EMBL-EBI

Highlights 2025

ebi.ac.uk

EMBL's European Bioinformatics Institute (EMBL-EBI)
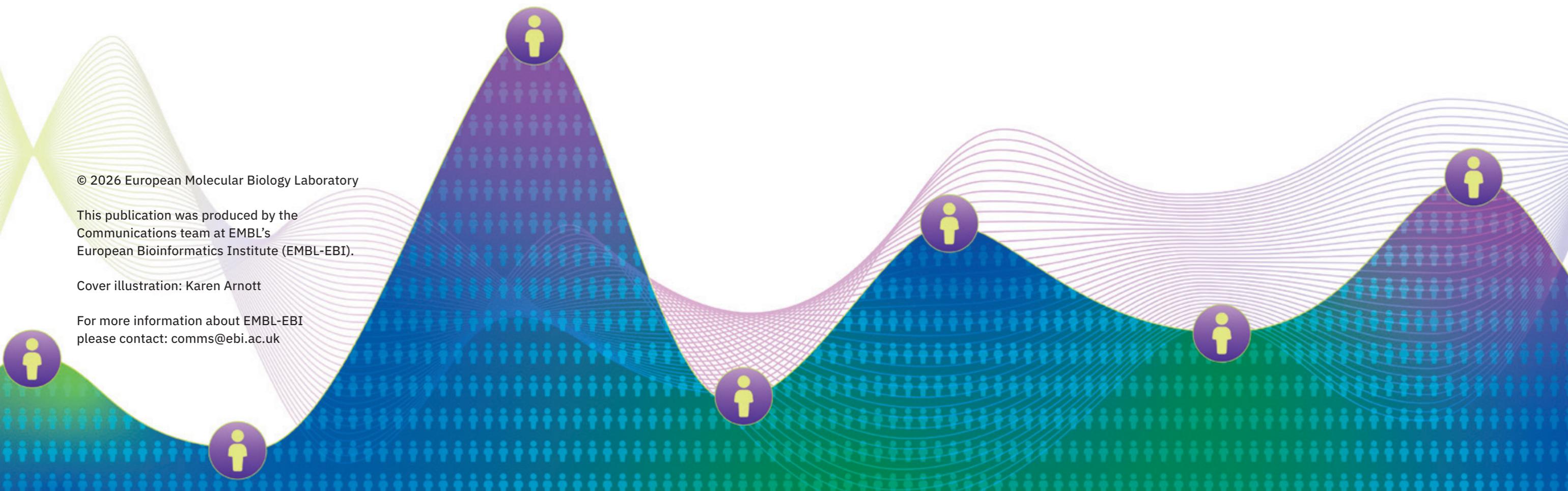
Contents

Foreword

Advances in AI are rapidly reshaping the way science is conducted. EMBL-EBI is proud to provide high-quality data, tools and expertise to help navigate this transformation. We make high-quality biomolecular data openly available and reusable, helping researchers apply AI more effectively for biological discovery, genomic medicine, and much more.

In 2025, new platforms and methods developed at EMBL-EBI helped researchers track antimicrobial resistance, explore rare disease at scale, and probe the mechanisms behind aggressive cancers. Our researchers combined large-scale sequence resources with deep learning to forecast disease risk and turn genomic data into timely intelligence for outbreak response.

We continued developing our vital technical infrastructure and expanding global training to support effective and responsible use of AI and bioinformatics in the life sciences. These efforts ensure our data resources remain robust, secure, and scalable for our global user community.

Our progress this year has been made possible through the commitment of our funders, including EMBL's Member States, UK Research and Innovation, Wellcome, the European Commission and the National Institutes of Health. Their investment allows us to maintain the open data infrastructure that scientists worldwide depend on.

EMBL-EBI Senior Management Team

Jo McEntyre
Interim Director
of EMBL-EBI

Julio Saez-Rodriguez
Head of Research
at EMBL-EBI

Rachel Curran
Head of EMBL-EBI
Administration & Operations
and EMBL's Deputy Chief
Operating Officer

Who we are

EMBL's European Bioinformatics Institute (EMBL-EBI) is the world's leading source of public biomolecular data. We enable life science research and its translation to medicine, agriculture, industry and society by providing biological data, tools and knowledge.

We are one of the six sites of the European Molecular Biology Laboratory (EMBL), Europe's life sciences laboratory. EMBL integrates fundamental research, advanced scientific services, and excellence in training to tackle complex biological and societal challenges. EMBL's ambitious scientific programme, Molecules to Ecosystems, seeks to understand life in context at a molecular level.

Our missions

- To freely provide data and bioinformatics services to the scientific community in ways that promote scientific progress.

- To contribute to the advancement of biology through investigator-driven research in bioinformatics.

- To provide bioinformatics training to scientists at all levels.

- To disseminate cutting-edge technologies to industry.

- To support, as an ELIXIR Node, the coordination of biomolecular data provision in Europe.

