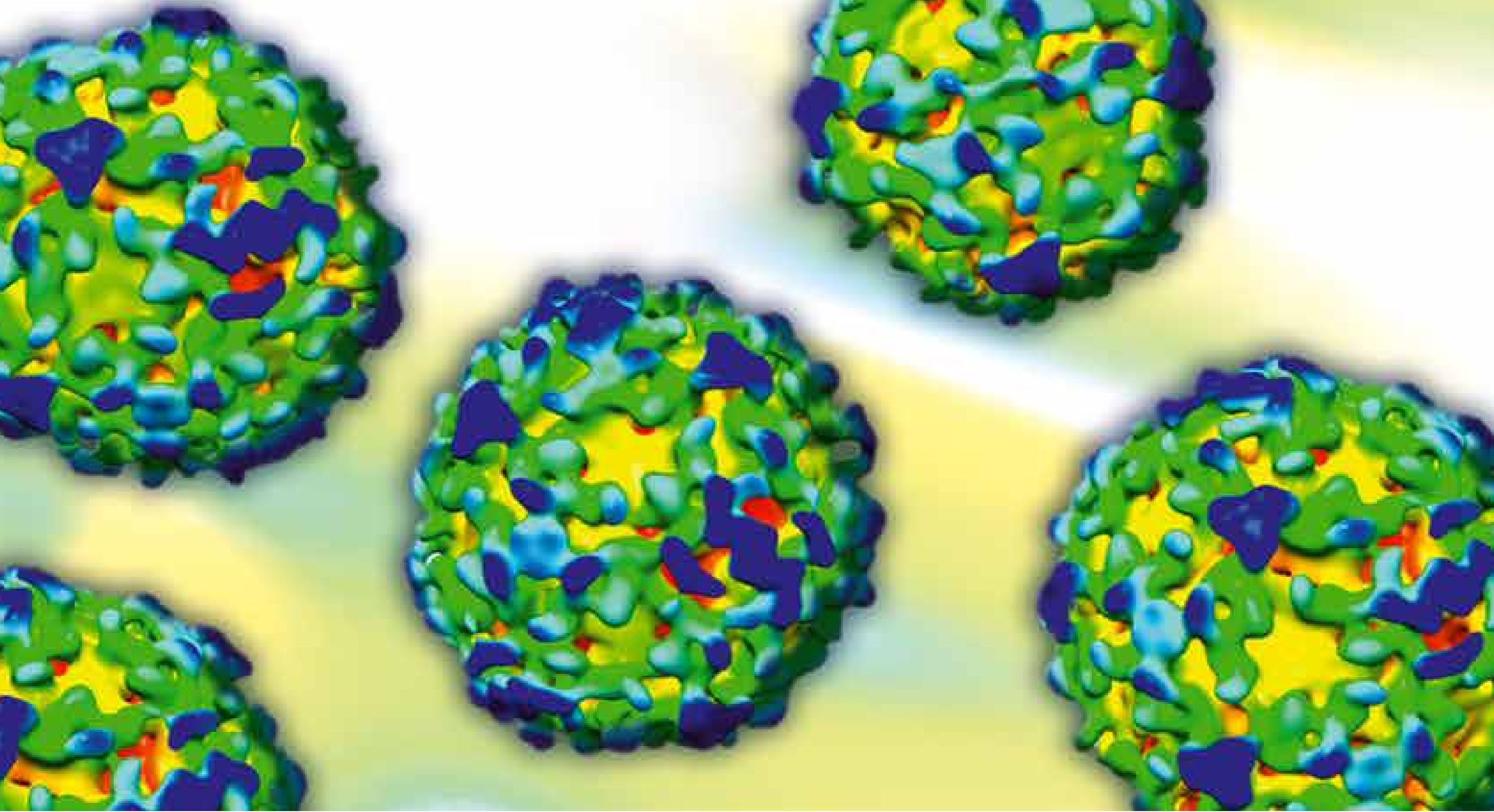
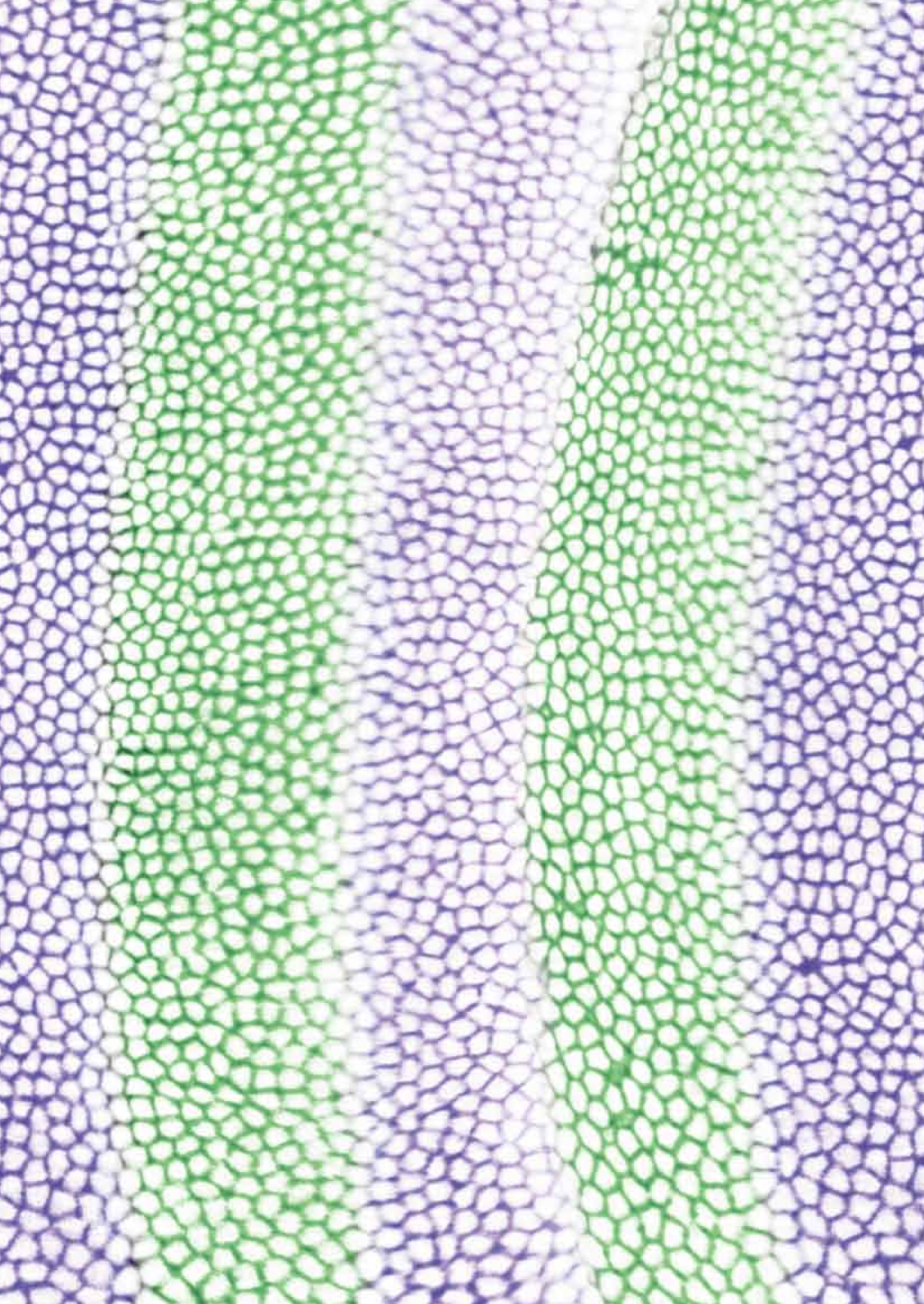


