at least three times. One-way ANOVA, *** $P < 0.0001$. RFU, relative fluorescence units. **e**, DPP8/DPP9 double-knockout 293T cells were co-transfected with 3×Flag hDPP9 variant together with CASP1-Myc and pro-IL-1 β . ASC-GFP speck formation was monitored for consistency with that in **c**. Lysates were analysed by immunoblotting. **f**, Left, inhibition of the protease activity of rDPP9 and rNLRP1–rDPP9 by VbP. Right, SDS-PAGE analyses of the rDPP9 and rNLRP1–rDPP9 complexes used in the activity assays. See Supplementary Figs. 2 and 4 for gel raw data. [Source data](#)

Extended Data Fig. 8 Additional evidence for the suppression of NLRP1 by DPP9 binding and enzymatic activity.

a, Immortalized keratinocytes stably transduced with Tet-ON-3×Flag hDPP9 lentiviruses were treated with or without doxycycline for 24 h. Conditioned media were subjected to anti-IL-1 β ELISA. The y axis represents the concentrations of IL-1 β in the medium. Bar graphs represent data from three inductions. Two-way ANOVA, *** $P < 0.0001$. **b**, Middle, representative images of 293T cells expressing ASC-GFP and hNLRP1 and transfected with wild-type hDPP9 or hDPP9 mutants. Left, the percentage of cells with ASC-GFP specks was quantified using more than 200 cells (left). Right, the expression of wild-type hDPP9 or hDPP9 mutants in 293T cells expressing ASC-GFP and NLRP1. The soluble lysate was blotted for hDPP9, hNLRP1-NT, ASC-GFP or GAPDH. Each experiment was repeated at least three times. **c**, Representative images of ASC-GFP expressing 293T cells transfected with the indicated hNLRP1 mutants. VbP (3 μ M) was added 24 h post transfection. Cells were fixed 48 h post transfection and the nuclei were counterstained with Hoescht 33342. The percentage of cells with ASC-GFP specks was quantified using more than 200 cells. **d**, ASC-GFP 293T cells transfected with C-terminally Flag-tagged hNLRP1 constructs. Anti-Flag immunoprecipitation was performed on approximately 1 mg whole-cell lysate 48 h after transfection. **e**, GST-fused rNLRP1 FIIND and rDPP9 were individually purified from insect cells. Wild-type or mutant rNLRP1 FIIND

proteins were used to pull down rDPP9 with GS4B resin in the presence or in the absence of VbP. **f**, Top, 293T cells were transfected with the indicated hNLRP1–Flag mutants. Anti-Flag immunoprecipitation was performed as in **d**. The hDPP9 band is indicated with the yellow arrow. The linker region between hNLRP1 UPA and CARD begins at amino acid 1360. Bottom, comparison of the linker regions of hNLRP1, rNLRP1, NLRP1B and hCARD8. The linker length and the sensitivity of the respective 2:1 complex to VbP are shown on the right. **g**, The gel filtration profile (left) and SDS–PAGE (right) of the rNLRP1 FIIND. See Supplementary Figs. [2](#) and [5](#) for gel raw data. [Source data](#)

Extended Data Table 1 Crystallography data collection and refinement statistics
[Full size table](#)

Extended Data Table 2 Cryo-EM data collection, 3D reconstruction and model statistics

[Full size table](#)

Supplementary information

[Supplementary Figures](#)

This file contains Supplementary Figures 1-5 displaying the uncropped scans.

[Reporting Summary](#)

Source data

[Source Data Fig. 3](#)

[Source Data Fig. 4](#)

[Source Data Extended Data Fig. 7](#)

[Source Data Extended Data Fig. 8](#)

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and

indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

Reprints and Permissions

About this article



Check for
updates

Cite this article

Huang, M., Zhang, X., Toh, G.A. *et al.* Structural and biochemical mechanisms of NLRP1 inhibition by DPP9. *Nature* **592**, 773–777 (2021).
<https://doi.org/10.1038/s41586-021-03320-w>

Download citation

- Received: 14 August 2020
- Accepted: 01 February 2021
- Published: 17 March 2021
- Issue Date: 29 April 2021
- DOI: <https://doi.org/10.1038/s41586-021-03320-w>

Further reading

- [**DPP9 restrains NLRP1 activation**](#)
 - Stefan Bauernfried
 - & Veit Hornung

Nature Structural & Molecular Biology (2021)

- **DPP9 sequesters the C terminus of NLRP1 to repress inflammasome activation**

- L. Robert Hollingsworth
- , Humayun Sharif
- , Andrew R. Griswold
- , Pietro Fontana
- , Julian Mintseris
- , Kevin B. Dagbay
- , Joao A. Paulo
- , Steven P. Gygi
- , Daniel A. Bachovchin
- & Hao Wu

Nature (2021)

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Download PDF](#)

Associated Content

DPP9 sequesters the C terminus of NLRP1 to repress inflammasome activation

- L. Robert Hollingsworth
- Humayun Sharif
- Hao Wu

Nature Article 17 Mar 2021

DPP9 restrains NLRP1 activation

- Stefan Bauernfried
- Veit Hornung

Nature Structural & Molecular Biology News & Views 30 Mar 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03320-w>

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

- Article
- [Published: 17 March 2021](#)

DPP9 sequesters the C terminus of NLRP1 to repress inflammasome activation

- [L. Robert Hollingsworth](#) ORCID: [orcid.org/0000-0002-4130-3580^{1,2,3 na1}](https://orcid.org/0000-0002-4130-3580),
- [Humayun Sharif](#) ORCID: [orcid.org/0000-0003-1308-3442^{1,2 na1}](https://orcid.org/0000-0003-1308-3442),
- [Andrew R. Griswold](#) ORCID: [orcid.org/0000-0003-2957-2355^{4,5 na1}](https://orcid.org/0000-0003-2957-2355),
- [Pietro Fontana^{1,2}](#),
- [Julian Mintseris⁶](#),
- [Kevin B. Dagbay](#) ORCID: [orcid.org/0000-0001-7002-7207^{1,2}](https://orcid.org/0000-0001-7002-7207),
- [Joao A. Paulo](#) ORCID: [orcid.org/0000-0002-4291-413X⁶](https://orcid.org/0000-0002-4291-413X),
- [Steven P. Gygi](#) ORCID: [orcid.org/0000-0001-7626-0034⁶](https://orcid.org/0000-0001-7626-0034),
- [Daniel A. Bachoychin](#) ORCID: [orcid.org/0000-0001-8210-1662^{5,7}](https://orcid.org/0000-0001-8210-1662)&
- [Hao Wu](#) ORCID: [orcid.org/0000-0002-7281-8579^{1,2}](https://orcid.org/0000-0002-7281-8579)

[Nature](#) volume **592**, pages 778–783(2021) [Cite this article](#)

- 7631 Accesses
- 1 Citations
- 43 Altmetric
- [Metrics details](#)

Subjects

- [Cryoelectron microscopy](#)
- [Inflammasome](#)
- [NOD-like receptors](#)

Abstract

Nucleotide-binding domain and leucine-rich repeat pyrin-domain containing protein 1 (NLRP1) is an inflammasome sensor that mediates the activation of caspase-1 to induce cytokine maturation and pyroptosis^{1,2,3,4}. Gain-of-function mutations of *NLRP1* cause severe inflammatory diseases of the skin^{4,5,6}. NLRP1 contains a function-to-find domain that auto-proteolyses into noncovalently associated subdomains^{7,8,9}, and proteasomal degradation of the repressive N-terminal fragment of NLRP1 releases its inflammatory C-terminal fragment (NLRP1 CT)^{10,11}. Cytosolic dipeptidyl peptidases 8 and 9 (hereafter, DPP8/DPP9) both interact with NLRP1, and small-molecule inhibitors of DPP8/DPP9 activate NLRP1 by mechanisms that are currently unclear^{10,12,13,14}. Here we report cryo-electron microscopy structures of the human NLRP1–DPP9 complex alone and with Val-boroPro (VbP), an inhibitor of DPP8/DPP9. The structures reveal a ternary complex that comprises DPP9, full-length NLRP1 and the NLRP1 CT. The binding of the NLRP1 CT to DPP9 requires full-length NLRP1, which suggests that NLRP1 activation is regulated by the ratio of NLRP1 CT to full-length NLRP1. Activation of the inflammasome by ectopic expression of the NLRP1 CT is consistently rescued by co-expression of autoproteolysis-deficient full-length NLRP1. The N terminus of the NLRP1 CT inserts into the DPP9 active site, and VbP disrupts this interaction. Thus, VbP weakens the NLRP1–DPP9 interaction and accelerates degradation of the N-terminal fragment¹⁰ to induce inflammasome activation. Overall, these data demonstrate that DPP9 quenches low levels of NLRP1 CT and thus serves as a checkpoint for activation of the NLRP1 inflammasome.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

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: Structure of the NLRP1–DPP9 complex.

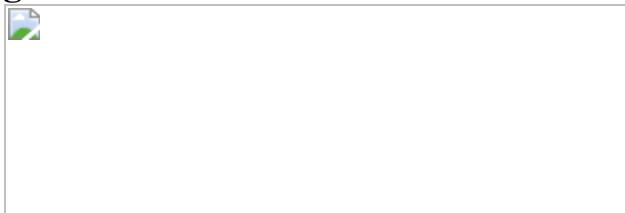


Fig. 2: Detailed interfaces in the NLRP1–DPP9 ternary complex and inhibition by VbP.



Fig. 3: Functional consequences of mutations that affect assembly of the NLRP1–DPP9 ternary complex.

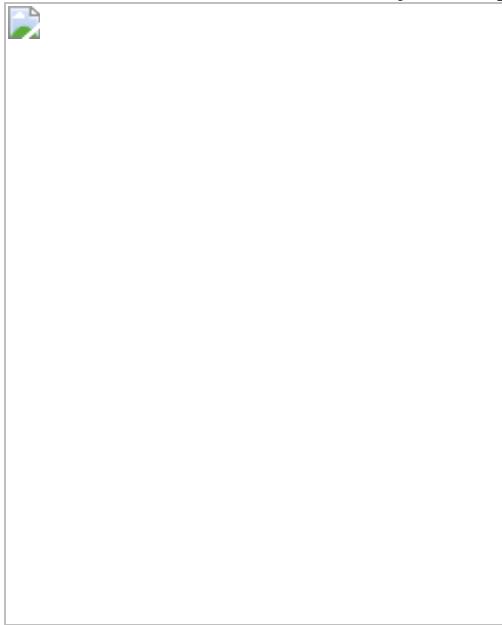
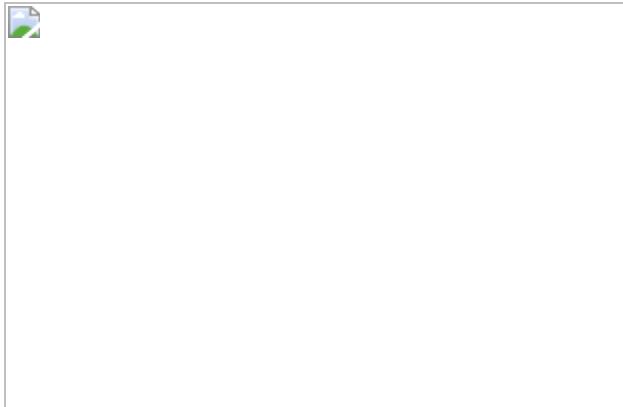


Fig. 4: Repression of activity of the NLRP1 CT inflammasome by formation of the ternary complex.



Data availability

Extended protein purification protocols are available on Protocols.io at <https://www.protocols.io/groups/hao-wu-lab>. Raw cryo-EM data are available on EMPIAR under the accession numbers EMPIAR-10594 (NLRP1-DPP9) and EMPIAR-10595 (NLRP1–DPP9–VbP). The cryo-EM maps are available on the Electron Microscopy Data Bank (EMDB) under the accession numbers [EMD-22074](#) (NLRP1–DPP9) and [EMD-22075](#) (NLRP1–DPP9–VbP). The atomic coordinates are available on the Protein Data Bank (PDB) under the accession numbers [6X6A](#) (NLRP1–DPP9) and [6X6C](#) (NLRP1–DPP9–VbP). Pymol session files and the image analysis macro are available from <https://doi.org/10.17605/OSF.IO/X7DV8>. All other data can be obtained from the corresponding authors upon reasonable request. [Source data](#) are provided with this paper.

References

1. 1.

Martinon, F., Burns, K. & Tschopp, J. The inflammasome: a molecular platform triggering activation of inflammatory caspases and processing of proIL- β . *Mol. Cell* **10**, 417–426 (2002).

[CAS](#) [PubMed](#) [Google Scholar](#)

2. 2.

Shen, C., Sharif, H., Xia, S. & Wu, H. Structural and mechanistic elucidation of inflammasome signaling by cryo-EM. *Curr. Opin. Struct. Biol.* **58**, 18–25 (2019).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

3. 3.

Mitchell, P. S., Sandstrom, A. & Vance, R. E. The NLRP1 inflammasome: new mechanistic insights and unresolved mysteries. *Curr. Opin. Immunol.* **60**, 37–45 (2019).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

4. 4.

Taabazuing, C. Y., Griswold, A. R. & Bachovchin, D. A. The NLRP1 and CARD8 inflammasomes. *Immunol. Rev.* **297**, 13–25 (2020).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

5. 5.

Soler, V. J. et al. Whole exome sequencing identifies a mutation for a novel form of corneal intraepithelial dyskeratosis. *J. Med. Genet.* **50**, 246–254 (2013).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

6. 6.

Zhong, F. L. et al. Germline NLRP1 mutations cause skin inflammatory and cancer susceptibility syndromes via inflammasome activation. *Cell* **167**, 187–202.e17 (2016).

[CAS](#) [PubMed](#) [Google Scholar](#)

7. 7.

D'Osualdo, A. et al. CARD8 and NLRP1 undergo autoproteolytic processing through a ZU5-like domain. *PLoS ONE* **6**, e27396 (2011).

[ADS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

8. 8.

Finger, J. N. et al. Autolytic proteolysis within the function to find domain (FIIND) is required for NLRP1 inflammasome activity. *J. Biol. Chem.* **287**, 25030–25037 (2012).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

9. 9.

Frew, B. C., Joag, V. R. & Mogridge, J. Proteolytic processing of Nlrp1b is required for inflammasome activity. *PLoS Pathog.* **8**, e1002659 (2012).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

10. 10.

Chui, A. J. et al. N-terminal degradation activates the NLRP1B inflammasome. *Science* **364**, 82–85 (2019).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

11. 11.

Sandstrom, A. et al. Functional degradation: a mechanism of NLRP1 inflammasome activation by diverse pathogen enzymes. *Science* **364**, eaau1330 (2019).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

12. 12.

Zhong, F. L. et al. Human DPP9 represses NLRP1 inflammasome and protects against autoinflammatory diseases via both peptidase activity and FIIND domain binding. *J. Biol. Chem.* **293**, 18864–18878 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

13. 13.

de Vasconcelos, N. M. et al. DPP8/DPP9 inhibition elicits canonical Nlrp1b inflammasome hallmarks in murine macrophages. *Life Sci. Alliance* **2**, e201900313 (2019).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

14. 14.

Okondo, M. C. et al. Inhibition of Dpp8/9 activates the Nlrp1b inflammasome. *Cell Chem. Biol.* **25**, 262–267.e5 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

15. 15.

Xia, S., Hollingsworth, L. R., IV & Wu, H. in *Cell Survival and Cell Death* (eds Newton, K. et al.) Ch. 7 (CSHL Press, 2019).

16. 16.

Okondo, M. C. et al. DPP8 and DPP9 inhibition induces pro-caspase-1-dependent monocyte and macrophage pyroptosis. *Nat. Chem. Biol.* **13**, 46–53 (2017).

[CAS](#) [PubMed](#) [Google Scholar](#)

17. 17.

Gai, K. et al. DPP8/9 inhibitors are universal activators of functional NLRP1 alleles. *Cell Death Dis.* **10**, 587 (2019).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

18. 18.

Griswold, A. R. et al. A chemical strategy for protease substrate profiling. *Cell Chem. Biol.* **26**, 901–907.e6 (2019).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

19. 19.

Geiss-Friedlander, R. et al. The cytoplasmic peptidase DPP9 is rate-limiting for degradation of proline-containing peptides. *J. Biol. Chem.* **284**, 27211–27219 (2009).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

20. 20.

Griswold, A. R. et al. DPP9's enzymatic activity and not its binding to CARD8 inhibits inflammasome activation. *ACS Chem. Biol.* **14**, 2424–2429 (2019).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

21. 21.

Chui, A. J. et al. Activation of the CARD8 inflammasome requires a disordered region. *Cell Rep.* **33**, 108264 (2020).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

22. 22.

Wang, R. et al. Autoinhibition of UNC5b revealed by the cytoplasmic domain structure of the receptor. *Mol. Cell* **33**, 692–703 (2009).

[CAS](#) [PubMed](#) [Google Scholar](#)

23. 23.

Wang, C., Yu, C., Ye, F., Wei, Z. & Zhang, M. Structure of the ZU5–ZU5–UPA–DD tandem of ankyrin-B reveals interaction surfaces necessary for ankyrin function. *Proc. Natl Acad. Sci. USA* **109**, 4822–4827 (2012).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

24. 24.

Ross, B. et al. Structures and mechanism of dipeptidyl peptidases 8 and 9, important players in cellular homeostasis and cancer. *Proc. Natl Acad. Sci. USA* **115**, E1437–E1445 (2018).

[CAS](#) [PubMed](#) [Google Scholar](#)

25. 25.

Roppongi, S. et al. Crystal structures of a bacterial dipeptidyl peptidase IV reveal a novel substrate recognition mechanism distinct from that of mammalian orthologues. *Sci. Rep.* **8**, 2714 (2018).

[ADS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

26. 26.

Grandemange, S. et al. A new autoinflammatory and autoimmune syndrome associated with NLRP1 mutations: NAIAD (*NLRP1*-associated autoinflammation with arthritis and dyskeratosis). *Ann. Rheum. Dis.* **76**, 1191–1198 (2017).

[CAS](#) [PubMed](#) [Google Scholar](#)

27. 27.

Robert Hollingsworth, L. et al. Mechanism of filament formation in UPA-promoted CARD8 and NLRP1 inflammasomes. *Nat. Commun.* **12**, 189 (2021).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

28. 28.

Gong, Q. et al. Structural basis for distinct inflammasome complex assembly by human NLRP1 and CARD8. *Nat. Commun.* **12**, 188 (2021).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

29. 29.

Bachmair, A., Finley, D. & Varshavsky, A. In vivo half-life of a protein is a function of its amino-terminal residue. *Science* **234**, 179–186 (1986).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

30. 30.

Baker, R. T. & Varshavsky, A. Inhibition of the N-end rule pathway in living cells. *Proc. Natl Acad. Sci. USA* **88**, 1090–1094 (1991).

[ADS](#) [CAS](#) [PubMed](#) [Google Scholar](#)

31. 31.

Nabet, B. et al. The dTAG system for immediate and target-specific protein degradation. *Nat. Chem. Biol.* **14**, 431–441 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

32. 32.

Johnson, D. C. et al. DPP8/DPP9 inhibitor-induced pyroptosis for treatment of acute myeloid leukemia. *Nat. Med.* **24**, 1151–1156 (2018).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

33. 33.

Mintseris, J. & Gygi, S. P. High-density chemical cross-linking for modeling protein interactions. *Proc. Natl Acad. Sci. USA* **117**, 93–102 (2020).

[CAS](#) [PubMed](#) [Google Scholar](#)

34. 34.

Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

35. 35.

Naydenova, K. & Russo, C. J. Measuring the effects of particle orientation to improve the efficiency of electron cryomicroscopy. *Nat. Commun.* **8**, 629 (2017).

[ADS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

36. 36.

Morin, A. et al. Collaboration gets the most out of software. *eLife* **2**, e01456 (2013).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

37. 37.

Towns, J. et al. XSEDE: accelerating scientific discovery. *Comput. Sci. Eng.* **16**, 62–74 (2014).

[Google Scholar](#)

38. 38.

Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

39. 39.

Rohou, A. & Grigorieff, N. CTFFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

40. 40.

Su, M. goCTF: geometrically optimized CTF determination for single-particle cryo-EM. *J. Struct. Biol.* **205**, 22–29 (2019).

[PubMed](#) [Google Scholar](#)

41. 41.

Wagner, T. et al. SPHIRE-crYOLO is a fast and accurate fully automated particle picker for cryo-EM. *Commun. Biol.* **2**, 218 (2019).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

42. 42.

Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* **7**, e42166 (2018).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

43. 43.

Zivanov, J., Nakane, T. & Scheres, S. H. W. Estimation of high-order aberrations and anisotropic magnification from cryo-EM data sets in

RELION-3.1. IUCrJ **7**, 253–267 (2020).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

44. 44.

Kucukelbir, A., Sigworth, F. J. & Tagare, H. D. Quantifying the local resolution of cryo-EM density maps. *Nat. Methods* **11**, 63–65 (2014).

[CAS](#) [PubMed](#) [Google Scholar](#)

45. 45.

Jacobson, M. P., Friesner, R. A., Xiang, Z. & Honig, B. On the role of the crystal environment in determining protein side-chain conformations. *J. Mol. Biol.* **320**, 597–608 (2002).

[CAS](#) [PubMed](#) [Google Scholar](#)

46. 46.

Huang, M. et al. Structural and biochemical mechanisms of NLRP1 inhibition by DPP9. *Nature* <https://doi.org/10.1038/s41586-021-03320-w> (2021).

47. 47.

Goddard, T. D., Huang, C. C. & Ferrin, T. E. Visualizing density maps with UCSF Chimera. *J. Struct. Biol.* **157**, 281–287 (2007).

[CAS](#) [PubMed](#) [Google Scholar](#)

48. 48.

Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66**, 486–501 (2010).

[CAS](#) [Google Scholar](#)

49. 49.

Klaholz, B. P. Deriving and refining atomic models in crystallography and cryo-EM: the latest Phenix tools to facilitate structure analysis. *Acta Crystallogr. D* **75**, 878–881 (2019).

[CAS](#) [Google Scholar](#)

50. 50.

Díaz, B. H. R. *Improvement of Protein Crystal Diffraction Using Post-Crystallization Methods: Infrared Laser Radiation Controls Crystal Order*. PhD thesis, Universität Hamburg (2018).

51. 51.

Robertson, M. J., van Zundert, G. C. P., Borrelli, K. & Skiniotis, G. GemSpot: a pipeline for robust modeling of ligands into cryo-EM maps. *Structure* **28**, 707–716.e3 (2020).

[CAS](#) [PubMed](#) [Google Scholar](#)

52. 52.

Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

53. 53.

Goddard, T. D. et al. UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.* **27**, 14–25 (2018).

[CAS](#) [PubMed](#) [Google Scholar](#)

54. 54.

Schrödinger. The PyMOL Molecular Graphics System, version 2.0, <https://pymol.org/2/> (2020).

55. 55.

Schrödinger. Maestro, release 2020-4, <https://www.schrodinger.com/products/maestro> (2020).

[Download references](#)

Acknowledgements

We thank T.-m. Fu of the laboratory of H.W. for initiating an apo-NLRP1 project; all members of the laboratories of H.W. and D.A.B. for helpful discussions; J. Chai for sharing their crystal structure of the FIIND of rat NLRP1 before submission; M. Su for his advice on running goCTF and CTF estimation for tilt datasets; M. Ericsson at the HMS electron microscopy facility for advice and training; R. Walsh, S. Sterling and S. Rawson at the Harvard Cryo-EM Center for Structural Biology for cryo-EM training and productive consultation; G. Skiniotis and M. Robertson for sharing their GemSpot script and for troubleshooting advice; T. Edwards and the National Cancer Institute's National Cryo-EM Facility at the Frederick National Laboratory for Cancer Research under contract HSSN261200800001E; and J. Myers and the Pacific Northwest Center for Cryo-EM at Oregon Health & Science University for screening and preliminary dataset collection, under the NIH grant U24GM129547 and accessed through EMSL (grid.436923.9), a DOE Office of Science User Facility sponsored by the Office of Biological and Environmental Research. The pcDNA3.1 LIC 6A and pcDNA 3.1 LIC 6D plasmids were gifts from S. Gradia (Addgene plasmids no. 30124 and no. 30127, respectively). The pLEX_305-N-dTAG plasmid was a gift from J. Bradner and B. Nabet (Addgene plasmid no. 91797). The pLEX_307 plasmid was a gift from D. Root (Addgene plasmid no. 41392). Our research used software and computing support at SBGrid and the Extreme Science and Engineering Discovery Environment (XSEDE) Bridges at the Pittsburgh Supercomputing Center, through allocation TG-MCB190086. This work was supported by the National Institutes of Health and the Cancer Research

Institute (DP1 HD087988 and R01 AI124491 to H.W., T32-GM007726 to L.R.H., the CRI Irvington Postdoctoral Fellowship to P.F., R01GM132129 to J.A.P., R01GM67945 to S.P.G., T32 GM007739-Andersen and F30 CA243444 to A.R.G. and the Memorial Sloan Kettering Cancer Center Core Grant P30 CA008748 and R01 AI137168 to D.A.B.). We apologize to authors whose work could not be cited because of space limitation.

Author information

Author notes

1. These authors contributed equally: L. Robert Hollingsworth, Humayun Sharif, Andrew R. Griswold

Affiliations

1. Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA, USA

L. Robert Hollingsworth, Humayun Sharif, Pietro Fontana, Kevin B. Dagbay & Hao Wu

2. Program in Cellular and Molecular Medicine, Boston Children's Hospital, Boston, MA, USA

L. Robert Hollingsworth, Humayun Sharif, Pietro Fontana, Kevin B. Dagbay & Hao Wu

3. Program in Biological and Biomedical Sciences, Harvard Medical School, Boston, MA, USA

L. Robert Hollingsworth

4. Tri-Institutional MD–PhD Program, Weill Cornell Medical College, Rockefeller University and Memorial Sloan Kettering Cancer Center, New York, NY, USA

Andrew R. Griswold

5. Pharmacology Program, Weill Cornell Graduate School of Medical Sciences and Memorial Sloan Kettering Cancer Center, New York, NY, USA

Andrew R. Griswold & Daniel A. Bachovchin

6. Department of Cell Biology, Harvard Medical School, Harvard University, Boston, MA, USA

Julian Mintseris, Joao A. Paulo & Steven P. Gygi

7. Chemical Biology Program, Memorial Sloan Kettering Cancer Center, New York, NY, USA

Daniel A. Bachovchin

Authors

1. L. Robert Hollingsworth
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Humayun Sharif
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Andrew R. Griswold
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Pietro Fontana
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Julian Mintseris

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Kevin B. Dagbay

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Joao A. Paulo

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Steven P. Gygi

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. Daniel A. Bachoychin

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. Hao Wu

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Contributions

L.R.H., H.S. and H.W. conceived the NLRP1–DPP9 complex study. L.R.H. designed constructs with input from H.S. L.R.H., H.S., P.F. and K.B.D. carried out preliminary expression and purification studies. L.R.H. purified the complexes. H.S. and L.R.H. made cryo-EM grids for data collection. L.R.H. screened cryo-EM grids and collected cryo-EM data. H.S. and L.R.H. analysed cryo-EM data. H.S. performed model building and

refinement. L.R.H. and H.S. designed mutants for in vitro and cell-based assays. L.R.H. and A.R.G. cloned mutants for functional study. A.R.G. performed all cell-based assays, peptide mass spectrometry and chemical enrichment of protease substrates analysis under the supervision of D.A.B. J.M. and J.A.P. performed crosslinking mass spectrometry and data analysis under the supervision of S.P.G. A.R.G., L.R.H., H.S., H.W. and D.A.B. designed the experiments. L.R.H., H.S., A.R.G., H.W. and D.A.B. wrote the manuscript with input from all of the other authors.

Corresponding authors

Correspondence to [Daniel A. Bachovchin](#) or [Hao Wu](#).

Ethics declarations

Competing interests

H.W. is a co-founder of Ventus Therapeutics. The other authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Edward Miao 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 Structural determination of the NLRP1–DPP9 complex.](#)

a, Purification of the NLRP1–DPP9 complex by ion-exchange chromatography. The ternary complex peak is shaded in green and labelled with an arrow. **b**, A representative (of >1,000 images) cryo-EM micrograph. **c**, Representative 2D class averages. **d**, Workflow for the determination of the structure of the NLRP1–DPP9 complex. **e**, Map–map and map–model FSC curves. **f**, Local-resolution distribution of the final map calculated with ResMap⁴⁴. [Source data](#)

Extended Data Fig. 2 Crosslinking mass spectrometry analysis of the NLRP1–DPP9 complex.

a, Summary of BS3 crosslinking between DPP9 and NLRP1. High-confidence crosslinked peptides are displayed, and residue ranges are labelled and colour-coded by domain. Crosslinked lysine pairs are indicated in red text. $\text{Ca}-\text{Ca}$ distances between lysine residues (in red) interpreted by the final NLRP1–DPP9 model are shown, along with the figure panels that show them in detail. All detected peptide pairs are tabulated in Source Data. **b**, Overview of BS3-mediated crosslinks. **c–h**, Close-up views highlighting crosslinked lysine pairs (red) interpreted by the final NLRP1–DPP9 model. [Source data](#)

Extended Data Fig. 3 Sequence and structural analysis of FIIND.

a, ClustalW multiple sequence alignment between human NLRP1 (hNLRP1), mouse NLRP1 (mNLRP1, different isoforms) and rat NLRP1 (rNLRP1, different isoforms). COP, Copenhagen; ZUC, Zucker; LEW, Lewis; SD, Sprague Dawley; and CDF, Fischer. Secondary structures and residue numbers are denoted on the basis of the human FIIND^A structure in the NLRP1–DPP9 ternary complex. Interface residues in the NLRP1–DPP9 complex are annotated with asterisks, and residues in the catalytic triad (H1186, E1195 and S1213) are boxed in black. **b**, FIIND^A overview with ZU5 (blue) and UPA (light pink) subdomains. The catalytic triad residues (H1186, E1195 and S1213) are shown in sticks. **c**, Topology of the FIIND with secondary structures labelled. **d**, Superimposition of FIIND^A onto the UPA^B. NLRP1^B must be free NLRP1 CT, because a ZU5 subdomain at site

B would have clashed with ZU5 and UPA at site A and with DPP9. **e**, The ZU5^A–UPA^A–UPA^B module that binds DPP9. UPA^A and UPA^B interact with each other in a front-to-back manner, with only a 9° rotation between them. **f**, Altered conformation of the UPA^B N terminus that binds in the DPP9 active-site tunnel in comparison to UPA^A in a complete FIIND^A.

Extended Data Fig. 4 Structural determination of the NLRP1–DPP9 complex with VbP.

a, Purification of the NLRP1–DPP9 complex in the presence of VbP by ion-exchange chromatography. The ternary complex peak is shaded in green and labelled with an arrow. **b**, A representative (of >1,000 images) cryo-EM micrograph. **c**, Representative 2D class averages. **d**, Workflow for the determination of the structure of the NLRP1–DPP9–VbP complex. **e**, Map–map and map–model FSC curves. **f**, Local-resolution distribution of the final map calculated with ResMap⁴⁴. [Source data](#)

Extended Data Fig. 5 VbP interactions in the DPP9 active site and comparison to a DPP substrate and NLRP1.

a, Schematic of covalent linkage between the S730 of DPP9 and VbP. **b**, Fit of VbP into the cryo-EM density. VbP is shown in stick, with carbon atoms in light brown. The charged amino group of VbP interacts with the DPP9 EE-loop (which also coordinates a substrate N terminus), and the carbonyl oxygen of VbP interacts with R133 of the R helix. The covalent linkage of VbP with S730 (the catalytic serine) is displayed. **c**, Structural alignment of the VbP-bound DPP9 model (green) and the crystal structure of bacterial DPP4 bound to the substrate Ile-Pro (PDB code 5YP3) (orange)²⁵. VbP assumes a pose that is notable similar to a model substrate. **d**, NLRP1 CT–DPP9 complex, in which DPP9 is coloured by Ca–Ca distances between NLRP1-bound and VbP-bound structures, as indicated. A distance scale bar is shown. VbP is displayed in sticks to mark the active site, with carbon atoms in green, oxygen atoms in red, nitrogen atoms in blue and boron atoms in orange. UPA of NLRP1 CT is shown in magenta.

Extended Data Fig. 6 Lack of cleavage of intact NLRP1 CT, but the cleavage of its isolated N-terminal peptide, by DPP9.

a, N-terminal sequencing of the purified NLRP1–DPP9 complex, showing that the NLRP1 CT is not cleaved by the co-expressed DPP9. **b**, Chemical enrichment of protease substrates assay, showing that DPP9 does not cleave NLRP1 CT. In brief, wild-type NLRP1 expressed in *DPP8 DPP9* double-knockout HEK293T cells was incubated with PBS or recombinant DPP9 before labelling with a 2PCA–biotin probe (which selectively biotinylates free N termini, except for those with a proline in the second position), followed by capture of biotinylated proteins. The inputs and the eluents were analysed by immunoblots using anti-NLRP1 CT (full-length NLRP1 and NLRP1 CT), anti-GAPDH and anti-streptavidin (biotinylated proteins) antibodies. DPP9 treatment did not increase the biotinylation of NLRP1 CT, as would be expected after the removal of the N-terminal Ser-Pro dipeptide. **c**, Evidence of cleavage of the isolated 15-residue N-terminal peptide in NLRP1 CT by recombinant DPP9 from mass spectrometry analysis. **d**, Inhibition of DPP9 catalytic activity against Ala-Pro-AMC by the isolated NLRP1 CT peptide. **e**, Schematic illustrates the ability of DPP9 to cleave an isolated UPA N-terminal peptide, but not dipeptides from an intact NLRP1 CT. **f**, Comparison of the binding modes of the UPA^B N-terminal peptide in the NLRP1–DPP9 complex and the Ile-Pro dipeptide in an acyl-enzyme intermediate²⁵. **g**, Theoretical dipeptide cleavage does not dampen activity of the NLRP1 inflammasome by LDH release or inflammasome signalling. $n = 3$ independent biological replicates. Data are mean \pm s.e.m. Anti-Flag (NLRP1 CT), anti-GSDMD and anti-GAPDH antibodies were used in the immunoblots. p30, GSDMD N-terminal fragment from caspase-1 cleavage. **h**, Theoretical dipeptide cleavage does not dampen activity of the NLRP1 inflammasome by formation of ASC specks. $n = 10$ quantified fields of view. Data are mean \pm s.e.m. Right, representative superimposed images of nuclei (blue), RFP (red) and GFP–ASC (green). All data are representative of >2 independent experiments. $*P < 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$ by unpaired two-sided *t*-test. Exact *P* values are provided in Source Data. [Source data](#)

Extended Data Fig. 7 Mutational analysis of the interactions in the NLRP1–DPP9 ternary complex.

a, Disorder-to-order transition of several DPP9 surface loops from the isolated DPP9 crystal structure (PDB code 6EOQ)²⁴ to the NLRP1-bound DPP9 cryo-EM structure. **b**, Genomic confirmation of DPP8 knockout generated in *DPP9*-knockout HEK293T cells that stably express Cas9²⁰, to create *DPP8 DPP9* double-knockout HEK293T cells. The single-guide RNA (sgRNA) sequence is highlighted. **c**, Immunoblots of the input lysates for the Flag co-immunoprecipitation with wild-type or mutant DPP9 and wild-type NLRP1–Flag, related to Fig. 2h. Anti-DPP9, anti-NLRP1 (full-length NLRP1 and NLRP1 CT) and anti-GAPDH antibodies were used in the immunoblots. **d**, Cleavage rate of a model DPP9 substrate, Gly-Pro-AMC, by wild-type DPP9 and its structure-guided mutants. Only the catalytically dead mutant DPP9(S730A) disrupts catalytic activity and sensitivity to VbP. $n = 3$ technical replicates. Data are mean \pm s.e.m. **e**, Immunoblots of the input lysates for the Flag co-immunoprecipitation with wild-type or mutant NLRP1–Flag and wild-type DPP9, related to Fig. 2i. Anti-DPP9, anti-NLRP1 (full-length NLRP1 and NLRP1 CT) and anti-GAPDH antibodies were used in the immunoblots. All data are representative of >2 independent experiments.

Extended Data Fig. 8 The ZU5 domain and DPP9 sterically hinder UPA polymerization.

a, Modelling of a FIIND polymer using the observed UPA^A–UPA^B relationship. Adjacent ZU5 molecules would clash, which suggests that UPA polymerization cannot occur in complete FIIND. **b**, Modelled recruitment of free UPA adjacent to UPA^A and UPA^B in the ternary complex with DPP9. The additional UPA subdomain next to FIIND^A clashes with the ZU5 subdomain, and the additional UPA next to UPA^B clashes with both DPP9 monomers in the complex, which suggests that DPP9 inhibits UPA oligomerization. **c**, A modelled UPA oligomer on the basis of the near front-to-back interaction in the NLRP1–DPP9 ternary complex. In the model, the N-terminal tails of free UPAs are shown in the UPA^A (pink) or UPA^B (magenta) conformation in complex with DPP9 but,

in reality, this conformation is likely to be different. **d**, Modelling of a UPA oligomer formed on one side of a NLRP1 FIIND^A–NLRP1 CT^B complex. The NLRP1 FIIND^A–NLRP1 CT^B binary complex can polymerize with freed NLRP1 CT. In **a–d**, DPP9 is coloured in green, and NLRP1 domains are coloured as indicated.

Extended Data Fig. 9 VbP displaces NLRP1 from DPP9 in vitro and in cells.

a, Schematic of the on-bead displacement experiment. The ternary complex is expressed in HEK293T cells, which are then lysed and incubated with Flag beads. Once bound, beads are split equally and washed with compounds or DMSO. The remainder of the protein is eluted off of the beads. MeBS, bestatin methyl ester. **b**, Two structurally distinct DPP9 inhibitors (VbP and 8J) displace DPP9 from NLRP1(S1213A) by the on-bead displacement assay. Anti-Flag (NLRP1(S1213A)), anti-MYC (NLRP1 CT) and anti-V5 (DPP9) antibodies were used in the immunoblots. Representative of two independent experiments. **c**, Schematic of the dTAG experiment. FKBP12 with the F36V mutation (dTAG) is fused to the N terminus of NLRP1. The dTAG13 ligand recruits an E3 ligase to FKBP12(F36V), leading to its ubiquitination and N-terminal degradation of the fusion protein. NLRP1 CT (UPA–CARD) that results from FIIND autoprocessing are released to assemble the inflammasome. **d**, NLRP1(FIIND/S1213A) expression in reconstituted HEK293T inflammasome system rescues GSDMD cleavage resulting from dTAG13-induced NLRP1 degradation. VbP prevents GSDMD rescue without inducing additional NLRP1 degradation. Anti-HA (dTAG–NLRP1, and dTAG–NLRP1 N-terminal fragment), anti-Flag (NLRP1(FIIND/S1213A)), anti-GSDMD and anti-GAPDH antibodies were used in the immunoblots. p30, GSDMD N-terminal fragment from caspase-1 cleavage. Representative of two independent experiments.

Extended Data Table 1 Cryo-EM data collection and refinement statistics of NLRP1–DPP9 structures

[Full size table](#)

Supplementary information

Supplementary Information

This file contains the uncropped scans.

Reporting Summary

Source data

Source Data Fig. 3

Source Data Fig. 4

Source Data Extended Data Fig. 1

Source Data Extended Data Fig. 2

Source Data Extended Data Fig. 4

Source Data Extended Data Fig. 6

Rights and permissions

Reprints and Permissions

About this article



Check for
updates

Cite this article

Hollingsworth, L.R., Sharif, H., Griswold, A.R. *et al.* DPP9 sequesters the C terminus of NLRP1 to repress inflammasome activation. *Nature* **592**, 778–783 (2021). <https://doi.org/10.1038/s41586-021-03350-4>

[Download citation](#)

- Received: 13 August 2020
- Accepted: 09 February 2021
- Published: 17 March 2021
- Issue Date: 29 April 2021
- DOI: <https://doi.org/10.1038/s41586-021-03350-4>

Further reading

- [**DPP9 restrains NLRP1 activation**](#)
 - Stefan Bauernfried
 - & Veit Hornung

Nature Structural & Molecular Biology (2021)

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

Associated Content

Structural and biochemical mechanisms of NLRP1 inhibition by DPP9

- Menghang Huang
- Xiaoxiao Zhang
- Jijie Chai

Nature Article 17 Mar 2021

DPP9 restrains NLRP1 activation

- Stefan Bauernfried
- Veit Hornung

Nature Structural & Molecular Biology News & Views 30 Mar 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03350-4>

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

- Article
- [Published: 21 April 2021](#)

High CO₂ levels drive the TCA cycle backwards towards autotrophy

- [Lydia Steffens](#) [ORCID: orcid.org/0000-0002-0546-2736](#)¹ natl,
- [Eugenio Pettinato](#) [ORCID: orcid.org/0000-0002-3174-020X](#)¹ natl,
- [Thomas M. Steiner](#) [ORCID: orcid.org/0000-0002-8783-8118](#)² natl,
- [Achim Mall](#)^{3,4},
- [Simone König](#)⁵,
- [Wolfgang Eisenreich](#) [ORCID: orcid.org/0000-0002-9832-8279](#)² &
- [Ivan A. Berg](#) [ORCID: orcid.org/0000-0002-1864-4969](#)¹

[Nature](#) volume **592**, pages 784–788(2021) [Cite this article](#)

- 8057 Accesses
- 1 Citations
- 330 Altmetric
- [Metrics details](#)

Subjects

- [Bacterial physiology](#)
- [Carbon cycle](#)
- [Enzymes](#)
- [Metabolic pathways](#)

Abstract

It has recently been shown that in anaerobic microorganisms the tricarboxylic acid (TCA) cycle, including the seemingly irreversible citrate synthase reaction, can be reversed and used for autotrophic fixation of carbon^{1,2}. This reversed oxidative TCA cycle requires ferredoxin-dependent 2-oxoglutarate synthase instead of the NAD-dependent dehydrogenase as well as extremely high levels of citrate synthase (more than 7% of the proteins in the cell). In this pathway, citrate synthase replaces ATP-citrate lyase of the reductive TCA cycle, which leads to the spending of one ATP-equivalent less per one turn of the cycle. Here we show, using the thermophilic sulfur-reducing delta-proteobacterium *Hippea maritima*, that this route is driven by high partial pressures of CO₂. These high partial pressures are especially important for the removal of the product acetyl coenzyme A (acetyl-CoA) through reductive carboxylation to pyruvate, which is catalysed by pyruvate synthase. The reversed oxidative TCA cycle may have been functioning in autotrophic CO₂ fixation in a primordial atmosphere that is assumed to have been rich in CO₂.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

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 roTCA cycle in *H. maritima*.

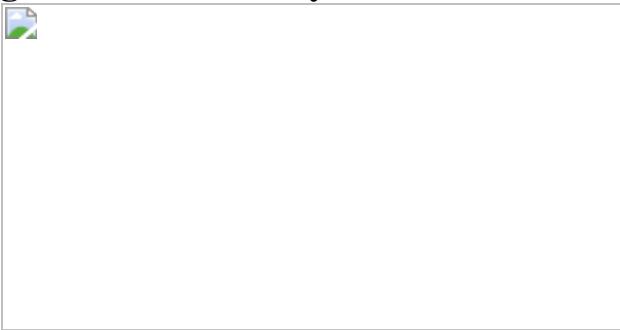


Fig. 2: Incorporation of the [1-¹³C]glutamate in *H. maritima* via the roTCA cycle.

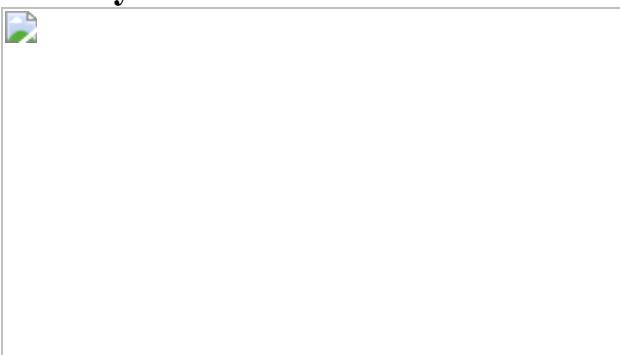
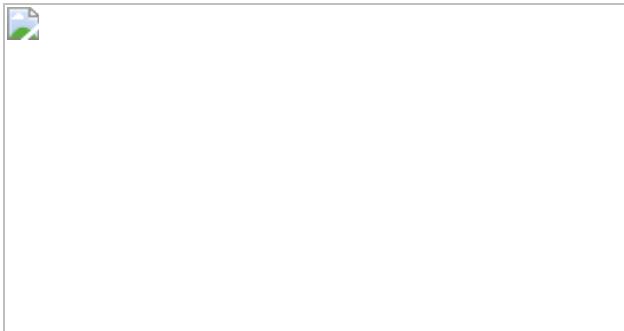


Fig. 3: Growth of different bacteria depending on the concentration of CO₂.