Our vision

To benefit humankind by advancing scientific discovery and impact through bioinformatics.
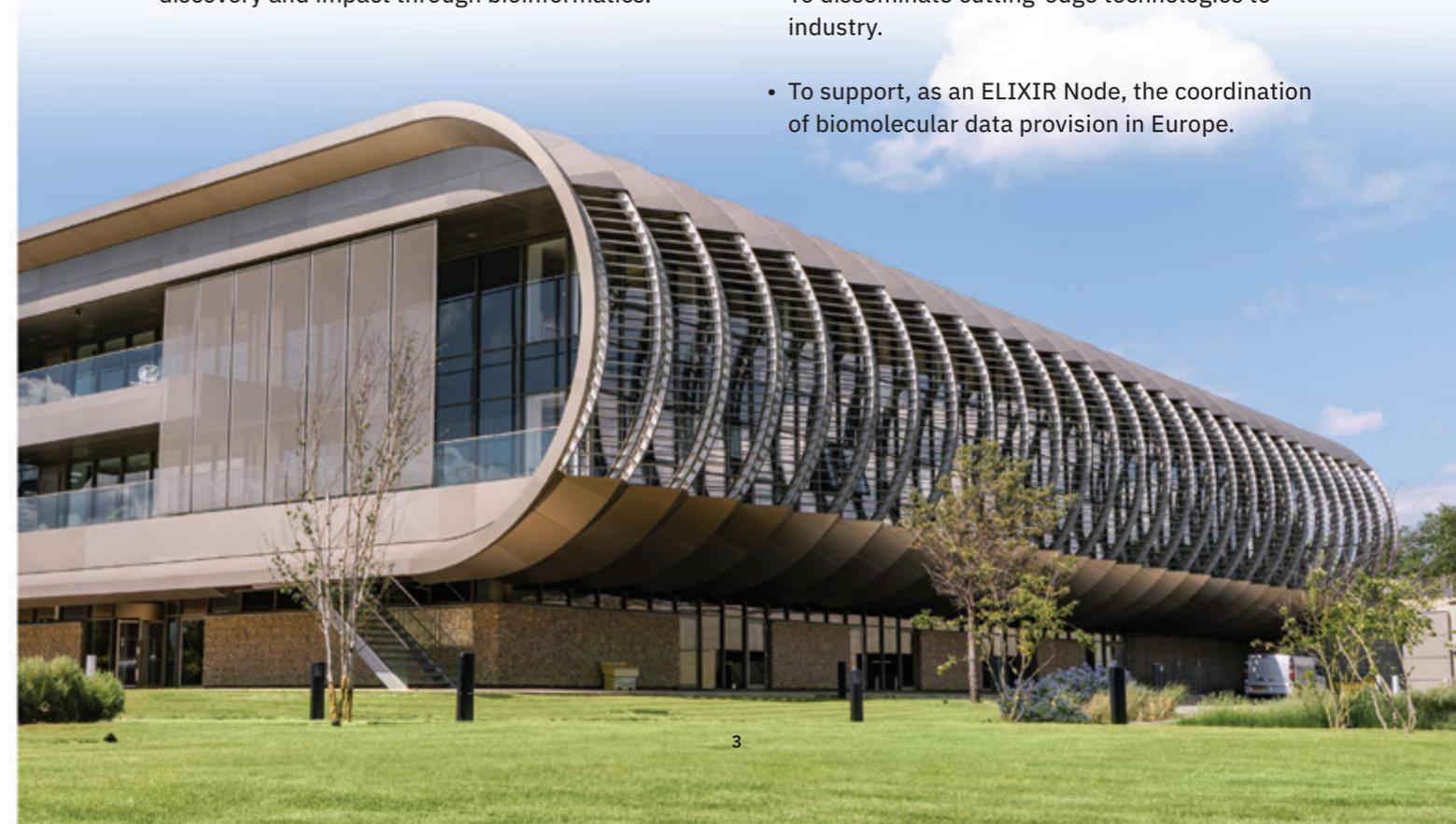
2025 in numbers

HR

646 established members of personnel*

from **65** countries

*in Full Time Equivalent includes staff, fellows, supernumeraries and trainees during 2025 calendar year

DATA RESOURCES

121 million requests to our data resources on an average day

from **127** million unique IP addresses annually*

>123,200 scientific publications enabled by EMBL-EBI data resources

*EMBL-EBI has seen a large increase in IP addresses in 2025, reflecting the increased use of AI-powered automated requests

RESEARCH

259 journal papers published

INDUSTRY

1,739

to our **10** Industry Programme workshops

attendees

GRANTS

173 active grants

146 collaborative grants with researchers

in **58** countries

from **860** institutes

TRAINING

430,223 users accessed our online training

85,155 webinar views

18 live courses

reaching **524** delegates

PUBLIC ENGAGEMENT

1,048 participants

18 public engagement events

26 staff and students

3 events organised with our community partners

Data resources

EMBL-EBI manages the world's most comprehensive suite of open data resources for the life sciences. Our 30+ data resources and dozens of tools span genetics, genomics, proteins, bioimaging, chemistry, literature and more.

A gateway to antimicrobial resistance data

Antimicrobial resistance (AMR) is one of the most pressing challenges in modern medicine, threatening to render life-saving medical treatments ineffective. To help address this, the global scientific community needs robust, standardised data to understand exactly how genetic changes in bacteria lead to drug resistance.

In response, EMBL-EBI launched the AMR Portal, a centralised hub designed to bring together bacterial genetics and experimental resistance data. Developed in collaboration with Imperial College London, this new resource connects bacterial genome sequences, phenotypic data, and functional annotations. By integrating data from many of EMBL-EBI's open data resources the portal allows scientists to trace resistance mutations to better understand how antimicrobial resistance evolves and spreads.

By providing high-quality, uniformly structured data at scale, the AMR Portal also serves as a foundation for training AI models capable of predicting resistance directly from DNA sequences. This innovation will help support the monitoring of antimicrobial resistance trends worldwide.

"The AMR Portal from EMBL-EBI provides an unprecedented, openly curated collection of antimicrobial resistance phenotype data. It makes it easier for researchers to understand the relationship between genotype and antimicrobial resistance, and will support the global scientific effort to combat AMR."
Kate Baker, Professor in the Department of Genetics at the University of Cambridge

UK-wide federated analysis of mental health data

Mental health conditions are complex, with underlying biological factors that vary significantly between individuals. One major barrier to understanding these conditions is the challenge of accessing and connecting the many different data types needed to understand mental health at a biological level.

The Open Psychiatry Project, a new initiative funded by the UKRI Medical Research Council and the National Institute for Health and Care Research, aims to address this. The project creates a secure, federated analysis platform, allowing researchers to study sensitive mental health data across multiple secure environments.

By integrating the resulting analyses into the Open Targets Platform, the project will link genetic evidence and real-world patient data to potential drug targets. This will enable scientists to prioritise the most promising biological leads, accelerating the development of new, more effective treatments for mental health conditions tailored to an individual's needs.

"This is the first major initiative to bring together mental health data for multiple stakeholders in an accessible way, and analysing this information collectively is a key step towards finding new targets for drug development. The findings from this project will be integrated into the Open Targets Platform, where they will enhance existing publicly available data to help identify and prioritise the most promising therapeutic targets."
Ellen McDonagh, Translational Informatics Director at Open Targets and EMBL-EBI co-lead for this project

Bridging rare disease research and clinical care

Rare diseases are individually scarce but collectively common, affecting over 300 million people worldwide. Because a clinician may only ever see one patient in their lifetime with a particular rare disease, access to relevant data at a global scale is necessary to enable faster diagnosis and develop appropriate treatment plans.