Annual Report

2012-2013

EMBL



Annual Report 2012-2013

European Molecular Biology Laboratory

Contents

The Director General's Report

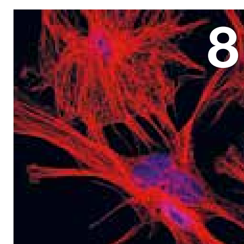
Foreword

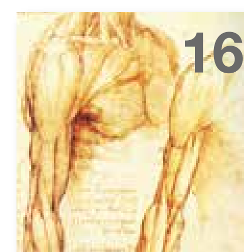
State of the Laboratory

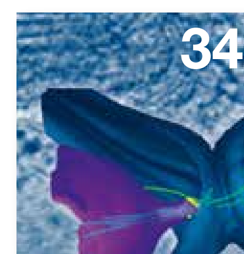
vi Research
ix Services
xii Technology Transfer
xii Training
xiii EMBL Alumni
xiv Outreach
xv Administration

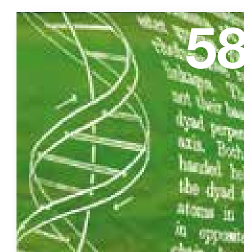
Integration of European Research

xv Member state relations
xvi EMBL partnerships
xviii European Research Infrastructures
xix Relations with the European Commission
xx Personnel statistics
xxii Financial report
xxiv Reviews of EMBL Scientific Units
xxv EMBL Genome Biology Unit Review
xxvii EMBL Monterotondo Review
xxix EMBL Grenoble Review
xxxi EMBL-EBI Research Activities Review









Science Highlights

2 ENCODE-ing the future
6 Turning back the clock
8 A touch of complexity
10 How viruses come of age
12 Plug-in puzzle
13 Perfect timing
14 Mysterious beginnings
16 Twisted communication
17 Bending borders
18 I am the one and only
22 Knotty problems
24 A force for change
27 Precious parcels
28 Faultless fit
32 Age matters
34 Let's get together
36 Better bugging
37 Just one drop
38 Unlocking the pore
42 Making waves
44 Bridging the gap
47 Context and variation
48 Natural capital
52 Shimmering crystals and bright beams
54 The greatest show on earth
56 A record catch
58 Perfect storage
62 Complex loading
64 Export economy
65 Picky partner
66 Back to normal