Data availability

All data generated in this manuscript are included within the paper (and its [Supplementary Information](#)). The raw data are presented in the manuscript and/or available from the corresponding authors upon reasonable request. For any further inquiries about our work please contact the corresponding authors. [Source data](#) are provided with this paper.

References

1. 1.

Nunoura, T. et al. A primordial and reversible TCA cycle in a facultatively chemolithoautotrophic thermophile. *Science* **359**, 559–563 (2018).

[ADS](#) [CAS](#) [Google Scholar](#)

2. 2.

Mall, A. et al. Reversibility of citrate synthase allows autotrophic growth of a thermophilic bacterium. *Science* **359**, 563–567 (2018).

[ADS](#) [CAS](#) [Google Scholar](#)

3. 3.

Fuchs, G. Alternative pathways of carbon dioxide fixation: insights into the early evolution of life? *Annu. Rev. Microbiol.* **65**, 631–658 (2011).

[CAS](#) [Google Scholar](#)

4. 4.

Berg, I. A. Ecological aspects of the distribution of different autotrophic CO₂ fixation pathways. *Appl. Environ. Microbiol.* **77**, 1925–1936 (2011).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

5. 5.

Song, Y. et al. Functional cooperation of the glycine synthase-reductase and Wood–Ljungdahl pathways for autotrophic growth of *Clostridium drakei*. *Proc. Natl Acad. Sci. USA* **117**, 7516–7523 (2020).

[CAS](#) [Google Scholar](#)

6. 6.

Könneke, M. et al. Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO₂ fixation. *Proc. Natl Acad. Sci. USA* **111**, 8239–8244 (2014).

[ADS](#) [Google Scholar](#)

7. 7.

Guynn, R. W., Gelberg, H. J. & Veech, R. L. Equilibrium constants of the malate dehydrogenase, citrate synthase, citrate lyase, and acetyl coenzyme A hydrolysis reactions under physiological conditions. *J. Biol. Chem.* **248**, 6957–6965 (1973).

[CAS](#) [Google Scholar](#)

8. 8.

Bisswanger, H. *Enzyme Kinetics: Principles and Methods* (Wiley, 2008).

9. 9.

Supek, F. & Vlahovicek, K. Comparison of codon usage measures and their applicability in prediction of microbial gene expressivity. *BMC Bioinformatics* **6**, 182 (2005).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

10. 10.

Supek, F. & Vlahovicek, K. INCA: synonymous codon usage analysis and clustering by means of self-organizing map. *Bioinformatics* **20**, 2329–2330 (2004).

[CAS](#) [Google Scholar](#)

11. 11.

Miroshnichenko, M. L., Rainey, F. A., Rhode, M. & Bonch-Osmolovskaya, E. A. *Hippea maritima* gen. nov., sp. nov., a new genus of thermophilic, sulfur-reducing bacterium from submarine hot vents. *Int. J. Syst. Bacteriol.* **49**, 1033–1038 (1999).

[CAS](#) [Google Scholar](#)

12. 12.

Schauder, R., Widdel, F. & Fuchs, G. Carbon assimilation pathways in sulfate-reducing bacteria. II. Enzymes of a reductive citric acid cycle in the autotrophic *Desulfobacter hydrogenophilus*. *Arch. Microbiol.* **148**, 218–225 (1987).

[CAS](#) [Google Scholar](#)

13. 13.

Bennett, B. D. et al. Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. *Nat. Chem. Biol.* **5**, 593–599 (2009).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

14. 14.

Greene, A. C. in *The Prokaryotes* (eds Rosenberg, E. et al.) 135–142 (Springer, 2014).

15. 15.

Dahle, H. et al. Energy landscapes in hydrothermal chimneys shape distributions of primary producers. *Front. Microbiol.* **9**, 1570 (2018).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

16. 16.

Zhao, W. *Diversity and Potential Geochemical Functions of Prokaryotes in Hot Springs of the Uzon Caldera, Kamchatka*. PhD thesis, Univ. Georgia, Athens (2008).

17. 17.

Obzhirov, A. I. Chemistry of free and dissolved gases of Matupit Bay, Rabaul Caldera, Papua New Guinea. *Geo-Mar. Lett.* **12**, 54–59 (1992).

[ADS](#) [CAS](#) [Google Scholar](#)

18. 18.

Zhang, T., Shi, X.-C., Ding, R., Xu, K. & Tremblay, P.-L. The hidden chemolithoautotrophic metabolism of *Geobacter sulfurreducens* uncovered by adaptation to formate. *ISME J.* **14**, 2078–2089 (2020).

[CAS](#) [Google Scholar](#)

19. 19.

Spieck, E. et al. Extremophilic nitrite-oxidizing *Chloroflexi* from Yellowstone hot springs. *ISME J.* **14**, 364–379 (2020).

[CAS](#) [Google Scholar](#)

20. 20.

Slobodkin, A. et al. Genomic insights into the carbon and energy metabolism of a thermophilic deep-sea bacterium *Deferrribacter autotrophicus* revealed new metabolic traits in the phylum Deferribacteres. *Genes* **10**, 849 (2019).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

21. 21.

Hua, Z. S. et al. Genomic inference of the metabolism and evolution of the archaeal phylum Aigarchaeota. *Nat. Commun.* **9**, 2832 (2018).

[ADS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

22. 22.

Berg, I. A. et al. Autotrophic carbon fixation in archaea. *Nat. Rev. Microbiol.* **8**, 447–460 (2010).

[CAS](#) [Google Scholar](#)

23. 23.

Zahnle, K. et al. Emergence of a habitable planet. *Space Sci. Rev.* **129**, 35–78 (2007).

[ADS](#) [CAS](#) [Google Scholar](#)

24. 24.

German, C. R. & Von Damm, K. L. in *Treatise on Geochemistry* (eds Turekian, K. K. & Holland, H. D.) 18–222 (Elsevier, 2006).

25. 25.

Pedersen, R. B. et al. Discovery of a black smoker vent field and vent fauna at the Arctic Mid-Ocean Ridge. *Nat. Commun.* **1**, 126 (2010).

[ADS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

26. 26.

Lupton, J. et al. Submarine venting of liquid carbon dioxide on a Mariana Arc volcano. *Geochem. Geophys. Geosyst.* **7**, Q08007 (2006).

[ADS](#) [Google Scholar](#)

27. 27.

Ajo-Franklin, J., Voltolini, M., Molins, S. & Yang, L. in *Geological Carbon Storage: Subsurface Seals and Caprock Integrity* (eds Vialle, S. et al.) 187–206 (Wiley, 2019).

28. 28.

Probst, A. J. et al. Lipid analysis of CO₂-rich subsurface aquifers suggests an autotrophy-based deep biosphere with lysolipids enriched in CPR bacteria. *ISME J.* **14**, 1547–1560 (2020).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

29. 29.

Martin, W. F. Older than genes: the acetyl CoA pathway and origins. *Front. Microbiol.* **11**, 817 (2020).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

30. 30.

Thauer, R. K., Kaster, A. K., Seedorf, H., Buckel, W. & Hedderich, R. Methanogenic archaea: ecologically relevant differences in energy conservation. *Nat. Rev. Microbiol.* **6**, 579–591 (2008).

[CAS](#) [Google Scholar](#)

31. 31.

Schuchmann, K. & Müller, V. Autotrophy at the thermodynamic limit of life: a model for energy conservation in acetogenic bacteria. *Nat. Rev. Microbiol.* **12**, 809–821 (2014).

[CAS](#) [Google Scholar](#)

32. 32.

Simon, E. J. & Shemin, D. The preparation of S-succinyl coenzyme A. *J. Am. Chem. Soc.* **75**, 2520 (1953).

[CAS](#) [Google Scholar](#)

33. 33.

Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254 (1976).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

34. 34.

Eylert, E. et al. Carbon metabolism of *Listeria monocytogenes* growing inside macrophages. *Mol. Microbiol.* **69**, 1008–1017 (2008).

[CAS](#) [Google Scholar](#)

35. 35.

Ahmed, Z. et al. ‘Isotopo’ a database application for facile analysis and management of mass isotopomer data. *Database* **2014**, bau077 (2014).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

36. 36.

Ausubel, F. M. et al. *Current Protocols in Molecular Biology* (Wiley, 1987).

37. 37.

Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).

[CAS](#) [PubMed](#) [Google Scholar](#)

38. 38.

Riddles, P. W., Blakeley, R. L. & Zerner, B. Ellman’s reagent: 5,5'-dithiobis(2-nitrobenzoic acid)—a reexamination. *Anal. Biochem.* **94**, 75–81 (1979).

[CAS](#) [Google Scholar](#)

39. 39.

Bergmeyer, H. U. Neue Werte für die molaren Extinktions-Koeffizienten von NADH und NADPH zum Gebrauch im Routine-Laboratorium. *Z. Klin. Chem. Klin. Biochem.* **13**, 507–508 (1975).

[CAS](#) [Google Scholar](#)

40. 40.

Hügler, M., Huber, H., Molyneaux, S. J., Vetriani, C. & Sievert, S. M. Autotrophic CO₂ fixation via the reductive tricarboxylic acid cycle in

different lineages within the phylum Aquificae: evidence for two ways of citrate cleavage. *Environ. Microbiol.* **9**, 81–92 (2007).

[Google Scholar](#)

41. 41.

Flint, D. H. Initial kinetic and mechanistic characterization of *Escherichia coli* fumarase A. *Arch. Biochem. Biophys.* **311**, 509–516 (1994).

[CAS](#) [Google Scholar](#)

42. 42.

Dawson, R. M. C., Elliott, D. C., Elliott, W. H. & Jones, K. M. *Data for Biochemical Research* (Clarendon, 1986).

43. 43.

Distler, U., Kuharev, J., Navarro, P. & Tenzer, S. Label-free quantification in ion mobility-enhanced data-independent acquisition proteomics. *Nat. Protoc.* **11**, 795–812 (2016).

[CAS](#) [Google Scholar](#)

44. 44.

Wildschütz, L. et al. Transcriptomic and proteomic analysis of iris tissue and aqueous humor in juvenile idiopathic arthritis-associated uveitis. *J. Autoimmun.* **100**, 75–83 (2019).

[Google Scholar](#)

45. 45.

Trüper, H. G. & Schlegel, H. G. Sulphur metabolism in Thiorhodaceae. I. Quantitative measurements on growing cells of *Chromatium okenii*. *Antonie van Leeuwenhoek* **30**, 225–238 (1964).

[Google Scholar](#)

46. 46.

Laemmli, U. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685 (1970).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

47. 47.

Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**, 1792–1797 (2004).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

48. 48.

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).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

49. 49.

Noor, E. et al. Pathway thermodynamics highlights kinetic obstacles in central metabolism. *PLOS Comput. Biol.* **10**, e1003483 (2014).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

50. 50.

Witt, A., Pozzi, R., Diesch, S., Hädicke, O. & Grammel, H. New light on ancient enzymes – in vitro CO₂ fixation by pyruvate synthase of *Desulfovibrio africanus* and *Sulfolobus acidocaldarius*. *FEBS J.* **286**, 4494–4508 (2019).

[CAS](#) [Google Scholar](#)

51. 51.

Furdui, C. & Ragsdale, S. W. The role of pyruvate ferredoxin oxidoreductase in pyruvate synthesis during autotrophic growth by the Wood–Ljungdahl pathway. *J. Biol. Chem.* **275**, 28494–28499 (2000).

[CAS](#) [Google Scholar](#)

52. 52.

Flamholz, A., Noor, E., Bar-Even, A. & Milo, R. eQuilibrator—the biochemical thermodynamics calculator. *Nucleic Acids Res.* **40**, D770–D775 (2012).

[CAS](#) [PubMed](#) [Google Scholar](#)

[Download references](#)

Acknowledgements

We thank G. Fuchs for discussions during this work and for critical reading of the manuscript, W. Schulz for his help with cloning of citrate synthase genes, A. M. Berg and D. Ackermann for technical assistance and A. Probst for the help with bioinformatics analysis. This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) Project-ID BE 4822/5-1 to I.A.B. and Project-ID 364653263–TRR 235 to W.E. as well as the Hans-Fischer Gesellschaft (Munich) to W.E.

Author information

Author notes

1. These authors contributed equally: Lydia Steffens, Eugenio Pettinato, Thomas M. Steiner

Affiliations

1. Institute for Molecular Microbiology and Biotechnology, University of Münster, Münster, Germany

Lydia Steffens, Eugenio Pettinato & Ivan A. Berg

2. Bavarian NMR Center–Structural Membrane Biochemistry, Department of Chemistry, Technische Universität München, Garching, Germany

Thomas M. Steiner & Wolfgang Eisenreich

3. K. G. Jebsen Centre for Deep Sea Research, University of Bergen, Bergen, Norway

Achim Mall

4. Department of Biological Sciences, University of Bergen, Bergen, Norway

Achim Mall

5. Core Unit Proteomics, Interdisciplinary Center for Clinical Research, Medical Faculty, University of Münster, Münster, Germany

Simone König

Authors

1. Lydia Steffens

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Eugenio Pettinato

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Thomas M. Steiner

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Achim Mall

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Simone König

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Wolfgang Eisenreich

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Ivan A. Berg

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Contributions

L.S. and E.P. performed growth experiments, cloning, purification and characterization of enzymes and enzyme assays. L.S performed codon usage analysis and sulfide determination. T.M.S. performed isotopologue profiling experiments and GC–MS analysis of *H. maritima* medium. S.K. conducted proteomics analyses. A.M. performed phylogenetic analysis and bioenergetics calculations. I.A.B. and W.E. wrote the manuscript with input from other authors. L.S., E.P., T.M.S., A.M., S.K. and W.E. prepared figures. The manuscript was reviewed and approved by all coauthors.

Corresponding authors

Correspondence to [Wolfgang Eisenreich](#) or [Ivan A. Berg](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

Peer review information *Nature* thanks William Martin 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 Phylogenetic tree of citrate-cleaving enzymes.](#)

Sequences were aligned using MUSCLE⁴⁷, the phylogeny was calculated with IQtree⁴⁸ using the maximum likelihood method (LG + R6 substitution model). The scale bar represents amino acid substitutions per site. Branch labels are SH-aLRT support/ultrafast bootstrap support (%) values from 1,000 replications. The sequences used to build the tree are available in Supplementary Table 11.

[Extended Data Fig. 2 SDS-PAGE \(12.5%\) of recombinant citrate synthases from *D. acetivorans* after purification with Ni-NTA column.](#)

M, molecular mass standard proteins; lane 1, Desace_08345 (2 µg); lane 2, Desace_06860 (2 µg); lane 3, Desace_09325 (2 µg). Their predicted

molecular masses are 49 kDa (Desace_08345) and 50 kDa (Desace_06860; Desace_09325). Proteins were stained with Coomassie blue. For gel source data, see Supplementary Fig. 1. The SDS–PAGE analysis was conducted three times and all results were similar to the ones shown here.

Extended Data Fig. 3 Protein quantification using MS.

a, CLUSTAL format alignment by MAFFT (v.7.452) of sequence parts of citrate synthase isoforms to illustrate the unique peptides chosen for quantification. Desace_06, Desace_06860; Desace_08, Desace_08345; Desace_09, Desace_09325. **b**, Calibration curves generated for each citrate synthase using tryptic digests of the recombinant proteins isolated by one-dimensional PAGE. Injection volume of standard solution (μl , *y* axis) versus MS peak area in MRM experiments (*x* axis) as calculated by Skyline.

Extended Data Fig. 4 Metabolites in *H. maritima* growth medium.

a, Sulfide production during growth under mixotrophic conditions (CO_2 , H_2 , S^0 and 0.2 g l^{-1} yeast extract). **b**, Concentrations of potential fermentation products and some amino acids in the medium after inoculation and after 1, 2 and 3 days of cultivation. Data are mean \pm s.e.m. of four biological replicates for sulfide determination and of three biological and three technical replicates for the potential fermentation products and amino acids in the medium. [Source data](#)

Extended Data Fig. 5 Growth of *H. maritima* with different substrates.

a, Growth of *H. maritima* in 80% H_2 , 20% CO_2 , 0.2 g l^{-1} yeast extract and 5 g l^{-1} acetate, in 80% N_2 , 20% CO_2 , 0.2 g l^{-1} yeast extract and 5 g l^{-1} acetate, and in 100% N_2 , 0.2 g l^{-1} yeast extract and 5 g l^{-1} acetate. **b**, Growth of *H. maritima* in 80% H_2 , 20% N_2 and 0.2 g l^{-1} yeast extract, in 80% N_2 , 20% CO_2 and 0.2 g l^{-1} yeast extract, in 100% N_2 and 0.2 g l^{-1}

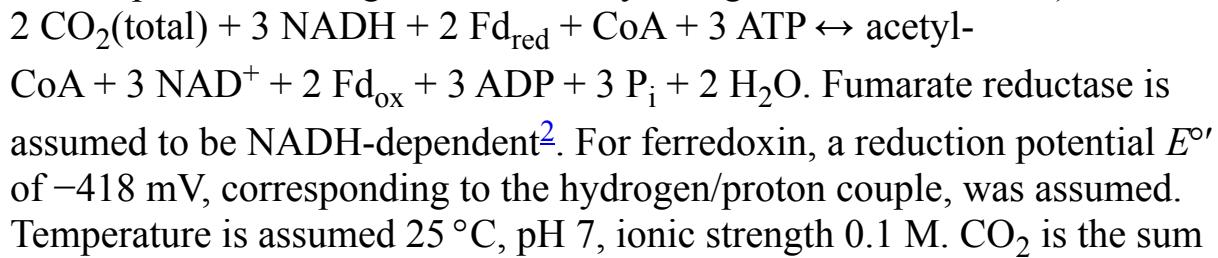
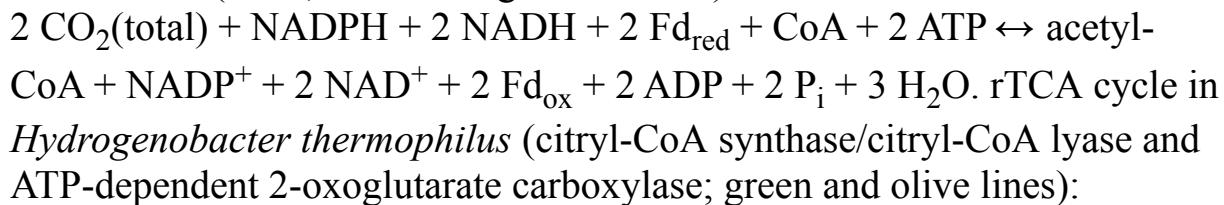
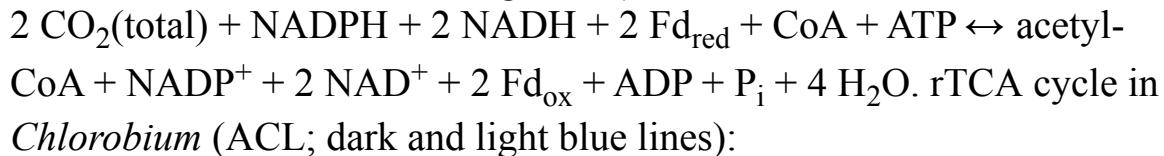
yeast extract, in 80% H₂ and 20% CO₂, and in 80% H₂, 20% CO₂ and 0.2 g l⁻¹ yeast extract. For all growth curves, data are presented as mean ± s.e.m. of three biological replicates. [Source data](#)

Extended Data Fig. 6 The labelling in C1 of Ala, Asp and Glu after growth in H₂, S⁰, 0.2 g l⁻¹ yeast extract and 4, 10 or 20% ¹³CO₂.

The labelling was calculated from the isotopologue composition of the corresponding amino acids and their fragments after the loss of the carboxylic atom (Supplementary Tables 4–6). Data are mean ± s.e.m. of six biological replicates for 4% ¹³CO₂ and of three biological replicates for 10% and 20% ¹³CO₂.

Extended Data Fig. 7 Dependency of energetic efficiencies of the rTCA cycle variants, pyruvate synthase and 2-oxoglutarate synthase reactions on CO₂ concentration.

Net equations are as follows². rTCA cycle (citrate synthase, as in Desulfurellaceae; red and orange lines):



of CO₂(aq) and its hydrated forms (H₂CO₃, HCO₃⁻ and CO₃²⁻). Under the *D. acetivorans* growth conditions (55 °C, pH 7, 2 bar pressure, 30.15 g l⁻¹ salinity), a CO₂ partial pressure of 40% corresponds to 147.6 mM CO₂, a partial pressure of 1% to 3.7 mM. **a**, Standard conditions (1 M concentrations for all reactants except CO₂). **b**, Assumed physiological metabolite concentrations (0.1 mM NADPH, 0.01 mM NADP⁺, 0.1 mM NADH, 1 mM NAD⁺, 5 mM ATP, 0.5 mM ADP, 10 mM phosphate⁴⁹, 1.07 mM CoA, 0.013 mM acetyl-CoA², 0.1 mM Fd_{red}, 0.001/0.0001 mM Fd_{ox} (refs. ^{3,50}), 0.23 mM succinyl-CoA, 0.44 mM 2-oxoglutarate¹³ and 0.18 mM pyruvate⁵¹). The calculations were done using eQuilibrator⁵².

Extended Data Fig. 8 Dependency of energetic efficiency of the reductive acetyl-CoA pathway in methanogens and acetogens on CO₂ concentration.

Net equation of the reductive acetyl-CoA pathway in methanogens and acetogens are as follows². Methanogens (*Methanothermobacter marburgiensis*; red and orange lines):

$$2 \text{ CO}_2(\text{total}) + 4 \text{ Fd}_{\text{red}} + \text{CoA} + 2 \text{ F}_{420}\text{H}_2 \leftrightarrow \text{acetyl-CoA} + 4 \text{ Fd}_{\text{ox}} + 2 \text{ F}_{420} + 5 \text{ H}_2\text{O}$$
. Acetogens (*Acetobacterium woodii*; dark and light blue lines):

$$2 \text{ CO}_2(\text{total}) + \text{NADPH} + \text{NADH} + 4 \text{ Fd}_{\text{red}} + \text{CoA} + \text{ATP} \leftrightarrow \text{acetyl-CoA} + \text{NADP}^+ + \text{NAD}^+ + 4 \text{ Fd}_{\text{ox}} + \text{ADP} + \text{P}_i + 4 \text{ H}_2\text{O}$$
. A free energy change ($\Delta G'$) of acetogenesis and methanogenesis is usually not sufficient to drive the synthesis of 1 mol ATP per mol of product under physiological conditions, and energy generation requires chemiosmotic coupling and entails flavin-based electron bifurcation^{30,31}. However, a high CO₂ partial pressure may enable the functioning of a hybrid reductive acetyl-CoA pathway that combines CO₂ reduction with ferredoxin, as in methanogens, and acetate formation from acetyl-CoA coupled with substrate phosphorylation, as in acetogens. Calculations were done as in Extended Data Fig. 7. **a**, Standard conditions (1 M concentrations for all reactants except CO₂). **b**, Assumed physiological metabolite concentrations (0.1 mM

NADPH, 0.01 mM NADP⁺, 0.1 mM NADH, 1 mM NAD⁺, 5 mM ATP, 0.5 mM ADP, 0.1 mM reduced coenzyme F₄₂₀, 0.1 mM oxidized coenzyme F₄₂₀, 10 mM phosphate⁴⁹, 0.28 mM CoA, 0.0104 mM acetyl-CoA⁵¹, 0.1 mM Fd_{red}, 0.001/0.0001 mM Fd_{ox} (refs. ^{3,50}).

Extended Data Table 1 Catalytic properties of *D. acetivorans* citrate synthase Desace_08345

[Full size table](#)

Extended Data Table 2 Enzymes of carbon metabolism in cell extracts of autotrophically grown *S. azorense*

[Full size table](#)

Supplementary information

[Supplementary Information](#)

This file contains Supplementary Tables 2, 3, 4, 5, 6, 7, 8, 10, a Supplementary Discussion and Supplementary References.

[Reporting Summary](#)

[Supplementary Figure 1](#)

Gel source data for Extended Data Fig. 2.

[Supplementary Table 1](#)

Proteome expression data for cell lysates of *D. acetivorans* grown under autotrophic (CO₂ + H₂ + S⁰) and heterotrophic (acetate + CO₂ + S⁰) conditions. Shown is the processing output as generated by Progenesis QIP containing One Way Anova *P* values, *Q* values and power as well as further parameters such as the number of unique peptides. Entries have been coloured for improved visualisation and according to criteria such as *p* = 0.05 and fold values < 2. A second Excel sheet was created by deleting all those entries not matching conditions *p* < 0.05, fold value > 2.

Supplementary Table 9

Proteome expression data for cell lysates of *H. maritima* grown with 4 and 20% CO₂ in the gas mixture under mixotrophic conditions (CO₂ + 80% H₂ + S⁰ + 0.2 g l⁻¹ yeast extract). Shown is the processing output as generated by Progenesis QIP containing One Way Anova *P* values, *Q* values and power as well as further parameters such as the number of unique peptides. Selection criteria were chosen as in Supplementary Table 1.

Supplementary Table 11

Alignment for Extended Data Fig. 1.

Peer Review File

Source data

Source Data Fig. 3

Source Data Extended Data Fig. 4

Source Data Extended Data Fig. 5

Rights and permissions

Reprints and Permissions

About this article



Check for
updates

Cite this article

Steffens, L., Pettinato, E., Steiner, T.M. *et al.* High CO₂ levels drive the TCA cycle backwards towards autotrophy. *Nature* **592**, 784–788 (2021). <https://doi.org/10.1038/s41586-021-03456-9>

[Download citation](#)

- Received: 12 August 2020
- Accepted: 15 March 2021
- Published: 21 April 2021
- Issue Date: 29 April 2021
- DOI: <https://doi.org/10.1038/s41586-021-03456-9>

Further reading

- [Life in a carbon dioxide world](#)
 - Martina Preiner
 - & William F. Martin

Nature (2021)

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

Associated Content

Life in a carbon dioxide world

- Martina Preiner
- William F. Martin

Nature News & Views 21 Apr 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03456-9>

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

- Article
- [Published: 14 April 2021](#)

CRL4^{AMBRA1} is a master regulator of D-type cyclins

- [Daniele Simoneschi](#) [ORCID: orcid.org/0000-0003-0448-4295](#)^{1,2},
- [Gergely Rona](#) [ORCID: orcid.org/0000-0003-3222-7261](#)^{1,2,3},
- [Nan Zhou](#) [ORCID: orcid.org/0000-0001-9524-4275](#)⁴,
- [Yeon-Tae Jeong](#)^{1,2},
- [Shaowen Jiang](#)^{1,2},
- [Giacomo Milletti](#) [ORCID: orcid.org/0000-0003-2391-342X](#)^{5,6},
- [Arnaldo A. Arbini](#)^{2,7},
- [Alfie O'Sullivan](#)^{1,2},
- [Andrew A. Wang](#) [ORCID: orcid.org/0000-0001-8018-3828](#)^{1,2},
- [Sorasicha Nithikasem](#)^{1,2},
- [Sarah Keegan](#)^{1,2,8},
- [Yik Siu](#)¹,
- [Valentina Cianfanelli](#) [ORCID: orcid.org/0000-0002-3857-9463](#)^{5,9},
- [Emiliano Maiani](#) [ORCID: orcid.org/0000-0003-1432-5394](#)^{9,10},
- [Francesca Nazio](#)⁵,
- [Francesco Cecconi](#) [ORCID: orcid.org/0000-0002-5614-4359](#)^{5,6,9},
- [Francesco Boccalatte](#) [ORCID: orcid.org/0000-0003-1732-2033](#)^{2,7},
- [David Fenyö](#)^{1,2,8},
- [Drew R. Jones](#)¹,
- [Luca Busino](#) [ORCID: orcid.org/0000-0001-6758-9276](#)⁴ &
- [Michele Pagano](#) [ORCID: orcid.org/0000-0003-3210-2442](#)^{1,2,3}

[Nature](#) volume **592**, pages 789–793(2021) [Cite this article](#)

- 8562 Accesses

- 2 Citations
- 196 Altmetric
- [Metrics details](#)

Subjects

- [Cell division](#)
- [Mechanisms of disease](#)
- [Ubiquitin ligases](#)

Abstract

D-type cyclins are central regulators of the cell division cycle and are among the most frequently deregulated therapeutic targets in human cancer¹, but the mechanisms that regulate their turnover are still being debated^{2,3}. Here, by combining biochemical and genetics studies in somatic cells, we identify CRL4^{AMBRA1} (also known as CRL4^{DCAF3}) as the ubiquitin ligase that targets all three D-type cyclins for degradation. During development, loss of *Ambra1* induces the accumulation of D-type cyclins and retinoblastoma (RB) hyperphosphorylation and hyperproliferation, and results in defects of the nervous system that are reduced by treating pregnant mice with the FDA-approved CDK4 and CDK6 (CDK4/6) inhibitor abemaciclib. Moreover, AMBRA1 acts as a tumour suppressor in mouse models and low *AMBRA1* mRNA levels are predictive of poor survival in cancer patients. Cancer hotspot mutations in D-type cyclins abrogate their binding to AMBRA1 and induce their stabilization. Finally, a whole-genome, CRISPR–Cas9 screen identified *AMBRA1* as a regulator of the response to CDK4/6 inhibition. Loss of *AMBRA1* reduces sensitivity to CDK4/6 inhibitors by promoting the formation of complexes of D-type cyclins with CDK2. Collectively, our results reveal the molecular mechanism that controls the stability of D-type cyclins during cell-cycle progression, in development and in human cancer, and implicate AMBRA1 as a critical regulator of the RB pathway.

[Access through your institution](#)

[Change institution](#)
[Buy or subscribe](#)

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: Three orthogonal screens identify CRL4^{AMBRA1} as a regulator of cyclin D1.

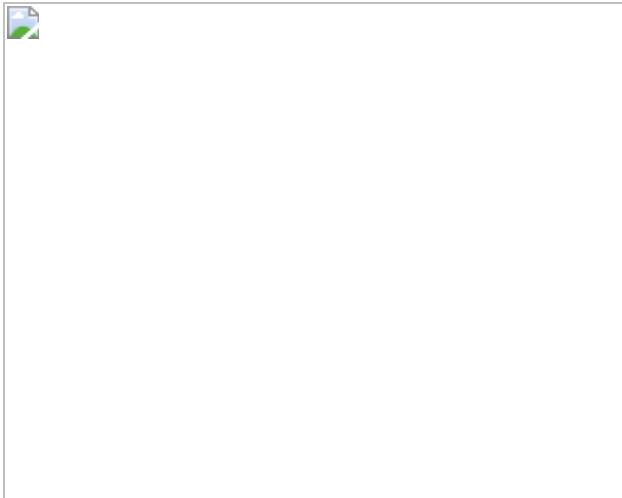


Fig. 2: AMBRA1 targets D-type cyclins for ubiquitin-mediated degradation, controlling cell cycle progression.

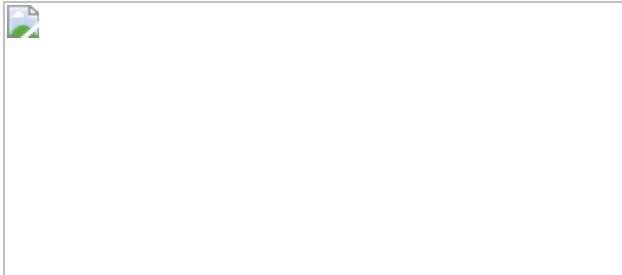


Fig. 3: D-type cyclins accumulate in *Ambral^{gt/gt}* mice and their hyperproliferative and neurological phenotypes are attenuated by CDK4/6 inhibition.

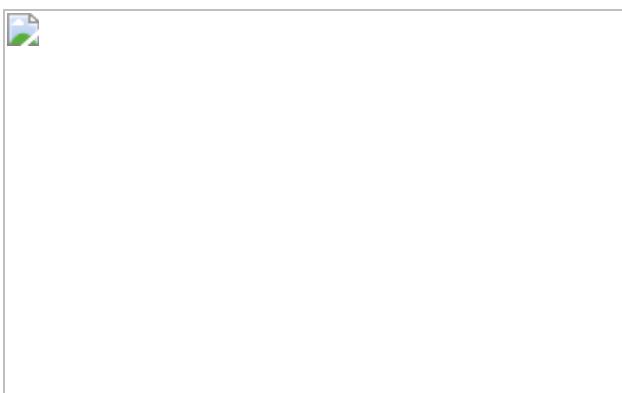


Fig. 4: AMBRA1 acts as a tumour suppressor and its absence reduces sensitivity to inhibitors of CDK4/6 by promoting the formation of D-type cyclins–CDK2 complexes.



Data availability

All accession codes, unique identifiers, and web links for publicly available datasets are available within the article and in the Reporting Summary. Those who carried out the original analysis and collection of the data generated by the DECIPHER community bear no responsibility for the further analysis or interpretation of the data. Mass spectrometry data for cyclin D1 immunopurifications are provided in Supplementary Table 1. The full scanned images obtained by electrophoretic separation are provided in Supplementary Fig. 1. All original data are available from the corresponding authors upon request. [Source data](#) are provided with this paper.

Code availability

Custom scripts for image analysis are provided at GitHub (https://github.com/FenyoLab/CRL4AMBRA1_img_analysis). Custom scripts for cancer-related analyses are provided at GitHub (https://github.com/Lesdormis/AMBRA1_analysis).

References

1. 1.

Malumbres, M. & Barbacid, M. Cell cycle, CDKs and cancer: a changing paradigm. *Nat. Rev. Cancer* **9**, 153–166 (2009).

[CAS](#) [Google Scholar](#)

2. 2.

Kanie, T. et al. Genetic reevaluation of the role of F-box proteins in cyclin D1 degradation. *Mol. Cell. Biol.* **32**, 590–605 (2012).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

3. 3.

Qie, S. & Diehl, J. A. Cyclin D degradation by E3 ligases in cancer progression and treatment. *Semin. Cancer Biol.* **67**, 159–170 (2020).

[CAS](#) [Google Scholar](#)

4. 4.

Reichermeier, K. M. et al. PIKES analysis reveals response to degraders and key regulatory mechanisms of the CRL4 network. *Mol. Cell* **77**, 1092–1106.e9 (2020).

[CAS](#) [Google Scholar](#)

5. 5.

Natsume, T., Kiyomitsu, T., Saga, Y. & Kanemaki, M. T. Rapid protein depletion in human cells by auxin-inducible degron tagging with short homology donors. *Cell Rep.* **15**, 210–218 (2016).

[CAS](#) [Google Scholar](#)

6. 6.

Barbash, O., Egan, E., Pontano, L. L., Kosak, J. & Diehl, J. A. Lysine 269 is essential for cyclin D1 ubiquitylation by the SCF^{Fbx4/αB-crystallin} ligase and subsequent proteasome-dependent degradation. *Oncogene* **28**, 4317–4325 (2009).

[CAS](#) [Google Scholar](#)

7. 7.

Fimia, G. M. et al. Ambra1 regulates autophagy and development of the nervous system. *Nature* **447**, 1121–1125 (2007).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

8. 8.

Davoli, T. et al. Cumulative haploinsufficiency and triplosensitivity drive aneuploidy patterns and shape the cancer genome. *Cell* **155**, 948–962 (2013).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

9. 9.

Behan, F. M. et al. Prioritization of cancer therapeutic targets using CRISPR–Cas9 screens. *Nature* **568**, 511–516 (2019).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

10. 10.

Rahrmann, E. P. et al. *Sleeping Beauty* screen identifies *RREB1* and other genetic drivers in human B-cell lymphoma. *Mol. Cancer Res.* **17**, 567–582 (2019).

[CAS](#) [Google Scholar](#)

11. 11.

Reddy, A. et al. Genetic and functional drivers of diffuse large B cell lymphoma. *Cell* **171**, 481–494 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

12. 12.

Chytil, A. et al. Construction of a cyclin D1–Cdk2 fusion protein to model the biological functions of cyclin D1–Cdk2 complexes. *J. Biol. Chem.* **279**, 47688–47698 (2004).

[CAS](#) [Google Scholar](#)

13. 13.

Asghar, U., Witkiewicz, A. K., Turner, N. C. & Knudsen, E. S. The history and future of targeting cyclin-dependent kinases in cancer therapy. *Nat. Rev. Drug Discov.* **14**, 130–146 (2015).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

14. 14.

Li, W. et al. MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol.* **15**, 554 (2014).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

15. 15.

Sanjana, N. E., Shalem, O. & Zhang, F. Improved vectors and genome-wide libraries for CRISPR screening. *Nat. Methods* **11**, 783–784 (2014).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

16. 16.

Reynders, M. et al. PHOTACs enable optical control of protein degradation. *Sci. Adv.* **6**, eaay5064 (2020).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

17. 17.

Lee, J. & Zhou, P. DCAF₅s, the missing link of the CUL4–DDB1 ubiquitin ligase. *Mol. Cell* **26**, 775–780 (2007).

[CAS](#) [Google Scholar](#)

18. 18.

Bern, M., Kil, Y. J. & Becker, C. Byonic: advanced peptide and protein identification software. *Curr. Protoc. Bioinformatics Chapter 13*, Unit13.20 (2012).

[Google Scholar](#)

19. 19.

Donato, V. et al. The TDH–GCN5L1–Fbxo15–KBP axis limits mitochondrial biogenesis in mouse embryonic stem cells. *Nat. Cell Biol.* **19**, 341–351 (2017).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

20. 20.

Jeong, Y. T. et al. The ULK1–FBXW5–SEC23B nexus controls autophagy. *eLife* **7**, e42253 (2018).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

21. 21.

Kuchay, S. et al. PTEN counteracts FBXL2 to promote IP3R3- and Ca²⁺-mediated apoptosis limiting tumour growth. *Nature* **546**, 554–

558 (2017).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

22. 22.

Ran, F. A. et al. Genome engineering using the CRISPR–Cas9 system. *Nat. Protoc.* **8**, 2281–2308 (2013).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

23. 23.

Dragicevic, P. in *Modern Statistical Methods for HCI* (eds Robertson, J. & Kaptein, M.) 291–330 (Springer, 2016).

24. 24.

Róna, G. et al. Dynamics of re-constitution of the human nuclear proteome after cell division is regulated by NLS-adjacent phosphorylation. *Cell Cycle* **13**, 3551–3564 (2014).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

25. 25.

Cecchetelli, A., Kwan, R., Krauter, S. & Chiriboga, L. Minimum concentration of cells required for the preparation of compact plasma-thrombin cell blocks. *J. Histotechnol.* **35**, 68–70 (2012).

[CAS](#) [Google Scholar](#)

26. 26.

Glinsmann-Gibson, B. et al. Recommendations for tissue microarray construction and quality assurance. *Appl. Immunohistochem. Mol. Morphol.* (2020).

27. 27.

Miki, Y. et al. AMBRA1, a novel α -synuclein-binding protein, is implicated in the pathogenesis of multiple system atrophy. *Brain Pathol.* **28**, 28–42 (2018).

[CAS](#) [Google Scholar](#)

28. 28.

Lewitowicz, P. et al. Tumor digital masking allows precise patient triaging: a study based on Ki-67 scoring in gastrointestinal stromal tumors. *Scanning* **2018**, 7807416 (2018).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

29. 29.

Lorenzo-Martín, L. F. et al. VAV2 signaling promotes regenerative proliferation in both cutaneous and head and neck squamous cell carcinoma. *Nat. Commun.* **11**, 4788 (2020).

[ADS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

30. 30.

Tuominen, V. J., Ruotoistenmäki, S., Viitanen, A., Jumppanen, M. & Isola, J. ImmunoRatio: a publicly available web application for quantitative image analysis of estrogen receptor (ER), progesterone receptor (PR), and Ki-67. *Breast Cancer Res.* **12**, R56 (2010).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

31. 31.

Shaknovich, R. et al. DNA methylation signatures define molecular subtypes of diffuse large B-cell lymphoma. *Blood* **116**, e81–e89 (2010).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

32. 32.

Lenz, G. et al. Molecular subtypes of diffuse large B-cell lymphoma arise by distinct genetic pathways. *Proc. Natl Acad. Sci. USA* **105**, 13520–13525 (2008).

[ADS](#) [CAS](#) [Google Scholar](#)

33. 33.

D'Angelo, F. et al. The molecular landscape of glioma in patients with Neurofibromatosis 1. *Nat. Med.* **25**, 176–187 (2019).

[Google Scholar](#)

34. 34.

Compagno, M. et al. Mutations of multiple genes cause deregulation of NF-κB in diffuse large B-cell lymphoma. *Nature* **459**, 717–721 (2009).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

35. 35.

Brune, V. et al. Origin and pathogenesis of nodular lymphocyte-predominant Hodgkin lymphoma as revealed by global gene expression analysis. *J. Exp. Med.* **205**, 2251–2268 (2008).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

36. 36.

Cerami, E. et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* **2**, 401–404 (2012).

[Google Scholar](#)

37. 37.

Gao, J. et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* **6**, pl1 (2013).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

38. 38.

Tate, J. G. et al. COSMIC: the Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res.* **47** (D1), D941–D947 (2019).

[CAS](#) [Google Scholar](#)

39. 39.

Schmitz, R. et al. Burkitt lymphoma pathogenesis and therapeutic targets from structural and functional genomics. *Nature* **490**, 116–120 (2012).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

40. 40.

Clijsters, L. et al. Cyclin F controls cell-cycle transcriptional outputs by directing the degradation of the three activator E2Fs *Mol. Cell* **74**, 1264–1277 (2019).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

41. 41.

Yada, M. et al. Phosphorylation-dependent degradation of c-Myc is mediated by the F-box protein Fbw7. *EMBO J.* **23**, 2116–2125 (2004).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

42. 42.

Maxwell, P. H. et al. The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. *Nature*

399, 271–275 (1999).

[ADS](#) [CAS](#) [Google Scholar](#)

43. 43.

Kubbutat, M. H., Jones, S. N. & Vousden, K. H. Regulation of p53 stability by Mdm2. *Nature* **387**, 299–303 (1997).

[ADS](#) [CAS](#) [Google Scholar](#)

44. 44.

Li, Y., Ge, D. & Lu, C. The SMART app: an interactive web application for comprehensive DNA methylation analysis and visualization. *Epigenetics Chromatin* **12**, 71 (2019).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

45. 45.

Lin, D. I. et al. Phosphorylation-dependent ubiquitination of cyclin D1 by the SCF^{FBX4-αB} crystallin complex. *Mol. Cell* **24**, 355–366 (2006).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

46. 46.

Okabe, H. et al. A critical role for FBXW8 and MAPK in cyclin D1 degradation and cancer cell proliferation. *PLoS ONE* **1**, e128 (2006).

[ADS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

47. 47.

Santra, M. K., Wajapeyee, N. & Green, M. R. F-box protein FBXO31 mediates cyclin D1 degradation to induce G1 arrest after DNA damage. *Nature* **459**, 722–725 (2009).

[ADS](#) [CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

48. 48.

Wei, S. et al. A novel mechanism by which thiazolidinediones facilitate the proteasomal degradation of cyclin D1 in cancer cells. *J. Biol. Chem.* **283**, 26759–26770 (2008).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

49. 49.

Yu, Z. K., Gervais, J. L. & Zhang, H. Human CUL-1 associates with the SKP1/SKP2 complex and regulates p21(CIP1/WAF1) and cyclin D proteins. *Proc. Natl Acad. Sci. USA* **95**, 11324–11329 (1998).

[ADS](#) [CAS](#) [Google Scholar](#)

50. 50.

Chen, B. B., Glasser, J. R., Coon, T. A. & Mallampalli, R. K. FBXL2 is a ubiquitin E3 ligase subunit that triggers mitotic arrest. *Cell Cycle* **10**, 3487–3494 (2011).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

51. 51.

Chen, B. B. et al. F-box protein FBXL2 targets cyclin D2 for ubiquitination and degradation to inhibit leukemic cell proliferation. *Blood* **119**, 3132–3141 (2012).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

52. 52.

Pawar, S. A. et al. C/EBPdelta targets cyclin D1 for proteasome-mediated degradation via induction of CDC27/APC3 expression. *Proc. Natl Acad. Sci. USA* **107**, 9210–9215 (2010).

[ADS](#) [CAS](#) [Google Scholar](#)

53. 53.

Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

54. 54.

Yang, H. W. et al. Stress-mediated exit to quiescence restricted by increasing persistence in CDK4/6 activation. *eLife* **9**, e44571 (2020).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

55. 55.

Cappuccio, G., Ugga, L., Parrini, E., D’Amico, A. & Brunetti-Pierri, N. Severe presentation and complex brain malformations in an individual carrying a CCND2 variant. *Mol. Genet. Genomic Med.* **7**, e708 (2019).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

56. 56.

Maini, I. et al. A novel CCND2 mutation in a previously reported case of megalencephaly and perisylvian polymicrogyria with postaxial polydactyly and hydrocephalus. *Neuropediatrics* **49**, 222–224 (2018).

[CAS](#) [Google Scholar](#)

57. 57.

McDermott, J. H. et al. Hypoglycaemia represents a clinically significant manifestation of PIK3CA- and CCND2-associated segmental overgrowth. *Clin. Genet.* **93**, 687–692 (2018).