To support rare disease research and improved patient care, EMBL-EBI's DECIPHER platform continues to be a vital resource, linking genetic variants to clinical phenotypes. A major 2025 update introduced better support for interpreting the impact of variants in mitochondrial DNA that cause mitochondrial disorders. This is a group of disorders that can cause severe symptoms that affect multiple body parts and progress over time

In 2025, EMBL-EBI supported efforts to make data from the Solve-RD project standardised and available in the European Genome-phenome Archive (EGA). The Solve-RD project is harmonising genomic and clinical data from partners across Europe, making previously fragmented rare disease datasets comparable and easier to analyse. This project enables researchers to work across cohorts and uncover genetic causes that would be impossible to detect in isolated datasets.

EMBL-EBI is also continuing to support the development of a federated data infrastructure that enables clinicians and researchers to access and analyse research study participant data across borders. The Federated European Genome-phenome Archive (FEGA), allows researchers to discover and access sensitive data stored securely within national borders. In 2025, this network expanded, adding Switzerland as the newest node.

"No single clinic has enough cases to quickly solve the puzzle of a rare disease diagnosis alone. By focusing on federating research data infrastructures and supporting interoperability with clinical resources such as DECIPHER, we are finally enabling the global collaboration needed to turn isolated data points into life-changing clinical care."
Mallory Freeberg, Human Genomics Team Leader at EMBL-EBI

DECIPHER IN 2025

>52,000 patient records
>208,000 phenotypes
>66,000 variants

shared by **>315** centres
(hospitals, universities, research institutes)

from **~35** countries

>4,300 citations in publications*

**278** citations in 2025

*since its inception in 2004

AI-driven protein annotations

The known protein universe is growing faster than researchers can classify new sequences. To keep up, EMBL-EBI is integrating AI into its core resources to analyse sequences at far greater scale and speed than manual methods allow.

As part of this effort, the InterPro team – which provides functional analysis of proteins – has introduced InterPro-N. This is a cutting-edge deep-learning system that can infer the function of protein sequences that previously lacked any meaningful classification. The update has increased our view of the protein world, delivering:

- 1.8 billion new protein annotations

- Over 90% coverage of the UniProt Knowledgebase (UniProtKB), filling long-standing gaps in domain classification

- Seamless integration of AI-generated insights alongside experimental data, clearly marked to help researchers distinguish between predicted and curated evidence

For the first time, the vast majority of proteins in the InterPro database now have a predicted function, family, or domain, accelerating discovery in fields ranging from drug design to evolutionary biology.
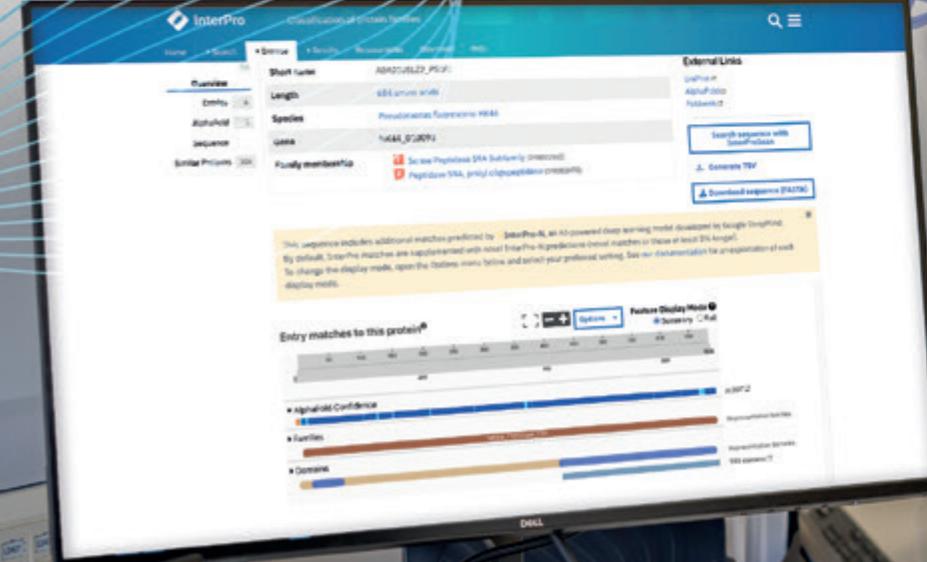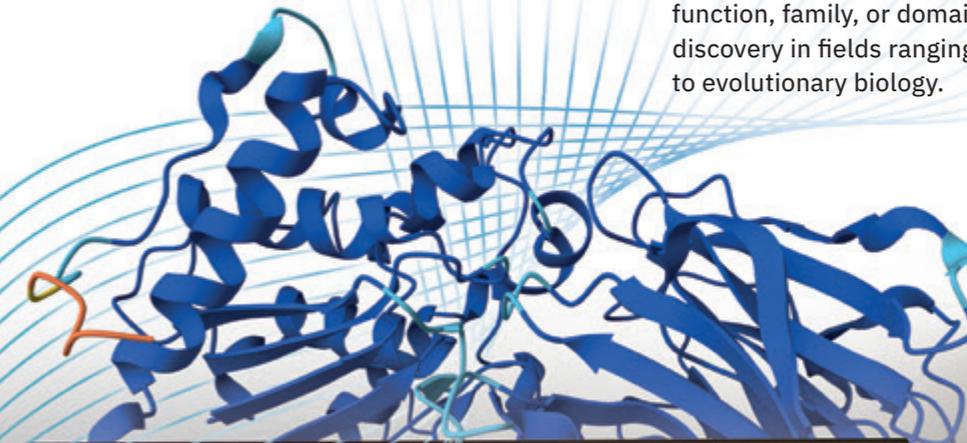
Building equitable biodata partnerships

Biological discovery is a global endeavour, yet participation in the biodata ecosystem remains unevenly distributed. Although open data resources are used by millions of scientists worldwide, analysis reveals that participation both in terms of data usage and sharing is heavily skewed towards high-income nations. Understanding the barriers facing researchers in under-represented regions is therefore essential to building an equitable biodata ecosystem.

In support of this effort, EMBL-EBI commissioned a report titled 'Towards global biodata partnerships', funded by Wellcome and produced by Research Consulting. The report analysed biodata usage patterns across the globe, revealing a direct correlation between data contribution and data usage. It found that the more a region contributes to open databases, the more its scientists use them.