68 A year in the life of EMBL
80 Index of group & team leaders

As our thoughts turn towards the celebration of EMBL's 40th anniversary in 2014, when there will be a natural tendency to look back, let's focus for this year on the present and the future. This year's report presents abundant evidence of the vitality of EMBL, in spite of our approaching 'middle age'. Our researchers were once again involved in a number of important breakthroughs, either as individual research teams or, increasingly, as an important part of a larger collaboration. The prediction made when EMBL was founded – that molecular biology would become 'big science' – has only recently been validated. 'Omics' and systems-based research, shorthand for global analyses of one or more aspects of a biological phenomenon that are being enabled by new technologies, mean that many projects can no longer be carried out by a single laboratory or institute, but demand many participants who bring very different expertise to bear on a common problem. The interdisciplinarity fostered in EMBL over the past decade has prepared us well for participation in such studies.

The requirement of the biological and medical communities for new or improved research infrastructures is driven significantly by the same technology developments and the resulting new opportunities. EMBL scientists have co-ordinated the preparatory phase of two of the ESFRI research infrastructure projects. As they approach realisation, these projects provide insight into how European research funders and ministries are struggling to find ways to make the investments necessary for future innovation at a time of current economic hardship. Ensuring the successful outcome of these projects will keep us, and many other people, occupied for some time to come. As part of these efforts, EMBL has actively interacted in multiple contexts with heads of state and government and key decision makers at the EU level concerning Horizon 2020, the proposed EU framework for research and innovation that will run between 2014 and 2020. EMBL issued a position paper on Horizon 2020 calling for further support of basic research, increases to the overall budget and greater support for research infrastructure. The European Commission has expressed its appreciation of EMBL's position paper and indicated that recommendations from key stakeholders will be taken into consideration. In addition, EMBL supported initiatives on the same topic taken by the EIROforum Directors General and the petition launched by the Initiative for Science in Europe.

Finally, I'd like to acknowledge a few of the senior researchers at EMBL whose performance has been recognised this year. EMBL-EBI Director Janet Thornton was made Dame Commander of the British Empire in recognition of her contribution to bioinformatics. The joint Heads of the Genome Biology Unit, Eileen Furlong and Lars Steinmetz, and the Head of the Grenoble outstation, Stephen Cusack, were each awarded a European Research Council (ERC) Advanced Investigators Grant. And I would like to welcome Ari Himma and Astrid von Soosten, who joined EMBL as Head of Human Resources and Head of Resource Development, respectively, and to thank and congratulate Keith Williamson who, after stints as Head of Personnel and Head of Finance, took over at a critical time at the end of 2012 in his new role as Administrative Director. I am convinced that these departments are all in capable hands and will develop positively in the coming years.

Iain Mattaj

State of the Laboratory

On 13 June 2012, work on a new bioinformatics technical hub on the Wellcome Trust Genome Campus in Hinxton officially started, with a landmark ceremony during which European Bioinformatics Institute (EMBL-EBI) Director Janet Thornton broke the ground. The new building, which has been funded by the UK government's Large Facilities Capital Fund, will house some 200 staff and is scheduled for completion in autumn 2013. It will house the ELIXIR Hub as well as an Industry and Innovation Suite that will promote the translation of biological information to applications in medicine, biotechnology and the environment.

Research

EMBL researchers consistently publish papers in high-impact journals and are heavily cited, allowing EMBL to be continuously ranked as one of the top research institutes worldwide. The quality of the research and services has continued to be recognised in the past year, with both junior and senior researchers receiving some of the most prestigious European grants and awards. For example, EMBL-EBI Director Janet Thornton was made Dame Commander of the British Empire in recognition of her contribution to bioinformatics. Also at EMBL-EBI, Rolf Apweiler, Ewan Birney and Sarah Teichmann were appointed members of EMBO. John Briggs and Carsten Schultz from EMBL Heidelberg received the Chica und Heinz Schaller prize and the HMLS investigator award, respectively. In the past year, the joint Heads of the Genome Biology Unit, Eileen Furlong and Lars Steinmetz, and Head of the Grenoble outstation, Stephen Cusack, were each awarded a European Research Council (ERC) Advanced Investigators Grant. EMBL's junior faculty has been equally successful. In 2012, group leaders Dónal O'Carroll and Martin Beck received ERC Starting Independent Research Grants, bringing the overall number of EMBL researchers holding an ERC grant to 14.