[CAS](#) [Google Scholar](#)

58. 58.

Mirzaa, G. et al. De novo CCND2 mutations leading to stabilization of cyclin D2 cause megalencephaly-polymicrogyria-polydactyly-hydrocephalus syndrome. *Nat. Genet.* **46**, 510–515 (2014).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

59. 59.

Firth, H. V. et al. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *Am. J. Hum. Genet.* **84**, 524–533 (2009).

[CAS](#) [PubMed](#) [PubMed Central](#) [Google Scholar](#)

60. 60.

Sameshima, T., Morisada, N., Egawa, T., Kugo, M. & Iijima, K. MPPH syndrome with aortic coarctation and macrosomia due to CCND2 mutations. *Pediatr. Int.* **62**, 115–117 (2020).

[Google Scholar](#)

[Download references](#)

Acknowledgements

We thank G. V. Pires, L. Moro, V. Donato, A. Zeke, A. Peschiaroli, M. Swartling, K. Avrampou, J. Wang, E. Hernandez and A. Tsirigos for their contribution to this work, and R. Abraham, S. Dann, K. Nakayama and W. Wei for critically reading the paper. This study makes use of data generated by the DECIPHER community. A full list of centres that contributed to the generation of the data is available from <https://decipher.sanger.ac.uk/about/stats> and via email from decipher@sanger.ac.uk. Funding for the DECIPHER project was provided

by Wellcome. We thank E. Rosser and M. Lees from the Great Ormond Street Hospital for Children NHS Foundation Trust in London, UK, for providing the information on two patients with mutations in *CCND2*, deposited at DECIPHER (<https://decipher.sanger.ac.uk>). The Proteomics Laboratory, the Genome Technology Center (GTC), the Experimental Pathology Research Laboratory, the Cytometry and Cell Sorting Laboratory, and the Metabolomics Laboratory at NYU Langone Health are partially supported by the Cancer Center Support Grant P30CA016087 at the Laura and Isaac Perlmutter Cancer Center. The NYULH Center for Biospecimen Research and Development, Histology and Immunohistochemistry Laboratory (RRID:SCR_018304) is supported in part by the Laura and Isaac Perlmutter Cancer Center Support Grant, NIH/NCI P30CA016087, and the National Institutes of Health S10 Grants, NIH/ORIP S10OD01058 and S10OD018338. The authors thank C. I. Silvescu for data acquisition and processing of the mass spectrometry images. D.S. thanks F. Simoneschi, P. Simoneschi, D. Guidi and E. Guidi for continuous support. M.P. thanks T. M. Thor and T. B. Balduur for continuous support. This work was funded by grants from the National Institutes of Health (R01-CA76584 and R35-GM136250) to M.P. and (R01-CA207513-01) to L.B. The F.C. laboratory is supported by grants from the Danish Cancer Society (KBVU R72-A4408, R146-A9364, R231-A14034 to F.C. and R146-A9471 to V.C.), the Novo Nordisk Foundation (NNF13OC0007559, NNF16OC0022544), the Lundbeck Foundation (R233-2016-3360 to F.C. and R209-2015-3505 to V.C.), the LEO Foundation (LF17024 to F.C.), Associazione Italiana per la Ricerca sul Cancro (AIRC project IG 2019 no. 23543 to F.C.), the Italian Ministry of Research (MIUR, project PRIN 2017 Radius), and the Italian Ministry of Health (Ricerca Corrente to F.N.). V.C. is supported by Fondazione Umberto Veronesi. E.M. is an adjunct professor at UniCamillus—Saint Camillus International University of Health Sciences in Rome, Italy. F.B. is supported by a Young Investigator Grant from the Alex's Lemonade Stand Foundation. M.P. is an investigator with the Howard Hughes Medical Institute.

Author information

Affiliations

1. Department of Biochemistry and Molecular Pharmacology, NYU Grossman School of Medicine, New York, NY, USA

Daniele Simoneschi, Gergely Rona, Yeon-Tae Jeong, Shaowen Jiang, Alfie O'Sullivan, Andrew A. Wang, Sorasicha Nithikasem, Sarah Keegan, Yik Siu, David Fenyö, Drew R. Jones & Michele Pagano

2. Laura and Isaac Perlmutter Cancer Center, NYU Grossman School of Medicine, New York, NY, USA

Daniele Simoneschi, Gergely Rona, Yeon-Tae Jeong, Shaowen Jiang, Arnaldo A. Arbini, Alfie O'Sullivan, Andrew A. Wang, Sorasicha Nithikasem, Sarah Keegan, Francesco Boccalatte, David Fenyö & Michele Pagano

3. Howard Hughes Medical Institute, NYU Grossman School of Medicine, New York, NY, USA

Gergely Rona & Michele Pagano

4. Department of Cancer Biology, University of Pennsylvania Perelman School of Medicine, Philadelphia, PA, USA

Nan Zhou & Luca Busino

5. Department of Pediatric Hemato-Oncology and Cell and Gene Therapy, IRCCS Bambino Gesù Children's Hospital, Rome, Italy

Giacomo Milletti, Valentina Cianfanelli, Francesca Nazio & Francesco Cecconi

6. Department of Biology, University of Rome Tor Vergata, Rome, Italy

Giacomo Milletti & Francesco Cecconi

7. Department of Pathology, NYU Grossman School of Medicine, New York, NY, USA

Arnaldo A. Arbini & Francesco Boccalatte

8. Institute for Systems Genetics, NYU Grossman School of Medicine,
New York, NY, USA

Sarah Keegan & David Fenyö

9. Cell Stress and Survival Unit, Danish Cancer Society Research Center,
Copenhagen, Denmark

Valentina Cianfanelli, Emiliano Maiani & Francesco Cecconi

10. Computational Biology Laboratory, Danish Cancer Society Research
Center, Copenhagen, Denmark

Emiliano Maiani

Authors

1. Daniele Simoneschi

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Gergely Rona

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Nan Zhou

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Yeon-Tae Jeong

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Shaowen Jiang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Giacomo Milletti

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Arnaldo A. Arbini

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Alfie O'Sullivan

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. Andrew A. Wang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. Sorasicha Nithikasem

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

11. Sarah Keegan

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

12. Yik Siu

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

13. Valentina Cianfanelli

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

14. Emiliano Maiani

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

15. Francesca Nazio

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

16. Francesco Cecconi

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

17. Francesco Boccalatte

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

18. David Fenyö

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

19. Drew R. Jones

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

20. Luca Busino

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

21. Michele Pagano

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Contributions

D.S. conceived, planned and performed most experiments, and co-wrote the manuscript. G.R., Y.-T.J., A.O., A.A.W. and S.N. helped with biochemical experiments. N.Z. and L.B. performed the DLBCL xenografts. S.J. helped with bioinformatics analyses. F.B. helped with the collection of mouse plasma, embryos and amniotic fluid. S.K. and D.F. performed the automated analysis of live-cell images. Y.S. and D.R.J. performed the LC-MS/MS detection of abemaciclib. G.M., V.C., E.M., F.N. and F.C. performed the experiments with *Ambra1^{+/+}* and *Ambra1^{gt/gt}* mice related to neural development. A.A.A. analysed and scored immunohistochemistry stainings. M.P. conceived, directed and coordinated the study, oversaw the results and co-wrote the manuscript. All authors discussed the results and commented on the manuscript.

Corresponding authors

Correspondence to [Luca Busino](#) or [Michele Pagano](#).

Ethics declarations

Competing interests

M.P. is a consultant for and has financial interests in Coho Therapeutics, CullGen, Kymera Therapeutics and SEED Therapeutics. M.P. is a cofounder of Coho Therapeutics, and is a member of the scientific advisory board of CullGen, Kymera Therapeutics and SEED Therapeutics, and is a consultant for Santi Therapeutics. The other authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Piotr Sicinski 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 CRL1 does not contribute to the degradation of D-type cyclins.](#)

a, hTERT RPE-1, U-2 OS and HCT-116 cells were treated with MG132 or MLN4924 for 4 h before collection. Cell extracts were immunoblotted for the indicated proteins. **b**, HCT-116 cells were treated for 3 h with vehicle (DMSO) or MLN4924 followed by cycloheximide for the indicated times before collection. Cell extracts were immunoblotted for the indicated proteins. **c**, HCT-116 cells were transfected with a non-targeting (N/T) siRNA or an siRNA against *DDB1* for three rounds before lysis and immunoblotting for the indicated proteins. **d**, HCT-116 cells were transfected with a non-targeting siRNA or an siRNA against *CUL1* for three rounds followed by treatment with cycloheximide for the indicated times before collection. Cell extracts were immunoblotted as indicated. **e**, U-2 OS cells were transfected with a non-targeting siRNA or an siRNA against *CUL1* for three rounds followed by treatment with cycloheximide for the indicated times before collection. Cell extracts were immunoblotted as indicated. **f**, PC-3 cells were transfected with a non-targeting (N/T) siRNA or an siRNA against *CUL1* for three rounds followed by treatment with cycloheximide for the indicated times before collection. Cell extracts were immunoblotted as indicated. **g**, hTERT RPE-1 cells were transfected with a non-targeting siRNA or an siRNA against *CUL1* for three rounds followed by treatment with cycloheximide for the indicated times before collection. Cell extracts were immunoblotted as indicated. **h**, U-2 OS cells were infected with a doxycycline-inducible lentivirus expressing a dominant-negative mutant of CUL1 (DN-CUL1). Cells were treated with doxycycline for the indicated times. Where indicated, cells were exposed to UV

(5 mJ cm⁻²) irradiation for 2 h before collection. Cell extracts were immunoblotted as indicated. **i**, Representative images of U-2 OS cells infected with a retrovirus expressing mAzG-CCND1 and treated with MLN4924 for the indicated times. Cells were monitored via live-cell imaging. Scale bar, 20 µm. **j**, Quantification of relative mAzG-CCND1 nuclear intensity fold change in arbitrary units (A.U.) from single cells tracked over time as in **i**. Data are mean ± s.d., $n = 34$ cells. **k**, Schematic representation of approach 1, aimed at measuring the accumulation of cyclin D1 upon silencing of all human DCAFs and the previously-reported regulators of D-type cyclins^{45,46,47,48,49,50,51,52} in HCT-116 and U-2 OS cells. **l**, Schematic representation of approach 2, aimed at identifying the binding partners of FFSS-tagged wild-type cyclin D1 and FFSS-tagged cyclin D1(T286A) compared with empty vector. HCT-116 cells were treated with DMSO, MG132 or MLN4924 for 4 h before lysis and immunoprecipitation with an anti-Flag resin, followed by elution with a 3×Flag peptide. Eluted samples were then analysed by LC-MS/MS for the identification of binding partners. **m**, Schematic representation of approach 3, aimed at identifying global regulators of cyclin D1 in HCT-116 cells using the whole-genome, CRISPR–Cas9 human GECKOv2 pooled library. Unless otherwise noted, experiments were performed at least three independent times. [Source data](#)

Extended Data Fig. 2 Validation of candidate regulators of D-type cyclins.

a, HCT-116 cells were transfected with a non-targeting siRNA or siRNAs against the indicated targets for two rounds before collection. Cell extracts were immunoblotted as indicated (upper panel). RNA extracts were also collected to measure the relative mRNA levels for each indicated target by quantitative PCR with reverse transcription (RT–qPCR). Each blue column represents the mRNA levels of the target indicated by the respective siRNA and normalized to its corresponding mRNA levels in cells transfected with a non-targeting siRNA (bottom). Data are mean mRNA levels ± s.d., $n = 3$ independent experiments. **b**, HCT-116 cells were transfected with a non-targeting siRNA or siRNAs against the indicated targets for two rounds followed by treatment with cycloheximide for the indicated times before collection. Cell extracts were immunoblotted as indicated. **c**, HCT-116 cells

were transfected with a non-targeting siRNA or an siRNA against *DCAF7* followed by treatment with cycloheximide for the indicated times before collection. Cell extracts were immunoblotted as indicated. **d**, U-2 OS cells were transfected with a non-targeting siRNA or an siRNA against *DCAF8* followed by treatment with cycloheximide for the indicated times before collection. Cell extracts were immunoblotted as indicated. **e**, U-2 OS cells were transfected with a non-targeting siRNA or an siRNA against *WDR61* followed by treatment with cycloheximide for the indicated times before collection. Cell extracts were immunoblotted as indicated (left). RNA extracts were also collected to measure the relative mRNA levels of *WDR61* by RT-qPCR (right). Data are mean mRNA levels \pm s.d., $n = 3$ independent experiments. **f**, hTERT RPE-1 cells were transfected with a non-targeting siRNA or an siRNA against *AMBRA1* followed by treatment with cycloheximide for the indicated times before collection. Cell extracts were immunoblotted as indicated. **g**, HeLa, HEK 293T, HCT-116, hTERT RPE-1 and U-2 OS cells were transfected with a non-targeting siRNA or an siRNA against *AMBRA1* before collection. Cell extracts were immunoblotted for the indicated proteins (l.e., long exposure; m.e., medium exposure; s.e., short exposure). **h**, Schematic representation of the *AMBRA1* genomic locus and gRNA target location. Exon 2 refers to the human *AMBRA1* gene in XM_005253009 (GRCh38.p13, NCBI gene: 55626). Representative *AMBRA1* wild-type genomic DNA template and knockout mutant sequences identified by TOPO-TA cloning of PCR products from HCT-116, U-2 OS and hTERT-RPE1 cells are depicted. **i**, Three *AMBRA1*^{-/-} and one *AMBRA1*^{+/+} hTERT RPE-1 clones were treated with cycloheximide for the indicated times, after which cell extracts were immunoblotted for the indicated proteins. **j**, Two *AMBRA1*^{-/-} and one *AMBRA1*^{+/+} T98G clones were treated with either cycloheximide for the indicated times, MG132 or MLN4924 (both for 4 h) before collection. Cell extracts were then immunoblotted for the indicated proteins. **k**, HCT-116 cells were infected with a retrovirus expressing mAzG-CCND1 and transfected with a non-targeting siRNA or an siRNA against *AMBRA1* for two rounds. Cells were then treated with cycloheximide for the indicated times, after which cell extracts were immunoblotted for the indicated proteins (endo, endogenous; exo, exogenous). **l**, *AMBRA1*^{-/-} U-2 OS cells were transiently transfected with a vector expressing a C-terminal 3 \times Flag-tagged *AMBRA1* or an empty vector for 96 h. Where indicated, cells were

treated with MLN4924 for 4 h before lysis. Cell extracts were immunoblotted as indicated. Unless otherwise noted, experiments were performed at least three independent times. [Source data](#)

Extended Data Fig. 3 Depletion of AMBRA1 using the auxin-inducible degron system.

a, Schematic representation of the CRISPR–Cas9-based strategy to insert an N-terminal, 2×Flag-tagged minimal AID (mAID) in the endogenous locus of AMBRA1 (5'-end of exon 2) in HCT-116 cells. Cells were infected with a doxycycline-inducible lentivirus expressing *Oryza sativa* TIR1 (osTIR1). **b**, Three knock-in clones were obtained as in **a** and compared to wild-type clones for proper endogenous tagging of *AMBRA1* by immunoblotting. Note the higher molecular weight of 2×Flag-mAID-tagged AMBRA1. **c**, Three knock-in clones were treated with doxycycline with or without increasing doses of auxin for 12 h. Cell extracts were immunoblotted as indicated. The asterisk indicates a non-specific band. **d**, One clone of 2×Flag-mAID-AMBRA1 HCT-116 cells (clone #F1) was treated with either DMSO or doxycycline for 12 h, and/or auxin, MG132 and MLN4924 (alone or in combination) for 4 h as indicated. Cell extracts were immunoblotted as indicated. **e**, Longer exposure from the cyclin D1 and phospho-cyclin D1 (T286) immunoblots described in **d**. **f**, Whole-cell extracts from 2×Flag-mAID-AMBRA1 HCT-116 cells pre-treated with either DMSO or doxycycline for 12 h, and exposed to combinations of auxin, MG132 and/or MLN4924 as indicated for 4 h, related to Fig. [2d](#). Cell extracts were immunoblotted as indicated. Unless otherwise noted, experiments were performed at least three independent times.

Extended Data Fig. 4 Mutations in the TP motif of D-type cyclins protect them from AMBRA1-mediated degradation.

a, Protein sequence alignment of the extreme C termini of D-type cyclins from different source organisms. Amino acids highlighted in red correspond to the highly conserved TP motif. Asterisks indicate positions that are fully conserved. Colons indicate conservation between groups of strongly similar properties. The dot indicates conservation between groups of weakly

similar properties. The alignment was performed in UniProt (<https://www.uniprot.org/>) using the Clustal Omega algorithm. UniProt identifiers are shown on the left. Source organisms: HUMAN, *H. sapiens*; BOVIN, *Bos taurus*; CANLF, *Canis familiaris*; MOUSE, *Mus musculus*; RAT, *Rattus norvegicus*; DANRE, *Danio rerio*; BRAFL, *Branchiostoma floridae*; STRPU, *Strongylocentrotus purpuratus*; CRAGI, *Crassostrea gigas*; LOTGI, *Lottia gigantean*; OCTBM, *Octopus bimaculoides*; STRMM, *Strigamia maritima*; BOMMO, *Bombyx mori*; ATTCE, *Atta cephalotes*; APIME, *Apis mellifera*; CULQU, *Culex quinquefasciatus*; HYDVU, *Hydra vulgaris*; NEMVE, *Nematostella vectensis*; TRIAD *Trichoplax adhaerens*. **b**, Multiple sequence logo of the alignment performed in **a** generated using WebLogo 3 (<http://weblogo.threepplusone.com>)⁵³. **c**, HCT-116 cells were infected with lentiviruses expressing either FFSS-tagged empty vector (EV) or FFSS-tagged wild-type or mutant cyclin D1. Cells were treated with MLN4924 before lysis and affinity purification (AP) with an anti-STREP resin. Proteins were immunoblotted as indicated. **d**, HEK 293T cells were transiently transfected with plasmids expressing either SF-tagged empty vector or SF-tagged D-type cyclins wild-type or mutants. Cells were treated with MLN4924 before lysis and immunoprecipitation with an anti-Flag resin. Proteins were immunoblotted as indicated. **e**, HCT-116 cells were infected with retroviruses expressing FFSS-tagged wild-type or mutant cyclin D1. Cells were treated with cycloheximide for the indicated times before collection. Cell extracts were immunoblotted as indicated. **f**, HCT-116 cells were infected with retroviruses expressing FFSS-tagged wild-type or mutant cyclin D2. Cells were treated with cycloheximide for the indicated times before collection. Cell extracts were immunoblotted as indicated. **g**, HCT-116 cells were infected with retroviruses expressing FFSS-tagged wild-type or mutant cyclin D3. Cells were treated with cycloheximide for the indicated times before collection. Cell extracts were immunoblotted as indicated. **h**, HCT-116 cells were transfected with vectors expressing SF-tagged CDC25A, SF-tagged p27, SF-tagged cyclin D1, SF-tagged cyclin D1(T286A) or empty vector. Cells were treated with MLN4924 for 4 h before collection. Lysates were used for an immunoprecipitation with an anti-Flag resin. Proteins were immunoblotted as indicated. **i**, HCT-116 cells were transfected with vectors expressing HA-tagged FBXO4, SF-tagged TRF1, SF-tagged cyclin D1, SF-tagged cyclin

D1(T286A) or empty vector. Cells were treated with MLN4924 for 4 h before collection. Lysates were used for an affinity purification with an anti-Strep resin. Proteins were immunoblotted as indicated. The asterisk indicates a non-specific band. **j**, HCT-116 cells were transfected with vectors expressing HA-tagged FBXW8, SF-tagged IRS1, SF-tagged cyclin D1, SF-tagged cyclin D1(T286A) or empty vector. Cells were treated with MLN4924 for 4 h before collection. Lysates were used for an immunoprecipitation with an anti-Flag resin. Proteins were immunoblotted as indicated. **k**, Parental and 2×Flag-mAID-AMBRA1 HCT-116 cells were lysed and their extracts were used for an immunoprecipitation with an anti-Flag resin. Proteins were immunoblotted as indicated. **l**, One clone of 2×Flag-mAID-AMBRA1 HCT-116 cells (clone F1) was infected with retroviruses expressing GFP-CCND1 or GFP-CCND1(T286A). Cells were treated with doxycycline for 12 h, before exposure to auxin for 4 h before collection as indicated. Control cells were treated with DMSO. Cell extracts were immunoblotted as indicated. **m**, Parental and *AMBRA1*^{-/-} HCT-116 and U-2 OS cells were lysed and their extracts were blotted as indicated. The accumulation of cyclin D1 and phosphorylated cyclin D1 (T286) is reported as their corresponding intensity ratios between *AMBRA1*^{-/-} and parental cells. Data are mean ± s.e.m. ($n = 3$ biological replicates per group). **n**, Three different *AMBRA1*^{-/-} HCT-116 clones (clones D2, F11 and G7), 2×Flag-mAID-AMBRA1 HCT-116 cells (clone F1) and HCT-116 cells transfected with a non-targeting siRNA or an siRNA against *AMBRA1* were lysed and their extracts were immunoblotted as indicated (either by employing fluorescently labelled secondary antibodies or by enhanced chemiluminescence). Untransfected *CCND1*^{-/-} HCT-116 cells or *CCND1*^{-/-} HCT-116 cells transfected with vectors expressing untagged cyclin D1(WT) or untagged cyclin D1(T286A) were used as control. Where indicated (black triangles), samples were serially diluted before electrophoresis and immunoblotting for easier visual comparison of cyclin D1 and phosphorylated cyclin D1 (T286) levels between groups. The analysis of normalized phosphorylated cyclin D1 (T286) over total cyclin D1 fluorescent intensity levels are provided as bar graphs. The ratios were calculated from lane 2 and lane 4 from each biological group in the fluorescent immunoblot. **o**, HCT-116 cells were infected with lentiviruses expressing FFSS-tagged cyclin D1 or FFSS-tagged cyclin D1(T286A). Cells were treated with DMSO (N/T) or MLN4924 for 4 h before lysis and

immunoprecipitation with an anti-Flag resin, followed by elution with a 3×Flag peptide. Eluted samples were then analysed by LC-MS/MS for the identification of post-translational modifications. The table reports all identified sites in wild-type cyclin D1 and cyclin D1(T286A) that were modified with a di-Gly/114.0429 signature. **p**, Synthetic, unphosphorylated (amino acids PKAAEEEEEEEVLDLACTPTDVRDVDI) and phosphorylated (T286) (amino acids PKAAEEEEEEEVLDLAC-pT-PTDVRDVDI) peptides (2.5 µg each) corresponding to the extreme C terminus of cyclin D1 were spotted on a nitrocellulose membrane before immunoblotting with an anti-phosphorylated cyclin D1 (T286) antibody. **q**, Synthetic, phosphorylated (T286) peptide corresponding to the sequence of cyclin D1 indicated in **p** was incubated with in vitro transcribed and translated AMBRA1 from rabbit reticulocytes extracts for the indicated times in the presence of wild-type ubiquitin (Ub) or a mutant of ubiquitin [Ub(K0)] in which all lysine residues have been mutated to arginine. Reactions were stopped with Laemmli buffer. Proteins were immunoblotted as indicated. Unless otherwise noted, experiments were performed at least three independent times. [Source data](#)

Extended Data Fig. 5 Loss of AMBRA1 regulates D-type cyclins during normal cell cycle, upon nutrient deprivation and after genotoxic stress.

a, Three *AMBRA1*^{-/-} and two *AMBRA1*^{+/+} clones were pulsed with EdU for 45 min before fixation with 4% paraformaldehyde, and immunofluorescent staining for the indicated proteins. Detection of EdU was accomplished using the Click-iT Plus EdU Alexa Fluor 647 Imaging Kit (Thermo Fisher Scientific). DAPI was used to detect DNA. Scale bars, 20 µm. **b**, *AMBRA1*^{+/+} and *AMBRA1*^{-/-} T98G cells were serum-deprived for 72 h. Cells were then released with the re-addition of 10% FBS and collected at the indicated times. Cell extracts were immunoblotted as indicated. **c**, Representative images of *AMBRA1*^{+/+} and *AMBRA1*^{-/-} U-2 OS cells stably infected with retroviruses expressing mAzG-CCND1 and mPlum-PCNA and monitored via live-cell imaging. Note that mPlum-PCNA forms nuclear foci during S phase. Scale bar, 20 µm. **d**, Quantification of mAzG-CCND1 changes in mean nuclear intensity in single cells tracked over time related

to **c**. Plots represent percentage change from $t = 0$ (PCNA foci formation). Lines show mean and the shaded regions represent the 95% confidence interval. *AMBRA1* WT, $n = 12$ cells; *AMBRA1* KO, $n = 34$ cells. **e**, U-2 OS cells were infected with a retrovirus expressing mAzG-cyclin D1, and transfected with a non-targeting siRNA or an siRNA against *AMBRA1* for two rounds. Cells were then exposed to UV (5 mJ cm^{-2}) irradiation for the indicated times before collection. Cell extracts were immunoblotted as indicated. **f**, HCT-116 cells were infected with a retrovirus expressing mAzG-cyclin D1, and transfected with a non-targeting siRNA or an siRNA against *AMBRA1* for two rounds. Cells were then exposed to UV (5 mJ cm^{-2}) irradiation for the indicated times before collection. Cell extracts were immunoblotted as indicated. **g**, *AMBRA1*^{+/+} and *AMBRA1*^{-/-} HCT-116 and U-2 OS cells were maintained in normal medium or nutrient-deprived in EBSS for 4 h before collection. Cell extracts were immunoblotted as indicated. **h**, HCT-116 cells were transfected with a non-targeting siRNA or combinations of siRNAs against *AMBRA1*, *E2F1*, *E2F2* and *E2F3* as indicated for three rounds before lysis and immunoblotting for the indicated proteins. **i**, T98G cells were transfected with an sgRNA control or an sgRNA targeting *AMBRA1*. *AMBRA1*^{-/-} clones positive for gene editing were pooled together. Parental and *AMBRA1*^{-/-} pooled T98G cells were then infected with a CDK4/6 activity reporter⁵⁴ and seeded in glass-bottom plates. Three days after seeding, cells were stained with Hoechst 33342 before imaging. Scale bars, 50 μm . **j**, *AMBRA1*^{+/+} and *AMBRA1*^{-/-} U-2 OS cells were fixed in 4% paraformaldehyde and immunostained as indicated. DAPI was used to detect DNA. Scale bars, 20 μm . **k**, HCT-116, U-2 OS and hTERT RPE-1 cells were treated with leptomycin B (LMB) for 4 h before exposure to cycloheximide for the indicated times. Cell extracts were immunoblotted as indicated. **l**, U-2 OS cells were infected with retroviruses expressing either mAZG-tagged wild-type cyclin D1 or cyclin D1 fused to a nuclear localization signal (NLS). Cells were then treated with cycloheximide for the indicated times before collection. Cell extracts were immunoblotted as indicated. Unless otherwise noted, experiments were performed at least three independent times. [Source data](#)

Extended Data Fig. 6 Report of MPPH-derived mutations in *CCND2* and evidence of the ability of abemaciclib to cross the

murine placental barrier.

a, Graphical representation of MPPH-associated *CCND2* mutations affecting T280 and P281 in cyclin D2^{55,56,57,58,59,60}. **b**, Genotype and clinical phenotypes of patient 295112 (<https://decipher.sanger.ac.uk>)⁵⁹. **c**, Genotype and clinical phenotypes of patient 305416 (<https://decipher.sanger.ac.uk>)⁵⁹. **d**, Representative immunofluorescent sections (20× magnification) of the mesencephalic ventricular zone of E13.5 *Ambra1*^{+/+} and *Ambra1*^{gt/gt} embryos stained for cyclin D1, cyclin D2 or p-RB (S807/811). Scale bars, 100 µm. **e**, Experimental timeline of the daily, oral administration of abemaciclib to pregnant *Ambra1*^{gt/+} mice (from E8.5 to E13.5) (1), and collection of amniotic fluid via a 28G insulin syringe from the amniotic sacs of E13.5 embryos for subsequent LC-MS/MS analysis (2). **f**, Mass spectrometry imaging of abemaciclib in mouse embryos. Left, optical registration image of the slide with each section before MALDI imaging. Right, extracted ion mobility image for abemaciclib (green) as detected based on authentic standards [M+H]⁺ at 507.2791 *m/z* (±15 ppm) with a mobility (1/K0) of 1.177 (±1%), maternal dose of 100 mg kg⁻¹. Scale bar, 2 mm. This experiment was performed once. **g**, **h**, MS intensity of abemaciclib in tissue samples (**g**), interpolated abemaciclib concentration in maternal plasma (**h**, left) and interpolated abemaciclib concentration in amniotic fluid (**h**, right). Maternal brain, embryonic tissue, amniotic fluid and maternal plasma samples were obtained from mice administrated with daily dosages of abemaciclib mesylate at 8 mg kg⁻¹ or 16 mg kg⁻¹ from E8.5 to E13.5, and at 100 mg kg⁻¹ from E12.5 to E13.5. The abemaciclib MS intensity shown for tissue samples was corrected based on a 3× signal-to-noise cut-off with a floor of 10,000 using the blank controls. The abemaciclib concentration shown in both the amniotic fluid and plasma sample was interpolated based on the standard curve prepared in blank amniotic fluid and blank plasma, respectively. N.D., not detected. This experiment was performed once. **i**, **j**, E13.5 *Ambra1*^{+/+} and *Ambra1*^{gt/gt} embryos were collected from pregnant mice treated with vehicle (PBS) or 8 mg kg⁻¹ abemaciclib. Immunofluorescence from sections was quantified to show area of p-RB (S807/811) (**i**) and cyclin D2 (**j**) staining in the mesencephalic ventricular zone (MVZ). Data are mean ± s.d. p-Rb (S807/811), *n* = 3; cyclin D2, *n* = 3.

Adjusted *P* values were calculated using a two-way ANOVA with Tukey's multiple comparisons test. **k**, Protein extracts derived from E9.5, E10.5 and E13.5 *Ambra1*^{+/+} and *Ambra1*^{gt/gt} embryo heads (four independent embryos per group) were immunoblotted for the indicated proteins. Unless otherwise noted, experiments were performed at least three independent times. [Source data](#)

Extended Data Fig. 7 AMBRA1 and D-type cyclins in human cancer.

a, Lollipop plots showing amino acid positions and numbers of mutations in *AMBRA1*, *FBXO11*, *CCNF* and *KEAP1* in pan-cancer samples generated from cBioPortal (<https://www.cbioperl.org/>) and using the curated set of non-redundant studies^{36,37} (last accessed: December 2019). **b**, Scatter plot of cyclin D1 protein levels (RPPA) and mRNA expression levels of *AMBRA1* (FPKM-UQ) in different TCGA cohorts. LGG, *n* = 435; OV, *n* = 257; PRAD, *n* = 344; TGCT, *n* = 124. Pearson correlation analysis was performed to calculate *P* and *R* values. The shaded area represents the 95% confidence interval of the regression line. **c**, Representative immunohistochemistry image fields of *AMBRA1*^{-/-} and *AMBRA1*^{+/+} HCT-116 plasma-thrombin cell blocks stained with either a rabbit polyclonal anti-human *AMBRA1* antibody (Proteintech Group 13762-1, lot nos. 00005112 and 00017852, RRID: AB_10642007), a mouse anti-human *AMBRA1* clone G-6 antibody (Santa Cruz Biotechnology sc-398204, lot no. G1217, RRID: AB_2861324), or a rabbit anti-human cyclin D1 clone SP4 antibody (Ventana Medical Systems 790-4508, lot no. B08078, RRID: AB_2335988). Scale bar, 50 µm. This experiment was performed once. **d**, A TMA containing normal and tumour ovarian human specimens was immunostained with antibodies to *AMBRA1* and cyclin D1. Violin plots show automated quantification of cyclin D1 intensity levels and pathological scoring of nuclear *AMBRA1* intensity levels (0, no expression; 1, faint expression; 2, mild or moderate expression) (*n* = 44 biologically independent samples). Adjusted *P* values were calculated using a one-way ANOVA with Dunnett's multiple comparisons test. **e**, The graph shows automated quantification of cyclin D1 and *AMBRA1* intensity levels from the TMA described in **d**. *n* = 44 biologically independent samples. Pearson

correlation analysis was performed to calculate P and R^2 values. The shaded area represents the 95% confidence interval of the regression line. **f**, Kaplan–Meier plots showing survival probability of human cancer patients partitioned in low *AMBRA1* mRNA levels ($n = 2,110$ for quantiles 0.2/0.8, and $n = 5,273$ for quantiles 0.5/0.5) and high *AMBRA1* mRNA levels ($n = 2,110$ for quantiles 0.2/0.8, and $n = 5,273$ for quantiles 0.5/0.5; TCGA (<https://cancer.gov/tcg>)). P values were calculated using the log-rank test. Meta P values: 0.0389 (for quantiles 0.2/0.8), 0.0475 (for quantiles 0.5/0.5). **g**, Priority targets from pan-cancer and cancer-type analyses, as previously described⁹. *CCND1*, *CCND2*, *CDK4* and *CDK6* are highlighted in red. Select oncogenes are in black. [Source data](#)

Extended Data Fig. 8 C-terminal cancer mutations affecting cyclin D1, cyclin D2 and cyclin D3.

a–c, Lollipop plots showing the amino acid positions and their mutation numbers in the degron motif of *CCND1* (**a**), *CCND2* (**b**) and *CCND3* (**c**) in pan-cancer, gathered and combined from cBioPortal (curated set of non-redundant studies)^{36,37}, COSMIC (<https://cancer.sanger.ac.uk/cosmic>)³⁸, and Schmitz et al.³⁹ (indicated by the dagger). Black arrows in **c** indicate mutations identified in haematopoietic and lymphoid malignancies. del, deletion; fs, frameshift; *, stop codon. Last accessed October 2019.

Extended Data Fig. 9 AMBRA1 CRISPR gene score distribution in DLBCL.

a–c, Average pan- (**a**), subtype-specific (**b**) and cell-type specific (**c**) DLBCL CRISPR gene score distribution, as reported¹¹. ABC, activated B cell-like; GCB, germinal centre B cell-like. Putative oncogenes are depicted in blue; putative tumour suppressors are depicted in red.

Extended Data Fig. 10 AMBRA1 is a tumour suppressor in DLBCL.

a, Differential expression of *AMBRA1* in DLBCL compared with normal B-lymphocytes from Compagno et al.³⁴. $P = 9.91 \times 10^{-6}$. **b**, Differential

expression of *AMBRA1* in DLBCL compared with normal B-lymphocytes from Brune et al.³⁵. $P = 0.003$. **a, b**, Dots indicate maxima and minima; whiskers indicate 90th and 10th percentiles; box edges indicate 75th and 25th percentiles; and centre lines indicate median; P values were calculated using a Student's *t*-test, as described (<https://www.oncomine.org>). **c**, TCGA-derived, CpG-aggregated methylation values (β -values) of *AMBRA1* and other established hypermethylated genes in cancer (depicted in red) compared with hypomethylated genes (depicted in green) in DLBCL. Solid lines denote the mean of β -values for each group. **d**, Kaplan–Meier plot showing survival probability of patients with DLBCL partitioned in low *AMBRA1* mRNA levels ($n = 50$) and high *AMBRA1* mRNA levels ($n = 49$), combining GEO number GSE23501³¹ and GEO number GSE10846³², using probe 52731_at (<https://www.oncomine.org>) (0.10/0.90 quantiles). P values by log-rank test. **e**, OncoPrint map of the genetic alterations in *AMBRA1* and selected RB-pathway genes; TCGA DLBC cohort. **f**, Heat map showing hierarchical clustering of transcriptional signatures of tumours bearing alterations in either *AMBRA1* or *CCND3*, derived from TCGA DLBC cohort. del, deletion; mut, mutation; amp, amplification; other, tumours bearing unaltered *AMBRA1* and *CCND3*. *AMBRA1* (RefSeq XM_005253009) mutations: R439K and D1287G. *CCND3* (RefSeq NM_001760) mutation: T283A. **g**, Gene enrichment analysis of the top 300 most differentially expressed genes in patients harbouring alterations in *AMBRA1* and *CCND3* from TCGA (DLBC cohort) was performed using Enrichr (<https://amp.pharm.mssm.edu/Enrichr/>) with BioPlant (2019), Reactome (2016) and NCI-Nature (2016) databases. Asterisks indicate pathways regulated by D-type cyclins. **h**, BJAB cells were infected with lentiviruses expressing either one of three independent sgRNAs targeting *AMBRA1* or one sgRNA targeting luciferase as control (CTRL). Upon FACS sorting, cell populations were treated with cycloheximide for the indicated times, MG132 or MLN4924 (for 4 h) before collection. Cell extracts were immunoblotted for the indicated proteins. **i**, BJAB cells expressing either control sgRNA targeting luciferase or sgRNA targeting *AMBRA1* were xenotransplanted in NSG mice via subcutaneous flank injections. Tumour volume was calculated by caliper measurement. Data are mean tumour volume \pm s.d. ($n = 3$ per group); nonlinear (exponential growth) fitted curves. P values by unpaired, multiple-comparison *t*-test using the Holm–Sidak method. Day 0: NS; day 13: $P = 0.1885$; day 14

$P = 0.0045$; day 15: $P = 0.0083$; day 16: $P = 0.0048$; day 17: $P = 0.0223$. * $P < 0.05$. **j**, The weight of tumours described in **i** was measured on excised tumours at the experimental end point. Data are mean tumour weight \pm s.e.m. ($n = 3$ per group). P values by unpaired *t*-test. **k**, U-2932 cells were infected with lentiviruses expressing either one of three independent sgRNAs targeting *AMBRA1* or one sgRNA targeting luciferase as control (CTRL). Upon FACS sorting, cell populations were collected, and cell extracts were immunoblotted as indicated. **l**, U-2932 expressing either control sgRNA targeting luciferase or sgRNA targeting *AMBRA1* were xenotransplanted in NSG mice via subcutaneous flank injections. Tumour volume was calculated by caliper measurement. Data are mean tumour volume \pm s.e.m. ($n = 5$ per group at day 0). P values by unpaired, multiple-comparison *t*-test using the Holm–Sidak method, until day 40 (the latest time point where at least 3 mice per group were available). Day 23: $P = 0.0008$; day 25: $P = 0.0191$; day 26: $P = 0.0133$; day 27: $P = 0.0041$; day 30: $P = 0.0079$; day 31: $P = 0.0292$; day 32: $P = 0.0148$; day 33: $P = 0.0002$; day 34: $P = 0.0001$; day 37: $P = 0.0165$; day 38: $P = 0.0306$; day 39: $P = 0.0240$; day 40: $P = 0.0401$. * $P < 0.05$. These data are from the same dataset shown in Fig. [4c](#). Unless otherwise noted, experiments were performed at least three independent times. [Source data](#)

Extended Data Fig. 11 Loss of *AMBRA1* promotes the formation of CDK2–D-type cyclins complexes.

a, Schematic representation of the experiment used to identify global regulators of the response to CDK4/6 inhibitors in U-2 OS cells using a whole-genome, CRISPR–Cas9 human GECKOv2 pooled library. **b**, Scatter plot of the candidate hits generated from **a**. This experiment was performed once, with four technical replicates. P values were calculated using the MAGeCK algorithm¹⁴. **c**, Parental and *AMBRA1*^{−/−} HCT-116, T98G or U-2 OS cells were treated with various concentrations of abemaciclib. Survival curves were generated using the AlamarBlue assay. Data are mean survival percentage \pm s.d. ($n = 3$ biological replicates per group). The signal from each condition was normalized to an untreated control to adjust for any relative growth difference. **d**, *AMBRA1*^{+/+} and *AMBRA1*^{−/−} T98G cells were treated with abemaciclib (75 nM) or DMSO for 24 h before collection and immunoblotting for the indicated proteins. **e**, HCT-116 cells were

transfected with a non-targeting siRNA or an siRNA against *AMBRA1* for three rounds before treatment with abemaciclib (75 nM) or DMSO for 24 h. Cells were collected and their extracts were immunoblotted for the indicated proteins. **f**, Parental and *AMBRA1*^{-/-} HCT-116 cells were collected, and their extracts were used for an immunoprecipitation with an anti-CDK4 antibody or control IgG isotype. Proteins were immunoblotted as indicated. **g**, HEK 293T were transfected with plasmids expressing either FFSS-tagged cyclin D1, cyclin D2 and cyclin D3 or empty vector before lysis and immunoprecipitation with an anti-Flag resin. Proteins were immunoblotted as indicated. **h**, *AMBRA1*^{-/-} HCT-116 cells were collected, and their extracts were used for an immunoprecipitation with an anti-cyclin D1 antibody or control IgG isotype. Proteins were immunoblotted as indicated. **i**, *AMBRA1*^{-/-} HCT-116 cells were collected, and their extracts were used for an immunoprecipitation with an anti-cyclin D3 antibody or control IgG isotype. Proteins were immunoblotted as indicated. **j**, HCT-116 cells were infected with retroviruses expressing either FH-tagged cyclin D1(WT) or cyclin D1(T286A), and their extracts were used for immunoprecipitation with anti-CDK2 antibody or control IgG isotype. Proteins were immunoblotted as indicated. Unless otherwise noted, experiments were performed at least three independent times. [Source data](#)

Extended Data Fig. 12 Loss of *AMBRA1* desensitizes cells to CDK4/6 inhibitors.

a, HEK 293T cells were (co)transfected for 48 h with FFSS-tagged cyclin D2 or FFSS-tagged cyclin D3, HA-tagged CDK2, or FH-tagged empty vector, in the combinations indicated. After a double-immunoprecipitation, cyclin D–CDK2 complexes were incubated with GST–RB(773-CT) for 30 min with or without 100 nM, 250 nM or 500nM palbociclib, or 500 nM flavopiridol, before stopping the reaction with Laemmli buffer. Proteins were immunoblotted as indicated. **b**, HEK 293T cells were (co)transfected for 48 h with FFSS-tagged cyclin D1, HA-tagged CDK2, HA-tagged CDK4 or FH-tagged empty vector in the combinations indicated. After a double immunoprecipitation, the purified cyclin D1–CDK4 and cyclin D1–CDK2 complexes were incubated with GST–RB(773-CT) for 30 min with or without 100 nM palbociclib or 100 nM flavopiridol, before stopping the reaction with Laemmli buffer. Proteins were immunoblotted as indicated. **c**,

HEK 293T cells were (co)transfected for 48 h with FFSS-tagged cyclin D3, HA-tagged CDK2, HA-tagged CDK4, or FH-tagged empty vector in the combinations indicated. After a double immunoprecipitation (anti-Flag immunoprecipitation, elution with 3×Flag peptide, then anti-HA immunoprecipitation), cyclin D3–CDK4 and cyclin D3–CDK2 complexes bound to the beads were incubated with GST–RB(773-CT) for 30 min with or without 100 nM, 250 nM or 500nM palbociclib, or 500 nM flavopiridol, before stopping the reaction with Laemmli buffer. Proteins were immunoblotted as indicated. **d**, Parental and *AMBRA1*^{-/-} HCT-116 cells were transfected with a non-targeting siRNA or an siRNA against *CDK2* for three rounds before treatment with various concentrations of palbociclib (left) or abemaciclib (right). Survival curves were generated using the AlamarBlue assay. Data are mean survival percentage ± s.d. ($n = 3$ biological replicates per group). The signal from each condition was normalized to an untreated control to adjust for any relative growth difference. **e**, Before treatment, cell extracts from **d** were immunoblotted as indicated. **f**, Working model: by modulating the levels of D-type cyclins, CRL4^{AMBRA1} controls the execution of cell-cycle progression, thus ensuring proper embryonic development and tumour suppression (top). Impairment of CRL4^{AMBRA1}-mediated degradation of D-type cyclins (for example, alterations or downregulation of *AMBRA1* or mutations in the TP motif of D-type cyclins) deregulates cell-cycle progression, and results in embryonic lethality, increased tumorigenic potential, and reduced sensitivity to CDK4/6 inhibitors (bottom). Unless otherwise noted, experiments were performed at least three independent times. [Source data](#)

Supplementary information

[Supplementary Figure 1](#)

This file contains the full scanned images obtained by electrophoretic separation.

[Reporting Summary](#)

[Supplementary Figures](#)

This file contains the representative FACS gating strategy (Supplementary Figure 2), additional information on the identification and detection of abemaciclib by LC-MS/MS (Supplementary Figures 3-5) and the representative image of a stripped membrane for immunoblotting (Supplementary Figure 6).

Peer Review File

Supplementary Table 1

The interactors of wild-type cyclin D1 (WT) or mutant cyclin D1(T286A) identified by LC-MS/MS and related to Fig. 1d. For each protein reported, the UniProt accession number is indicated in the first column. # Peptides = total number of PSMs, discounting duplicates. # PSM = total number of PSMs, including duplicate PSMs.

Supplementary Table 2

All cases of MPPH-derived mutations in *CCND2* described in the literature^{55–60} or deposited in Decipher (<https://decipher.sanger.ac.uk>).