Although one in two visits to EMBL-EBI resources originates from low- and middle-income countries (LMICs), scientists in these regions don't use data resources as intensively as scientists in higher-resource environments. This can be due to a range of factors, including infrastructure, policy and funding challenges. The report offers a roadmap for change, urging funders and major data hubs to move towards active, equitable engagement that enables scientists in low-resource settings to participate fully in the global biodata ecosystem.

"This report explores a number of challenges and opportunities that must be addressed if we are to improve the global biodata ecosystem."
Chiann-Mun Chen, Senior Research Manager for Discovery Research at Wellcome
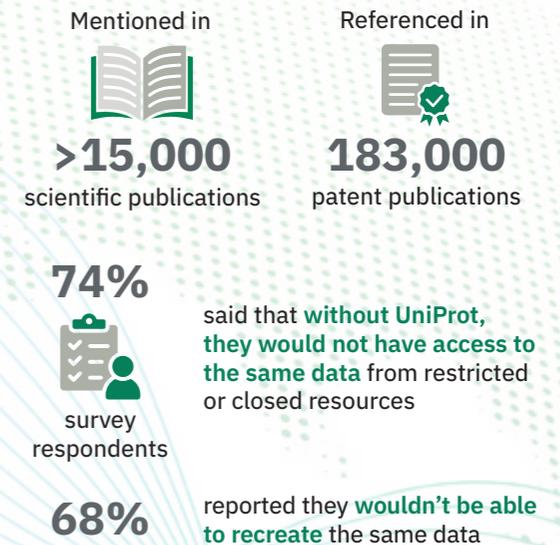
Measuring the value of UniProt

A comprehensive impact study conducted by CSIL, an Italian economic research institute, demonstrates that UniProt, the universal protein resource co-developed by EMBL-EBI, delivers substantial economic and scientific returns that far exceed its operational expenses.

The study estimates that UniProt generates annual gains per user ranging from €3,000 to €5,000, largely by saving researchers time they would otherwise spend recreating or reprocessing data. Beyond these efficiencies, the report highlights UniProt's essential role in driving innovation.

"At a time when many data resources face serious funding challenges, this case study sends a clear message: UniProt is a strategic, high-impact asset. Sustaining it is not just good for life science – it's an essential investment in innovation, economic growth, and human and planetary well-being."
Alan Bridge, Director of SIB's Swiss-Prot group and UniProt Principal Investigator

UNIPROT IN NUMBERS

Mentioned in
>15,000
scientific publications

Referenced in
183,000
patent publications

74%
survey respondents
said that **without UniProt, they would not have access to the same data** from restricted or closed resources

68%
reported they **wouldn't be able to recreate** the same data

Sequencing all life on Earth

We are living through a period of rapid biodiversity loss that scientists are describing as a sixth mass extinction. To preserve the biological diversity of our planet, the global scientific community has embarked on a biological moonshot, the Earth BioGenome Project (EBP). This ambitious initiative aims to sequence the genomes of all 1.67 million known eukaryotic species by 2035.

In 2025, the EBP released a plan for its next phase, aiming to scale up efforts tenfold. The goal is to sequence 3,000 new genomes every month, a pace necessary to capture an additional 150,000 species over the next four years. This acceleration will also prioritise species vital for ecosystem health, food security, and pandemic prevention.

Although the raw genome sequences generated by the project contain immense biological information, they hold little meaning until the genes and features within them are identified. EMBL-EBI researchers within the Ensembl project are leading the annotation of the incoming EBP genomes. The annotated genomes will be stored and made openly available in the new Ensembl platform.

"Annotation is what makes these data truly valuable; it allows researchers to understand which genes are present, what they do, and how species have evolved and adapted over time. By making these annotated genomes openly available, we're giving scientists around the world the tools they need to address urgent challenges in biodiversity, conservation, and more."
Fergal Martin, Eukaryotic Annotation Team Leader at EMBL-EBI

ALEENA MUSHTAQ
SENIOR ENSEMBL OUTREACH OFFICER

After completing her PhD in molecular nutrition at the Quadram Institute in Norwich, UK, Aleena joined EMBL-EBI as a Senior Ensembl Outreach Officer. Aleena helps scientists make the most of the Ensembl genome browser. Ensembl integrates gene, variation and regulatory data for all species, as well as genomics analysis tools, all within a user-friendly platform.

Ensembl's Outreach team delivers workshops and tailored training to scientists worldwide, ensuring the project's tools and data are accessible to a diverse research community. In 2025 alone, Aleena and her colleagues delivered 95 training courses in 16 countries.

"An exciting recent project is adapting training materials for different levels and languages. With the help of AI, we can simplify content and translate our online resources, making them more accessible to a global audience," said Aleena.

DALJIT NIJJAR
LITERATURE SERVICES COORDINATOR

With a background spanning digital marketing and the charity sector, Dal brings a user-centred perspective to EMBL-EBI. Having spent his career modernising systems and optimising digital services to better serve the public, he now applies his expertise in digital transformation to the world of open science as a Product Manager and Coordinator for Europe PMC.

Dal leads a multidisciplinary team focused on reducing the cognitive load associated with literature search. He champions the use of agile methodologies and explores how AI can shift search from simple keyword matching to understanding user intent. His work ensures that millions of biomedical research articles remain discoverable and accessible to researchers worldwide.

"I enjoy the challenge of working in open science; knowing we are making the platform better for the scientific community is very motivating," said Dal.

Research

EMBL-EBI's research groups use computational methods and artificial intelligence to make sense of vast, complex datasets. Our researchers collaborate with experimental scientists to build practical solutions for real-world challenges linked to human health and the environment.

AI model forecasts disease risk

The ability to accurately forecast an individual's future health trajectory is a long-standing goal in genomic medicine. While clinicians can estimate risks for some health conditions, predicting the onset of multiple, interconnected diseases remains a significant challenge.

Researchers at EMBL-EBI, in collaboration with the German Cancer Research Center (DKFZ) and the University of Copenhagen, have developed a pioneering AI model for disease prediction. The model was trained on anonymised health records from more than 400,000 participants in the UK Biobank, and validated using data from 1.9 million patients in the Danish National Patient Registry.