Inter-Unit Collaborations

In 2012, EMBL scientists produced more than 600 scholarly papers, most of which were produced in collaboration with other institutions from around the globe. In addition, there are many examples of teamwork at EMBL that have led to scientific breakthroughs. For example, the collaboration between Associate Director Matthias Hentze's group, several groups in the Genome Biology Unit and external collaborators resulted in a publication in *Cell* identifying 300 proteins previously not known to bind to RNA. To do so, the scientists developed a new method to identify and isolate all proteins that bind to RNA in living cells. These findings raise new possibilities for researchers to investigate, such as the prospect that conditions ranging from diabetes to glaucoma could be caused by a malfunction in the ability of some proteins to bind to and regulate RNA (page 56). The large international 1000 Genomes consortium, which includes many researchers from EMBL-EBI, the Genome Biology Unit and the Genomics Core Facility, published the long-awaited map of variation in the

The new South building on the Genome Campus in Hinxton.

EUROPEAN RESEARCH COUNCIL GRANTS

ERC Advanced Grant

Detlev Arendt, Heidelberg
Peer Bork, Heidelberg
Stephen Cusack, Grenoble
Eileen Furlong, Heidelberg
Matthias Hentze, Heidelberg
Lars Steinmetz, Heidelberg

ERC Starting Grant

Martin Beck, Heidelberg
Marcus Heisler, Heidelberg
Takashi Hiiragi, Heidelberg
Dónal O'Carroll, Monterotondo
Francesca Peri, Heidelberg
Ramesh Pillai, Grenoble
Christiane Schaffitzel, Grenoble
Rocio Sotillo, Monterotondo

human genome in *Nature*. This was a major milestone for Paul Flicek's team, who co-led the project's data co-ordination centre with the US National Centre for Biotechnology Information (NCBI) in the USA. Jan Korbel, who co-leads the project's study of structural variation, noted that the integrated view of genome variation will be extremely useful for identifying genetic causes of diseases. This vast dataset is freely available from several sources including Ensembl, the Amazon Web Services Cloud as well as on the 1000 Genomes browser, which also provides a suite of tools to help scientists make the most of the data (page 47).

Computational Biology and Bioinformatics

The EMBL-EBI research activities were reviewed in March 2013 and the panel rated the overall performance as outstanding (page xxxi). In the past years, under the inspiring leadership of Janet Thornton, EMBL-EBI has firmly established itself as one of the leading bioinformatics institutes worldwide, providing services and carrying out research of the highest calibre. The past year has been very successful for EMBL-EBI, with several papers attracting attention from both the scientific community and the media. In January 2013, Nick Goldman's and Ewan Birney's groups published a scalable method to reliably store information in DNA, offering a realistic, although currently expensive, technology for large-scale, long-term archiving of infrequently accessed data (page 58). In September 2012, the ENCODE project dominated the scientific news worldwide. Co-led by Ewan Birney at EMBL-EBI and funded by the National Human Genome Research Institute in the USA, the project comprised over 400 scientists in 32 labs throughout the world, and produced a detailed map of genome function. ENCODE's staggering output inspired a new publishing model: upwards of 30 papers were published under open-access license in several different peer-reviewed journals, with the contents linked by topic and united for optimum exploration in a single interface provided by *Nature*. The launch of ENCODE was all about public engagement, with a press conference featuring an exhibit at the Science Museum in London, UK and aerial acrobats re-enacting gene expression. Media coverage of the story was widespread: more than 13 000 articles devoted to the project were published online, almost 500 of which were in high-level publications; it made front-page news in the UK's *The Guardian*, *The New York Times*, and major national newspapers throughout Europe; feature articles appeared in magazines such as *The Economist* and *Der Spiegel*, and countless discussions were generated (page 2).

The Structural and Computational Biology Unit at EMBL Heidelberg received media attention for their analysis of the gut metagenome. This was performed at such a high resolution that more than 10 million mutations in the various bacterial strains in the gut of 207 individuals could be identified. The findings by Peer Bork and co-authors showed that, at least when healthy, each of us carries a unique set of bacterial strains, defined by their mutations, in our gut (page 21).

Structural Biology

Research in John Briggs' group at EMBL Heidelberg resulted in three papers being published in leading journals in as many months. One, published in *Cell*, resulted from a collaboration with Marko Kaksonen's group in the Cell Biology and Biophysics Unit, by combining the real-time, multi-colour imaging of fluorescence microscopy with the high resolution of electron microscopy (a technique developed by an EMBL Interdisciplinary Postdoctoral (EIPOD) fellow shared by the two groups, and featured in the Annual Report 2010-2011, page 4). This enabled them to study the changes in and around the cell membrane as an individual vesicle forms at the cell surface (page 17). Another study, published in *Nature,* showed the detailed structure of the shell that surrounds the genetic material of retroviruses, such as HIV (page 10).

EMBL Grenoble was reviewed in February 2013 (page xxix). The panel rated the research, technology development and service activities as outstanding. All

groups at the outstation work on problems of great scientific interest and have been able to bring about remarkable advances in structural and molecular biology. For instance, scientists in Imre Berger's and Christiane Schaffitzel's groups and their collaborators described the architecture of the central scaffold of TFIID, the human protein complex essential for the transcription of DNA to messenger RNA (page 12). In turn, Stephen Cusack and his group determined the 3D structure of part of the flu virus' RNA polymerase in complex with small chemical compounds that can inhibit this enzyme's crucial activity in viral replication. The study will help on-going efforts to design innovative drugs that are effective against all strains of the flu virus (page 28).