Supplementary Table 3

The mutations in the degron motif of *CCND1*, *CCND2* and *CCND3* gathered and combined from cBioPortal (curated set of non-redundant studies, <https://www.cbioportal.org/>)^{36,37}, COSMIC (<https://cancer.sanger.ac.uk/cosmic>)³⁸ and, and Schmitz *et al.*³⁹. “del” = deletion, “fs” = frameshift, “BL” = Burkitt’s lymphoma, “HIV+” = Human immunodeficiency virus positive, “ABC” = Activated B cell-like, “GCB” = Germinal center B cell-like, “DLBL” = Diffuse large B-cell lymphoma, “NOS” = Not otherwise specified, “AML” = Acute myeloid leukemia, “FL” = Follicular lymphoma. Note, some sample IDs refer to patient-derived cell lines.

Supplementary Table 4

Sequences of primers, siRNA oligos, and gRNAs used in this study. Where applicable, sources and catalog numbers are indicated.

Source data

[Source Data Fig. 1](#)

[Source Data Fig. 2](#)

[Source Data Fig. 3](#)

[Source Data Fig. 4](#)

[Source Data Extended Data Fig. 1](#)

[Source Data Extended Data Fig. 2](#)

[Source Data Extended Data Fig. 4](#)

[Source Data Extended Data Fig. 5](#)

[Source Data Extended Data Fig. 6](#)

[Source Data Extended Data Fig. 7](#)

[Source Data Extended Data Fig. 10](#)

[Source Data Extended Data Fig. 11](#)

[Source Data Extended Data Fig. 12](#)

Rights and permissions

[Reprints and Permissions](#)

About this article



Cite this article

Simoneschi, D., Rona, G., Zhou, N. *et al.* CRL4^{AMBRA1} is a master regulator of D-type cyclins. *Nature* **592**, 789–793 (2021). <https://doi.org/10.1038/s41586-021-03445-y>

[Download citation](#)

- Received: 08 March 2020
- Accepted: 11 March 2021
- Published: 14 April 2021
- Issue Date: 29 April 2021
- DOI: <https://doi.org/10.1038/s41586-021-03445-y>

Further reading

- [**The path to destruction for D-type cyclin proteins**](#)
 - Charupong Saengboonmee
 - & Piotr Sicinski
- [**AMBRA1 regulates cyclin D to guard S-phase entry and genomic integrity**](#)

- Emiliano Maiani
- , Giacomo Milletti
- , Francesca Nazio
- , Søs Grønbæk Holdgaard
- , Jirina Bartkova
- , Salvatore Rizza
- , Valentina Cianfanelli
- , Mar Lorente
- , Daniele Simoneschi
- , Miriam Di Marco
- , Pasquale D'Acunzo
- , Luca Di Leo
- , Rikke Rasmussen
- , Costanza Montagna
- , Marilena Raciti
- , Cristiano De Stefanis
- , Estibaliz Gabicagogeascoa
- , Gergely Rona
- , Nélida Salvador
- , Emanuela Pupo
- , Joanna Maria Merchut-Maya
- , Colin J. Daniel
- , Marianna Carinci
- , Valeriana Cesarini
- , Alfie O'sullivan
- , Yeon-Tae Jeong
- , Matteo Bordi
- , Francesco Russo
- , Silvia Campello
- , Angela Gallo
- , Giuseppe Filomeni
- , Letizia Lanzetti
- , Rosalie C. Sears
- , Petra Hamerlik
- , Armando Bartolazzi
- , Robert E. Hynds
- , David R. Pearce

- , Charles Swanton
- , Michele Pagano
- , Guillermo Velasco
- , Elena Papaleo
- , Daniela De Zio
- , Apolinar Maya-Mendoza
- , Franco Locatelli
- , Jiri Bartek
- & Francesco Cecconi

Nature (2021)

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

Associated Content

[The path to destruction for D-type cyclin proteins](#)

- Charupong Saengboonmee
- Piotr Sicinski

Nature News & Views 14 Apr 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03445-y>.

- Article
- [Published: 14 April 2021](#)

The AMBRA1 E3 ligase adaptor regulates the stability of cyclin D

- [Andrea C. Chaikovsky](#) ORCID: [orcid.org/0000-0003-1842-5384^{1,2}](https://orcid.org/0000-0003-1842-5384),
- [Chuan Li³](#),
- [Edwin E. Jeng²](#),
- [Samuel Loebell^{1,2}](#),
- [Myung Chang Lee](#) ORCID: [orcid.org/0000-0002-7051-5822^{1,2}](https://orcid.org/0000-0002-7051-5822),
- [Christopher W. Murray^{2,4}](#),
- [Ran Cheng](#) ORCID: [orcid.org/0000-0002-7617-3018³](https://orcid.org/0000-0002-7617-3018),
- [Janos Demeter](#) ORCID: [orcid.org/0000-0002-7301-8055⁵](https://orcid.org/0000-0002-7301-8055),
- [Danielle L. Swaney^{6,7,8}](#),
- [Si-Han Chen](#) ORCID: [orcid.org/0000-0002-2890-9764^{6,7,8}](https://orcid.org/0000-0002-2890-9764),
- [Billy W. Newton^{6,7,8}](#),
- [Jeffrey R. Johnson^{6,7,8}](#),
- [Alexandros P. Drainas^{1,2}](#),
- [Yan Ting Shue^{1,2}](#),
- [Jose A. Seoane](#) ORCID: [orcid.org/0000-0002-3856-9177^{2,9}](https://orcid.org/0000-0002-3856-9177),
- [Preethi Srinivasan^{2,9}](#),
- [Andy He^{1,2}](#),
- [Akihiro Yoshida^{10,11}](#),
- [Susan Q. Hipkins^{1,2}](#),
- [Edel McCrea^{1,2}](#),
- [Carson D. Poltorack^{1,2}](#),
- [Nevan J. Krogan](#) ORCID: [orcid.org/0000-0003-4902-337X^{6,7,8}](https://orcid.org/0000-0003-4902-337X),
- [J. Alan Diehl](#) ORCID: [orcid.org/0000-0001-9854-5049^{10,11}](https://orcid.org/0000-0001-9854-5049),
- [Christina Kong⁴](#),

- [Peter K. Jackson⁵](#),
- [Christina Curtis](#) [ORCID: orcid.org/0000-0003-0166-3802^{2,9}](#),
- [Dmitri A. Petrov³](#),
- [Michael C. Bassik](#) [ORCID: orcid.org/0000-0001-5185-8427²](#),
- [Monte M. Winslow](#) [ORCID: orcid.org/0000-0002-5730-9573^{2,4}](#) &
- [Julien Sage](#) [ORCID: orcid.org/0000-0002-8928-9968^{1,2}](#)

[Nature](#) volume **592**, pages 794–798(2021) [Cite this article](#)

- 6593 Accesses
- 2 Citations
- 66 Altmetric
- [Metrics details](#)

Subjects

- [Cell division](#)
- [Ubiquitin ligases](#)

Abstract

The initiation of cell division integrates a large number of intra- and extracellular inputs. D-type cyclins (hereafter, cyclin D) couple these inputs to the initiation of DNA replication¹. Increased levels of cyclin D promote cell division by activating cyclin-dependent kinases 4 and 6 (hereafter, CDK4/6), which in turn phosphorylate and inactivate the retinoblastoma tumour suppressor. Accordingly, increased levels and activity of cyclin D–CDK4/6 complexes are strongly linked to unchecked cell proliferation and cancer^{2,3}. However, the mechanisms that regulate levels of cyclin D are incompletely understood^{4,5}. Here we show that autophagy and beclin 1 regulator 1 (AMBRA1) is the main regulator of the degradation of cyclin D. We identified *AMBRA1* in a genome-wide screen to investigate the genetic basis of the response to CDK4/6 inhibition. Loss of AMBRA1 results in

high levels of cyclin D in cells and in mice, which promotes proliferation and decreases sensitivity to CDK4/6 inhibition. Mechanistically, AMBRA1 mediates ubiquitylation and proteasomal degradation of cyclin D as a substrate receptor for the cullin 4 E3 ligase complex. Loss of AMBRA1 enhances the growth of lung adenocarcinoma in a mouse model, and low levels of AMBRA1 correlate with worse survival in patients with lung adenocarcinoma. Thus, AMBRA1 regulates cellular levels of cyclin D, and contributes to cancer development and the response of cancer cells to CDK4/6 inhibitors.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

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: AMBRA1 loss regulates the response to CDK4/6 inhibition as well as levels of cyclin D.

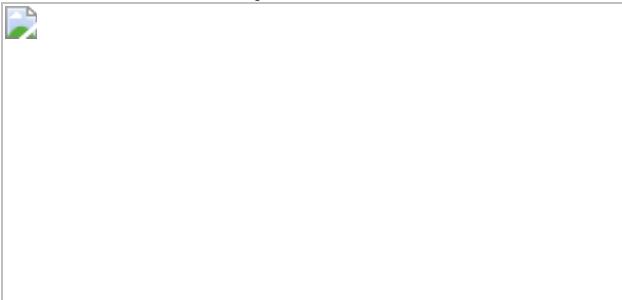


Fig. 2: AMBRA1 regulates the stability of cyclin D.

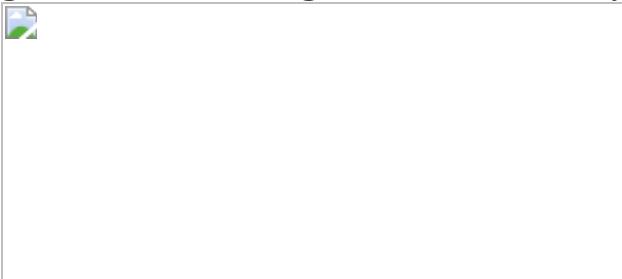
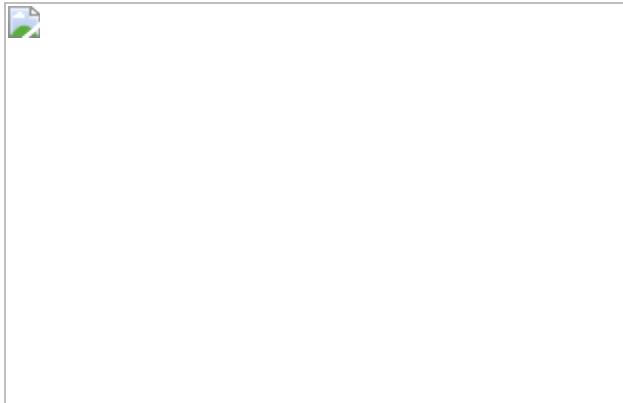


Fig. 3: CRL4^{AMBRA1} ubiquitylates cyclin D.



Fig. 4: AMBRA1 is a tumour suppressor in KRAS-mutant lung adenocarcinoma.



Data availability

Sequencing data from Tuba-seq experiments and RNA-sequencing data from *AMBRA1*-knockout U2OS cells are available from the Gene Expression Omnibus under accession numbers [GSE146303](#) and [GSE159920](#), respectively. Mass spectrometry data from shotgun proteomics experiments and analysis of ubiquitylated proteins are available through the ProteomeXchange Consortium, with dataset identifiers PXD021789 and PXD022111, respectively. Public nonprotected RNA-sequencing, copy number alteration, exome sequencing and reverse-phase protein array lung adenocarcinoma datasets from the TCGA were downloaded from <https://gdc.cancer.gov/>. Clinical data were obtained from a previous publication³⁶ (PMID: 29625055). Gene dependency data from the Cancer Dependency Map are publicly available at www.depmap.org. Protein sequences for mass spectrometry analysis were obtained from the NCBI *Homo sapiens* protein database (ftp://ftp.ncbi.nlm.nih.gov/refseq/release/release-notes/archive/RefSeq-release86.txt, downloaded 05/11/2018) (shotgun mass spectrometry) and from Uniprot (<https://www.uniprot.org/uniprot/?query=proteome:UP000005640%20reviewed:yes>, downloaded 02/28/2020) (ubiquitin remnant profiling). All other data are available in the Article and [Supplementary Information](#), or from the corresponding author upon reasonable request. [Source data](#) are provided with this paper.

References

1. 1.

Deshpande, A., Sicinski, P. & Hinds, P. W. Cyclins and cdks in development and cancer: a perspective. *Oncogene* **24**, 2909–2915 (2005).

[CAS](#) [Article](#) [Google Scholar](#)

2. 2.

Sherr, C. J. Cancer cell cycles. *Science* **274**, 1672–1677 (1996).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

3. 3.

Malumbres, M. & Barbacid, M. Cell cycle, CDKs and cancer: a changing paradigm. *Nat. Rev. Cancer* **9**, 153–166 (2009).

[CAS](#) [Article](#) [Google Scholar](#)

4. 4.

Kanie, T. et al. Genetic reevaluation of the role of F-box proteins in cyclin D1 degradation. *Mol. Cell. Biol.* **32**, 590–605 (2012).

[CAS](#) [Article](#) [Google Scholar](#)

5. 5.

Qie, S. & Diehl, J. A. Cyclin D degradation by E3 ligases in cancer progression and treatment. *Semin. Cancer Biol.* **67**, 159–170 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

6. 6.

Sherr, C. J., Beach, D. & Shapiro, G. I. Targeting CDK4 and CDK6: from discovery to therapy. *Cancer Discov.* **6**, 353–367 (2016).

[CAS](#) [Article](#) [Google Scholar](#)

7. 7.

Walter, D. M. et al. RB constrains lineage fidelity and multiple stages of tumour progression and metastasis. *Nature* **569**, 423–427 (2019).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

8. 8.

Wander, S. A. et al. The Genomic landscape of intrinsic and acquired resistance to cyclin-dependent kinase 4/6 inhibitors in patients with hormone receptor-positive metastatic breast cancer. *Cancer Discov.* **10**, 1174–1193 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

9. 9.

Schoninger, S. F. & Blain, S. W. The ongoing search for biomarkers of CDK4/6 inhibitor responsiveness in breast cancer. *Mol. Cancer Ther.* **19**, 3–12 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

10. 10.

Fimia, G. M. et al. Ambra1 regulates autophagy and development of the nervous system. *Nature* **447**, 1121–1125 (2007).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

11. 11.

Antonioli, M. et al. AMBRA1 interplay with cullin E3 ubiquitin ligases regulates autophagy dynamics. *Dev. Cell* **31**, 734–746 (2014).

[CAS](#) [Article](#) [Google Scholar](#)

12. 12.

Cianfanelli, V. et al. AMBRA1 links autophagy to cell proliferation and tumorigenesis by promoting c-Myc dephosphorylation and degradation. *Nat. Cell Biol.* **17**, 706 (2015).

[CAS](#) [Article](#) [Google Scholar](#)

13. 13.

Montaudon, E. et al. PLK1 inhibition exhibits strong anti-tumoral activity in CCND1-driven breast cancer metastases with acquired palbociclib resistance. *Nat. Commun.* **11**, 4053 (2020).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

14. 14.

Alt, J. R., Cleveland, J. L., Hannink, M. & Diehl, J. A. Phosphorylation-dependent regulation of cyclin D1 nuclear export and cyclin D1-dependent cellular transformation. *Genes Dev.* **14**, 3102–3114 (2000).

[CAS](#) [Article](#) [Google Scholar](#)

15. 15.

Lin, D. I. et al. Phosphorylation-dependent ubiquitination of cyclin D1 by the SCF^{FBX4-αB} crystallin complex. *Mol. Cell* **24**, 355–366 (2006).

[CAS](#) [Article](#) [Google Scholar](#)

16. 16.

Ewen, M. E. et al. Functional interactions of the retinoblastoma protein with mammalian D-type cyclins. *Cell* **73**, 487–497 (1993).

[CAS](#) [Article](#) [Google Scholar](#)

17. 17.

Jahn, S. C. et al. Assembly, activation, and substrate specificity of cyclin D1/Cdk2 complexes. *Biochemistry* **52**, 3489–3501 (2013).

[CAS](#) [Article](#) [Google Scholar](#)

18. 18.

Chytíl, A. et al. Construction of a cyclin D1–Cdk2 fusion protein to model the biological functions of cyclin D1–Cdk2 complexes. *J. Biol. Chem.* **279**, 47688–47698 (2004).

[CAS](#) [Article](#) [Google Scholar](#)

19. 19.

Jansen, V. M. et al. Kinome-wide RNA interference screen reveals a role for PDK1 in acquired resistance to CDK4/6 inhibition in ER-positive breast cancer. *Cancer Res.* **77**, 2488–2499 (2017).

[CAS](#) [Article](#) [Google Scholar](#)

20. 20.

Herrera-Abreu, M. T. et al. Early adaptation and acquired resistance to CDK4/6 inhibition in estrogen receptor-positive breast cancer. *Cancer Res.* **76**, 2301–2313 (2016).

[CAS](#) [Article](#) [Google Scholar](#)

21. 21.

James, M. K., Ray, A., Leznova, D. & Blain, S. W. Differential modification of p27^{Kip1} controls its cyclin D-cdk4 inhibitory activity. *Mol. Cell. Biol.* **28**, 498–510 (2008).

[CAS](#) [Article](#) [Google Scholar](#)

22. 22.

Guiley, K. Z. et al. p27 allosterically activates cyclin-dependent kinase 4 and antagonizes palbociclib inhibition. *Science* **366**, eaaw2106 (2019).

[CAS](#) [Article](#) [Google Scholar](#)

23. 23.

Pagano, M., Theodoras, A. M., Tam, S. W. & Draetta, G. F. Cyclin D1-mediated inhibition of repair and replicative DNA synthesis in human fibroblasts. *Genes Dev.* **8**, 1627–1639 (1994).

[CAS](#) [Article](#) [Google Scholar](#)

24. 24.

Chen, S. H. et al. CRL4^{AMBRA1} targets elongin C for ubiquitination and degradation to modulate CRL5 signaling. *EMBO J.* **37**, e97508 (2018).

[PubMed](#) [PubMed Central](#) [Google Scholar](#)

25. 25.

Xia, P. et al. WASH inhibits autophagy through suppression of beclin 1 ubiquitination. *EMBO J.* **32**, 2685–2696 (2013).

[CAS](#) [Article](#) [Google Scholar](#)

26. 26.

Rogers, Z. N. et al. Mapping the *in vivo* fitness landscape of lung adenocarcinoma tumor suppression in mice. *Nat. Genet.* **50**, 483–486 (2018).

[CAS](#) [Article](#) [Google Scholar](#)

27. 27.

Simoneschi, D. & Pagano, D. CRL4^{AMBRA1} is a regulator of D-type cyclins. *Nature* (2021).

28. 28.

Maiani, E., Milletti, G., Bartek, J. & Cecconi, F. AMBRA1 regulates cyclin D to guard S-phase entry and genomic integrity. *Nature* (2021).

29. 29.

Santra, M. K., Wajapeyee, N. & Green, M. R. F-box protein FBXO31 mediates cyclin D1 degradation to induce G1 arrest after DNA damage. *Nature* **459**, 722–725 (2009).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

30. 30.

Burkhart, D. L. & Sage, J. Cellular mechanisms of tumour suppression by the retinoblastoma gene. *Nat. Rev. Cancer* **8**, 671–682 (2008).

[CAS](#) [Article](#) [Google Scholar](#)

31. 31.

Finn, R. S. et al. PD 0332991, a selective cyclin D kinase 4/6 inhibitor, preferentially inhibits proliferation of luminal estrogen receptor-positive human breast cancer cell lines in vitro. *Breast Cancer Res.* **11**, R77 (2009).

[Article](#) [Google Scholar](#)

32. 32.

Gong, X. et al. Genomic aberrations that activate D-type cyclins are associated with enhanced sensitivity to the CDK4 and CDK6 inhibitor abemaciclib. *Cancer Cell* **32**, 761–776.e6 (2017).

[CAS](#) [Article](#) [Google Scholar](#)

33. 33.

Finn, R. S. et al. Biomarker analyses of response to cyclin-dependent kinase 4/6 inhibition and endocrine therapy in women with treatment-naïve metastatic breast cancer. *Clin. Cancer Res.* **26**, 110–121 (2020).

[CAS](#) [Article](#) [Google Scholar](#)

34. 34.

DeMichele, A. et al. CDK 4/6 inhibitor palbociclib (PD0332991) in Rb+ advanced breast cancer: phase II activity, safety, and predictive biomarker assessment. *Clin. Cancer Res.* **21**, 995–1001 (2015).

[CAS](#) [Article](#) [Google Scholar](#)

35. 35.

Xue, Y. et al. CDK4/6 inhibitors target SMARCA4-determined cyclin D1 deficiency in hypercalcemic small cell carcinoma of the ovary. *Nat. Commun.* **10**, 558 (2019).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

36. 36.

Liu, J. et al. An integrated TCA pan-cancer clinical data resource to drive high-quality survival outcome analytics. *Cell* **173**, 400–416 (2018).

[CAS](#) [Article](#) [Google Scholar](#)

[Download references](#)

Acknowledgements

We thank S. Rubin and J. Skotheim for critical reading of the manuscript; A. Koff, V. Dulic and K. Keyomarsi for helpful discussions; and all of the members of the laboratory of J.S., and especially G. Coles, for their help and support throughout this study. Research reported in this Article was supported by the NIH (J.S., R01CA228413 and 1R35CA231997; A.C.C., 1F99CA245471-01; E.E.J., 2T32CA009302; R.C., 5T32GM007276; M.M.W. and D.A.P., R01CA2344349; and N.J.K., P50AI150476 and U54CA209891), the California TRDRP (J.S., 28IR-0046) and the NSF (GRFP, to A.C.C.). E.E.J. and M.C.L. were supported by a Stanford Graduate Fellowship. S.L. was supported by a Boehringer Ingelheim Fonds MD Fellowship. C.L. is the Connie and Bob Lurie Fellow of the Damon Runyon Cancer Research Foundation (DRG-2331). J.S. is the Harriet and Mary Zelencik Scientist in Children's Cancer and Blood Diseases and the Elaine and John Chambers Professor in Pediatric Cancer.

Author information

Affiliations

1. Department of Pediatrics, Stanford University, Stanford, CA, USA

Andrea C. Chaikovsky, Samuel Loebell, Myung Chang
Lee, Alexandros P. Drainas, Yan Ting Shue, Andy He, Susan Q.
Hipkins, Edel McCrea, Carson D. Poltorack & Julien Sage

2. Department of Genetics, Stanford University, Stanford, CA, USA

Andrea C. Chaikovsky, Edwin E. Jeng, Samuel Loebell, Myung Chang
Lee, Christopher W. Murray, Alexandros P. Drainas, Yan Ting
Shue, Jose A. Seoane, Preethi Srinivasan, Andy He, Susan Q.
Hipkins, Edel McCrea, Carson D. Poltorack, Christina Curtis, Michael
C. Bassik, Monte M. Winslow & Julien Sage

3. Department of Biology, Stanford University, Stanford, CA, USA

Chuan Li, Ran Cheng & Dmitri A. Petrov

4. Department of Pathology, Stanford University, Stanford, CA, USA

Christopher W. Murray, Christina Kong & Monte M. Winslow

5. Department of Microbiology and Immunology, Stanford University, Stanford, CA, USA

Janos Demeter & Peter K. Jackson

6. Department of Cellular and Molecular Pharmacology, University of California, San Francisco, San Francisco, CA, USA

Danielle L. Swaney, Si-Han Chen, Billy W. Newton, Jeffrey R. Johnson & Nevan J. Krogan

7. Quantitative Biosciences Institute (QBI), University of California, San Francisco, San Francisco, CA, USA

Danielle L. Swaney, Si-Han Chen, Billy W. Newton, Jeffrey R. Johnson & Nevan J. Krogan

8. Gladstone Institutes, San Francisco, CA, USA

Danielle L. Swaney, Si-Han Chen, Billy W. Newton, Jeffrey R. Johnson & Nevan J. Krogan

9. Department of Medicine, Stanford University, Stanford, CA, USA

Jose A. Seoane, Preethi Srinivasan & Christina Curtis

10. Department of Biochemistry, School of Medicine, Case Western Reserve University, Cleveland, OH, USA

Akihiro Yoshida & J. Alan Diehl

11. Case Comprehensive Cancer Center, School of Medicine, Case Western Reserve University, Cleveland, OH, USA

Akihiro Yoshida & J. Alan Diehl

Authors

1. Andrea C. Chaikovsky

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Chuan Li

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Edwin E. Jeng

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Samuel Loebell

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Myung Chang Lee

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Christopher W. Murray

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Ran Cheng

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Janos Demeter

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. Danielle L. Swaney

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. Si-Han Chen

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

11. Billy W. Newton

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

12. Jeffrey R. Johnson

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

13. Alexandros P. Drainas

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

14. Yan Ting Shue

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

15. Jose A. Seoane

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

16. Preethi Srinivasan

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

17. Andy He

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

18. Akihiro Yoshida

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

19. Susan Q. Hipkins

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

20. Edel McCrea

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

21. Carson D. Poltorack

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

22. Nevan J. Krogan

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

23. J. Alan Diehl

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

24. Christina Kong

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

25. Peter K. Jackson

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

26. Christina Curtis

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

27. Dmitri A. Petrov

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

28. Michael C. Bassik

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

29. Monte M. Winslow

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

30. Julien Sage

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Contributions

A.C.C. and J.S. designed most of the experiments and interpreted the results. A.C.C. and E.E.J. performed and analysed the CRISPR–Cas9 screen under the supervision of M.C.B.; A.C.C., M.C.L. and C.W.M. performed the Tuba-seq experiments under the supervision of M.M.W.; C.L. performed the computational analysis of the Tuba-seq experiments

under the supervision of M.M.W. and D.P.; A.C.C. and A.H. performed the xenograft experiments. Y.T.S., S.Q.H. and A.H. performed immunostaining; Y.T.S. dissected mouse embryos; C.K. performed the histopathological analysis. J.A.S. and P.S. performed the analysis of human lung cancer data under the supervision of C.C.; S.L., E.M. and C.P. performed experiments related to cell cycle phenotypes under the guidance of A.C.C.; A.P.D. analysed the RNA-sequencing data; A.Y. and J.A.D. helped to prepare and design the protein stability and ubiquitylation experiments; A.C.C, R.C., J.D. and P.K.J. performed and analysed the shotgun mass-spectrometry experiments; D.L.S., S.-H.C., B.W.N., J.R.J. and N.J.K. performed and analysed the ubiquitylation mass-spectrometry experiments; A.C.C. and J.S. wrote the manuscript, with contributions from all authors.

Corresponding author

Correspondence to [Julien Sage](#).

Ethics declarations

Competing interests

J.S. has received research funding from Stemcentrx/Abbvie, Pfizer and Revolution Medicines. M.M.W. and D.P. have equity in, and are advisors for, D2GOncology. C.C. is a scientific advisor to GRAIL and reports stock options as well as consulting for GRAIL and Genentech. N.J.K. has received research support from Vir Biotechnology and F. Hoffmann-La Roche. The authors declare no other competing interests.

Additional information

Peer review information *Nature* thanks Marianne Bronner, Piotr Sicinski 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 AMBRA1 and other factors involved in the response of cells to CDK4/6 inhibitors.

a, Proliferation of U937 cells in the presence of 0.5 μM palbociclib (palbo) over 6 d, determined by cell counting ($n = 1$ experiment). **b**, Immunoassay of total RB and RB phosphorylated at S807 and S811 (p-RB S807/811) in U937 cells over 36 h of palbociclib treatment. **c**, Quantification of phosphorylated RB relative to total RB from **b** ($n = 1$ experiment). **d**, Schematic of the CRISPR–Cas9 screen in U937 cells. **e**, Protein–protein interaction map of screen results, generated using Metascape. Coloured nodes represent densely connected gene neighbourhoods. Legend indicates the gene ontology term that is most significantly enriched within each neighbourhood. Node size indicates the degree of connectedness. Gene names can be found in Supplementary Table 3. **f**, Schematic of the screen results among RB-pathway genes expressed in U937 cells. **g**, Number of control and knockout U937 cells treated with 0.5 μM palbociclib or DMSO control for 48 h. Each symbol is an isogenic clone ($n = 3$ biological replicates per clone). **h**, Left, schematic of the competition assay between GFP-negative parental U937 cells and GFP-positive knockout cell populations. Right, example of flow cytometry analysis for one experiment with *AMBRA1*-knockout cells. **i**, Percentage of GFP-positive control or knockout populations in competition assays as in **h** ($n = 3$ biological replicates). **j**, Representative flow cytometry plots of annexin V and propidium iodide (PI) staining in U937 cells treated with 0.5 μM palbociclib for 24 h. **k**, Percentage of apoptotic (annexinV⁺PI⁺) U937 cells after a 24-h palbociclib treatment ($n = 3$ biological replicates per clone). Palbociclib does not induce apoptosis in any genotype. **l**, Representative flow cytometry plots of BrdU and PI staining in U937 cells treated with 0.5 μM palbociclib for 24 h. **m**, Percentage of S-phase cells by BrdU and PI staining in U937 cells treated with 1 μM abemaciclib for 24 h ($n = 3$ biological replicates per clone). **n**, Immunoassay for AMBRA1 and RB in control and knockout cancer cell lines generated by CRISPR–Cas9. For U2OS (osteosarcoma), NCI-H1792 (lung adenocarcinoma) and NCI-H460 (large cell lung cancer), each lane is an isogenic clone. MCF7 cells

(breast cancer) are populations. **o**, Percentage of cycling S-phase cells from **n** after a 24-h treatment with palbociclib (0.5 μ M for all cell lines except for MCF7 cells, 0.04 μ M). U2OS, NCI-H1792 and NCI-H460 cells were analysed by BrdU and PI staining, and each symbol is an isogenic clone ($n = 3$ biological replicates per clone). MCF7 cells were analysed by PI staining ($n = 3$ biological replicates). **p**, Quantification of RB phosphorylated at S795 (p-RB S795) over total RB in U937 cells treated with increasing doses of palbociclib for 24 h, measured by immunoassay ($n = 4$ biological replicates). **q**, Fold-change in mRNA levels of E2F target genes in U937 cells treated with 0.5 μ M palbociclib for 24 h, measured by quantitative PCR with reverse transcription (RT-qPCR) ($n = 3$ biological replicates). All data are mean \pm s.d. P values calculated by two-sided unpaired *t*-test (**g, k, m, o**) and two-sided paired *t*-test (**i, p, q**). Tubulin, HSP90 and actin are loading controls. [Source data](#)

Extended Data Fig. 2 AMBRA1 loss regulates cyclin D post-transcriptionally and dependency on AMBRA1 correlates with cyclin D signalling networks.

a, b, RT-qPCR analysis of the genes encoding D-type cyclins (CCND genes) and CDK4 in U937 cells (**a**) ($n = 3$ biological replicates per clone) or expressed D-type cyclins in other cancer cell lines (**b**). For U2OS, NCI-H1792 and NCI-H460 cells, each symbol is an isogenic clone ($n = 2$ biological replicates per clone). MCF7 cells are populations ($n = 3$ biological replicates). P values evaluate differences between knockout cells and controls for each gene. **c**, Immunoassay of D-type cyclins in cancer cell lines in **b**. **d**, Immunoassay of AMBRA1, cyclin D1 and CDK4 in U2OS cells after 48 h of *AMBRA1* knockdown by siRNA pools. **e**, Quantification of cyclin D1 and CDK4 protein levels in **d** ($n = 3$ biological replicates). **f**, Correlation of gene dependency scores between *AMBRA1*, RB pathway genes and additional cancer drivers, according to DepMap. Red lines mark the top and bottom 5% of genes. **g**, The 20 most significantly enriched gene ontology terms among the top 100 genes, the loss of which best correlate with loss of *AMBRA1* in DepMap. **h**, Principal component (PC) analysis of RNA-sequencing data from U2OS cells, three biological replicates per cell line. **i**, Volcano plot of RNA-sequencing results comparing control and *AMBRA1*-knockout U2OS cells.

Significantly differentially expressed genes ($P < 0.01$) are in red. All data shown as mean \pm s.d. P values calculated by two-sided unpaired t -test (**a**, **b**), two-sided paired t -test (**e**), hypergeometric test (**g**) and Wald test (**i**).

[Source data](#)

Extended Data Fig. 3 AMBRA1 deletion in mouse embryos results in increased cyclin D levels.

a, sgRNA design to knockout *Ambra1* in mouse zygotes by microinjection of sgRNAs and Cas9. Controls were injected with a non-targeting sgRNA. **b**, Representative bright-field images of control ($n = 5$) and mutant ($n = 3$) embryos at embryonic day (E)13.5. Similar to previous reports¹⁰, the *Ambra1*-mutant embryos generated here have neural tube defects with midbrain and hindbrain exencephaly and/or spina bifida (arrows). Scale bar, 2 mm. **c**, Representative cyclin D immunofluorescence (red signal, the antibody recognizes cyclin D1 and cyclin D2) in control and *Ambra1*-mutant E13.5 embryos (from $n = 3$ embryos per sgRNA). DAPI shows DNA. The liver is autofluorescent. Scale bar, 1 mm. **d**, High-magnification view of the developing brain from one control and one *Ambra1*-mutant embryo (asterisks in **c**). v, ventricle, cp, choroid plexus. Scale bar, 500 μ m. Representative of three embryos per sgRNA.

Extended Data Fig. 4 Pathways previously associated with AMBRA1 do not explain tolerance to CDK4/6 inhibitors.

a, Immunoblot analysis of autophagy flux by LC3 conversion (LC3-I to LC3-II, which occurs during autophagosome formation) and RB phosphorylation (p-RB S795) in U937 cells treated with 0.5 μ M palbociclib for 24 h and acutely treated with 25 μ M chloroquine (CQ) (an autophagy inhibitor) for the final 4 h. **b**, Quantification of LC3-II levels with 4 h of chloroquine treatment, indicating autophagy flux, from cells in **a** ($n = 3$ biological replicates). No significant differences were identified by two-way ANOVA ($P_{\text{cell line}} = 0.44$, $P_{\text{treatment}} = 0.38$, $P_{\text{interaction}} = 0.92$). **c**, Immunoblot of total and phosphorylated RB and LC3 conversion in wild-type U937 cells treated with 0.5 μ M palbociclib, 25 μ M chloroquine or both for 24 h. Representative of three independent experiments. **d**,

Representative flow cytometry plots of BrdU and PI staining in cells from **c**. **e**, Quantification of S-phase cells from **d** ($n = 3$ biological replicates). Autophagy inhibition does not alter palbociclib response. **f**, Immunoassay of the mTORC1 target phosphorylation sites (S2448 of mTOR, and T37 and T46 of 4EBP1) in U937 cells following amino acid starvation. Representative of two independent experiments. **g**, Immunoassay of MYC in U937 clones. **h**, Quantification of MYC from **g**. Each symbol is an isogenic clone ($n = 3$ biological replicates per clone). **i**, Immunoassay of PLK1 and AURKA and immunoblot of AURKB in control and *AMBRA1*-knockout U2OS cells. Each lane is a biological replicate. **j**, Quantification of **i** ($n = 3$ biological replicates). All data are mean \pm s.d. P values calculated by two-way ANOVA (**b**), two-sided paired *t*-test (**e, j**), and two-sided unpaired *t*-test (**h**). HSP90, tubulin and actin are loading controls.

[Source data](#)

Extended Data Fig. 5 Cyclin D mediates the phenotypes of *AMBRA1*-mutant cells.

a, Immunoassay of cyclin D1, D2, and D3 in wild-type U2OS cells overexpressing all three D-type cyclins from the same lentiviral vector or RFP as a control. **b**, Representative flow cytometry plots of BrdU and PI staining in cells from **a** treated with increasing doses of palbociclib for 24 h. **c**, Percentage of cycling S-phase cells from **b** ($n = 3$ biological replicates). Data are mean \pm s.d. P values calculated by two-way ANOVA ($P_{\text{cell line}} < 0.0001$) with post hoc Sidak test. **d**, Representative flow cytometry plots of BrdU and PI staining in U2OS cells overexpressing stabilized cyclin D1(T286A)-HA or RFP control, treated with increasing doses of palbociclib for 24 h. **e**, Representative flow cytometry plots of BrdU and PI staining in control and *AMBRA1*-knockout U2OS clones after 48 h of cyclin D1 (*CCND1*) knockdown with siRNA pools. **f**, Co-immunoprecipitation of p27 in control, knockout and cyclin-D1(T286A)-overexpressing U2OS cells, and immunoassay of relevant protein complexes ($n = 2$ biological replicates). HSP90 is a loading control. [Source data](#)

Extended Data Fig. 6 *AMBRA1* regulates the ubiquitylation of D-type cyclins.

a, Immunoblot analysis of cyclin D3 in wild-type U937 cells (left) or cyclin D1 in wild-type U2OS cells (right) treated with 0.5 μ M palbociclib, 25 μ M chloroquine or both for 24 h. LC3 and HSP90 blots for U937 cells are the same as in Extended Data Fig. 4c, as the experiments were performed simultaneously. Untreated *AMBRA1*-knockout cells serve as a control for increased cyclin D expression. Asterisk, unspecific band. $n = 3$ (U937) or $n = 1$ (U2OS) biological replicates. **b, c**, Immunoassay quantification of cyclin D2 (**b**) and cyclin D3 (**c**) in U2OS cells treated with 1 μ M bortezomib for 4 h ($n = 4$ biological replicates). **d**, Quantification of ubiquitylated cyclin D1 relative to total cyclin D1 isolated from U2OS clones pretreated with 1 μ M bortezomib for 4 h using TUBEs. Each symbol is an isogenic clone ($n = 3$ (sgCtrl) or $n = 5$ (sg*AMBRA1*)). **e, f**, Immunoassay of ubiquitylated cyclin D1 isolated using TUBEs following *AMBRA1* knockdown in U2OS cells (**e**) or in populations of control and *AMBRA1*-knockout MCF7 cells (**f**). **g, h**, Quantification of ubiquitylated cyclin D1 relative to total cyclin D1 in *AMBRA1*-knockdown U2OS cells (**g**) ($n = 2$ biological replicates) or *AMBRA1*-knockout MCF7 cells (**h**) ($n = 2$ (sgCtrl) or $n = 3$ (sg*AMBRA1*) biological replicates) as shown in **e, f**, respectively. For all TUBE experiments, only quantification of samples with similar levels of ubiquitin pull down are shown. See Supplementary Table 9 for all data. **i**, Immunoblot analysis of *AMBRA1* in 293T cells expressing control or *AMBRA1*-targeting shRNAs, pretreated with 10 μ M MG132 for 4 h. ($n = 1$ experiment). **j**, Principal component analysis of mass spectrometry data from cells in **i** (two replicates each of shNT no. 1 and sh*AMBRA1* no. 1 and no. 2) after enriching for ubiquitylated peptides. **k**, Volcano plot of mass-spectrometry data comparing ubiquitylated peptides in control and *AMBRA1* knockdown 293T cells. Each dot is a peptide. Red symbols, significantly upregulated peptides; blue symbols, significantly downregulated peptides, with the indicated cut-offs. All other data are mean \pm s.d. P values calculated by two-sided paired *t*-test (**b, c**), two-sided unpaired *t*-test (**d**) and two-sided unpaired *t*-test followed by Benjamini–Hochberg correction (**k**). HSP90 and GAPDH are loading controls. [Source data](#)

Extended Data Fig. 7 *AMBRA1* binding to CUL4 is required for regulating cyclin D.

a, Co-immunoprecipitation of transfected AMBRA1–Myc–Flag and cyclin D–HA (D1, D2 or D3) in 293T cells, analysed by immunoassay. **b**, Co-immunoprecipitation of transfected Myc-tagged cullin proteins with endogenous AMBRA1 in U2OS cells, analysed by immunoassay. **c**, RT–qPCR analysis of *CCND1* mRNA expression in U2OS cells following knockdown of *AMBRA1* or various cullin genes by siRNA pools ($n = 3$ biological replicates). **d**, Co-immunoprecipitation of transfected wild-type (WT) AMBRA1 and AMBRA1(Δ H) with endogenous CUL4A and CUL4B in 293T cells. **e**, Immunoassay of AMBRA1 in control and *AMBRA1*-knockout U2OS cells with doxycycline-inducible expression of wild-type AMBRA1, AMBRA1(Δ H) or GFP control, after treatment with 500 ng ml $^{-1}$ doxycycline (+Dox) or DMSO (−Dox) for 2 d. **f**, Immunoassay of cyclin D1 ubiquitylation in 293T cells with overexpression of wild-type AMBRA1 or AMBRA1(Δ H). Cells were pretreated with 1 μ M bortezomib for 3 h and lysed in denaturing conditions before immunoprecipitation of cyclin D1. Representative of two independent experiments. **g**, Immunoassay of cyclin D1–HA in U2OS cells expressing wild-type cyclin D1 or phosphomutant cyclin D1 (cyclin D1(T286A)) treated with 10 μ g ml $^{-1}$ cycloheximide for up to 2 h. Cells were transfected with control or *AMBRA1*-targeted siRNA pools 3 d previously. **h**, Quantification of cyclin D1–HA protein levels in U2OS cells from **g** with best-fit curves for one-phase decay ($n = 3$ biological replicates). **i**, Co-immunoprecipitation of cyclin D1–HA (wild-type or T286A) and endogenous AMBRA1 in U2OS cells. CDK4 serves as a positive control for cyclin D1 binding. Representative of two independent experiments. All data are mean \pm s.d. P values calculated by two-sided paired *t*-test (**c**) and two-way ANOVA (**h**). HSP90 and actin are loading controls. [Source data](#)

Extended Data Fig. 8 AMBRA1 ubiquitylates cyclin D.

a, Coomassie-blue-stained gel with protein extracts from insect Sf9 cells (−) or Sf9 cells expressing cyclin D1 and CDK4 (arrows). **b**, Immunoblot for cyclin D1, cyclin D1 phosphorylated on T286 (P-T286) and CDK4 in protein extracts, similar to **a**. **c**, Immunoassay of Flag and Myc tag in untransfected 293T cells (−) or 293T cells transfected with AMBRA1–3×Flag or Myc3–CUL4B. Actin is a loading control. $n = 1$ experiment. [Source data](#)

Extended Data Fig. 9 AMBRA1 is a tumour suppressor in KRAS-mutant mouse lung adenocarcinoma.

a, Lollipop plot for *RBI* and *AMBRA1* mutations in 10,953 patients (10,967 samples) in 32 studies from TCGA (data downloaded from <https://cBioPortal.org> in September 2020). **b**, Immunoassay of AMBRA1 and cyclin D1 in *AMBRA1*-knockout U2OS cells upon stable expression of GFP, wild-type AMBRA1 (WT) or two mutant forms of AMBRA1 from **a** (stop codons at the position indicated by an asterisk). HSP90 is a loading control. Expression of 217* was not detected, suggesting an unstable protein. **c**, Quantification of cyclin D1 in **b** ($n = 3$ biological replicates). Data are mean \pm s.d. P values calculated by two-sided paired *t*-test. **d**, **e**, Relative tumour sizes for each sgRNA in KT mice (lacking Cas9) (**d**) ($n = 4$ mice) and KPTC mice (**e**) ($n = 5$ mice). Tumour sizes were calculated from merged data from all tumours in all mice and normalized to inert sgRNAs 15 (**d**) or 14 (**e**) weeks after cancer initiation. **f**, **g**, Tumour number for each sgRNA in KTC mice (**f**) ($n = 9$ mice) and KPTC mice (**g**) ($n = 5$ mice). Data from all tumours in all mice were merged and normalized to the average tumour number across inert sgRNAs. For **d–g**, error bars denote 95% confidence intervals determined by bootstrap sampling. **h**, Representative H&E staining of tumours from KC mice infected with lentiviral vectors encoding Cre recombinase and either a control or *Ambra1*-targeted sgRNA. Scale bar, 100 μ m. $n = 6$ (*Neo* no. 1) or $n = 5$ (*Ambra1* no. 1) mice. **i**, Representative immunofluorescence for cyclin D in control and *Ambra1*-knockout KC tumours. The cyclin D antibody used recognizes cyclin D1 and D2. Scale bars, 100 μ m. From $n = 2$ mice per sgRNA). [Source data](#)

Extended Data Fig. 10 AMBRA1 is a tumour suppressor in KRAS-mutant human lung adenocarcinoma.

a, Immunoassay of AMBRA1, RB and cyclin D1 in control and knockout human A549 lung adenocarcinoma cells. Actin is a loading control. **b**, Growth of control and mutant A549 xenografts in NOD-SCID-gamma (NSG) mice ($n = 8$ tumours per sgRNA). $****P_{\text{interaction}} < 0.0001$ by two-way ANOVA comparing the *AMBRA1*-knockout curve with control.