The model uses generative AI, similar to the technology behind large language models, to analyse the sequence and timing of medical events. By learning the underlying grammar of human health, the model can estimate the probability of over 1,000 different diseases occurring over a decade in advance.

Like weather forecasts, this new AI model provides probabilities, not certainties. It doesn't predict exactly what will happen to an individual, but it offers well-calibrated estimates of how likely certain conditions are to occur over a given period. By providing these risk estimates, the model could help healthcare systems to shift from reactive treatment to proactive prevention, allowing for earlier interventions tailored to a patient's medical history.

"Medical events often follow predictable patterns. Our AI model learns those patterns and can forecast future health outcomes. It gives us a way to explore what might happen based on a person's medical history and other key factors. Crucially, this is not a certainty, but an estimate of the potential risks."
Tom Fitzgerald, Staff Scientist at EMBL-EBI

Improving outbreak response for NHS Scotland

Bacterial infections can spread rapidly in hospitals where vulnerable patients are treated in close proximity. To stop these outbreaks, clinical teams need to know how an infection is moving between patients. However, analysing bacterial DNA to spot these transmission chains is often a complex, time-consuming process requiring specialist bioinformatics expertise.

Researchers at EMBL-EBI developed the Population analysis PIPEline (PopPIPE), a user-friendly genomic analysis tool designed for clinical use. In collaboration with the University of St Andrews, the tool was deployed across six hospital wards in Scotland to investigate antibiotic-resistant infections. By automating the analysis, PopPIPE allowed clinicians to confirm suspected links between patients and reveal hidden transmission clusters that manual methods had missed.

Following this success in hospitals, the tool is now being piloted in wider national workflows, for example tracking gonorrhoea strains between cities. By turning complex genomic data into actionable insights, PopPIPE is helping public health teams stay ahead of outbreaks.

"PopPIPE significantly simplifies genomic epidemiology. It speeds up our outbreak response and improves accuracy, helping us manage infections more effectively. It gives us a clearer picture of how infections are spreading and where we need to focus control measures."
Martin McHugh, Principal Clinical Scientist at NHS Lothian

New mutation mechanism driving aggressive bone cancer

Osteosarcoma is a highly aggressive bone cancer that primarily affects children and young adults. It presents a significant clinical challenge due to the complexity of its genome, which obscures the underlying mechanisms driving the disease and slows the development of treatments.

A collaborative study by researchers from EMBL-EBI, University College London (UCL), the Royal National Orthopaedic Hospital, and Genomics England has shed new light on these underlying mechanisms. By applying advanced long-read sequencing to different regions within individual tumours, the team identified a previously unrecognised mutational process, termed loss–translocation–amplification (LTA) chromothripsis.

The study showed that chromosomes affected by LTA chromothripsis remain genomically unstable, continuing to accumulate structural abnormalities as the cancer evolves. This ongoing instability helps explain tumour adaptability and its ability to resist treatment. Importantly, an analysis of more than 5,300 cancer genomes revealed that such instability is present in several other cancer types.

The study also identified a new prognostic biomarker for osteosarcoma: loss of heterozygosity (LOH). LOH occurs when one copy of a genomic region is lost. In osteosarcoma, a high degree of LOH across the genome predicts a lower survival probability.

"This biomarker (LOH) could help us identify patients who are unlikely to benefit from treatment which can have very unpleasant effects and which patients find difficult to tolerate. This is invaluable for providing patients with more tailored treatments and help spare unnecessary effects of toxic therapies."
Adrienne Flanagan, Professor at University College London and Consultant Histopathologist at the Royal National Orthopaedic Hospital

LEONIE JOHANNA LORENZ
PREDOCTORAL FELLOW

After studying biomathematics in Germany, Leonie joined EMBL-EBI as a Predoctoral Fellow. She combines genomics with equations to predict how bacteria adapt to new pressures. Focusing on the bacterium *Streptococcus pneumoniae*, Leonie uses mathematical models to untangle how vaccination alters bacterial diversity and drives the emergence of new strains.

Beyond her research, Leonie is passionate about making mathematical modeling accessible to the wider community and has developed an online learning resource for the EMBL-EBI training catalogue. She values the institute's interdisciplinary culture, where methodological overlaps allow for unique problem-solving across different biological fields.

"I can chat with people working on cancer signalling or human genetics who face similar data problems to me, even though I work on bacteria. We get to solve these issues together," said Leonie.

BERTHA RUEDA
POSTDOCTORAL FELLOW

After completing her PhD in cancer genomics at the National Autonomous University of Mexico, Bertha joined EMBL-EBI as a Postdoctoral Fellow in the Cortes-Ciriano group. She works on the SAMBAI project, part of Cancer Grand Challenges, where she investigates how social, molecular and ancestry-related factors shape cancer outcomes in populations with African ancestry.

"We must explore underserved populations to create fairer medicine. I hope to be part of this change, helping to reduce inequalities in access to health, knowledge and science, and having a direct impact on patient's lives."

Outside her research, Bertha's world comes alive with salsa, painting and short films.

Training and industry partnerships

EMBL-EBI delivers world-leading bioinformatics training for all levels, including predoctoral degrees awarded by the University of Cambridge. Our approach to democratising global life science research and building bioinformatics capacity is reflected in our training portfolio.

AlphaFold education summit

The AlphaFold AI, developed by Google DeepMind, predicts the structure of millions of proteins. It reshaped structural biology, enabling new discoveries in fields ranging from drug development to environmental science. In 2021, EMBL-EBI and Google DeepMind launched the AlphaFold Database, which made all AlphaFold predictions freely accessible worldwide, and in 2024, they released a free, online AlphaFold training course.

Despite this, a significant adoption gap remains. Researchers and educators in low-resource settings face many barriers including limited access to computing, low-bandwidth internet, language constraints and a shortage of trained instructors.

To enable more people to use AlphaFold efficiently and responsibly, Google DeepMind and EMBL-EBI convened the AlphaFold Education Summit in January 2025. Bringing together participants from 14 countries, with a strong focus on low-resource settings, the summit centred on a train-the-trainer model, equipping local experts to deliver AlphaFold training within their own communities.

Follow-on work with participants in the summit has nucleated a community of AlphaFold trainers, who are delivering their own workshops, sharing materials and case studies, and translating training materials into local languages, now supported by an online community space.