The research group of the Head of EMBL Hamburg, Matthias Wilmanns, has solved the high-resolution crystal structures of two master transcription factors in the presence of specific cognate DNA elements, including a master regulator of melanocyte development. Mutations in this protein can lead to either hereditary diseases or cancer (page 65). Matthias' group also determined the structure of a complex of two transcription factors involved in embryonic stem-cell differentiation, Oct4 and Sox2, in the presence of a shared DNA-response motif that allows the formation of a DNA-mediated protein/protein complex. By visualising a previously unseen connecting loop, an important step was made in advancing our understanding of the specific roles of Oct4 and Sox2 in pluripotency of stem cells (page 44). The establishment of a new Centre for Structural Systems Biology (CSSB) on the German Synchrotron Research Centre (DESY) campus has continued to advance. An architect has been appointed, the building terrain has been cleared, and recruitment of staff has started. The CSSB will be an interdisciplinary centre with partners from several universities and research facilities from northern Germany, including EMBL, aiming to bring together state-of-the-art structural biology, infection biology and systems biology approaches. The CSSB Scientific Advisory Board expressed great enthusiasm for this initiative and its potential for discovery and innovation in the field of infectious diseases.

The architects' model of the new CSSB building on the DESY campus in Hamburg.

Genome Biology

The Genome Biology Unit in Heidelberg underwent its four-yearly review in May 2012 (page xxv). The panel rated the research of the highest calibre, comparing it favourably with the best work of this type in the world. The Unit combines an effective managerial team with an outstanding group of young scientists whose broad expertise collectively creates a fertile environment in which to carry out the interdisciplinary research now at the heart of the Unit's endeavours. This fertile environment has enabled the group of Jeroen Krijgsveld, Head of the Proteomics Core Facility, to develop a new method to analyse secreted proteins. This approach has generated wide interest in the community thanks to its multiple potential applications, from testing the cellular response to drugs to analysing samples from patients (page 36). As part of the International Cancer Genome Consortium, Jan Korbel's group discovered that early-onset prostate cancers are triggered by a different mechanism from those that develop at a later age. This finding could have important consequences for the diagnosis and treatment of this cancer (page 32).

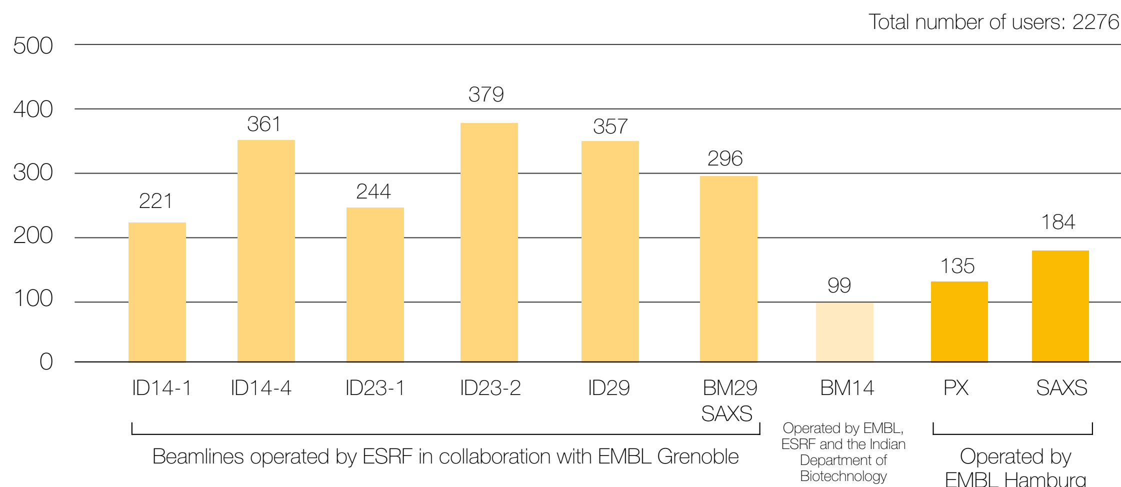
Cell Biology and Biophysics

Using a systematic approach for silencing genes that they developed, Rainer Pepperkok's group together with Jan Ellenberg's group observed that 15% of human genes somehow influence the cell's ability to secrete a protein, revealing an unexpected variety of genes involved in transporting molecules to the cell membrane and beyond (page 64). The team of Lars Hufnagel took the Selective-Plane Illumination Microscope (SPIM) previously developed at EMBL (see AR 2007-2008, page 74) to a new level, which they called Multi-View SPIM. The new microscope takes four full images from different angles, eliminating the need to rotate samples, and enables scientists to merge the images straightforwardly into a single high-quality 3D image (page 54).