Tumour volume measurements for *RBL*-knockout tumours were staggered 1 d behind control and *AMBRA1*-knockout tumours, which precludes two-way ANOVA. Data are mean \pm s.e.m., with best-fit curves for exponential growth. **c**, Final tumour weights from **b**. Each symbol is one tumour ($n = 8$ per sgRNA). Data are mean \pm s.d. **d, g, j**, Cyclin D1 protein levels as measured by reverse phase protein array in relation to the mRNA expression as measured by RNA sequencing (upper quartile of fragments per kilobase of transcript per million mapped reads (FPKM-UQ)) of RB pathway genes that best predict cyclin D1 protein in TCGA *KRAS* G12-mutant lung adenocarcinoma (**d**) ($n = 90$ samples), *KRAS* wild-type lung adenocarcinoma (**g**) ($n = 257$ samples) and *EGFR*-mutant lung adenocarcinoma (**j**) ($n = 41$ samples), using a step-wise regression model. For **g, j**, *AMBRA1* was not selected in the final model but is shown for comparison. Each column is an individual sample, and samples are sorted by cyclin D1 protein levels. **e, h**, Kaplan–Meier plot of *AMBRA1* expression (high, upper third; low, bottom third) in TCGA *KRAS* wild-type lung adenocarcinoma (**e**) ($n = 361$ patients) and *EGFR*-mutant lung adenocarcinoma (**h**) ($n = 60$ patients). **f, i**, Forest plot of Cox proportional hazard model of TCGA *KRAS* wild-type lung adenocarcinoma (**f**) ($n = 340$ patients) and *EGFR*-mutant lung adenocarcinoma (**i**) ($n = 60$ patients). Model is adjusted by stage, age and gender. *P* values calculated by two-way ANOVA (**b**), two-sided unpaired *t*-test (**c**), *F*-test (**d, g, j**), log-rank test (**e, h**) and Wald test (**f, i**). [Source data](#)

Extended Data Fig. 11 AMBRA1 regulates cyclin D protein stability and signalling through the RB pathway.

AMBRA1 limits CDK4/6 activity by mediating ubiquitylation and degradation of D-type cyclins as part of the CRL4 E3 ligase complex. Loss of AMBRA1 leads to accumulation of cyclin D protein and decreased sensitivity to CDK4/6 inhibitors, owing to sustained RB phosphorylation and therefore persistent cell cycle progression.

Supplementary information

Supplementary Information

This file contains Methods, References, Supplementary Figure 1: gating strategies for flow cytometry experiments, and Supplementary Methods.

Reporting Summary

Peer Review File

Supplementary Tables 1-3

Analysis of a genome-wide CRISPR/Cas9 screen to identify genetic determinants of palbociclib response in U937 cells. Table 1: casTLE analysis of the screen results. Table 2: Functional enrichment analysis of significant hits by Metascape. Table 3: MCODE clusters identified by Metascape following protein-protein interaction network analysis of significant hits.

Supplementary Table 4

The top 100 genes most highly correlated with *AMBRA1* based on dependency score in the CRISPR (Avana) Public 19Q4 dataset from the Cancer Dependency Map (downloaded January 23, 2020).

Supplementary Table 5

Differential gene expression analysis of *AMBRA1* knock-out U2OS cells compared to controls, determined by RNA-seq.

Supplementary Table 6-8

Shotgun mass spectrometry analysis of wild-type and *AMBRA1* knock-out (KO) U2OS cells, with or without 24 hours of palbociclib treatment. Table 6: Peptide counts. Table 7: Fold-change between wild-type and KO cells. Table 8: Fold-change between untreated and palbociclib-treated conditions.

Supplementary Table 9

Immunoassay quantification of cyclin D1 ubiquitylation upon loss of AMBRA1 in multiple cellular contexts.

Supplementary Table 10

Mass spectrometry analysis of enriched ubiquitylated peptides from wild-type and AMBRA1 knock-down 293T cells.

Source data

[**Source Data Fig. 1**](#)

[**Source Data Fig. 2**](#)

[**Source Data Fig. 3**](#)

[**Source Data Fig. 4**](#)

[**Source Data Extended Data Fig. 1**](#)

[**Source Data Extended Data Fig. 2**](#)

[**Source Data Extended Data Fig. 4**](#)

[**Source Data Extended Data Fig. 5**](#)

[**Source Data Extended Data Fig. 6**](#)

[**Source Data Extended Data Fig. 7**](#)

[**Source Data Extended Data Fig. 8**](#)

[**Source Data Extended Data Fig. 9**](#)

Source Data Extended Data Fig. 10

Rights and permissions

[Reprints and Permissions](#)

About this article



Check for
updates

Cite this article

Chaikovsky, A.C., Li, C., Jeng, E.E. *et al.* The AMBRA1 E3 ligase adaptor regulates the stability of cyclin D. *Nature* **592**, 794–798 (2021).
<https://doi.org/10.1038/s41586-021-03474-7>

[Download citation](#)

- Received: 06 March 2020
- Accepted: 18 March 2021
- Published: 14 April 2021
- Issue Date: 29 April 2021
- DOI: <https://doi.org/10.1038/s41586-021-03474-7>

Further reading

- [The path to destruction for D-type cyclin proteins](#)
 - Charupong Saengboonmee

- & Piotr Sicinski

Nature (2021)

- **AMBRA1 regulates cyclin D to guard S-phase entry and genomic integrity**

- Emiliano Maiani
- , Giacomo Milletti
- , Francesca Nazio
- , SÃ¢s GrÃ¢nbÃ¶k Holdgaard
- , Jirina Bartkova
- , Salvatore Rizza
- , Valentina Cianfanelli
- , Mar Lorente
- , Daniele Simoneschi
- , Miriam Di Marco
- , Pasquale DâAcunzo
- , Luca Di Leo
- , Rikke Rasmussen
- , Costanza Montagna
- , Marilena Raciti
- , Cristiano De Stefanis
- , Estibaliz Gabicagogeascoa
- , Gergely Rona
- , NÃ©lida Salvador
- , Emanuela Pupo
- , Joanna Maria Merchut-Maya
- , Colin J. Daniel
- , Marianna Carinci
- , Valeriana Cesarini
- , Alfie Oâsullivan
- , Yeon-Tae Jeong
- , Matteo Bordi
- , Francesco Russo
- , Silvia Campello
- , Angela Gallo

- , Giuseppe Filomeni
- , Letizia Lanzetti
- , Rosalie C. Sears
- , Petra Hamerlik
- , Armando Bartolazzi
- , Robert E. Hynds
- , David R. Pearce
- , Charles Swanton
- , Michele Pagano
- , Guillermo Velasco
- , Elena Papaleo
- , Daniela De Zio
- , Apolinar Maya-Mendoza
- , Franco Locatelli
- , Jiri Bartek
- & Francesco Cecconi

Nature (2021)

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

Associated Content

[The path to destruction for D-type cyclin proteins](#)

- Charupong Saengboonmee
- Piotr Sicinski

Nature News & Views 14 Apr 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03474-7>

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

- Article
- [Published: 14 April 2021](#)

AMBRA1 regulates cyclin D to guard S-phase entry and genomic integrity

- [Emiliano Maiani](#) [ORCID: orcid.org/0000-0003-1432-5394](#)^{1,2} na1,
- [Giacomo Milletti](#) [ORCID: orcid.org/0000-0003-2391-342X](#)^{3,4} na1,
- [Francesca Nazio](#)³,
- [Søs Grønbæk Holdgaard](#)¹,
- [Jirina Bartkova](#)^{5,6},
- [Salvatore Rizza](#) [ORCID: orcid.org/0000-0001-7335-4684](#)⁷,
- [Valentina Cianfanelli](#) [ORCID: orcid.org/0000-0002-3857-9463](#)^{1,3},
- [Mar Lorente](#)^{8,9},
- [Daniele Simoneschi](#) [ORCID: orcid.org/0000-0003-0448-4295](#)^{10,11,12},
- [Miriam Di Marco](#)²,
- [Pasquale D'Acunzo](#) [ORCID: orcid.org/0000-0001-7237-0076](#)^{13,14},
- [Luca Di Leo](#) [ORCID: orcid.org/0000-0002-2419-2196](#)¹⁵,
- [Rikke Rasmussen](#)¹⁶,
- [Costanza Montagna](#)^{7,17,18},
- [Marilena Raciti](#)¹,
- [Cristiano De Stefanis](#)¹⁹,
- [Estibaliz Gabicagogeascoa](#)^{8,9},
- [Gergely Rona](#)^{10,11,12},
- [Nélida Salvador](#)^{8,9},
- [Emanuela Pupo](#)²⁰,
- [Joanna Maria Merchut-Maya](#)^{5,21},
- [Colin J. Daniel](#)²²,
- [Marianna Carinci](#)^{3,23},

- [Valeriana Cesarini](#)^{3,24},
- [Alfie O'sullivan](#)^{10,11,12},
- [Yeon-Tae Jeong](#)^{10,11,12},
- [Matteo Bordi](#) ORCID: orcid.org/0000-0001-8207-8546^{3,4},
- [Francesco Russo](#)²⁵,
- [Silvia Campello](#) ORCID: orcid.org/0000-0003-0536-2484⁴,
- [Angela Gallo](#)³,
- [Giuseppe Filomeni](#) ORCID: orcid.org/0000-0002-2719-1412⁷,
- [Letizia Lanzetti](#) ORCID: orcid.org/0000-0001-7541-5524^{20,26},
- [Rosalie C. Sears](#) ORCID: orcid.org/0000-0003-1558-2413^{22,27},
- [Petra Hamerlik](#)^{16,28},
- [Armando Bartolazzi](#) ORCID: orcid.org/0000-0001-7820-5127²⁹,
- [Robert E. Hynds](#) ORCID: orcid.org/0000-0002-2170-8791^{30,31},
- [David R. Pearce](#) ORCID: orcid.org/0000-0001-8443-2924³⁰,
- [Charles Swanton](#) ORCID: orcid.org/0000-0002-4299-3018^{30,31},
- [Michele Pagano](#)^{10,11,12},
- [Guillermo Velasco](#) ORCID: orcid.org/0000-0002-1994-2386^{8,9},
- [Elena Papaleo](#)^{2,32},
- [Daniela De Zio](#) ORCID: orcid.org/0000-0002-9454-402X¹⁵,
- [Apolinar Maya-Mendoza](#) ORCID: orcid.org/0000-0001-7452-9896^{5,21},
- [Franco Locatelli](#)^{3,33},
- [Jiri Bartek](#) ORCID: orcid.org/0000-0003-2013-7525^{5,6} &
- [Francesco Cecconi](#) ORCID: orcid.org/0000-0002-5614-4359^{1,3,4}

[Nature](#) volume **592**, pages 799–803(2021) [Cite this article](#)

- 8253 Accesses
- 1 Citations
- 74 Altmetric
- [Metrics details](#)

Subjects

- [Checkpoints](#)
- [Tumour-suppressor proteins](#)

Abstract

Mammalian development, adult tissue homeostasis and the avoidance of severe diseases including cancer require a properly orchestrated cell cycle, as well as error-free genome maintenance. The key cell-fate decision to replicate the genome is controlled by two major signalling pathways that act in parallel—the MYC pathway and the cyclin D–cyclin-dependent kinase (CDK)–retinoblastoma protein (RB) pathway^{1,2}. Both MYC and the cyclin D–CDK–RB axis are commonly deregulated in cancer, and this is associated with increased genomic instability. The autophagic tumour-suppressor protein AMBRA1 has been linked to the control of cell proliferation, but the underlying molecular mechanisms remain poorly understood. Here we show that AMBRA1 is an upstream master regulator of the transition from G1 to S phase and thereby prevents replication stress. Using a combination of cell and molecular approaches and in vivo models, we reveal that AMBRA1 regulates the abundance of D-type cyclins by mediating their degradation. Furthermore, by controlling the transition from G1 to S phase, AMBRA1 helps to maintain genomic integrity during DNA replication, which counteracts developmental abnormalities and tumour growth. Finally, we identify the CHK1 kinase as a potential therapeutic target in AMBRA1-deficient tumours. These results advance our understanding of the control of replication-phase entry and genomic integrity, and identify the AMBRA1–cyclin D pathway as a crucial cell-cycle-regulatory mechanism that is deeply interconnected with genomic stability in embryonic development and tumorigenesis.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

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: AMBRA1 regulates cell proliferation by affecting the stability of D-type cyclins through interaction with DDB1 and CLR4.

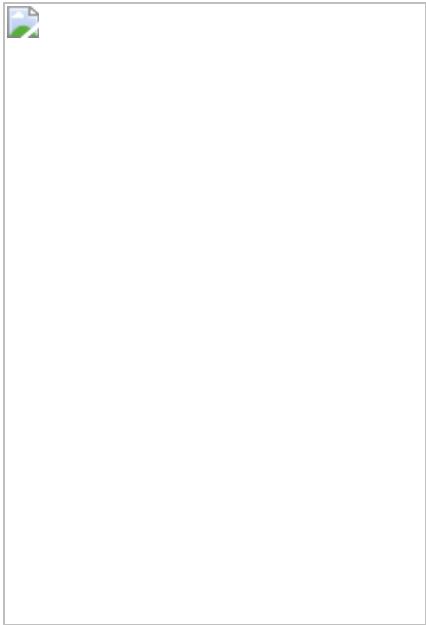


Fig. 2: Depletion of AMBRA1 causes replication stress.

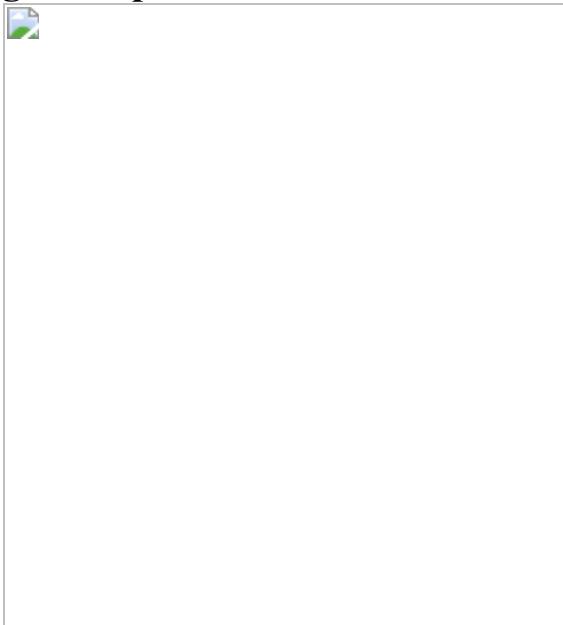
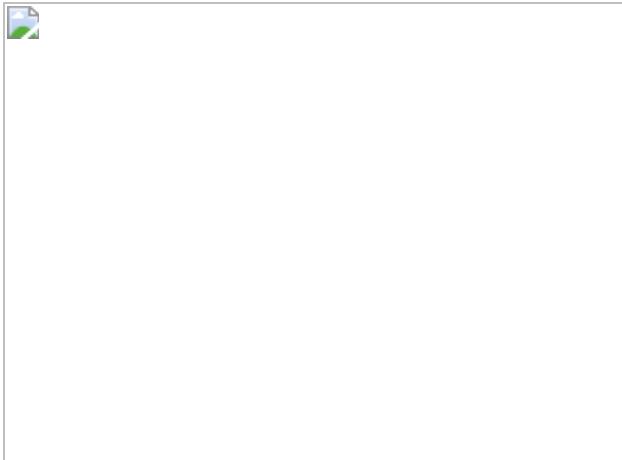


Fig. 3: AMBRA1 is a tumour suppressor and its loss is synthetic lethal with CHK1 inhibition.



Data availability

Data from the Kaplan–Meier analysis in Extended Data Fig. 6g,h referenced during the study are available in a public repository from the websites (<http://kmplot.com/> and <http://gepia2.cancer-pku.cn/#analysis>). AMBRA1 expression data and the stemness score (RNA-based) were downloaded from the Xena platform (<http://xena.ucsc.edu/>). The graph and map of *AMBRA1* mutations in TCGA Pan-Cancer Atlas studies were downloaded from cBioPortal (<https://www.cbioportal.org/>). The original uncropped immunoblot data that support the findings of this study are available in Supplementary Fig. 1. A representative gating strategy for fluorescence-activated cell sorting (FACS) analysis is included in Supplementary Fig. 2. [Source data](#) are provided with this paper.

Code availability

All of the computer scripts and source codes used to generate and analyse the results from The Cancer Genome Atlas (TCGA) analyses presented in Extended Data Figs. 6a, 8a are available at https://github.com/ELELAB/AMBRA_low.

References

1. 1.

Sherr, C. J. & Sicinski, P. in *D-type Cyclins and Cancer* (eds Hinds, P. W. & Brown, N. E.) 1–26 (2018).

2. 2.

Otto, T. & Sicinski, P. Cell cycle proteins as promising targets in cancer therapy. *Nat. Rev. Cancer* **17**, 93–115 (2017).

[CAS](#) [Article](#) [Google Scholar](#)

3. 3.

Cianfanelli, V. et al. Ambra1 at a glance. *J. Cell Sci.* **128**, 2003–2008 (2015).

[CAS](#) [Article](#) [Google Scholar](#)

4. 4.

Antonioli, M. et al. AMBRA1 interplay with cullin E3 ubiquitin ligases regulates autophagy dynamics. *Dev. Cell* **31**, 734–746 (2014).

[CAS](#) [Article](#) [Google Scholar](#)

5. 5.

Nazio, F. et al. mTOR inhibits autophagy by controlling ULK1 ubiquitylation, self-association and function through AMBRA1 and TRAF6. *Nat. Cell Biol.* **15**, 406–416 (2013).

[CAS](#) [Article](#) [Google Scholar](#)

6. 6.

Cianfanelli, V. et al. AMBRA1 links autophagy to cell proliferation and tumorigenesis by promoting c-Myc dephosphorylation and degradation. *Nat. Cell Biol.* **17**, 20–30 (2015).

[CAS](#) [Article](#) [Google Scholar](#)

7. 7.

Fimia, G. M. et al. Ambra1 regulates autophagy and development of the nervous system. *Nature* **447**, 1121–1125 (2007).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

8. 8.

Baldin, V., Lukas, J., Marcote, M. J., Pagano, M. & Draetta, G. Cyclin D1 is a nuclear protein required for cell cycle progression in G1. *Genes Dev.* **7**, 812–821 (1993).

[CAS](#) [Article](#) [Google Scholar](#)

9. 9.

Mateyak, M. K., Obaya, A. J. & Sedivy, J. M. c-Myc regulates cyclin D-Cdk4 and -Cdk6 activity but affects cell cycle progression at multiple independent points. *Mol. Cell. Biol.* **19**, 4672–4683 (1999).

[CAS](#) [Article](#) [Google Scholar](#)

10. 10.

Lange, C., Huttner, W. B. & Calegari, F. Cdk4/cyclinD1 overexpression in neural stem cells shortens G1, delays neurogenesis, and promotes the generation and expansion of basal progenitors. *Cell Stem Cell* **5**, 320–331 (2009).

[CAS](#) [Article](#) [Google Scholar](#)

11. 11.

Di Rita, A. et al. HUWE1 E3 ligase promotes PINK1/PARKIN-independent mitophagy by regulating AMBRA1 activation via IKK α . *Nat. Commun.* **9**, 3755 (2018).

[ADS](#) [Article](#) [Google Scholar](#)

12. 12.

Diehl, J. A., Zindy, F. & Sherr, C. J. Inhibition of cyclin D1 phosphorylation on threonine-286 prevents its rapid degradation via the ubiquitin-proteasome pathway. *Genes Dev.* **11**, 957–972 (1997).

[CAS](#) [Article](#) [Google Scholar](#)

13. 13.

Di Bartolomeo, S. et al. The dynamic interaction of AMBRA1 with the dynein motor complex regulates mammalian autophagy. *J. Cell Biol.* **191**, 155–168 (2010).

[Article](#) [Google Scholar](#)

14. 14.

Liu, E. Y. et al. Loss of autophagy causes a synthetic lethal deficiency in DNA repair. *Proc. Natl Acad. Sci. USA* **112**, 773–778 (2015).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

15. 15.

Burrell, R. A. et al. Replication stress links structural and numerical cancer chromosomal instability. *Nature* **494**, 492–496 (2013).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

16. 16.

Fragkos, M. & Naim, V. Rescue from replication stress during mitosis. *Cell Cycle* **16**, 613–633 (2017).

[CAS](#) [Article](#) [Google Scholar](#)

17. 17.

Wilhelm, T. et al. Spontaneous slow replication fork progression elicits mitosis alterations in homologous recombination-deficient mammalian cells. *Proc. Natl Acad. Sci. USA* **111**, 763–768 (2014).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

18. 18.

Lukas, C. et al. 53BP1 nuclear bodies form around DNA lesions generated by mitotic transmission of chromosomes under replication stress. *Nat. Cell Biol.* **13**, 243–253 (2011).

[CAS](#) [Article](#) [Google Scholar](#)

19. 19.

Maya-Mendoza, A. et al. High speed of fork progression induces DNA replication stress and genomic instability. *Nature* **559**, 279–284 (2018).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

20. 20.

Petermann, E., Woodcock, M. & Helleday, T. Chk1 promotes replication fork progression by controlling replication initiation. *Proc. Natl Acad. Sci. USA* **107**, 16090–16095 (2010).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

21. 21.

Feijoo, C. et al. Activation of mammalian Chk1 during DNA replication arrest: a role for Chk1 in the intra-S phase checkpoint monitoring replication origin firing. *J. Cell Biol.* **154**, 913–923 (2001).

[CAS](#) [Article](#) [Google Scholar](#)

22. 22.

Bartkova, J. et al. Oncogene-induced senescence is part of the tumorigenesis barrier imposed by DNA damage checkpoints. *Nature* **444**, 633–637 (2006).

[ADS](#) [Article](#) [Google Scholar](#)

23. 23.

Lukas, J., Bartkova, J., Rohde, M., Strauss, M. & Bartek, J. Cyclin D1 is dispensable for G1 control in retinoblastoma gene-deficient cells independently of cdk4 activity. *Mol. Cell. Biol.* **15**, 2600–2611 (1995).

[CAS](#) [Article](#) [Google Scholar](#)

24. 24.

Bartek, J. & Lukas, J. Pathways governing G1/S transition and their response to DNA damage. *FEBS Lett.* **490**, 117–122 (2001).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

25. 25.

Chaikovsky, A. C. et al. The AMBRA1 E3 ligase adaptor regulates the stability of cyclin D. *Nature* <https://doi.org/10.1038/s41586-021-03474-7> (2021).

26. 26.

Simoneschi, D. et al. CRL4^{AMBRA1} is a master regulator of D-type cyclins. *Nature* <https://doi.org/10.1038/s41586-021-03445-y> (2021).

27. 27.

Murga, M. et al. Exploiting oncogene-induced replicative stress for the selective killing of Myc-driven tumors. *Nat. Struct. Mol. Biol.* **18**, 1331–1335 (2011).

[CAS](#) [Article](#) [Google Scholar](#)

28. 28.

Syljuåsen, R. G. et al. Inhibition of human Chk1 causes increased initiation of DNA replication, phosphorylation of ATR targets, and DNA breakage. *Mol. Cell. Biol.* **25**, 3553–3562 (2005).

[Article](#) [Google Scholar](#)

29. 29.

Sen, T. et al. CHK1 inhibition in small-cell lung cancer produces single-agent activity in biomarker-defined disease subsets and combination activity with cisplatin or olaparib. *Cancer Res.* **77**, 3870–3884 (2017).

[CAS](#) [Article](#) [Google Scholar](#)

30. 30.

Bartek, J., Mistrik, M. & Bartkova, J. Thresholds of replication stress signaling in cancer development and treatment. *Nat. Struct. Mol. Biol.* **19**, 5–7 (2012).

[CAS](#) [Article](#) [Google Scholar](#)

31. 31.

Toledo, L. I., Murga, M. & Fernandez-Capetillo, O. Targeting ATR and Chk1 kinases for cancer treatment: a new model for new (and old) drugs. *Mol. Oncol.* **5**, 368–373 (2011).

[CAS](#) [Article](#) [Google Scholar](#)

32. 32.

Dere, E. et al. Heterozygous Ambra1 deficiency in mice: a genetic trait with autism-like behavior restricted to the female gender. *Front. Behav. Neurosci.* **8**, 181 (2014).

[Google Scholar](#)

33. 33.

La Barbera, L., Vedele, F., Nobili, A., D'Amelio, M. & Krashia, P. Neurodevelopmental disorders: functional role of ambra1 in autism and schizophrenia. *Mol. Neurobiol.* **56**, 6716–6724 (2019).

[Article](#) [Google Scholar](#)

34. 34.

Kuleshov, M. V. et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res.* **44**, W90–W97 (2016).

[CAS](#) [Article](#) [Google Scholar](#)

35. 35.

Chen, E. Y. et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**, 128 (2013).

[Article](#) [Google Scholar](#)

36. 36.

Tang, Z. et al. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res.* **45**, W98–W102 (2017).

[CAS](#) [Article](#) [Google Scholar](#)

37. 37.

Nagy, A., Gyöngyi, M. & Győrffy, B. Pancancer survival analysis of cancer hallmark genes. *Sci. Rep.* **11**, 6047 (2021).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

[Download references](#)

Acknowledgements

E. Maiani is an Adjunct Professor at UniCamillus—Saint Camillus International University of Health Sciences. The F.C. laboratory is supported by grants from the Danish Cancer Society (KBVU R72-A4408, R146-A9364, R231-A14034 to F.C.; R146-A9471 to V. Cianfanelli; R146-A9414 to G.F.; R204-A12424 to D.D.Z.), the Novo Nordisk Foundation (NNF13OC0007559, NNF16OC0022544), the Lundbeck Foundation (R233-2016-3360 to F.C.; R209-2015-3505 to V. Cianfanelli), the LEO Foundation (LF17024 to F.C. and E. Papaleo; LF-OC-19-000004 to D.D.Z.), the Associazione Italiana per la Ricerca sul Cancro (AIRC project IG 2019 #23543 to F.C.; #22811 to L.L.; 5x1000 #9962 and AIRC IG 2018 #21724 to F.L.), the Italian Ministry of Research (MIUR, project PRIN 2017 FS5SHL Radius) and the Italian Ministry of Health (Ricerca Corrente to F.L. and F.N.). This work was also supported by the European Union’s Horizon 2020 research and innovation program (Marie Skłodowska-Curie grant agreement 642295 (MEL-PLEX)). D.D.Z. is supported by the Melanoma Research Alliance (MRA 620385). The F.C., J. Bartek and E. Papaleo laboratories in Copenhagen are part of the Center of Excellence for Autophagy, Recycling and Disease (CARD), funded by the Danmarks Grundforskningsfond (DNRF125). L.L. is supported by FPRC 5x1000 Ministero della Salute 2015. V. Cianfanelli, C.M. and M.B. are supported by the Fondazione Umberto Veronesi. M.P. is funded by grants from the National Institute of Health (R01-CA76584 and R35-GM136250) and is an investigator with the Howard Hughes Medical Institute; the work of E. Papaleo is supported by the Carlsberg Foundation Distinguished Fellowship (CF18-0314); work in the G.V. group is supported by the PI15/00339 grant, integrated into the State Plan for R&D + I2013-2016 and funded by the Instituto de Salud Carlos III (ISCIII) and the European Regional Development Fund (ERDF), by the Marie Skłodowska-Curie Innovative Training Network (ITN) action TRAIN (GA 721532) funded by the European Commission (H2020) and by grants from Voices Against Brain Cancer and ‘Fundació La Marató de TV3’ (20134031). J. Bartek, J. Bartkova and A.M.-M. are supported by grants from the Danish Cancer Society (R204-A12617-B153), the Novo Nordisk Foundation (16854 and

0060590), the Danish Council for Independent Research (DFF-7016-00313), the Lundbeck Foundation (R266-2017-4289), the Swedish Research Council (VR-MH 2014-46602-117891-30) and the Swedish Cancerfonden (170176). We thank P. Bonaldo and P. Braghetta for the generation of the *Ambra1*^{flox/flox} mouse model, V. Turcanova for help with cloning and mutagenesis, Plaisant S.r.l. (Castel Romano) and the Danish Cancer Society animal facilities for help with in vivo experiments, V. Tocco and the FACS facility members for technical help in the flow cytometry analysis and C. Rodolfo for his help and support. G.M. is grateful to A. M. Gatta and V. Milletti for their support.

Author information

Author notes

1. These authors contributed equally: Emiliano Maiani, Giacomo Milletti

Affiliations

1. Cell Stress and Survival Unit, Center for Autophagy, Recycling and Disease (CARD), Danish Cancer Society Research Center, Copenhagen, Denmark

Emiliano Maiani, Søs Grønbæk Holdgaard, Valentina Cianfanelli, Marilena Raciti & Francesco Cecconi

2. Computational Biology Laboratory, Center for Autophagy, Recycling and Disease (CARD), Danish Cancer Society Research Center, Copenhagen, Denmark

Emiliano Maiani, Miriam Di Marco & Elena Papaleo

3. Department of Pediatric Onco-Hematology and Cell and Gene Therapy, IRCCS Bambino Gesù Children's Hospital, Rome, Italy

Giacomo Milletti, Francesca Nazio, Valentina Cianfanelli, Marianna Carinci, Valeriana Cesarini, Matteo Bordi, Angela Gallo, Franco

Locatelli & Francesco Cecconi

4. Department of Biology, University of Rome ‘Tor Vergata’, Rome, Italy

Giacomo Milletti, Matteo Bordi, Silvia Campello & Francesco Cecconi

5. Genome Integrity Unit, Danish Cancer Society Research Center, Copenhagen, Denmark

Jirina Bartkova, Joanna Maria Merchut-Maya, Apolinar Maya-Mendoza & Jiri Bartek

6. Division of Genome Biology, Department of Medical Biochemistry and Biophysics, Science for Life Laboratory, Karolinska Institute, Stockholm, Sweden

Jirina Bartkova & Jiri Bartek

7. Redox Biology Group, Danish Cancer Society Research Center, Copenhagen, Denmark

Salvatore Rizza, Costanza Montagna & Giuseppe Filomeni

8. Department of Biochemistry and Molecular Biology, School of Biology, Complutense University, Madrid, Spain

Mar Lorente, Estibaliz Gabicagogeasca, Nélida Salvador & Guillermo Velasco

9. Instituto de Investigación Sanitaria San Carlos (IdISSC), Madrid, Spain

Mar Lorente, Estibaliz Gabicagogeasca, Nélida Salvador & Guillermo Velasco

10. Department of Biochemistry and Molecular Pharmacology, NYU Grossman School of Medicine, New York, NY, USA

Daniele Simoneschi, Gergely Rona, Alfie O'sullivan, Yeon-Tae Jeong & Michele Pagano

11. Laura and Isaac Perlmutter Cancer Center, NYU Grossman School of Medicine, New York, NY, USA

Daniele Simoneschi, Gergely Rona, Alfie O'sullivan, Yeon-Tae Jeong & Michele Pagano

12. Howard Hughes Medical Institute, NYU Grossman School of Medicine, New York, NY, USA

Daniele Simoneschi, Gergely Rona, Alfie O'sullivan, Yeon-Tae Jeong & Michele Pagano

13. Center for Dementia Research, Nathan S. Kline Institute for Psychiatric Research, Orangeburg, NY, USA

Pasquale D'Acunzo

14. Department of Psychiatry, New York University School of Medicine, New York, NY, USA

Pasquale D'Acunzo

15. Melanoma Research Team, Cell Stress and Survival Unit, Danish Cancer Society Research Center, Copenhagen, Denmark

Luca Di Leo & Daniela De Zio

16. Brain Tumor Biology Group, Danish Cancer Society Research Center, Copenhagen, Denmark

Rikke Rasmussen & Petra Hamerlik

17. UniCamillus—Saint Camillus International University of Health Sciences, Rome, Italy

Costanza Montagna

18. Institute of Sports Medicine Copenhagen, Department of Orthopedic Surgery, Copenhagen University Hospital – Bispebjerg and Frederiksberg, Copenhagen, Denmark

Costanza Montagna

19. Research Laboratories, IRCCS Bambino Gesù Children's Hospital, Rome, Italy

Cristiano De Stefanis

20. Candiolo Cancer Institute, FPO – IRCCS, Turin, Italy

Emanuela Pupo & Letizia Lanzetti

21. DNA Replication and Cancer Group, Genome Integrity Unit, Danish Cancer Society Research Center, Copenhagen, Denmark

Joanna Maria Merchut-Maya & Apolinar Maya-Mendoza

22. Department of Molecular and Medical Genetics, Oregon Health & Science University, Portland, OR, USA

Colin J. Daniel & Rosalie C. Sears

23. Department of Medical Sciences, Laboratory for Technologies of Advanced Therapies (LTTA), University of Ferrara, Ferrara, Italy

Marianna Carinci

24. Department of Biomedical Sciences, Institute of Translational Pharmacology, National Research Council of Italy (CNR), Rome, Italy

Valeriana Cesarini

25. Section for Clinical Mass Spectrometry, Danish Center for Neonatal Screening, Department of Congenital Disorders, Statens Serum Institut, Copenhagen, Denmark

Francesco Russo

26. Department of Oncology, University of Torino Medical School, Turin, Italy

Letizia Lanzetti

27. Knight Cancer Institute, Oregon Health & Science University, Portland, OR, USA

Rosalie C. Sears

28. Department of Drug Design and Pharmacology, Copenhagen University, Copenhagen, Denmark

Petra Hamerlik

29. Department of Pathology and Pathology Research Laboratory, Sant'Andrea Hospital, Rome, Italy

Armando Bartolazzi

30. Cancer Research UK Lung Cancer Centre of Excellence, UCL Cancer Institute, University College London, London, UK

Robert E. Hynds, David R. Pearce & Charles Swanton

31. Cancer Evolution and Genome Instability Laboratory, The Francis Crick Institute, London, UK

Robert E. Hynds & Charles Swanton

32. Novo Nordisk Foundation Center for Protein Research, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

Elena Papaleo

33. Department of Gynecology–Obstetrics and Pediatrics, Sapienza University, Rome, Italy

Franco Locatelli

Authors

1. Emiliano Maiani

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Giacomo Milletti

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Francesca Nazio

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Søs Grønbæk Holdgaard

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Jirina Bartkova

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Salvatore Rizza

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Valentina Cianfanelli

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Mar Lorente

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. Daniele Simoneschi

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. Miriam Di Marco

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

11. Pasquale D'Acunzo

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

12. Luca Di Leo

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

13. Rikke Rasmussen

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

14. Costanza Montagna

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

15. Marilena Raciti

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

16. Cristiano De Stefanis

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

17. Estibaliz Gabicagogeascoa

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

18. Gergely Rona

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

19. Nélida Salvador

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

20. Emanuela Pupo

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

21. Joanna Maria Merchut-Maya

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

22. Colin J. Daniel

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

23. Marianna Carinci

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

24. Valeriana Cesarini

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

25. Alfie O'sullivan

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

26. Yeon-Tae Jeong

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

27. Matteo Bordi

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

28. Francesco Russo

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

29. Silvia Campello

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

30. Angela Gallo

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

31. Giuseppe Filomeni

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

32. Letizia Lanzetti

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

33. Rosalie C. Sears

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

34. Petra Hamerlik

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

35. Armando Bartolazzi

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

36. Robert E. Hynds

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

37. David R. Pearce

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

38. Charles Swanton

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

39. Michele Pagano

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

40. Guillermo Velasco

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

41. Elena Papaleo

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

42. Daniela De Zio

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

43. Apolinar Maya-Mendoza

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

44. Franco Locatelli

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

45. Jiri Bartek

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

46. Francesco Cecconi

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Contributions

E.M., G.M., J. Bartek and F.C. conceived the study and designed experiments. G.M., P.D., C.D.S., M.C. and V. Cesarini performed all analyses regarding the development of *Ambra1* cKO mice. E.M., G.M., C.M. and S.G.H. carried out the biochemical and microscopy experiments linking AMBRA1 to genomic instability and synthetic lethality. S.G.H., J. Bartkova, V. Cianfanelli, D.D.Z. and A.B. carried out the experiments and analysis of the *Ambra1*-deficient KRAS lung model. F.N. performed the experiments involving N-MYC. S.G.H., E. Pupo and L.L. performed analyses of mitotic cells. E.M., S.R. and L.d.L., evaluated U2OS-FUCCI dynamics upon AMBRA1 deficiency. E.M., M.D.M., F.R. and E. Papaleo performed all related bioinformatics. R.R. performed traffic light experiments. E.M. and M.R. performed the experiments and analysis of xenograft SKUT-1B experiments. M.L., E.G., N.S. and G.V. carried out the xenograft experiments with transformed MEFs. C.J.D. and R.C.S. generated and validated the MYC(pS62) antibody used in the immunohistochemistry experiments. A.M.-M. and J.M.M.-M. performed experiments and analysis regarding fork speed and symmetry. D.S., G.R., Y.-T.J. and M.P. provided key information about AMBRA1 substrates, as well as some cDNAs. A.O. helped with some biochemical experiments. R.E.H. and D.R.P. gave experimental support for lung cancer cell models. M.B., S.C., A.G., G.F., L.L., P.H., A.B., C.S., M.P., E. Papaleo, D.D.Z., A.M.-M. and F.L. provided critical support, key data analyses and conceptual advice. E.M., G.M., J. Bartek and F.C. wrote the original draft. All authors took part in writing, reviewing and editing the final manuscript. All authors read and accepted the manuscript.

Corresponding authors

Correspondence to [Jiri Bartek](#) or [Francesco Cecconi](#).

Ethics declarations

Competing interests

M.P. is a consultant for and has financial interests in Coho Therapeutics, CullGen, Kymera Therapeutics and SEED Therapeutics. M.P. is a

cofounder of Coho Therapeutics, is on the Scientific Advisory Board of CullGen and Kymera Therapeutics, and is a consultant for Santi Therapeutics. The other authors declare no competing interests.

Additional information

Peer review information *Nature* thanks Piotr Sicinski 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 AMBRA1 regulates cyclin D stability in NSCs.](#)

a, Schematic for production of the conditional knockout mouse model. **b, c**, Images of wild-type and *Ambra1* cKO P21 mice (**b**) and brains (**c**). **c**, bottom, representative image of PCR amplification of *Tm1c*, *Ambra1* and *Cre*. **d**, Wild-type and *Ambra1* cKO olfactory bulbs in sagittal sections of E18.5 embryos, stained for Ki67 antibody and Hoechst ($n = 3$). **e**, Quantification of Ki67⁺ cell area in the whole brain of wild-type and *Ambra1* cKO E13.5 embryos (sagittal sections shown in Fig. [1b](#)) ($n = 5$). *P* value by two-tailed unpaired *t*-test. **f**, Representative scheme of NSCs extraction and cell culturing from mouse embryo medial ganglionic eminences (MGE). LGE, lateral ganglionic eminences. **g**, Densitometry quantification of normalized protein levels in wild-type and *Ambra1* cKO NSCs shown in Fig. [1c](#) ($n = 4$). *P* values by two-sided one-way ANOVA followed by Sidak's multiple comparisons test. **h**, Left, representative images of NSCs extracted from mouse embryo medial ganglionic eminences. Right, violin plot of clonal neurosphere diameters in wild-type and *Ambra1* cKO NSCs ($n = 3$; total of 128 neurospheres analysed for each condition). *P* value by two-tailed unpaired *t*-test. **i**, Whole-brain quantification of E13.5 wild-type and *Ambra1* cKO cyclin D1 staining

normalized over DAPI, represented in Fig. 1e. *P* value by two-tailed unpaired *t*-test (*n* = 5). j, Wild-type and *Ambra1* cKO olfactory bulbs in sagittal sections of E18.5 embryos, stained for cyclin D1 antibody and Hoechst (*n* = 3). k, Sagittal sections of wild-type and *Ambra1* cKO E13.5 embryos, stained for cyclin D2 (*n* = 3). l, Left, representative images of sagittal sections of the mesencephalic ventricular zone in wild-type and *Ambra1* cKO E13.5 embryos, stained for RB(pS807/811) (*n* = 5). Right, quantification of RB(pS807/811)-positive area in the mesencephalic ventricular zone of E13.5 wild-type and *Ambra1* cKO embryos (*n* = 5). *P* value by two-tailed unpaired *t*-test. m, Left, representative images of RB(pS807/811) in sagittal sections of the olfactory bulb in wild-type and *Ambra1* cKO E18.5 embryos. Right, quantification of the number of RB(pS807/811)-positive cells (*n* = 3). *P* value by two-tailed unpaired *t*-test. n, Immunoblot of N-MYC after cycloheximide treatment in wild-type and *Ambra1* cKO NSCs (*n* = 3). o, Quantitative PCR with reverse transcription (qRT–PCR) of NSCs; the investigated genes are at the bottom of the graph (*n* = 5). *P* values by two-tailed unpaired *t*-test. p, Immunoblot of control and *AMBRA1*-silenced SH-SY5Y cells (*n* = 3). q, Immunoblot of *AMBRA1* immunoprecipitation in SH-SY5Y cells. r, Immunoblot for *AMBRA1*, PP2AC and N-MYC in SH-SY5Y cells silenced for the indicated genes (*n* = 3). Unless otherwise stated, *n* refers to biologically independent samples. For immunoblots, actin was used as loading control. Data are mean ± s.e.m. Scale bars, 250 μ m. [Source data](#)

Extended Data Fig. 2 *Ambra1* deficiency affects the cell cycle, cell death and neuronal differentiation.

a, Densitometric quantification of cyclin D1 and D2 protein levels in the cycloheximide time course normalized over actin (*n* = 3). *P* values by two-sided one-way ANOVA followed by Sidak's multiple comparisons test. b, Immunoblot of wild-type or *Ambra1* cKO NSCs treated with cycloheximide and/or MG132 for the indicated times. (*n* = 3). c, Distribution of cell-cycle phases in NSCs after release from nocodazole treatment (*n* = 3). *P* values by two-sided one-way ANOVA followed by Sidak's multiple comparisons test. d, Six-hour BrdU incorporation of passage-2 wild-type and *Ambra1* cKO NSCs with or without abemaciclib treatment (*n* = 3). *P* values by two-sided one-way ANOVA followed by Tukey's multiple comparisons test. e,

Percentage of apoptotic cells in wild-type and *Ambra1* cKO NSCs ($n = 3$). EA, early apoptotic; LA, late apoptotic. P values by two-sided one-way ANOVA followed by Sidak's multiple comparisons test. **f**, Left, immunoblot of the indicated proteins in NSCs after abemaciclib treatment. Right, densitometry quantification of the indicated proteins ($n = 3$). P values by two-sided one-way ANOVA followed by Sidak's multiple comparisons test. **g**, Left, representative images of sagittal sections from wild-type and *Ambra1* cKO E13.5 embryos, stained for SOX2 and TBR2. Right, quantification of immunostained positive area (SOX, $n = 6$; TBR2, $n = 4$). P values by two-tailed unpaired t -test. **h**, Left, representative images of sagittal sections of wild-type and *Ambra1* cKO E18.5 embryos, stained for TBR2. Right, quantification of immunostained positive area ($n = 6$). P value by two-tailed unpaired t -test. Arrows indicate TBR2⁺ cells in the subventricular zone. **i, j**, Representative images of sagittal sections of wild-type and *Ambra1* cKO E18.5 embryos, stained for the neuronal marker NeuN. **i**, Left, higher magnification of the mesencephalic alar plate. Right, quantification of immunostained positive cells ($n = 3$). P value by two-tailed unpaired t -test. **j**, Lower magnification to better appreciate the uncropped quantified area ($n = 3$). Scale bar, 500 μ m. Unless otherwise stated, n refers to biologically independent samples. For immunoblots, actin was used as loading control. Data are mean \pm s.e.m. Unless otherwise noted, scale bars represent 250 μ m. [Source data](#)

Extended Data Fig. 3 The AMBRA1–cyclin D1 axis affects the cell cycle.