Meet Ada, your AI training companion

EMBL-EBI offers a comprehensive library of bioinformatics training. With over 300 on-demand courses and materials available, finding the right resource is not always straightforward.

The EMBL-EBI Training team has developed Ada, a new AI-powered assistant designed to transform how learners interact with the institute's educational archive. Named after the pioneering mathematician Ada Lovelace, the tool uses a large language model (LLM) to offer a conversational alternative to traditional search bars. Instead of searching by keywords, users can now ask natural language questions and receive guided, context-aware recommendations drawn directly from EMBL-EBI's training catalogue.

Since launch, the chatbot has been accessed by users from 85 countries. The tool provides instant, round-the-clock support, ensuring learners can find the bioinformatics training they need, whenever they need it.

"We saw an opportunity to use AI to fundamentally improve how users access our training. Taking that initial concept and engineering it into a practical, reliable tool required a proactive and agile approach. Our AI companion now actively helps researchers navigate our training resources."
Prakash Singh Gaur, Web Developer at EMBL-EBI

AJAY MISHRA
SENIOR SCIENTIFIC TRAINING MANAGER

For Ajay Mishra, learning is not a phase but a lifelong journey. Ajay supports EMBL-EBI's global bioinformatics training by coordinating the institute's e-learning offering.

Originally from India, Ajay has a background in biochemistry research. He brings a learner-first mindset and works with subject matter experts from across EMBL-EBI to build training that is findable, accessible, interoperable and reusable (FAIR).

Together with his colleagues in the EMBL-EBI Training team, Ajay helps broaden access to bioinformatics skills, empower diverse scientific communities, and ensure that training keeps pace with rapidly evolving life sciences data, tools and the latest AI advancements, including EMBL-EBI's first training chatbot, Ada.

"We're exploring ways to improve EMBL-EBI's on-demand training to make it more interactive and enrich the overall learning experience for our global audience," said Ajay.

Industry Programme and benefits

The EMBL-EBI Industry Programme is a pre-competitive, subscription-based programme for companies that use EMBL-EBI data resources and tools as a core part of their research and development. It offers a structured forum where member companies from pharma, agri-food, nutrition and healthcare can engage directly with EMBL-EBI scientists and discuss shared data challenges.

Member benefits include:

- Knowledge exchange and a chance to shape data resources, standards and policies

- Connecting with EMBL-EBI experts and peers at other companies in a trusted environment

- Co-creation of bespoke bioinformatics workshops and quarterly strategy meetings

- Opportunities to benchmark, learn from peers and reduce duplication

By bringing together industry leaders and EMBL-EBI specialists, the programme supports innovation and helps companies integrate open biodata into their research pipelines. The programme welcomes new members who wish to contribute to and benefit from this community.

JANA BROECKER
SENIOR SCIENTIFIC OFFICER

With a background spanning biotechnology, biophysics, and structural biology, Jana has lived and worked across Germany, Canada, and the UK. She joined EMBL-EBI from the pharmaceutical industry, bringing a wealth of experience in leading cross-functional collaborations.

As a Senior Scientific Officer in the Industry Partnerships team, Jana builds vital bridges between academic research and commercial application. She facilitates knowledge exchange and drives public-private partnerships such as the new Agri-Tech Partnership helping disparate groups tackle shared data challenges and accelerate innovation.

"This role combines everything I care about: connecting people, enabling collaboration, and helping turn science into real-world solutions," said Jana.

Culture and infrastructure

To fulfil its missions, EMBL-EBI relies on a deeply collaborative culture, a diverse workforce and robust technical infrastructure. Engaging with our diverse network of stakeholders, from funders, policymakers and collaborators is essential for sustaining EMBL-EBI's missions.

A hub for collaboration and innovation

The official opening of EMBL-EBI's Thornton Building took place in the presence of key stakeholders including the UK Minister for Science, Research and Innovation, Lord Vallance of Balham.

Named in honour of Professor Dame Janet Thornton, the institute's Director Emeritus, the new building was funded by the UK Government via UK Research and Innovation (UKRI), the Biotechnology and Biological Sciences Research Council (BBSRC), and Wellcome.

The facility now houses the Open Targets public-private partnership, the EMBL-EBI Industry Programme, the institute's research groups and technical teams, providing a dedicated space for translation and partnership.

"EMBL-EBI's work speeds up scientific discovery and enables millions of scientists, engineers, and innovators to develop solutions for urgent challenges such as infectious disease, food insecurity, and the biodiversity crisis. The United Kingdom is proud to host and support this unique research infrastructure."
Lord Vallance of Balham, Minister for Science, Research and Innovation

Supporting science strategies

Around the world, policymakers are placing more emphasis on the value of open data.

- The UK Government's new AI for Science strategy recognised the vital role that high-quality, foundational datasets play in enabling AI development, citing EMBL-EBI for the strength of our data foundations.

- The European Commission's new Strategy for Artificial Intelligence in Science highlighted the EMBL-Google DeepMind AlphaFold Database collaboration as a blueprint for public-private partnership for driving AI innovation.

- The World Health Organisation adopted its Pandemic Agreement in May 2025, incorporating EMBL-EBI recommendations to prioritise open access to data for future outbreak response.

These developments make open science a policy priority, reinforcing the importance of sustained investment in open data that remains accessible when the world needs it most.

EMBL-EBI discusses artificial general intelligence

Following the story of AlphaFold, The Thinking Game documentary explores the journey of Google DeepMind and its co-founder, Demis Hassabis, in their quest to develop artificial general intelligence. It features guest appearances from EMBL-EBI's Director Emeritus Janet Thornton and EMBL's Interim Director General Ewan Birney.

To celebrate the film's release, we transformed our Kendrew Lecture Theatre into a cinema for the afternoon. The event concluded with a panel discussion on how AI is reshaping the life sciences. The film has since become a global success, reaching hundreds of millions of views on YouTube.

Data centre migration

In 2025, the IT & Technical Services department undertook a major migration of critical IT infrastructure from the Hinxton campus to two state-of-the-art data centres in Harlow and Enfield. The migration encompassed fallback web services, storage, database backups, desktop network services, and components of the File Replication archive.