Developmental Biology

Alexander Aulehla's group found that the size of the future vertebrae in a mammalian embryo is influenced by a wavelike gene expression pattern. A new assay performed on a single layer of cells allowed the group to identify a clear link between the size of the embryo, the segment size and the speed of those expression waves (page 42). Looking at a gene that is crucial for embryonic development, François Spitz's group showed that it is controlled by a series of regulatory elements, the interplay of which is determined not by their DNA sequence but by where they are located on a chromosome. This means it is likely that the way DNA folds in three dimensions can, under certain circumstances, bring together different sets of

Beamline Users EMBL 2012



regulatory elements to trigger or prevent gene expression (page 16).

Mouse Biology

EMBL's Mouse Biology Unit was reviewed in October 2012, following its first major wave of staff turnover (page xxvii). The productivity of this relatively small Unit was found to be outstanding. The panel acknowledged the significant contribution of former Head Nadia Rosenthal to the success of the Unit. To fully accomplish its goals, the new Head of Unit, Phil Avner, is planning to more tightly focus the Unit's research efforts. He is also very interested in developing a better integration of EMBL Monterotondo with the Italian scientific community, particularly through more collaboration between the institute and local researchers. Paul Heppenstall's group demonstrated how extending findings obtained in other animal models to the mouse model can lead to interesting surprises. Previous research had found that worms lacking a specific protein were practically unable to respond to touch, but when Paul's group silenced the equivalent gene in mice, the animals seemed unaffected. These findings have provided new insights into the protein's role and its interactions with the cell's scaffold (page 8).

Services

Structural Biology Services

In addition to running successful research programmes in structural biology, EMBL's outstations in Grenoble and Hamburg both provide access to cutting-edge research infrastructures. The Hamburg teams work closely with the German Synchrotron Research Centre (DESY) and the Grenoble teams with the European Synchrotron Radiation Facility (ESRF) to make the powerful X-ray sources available for applications in the life sciences. In 2012, EMBL Grenoble and Hamburg jointly registered 2276 users.

At precisely 8 am on Monday 22 October 2012, the DORIS synchrotron ring at DESY finished work on its very last experiment. Ever since EMBL Hamburg opened its doors for scientific business in 1974 as EMBL's first outstation, DORIS has been a central part of its life and work. In recent years, considerable effort has gone into the design and construction of beamlines at the newly refurbished PETRA III ring. The recently completed conversion of the PETRA storage ring into a leading synchrotron radiation facility has allowed EMBL Hamburg to create an integrated facility for future state-of-the-art applications in structural biology. EMBL operates three undulator beamlines, two of which are dedicated to macromolecular X-ray crystallography (P13 and P14) and one to small angle X-ray scattering applications (P12) of biological samples (BioSAXS). The beamlines are integrated into advanced facilities for biological sample preparation, characterisation and crystallisation as well as for X-ray data processing and evaluation. A key transition phase came to an end in 2012 as construction, commissioning and test user operation at the PETRA III beamlines was completed. Key milestones in novel technical implementations were:

- The establishment of a new combined on-line chromatography system at the BioSAXS beamline, allowing parallel biophysical characterisation of various parameters of biological samples.
- The installation of two PILATUS 6M detectors on both X-ray crystallography beamlines P13 and P14.
- The installation of a state-of-the-art MD3 diffractometer, in collaboration with EMBL Grenoble enabling continuous 4D-scans on long crystals between two centering points.
- The installation of vertical and horizontally focussing bimorph mirrors, allowing very hot microbeams to be generated.

To date, two calls for proposals, one in early 2012 and the other in early 2013, have been launched and almost 200 applications were received.

EMBL Hamburg also offers a variety of computational services and software packages. Last January, a novel software server for *in silico* lead and ligand prediction (ViCi) was made available free of charge to the worldwide academic community by Victor Lamzin's group. The software aids in the prioritisation of compounds for experimental screening.

Aside from the synchrotron-related activities, EMBL Hamburg is co-ordinating a project to construct, commission, and operate a biological sample preparation facility at the European X-Ray Free Electron Laser (XFEL) that is currently under construction on campus. Called XFEL-based Biological Infrastructure (XBI), the preliminary project, initially submitted by EMBL and its international collaborators, was recommended for implementation by the XFEL Council and resulted in the submission of a full XBI-project proposal.

In Grenoble, the ESRF is halfway through phase I of its upgrade programme, which entails the construction of five new beamlines, the refurbishment of many existing beamlines and major new developments in synchrotron radiation instrumentation. As part of the upgrade, the BioSAXS station has been rebuilt complete with inline high-performance liquid chromatography (HPLC) and the high throughput, small volume sample changer designed by Florent Cipriani's group in collaboration with EMBL Hamburg. Users can now benefit from increased flux, improved focus (100 μm^2 beamsize), shorter data collection times and automated sample handling and data reduction. In addition, EMBL Grenoble is involved in the construction of a new beamline complex on ID30 comprising the Massively Automated Sample Selection Integration Facility (MASSIF, ID30A) and a state-of-the-art re-incarnation of the ID14-4 beamline called ID30B. MASSIF is designed to allow automated screening of large numbers of crystals for difficult projects – such as membrane proteins and large complexes – in order to find the best spot on the best crystal for data collection.