a, Immunoblot of control or *AMBRA1*-silenced U87-MG cells for the indicated proteins ($n = 3$). **b**, Immunoblot of cyclin D1 in control or *AMBRA1*-silenced U87-MG cells treated with cycloheximide and/or MG132 for the indicated times ($n = 3$). **c**, Analysis of densitometry for the cyclin D immunoblot in U87-MG cells, silenced for the indicated genes, shown in Fig. 1g ($n = 4$). P values by one-way ANOVA followed by Dunnett's multiple comparisons test. **d**, Left, immunoblot of cyclin D1 in U87-MG cells silenced for *AMBRA1* expression and overexpressing empty vector (pcDNA), wild-type *AMBRA1* or *AMBRA1*(ΔWD40). Right, analysis from densitometry ($n = 3$). P values by one-way ANOVA followed by Tukey's multiple comparisons test. **e**, Immunoblot analysis of cyclin D1

immunoprecipitation from protein extracts of control and *AMBRA1*-silenced U87-MG cells ($n = 3$). **f**, Co-immunoprecipitation of AMBRA1 in U87-MG cells transiently overexpressing empty vector, cyclin D1–Flag or cyclin D1(T286A)–Flag. Cells were treated with MG132 for 3 h before lysis ($n = 3$). **g**, Fold change in the number of cells in control or *AMBRA1*-silenced U87-MG cells ($n = 11$). P value by two-tailed unpaired *t*-test. **h**, Immunoblot of the indicated proteins of U87-MG, BJ-hTERT and U2OS cells that were untreated or treated with MLN4924 for 4 h ($n = 3$). **i, j**, Cells immunostained with cyclin D1, EdU antibody and counterstained with Hoechst. **i**, Scatter plots reporting single-cell total nuclear intensities of EdU versus Hoechst (cells examined over three independent experiments: siSCR, $n = 3,279$; siAMBRA1, $n = 3,608$ cells). **j**, Box plots (centre line, median; box limits, 25th and 75th percentile) indicating total cyclin D1 nuclear intensities (siSCR, $n = 3,279$; siAMBRA1, $n = 3,608$ cells. median siSCR = 169,654; siAMBRA1 = 429,623). **k, l**, Immunoblot of cell-cycle markers in control and *AMBRA1*-silenced BJ-hTERT cells (**k**) and cell-cycle-sorted U2OS-FUCCI cells (**l**) ($n = 3$). **m**, Immunoblot of the indicated proteins in *AMBRA1*-silenced BJ-hTERT cells synchronized by 24-h serum starvation. Cells were collected after the indicated starvation recovery time points ($n = 3$). **n**, Representative images of live-cell imaging of control and *AMBRA1*-silenced U2OS-FUCCI cells from 0 to 14 h with a 2-h interval between different images. The length of the G1 phase is shown in Fig. [1k](#) ($n = 3$). Scale bar, 5 μm . **o**, Cell proliferation in control or *AMBRA1*-silenced BJ-hTERT cells (24 h and 48 h $n = 6$; 72 h siSCR $n = 6$, siAMBRA1 $n = 5$). **p, q**, Control or *AMBRA1*-silenced U2OS-FUCCI cells. **p**, Representative contour plot. **q**, Fold increase of cells present in S–G2 phase in *AMBRA1*-downregulated cells with respect to control cells ($n = 10$). **r**, Box plots (centre line, median; box limits, 25th and 75th percentile; whiskers, minimum and maximum) showing the cell-cycle length of siSCR ($n = 65$; median = 13) or siAMBRA1 ($n = 65$; median = 8.5) U2OS-FUCCI cells examined over three independent experiments. Unless otherwise stated, n refers to biologically independent samples; data are mean \pm s.e.m. Data were analysed using a two-tailed unpaired *t*-test (**g, o, q**) or two-tailed Mann–Whitney test (**j, r**). For immunoblots, actin or β -tubulin were used as loading control. [Source data](#)

Extended Data Fig. 4 AMBRA1 deficiency causes replication stress.

a, Total γH2AX nuclear intensity in the different cell-cycle phases of BJ-hTERT cells ($n = 3$). Data are mean ± s.d. **b**, Average number of γH2AX foci in control or *AMBRA1*-silenced U2OS cells ($n = 3$). **c**, Alkaline comet assay of control, *AMBRA1*- and *ATG7*-silenced U2OS cells ($n = 3$). **d**, Scatter plots showing γH2AX versus Hoechst total nuclear intensities from immunostainings of control, *AMBRA1*- and *ATG7*-silenced BJ-hTERT cells. The proportion of γH2AX-positive cells (red, arbitrary cut-off) is indicated (si*SCR*, $n = 721$; si*AMBRA1*, $n = 725$; si*ATG7*, $n = 733$ cells examined over 3 independent experiments). **e**, Immunoblot of γH2AX in control, *AMBRA1*- and *ATG7*-silenced BJ-hTERT cells ($n = 3$). **f**, Homologous recombination (HR) efficiency in control, *AMBRA1*- and *ATG7*-silenced U2OS cells ($n = 3$). Data are mean ± s.d. **g**, Number of BRCA1 foci per nucleus in control and *AMBRA1*-silenced U2OS cells either untreated or treated with 3-Gy irradiation, stained against BRCA1 ($n = 500$ cells examined over 3 independent experiments, centre indicates the mean). **h**, Time in mitosis in control ($n = 91$ cells examined over 3 independent experiments) or *AMBRA1*-silenced ($n = 72$ cells examined over 3 independent experiments) cells. Bars represent median and interquartile range. **i**, Dying cells upon mitotic exit as evaluated by time-lapse imaging ($n = 2$ independent experiments; more than 60 cells per condition). **j**, Distribution of 53BP1 nuclear foci in G1 U2OS cells ($n = 3$). **k**, Total γH2AX versus Hoechst intensity in control and *AMBRA1*-silenced BJ-hTERT cells that were untreated or treated with 2 mM hydroxyurea (HU) for 2 h (si*SCR*, $n = 2,481$; si*AMBRA1*, $n = 2,237$; si*SCR* + HU, $n = 2,484$; si*AMBRA1* + HU, $n = 2,281$ cells; scatter plots are representative of $n = 3$ independent experiments). **l**, Quantification of normalized protein levels of CHK1 represented in Fig. 2e ($n = 3$). **m**, **n**, BJ-hTERT cells as in Extended Data Fig. 4k treated with cycloheximide or with cycloheximide and 2 mM hydroxyurea. **m**, Immunoblot analysis of the indicated proteins in total cell lysates. **n**, Quantification of normalized CHK1 protein expression levels ($n = 4$). **o**, **p**, qRT-PCR analyses of the indicated genes in control or *AMBRA1*-silenced BJ-hTERT (**o**) and U2OS (**p**) cells, respectively (*CCNA2*, *E2F1* and *RAD51* $n = 5$; *BRCA1* $n = 4$; *CHEK1* $n = 3$). **q**, **r**, Immunoblot analysis of the indicated proteins in control or *AMBRA1*-

silenced U2OS (**q**) and BJ-hTERT (**r**) cells ($n = 3$ in both conditions). **s**, Gene ontology (GO) biological processes (2018) from enrichment analysis of DEA (Differential Expression Analysis) genes from RNA sequencing (RNA-seq) experiments. DEA originating from three RNA-seq independent experiments was used as input for the web-based software EnrichR^{34,35}. P values computed using Fisher's exact test; clearer bars show a smaller P value. Unless otherwise stated, n refers to biologically independent samples; data are mean \pm s.e.m. Data were analysed using a two-tailed unpaired t -test (**a**, **b**, **c**, **f**, **j**, **l**, **n**, **o**, **p**) or two-tailed Mann–Whitney test (**g**, **h**). For immunoblots, β -tubulin, SOD1 or GADPH were used as loading control. [Source data](#)

Extended Data Fig. 5 AMBRA1 deficiency causes replication stress.

a, Analysis of DEA genes (from $n = 3$ independent RNA-seq experiments) predicting the transcription factor activated after depletion of AMBRA1. **b**, qRT–PCR analyses of the indicated genes in control or *AMBRA1*-silenced U2OS-FUCCI cells sorted for the different cell-cycle phases ($n = 3$). **c**, Immunoblot for the indicated proteins in U2OS cells interfered for the indicated autophagy regulators ($n = 3$). **d**, Left, violin plot of γ H2AX nuclear mean intensity in control and *AMBRA1*-silenced BJ-hTERT cells that were untreated or treated with 0.1 μ M abemaciclib for 48 h. Right, representative scatter plot of single-cell γ H2AX nuclear mean intensity versus Hoechst, and cell cycle phase gating strategies from control and *AMBRA1*-silenced BJ-hTERT cells treated with abemaciclib ($n = 643$ cells). **e**, Cell count of control U87-MG cells or U87-MG cells with inducible cyclin D1 expression, three days after stimulation with dox ($n = 3$). **f**, Cell count of control BJ-hTERT cells or BJ-hTERT cells with inducible cyclin D1 expression at the indicated time points after stimulation with dox, normalized over non-induced cells (1-d V15+, $n = 6$; 3-d V15+, 3-d E30+, $n = 4$; 1-d E30+, 4-d V15+, 4-d E30+, 6-d V15+ and 6-d E30+, $n = 5$). V15+: dox-treated control cells; E30+: dox-treated cyclin D1-inducible cells. **g**, **h**, Percentage of cells in each cell-cycle phase in U87-MG (**g**) and BJ-hTERT (**h**) cells, control or with inducible cyclin D1 expression, untreated or 48 h after doxycycline stimulation ($n = 3$). **i**, Immunoblot for the indicated proteins in control U87-MG cells or U87-MG

cells with inducible cyclin D1 expression at the indicated time points with or without dox stimulation ($n = 3$). **j, k**, Mean fork speed (**j**) (kb min $^{-1}$) and fork symmetry analysis (**k**) of DNA fibres from control BJ-hTERT cells and BJ-hTERT cells with inducible cyclin D1 expression treated as in Fig. [2d](#) (scored forks: –dox, $n = 312$; 3-d dox, $n = 449$; 4-d dox, $n = 429$; 6-d dox, $n = 426$). Data are mean \pm s.d. **l**, Quantification of immunohistochemistry staining in Fig. [2g](#) ($n = 3$ mice). **m**, Immunoblot for the indicated proteins in wild-type or *Ambra1* cKO NSCs ($n = 3$). Unless otherwise stated, n refers to biologically independent samples; data are mean \pm s.e.m. Data were analysed using a two-tailed unpaired *t*-test (**b, l**), two-tailed Mann–Whitney test (**d, j, k**), two-way ANOVA followed by Sidak’s multiple comparisons test (**e, g, h**) or one-way ANOVA followed by Sidak’s multiple comparisons test (**f**). Exact *P* values are provided in the ‘Statistical analysis and data reproducibility’ section of the [Supplementary Methods](#). [Source data](#)

[Extended Data Fig. 6 Bioinformatics analysis of *AMBRA1* in cancer.](#)

a, Bioinformatics analysis of expression data from the TCGA database. Pie charts show the percentage of *AMBRA1*-low cancers (light blue) with respect to the total (grey) in the indicated datasets. BLCA, bladder urothelial carcinoma; COAD, colon adenocarcinoma; KIRC, kidney renal clear cell carcinoma; KIRP, kidney renal papillary cell carcinoma; LUSC, lung squamous cell carcinoma; PRAD, prostate adenocarcinoma; UCEC, uterine corpus endometrial carcinoma. **b**, Xena correlation analysis of *AMBRA1* mRNA expression and stemness score. The shaded area in the plot indicates the confidence interval (95%). **c**, Lollipop plots showing the distribution of *AMBRA1* mutations annotated in TCGA Pan-Cancer Atlas Studies datasets. **d**, Frequency of *AMBRA1* mutations (expressed as a percentage) in TCGA Pan-Cancer Atlas Studies datasets. The cut-off was selected at 2%. **e**, Oncoprint of *AMBRA1* alterations (homodeletions, shallow deletions, mutations), and *TP53* and *EGFR* mutations from TCGA Pan-Lung Cancer datasets. **f**, Mutual exclusivity and co-occurrence analysis of the indicated genes from TCGA Pan-Lung Cancer datasets. *P* values derived from one-sided Fisher’s exact test. **g**, Kaplan–Meier analysis of patients in the Pan-Cancer Atlas Studies database was generated based on the expression level of *AMBRA1* (low, below 20%; high, above 80%). Plot

was downloaded from the online database GEPIA³⁶ (<http://gepia2.cancer-pku.cn/#analysis>). *P* values derived from one-sided log-rank Mantel–Cox test). **h**, Kaplan–Meier analysis of overall survival based on RNA-seq analysis of *AMBRA1* mRNA levels using the KM-plotter³⁷ lung adenocarcinoma database.

Extended Data Fig. 7 AMBRA1 controls tumour growth in a mouse model of lung cancer.

a, Schematic representations of the mouse model and initial testing of the system. The *Kras*^{G12D} transgenic mouse is mated with the conditional *Ambra1*^{flox/flox} mouse to produce the *Ambra1*^{+/+}::*Kras*^{G12D/+} and the *Ambra1*^{flox/flox}::*Kras*^{G12D/+} genotypes. Lung-specific expression of oncogenic *Kras*^{G12D} and deletion of *Ambra1* is induced by intranasal inoculation with defective adenoviral particles carrying the Cre recombinase. **b**, Immunoblot analysis of AMBRA1 immunoprecipitation from tissue lung samples from *Ambra1*^{flox/flox}::*Kras*^{G12D/+} mice 16 weeks after administration of AdenoCre ($n = 3$). **c**, The expression of the *Ambra1* floxed allele after Cre administration was verified by RT–PCR performed in lung tissue samples as in **c** ($n = 3$). Primers were designed to distinguish wild-type and floxed alleles. **d**, Representative examples of H&E images of fixed lungs. Bottom, Magnification of the bronchus, highlighting the tumour initiation site. Scale bar, 1 mm. **e**, Quantification of immunohistochemistry staining in Fig. **3b** (Ki67, $n = 3$; γH2AX, $n = 3$; RPA(pS4/8), $n = 3$; cyclin D1, $n = 4$; c-Myc(pS62), $n = 3$ in two independent tumours for each condition). Unless otherwise stated, n refers to biologically independent samples; data are mean ± s.e.m. *P* values for γH2AX and cyclin D1 by two-tailed Welch *t*-test; *P* values for Ki67, c-MYC(pS62), RPA(pS4/8) by two-tailed unpaired *t*-test. [Source data](#)

Extended Data Fig. 8 AMBRA1 deficiency is synthetic lethal with CHK1 inhibitors.

a, Ratio between *CHEK1* expression in the *AMBRA1*-low subpopulation of cancers with respect to normal tissue. **b**, Gating strategy for Fig. **3c**.

Bottom, scatter plots of total nuclear DNA intensity versus γ H2AX intensity (si*SCR* DMSO, $n = 1,850$; si*SCR* AZD, $n = 1,716$; si*AMBRA1* DMSO, $n = 1,866$; si*AMBRA1* AZD, $n = 1,731$ cells; representative of three independent experiments). The γ H2AX-positive cells (arbitrary cut-off) are indicated in red. Top, Hoechst nuclear intensity versus counts. γ H2AX-positive cells are indicated by the red line. **c**, Immunoblot of AMBRA1, RPA(pS4/8) and β -tubulin in control or *AMBRA1*-silenced BJ-hTERT cells that were untreated or treated with AZD7762. **d**, Left, gating strategy for the quantification on the right. Top, Hoechst nuclear intensity versus counts. Bottom, scatter plots reporting single-cell total nuclear DNA intensity versus TUNEL intensity. Right, TUNEL-positive cells in the different cell phases calculated based on Hoechst intensity ($n = 3$). *P* values by two-tailed unpaired *t*-test. **e**, Viability analysis of control and *AMBRA1*-silenced BJ-hTERT cells treated with the indicated concentrations of LY2603618 for 24 h ($n = 3$). *P* values by two-stage step-up (Benjamini, Krieger and Yekutieli). **f**, Cell viability in control, *AMBRA1*- and *ATG7*-silenced BJ-hTERT cells that were untreated or treated with AZD7762 for 24 h ($n = 4$ for Control and treatments with 100 nM AZD7762). *P* value by two-tailed unpaired *t*-test. **g**, Cell viability in control and *AMBRA1*-silenced BJ-hTERT cells that were untreated or treated with olaparib for 24 h ($n = 3$). Analysis by two-tailed unpaired *t*-test. **h**, Fork symmetry analysis from BJ-hTERT cells treated for 24 h with 100 nM AZD7762 or 5 μ M LY2603618 (scored forks: si*SCR* DMSO, $n = 533$; si*SCR* AZD, $n = 560$; si*SCR* LY, $n = 548$; si*AMBRA1* DMSO, $n = 601$; si*AMBRA1* AZD, $n = 543$; si*AMBRA1* LY, $n = 548$). *P* values by two-tailed Mann–Whitney test. Data are mean \pm s.d. **i**, Cell viability analysis of control and *AMBRA1*-silenced A549, H11299 and HCC827 lung cancer cell lines treated with the indicated concentrations of AZD7762 and LY2603618 ($n = 3$) for 24 h. *P* values by two-stage step-up (Benjamini, Krieger and Yekutieli). Data are mean \pm s.d. **j**, Immunoblot of the indicated proteins in control or *AMBRA1*-silenced A549, HCC827 and H1299 cells ($n = 3$). **k**, Immunoblot of sarcoma cell lines ($n = 3$). **l**, Immunoblot of SKUT-1B cells treated with the inhibitor MLN4924 for 4 h ($n = 3$). **m**, Left, immunoblot of SKUT-1B cells reconstituted with wild-type AMBRA1 or mutant AMBRA1(Δ WD40) or AMBRA1(PXP). Right, densitometry quantification of the indicated normalized protein levels ($n = 3$). *P* values by two-sided one-way ANOVA followed by Sidak's multiple comparisons test. **n**, Late apoptosis analysis in SKUT-1B cells

reconstituted with wild-type AMBRA1, AMBRA1(Δ WD40) or AMBRA1(PXP) and treated with 200 nM AZD7762 for 24 h ($n = 3$). P values by two-sided one-way ANOVA followed by Tukey's multiple comparisons test. Unless otherwise stated, n refers to biologically independent samples; data are mean \pm s.e.m. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$. Exact P values are provided in the 'Statistical analysis and data reproducibility' section of the [Supplementary Methods](#). [Source data](#)

Extended Data Fig. 9 AMBRA1 deficiency is synthetic lethal with CHK1 inhibitors in vivo.

a, Cell viability of *Ambra1*^{+/+} and *Ambra1*^{gt/gt} MEFs treated with AZD7762 or vehicle for 24 h ($n = 4$ independent experiments). P values by two-tailed unpaired *t*-test. **b**, Box plots (centre line, median; box limits, 25th and 75th percentile; whiskers, minimum and maximum) indicating weight of *Ambra1*^{+/+} and *Ambra1*^{gt/gt} MEF xenografts referred to in Fig. 3f (*Ambra1*^{+/+} + vehicle, $n = 8$; *Ambra1*^{+/+} + AZD7762, $n = 8$; *Ambra1*^{gt/gt} + vehicle, $n = 10$; *Ambra1*^{gt/gt} + AZD7762, $n = 11$ mice). P values by two-tailed unpaired *t*-test. **c**, Cell death percentage in control U87-MG cells or overexpressing cyclin D1, either untreated or treated with AZD7762 for 24 h; mean \pm s.e.m. ($n = 3$ independent experiments). P values by two-sided one-way ANOVA followed by Tukey's multiple comparisons test. Unless otherwise stated, data are mean \pm s.d. [Source data](#)

Supplementary information

Supplementary Information

This file contains Supplementary Methods and References.

Reporting Summary

Supplementary Figure 1

Original uncropped Immunoblot Data.

Supplementary Figure 2

Representative gating strategy of experiments shown in Fig. 1d, Extended Data Fig. 2c, 2d, 3l, 3f, 5b, 5g, 5h, 8n.

Supplementary Table 1

Raw count of all samples analysed by RNASeq.

Supplementary Table 2

RNASeq differential analysis of all *siAMBRA1* tRNA vs all *siSCR* samples.

Supplementary Table 3

Bioinformatical analysis of *AMBRA1* and *CHEK1* expression levels from TCGA database.

Video 1

Time-lapse imaging of mitotic H2B-GFP U2OS control cells.

Video 2

Time-lapse imaging of mitotic H2B-GFP U2OS cells depleted for *AMBRA1*.

Source data

Source Data Fig. 1

Source Data Fig. 2

[Source Data Fig. 3](#)

[Source Data Extended Data Fig. 1](#)

[Source Data Extended Data Fig. 2](#)

[Source Data Extended Data Fig. 3](#)

[Source Data Extended Data Fig. 4](#)

[Source Data Extended Data Fig. 5](#)

[Source Data Extended Data Fig. 7](#)

[Source Data Extended Data Fig. 8](#)

[Source Data Extended Data Fig. 9](#)

Rights and permissions

[Reprints and Permissions](#)

About this article



Check for
updates

Cite this article

Maiani, E., Milletti, G., Nazio, F. *et al.* AMBRA1 regulates cyclin D to guard S-phase entry and genomic integrity. *Nature* **592**, 799–803 (2021).
<https://doi.org/10.1038/s41586-021-03422-5>

Download citation

- Received: 09 March 2020
- Accepted: 04 March 2021
- Published: 14 April 2021
- Issue Date: 29 April 2021
- DOI: <https://doi.org/10.1038/s41586-021-03422-5>

Further reading

- [The path to destruction for D-type cyclin proteins](#)
 - Charupong Saengboonmee
 - & Piotr Sicinski

Nature (2021)

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

Associated Content

[The path to destruction for D-type cyclin proteins](#)

- Charupong Saengboonmee
- Piotr Sicinski

Nature News & Views 14 Apr 2021

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03422-5>

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

- Matters Arising
- [Published: 28 April 2021](#)

Quantifying forest change in the European Union

- [Iddo K. Wernick](#) [ORCID: orcid.org/0000-0002-7406-2365¹](#),
- [Philippe Ciais](#) [ORCID: orcid.org/0000-0001-8560-4943²](#),
- [Jonas Fridman](#) [ORCID: orcid.org/0000-0002-8295-665X³](#),
- [Peter Höglberg](#) [ORCID: orcid.org/0000-0002-2849-2719⁴](#),
- [Kari T. Korhonen⁵](#),
- [Annika Nordin⁴](#) &
- [Pekka E. Kauppi](#) [ORCID: orcid.org/0000-0003-1735-756X^{4,6}](#)

[Nature](#) volume **592**, pagesE13–E14(2021)[Cite this article](#)

- 773 Accesses
- 1 Citations
- 8 Altmetric
- [Metrics details](#)

Subjects

- [Climate sciences](#)
- [Environmental sciences](#)

[Matters Arising](#) to this article was published on 28 April 2021

The [Original Article](#) was published on 01 July 2020

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

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](#)

References

1. 1.

Ceccherini, G. et al. Abrupt increase in harvested forest area over Europe after 2015. *Nature* **583**, 72–77 (2020).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

2. 2.

E-yearbook of Food and Natural Resource Statistics for 2018 (eds Kortesmaa, T. & Salo-Kauppinen, R) https://stat.luke.fi/en/e-yearbook-food-and-naturalresource-statistics-2018-2018_en (Natural Resources Institute Finland, 2018).

3. 3.

Swedish Forest Agency. *The Swedish National Forest Inventory* (accessed 15 August 2020); <https://www.slu.se/en/Collaborative-Centres-and-Projects/the-swedish-national-forest-inventory/>

4. 4.

Hansen, M. C. et al. High-resolution global maps of 21st-century forest cover change. *Science* **342**, 850–853 (2013).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

5. 5.

Kauppi, P. E. et al. Carbon benefits from Forest Transitions promoting biomass expansions and thickening. *Glob. Change Biol.* **26**, 5365–5370 (2020).

[ADS](#) [Article](#) [Google Scholar](#)

6. 6.

European Pellet Council. *Bioenergy Europe Statistical Report 2019* https://epc.bioenergyeurope.org/wp-content/uploads/2020/02/SR19_Pellet_final-web-1.pdf (2019).

[Download references](#)

Author information

Affiliations

1. Program for the Human Environment, The Rockefeller University, New York, NY, USA

Iddo K. Wernick

2. Laboratoire des Sciences du Climat et de l'Environnement, CEA-CNRS-UVSQ-8 UPSACLAY, Gif sur Yvette, France

Philippe Ciais

3. Department of Forest Resource Management, Swedish University of Agricultural Sciences (SLU), Umeå, Sweden

Jonas Fridman

4. Department of Forest Ecology and Management, Swedish University of Agricultural Sciences, Umeå, Sweden

Peter Höglberg, Annika Nordin & Pekka E. Kauppi

5. Natural Resources Institute Finland, Helsinki, Finland

Kari T. Korhonen

6. Department of Forest Sciences, University of Helsinki, Helsinki, Finland

Pekka E. Kauppi

Authors

1. Iddo K. Wernick

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Philippe Ciais

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Jonas Fridman

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Peter Höglberg

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Kari T. Korhonen

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Annika Nordin

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Pekka E. Kauppi

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Contributions

This Comment was initiated by I.K.W. and P.E.K. P.C., J.F., P.H., K.T.K. and A.N. all contributed to the writing and editing of the Comment.

Corresponding author

Correspondence to [Iddo K. Wernick](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

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

Rights and permissions

[Reprints and Permissions](#)

About this article



Check for
updates

Cite this article

Wernick, I.K., Ciais, P., Fridman, J. *et al.* Quantifying forest change in the European Union. *Nature* **592**, E13–E14 (2021).

<https://doi.org/10.1038/s41586-021-03293-w>

[Download citation](#)

- Received: 24 August 2020
- Accepted: 26 January 2021
- Published: 28 April 2021
- Issue Date: 29 April 2021
- DOI: <https://doi.org/10.1038/s41586-021-03293-w>

Further reading

- [Reply to Wernick, I. K. et al.; PalahÃ, M. et al.](#)
 - Guido Ceccherini
 - , Gregory Duveiller
 - , Giacomo Grassi
 - , Guido Lemoine
 - , Valerio Avitabile
 - , Roberto Pilli
 - & Alessandro Cescatti

Nature (2021)

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03293-w>

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

- Matters Arising
- [Published: 28 April 2021](#)

Concerns about reported harvests in European forests

- [Marc Palahí](#) [ORCID: orcid.org/0000-0002-5587-1597¹](#),
[Rubén Valbuena](#) [ORCID: orcid.org/0000-0003-0493-7581²](#),
- [Cornelius Senf³](#),
- [Nezha Acil^{4,5}](#),
- [Thomas A. M. Pugh](#) [ORCID: orcid.org/0000-0002-6242-7371^{4,5,6}](#),
- [Jonathan Sadler^{4,5}](#),
- [Rupert Seidl](#) [ORCID: orcid.org/0000-0002-3338-3402³](#),
- [Peter Potapov⁷](#),
- [Barry Gardiner](#) [ORCID: orcid.org/0000-0002-4106-9026⁸](#),
- [Lauri Hetemäki](#) [ORCID: orcid.org/0000-0001-7051-8874¹](#),
- [Gherardo Chirici⁹](#),
- [Saverio Francini^{9,10}](#),
- [Tomáš Hlásny¹¹](#),
- [Bas Jan Willem Lerink¹²](#),
- [Håkan Olsson](#) [ORCID: orcid.org/0000-0002-3476-1401¹³](#),
- [José Ramón González Olabarria](#) [ORCID: orcid.org/0000-0002-5040-712X¹⁴](#),
- [Davide Ascoli¹⁵](#),
- [Antti Asikainen¹⁶](#),
- [Jürgen Bauhus](#) [ORCID: orcid.org/0000-0002-9673-4986¹⁷](#),
- [Göran Berndes¹⁸](#),
- [Janis Donis¹⁹](#),
- [Jonas Fridman](#) [ORCID: orcid.org/0000-0002-8295-665X¹³](#),
- [Marc Hanewinkel](#) [ORCID: orcid.org/0000-0003-4081-6621¹⁷](#),
- [Hervé Jactel²⁰](#),

- [Marcus Lindner](#) ORCID: [orcid.org/0000-0002-0770-003X²¹](https://orcid.org/0000-0002-0770-003X),
- [Marco Marchetti](#)²²,
- [Róbert Marušák](#)¹¹,
- [Douglas Sheil](#) ORCID: [orcid.org/0000-0002-1166-6591²³](https://orcid.org/0000-0002-1166-6591),
- [Margarida Tomé](#)²⁴,
- [Antoni Trasobares](#)²⁵,
- [Pieter Johannes Verkerk](#) ORCID: [orcid.org/0000-0001-5322-8007¹](https://orcid.org/0000-0001-5322-8007),
- [Minna Korhonen](#)¹ &
- [Gert-Jan Nabuurs](#) ORCID: [orcid.org/0000-0002-9761-074X^{12,23}](https://orcid.org/0000-0002-9761-074X)

[Nature](#) volume 592, pages E15–E17 (2021) [Cite this article](#)

- 1607 Accesses
- 1 Citations
- 231 Altmetric
- [Metrics details](#)

Subjects

- [Environmental health](#)
- [Environmental impact](#)
- [Forestry](#)

[Matters Arising](#) to this article was published on 28 April 2021

The [Original Article](#) was published on 01 July 2020

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

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: Abrupt changes in GFC after 2015 are visible in many temperate regions.

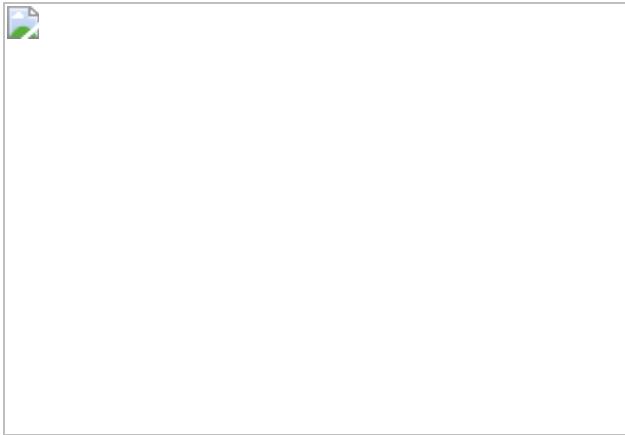
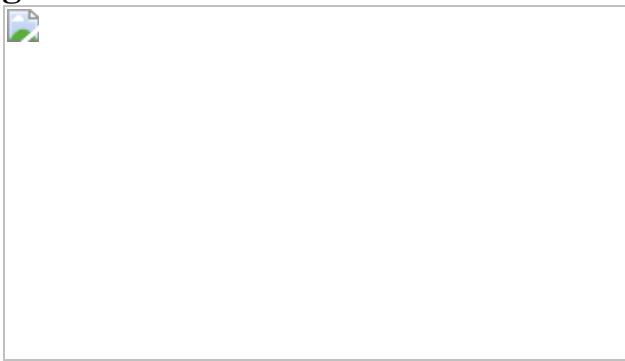


Fig. 2: Areas identified as natural disturbances.



References

1. 1.

Ceccherini, G. et al. Abrupt increase in harvested forest area over Europe after 2015. *Nature* **583**, 72–77 (2020).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

2. 2.

Tree Cover Loss (Hansen/UMD/Google/USGS/NASA) (Global Forest Watch, accessed 1 July 2020);
<http://data.globalforestwatch.org/datasets/14228e6347c44f5691572169e9e107ad>.

3. 3.

Santoro, M. et al. GlobBiomass - global datasets of forest biomass. *PANGAEA* <https://doi.org/10.1594/PANGAEA.894711> (2018).

4. 4.

Hansen, M. C. et al. High-resolution global maps of 21st-century forest cover change. *Science* **342**, 850–853 (2013).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

5. 5.

Chirici, G. et al. Monitoring clearcutting and subsequent rapid recovery in Mediterranean coppice forests with Landsat time series. *Ann. For. Sci.* **77**, 40 (2020).

[Article](#) [Google Scholar](#)

6. 6.

Turubanova, S. et al. Ongoing primary forest loss in Brazil, Democratic Republic of the Congo, and Indonesia. *Environ. Res. Lett.* **13**, 074028 (2018).

[ADS](#) [Article](#) [Google Scholar](#)

7. 7.

Olofsson, P. et al. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* **148**, 42–57 (2014).

[ADS](#) [Article](#) [Google Scholar](#)

8. 8.

Forest Europe. *State of Europe's Forests 2015* (Ministerial Conference on the Protection of Forests in Europe, 2015).

9. 9.

Rozendaal, D. M. A., Santoro, M. & Schepaschenko, D. *DUE GlobBiomass*. D24 Validation Report (FSU Jena, 2017).

10. 10.

Senf, C. et al. Canopy mortality has doubled in Europe's temperate forests over the last three decades. *Nat. Commun.* **9**, 4978 (2018).

[ADS](#) [Article](#) [Google Scholar](#)

11. 11.

Olofsson, P. et al. Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. *Remote Sens. Environ.* **129**, 122–131 (2013).

[ADS](#) [Article](#) [Google Scholar](#)

12. 12.

Bullock, E. L., Woodcock, C. E., Souza, C., Jr & Olofsson, P. Satellite-based estimates reveal widespread forest degradation in the Amazon. *Glob. Change Biol.* **26**, 2956–2969 (2020).

[ADS](#) [Article](#) [Google Scholar](#)

13. 13.

Forzieri, G. et al. A spatially explicit database of wind disturbances in European forests over the period 2000–2018. *Earth Syst. Sci. Data* **12**, 257–276 (2020).

[ADS](#) [Article](#) [Google Scholar](#)

14. 14.

Hurmekoski, E. et al. Diversification of the forest industries: Role of new wood-based products. *Can. J. For. Res.* **48**, 1417–1432 (2018).

[Article](#) [Google Scholar](#)

15. 15.

Levers, C. et al. Drivers of forest harvesting intensity patterns in Europe. *For. Ecol. Manage.* **315**, 160–172 (2014).

[Article](#) [Google Scholar](#)

16. 16.

FAOSTAT. *Forestry Production and Trade* (Food and Agriculture Organization of the United Nations, accessed 15 Dec 2019);
<http://www.fao.org/faostat/en/#data/FQ>.

17. 17.

Seidl, R., Schelhaas, M.-J., Rammer, W. & Verkerk, P. J. Increasing forest disturbances in Europe and their impact on carbon storage. *Nat. Clim. Change* **4**, 806–810 (2014).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

18. 18.

Verkerk, P. J. et al. Climate-Smart Forestry: the missing link. *For. Policy Econ.* **115**, 102164 (2020).

[Article](#) [Google Scholar](#)

[Download references](#)

Acknowledgements

We thank G. Ceccherini and co-authors for immediately making available all original material, processing codes and results of their study upon request. N.A. and T.A.M.P. acknowledge funding from the European

Research Council (ERC) Programme Grant Agreement no. 758873 (TreeMort).

Author information

Author notes

1. These authors contributed equally: Marc Palahí, Ruben Valbuena

Affiliations

1. European Forest Institute, Joensuu, Finland

Marc Palahí, Lauri Hetemäki, Pieter Johannes Verkerk & Minna Korhonen

2. School of Natural Sciences, Bangor University, Bangor, UK

Rubén Valbuena

3. Ecosystem Dynamics and Forest Management Group, Technical University of Munich, Munich, Germany

Cornelius Senf & Rupert Seidl

4. School of Geography, Earth and Environmental Sciences, University of Birmingham, Birmingham, UK

Nezha Acil, Thomas A. M. Pugh & Jonathan Sadler

5. Birmingham Institute of Forest Research, University of Birmingham, Birmingham, UK

Nezha Acil, Thomas A. M. Pugh & Jonathan Sadler

6. Department of Physical Geography and Ecosystem Science, Lund University, Lund, Sweden

Thomas A. M. Pugh

7. Department of Geographical Sciences, University of Maryland,
College Park, MD, USA

Peter Potapov

8. Institut Européen de la Forêt Cultivée, Cestas, France

Barry Gardiner

9. Dipartimento di Scienze e Tecnologie Agrarie, Alimentari, Ambientali
e Forestali, Università degli Studi di Firenze, Florence, Italy

Gherardo Chirici & Saverio Francini

10. Dipartimento per l’Innovazione dei Sistemi Biologici, Agroalimentari
e Forestali, Università degli Studi della Tuscia, Viterbo, Italy

Saverio Francini

11. Faculty of Forestry and Wood Sciences, Czech University of Life
Sciences Prague, Prague, Czech Republic

Tomáš Hlásny & Róbert Marušák

12. Wageningen Environmental Research, Wageningen University and
Research, Wageningen, The Netherlands

Bas Jan Willem Lerink & Gert-Jan Nabuurs

13. Department of Forest Resource Management, Swedish University of
Agricultural Sciences (SLU), Umeå, Sweden

Håkan Olsson & Jonas Fridman

14. Joint Research Unit CTFC - AGROTECNIO, Solsona, Spain

José Ramón González Olabarria

15. Department of Agricultural, Forest and Food Sciences, University of Turin, Grugliasco, Italy

Davide Ascoli

16. Natural Resources Institute Finland, Joensuu, Finland

Antti Asikainen

17. Albert-Ludwigs-University of Freiburg, Freiburg, Germany

Jürgen Bauhus & Marc Hanewinkel

18. Department of Space, Earth and Environment, Chalmers University of Technology, Gothenburg, Sweden

Göran Berndes

19. Latvian State Forest Research Institute Silava, Salaspils, Latvia

Janis Donis

20. INRAE, University of Bordeaux, BIOGECO, Cestas, France

Hervé Jactel

21. European Forest Institute, Bonn, Germany

Marcus Lindner

22. University of Molise, Campobasso, Italy

Marco Marchetti

23. Forest Ecology and Forest Management Group, Wageningen University and Research, Wageningen, The Netherlands

Douglas Sheil & Gert-Jan Nabuurs

24. Centro de Estudos Florestais, Instituto Superior de Agronomia,
Universidade de Lisboa, Lisbon, Portugal

Margarida Tomé

25. Forest Science and Technology Centre of Catalonia, CTFC, Solsona,
Spain

Antoni Trasobares

Authors

1. Marc Palahí

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Rubén Valbuena

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Cornelius Senf

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Nezha Acil

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Thomas A. M. Pugh

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Jonathan Sadler

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Rupert Seidl

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Peter Potapov

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. Barry Gardiner

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. Lauri Hetemäki

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

11. Gherardo Chirici

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

12. Saverio Francini

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

13. Tomáš Hlásny

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

14. Bas Jan Willem Lerink

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

15. Håkan Olsson

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

16. José Ramón González Olabarria

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

17. Davide Ascoli

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

18. Antti Asikainen

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

19. Jürgen Bauhus

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

20. Göran Berndes

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

21. Janis Donis

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

22. Jonas Fridman

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

23. Marc Hanewinkel

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

24. Hervé Jactel

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

25. Marcus Lindner

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

26. Marco Marchetti

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

27. Róbert Marušák

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

28. Douglas Sheil

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

29. Margarida Tomé

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

30. Antoni Trasobares

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

31. Pieter Johannes Verkerk
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

32. Minna Korhonen
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

33. Gert-Jan Nabuurs
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Contributions

M.P. and G.-J.N. conceived and initiated the study. R.V., C.S., N.A., T.A.M.P., G.C., S.F., T.H., B.J.W.L. and D.A. ran different parts of the analyses and demonstrations. M.P., R.V., G.-J.N., C.S., T.A.M.P., J.S., R.S., B.G. and L.H. drafted the initial version of the manuscript. P.P. provided first-hand experience of the algorithms involved in the production of GFC data. All authors offered insights from their own national statistics and local knowledge, which focused the analyses and the argumentation, and contributed critically to the interpretation of the results, revising and approving the final version of the manuscript.

Corresponding authors

Correspondence to [Marc Palahí](#) or [Rubén Valbuena](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

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

Rights and permissions

[Reprints and Permissions](#)

About this article



Check for
updates

Cite this article

Palahí, M., Valbuena, R., Senf, C. *et al.* Concerns about reported harvests in European forests. *Nature* **592**, E15–E17 (2021).
<https://doi.org/10.1038/s41586-021-03292-x>

[Download citation](#)

- Received: 03 July 2020
- Accepted: 26 January 2021
- Published: 28 April 2021
- Issue Date: 29 April 2021
- DOI: <https://doi.org/10.1038/s41586-021-03292-x>

Further reading

- [Reply to Wernick, I. K. et al.; PalahÃ, M. et al.](#)

- Guido Ceccherini
- , Gregory Duveiller
- , Giacomo Grassi
- , Guido Lemoine
- , Valerio Avitabile
- , Roberto Pilli
- & Alessandro Cescatti

Nature (2021)

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03292-x>

- Matters Arising
- [Published: 28 April 2021](#)

Reply to Wernick, I. K. et al.; Palahí, M. et al.

- [Guido Ceccherini](#) ORCID: orcid.org/0000-0003-1772-9554¹,
- [Gregory Duveiller](#)¹,
- [Giacomo Grassi](#)¹,
- [Guido Lemoine](#)²,
- [Valerio Avitabile](#)¹,
- [Roberto Pilli](#)¹ &
- [Alessandro Cescatti](#)¹

[Nature](#) volume **592**, pages E18–E23(2021) [Cite this article](#)

- 834 Accesses
- 10 Altmetric
- [Metrics details](#)

The [Original Article](#) was published on 28 April 2021

The [Original Article](#) was published on 28 April 2021

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

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: Large and small forest-loss patches, and *t*-test on cloud-free observations, for the periods 2013–2015 and 2016–2018.

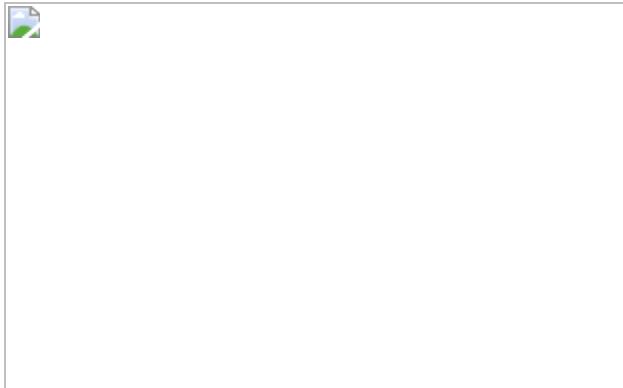


Fig. 2: Comparison between pixel-based and sample-based estimates of forest change in Sweden and Finland.

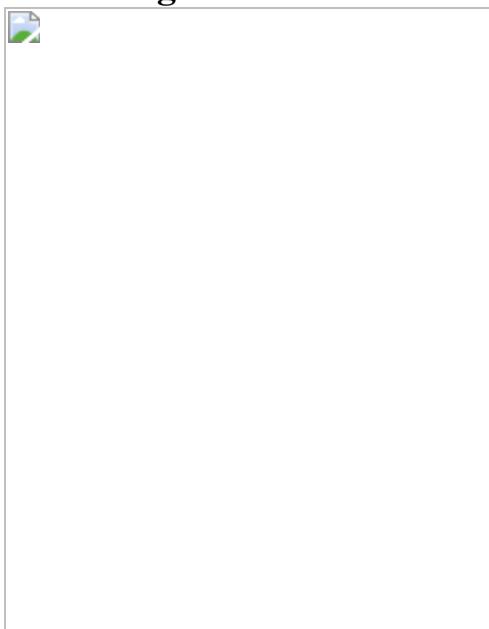
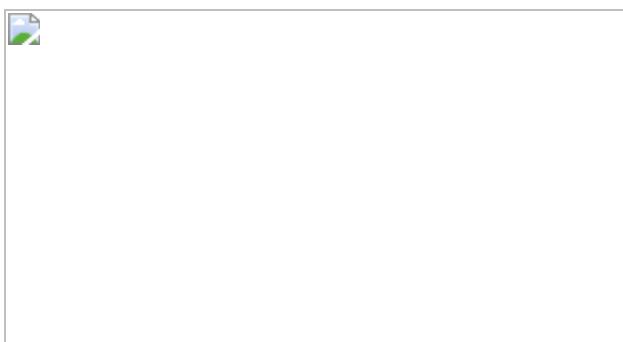


Fig. 3: Forest loss from 'extreme events' stated in our study and in the FORWIND database.



Data availability

To ensure full reproducibility and transparency of our research, all of the data and the scripts used in our analysis have been made available or can be obtained from the corresponding author upon request. Codes used for this study (Google Earth Engine and R scripts, and data synthesis on the validation for GFC stable forests and loss) are available on GitHub at <https://github.com/guidoceccherini/NatureCommentary>.

References

1. 1.

Wernick, I. K. et al. Quantifying forest change in the European Union. *Nature* <https://doi.org/10.1038/s41586-021-03293-w> (2021).

2. 2.

Palahí, M. et al. Reported harvested area and biomass loss in European forests. *Nature* <https://doi.org/10.1038/s41586-021-03292-x> (2021).

3. 3.

Ceccherini, G. et al. Abrupt increase in harvested forest area over Europe after 2015. *Nature* **583**, 72–77 (2020).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

4. 4.

Hansen, M. C. et al. High-resolution global maps of 21st-century forest cover change. *Science* **342**, 850–853 (2013).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

5. 5.

Potapov, P. V. et al. Eastern Europe's forest cover dynamics from 1985 to 2012 quantified from the full Landsat archive. *Remote Sens. Environ.* **159**, 28–43 (2015).

[ADS](#) [Article](#) [Google Scholar](#)

6. 6.

Hansen, M. C. et al. High-resolution global maps of 21st-century forest cover change. *Science* **342**, 850–853 (2013).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

7. 7.

Goldman, L. & Weisse, M. Technical blog: Global Forest Watch's 2018 data update explained. *Global Forest Watch* <https://blog.globalforestwatch.org/data-and-research/technical-blog-global-forest-watchs-2018-data-update-explained/> (2019).

8. 8.

Curtis, P. G., Slay, C. M., Harris, N. L., Tyukavina, A. & Hansen, M. C. Classifying drivers of global forest loss. *Science* **361**, 1108–1111 (2018).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

9. 9.

Harris, N. L. et al. Global maps of twenty-first century forest carbon fluxes. *Nat. Clim. Change* **11**, 234–240 (2021).

[ADS](#) [Article](#) [Google Scholar](#)

10. 10.

Tree Cover Loss by Dominant Driver (Global Forest Watch, accessed August 2020);
<http://data.globalforestwatch.org/datasets/f2b7de1bdde04f7a9034ecb363d71f0e>.

11. 11.

Rossi, F., Breidenbach, J., Puliti, S., Astrup, R. & Talbot, B. Assessing harvested sites in a forested boreal mountain catchment through Global Forest Watch. *Remote Sens.* **11**, 543 (2019).

[ADS](#) [Article](#) [Google Scholar](#)

12. 12.

Alkama, R. & Cescatti, A. Biophysical climate impacts of recent changes in global forest cover. *Science* **351**, 600–604 (2016).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

13. 13.