As part of this initiative, EMBL-EBI joined the UK government's Crown Hosting Data Centres framework, delivering both cost efficiencies and assured standards for security, performance, connectivity, and environmental sustainability. The new facilities offer a substantial long-term reduction in power costs, with the Harlow site utilising 100% certified renewable energy and Enfield drawing its power from a neighbouring energy-from-waste facility.

SIMONE MATIAS VEGH
MAJOR INITIATIVES PROGRAMME MANAGER

Born in Brazil, Simone Vegh is a Major Initiatives Programme Manager at EMBL-EBI, where she coordinates long-term, large-scale funding programmes that underpin the institute's technical infrastructure. Working closely with key funders such as UK Research and Innovation, Simone supports the delivery of the Data Infrastructure Programme, a ten-year investment that enables EMBL-EBI to meet the evolving needs of the global life science community.

This funding supports critical infrastructure that underlies all EMBL-EBI data resources, from biodata to bioimaging, and ensures that open, high-quality biological data remain accessible worldwide. Simone also developed EMBL-EBI's current risk management framework, strengthening governance, resilience, and cybersecurity.

"The success of these programmes reflects the trust and commitment of EMBL-EBI's funders. Their sustained support makes it possible for the institute to deliver global scientific impact and accelerate discovery for researchers everywhere," said Simone.

MILA CONIFF
SENIOR SERVICE DESK OFFICER

Originally from Kazakhstan, Mila Coniff has a background in computer systems. She joined EMBL-EBI on a graduate position and has since progressed to Senior Service Desk Officer, becoming an essential member of the institute's IT support team.

Mila acts as the first point of contact for technical issues, ensuring that over 700 staff and students remain connected and operational. Beyond troubleshooting, her role involves onboarding new starters and managing complex queries, requiring a unique blend of technical expertise and soft skills like active listening and empathy.

"My team is incredibly collaborative and has taught me everything I know. We all help each other and there's no such thing as a silly question," said Mila.

Explore the World of Molecular Biology

The EMBL-EBI Thornton Building welcomed a vibrant addition to its shared space: the World of Molecular Biology exhibit. This interactive display offers visitors and staff a fascinating glimpse into the converging worlds of bioimaging, bioinformatics, and AI.

Adapted using select exhibits from the World of Molecular Biology exhibition at the EMBL Imaging Centre in Heidelberg, the installation is designed to introduce non-scientific audiences to the inner workings of life at all scales. Through stunning imagery, animations, and interactive elements, it demystifies complex technologies and concepts. The exhibit serves as an entry point for our visitors, illustrating how the institute's research and services shapes our understanding of the living world.

"Our public engagement strategy is built on the belief that science belongs to everyone. This exhibit brings that mission to life by transforming the invisible world of data into something tangible and visually engaging. It creates a welcoming space where anyone regardless of their background can step inside, spark their curiosity, and see how molecular biology connects to their own lives."
Briony Jackson, Head of Science Education and Public Engagement at EMBL

SAM SHINGLES
PUBLIC ENGAGEMENT OFFICER

With a background in planetary science, Public Engagement Officer Sam Shingles brings a unique perspective to EMBL-EBI. Moving from studying the stars to molecular biology, Sam focuses on demystifying complex concepts for non-specialist audiences and embedding a culture of open dialogue across the institute.

Sam champions a public engagement strategy based on two-way conversation. He collaborates closely with local community partners to co-develop accessible activities. This helps community groups navigate complex topics such as coding and AI. His work ensures that EMBL's commitment to open science extends beyond the lab, by supporting the public to understand the data-driven technologies that shape their lives.

"Hearing public questions, concerns and challenges is vital for scientists. It reminds them that their work is deeply personal to people and helps them understand the communities they serve," said Sam.

Building an inclusive community

EMBL-EBI's focus on strengthening its community continued in 2025, with several new initiatives bringing colleagues together to socialise across teams and departments, mark important moments and support staff well-being.

- **Take 5 outside:** A campaign that encourages staff to take regular breaks outside, and capture and share a photo of the surrounding nature.

- **Campus expansion:** The expansion of the Wellcome Genome Campus, where EMBL-EBI is located, is creating new opportunities for connection across academia, industry, and the wider community. Over the coming years, the campus will grow significantly, welcoming more research organisations and tech companies. This growth will strengthen EMBL-EBI's links with partners working across genomics, biodata, health, AI and more.

- **John Jumper visits EMBL-EBI:** We were pleased to welcome John Jumper, Distinguished Scientist at Google DeepMind and 2024 Chemistry Nobel Prize winner, to EMBL-EBI. Jumper and colleagues from Google DeepMind met with EMBL-EBI's AlphaFold Database and Protein Data Bank in Europe (PDBe) teams before delivering a seminar on the development of AlphaFold.

- **Multicultural day:** This EMBL-EBI Staff Association celebration brought the EMBL-EBI community together to share culinary traditions, music, and personal stories, reflecting the incredible diversity that defines our institute. The event provided a warm, welcoming space for colleagues to celebrate their unique backgrounds. It was a spirited reminder that our scientific strength is underpinned by the rich variety of cultures and perspectives we are lucky to share.

- **Farewell to Rolf Apweiler:** After a remarkable career spanning four decades, EMBL-EBI Director Emeritus Rolf Apweiler retired in 2025. Rolf has been a driving force in making biological data open and accessible to the world and his impact remains visible from the global success of UniProt to his role in guiding the institute through the pandemic.

New leadership

Jo McEntyre

Jo McEntyre has stepped up to become Interim Director of EMBL-EBI. She leads the institute, overseeing its strategic direction, services, and research, while ensuring continuity in our mission to manage the world's biological data.

Emily Clark

Emily Clark is the new Genome Analysis Team Leader. Her team manages the portals that make agricultural and biodiversity data accessible, including for the Earth BioGenome Project. They also develop the comparative genomics tools within Ensembl.

Tim Coorens

Tim Coorens joined as a Research Group Leader to investigate the hidden histories of human cells. His group uses somatic mutations as genetic barcodes to trace cellular lineages, helping to uncover how normal tissues develop and how cancers originate.

Noel O'Boyle

Appointed as Chemical Biology Resources Team Leader, Noel O'Boyle oversees critical data resources including ChEMBL and ChEBI. His focus includes expanding data coverage to include new biologics and using AI to streamline data curation and search.