EMBL Grenoble has been an essential contributor to the development of automation technology not only at ESRF but also at other beamlines worldwide. For example, the pioneering MD2 micro-diffractometer has now been sold by MAATEL for use at thirteen synchrotrons from California to Shanghai. Recently, an ultra-high precision, vertical diffractometer, MD3, has been developed as a result of a collaboration between the Grenoble and Hamburg outstations, led by Florent, with the first such instrument now operational at PETRA III. Another example is the new technology developed by Florent's and José Márquez's teams, called CrystalDirect™, which allows automated mounting of crystals from the solution in which they grow, a bottleneck step that was not previously automated (page 53).

In November 2012, BM14, the beamline operated by an EMBL-India-ESRF consortium and operationally managed by EMBL, was reviewed by an external panel. BM14 gives European and Indian users access to an extremely stable, largely radiation damage-free, bending magnet beamline renowned for the quality of its anomalous data measurements. The beamline is also used by EMBL as a testing ground for new instrumentation. The review was very positive, with the panel highlighting the benefits accrued by all partners. The panel that reviewed EMBL Grenoble activities was also very supportive of the new Unité Mixte de Service (UMS), a joint venture by the CNRS, CEA, UJF and EMBL to organise the IBS and UVHCI technical platforms for access to beamlines by local, national, industrial and international users.

In addition to the beamlines, EMBL Grenoble runs three other user platforms, which are available to in-house users, members of the Grenoble Partnership for Structural Biology and external users under the BioStruct-X programme and its predecessor, the oversubscribed P-Cube project. These are the High-Throughput Crystallisation Facility, the unique ESPRIT platform for the identification of soluble domains in hard-to-express proteins, and the Eukaryotic Expression Facility for expressing multi-protein complexes in insect cells using the MultiBac method.

EMBL Hamburg and EMBL Grenoble both participate in a new Seventh Framework Programme (FP7) initiative called BioStruct-X, coordinated by Matthias Wilmanns, which received almost 500 project applications from structural biologists during its first year. BioStruct-X offers access to structural biology applications at 44 installations in four key areas: small angle X-ray scattering, macromolecular X-ray crystallography, biological X-ray imaging, and protein production and high-throughput crystallisation. BioStruct-X cooperates with the ESRFI project Instruct (Integrated Structural Biology Infrastructure for Europe), in which EMBL is also involved, in aiming to provide an integrated and co-ordinated technology platform.

Bioinformatics Services

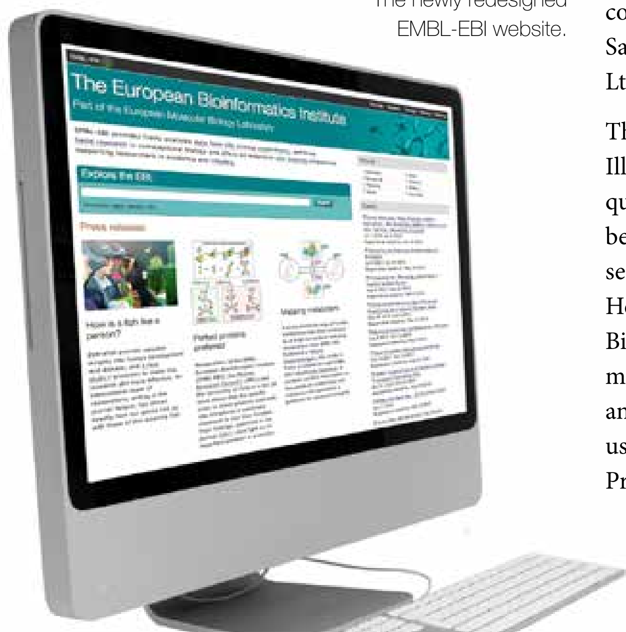
EMBL-EBI hosts Europe's most comprehensive biomedical data resources and makes them freely available to the scientific community in ways that promote scientific progress. The data are heavily used by scientists around the world working in both academia and industry. In 2012, there were on average 7 million webhits on the services per day, which is up from 5.3 million in 2011. Continuing the trend from previous

years, all core data resources have grown substantially in 2012. For instance, the nucleotide sequence databases currently have a doubling time of less than one year, which means that more than 50% of the sequence data have been in the archive for less than a year. In order to keep up with these growing data volumes, scientists at EMBL-EBI spend a lot of their time developing effective mechanisms for data compression. The recent first release of CRAM, an open software toolkit and file format for compressing sequence data, has been widely embraced. In the past years, EMBL-EBI has received funds from the UK government Large Facilities Capital Fund to enable data service provision, including acquisition of space and equipment. The funds are being used to construct a new building on campus and to support the acquisition of high-security data centre space and computing equipment.