Hansen, M. C. et al. The fate of tropical forest fragments. *Sci. Adv.* **6**, eaax8574 (2020).

14. 14.

NYDF Assessment Partners. *Goal 1 assessment: Striving to end natural forest loss*. New York Declaration on Forests Progress Assessment. Climate Focus (coordinator and editor) (2020).

15. 15.

Olofsson, P. et al. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* **148**, 42–57 (2014).

[ADS](#) [Article](#) [Google Scholar](#)

16. 16.

Tyukavina, A. et al. Aboveground carbon loss in natural and managed tropical forests from 2000 to 2012. *Environ. Res. Lett.* **10**, 074002 (2015); correction **13**, 109501 (2018).

[ADS](#) [Article](#) [Google Scholar](#)

17. 17.

Tyukavina, A. et al. Types and rates of forest disturbance in Brazilian Legal Amazon, 2000–2013. *Sci. Adv.* **3**, e1601047 (2017).

[ADS](#) [Article](#) [Google Scholar](#)

18. 18.

Olofsson, P. et al. Mitigating the effects of omission errors on area and area change estimates. *Remote Sens. Environ.* **236**, 111492 (2020).

[ADS](#) [Article](#) [Google Scholar](#)

19. 19.

Rozendaal, D. M. A., Santoro, M. & Schepaschenko, D. *DUE GlobBiomass D24 Validation Report* (FSU Jena, 2017).

20. 20.

Forzieri, G. et al. A spatially explicit database of wind disturbances in European forests over the period 2000–2018. *Earth Syst. Sci. Data* **12**, 257–276 (2020).

[Google Scholar](#)

21. 21.

Camia, A. et al. *The use of woody biomass for energy production in the EU and impacts on forests*. JRC Science for Policy Report EUR 30548 EN (Publications Office of the European Union, 2020).

22. 22.

Seidl, R., Schelhaas, M.-J., Rammer, W. & Verkerk, P. J. Increasing forest disturbances in Europe and their impact on carbon storage. *Nat. Clim. Change* **4**, 806–810 (2014).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

23. 23.

Senf, C. & Seidl, R. Mapping the forest disturbance regimes of Europe. *Nat. Sustain.* **4**, 63–70 (2021).

[Article](#) [Google Scholar](#)

24. 24.

European Commission. *A sustainable bioeconomy for Europe: strengthening the connection between economy, society and the environment*. COM(2018) 673 (2018).

25. 25.

FAOSTAT. *Forestry Production and Trade* (Food and Agriculture Organization of the United Nations, accessed July 2020);
<http://www.fao.org/faostat/en/#data/FO>.

26. 26.

Swedish Forest Agency. *Forest harvesting development in Sweden*.
<https://www.skogsstyrelsen.se/globalassets/statistik/bruttoavverkning/forest-harvesting-development-in-sweden.pdf> (2020).

27. 27.

Camia, A., Robert, N. & Pilli, R. *Biomass production, supply, uses and flows in the European Union: first results from an integrated assessment*. JRC Science for Policy Report EUR 28993 EN (Publications Office of the European Union, 2018).

28. 28.

Nabuurs, G.-J. et al. First signs of carbon sink saturation in European forest biomass. *Nat. Clim. Change* **3**, 792–796 (2013).

[ADS](#) [CAS](#) [Article](#) [Google Scholar](#)

29. 29.

Fernandez, R. et al. *Annual European Union greenhouse gas inventory 1990–2016 and inventory report 2018*. EEA Report No 5/2018 (European Environment Agency, 2018).

30. 30.

Korosuo, A. et al. *Forest reference levels under Regulation (EU) 2018/841 for the period 2021–2025: overview and main findings of the technical assessment*. JRC Science for Policy Report EUR 30403 EN (Publications Office of the European Union, 2020).

31. 31.

European Commission. *Stepping up Europe's 2030 climate ambition. Investing in a climate-neutral future for the benefit of our people*. COM/2020/562 (2020).

32. 32.

Santoro, M. GlobBiomass - global datasets of forest biomass. PANGAEA <https://doi.org/10.1594/PANGAEA.894711> (2018).

33. 33.

Forest Europe. *State of Europe's Forests 2015*. (Ministerial Conference on the Protection of Forests in Europe, 2015).

34. 34.

Gorelick, N. et al. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **202**, 18–27 (2017).

[ADS](#) [Article](#) [Google Scholar](#)

[Download references](#)

Acknowledgements

We thank F. Achard, H. Eva, F. Gianoli, M. Pickering, M. Piccardo, R. Abad Vinas, B. Eckhart, S. Rossi and B. Marcolla for the validation exercise; and A. Korosuo for clarifications of specific country circumstances and for useful editorial suggestions. The salvage loggings dataset was collected in 15 EU Member States by searching national datasets and/or consulting with national experts. Through the Standing Forestry Committee under the European Commission, the following Member States validated and/or acknowledged data on salvage logging and/or provided additional information: Austria, Bulgaria, Croatia, Czechia, Estonia, France, Finland, Germany, Hungary, Lithuania, Romania, Poland, Slovakia, Slovenia and Sweden. It is important to note that the time series for which annual data on salvage loggings are available varies among Member States.

Author information

Affiliations

1. Bio-Economy Unit, European Commission Joint Research Centre, Ispra, Italy

Guido Ceccherini, Gregory Duveiller, Giacomo Grassi, Valerio Avitabile, Roberto Pilli & Alessandro Cescatti

2. Food Security Unit, European Commission Joint Research Centre, Ispra, Italy

Guido Lemoine

Authors

1. Guido Ceccherini
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Gregory Duveiller

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Giacomo Grassi

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Guido Lemoine

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Valerio Avitabile

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Roberto Pilli

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Alessandro Cescatti

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Contributions

G.C., V.A. and A.C. designed the validation methodology. G.C. and G.D. analysed the data and wrote the Google Earth Engine and R scripts. G.C., V.A., G.G. and A.C. wrote the manuscript with contributions from G.D., G.L. and R.P. All authors contributed critically to the interpretation of the results and gave final approval for publication.

Corresponding author

Correspondence to [Guido Ceccherini](#).

Ethics declarations

Competing interests

The authors declare no competing interests.

Additional information

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

Supplementary information

[Reporting Summary](#)

Rights and permissions

[Reprints and Permissions](#)

About this article



Check for
updates

Cite this article

Ceccherini, G., Duveiller, G., Grassi, G. *et al.* Reply to Wernick, I. K. *et al.*; Palahí, M. *et al.*. *Nature* **592**, E18–E23 (2021).
<https://doi.org/10.1038/s41586-021-03294-9>

Download citation

- Published: 28 April 2021
- Issue Date: 29 April 2021
- DOI: <https://doi.org/10.1038/s41586-021-03294-9>

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Access through your institution](#)

[Change institution](#)

[Buy or subscribe](#)

Associated Content

[Abrupt increase in harvested forest area over Europe after 2015](#)

- Guido Ceccherini
- Gregory Duveiller
- Alessandro Cescatti

Nature Article 01 Jul 2020

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03294-9>

Amendments & Corrections

- [**Author Correction: Dense sampling of bird diversity increases power of comparative genomics**](#) [08 April 2021]
Author Correction •
- [**Author Correction: Protecting the global ocean for biodiversity, food and climate**](#) [08 April 2021]
Author Correction •
- [**Author Correction: High-depth African genomes inform human migration and health**](#) [12 April 2021]
Author Correction •
- [**Author Correction: poly\(UG\)-tailed RNAs in genome protection and epigenetic inheritance**](#) [14 April 2021]
Author Correction •
- [**Author Correction: Assessing progress towards sustainable development over space and time**](#) [14 April 2021]
Author Correction •
- [**Author Correction: AIM2 in regulatory T cells restrains autoimmune diseases**](#) [14 April 2021]
Author Correction •
- [**Addendum: Active superelasticity in three-dimensional epithelia of controlled shape**](#) [12 April 2021]
Addendum •

Author Correction: Dense sampling of bird diversity increases power of comparative genomics

[Download PDF](#)

- Author Correction
- Open Access
- [Published: 08 April 2021](#)

Author Correction: Dense sampling of bird diversity increases power of comparative genomics

- [Shaohong Feng](#) ORCID: [orcid.org/0000-0002-2462-7348^{1,2,3 na1}](https://orcid.org/0000-0002-2462-7348),
- [Josefin Stiller](#) ORCID: [orcid.org/0000-0001-6009-9581^{4 na1}](https://orcid.org/0000-0001-6009-9581),
- [Yuan Deng](#)^{1,3,4 na1},
- [Joel Armstrong](#) ORCID: [orcid.org/0000-0003-2077-4671^{5 na1}](https://orcid.org/0000-0003-2077-4671),
- [Qi Fang](#) ORCID: [orcid.org/0000-0002-9181-8689^{1,3,4}](https://orcid.org/0000-0002-9181-8689),
- [Andrew Hart Reeve](#) ORCID: [orcid.org/0000-0001-5233-6030⁶](https://orcid.org/0000-0001-5233-6030),
- [Duo Xie](#) ORCID: [orcid.org/0000-0001-7010-3601^{1,3,7}](https://orcid.org/0000-0001-7010-3601),
- [Guangji Chen](#) ORCID: [orcid.org/0000-0002-9441-1155^{1,3,7}](https://orcid.org/0000-0002-9441-1155),
- [Chunxue Guo](#)^{1,3},
- [Brant C. Faircloth](#) ORCID: [orcid.org/0000-0002-1943-0217^{8,9}](https://orcid.org/0000-0002-1943-0217),
- [Bent Petersen](#) ORCID: [orcid.org/0000-0002-2472-8317^{10,11}](https://orcid.org/0000-0002-2472-8317),
- [Zongji Wang](#) ORCID: [orcid.org/0000-0002-5107-6246^{1,3,12,13}](https://orcid.org/0000-0002-5107-6246),
- [Qi Zhou](#) ORCID: [orcid.org/0000-0002-7419-2047^{12,13,14}](https://orcid.org/0000-0002-7419-2047),
- [Mark Diekhans](#) ORCID: [orcid.org/0000-0002-0430-0989⁵](https://orcid.org/0000-0002-0430-0989),
- [Wanjun Chen](#) ORCID: [orcid.org/0000-0002-4418-7650^{1,3}](https://orcid.org/0000-0002-4418-7650),
- [Sergio Andreu-Sánchez](#)⁴,
- [Ashot Margaryan](#) ORCID: [orcid.org/0000-0002-2576-2429^{11,15}](https://orcid.org/0000-0002-2576-2429),
- [Jason Travis Howard](#)¹⁶,
- [Carole Parent](#)¹⁷,

- [George Pacheco](#) ORCID: orcid.org/0000-0002-9367-6813¹¹,
- [Mikkel-Holger S. Sinding](#) ORCID: orcid.org/0000-0003-1371-219X¹¹,
- [Lara Puetz](#) ORCID: orcid.org/0000-0002-0323-799X¹¹,
- [Emily Cavill](#) ORCID: orcid.org/0000-0003-2198-8102¹¹,
- [Angela M. Ribeiro](#)⁶,
- [Leopold Eckhart](#) ORCID: orcid.org/0000-0002-5645-2036¹⁸,
- [Jon Fjeldså](#) ORCID: orcid.org/0000-0003-0790-3600^{6,19},
- [Peter A. Hosner](#) ORCID: orcid.org/0000-0001-7499-6224^{6,19},
- [Robb T. Brumfield](#) ORCID: orcid.org/0000-0003-2307-0688^{8,9},
- [Les Christidis](#) ORCID: orcid.org/0000-0002-3345-6034²⁰,
- [Mads F. Bertelsen](#)²¹,
- [Thomas Sicheritz-Ponten](#)^{10,11},
- [Dieter Thomas Tietze](#) ORCID: orcid.org/0000-0001-6868-227X²²,
- [Bruce C. Robertson](#) ORCID: orcid.org/0000-0002-5348-2731²³,
- [Gang Song](#) ORCID: orcid.org/0000-0002-0190-8315^{24,25},
- [Gerald Borgia](#)²⁶,
- [Santiago Claramunt](#) ORCID: orcid.org/0000-0002-8926-5974^{27,28},
- [Irby J. Lovette](#) ORCID: orcid.org/0000-0003-0268-4875²⁹,
- [Saul J. Cowen](#) ORCID: orcid.org/0000-0002-1045-5637³⁰,
- [Peter Njoroge](#) ORCID: orcid.org/0000-0003-1165-3579³¹,
- [John Philip Dumbacher](#) ORCID: orcid.org/0000-0001-8942-1554³²,
- [Oliver A. Ryder](#) ORCID: orcid.org/0000-0003-2427-763X^{33,34},
- [Jérôme Fuchs](#) ORCID: orcid.org/0000-0003-1972-0078³⁵,
- [Michael Bunce](#) ORCID: orcid.org/0000-0002-0302-4206³⁶,
- [David W. Burt](#) ORCID: orcid.org/0000-0002-9991-1028³⁷,
- [Joel Cracraft](#) ORCID: orcid.org/0000-0001-7587-8342³⁸,
- [Guanliang Meng](#) ORCID: orcid.org/0000-0002-6488-1527¹,
- [Shannon J. Hackett](#) ORCID: orcid.org/0000-0002-1404-0332³⁹,
- [Peter G. Ryan](#)⁴⁰,
- [Knud Andreas Jönsson](#) ORCID: orcid.org/0000-0002-1875-9504⁶,
- [Ian G. Jamieson](#)^{23 na2},
- [Rute R. da Fonseca](#) ORCID: orcid.org/0000-0002-2805-4698¹⁹,
- [Edward L. Braun](#) ORCID: orcid.org/0000-0003-1643-5212⁴¹,

- [Peter Houde](#) ORCID: orcid.org/0000-0003-4541-5974⁴²,
- [Siavash Mirarab](#) ORCID: orcid.org/0000-0001-5410-1518⁴³,
- [Alexander Suh](#) ORCID: orcid.org/0000-0002-8979-9992^{44,45,46},
- [Bengt Hansson](#) ORCID: orcid.org/0000-0001-6694-8169⁴⁷,
- [Suvi Ponnikas](#) ORCID: orcid.org/0000-0003-3526-2118⁴⁷,
- [Hanna Sigeman](#) ORCID: orcid.org/0000-0002-1457-4174⁴⁷,
- [Martin Stervander](#) ORCID: orcid.org/0000-0002-6139-7828^{47,48},
- [Paul B. Frandsen](#) ORCID: orcid.org/0000-0002-4801-7579^{49,50},
- [Henriette van der Zwan](#) ORCID: orcid.org/0000-0001-5288-7609⁵¹,
- [Rencia van der Sluis](#) ORCID: orcid.org/0000-0003-3463-1700⁵¹,
- [Carina Visser](#)⁵²,
- [Christopher N. Balakrishnan](#) ORCID: orcid.org/0000-0002-0788-0659⁵³,
- [Andrew G. Clark](#) ORCID: orcid.org/0000-0001-7159-8511⁵⁴,
- [John W. Fitzpatrick](#)²⁹,
- [Reed Bowman](#) ORCID: orcid.org/0000-0002-6027-798X⁵⁵,
- [Nancy Chen](#) ORCID: orcid.org/0000-0001-8966-3449⁵⁶,
- [Alison Cloutier](#)^{57,58},
- [Timothy B. Sackton](#) ORCID: orcid.org/0000-0003-1673-9216⁵⁹,
- [Scott V. Edwards](#)^{57,58},
- [Dustin J. Foote](#)^{53,60},
- [Subir B. Shakya](#) ORCID: orcid.org/0000-0002-3561-2336^{8,9},
- [Frederick H. Sheldon](#) ORCID: orcid.org/0000-0002-0815-8780^{8,9},
- [Alain Vignal](#)⁶¹,
- [André E. R. Soares](#) ORCID: orcid.org/0000-0002-7768-2199^{62,63},
- [Beth Shapiro](#) ORCID: orcid.org/0000-0002-2733-7776^{63,64},
- [Jacob González-Solís](#) ORCID: orcid.org/0000-0002-8691-9397^{65,66},
- [Joan Ferrer-Obiol](#) ORCID: orcid.org/0000-0002-1184-5434^{65,67},
- [Julio Rozas](#) ORCID: orcid.org/0000-0002-6839-9148^{65,67},
- [Marta Riutort](#) ORCID: orcid.org/0000-0002-2134-7674^{65,67},
- [Anna Tigano](#) ORCID: orcid.org/0000-0001-9240-3058^{68,69},
- [Vicki Friesen](#)⁶⁹,
- [Love Dalén](#) ORCID: orcid.org/0000-0001-8270-7613^{70,71},
- [Araxi O. Urrutia](#) ORCID: orcid.org/0000-0001-9011-8675^{72,73},

- [Tamás Székely](#) ORCID: [orcid.org/0000-0003-2093-0056⁷²](https://orcid.org/0000-0003-2093-0056),
- [Yang Liu](#) ORCID: [orcid.org/0000-0003-4580-5518⁷⁴](https://orcid.org/0000-0003-4580-5518),
- [Michael G. Campana](#) ORCID: [orcid.org/0000-0003-0461-6462⁷⁵](https://orcid.org/0000-0003-0461-6462),
- [André Corvelo](#) ORCID: [orcid.org/0000-0003-0989-7806⁷⁶](https://orcid.org/0000-0003-0989-7806),
- [Robert C. Fleischer](#) ORCID: [orcid.org/0000-0002-2792-7055⁷⁵](https://orcid.org/0000-0002-2792-7055),
- [Kim M. Rutherford](#) ORCID: [orcid.org/0000-0001-6277-726X⁷⁷](https://orcid.org/0000-0001-6277-726X),
- [Neil J. Gemmell](#) ORCID: [orcid.org/0000-0003-0671-3637⁷⁷](https://orcid.org/0000-0003-0671-3637),
- [Nicolas Dussex](#) ORCID: [orcid.org/0000-0002-9179-8593^{70,71,77}](https://orcid.org/0000-0002-9179-8593),
- [Henrik Mouritsen](#) ORCID: [orcid.org/0000-0001-7082-4839⁷⁸](https://orcid.org/0000-0001-7082-4839),
- [Nadine Thiele](#) ORCID: [orcid.org/0000-0003-3399-5908⁷⁸](https://orcid.org/0000-0003-3399-5908),
- [Kira Delmore](#) ORCID: [orcid.org/0000-0003-4108-9729^{79,80}](https://orcid.org/0000-0003-4108-9729),
- [Miriam Liedvogel](#) ORCID: orcid.org/0000-0002-8372-8560,
- [Andre Franke](#) ORCID: [orcid.org/0000-0003-1530-5811⁸¹](https://orcid.org/0000-0003-1530-5811),
- [Marc P. Hoeppner](#)⁸¹,
- [Oliver Krone](#) ORCID: [orcid.org/0000-0002-4507-5124⁸²](https://orcid.org/0000-0002-4507-5124),
- [Adam M. Fudickar](#)⁸³,
- [Borja Milá](#) ORCID: [orcid.org/0000-0002-6446-0079⁸⁴](https://orcid.org/0000-0002-6446-0079),
- [Ellen D. Ketterson](#)⁸⁵,
- [Andrew Eric Fidler](#) ORCID: [orcid.org/0000-0003-2992-658X⁸⁶](https://orcid.org/0000-0003-2992-658X),
- [Guillermo Friis](#) ORCID: [orcid.org/0000-0002-0731-6468⁸⁷](https://orcid.org/0000-0002-0731-6468),
- [Ángela M. Parody-Merino](#) ORCID: [orcid.org/0000-0001-9945-1191⁸⁸](https://orcid.org/0000-0001-9945-1191),
- [Phil F. Battley](#) ORCID: [orcid.org/0000-0002-8590-8098⁸⁸](https://orcid.org/0000-0002-8590-8098),
- [Murray P. Cox](#) ORCID: [orcid.org/0000-0003-1936-0236⁸⁹](https://orcid.org/0000-0003-1936-0236),
- [Nicholas Costa Barroso Lima](#) ORCID: [orcid.org/0000-0002-3011-010X^{62,90}](https://orcid.org/0000-0002-3011-010X),
- [Francisco Prosdocimi](#) ORCID: [orcid.org/0000-0002-6761-3069⁹¹](https://orcid.org/0000-0002-6761-3069),
- [Thomas Lee Parchman](#) ORCID: [orcid.org/0000-0003-1771-1514⁹²](https://orcid.org/0000-0003-1771-1514),
- [Barney A. Schlinger](#)^{93,94},
- [Bette A. Loiselle](#) ORCID: [orcid.org/0000-0003-1434-4173^{95,96}](https://orcid.org/0000-0003-1434-4173),
- [John G. Blake](#) ORCID: [orcid.org/0000-0003-4668-2636⁹⁵](https://orcid.org/0000-0003-4668-2636),
- [Haw Chuan Lim](#) ORCID: [orcid.org/0000-0002-6420-1667^{75,97}](https://orcid.org/0000-0002-6420-1667),
- [Lainy B. Day](#) ORCID: [orcid.org/0000-0002-9122-3044⁹⁸](https://orcid.org/0000-0002-9122-3044),
- [Matthew J. Fuxjager](#)⁹⁹,

- [Maude W. Baldwin](#) ORCID: orcid.org/0000-0002-7601-8894¹⁰⁰,
- [Michael J. Braun](#) ORCID: orcid.org/0000-0001-8844-1756^{101,102},
- [Morgan Wirthlin](#) ORCID: orcid.org/0000-0001-7967-7070¹⁰³,
- [Rebecca B. Dikow](#) ORCID: orcid.org/0000-0001-9668-3683⁵⁰,
- [T. Brandt Ryder](#) ORCID: orcid.org/0000-0002-5517-6607¹⁰⁴,
- [Glauco Camenisch](#)¹⁰⁵,
- [Lukas F. Keller](#) ORCID: orcid.org/0000-0002-0149-0174¹⁰⁵,
- [Jeffrey M. DaCosta](#) ORCID: orcid.org/0000-0002-8590-2083¹⁰⁶,
- [Mark E. Hauber](#) ORCID: orcid.org/0000-0003-2014-4928¹⁰⁷,
- [Matthew I. M. Louder](#) ORCID: orcid.org/0000-0003-4421-541X^{53,107,108},
- [Christopher C. Witt](#) ORCID: orcid.org/0000-0003-2781-1543¹⁰⁹,
- [Jimmy A. McGuire](#) ORCID: orcid.org/0000-0002-9562-5585¹¹⁰,
- [Joann Mudge](#) ORCID: orcid.org/0000-0002-2972-1640¹¹¹,
- [Libby C. Megna](#)¹¹²,
- [Matthew D. Carling](#) ORCID: orcid.org/0000-0002-9401-7844¹¹²,
- [Biao Wang](#)¹¹³,
- [Scott A. Taylor](#) ORCID: orcid.org/0000-0001-9580-9125¹¹⁴,
- [Glacia Del-Rio](#)⁹,
- [Alexandre Aleixo](#) ORCID: orcid.org/0000-0002-7816-9725¹¹⁵,
- [Ana Tereza Ribeiro Vasconcelos](#) ORCID: orcid.org/0000-0002-4632-2086⁶²,
- [Claudio V. Mello](#) ORCID: orcid.org/0000-0002-9826-8421¹¹⁶,
- [Jason T. Weir](#) ORCID: orcid.org/0000-0001-8372-9937^{27,28,117},
- [David Haussler](#) ORCID: orcid.org/0000-0003-1533-4575⁵,
- [Qiye Li](#) ORCID: orcid.org/0000-0002-5993-0312^{1,3},
- [Huanming Yang](#)^{3,118},
- [Jian Wang](#)³,
- [Fumin Lei](#)^{24,119},
- [Carsten Rahbek](#)^{19,120,121,122},
- [M. Thomas P. Gilbert](#) ORCID: orcid.org/0000-0002-5805-7195^{11,123},
- [Gary R. Graves](#) ORCID: orcid.org/0000-0003-1406-5246^{19,101},
- [Erich D. Jarvis](#) ORCID: orcid.org/0000-0001-8931-5049^{17,124,125},
- [Benedict Paten](#) ORCID: orcid.org/0000-0001-8863-3539⁵ &

- [Guojie Zhang](#) [ORCID: orcid.org/0000-0001-6860-1521^{1,2,4,119}](#)

[Nature volume 592](#), page E24(2021) [Cite this article](#)

- 945 Accesses
- 4 Altmetric
- [Metrics details](#)

Subjects

- [Comparative genomics](#)
- [Evolutionary genetics](#)

The [Original Article](#) was published on 11 November 2020

[Download PDF](#)

Correction to: *Nature* <https://doi.org/10.1038/s41586-020-2873-9> Published online 11 November 2020

In [Supplementary Table 1](#) of this Article, 23 samples (B10K-DU-029-32, B10K-DU-029-33, B10K-DU-029-36 to B10K-DU-029-44, B10K-DU-029-46, B10K-DU-029-47, B10K-DU-029-49 to B10K-DU-029-53, B10K-DU-029-75 to B10K-DU-029-77, B10K-DU-029-80, and B10K-DU-030-03; styled in boldface in the revised table) were assigned to the incorrect institution. [Supplementary Table 1](#) has been amended to reflect the correct source institution for these samples, and associated data (tissue, museum ID/source specimen ID, site, state/province, latitude, longitude, date collected and sex) have been updated accordingly. The original table is provided as [Supplementary Information](#) to this Amendment, and the original Article has been corrected online.

Supplementary information is available in the online version of this Amendment.

Author information

Author notes

1. These authors contributed equally: Shaohong Feng, Josefina Stiller, Yuan Deng, Joel Armstrong
2. Deceased: Ian G. Jamieson

Affiliations

1. China National GeneBank, BGI-Shenzhen, Shenzhen, China

Shaohong Feng, Yuan Deng, Qi Fang, Duo Xie, Guangji Chen, Chunxue Guo, Zongji Wang, Wanjun Chen, Guanliang Meng, Qiye Li & Guojie Zhang

2. State Key Laboratory of Genetic Resources and Evolution, Kunming Institute of Zoology, Chinese Academy of Sciences, Kunming, China

Shaohong Feng & Guojie Zhang

3. BGI-Shenzhen, Shenzhen, China

Shaohong Feng, Yuan Deng, Qi Fang, Duo Xie, Guangji Chen, Chunxue Guo, Zongji Wang, Wanjun Chen, Qiye Li, Huanming Yang & Jian Wang

4. Villum Centre for Biodiversity Genomics, Section for Ecology and Evolution, Department of Biology, University of Copenhagen, Copenhagen, Denmark

Josefin Stiller, Yuan Deng, Qi Fang, Sergio Andreu-Sánchez & Guojie Zhang

5. UC Santa Cruz Genomics Institute, UC Santa Cruz, Santa Cruz, CA, USA

Joel Armstrong, Mark Diekhans, David Haussler & Benedict Paten

6. Natural History Museum of Denmark, University of Copenhagen,
Copenhagen, Denmark

Andrew Hart Reeve, Ångela M. Ribeiro, Jon Fjeldså, Peter A.
Hosner & Knud Andreas Jønsson

7. BGI Education Center, University of Chinese Academy of Sciences,
Shenzhen, China

Duo Xie & Guangji Chen

8. Department of Biological Sciences, Louisiana State University, Baton
Rouge, LA, USA

Brant C. Faircloth, Robb T. Brumfield, Subir B. Shakya & Frederick
H. Sheldon

9. Museum of Natural Science, Louisiana State University, Baton Rouge,
LA, USA

Brant C. Faircloth, Robb T. Brumfield, Subir B. Shakya, Frederick H.
Sheldon & Glaucia Del-Rio

10. Centre of Excellence for Omics-Driven Computational Biodiscovery
(COMBio), Faculty of Applied Sciences, AIMST University, Kedah,
Malaysia

Bent Petersen & Thomas Sicheritz-Ponten

11. Section for Evolutionary Genomics, The GLOBE Institute, Faculty of
Health and Medical Sciences, University of Copenhagen, Copenhagen,
Denmark

Bent Petersen, Ashot Margaryan, George Pacheco, Mikkel-Holger S.
Sinding, Lara Puetz, Emily Cavill, Thomas Sicheritz-Ponten & M.
Thomas P. Gilbert

12. MOE Laboratory of Biosystems Homeostasis and Protection, Life Sciences Institute, Zhejiang University, Hangzhou, China

Zongji Wang & Qi Zhou

13. Department of Neuroscience and Developmental Biology, University of Vienna, Vienna, Austria

Zongji Wang & Qi Zhou

14. Center for Reproductive Medicine, The 2nd Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, China

Qi Zhou

15. Institute of Molecular Biology, National Academy of Sciences, Yerevan, Armenia

Ashot Margaryan

16. Novogene, Durham, NC, USA

Jason Travis Howard

17. Duke University Medical Center, Durham, NC, USA

Carole Parent & Erich D. Jarvis

18. Department of Dermatology, Medical University of Vienna, Vienna, Austria

Leopold Eckhart

19. Center for Macroecology, Evolution, and Climate, GLOBE Institute, University of Copenhagen, Copenhagen, Denmark

Jon Fjeldså, Peter A. Hosner, Rute R. da Fonseca, Carsten Rahbek & Gary R. Graves

20. Southern Cross University, Coffs Harbour, New South Wales,
Australia

Les Christidis

21. Centre for Zoo and Wild Animal Health, Copenhagen Zoo,
Frederiksberg, Denmark

Mads F. Bertelsen

22. Center of Natural History, Universität Hamburg, Hamburg, Germany

Dieter Thomas Tietze

23. Department of Zoology, University of Otago, Dunedin, New Zealand

Bruce C. Robertson & Ian G. Jamieson

24. Key Laboratory of Zoological Systematics and Evolution, Institute of
Zoology, Chinese Academy of Sciences, Beijing, China

Gang Song & Fumin Lei

25. Environmental Futures Research Institute, Griffith University, Nathan,
Queensland, Australia

Gang Song

26. Department of Biology, University of Maryland, College Park, MD,
USA

Gerald Borgia

27. Department of Natural History, Royal Ontario Museum, Toronto,
Ontario, Canada

Santiago Claramunt & Jason T. Weir

28. Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Ontario, Canada
- Santiago Claramunt & Jason T. Weir
29. Cornell Lab of Ornithology, Cornell University, Ithaca, NY, USA
- Irby J. Lovette & John W. Fitzpatrick
30. Biodiversity and Conservation Science, Department of Biodiversity Conservation and Attractions, Perth, Western Australia, Australia
- Saul J. Cowen
31. Ornithology Section, Zoology Department, National Museums of Kenya, Nairobi, Kenya
- Peter Njoroge
32. Ornithology and Mammalogy, California Academy of Sciences, San Francisco, CA, USA
- John Philip Dumbacher
33. San Diego Zoo Institute for Conservation Research, Escondido, CA, USA
- Oliver A. Ryder
34. Evolution, Behavior, and Ecology, Division of Biology, University of California San Diego, La Jolla, CA, USA
- Oliver A. Ryder
35. Institut de Systématique, Evolution, Biodiversité (ISYEB), Muséum National d'Histoire Naturelle, CNRS, Sorbonne Université, EPHE, Université des Antilles, Paris, France
- Jérôme Fuchs

36. Trace and Environmental DNA (TrEnD) Laboratory, School of Molecular and Life Sciences, Curtin University, Western Australia, Perth, Australia

Michael Bunce

37. UQ Genomics, University of Queensland, Brisbane, Queensland, Australia

David W. Burt

38. Department of Ornithology, American Museum of Natural History, New York, NY, USA

Joel Cracraft

39. Integrative Research Center, Field Museum of Natural History, Chicago, IL, USA

Shannon J. Hackett

40. FitzPatrick Institute of African Ornithology, University of Cape Town, Cape Town, South Africa

Peter G. Ryan

41. Department of Biology, University of Florida, Gainesville, FL, USA

Edward L. Braun

42. Department of Biology, New Mexico State University, Las Cruces, NM, USA

Peter Houde

43. Department of Electrical and Computer Engineering, University of California San Diego, La Jolla, CA, USA

Siavash Mirarab

44. Department of Ecology and Genetics – Evolutionary Biology,
Evolutionary Biology Centre (EBC), Science for Life Laboratory,
Uppsala University, Uppsala, Sweden

Alexander Suh

45. Department of Organismal Biology – Systematic Biology,
Evolutionary Biology Centre (EBC), Science for Life Laboratory,
Uppsala University, Uppsala, Sweden

Alexander Suh

46. School of Biological Sciences, University of East Anglia, Norwich,
UK

Alexander Suh

47. Department of Biology, Lund University, Lund, Sweden

Bengt Hansson, Suvi Ponnikas, Hanna Sigeman & Martin Stervander

48. Institute of Ecology and Evolution, University of Oregon, Eugene, OR,
USA

Martin Stervander

49. Department of Plant and Wildlife Sciences, Brigham Young
University, Provo, UT, USA

Paul B. Frandsen

50. Data Science Lab, Office of the Chief Information Officer,
Smithsonian Institution, Washington, DC, USA

Paul B. Frandsen & Rebecca B. Dikow

51. Focus Area for Human Metabolomics, North-West University,
Potchefstroom, South Africa

Henriette van der Zwan & Rencia van der Sluis

52. Department of Animal Sciences, University of Pretoria, Pretoria, South Africa

Carina Visser

53. Department of Biology, East Carolina University, Greenville, NC, USA

Christopher N. Balakrishnan, Dustin J. Foote & Matthew I. M. Louder

54. Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY, USA

Andrew G. Clark

55. Avian Ecology Program, Archbold Biological Station, Venus, FL, USA

Reed Bowman

56. Department of Biology, University of Rochester, Rochester, NY, USA

Nancy Chen

57. Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA

Alison Cloutier & Scott V. Edwards

58. Museum of Comparative Zoology, Harvard University, Cambridge, MA, USA

Alison Cloutier & Scott V. Edwards

59. Informatics Group, Harvard University, Cambridge, MA, USA

Timothy B. Sackton

60. Sylvan Heights Bird Park, Scotland Neck, NC, USA
Dustin J. Foote
61. GenPhySE, INRA, INPT, INP-ENVT, Université de Toulouse,
Castanet-Tolosan, France
Alain Vignal
62. Laboratório Nacional de Computação Científica, Petrópolis, Brazil
André E. R. Soares, Nicholas Costa Barroso Lima & Ana Tereza
Ribeiro Vasconcelos
63. Department of Ecology and Evolutionary Biology, University of
California Santa Cruz, Santa Cruz, CA, USA
André E. R. Soares & Beth Shapiro
64. Howard Hughes Medical Institute, University of California Santa
Cruz, Santa Cruz, CA, USA
Beth Shapiro
65. Institut de Recerca de la Biodiversitat (IRBio), Universitat de
Barcelona, Barcelona, Spain
Jacob González-Solís, Joan Ferrer-Obiol, Julio Rozas & Marta Riutort
66. Departament de Biologia Evolutiva, Ecologia i Ciències Ambientals
(BEECA), Universitat de Barcelona, Barcelona, Spain
Jacob González-Solís
67. Departament de Genètica, Microbiologia i Estadística, Universitat de
Barcelona, Barcelona, Spain
Joan Ferrer-Obiol, Julio Rozas & Marta Riutort

68. Department of Molecular, Cellular and Biomedical Sciences,
University of New Hampshire, Durham, NH, USA

Anna Tigano

69. Department of Biology, Queen's University, Kingston, Ontario,
Canada

Anna Tigano & Vicki Friesen

70. Department of Bioinformatics and Genetics, Swedish Museum of
Natural History, Stockholm, Sweden

Love Dalén & Nicolas Dussex

71. Centre for Palaeogenetics, Stockholm, Sweden

Love Dalén & Nicolas Dussex

72. Milner Centre for Evolution, University of Bath, Bath, UK

Araxi O. Urrutia & Tamás Székely

73. Instituto de Ecología, UNAM, Mexico City, Mexico

Araxi O. Urrutia

74. State Key Laboratory of Biocontrol, School of Ecology, Sun Yat-sen
University, Guangzhou, China

Yang Liu

75. Center for Conservation Genomics, Smithsonian Conservation Biology
Institute, Smithsonian Institution, Washington, DC, USA

Michael G. Campana, Robert C. Fleischer & Haw Chuan Lim

76. New York Genome Center, New York, NY, USA

André Corvelo

77. Department of Anatomy, University of Otago, Dunedin, New Zealand

Kim M. Rutherford, Neil J. Gemmell & Nicolas Dussex

78. AG Neurosensory Sciences, Institut für Biologie und
Umweltwissenschaften, University of Oldenburg, Oldenburg,
Germany

Henrik Mouritsen & Nadine Thiele

79. Biology Department, Texas A&M University, College Station, TX,
USA

Kira Delmore

80. MPRG Behavioural Genomics, Max Planck Institute for Evolutionary
Biology, Plön, Germany

Kira Delmore

81. Institute of Clinical Molecular Biology, Christian-Albrechts-
University of Kiel, Kiel, Germany

Andre Franke & Marc P. Hoeppner

82. Department of Wildlife Diseases, Leibniz Institute for Zoo and
Wildlife Research, Berlin, Germany

Oliver Krone

83. Environmental Resilience Institute, Indiana University, Bloomington,
IN, USA

Adam M. Fudickar

84. National Museum of Natural Sciences, Spanish National Research
Council (CSIC), Madrid, Spain

Borja Milá

85. Department of Biology, Indiana University, Bloomington, IN, USA

Ellen D. Ketterson

86. Institute of Marine Science, University of Auckland, Auckland, New Zealand

Andrew Eric Fidler

87. Center for Genomics and Systems Biology, Department of Biology, New York University – Abu Dhabi, Abu Dhabi, UAE

Guillermo Friis

88. Wildlife and Ecology Group, Massey University, Palmerston North, New Zealand

Ángela M. Parody-Merino & Phil F. Battley

89. School of Fundamental Sciences, Massey University, Palmerston North, New Zealand

Murray P. Cox

90. Departamento de Bioquímica e Biologia Molecular, Centro de Ciências, Universidade Federal do Ceará, Fortaleza, Brazil

Nicholas Costa Barroso Lima

91. Laboratório de Genômica e Biodiversidade, Instituto de Bioquímica Médica Leopoldo de Meis, Rio de Janeiro, Brazil

Francisco Prosdocimi

92. Department of Biology, University of Nevada Reno, Reno, NV, USA

Thomas Lee Parchman

93. Department of Integrative Biology and Physiology, UCLA, Los Angeles, CA, USA
Barney A. Schlinger
94. Smithsonian Tropical Research Institute, Panama City, Panama
Barney A. Schlinger
95. Department of Wildlife Ecology and Conservation, University of Florida, Gainesville, FL, USA
Bette A. Loiselle & John G. Blake
96. Center for Latin American Studies, University of Florida, Gainesville, FL, USA
Bette A. Loiselle
97. Department of Biology, George Mason University, Fairfax, VA, USA
Haw Chuan Lim
98. Department of Biology and Neuroscience Minor, University of Mississippi, University, MS, USA
Lainy B. Day
99. Department of Ecology and Evolutionary Biology, Brown University, Providence, RI, USA
Matthew J. Fuxjager
100. Max Planck Institute for Ornithology, Seewiesen, Germany
Maude W. Baldwin
101. Department of Vertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington, DC, USA

Michael J. Braun & Gary R. Graves

102. Behavior, Ecology, Evolution and Systematics Program, University of Maryland, College Park, MD, USA

Michael J. Braun

103. Computational Biology Department, Carnegie Mellon University, Pittsburgh, PA, USA

Morgan Wirthlin

104. Migratory Bird Center, Smithsonian National Zoological Park and Conservation Biology Institute, Washington, DC, USA

T. Brandt Ryder

105. Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland

Glauco Camenisch & Lukas F. Keller

106. Biology Department, Boston College, Chestnut Hill, MA, USA

Jeffrey M. DaCosta

107. Department of Evolution, Ecology, and Behavior, School of Integrative Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA

Mark E. Hauber & Matthew I. M. Louder

108. International Research Center for Neurointelligence, University of Tokyo, Tokyo, Japan

Matthew I. M. Louder

109. Museum of Southwestern Biology, Department of Biology, University of New Mexico, Albuquerque, NM, USA

Christopher C. Witt

110. Museum of Vertebrate Zoology, Department of Integrative Biology,
University of California, Berkeley, Berkeley, CA, USA

Jimmy A. McGuire

111. National Center for Genome Resources, Santa Fe, NM, USA

Joann Mudge

112. Department of Zoology and Physiology, University of Wyoming,
Laramie, WY, USA

Libby C. Megna & Matthew D. Carling

113. School of BioSciences, The University of Melbourne, Melbourne,
Victoria, Australia

Biao Wang

114. Department of Ecology and Evolutionary Biology, University of
Colorado Boulder, Boulder, CO, USA

Scott A. Taylor

115. Finnish Museum of Natural History, University of Helsinki, Helsinki,
Finland

Alexandre Aleixo

116. Department of Behavioral Neuroscience, Oregon Health and Science
University, Portland, OR, USA

Claudio V. Mello

117. Department of Biological Sciences, University of Toronto
Scarborough, Toronto, Ontario, Canada

Jason T. Weir

118. James D. Watson Institute of Genome Sciences, Hangzhou, China

Huanming Yang

119. Center for Excellence in Animal Evolution and Genetics, Chinese Academy of Sciences, Kunming, China

Fumin Lei & Guojie Zhang

120. Danish Institute for Advanced Study, University of Southern Denmark, Odense, Denmark

Carsten Rahbek

121. Institute of Ecology, Peking University, Beijing, China

Carsten Rahbek

122. Department of Life Sciences, Imperial College London, Ascot, UK

Carsten Rahbek

123. University Museum, Norwegian University of Science and Technology, Trondheim, Norway

M. Thomas P. Gilbert

124. The Rockefeller University, New York, NY, USA

Erich D. Jarvis

125. Howard Hughes Medical Institute, Chevy Chase, MD, USA

Erich D. Jarvis

Authors

1. Shaohong Feng
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Josefina Stiller
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Yuan Deng
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Joel Armstrong
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Qi Fang
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Andrew Hart Reeve
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Duo Xie
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Guangji Chen
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. Chunxue Guo

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. Brant C. Faircloth

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

11. Bent Petersen

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

12. Zongji Wang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

13. Qi Zhou

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

14. Mark Diekhans

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

15. Wanjun Chen

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

16. Sergio Andreu-Sánchez

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

17. Ashot Margaryan

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

18. Jason Travis Howard

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

19. Carole Parent

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

20. George Pacheco

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

21. Mikkel-Holger S. Sinding

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

22. Lara Puetz

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

23. Emily Cavill

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

24. Ângela M. Ribeiro

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

25. Leopold Eckhart

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

26. Jon Fjeldså

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

27. Peter A. Hosner

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

28. Robb T. Brumfield

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

29. Les Christidis

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

30. Mads F. Bertelsen

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

31. Thomas Sicheritz-Ponten

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

32. Dieter Thomas Tietze

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

33. Bruce C. Robertson

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

34. Gang Song

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

35. Gerald Borgia

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

36. Santiago Claramunt

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

37. Irby J. Lovette

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

38. Saul J. Cowen

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

39. Peter Njoroge

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

40. John Philip Dumbacher

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

41. Oliver A. Ryder

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

42. Jérôme Fuchs

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

43. Michael Bunce

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

44. David W. Burt

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

45. Joel Cracraft

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

46. Guanliang Meng

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

47. Shannon J. Hackett

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

48. Peter G. Ryan

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

49. Knud Andreas Jønsson

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

50. Ian G. Jamieson

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

51. Rute R. da Fonseca

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

52. Edward L. Braun

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

53. Peter Houde

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

54. Siavash Mirarab

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

55. Alexander Suh

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

56. Bengt Hansson

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

57. Suvi Ponnikas

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

58. Hanna Sigeman

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

59. Martin Stervander

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

60. Paul B. Frandsen

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

61. Henriette van der Zwan

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

62. Rencia van der Sluis

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

63. Carina Visser

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

64. Christopher N. Balakrishnan

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

65. Andrew G. Clark

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

66. John W. Fitzpatrick

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

67. Reed Bowman

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

68. Nancy Chen

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

69. Alison Cloutier

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

70. Timothy B. Sackton

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

71. Scott V. Edwards

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

72. Dustin J. Foote

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

73. Subir B. Shakya

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

74. Frederick H. Sheldon

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

75. Alain Vignal

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

76. André E. R. Soares

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

77. Beth Shapiro

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

78. Jacob González-Solís

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

79. Joan Ferrer-Obiol

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

80. Julio Rozas

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

81. Marta Riutort

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

82. Anna Tigano

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

83. Vicki Friesen

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

84. Love Dalén

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

85. Araxi O. Urrutia

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

86. Tamás Székely

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

87. Yang Liu

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

88. Michael G. Campana

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

89. André Corvelo

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

90. Robert C. Fleischer

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

91. Kim M. Rutherford

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

92. Neil J. Gemmell

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

93. Nicolas Dussex

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

94. Henrik Mouritsen

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

95. Nadine Thiele

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

96. Kira Delmore

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

97. Miriam Liedvogel

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

98. Andre Franke

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

99. Marc P. Hoeppner

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

100. Oliver Krone

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

101. Adam M. Fudickar

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

102. Borja Milá

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

103. Ellen D. Ketterson

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

104. Andrew Eric Fidler

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

105. Guillermo Friis

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

106. Ángela M. Parody-Merino

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

107. Phil F. Battley

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

108. Murray P. Cox

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

109. Nicholas Costa Barroso Lima

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

110. Francisco Prosdocimi

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

111. Thomas Lee Parchman

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

112. Barney A. Schlinger

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

113. Bette A. Loiselle

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

114. John G. Blake

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

115. Haw Chuan Lim

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

116. Lainy B. Day

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

117. Matthew J. Fuxjager

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

118. Maude W. Baldwin

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

119. Michael J. Braun

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

120. Morgan Wirthlin

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

121. Rebecca B. Dikow

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

122. T. Brandt Ryder

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

123. Glauco Camenisch

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

124. Lukas F. Keller

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

125. Jeffrey M. DaCosta

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

126. Mark E. Hauber

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

127. Matthew I. M. Louder

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

128. Christopher C. Witt

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

129. Jimmy A. McGuire

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

130. Joann Mudge

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

131. Libby C. Megna

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

132. Matthew D. Carling

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

133. Biao Wang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

134. Scott A. Taylor

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

135. Glaucia Del-Rio

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

136. Alexandre Aleixo

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

137. Ana Tereza Ribeiro Vasconcelos

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

138. Claudio V. Mello

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

139. Jason T. Weir

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

140. David Haussler

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

141. Qiye Li

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

142. Huanming Yang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

143. Jian Wang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

144. Fumin Lei

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

145. Carsten Rahbek

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

146. M. Thomas P. Gilbert

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

147. Gary R. Graves

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

148. Erich D. Jarvis

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

149. Benedict Paten

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

150. Guojie Zhang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Corresponding authors

Correspondence to [Benedict Paten](#) or [Guojie Zhang](#).