Jess Ewald

Appointed as a new Research Group Leader, Jess Ewald and her group are working at the intersection of toxicology and machine learning. They use different data types to profile cells and predict how chemical contaminants impact both human health and ecosystems.

Alex Picchianti

Alex Picchianti joined as the Head of Infrastructure Services. He leads the technical teams responsible for the scalable, secure, and reliable high-performance computing and storage systems that underpin the institute's massive data growth.

Interim Director

Johanna McEntyre

Director's Office

Research Management Office

IT & Technical Services (ITS)

Andy Cafferkey

Web Development
Geetika Malhotra

IT Applications
David Ocaña

Systems Infrastructure
Alex Picchianti

IT Operations
Tim Porter

Genome Assembly and Annotation

Rob Finn

Genome Analysis
Emily Clarke

European Nucleotide Archive
Guy Cochrane
(outgoing)

Non-vertebrate Genomics
Sarah Dyer

Microbiome Informatics
Rob Finn

Human Genomics
Mallory Freeberg

Eukaryotic Annotation
Fergal Martin

Genomics Technology Infrastructure
Andy Yates

Knowledge Management

Helen Parkinson

Literature Services
Melissa Harrison

Molecular Systems
Henning Hermjakob

Samples, Phenotypes & Ontologies
Helen Parkinson

BioStudies
Ugis Sarkans

Macromolecules, Structure, Chemistry and Bioimaging

Sameer Velankar

Protein Family Resources
Alex Bateman

Functional Genomics
Christina Ernst

BioImage Archive
Matthew Hartley

Chemical Biology Services
Noel O'Boyle

Protein Function Content
Sandra Orchard
(outgoing)

Protein Function Development
Maria-Jesus Martin

Electron Microscopy Data Bank
Kyle Morris

Cellular Structure and 3D Bioimaging
Ardan Patwardhan

Protein Data Bank in Europe
Sameer Velankar

Proteomics and Metabolomics
Juan Antonio Vizcaíno

Research

Julio Saez-Rodriguez

Research Office

Alex Bateman

Isidro Cortes-Ciriano

Jess Ewald

Nick Goldman

Thomas Keane
(visiting)

John Marioni
(visiting)

Evangelia Petsalaki

Janet Thornton
(visiting)

Ewan Birney

Tim Coorens

Rob Finn

Zamin Iqbal
(visiting)

John Lees

Irene Papatheodorou
(visiting)

Julio Saez-Rodriguez

Virginie Uhlmann
(visiting)

Admin & Operations

Rachel Curran

Admin Operations Office

Campus Operations and Health & Safety
Dan Trigg

Finance, Procurement and Accounting
Nirav Dalal

Human Resources
Manu Morgan

Legal Services
Alannah Lockwood

Grants Office
Ruth Sandland

Training

Cath Brooksbank

Major Initiatives
Mary Barlow

Communications
Gemma Wood

Industry Programme
David Hulcoop

Open Targets
David Hulcoop

This organogram shows EMBL-EBI leadership in 2025.

Governance, management, and funders

EMBL-EBI is part of the European Molecular Biology Laboratory (EMBL), Europe's life sciences laboratory, enabling and advancing discovery to understand life. EMBL integrates fundamental research, advanced scientific services and excellence in training and innovation to tackle complex biological and societal challenges.

In 2025, EMBL was led by Interim Director General Peer Bork and Interim Executive Director Ewan Birney, both appointed by the EMBL Council. During the same year, EMBL-EBI was led by Interim Director Jo McEntyre, Head of Research Julio Saez-Rodriguez, Head of Admin and Operations Rachel Curran and Associate Director Rolf Apweiler, who has now retired.

The EMBL Council is composed of representatives from all member states of the Laboratory and determines its policy in scientific, technical and administrative matters by giving guidelines to EMBL leadership. The Council ensures that the financial requirements of the agreement establishing EMBL, and of the agreements with host member states, are complied with.

We would like to thank the EMBL member states as well as our other funders listed below for their continued support.

- Biotechnology and Biological Sciences Research Council
- Bill & Melinda Gates Foundation
- Canadian Institutes of Health Research
- Cancer Research UK
- Chan Zuckerberg Initiative
- European Association for Cancer Research
- European Commission
- Engineering and Physical Sciences Research Council
- Economic and Social Research Council
- Fonds National de la Recherche
- International Society of Pneumonia and Pneumococcal Diseases
- Leverhulme Trust
- Mark Foundation for Cancer Research
- Gordon and Betty Moore Foundation
- Medical Research Council
- National Institutes of Health
- National Institute for Health Research
- Novo Nordisk
- Osteosarcoma Institute
- Research Council RCUK
- Pezcoller Foundation
- Wellcome Trust

Imprint

Publisher
EMBL-EBI

Portrait image credits
p. 6 (Kate Baker): Kate Baker/Department of Genetics at the University of Cambridge
p. 15 (Martin McHugh): School of Medicine at the University of St Andrews
p. 16 (Adrienne Flannigan): Adrienne Flannigan/University College London

All portraits not listed here are from personal collections or taken on behalf of EMBL-EBI by Jeff Dowling.

Other image credits
p. 3 (EMBL-EBI South Building): Jeff Dowling/EMBL-EBI
p. 18-19 (AlphaFold Education Summit): Laura Lewis/AlphaFold Google DeepMind
p. 19 (AlphaFold protein): AF-0000000365762994: Spike protein VP9
p. 22-23 (Thornton Building opening event): Mike Thornton/Stillvision Photography
p. 27 (Spring view/Take 5 outside): Simone Vegh/EMBL-EBI
p. 27 (John Jumper): Jeff Dowling/EMBL-EBI
p. 27 (Rolf Apweiler's retirement): Jeff Dowling/EMBL-EBI

Printer
Healeys. healeys-printers.co.uk

EMBL-EBI is a part of the European Molecular Biology Laboratory.

A digital version of this publication is available on
www.ebi.ac.uk/about/our-impact

EMBL member states and associate member states: Austria,
Belgium, Croatia, Czech Republic, Denmark, Estonia, Finland,
France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy,
Latvia, Lithuania, Luxembourg, Malta, Montenegro, Netherlands,
Norway, Poland, Portugal, Slovakia, Spain, Sweden, Switzerland,
United Kingdom, Australia

Prospect member states: Bulgaria, Serbia