Supplementary information

Supplementary Information

This file contains the original incorrect Supplementary Table 1.

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



Check for
updates

Cite this article

Feng, S., Stiller, J., Deng, Y. *et al.* Author Correction: Dense sampling of bird diversity increases power of comparative genomics. *Nature* **592**, E24 (2021). <https://doi.org/10.1038/s41586-021-03473-8>

Download citation

- Published: 08 April 2021
- Issue Date: 29 April 2021
- DOI: <https://doi.org/10.1038/s41586-021-03473-8>

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

Download PDF

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03473-8>

Author Correction: Protecting the global ocean for biodiversity, food and climate

[Download PDF](#)

- Author Correction
- [Published: 08 April 2021](#)

Author Correction: Protecting the global ocean for biodiversity, food and climate

- [Enric Sala](#) [ORCID: orcid.org/0000-0003-4730-3570¹](#),
- [Juan Mayorga](#) [ORCID: orcid.org/0000-0003-1961-8313^{1,2}](#),
- [Darcy Bradley](#) [ORCID: orcid.org/0000-0003-2581-8768²](#),
- [Reniel B. Cabral](#) [ORCID: orcid.org/0000-0002-1137-381X²](#),
- [Trisha B. Atwood](#) [ORCID: orcid.org/0000-0001-7153-5190³](#),
- [Arnaud Auber](#) [ORCID: orcid.org/0000-0002-8415-1652⁴](#),
- [William Cheung](#) [ORCID: orcid.org/0000-0001-9998-0384⁵](#),
- [Christopher Costello](#) [ORCID: orcid.org/0000-0002-9646-7806²](#),
- [Francesco Ferretti⁶](#),
- [Alan M. Friedlander^{1,7}](#),
- [Steven D. Gaines](#) [ORCID: orcid.org/0000-0002-7604-3483²](#),
- [Cristina Garilao¹⁸](#),
- [Whitney Goodell^{1,7}](#),
- [Benjamin S. Halpern](#) [ORCID: orcid.org/0000-0001-8844-2302⁹](#),
- [Audra Hinson](#) [ORCID: orcid.org/0000-0002-4231-4820³](#),
- [Kristin Kaschner⁸](#),
- [Kathleen Kesner-Reyes¹⁰](#),
- [Fabien Leprieur¹¹](#),
- [Jennifer McGowan](#) [ORCID: orcid.org/0000-0001-9061-3465¹²](#),
- [Lance E. Morgan¹³](#),

- [David Mouillot](#) ORCID: [orcid.org/0000-0003-0402-2605¹¹](https://orcid.org/0000-0003-0402-2605),
- [Juliano Palacios-Abrantes](#) ORCID: [orcid.org/0000-0001-8969-5416⁵](https://orcid.org/0000-0001-8969-5416),
- [Hugh P. Possingham](#) ORCID: [orcid.org/0000-0001-7755-996X¹⁴](https://orcid.org/0000-0001-7755-996X),
- [Kristin D. Rechberger](#)¹⁵,
- [Boris Worm](#)¹⁶ &
- [Jane Lubchenco](#) ORCID: [orcid.org/0000-0003-3540-5879¹⁷](https://orcid.org/0000-0003-3540-5879)

[Nature](#) volume 592, page E25(2021) [Cite this article](#)

- 921 Accesses
- 3 Altmetric
- [Metrics details](#)

Subjects

- [Environmental sciences](#)
- [Ocean sciences](#)

The [Original Article](#) was published on 17 March 2021

[Download PDF](#)

Correction to: *Nature* <https://www.nature.com/articles/s41586-021-03371-z>
Published online 17 March 2021

In this Article, the affiliation of author Cristina Garilao was incorrectly given as ‘Evolutionary Biology and Ecology Laboratory, Albert Ludwigs University, Freiburg, Germany.’ (affiliation 8). It should be ‘GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany’ (affiliation 18). The original Article has been corrected online.

Author information

Affiliations

1. Pristine Seas, National Geographic Society, Washington, DC, USA
Enric Sala, Juan Mayorga, Alan M. Friedlander & Whitney Goodell
2. Environmental Market Solutions Lab, University of California Santa Barbara, Santa Barbara, CA, USA
Juan Mayorga, Darcy Bradley, Reniel B. Cabral, Christopher Costello & Steven D. Gaines
3. Department of Watershed Sciences and Ecology Center, Utah State University, Logan, UT, USA
Trisha B. Atwood & Audra Hinson
4. IFREMER, Unité Halieutique de Manche et Mer du Nord, Boulogne-sur-Mer, France
Arnaud Auber
5. Changing Ocean Research Unit, Institute for the Oceans and Fisheries, The University of British Columbia, Vancouver, British Columbia, Canada
William Cheung & Juliano Palacios-Abrantes
6. Department of Fish and Wildlife Conservation, Virginia Polytechnic Institute and State University, Blacksburg, VA, USA
Francesco Ferretti
7. Hawai‘i Institute of Marine Biology, Kāne‘ohe, HI, USA
Alan M. Friedlander & Whitney Goodell
8. Evolutionary Biology and Ecology Laboratory, Albert Ludwigs University, Freiburg, Germany

Kristin Kaschner

9. National Center for Ecological Analysis and Synthesis (NCEAS),
University of California, Santa Barbara, CA, USA

Benjamin S. Halpern

10. Quantitative Aquatics, Los Baños, The Philippines

Kathleen Kesner-Reyes

11. MARBEC, Université de Montpellier, Montpellier, France

Fabien Leprieur & David Mouillot

12. The Nature Conservancy, Arlington, VA, USA

Jennifer McGowan

13. Marine Conservation Institute, Seattle, WA, USA

Lance E. Morgan

14. Centre for Biodiversity and Conservation Science (CBCS), The
University of Queensland, Brisbane, Queensland, Australia

Hugh P. Possingham

15. Dynamic Planet, Washington, DC, USA

Kristin D. Rechberger

16. Ocean Frontiers Institute, Dalhousie University, Halifax, Nova Scotia,
Canada

Boris Worm

17. Oregon State University, Corvallis, OR, USA

Jane Lubchenco

18. GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany

Cristina Garilao

Authors

1. Enric Sala

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Juan Mayorga

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Darcy Bradley

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Reniel B. Cabral

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Trisha B. Atwood

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Arnaud Auber

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. William Cheung

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Christopher Costello

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. Francesco Ferretti

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. Alan M. Friedlander

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

11. Steven D. Gaines

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

12. Cristina Garilao

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

13. Whitney Goodell

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

14. Benjamin S. Halpern

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

15. Audra Hinson

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

16. Kristin Kaschner

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

17. Kathleen Kesner-Reyes

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

18. Fabien Leprieur

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

19. Jennifer McGowan

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

20. Lance E. Morgan

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

21. David Mouillot

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

22. Juliano Palacios-Abrantes

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

23. Hugh P. Possingham

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

24. Kristin D. Rechberger

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

25. Boris Worm

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

26. Jane Lubchenco

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Corresponding author

Correspondence to [Enric Sala](#).

Rights and permissions

[Reprints and Permissions](#)

About this article



Check for
updates

Cite this article

Sala, E., Mayorga, J., Bradley, D. *et al.* Author Correction: Protecting the global ocean for biodiversity, food and climate. *Nature* **592**, E25 (2021). <https://doi.org/10.1038/s41586-021-03496-1>

[Download citation](#)

- Published: 08 April 2021
- Issue Date: 29 April 2021
- DOI: <https://doi.org/10.1038/s41586-021-03496-1>

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Download PDF](#)

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03496-1>

Author Correction: High-depth African genomes inform human migration and health

[Download PDF](#)

- Author Correction
- Open Access
- [Published: 12 April 2021](#)

Author Correction: High-depth African genomes inform human migration and health

- [Ananyo Choudhury](#)¹,
- [Shaun Aron](#)¹,
- [Laura R. Botigué](#)²,
- [Dhriti Sengupta](#)¹,
- [Gerrit Botha](#)³,
- [Taoufik Bensellak](#) [ORCID: orcid.org/0000-0002-6151-6276](#)⁴,
- [Gordon Wells](#)^{5,6,24},
- [Judit Kumuthini](#)^{5,6},
- [Daniel Shriner](#) [ORCID: orcid.org/0000-0003-1928-5520](#)⁷,
- [Yasmina J. Fakim](#)^{8,9},
- [Anisah W. Ghoorah](#) [ORCID: orcid.org/0000-0001-8273-6000](#)⁹,
- [Eileen Dareng](#) [ORCID: orcid.org/0000-0003-0802-419X](#)^{10,11},
- [Trust Odia](#)¹²,
- [Oluwadamilare Falola](#)¹²,
- [Ezekiel Adebisi](#)^{12,13},
- [Scott Hazelhurst](#) [ORCID: orcid.org/0000-0002-0581-149X](#)^{1,14},
- [Gaston Mazandu](#)³,
- [Oscar A. Nyangiri](#) [ORCID: orcid.org/0000-0003-2741-2921](#)¹⁵,
- [Mamana Mbiyavanga](#)³,

- [Alia Benkahla¹⁶](#),
- [Samar K. Kassim](#) ORCID: [orcid.org/0000-0002-4359-6620¹⁷](https://orcid.org/0000-0002-4359-6620),
- [Nicola Mulder³](#),
- [Sally N. Adebamowo^{18,19}](#),
- [Emile R. Chimusa²⁰](#),
- [Donna Muzny²¹](#),
- [Ginger Metcalf²¹](#),
- [Richard A. Gibbs^{21,22}](#),
- [TrypanoGEN Research Group](#),
- [Charles Rotimi](#) ORCID: [orcid.org/0000-0001-5759-053X⁷](https://orcid.org/0000-0001-5759-053X),
- [Michèle Ramsay](#) ORCID: [orcid.org/0000-0002-4156-4801^{1,23}](https://orcid.org/0000-0002-4156-4801),
- [H3Africa Consortium](#),
- [Adebowale A. Adeyemo](#) ORCID: [orcid.org/0000-0002-3105-3231⁷](https://orcid.org/0000-0002-3105-3231),
- [Zané Lombard](#) ORCID: [orcid.org/0000-0002-7997-2616²³](https://orcid.org/0000-0002-7997-2616) &
- [Neil A. Hanchard](#) ORCID: [orcid.org/0000-0003-1925-2665²²](https://orcid.org/0000-0003-1925-2665)

[Nature](#) volume **592**, page E26(2021) [Cite this article](#)

- 515 Accesses
- 2 Altmetric
- [Metrics details](#)

Subjects

- [Evolutionary genetics](#)
- [Genetic variation](#)
- [Genetics research](#)
- [Next-generation sequencing](#)

The [Original Article](#) was published on 28 October 2020

[Download PDF](#)

Correction to: *Nature* <https://doi.org/10.1038/s41586-020-2859-7> Published online 28 October 2020

In this Article, the affiliation ‘South African National Bioinformatics Network, University of the Western Cape, Bellville, South Africa’, which is associated with authors Gordon Wells and Judit Kumuthini, should be ‘South African National Bioinformatics **Institute**, University of the Western Cape, Bellville, South Africa’. In addition, author Gordon Wells should be associated with an additional affiliation, ‘Africa Health Research Institute, Durban, South Africa’. In Fig. 1a, the circle representing the data for South Africa should be dark grey (indicating the AGVP data source) rather than light grey (indicating the 1000G data source). The Article has been corrected online.

Author information

Affiliations

1. Sydney Brenner Institute for Molecular Bioscience, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa
Ananyo Choudhury, Shaun Aron, Dhriti Sengupta, Scott Hazelhurst, Michèle Ramsay, Ananyo Choudhury, Shaun Aron, Dhriti Sengupta, Scott Hazelhurst, Michèle Ramsay, Godfred Agongo & Romuald P. Boua
2. Center for Research in Agricultural Genomics (CRAG), Plant and Animal Genomics Program, CSIC-IRTA-UAB-UB, Barcelona, Spain
Laura R. Botigué & Laura Botigué
3. Computational Biology Division and H3ABioNet, Department of Integrative Biomedical Sciences, IDM, University of Cape Town, Cape Town, South Africa
Gerrit Botha, Gaston Mazandu, Mamana Mbiyavanga, Nicola Mulder, Gerrit Botha, Gaston Mazandu, Mamana

Mbiyavanga & Nicola Mulder

4. System and Data Engineering Team, Abdelmalek Essaadi University, ENSA, Tangier, Morocco

Taoufik Bensellak, Taoufik Bensellak & Ahmed Moussa

5. Centre for Proteomic and Genomic Research (CPGR), Cape Town, South Africa

Gordon Wells, Judit Kumuthini & Gordon Wells

6. South African National Bioinformatics Institute, University of the Western Cape, Bellville, South Africa

Gordon Wells, Judit Kumuthini, Gordon Wells & Judit Kumuthini

7. Center for Research on Genomics and Global Health, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD, USA

Daniel Shriner, Charles Rotimi, Daniel Shriner, Charles Rotimi, Adebawale A. Adeyemo & Adebawale A. Adeyemo

8. Department of Agriculture and Food Science, Faculty of Agriculture, University of Mauritius, Reduit, Mauritius

Yasmina J. Fakim & Yasmina J. Fakim

9. Department of Digital Technologies, Faculty of Information, Communication & Digital Technologies, University of Mauritius, Reduit, Mauritius

Yasmina J. Fakim, Anisah W. Ghoorah, Yasmina J. Fakim & Anisah W. Ghoorah

10. Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK

Eileen Dareng & Eileen Dareng

11. Institute of Human Virology Nigeria, Abuja, Nigeria

Eileen Dareng, Eileen Dareng & Clement Adebamowo

12. Covenant University Bioinformatics Research (CUBRe), Covenant University, Ota, Nigeria

Trust Odia, Oluwadamilare Falola, Ezekiel Adebiyi, Trust Odia, Oluwadamilare Falola & Ezekiel Adebiyi

13. Department of Computer and Information Sciences, Covenant University, Ota, Nigeria

Ezekiel Adebiyi & Ezekiel Adebiyi

14. School of Electrical and Information Engineering, University of the Witwatersrand, Johannesburg, South Africa

Scott Hazelhurst & Scott Hazelhurst

15. College of Veterinary Medicine, Animal Resources and Biosecurity, Makerere University, Kampala, Uganda

Oscar A. Nyangiri, Enock Matovu, Oscar A. Nyangiri, Oscar A. Nyangiri, Enock Matovu & Oscar A. Nyangiri

16. Laboratory of Bioinformatics, Biomathematics and Biostatistics (BIMS), Institute Pasteur of Tunis, Tunis, Tunisia

Alia Benkahla & Alia Benkahla

17. Medical Biochemistry and Molecular Biology Department, Faculty of Medicine, Ain Shams University, Abbaseya, Cairo, Egypt

Samar K. Kassim & Samar K. Kassim

18. Department of Epidemiology and Public Health, University of Maryland School of Medicine, University of Maryland Baltimore, Baltimore, MD, USA

Sally N. Adebamowo, Sally N. Adebamowo & Clement Adebamowo

19. University of Maryland Greenebaum Comprehensive Cancer Center, University of Maryland School of Medicine, University of Maryland Baltimore, Baltimore, MD, USA

Sally N. Adebamowo, Sally N. Adebamowo & Clement Adebamowo

20. Division of Human Genetics, Department of Pathology, Faculty of Health Sciences, Institute for Infectious, Disease and Molecular Medicine, University of Cape Town, Cape Town, South Africa

Emile R. Chimusa, Emile R. Chimusa & Ambroise Wonkam

21. Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX, USA

Donna Muzny, Ginger Metcalf & Richard A. Gibbs

22. Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, Texas, USA

Richard A. Gibbs, Neil A. Hanchard, Graeme Mardon & Neil A. Hanchard

23. Division of Human Genetics, National Health Laboratory Service, and School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

Michèle Ramsay, Michèle Ramsay, Zané Lombard & Zané Lombard

24. Africa Health Research Institute, Durban, South Africa

Gordon Wells & Gordon Wells

25. Institut de Recherche pour le Développement, Montpellier, France

Bruno Bucheton & Bruno Bucheton

26. Institute of Integrative Biology, University of Liverpool, Liverpool, UK

Christiane Hertz-Fowler, Harry Noyes, Christiane Hertz-Fowler & Harry Noyes

27. Wellcome Trust, London, UK

Christiane Hertz-Fowler & Christiane Hertz-Fowler

28. Jean Lorougnon Guede University, Daloa, Côte d'Ivoire

Mathurin Koffi & Mathurin Koffi

29. Institute of Biodiversity Animal Health and Comparative Medicine, University of Glasgow, Glasgow, UK

Annette Macleod & Annette Macleod

30. Institut National de Recherche Biomedicale, Kinshasa, Democratic Republic of Congo

Dieudonne Mumba-Ngoyi & Dieudonne Mumba-Ngoyi

31. Faculty of Science, University of Dschang, Dschang, Cameroon

Gustave Simo & Gustave Simo

32. School of Veterinary Medicine, University of Zambia, Lusaka, Zambia

Martin Simunza & Martin Simunza

33. Institute of Human Virology and Greenebaum Cancer Center, University of Maryland School of Medicine, Baltimore, MD, USA

Clement Adebamowo

34. Navrongo Health Research Centre, Navrongo, Ghana

Godfred Agongo & Abraham Oduro

35. Clinical Research Unit of Nanoro, Institut de Recherche en Sciences de la Sante, Bobo-Dioulasso, Burkina Faso

Romuald P. Boua & Hermann Sorgho

36. Faculty of Medicine and Odontostomatology, University of Science, Techniques and Technologies of Bamako (USTTB), Bamako, Mali

Guida Landouré & Oumar Samassékou

37. Service de Neurologie, Centre Hospitalier Universitaire du Point “G”, Bamako, Mali

Guida Landouré, Lassana Cissé & Salimata Diarra

38. Neurogenetics Branch, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD, USA

Salimata Diarra

39. Botswana-Baylor Children’s Clinical Centre of Excellence, Gaborone, Botswana

Gabriel Anabwani & Mogomotsi Matshaba

40. Medical Microbiology, College of Health Sciences, Makerere University, Kampala, Uganda

Moses Joloba, Samuel Kyobe & Savannah Mwesigwa

41. Baylor College of Medicine Children’s Foundation, Kampala, Uganda

Adeodata Kekitiinwa

42. Department of Pathology and Immunology, Baylor College of Medicine, Houston, TX, USA

Graeme Mardon

43. Department of Biological Sciences, University of Botswana, Gaborone, Botswana

Sununguko W. Mpoloka, Busisiwe Mlotshwa, Gaone Retshabile & Lesedi Williams

44. University of Ghana Medical School, Accra, Ghana

Dwomoa Adu

45. The University of Kansas School of Medicine, Kansas City, KS, USA

Akinlolu Ojo

46. Department of Human Genetics, University of Michigan Medical School, Ann Arbor, MI, USA

David Burke

47. College of Medicine, University of Ibadan, Ibadan, Nigeria

Babatunde O. Salako

48. Department of Molecular Genetics, University of Toronto, Toronto, Ontario, Canada

Philip Awadalla

49. Ontario Institute for Cancer Research, Toronto, Ontario, Canada

Philip Awadalla, Vanessa Bruat & Elias Gbeha

Authors

1. Ananyo Choudhury
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Shaun Aron
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Laura R. Botigué
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Dhriti Sengupta
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Gerrit Botha
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Taoufik Bensellak
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Gordon Wells
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Judit Kumuthini
[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. Daniel Shriner

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. Yasmina J. Fakim

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

11. Anisah W. Ghoorah

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

12. Eileen Dareng

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

13. Trust Odia

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

14. Oluwadamilare Falola

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

15. Ezekiel Adebisi

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

16. Scott Hazelhurst

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

17. Gaston Mazandu

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

18. Oscar A. Nyangiri

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

19. Mamana Mbiyavanga

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

20. Alia Benkahla

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

21. Samar K. Kassim

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

22. Nicola Mulder

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

23. Sally N. Adebamowo

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

24. Emile R. Chimusa

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

25. Donna Muzny

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

26. Ginger Metcalf

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

27. Richard A. Gibbs

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

28. Charles Rotimi

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

29. Michèle Ramsay

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

30. Adebowale A. Adeyemo

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

31. Zané Lombard

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

32. Neil A. Hanchard

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Consortia

TrypanoGEN Research Group

- Enock Matovu
- , Bruno Bucheton
- , Christiane Hertz-Fowler
- , Mathurin Koffi
- , Annette Macleod
- , Dieudonne Mumba-Ngoyi
- , Harry Noyes
- , Oscar A. Nyangiri
- , Gustave Simo
- & Martin Simuunza

H3Africa Consortium

- Ananyo Choudhury
- , Shaun Aron
- , Laura Botigué
- , Dhriti Sengupta
- , Gerrit Botha
- , Taoufik Bensellak
- , Gordon Wells
- , Judit Kumuthini
- , Daniel Shriner
- , Yasmina J. Fakim
- , Anisah W. Ghoorah
- , Eileen Dareng
- , Trust Odia
- , Oluwadamilare Falola
- , Ezekiel Adebiyi
- , Scott Hazelhurst
- , Gaston Mazandu
- , Oscar A. Nyangiri
- , Mamana Mbiyavanga
- , Alia Benkahla

- , Samar K. Kassim
- , Nicola Mulder
- , Sally N. Adebamowo
- , Emile R. Chimusa
- , Charles Rotimi
- , Michèle Ramsay
- , Adebowale A. Adeyemo
- , Zané Lombard
- , Neil A. Hanchard
- , Clement Adebamowo
- , Godfred Agongo
- , Romuald P. Boua
- , Abraham Oduro
- , Hermann Sorgho
- , Guida Landouré
- , Lassana Cissé
- , Salimata Diarra
- , Oumar Samassékou
- , Gabriel Anabwani
- , Mogomotsi Matshaba
- , Moses Joloba
- , Adeodata Kekitiinwa
- , Graeme Mardon
- , Sununguko W. Mpoloka
- , Samuel Kyobe
- , Busisiwe Mlotshwa
- , Savannah Mwesigwa
- , Gaone Retshabile
- , Lesedi Williams
- , Ambroise Wonkam
- , Ahmed Moussa
- , Dwomoa Adu
- , Akinlolu Ojo
- , David Burke
- , Babatunde O. Salako
- , Enock Matovu
- , Bruno Bucheton

- , Christiane Hertz-Fowler
- , Mathurin Koffi
- , Annette Macleod
- , Dieudonne Mumba-Ngoyi
- , Harry Noyes
- , Oscar A. Nyangiri
- , Gustave Simo
- , Martin Simuunza
- , Philip Awadalla
- , Vanessa Bruat
- & Elias Gbeha

Corresponding authors

Correspondence to [Adebowale A. Adeyemo](#) or [Zané Lombard](#) or [Neil A. Hanchard](#).

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



Check for
updates

Cite this article

Choudhury, A., Aron, S., Botigué, L.R. *et al.* Author Correction: High-depth African genomes inform human migration and health. *Nature* **592**, E26 (2021). <https://doi.org/10.1038/s41586-021-03286-9>

[Download citation](#)

- Published: 12 April 2021
- Issue Date: 29 April 2021
- DOI: <https://doi.org/10.1038/s41586-021-03286-9>

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Download PDF](#)

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03286-9>

Author Correction: poly(UG)-tailed RNAs in genome protection and epigenetic inheritance

[Download PDF](#)

- Author Correction
- [Published: 14 April 2021](#)

Author Correction: poly(UG)-tailed RNAs in genome protection and epigenetic inheritance

- [Aditi Shukla](#) ORCID: orcid.org/0000-0001-9743-4128^{1 na1},
- [Jenny Yan](#)^{1 na1},
- [Daniel J. Pagano](#) ORCID: orcid.org/0000-0002-6513-3870¹,
- [Anne E. Dodson](#) ORCID: orcid.org/0000-0001-7719-0496¹,
- [Yuhan Fei](#)^{1,2},
- [Josh Gorham](#)¹,
- [J. G. Seidman](#)¹,
- [Marvin Wickens](#) ORCID: orcid.org/0000-0002-0593-5740³ &
- [Scott Kennedy](#) ORCID: orcid.org/0000-0002-7974-8155¹

[Nature](#) volume 592, pageE27(2021)[Cite this article](#)

- 413 Accesses
- 2 Altmetric
- [Metrics details](#)

Subjects

- [Epigenetic memory](#)

- [Gene silencing](#)
- [Post-translational modifications](#)
- [RNAi](#)
- [siRNAs](#)

The [Original Article](#) was published on 20 May 2020

[Download PDF](#)

Correction to: *Nature* <https://doi.org/10.1038/s41586-020-2323-8> Published online 20 May 2020

In this Article, the following grants were omitted from the Acknowledgements section: NIH grants GM088289 and GM132286 to S.K., and NIH grant GM50942 to M.W. The original Article has been corrected online.

Author information

Author notes

1. These authors contributed equally: Aditi Shukla, Jenny Yan

Affiliations

1. Department of Genetics, Blavatnik Institute at Harvard Medical School, Boston, MA, USA

Aditi Shukla, Jenny Yan, Daniel J. Pagano, Anne E. Dodson, Yuhan Fei, Josh Gorham, J. G. Seidman & Scott Kennedy

2. College of Agriculture, Nanjing Agricultural University, Nanjing, China

Yuhan Fei

3. Department of Biochemistry, University of Wisconsin–Madison,
Madison, WI, USA

Marvin Wickens

Authors

1. Aditi Shukla

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Jenny Yan

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Daniel J. Pagano

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Anne E. Dodson

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Yuhang Fei

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Josh Gorham

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. J. G. Seidman

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Marvin Wickens

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. Scott Kennedy

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Corresponding author

Correspondence to [Scott Kennedy](#).

Rights and permissions

[Reprints and Permissions](#)

About this article



Check for
updates

Cite this article

Shukla, A., Yan, J., Pagano, D.J. *et al.* Author Correction: poly(UG)-tailed RNAs in genome protection and epigenetic inheritance. *Nature* **592**, E27 (2021). <https://doi.org/10.1038/s41586-021-03364-y>

[Download citation](#)

- Published: 14 April 2021

- Issue Date: 29 April 2021
- DOI: <https://doi.org/10.1038/s41586-021-03364-y>

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Download PDF](#)

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03364-y>.

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

Author Correction: Assessing progress towards sustainable development over space and time

[Download PDF](#)

- Author Correction
- [Published: 14 April 2021](#)

Author Correction: Assessing progress towards sustainable development over space and time

- [Zhenci Xu](#)¹,
- [Sophia N. Chau](#)¹,
- [Xiuzhi Chen](#)²,
- [Jian Zhang](#)³,
- [Yingjie Li](#)¹,
- [Thomas Dietz](#)^{1,4},
- [Jinyan Wang](#)²,
- [Julie A. Winkler](#)⁵,
- [Fan Fan](#)⁶,
- [Baorong Huang](#)⁷,
- [Shuxin Li](#)¹,
- [Shaohua Wu](#)⁸,
- [Anna Herzberger](#)¹,
- [Ying Tang](#)^{1,5},
- [Dequ Hong](#)⁹,
- [Yunkai Li](#)² &
- [Jianguo Liu](#) ORCID: orcid.org/0000-0001-6344-0087¹

[Nature](#) volume 592, page E28(2021) [Cite this article](#)

- 405 Accesses
- 2 Altmetric
- [Metrics details](#)

Subjects

- [Developing world](#)
- [Sustainability](#)

The [Original Article](#) was published on 01 January 2020

[Download PDF](#)

Correction to: *Nature* <https://doi.org/10.1038/s41586-019-1846-3> Published online 01 January 2020

In this Article, there were several minor errors to the main text and Supplementary Information, as detailed below. In the sentence “The change in SDG Index score among provinces from 2000 to 2015 ranged from a 11.1% increase (**Shandong**) to a 51.8% increase (Ningxia).”, ‘Shandong’ was incorrectly shown as ‘Shanghai’. In the sentence “The three SDGs that improved the most, in order of greatest to least improvement, were SDG 9 (industry, innovation and infrastructure), SDG 10 (reduced inequalities), and SDG 17 (**partnerships for the goals**).”, ‘partnerships for the goals’ was incorrectly shown as ‘affordable and clean energy’.

In the Methods, ‘for SDG scores’ should be removed from the heading ‘Uncertainty and sensitivity analysis for SDG scores’. In this paragraph, the asterisk should be removed from the following equation so it appears as:

$$\$\{ C\}_{n}^m = \frac{n!}{m!(n-m)!}$$

Instead of:

$$\$\{ C\}_{n}^m = \frac{n!}{m! \ast (n-m)!}$$

In the Methods paragraph starting “We also ran a sensitivity analysis³⁷...”, seven occurrences of ‘SDG scores’ should have been ‘SDG **Index** scores’. Similarly, in the legend to Extended Data Fig. 3, four occurrences of ‘SDG scores’ should have been ‘SDG **Index** scores’. In the Supplementary Information, several minor changes have been made, including changing “smaller (larger)” to “larger (smaller)”; replacing “Shanxi” with “Jiangxi”; and replacing “the coefficients” with “the **multiple regression** coefficients”. The original Article has been corrected online.

Author information

Affiliations

1. Center for Systems Integration and Sustainability, Department of Fisheries and Wildlife, Michigan State University, East Lansing, MI, USA
Zhenci Xu, Sophia N. Chau, Yingjie Li, Thomas Dietz, Shuxin Li, Anna Herzberger, Ying Tang & Jianguo Liu
2. College of Water Resources and Civil Engineering, China Agricultural University, Beijing, China
Xiuzhi Chen, Jinyan Wang & Yunkai Li
3. School of Life Sciences, State Key Laboratory of Grassland Agro-ecosystems, Lanzhou University, Lanzhou, China
Jian Zhang
4. Environmental Science and Policy Program, Department of Sociology and Animal Studies Program, Michigan State University, East Lansing, MI, USA
Thomas Dietz

5. Department of Geography, Environment, and Spatial Sciences,
Michigan State University, East Lansing, MI, USA

Julie A. Winkler & Ying Tang

6. School of Economics, Renmin University of China, Beijing, China

Fan Fan

7. Institutes of Science and Development, Chinese Academy of Sciences,
Beijing, China

Baorong Huang

8. Institute of Land and Urban-Rural Development, Zhejiang University
of Finance and Economics, Hangzhou, China

Shaohua Wu

9. Department of Humanities and Information, Zhejiang College of
Construction, Hangzhou, China

Dequ Hong

Authors

1. Zhenci Xu

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Sophia N. Chau

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Xiuzhi Chen

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Jian Zhang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Yingjie Li

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Thomas Dietz

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Jinyan Wang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Julie A. Winkler

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. Fan Fan

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. Baorong Huang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

11. Shuxin Li

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

12. Shaohua Wu

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

13. Anna Herzberger

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

14. Ying Tang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

15. Dequ Hong

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

16. Yunkai Li

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

17. Jianguo Liu

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Corresponding authors

Correspondence to [Yunkai Li](#) or [Jianguo Liu](#).

Rights and permissions

[Reprints and Permissions](#)

About this article



Check for
updates

Cite this article

Xu, Z., Chau, S.N., Chen, X. *et al.* Author Correction: Assessing progress towards sustainable development over space and time. *Nature* **592**, E28 (2021). <https://doi.org/10.1038/s41586-021-03479-2>

[Download citation](#)

- Published: 14 April 2021
- Issue Date: 29 April 2021
- DOI: <https://doi.org/10.1038/s41586-021-03479-2>

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Download PDF](#)

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03479-2>

Author Correction: AIM2 in regulatory T cells restrains autoimmune diseases

[Download PDF](#)

- Author Correction
- [Published: 14 April 2021](#)

Author Correction: AIM2 in regulatory T cells restrains autoimmune diseases

- [Wei-Chun Chou](#)^{1,2 na1},
- [Zengli Guo](#) ORCID: orcid.org/0000-0003-2950-4790^{1,3 na1},
- [Hao Guo](#) ORCID: orcid.org/0000-0001-8374-0132^{1,2},
- [Liang Chen](#)^{1,3},
- [Ge Zhang](#)^{1,3},
- [Kaixin Liang](#) ORCID: orcid.org/0000-0002-6751-3367^{1,4},
- [Ling Xie](#)⁵,
- [Xianming Tan](#) ORCID: orcid.org/0000-0002-5478-2269¹,
- [Sara A. Gibson](#)¹,
- [Elena Rampanelli](#) ORCID: orcid.org/0000-0002-7742-0092^{1,2},
- [Yan Wang](#) ORCID: orcid.org/0000-0001-5000-960X¹,
- [Stephanie A. Montgomery](#) ORCID: orcid.org/0000-0001-8012-5302^{1,6},
- [W. June Brickey](#)^{1,3},
- [Meng Deng](#) ORCID: orcid.org/0000-0002-9051-7534^{1,4},
- [Leslie Freeman](#)^{1,2},
- [Song Zhang](#)^{1,3},
- [Maureen A. Su](#)^{1,7 nAff8},
- [Xian Chen](#)⁵,
- [Yisong Y. Wan](#) ORCID: orcid.org/0000-0003-4346-9717^{1,3} &

- [Jenny P.-Y. Ting](#) ORCID: [orcid.org/0000-0002-7846-0395^{1,2,3,4}](https://orcid.org/0000-0002-7846-0395)

[Nature](#) volume **592**, page E29(2021) [Cite this article](#)

- 443 Accesses
- 2 Altmetric
- [Metrics details](#)

Subjects

- [Autoimmunity](#)
- [Immunology](#)

The [Original Article](#) was published on 27 January 2021

[Download PDF](#)

Correction to: *Nature* <https://doi.org/10.1038/s41586-021-03231-w> Published online 27 January 2021

In Fig. 3h of this Article, along the bottom row labelled ‘IL-2’, the third lane (above the ‘WT’ bracket) should be a plus and not a minus symbol, and the fourth lane (above the ‘*Aim2*^{-/-}’ bracket) should be a minus and not a plus; from left to right, the symbols for the ‘IL-2’ row should therefore be: ‘– + – +’. In Extended Data Fig. 2k, the first label should be ‘IL-17A⁻IFN γ ⁺’ rather than ‘IL-17A⁺IFN γ ⁻’, and the third label should be ‘IL-17A⁺IFN γ ⁻’ rather than ‘IL-17A⁻IFN γ ⁺’. In addition, in the final sentence of the Fig. 4 legend, the text should be ‘two-sided *t*-test (**a–c**, **e** right, **m** right)’ and not ‘two-sided *t*-test (**a–c**, **e** right)’. The original Article has been corrected online.

Author information

Author notes

1. Maureen A. Su

Present address: Department of Microbiology Immunology and Medical Genetics and Pediatrics, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA

2. These authors contributed equally: Wei-Chun Chou, Zengli Guo

Affiliations

1. Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Wei-Chun Chou, Zengli Guo, Hao Guo, Liang Chen, Ge Zhang, Kaixin Liang, Xianming Tan, Sara A. Gibson, Elena Rampanelli, Yan Wang, Stephanie A. Montgomery, W. June Brickey, Meng Deng, Leslie Freeman, Song Zhang, Maureen A. Su, Yisong Y. Wan & Jenny P.-Y. Ting

2. Department of Genetics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Wei-Chun Chou, Hao Guo, Elena Rampanelli, Leslie Freeman & Jenny P.-Y. Ting

3. Department of Microbiology and Immunology, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Zengli Guo, Liang Chen, Ge Zhang, W. June Brickey, Song Zhang, Yisong Y. Wan & Jenny P.-Y. Ting

4. Oral and Craniofacial Biomedicine Program, School of Dentistry, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Kaixin Liang, Meng Deng & Jenny P.-Y. Ting

5. Department of Biochemistry and Biophysics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Ling Xie & Xian Chen

6. Department of Pathology and Laboratory Medicine, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Stephanie A. Montgomery

7. Department of Pediatrics, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

Maureen A. Su

Authors

1. Wei-Chun Chou

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Zengli Guo

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Hao Guo

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Liang Chen

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Ge Zhang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Kaixin Liang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Ling Xie

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Xianming Tan

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. Sara A. Gibson

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. Elena Rampanelli

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

11. Yan Wang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

12. Stephanie A. Montgomery

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

13. W. June Brickey

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

14. Meng Deng

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

15. Leslie Freeman

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

16. Song Zhang

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

17. Maureen A. Su

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

18. Xian Chen

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

19. Yisong Y. Wan

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

20. Jenny P.-Y. Ting

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Corresponding authors

Correspondence to [Yisong Y. Wan](#) or [Jenny P.-Y. Ting](#).

Rights and permissions

[Reprints and Permissions](#)

About this article



Check for
updates

Cite this article

Chou, WC., Guo, Z., Guo, H. *et al.* Author Correction: AIM2 in regulatory T cells restrains autoimmune diseases. *Nature* **592**, E29 (2021).
<https://doi.org/10.1038/s41586-021-03490-7>

[Download citation](#)

- Published: 14 April 2021
- Issue Date: 29 April 2021
- DOI: <https://doi.org/10.1038/s41586-021-03490-7>

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Download PDF](#)

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03490-7>

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

Addendum: Active superelasticity in three-dimensional epithelia of controlled shape
[Download PDF](#)

- Addendum
- [Published: 12 April 2021](#)

Addendum: Active superelasticity in three-dimensional epithelia of controlled shape

- [Ernest Latorre](#)^{1,2},
- [Sohan Kale](#)²,
- [Laura Casares](#)¹,
- [Manuel Gómez-González](#) ORCID: [orcid.org/0000-0002-5091-0411](#)¹,
- [Marina Uroz](#)¹,
- [Léo Valon](#)¹,
- [Roshna V. Nair](#)³,
- [Elena Garreta](#)¹,
- [Nuria Montserrat](#)^{1,4},
- [Aránzazu del Campo](#) ORCID: [orcid.org/0000-0001-5725-2135](#)^{3,5},
- [Benoit Ladoux](#) ORCID: [orcid.org/0000-0003-2086-1556](#)^{6,7},
- [Marino Arroyo](#) ORCID: [orcid.org/0000-0003-1647-940X](#)^{1,2} &
- [Xavier Trepat](#) ORCID: [orcid.org/0000-0002-7621-5214](#)^{1,4,8,9}

[Nature](#) volume 592, page E30(2021) [Cite this article](#)

- 847 Accesses
- 4 Altmetric
- [Metrics details](#)

Subjects

- [Biological physics](#)
- [Biomaterials – cells](#)
- [Biophysics](#)

- [Cytoskeleton](#)
- [Mechanical engineering](#)

The [Original Article](#) was published on 31 October 2018

[Download PDF](#)

Addendum to: *Nature* <https://doi.org/10.1038/s41586-018-0671-4> Published online 31 October 2018

In our original Article, we developed a phototriggerable derivative of the actin disruptor cytochalasin D, denoted photo cytoD or Nvoc-CytoD, to show that a local perturbation of the actin cortex increases cell area (Extended Data Fig. 8 of the original Article). Through direct visualization, we showed that the F-actin cytoskeleton of cells that were incubated with Nvoc-CytoD (50 μ M) was rapidly disrupted upon light exposure (Extended Data Fig. 8c). Cells incubated with Nvoc-CytoD without light exposure, and cells exposed to light without Nvoc-CytoD, did not show changes in their cytoskeleton or in cell area (Extended Data Fig. 8a–d). A detailed characterization of the synthesis and chemical validation of Nvoc-CytoD was not included in our original Article, and we would like to provide it here.

Nvoc-CytoD includes a nitroveratryloxycarbonyl (Nvoc) photoremovable protecting group at the hydroxyl group at C7 of CytoD (Fig. 1a of this Addendum), which is relevant for binding to actin. The attachment of the photoremovable protecting group is expected to inhibit binding of Nvoc-CytoD to actin temporarily, and enable light-dependent delivery of the active drug CytoD to living cells.

Fig. 1: The structure of Nvoc-CytoD and the binding of CytoD to actin.

 **figure1**

a, Structure of Nvoc-CytoD and its photochemical activation reaction. **b**, Expanded view of CytoD (cyan sticks) bound to actin monomer (grey). The red dotted lines depict hydrogen-bonding interactions. This figure is recreated from Protein Data Bank (PDB) ID 3EKU (ref. [1](#)) and is rendered using PyMOL. **c**, The hydrogen-bonding interactions involved in CytoD–actin binding are shown as red dotted lines. Note that the hydroxyl at C7 interacts with Ile136(NH) and V134(CO).

[Full size image](#)

The [Supplementary Information](#) to this Addendum contains information on the molecular design of Nvoc-CytoD, as well as its synthesis, structural characterization and photochemical properties.

Supplementary information is available in the online version of this Amendment.

References

1. 1.

Nair, U. B., Joel, P. B., Wan, Q., Lowey, S., Rould, M. A. & Trybus, K. M. Crystal structures of monomeric actin bound to cytochalasin D. *J. Mol. Biol.* **384**, 848–864 (2008).

[CAS Article](#) [Google Scholar](#)

[Download references](#)

Author information

Affiliations

1. Institute for Bioengineering of Catalonia (IBEC), The Barcelona Institute for Science and Technology (BIST), Barcelona, Spain

Ernest Latorre, Laura Casares, Manuel Gómez-González, Marina Uroz, Léo Valon, Elena Garreta, Nuria Montserrat, Marino Arroyo & Xavier Trepat

2. LaCàN, Universitat Politècnica de Catalunya-BarcelonaTech, Barcelona, Spain

Ernest Latorre, Sohan Kale & Marino Arroyo

3. INM-Leibniz Institut für Neue Materialien, Saarbrücken, Germany

Roshna V. Nair & Aránzazu del Campo

4. Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina, Barcelona, Spain

Nuria Montserrat & Xavier Trepat

5. Chemistry Department, Saarland University, Saarbrücken, Germany

Aránzazu del Campo

6. CNRS UMR 7592, Institut Jacques Monod (IJM), Université Paris Diderot, Paris, France

Benoit Ladoux

7. Mechanobiology Institute (MBI), National University of Singapore, Singapore, Singapore

Benoit Ladoux

8. Unitat de Biofísica i Bioenginyeria, Universitat de Barcelona, Barcelona, Spain

Xavier Trepot

9. Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain

Xavier Trepot

Authors

1. Ernest Latorre

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

2. Sohan Kale

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

3. Laura Casares

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

4. Manuel Gómez-González

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

5. Marina Uroz

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

6. Léo Valon

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

7. Roshna V. Nair

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

8. Elena Garreta

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

9. Nuria Montserrat

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

10. Aránzazu del Campo

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

11. Benoit Ladoux

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

12. Marino Arroyo

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

13. Xavier Trepaut

[View author publications](#)

You can also search for this author in [PubMed](#) [Google Scholar](#)

Corresponding authors

Correspondence to [Marino Arroyo](#) or [Xavier Trepaut](#).

Supplementary information

[Supplementary Information](#)

Rights and permissions

[Reprints and Permissions](#)

About this article



Check for
updates

Cite this article

Latorre, E., Kale, S., Casares, L. *et al.* Addendum: Active superelasticity in three-dimensional epithelia of controlled shape. *Nature* **592**, E30 (2021).
<https://doi.org/10.1038/s41586-021-03281-0>

[Download citation](#)

- Published: 12 April 2021
- Issue Date: 29 April 2021
- DOI: <https://doi.org/10.1038/s41586-021-03281-0>

Comments

By submitting a comment you agree to abide by our [Terms](#) and [Community Guidelines](#). If you find something abusive or that does not comply with our terms or guidelines please flag it as inappropriate.

[Download PDF](#)

This article was downloaded by **calibre** from <https://www.nature.com/articles/s41586-021-03281-0>

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