The EMBL-EBI services have been reorganised into 'clusters' to reflect the research communities they serve. The clusters provide a better framework to make strategic decisions whilst still enabling a deep engagement with the communities. In addition, the new EMBL-EBI website has been launched. The redesign strategy focused on keeping users at the centre of the process in order to provide an intuitive interface that encourages people to explore the numerous bioinformatics services.

Other bioinformatics service highlights of the past year include:

- UK PubMed Central re-launched as Europe PubMed Central, as three European funders, including the ERC, joined forces with 18 British funders;
- the community of neglected diseases received a boost from the online drug-discovery database ChEMBL, as John Overington's team collaborated with the Medicines for Malaria Venture to provide one-stop access to MalariaBox and other open-access data, including new, high-value malaria and tuberculosis datasets;



The newly redesigned EMBL-EBI website.

- the mountain gorilla genomics dataset and four other genomes entered Ensembl in 2012;
- the EMBL-EBI Metagenomics resource was officially launched by Sarah Hunter's and Guy Cochrane's teams, enabling users to submit, archive and analyse genomic information from environments containing many species;
- UniProt celebrated its 10th birthday. Both UniProt and InterPro launched new interfaces addressing users' needs;
- the newly launched Enzyme Portal draws together data that previously resided in ten different databases;
- EMBL-EBI and colleagues in the International Molecular Exchange (IMEx) consortium are offering researchers a freely available set of experimental interaction data that can be queried from a single interface;
- MetaboLights, which gives a home to metabolomics experiments and derived information, was launched in 2012 by Christoph Steinbeck's team. In October 2012, the European Co-ordination of Standards in Metabolomics (COSMOS) consortium officially started work on metabolomics data standardisation, publication and dissemination workflows.

Core Facilities

EMBL Heidelberg operates seven Core Facilities, which are central components of EMBL's research network. They offer cutting-edge technology and services in the areas of advanced light microscopy, electron microscopy, genomics, proteomics, protein expression and purification, flow cytometry, and chemical biology and are heavily used by many research groups at EMBL and, to a lesser extent, by external academic users from our member states as well as by EMBO Young Investigators. At the end of August 2012, the activities of the Monoclonal Antibody Facility at EMBL Monterotondo were turned into a spin-off company. Led by the Head of the former Facility, Alan Sawyer, the company is now established as Paratopes Ltd in the UK.

The Genomics Core Facility in Heidelberg added an Illumina Miseq device, upgraded some of the other sequencing and qPCR devices, and increased the number of liquid handling robots used for next generation sequencing sample preparation. The University of Heidelberg joined as the third partner in the Chemical Biology Core Facility, which also launched a focussed medicinal chemistry effort. The new BD Fortessa analyser in the Flow Cytometry Core Facility allows users to operate on a self-service basis. The EMBL Protein Expression and Purification Core Facility has

interacted with other facilities in the Protein Production and Purification Partnership in Europe (P4EU) network in a continuous exchange of ideas and discussions. The Advanced Light Microscopy Facility has extended its capacity for super resolution microscopy, and started a collaboration with Perkin Elmer to offer its users the high volume data storage and large image data analysis software package Columbus.

IT and Biocomputing Services

A major IT project is the FP7-funded 'Helix Nebula – the Science Cloud' in which EMBL is part of a consortium involving over 30 partners including EIROforum members such as CERN and the European Space Agency (ESA), as well as leading IT industry providers. The project aims to develop a technical solution to the enormous and ever-growing IT requirements of European scientists by launching a sustainable European cloud computing platform. This will provide stable computing capacities and services that elastically meet the demands of the European scientific community.

An EMBL Heidelberg-wide initiative to enhance the biocomputing activities has been triggered by the increasing importance of biocomputing for all EMBL groups, including wet labs. The Bio-IT project is being developed as a joint effort by several groups at EMBL Heidelberg to promote interactions amongst the bioinformatics community, and to provide a focal point for EMBL biocomputing resources.

Technology Transfer

EMBL Enterprise Management Technology Transfer GmbH (EMBLEM), the limited liability company wholly owned by EMBL, ensures the rapid commercial development of promising innovations while concomitantly securing the free dissemination of knowledge for basic research purposes. The success of the technology transfer activities is reflected both in the broad engagement of scientific staff – over 460 EMBL scientists are on record as inventors – as well as in the more than 250 satisfied commercial licensees of EMBL technologies, more than half of whom are recurring customers interested in establishing a long-term relationship with EMBL and EMBLEM. EMBLEM manages a portfolio of more than 200 granted patents and patent applications, 120 copyrights and trademarks and 16 spin-off companies. In 2012, the EMBL spin-off company Cellzome was acquired by GlaxoSmithKline (GSK), whereas another EMBL spin-off, Savira Pharmaceuticals, signed a collaboration agreement with Roche to develop RNA polymerase inhibitors for the treatment of influenza virus infections (page 30).

EMBLEM TECHNOLOGY TRANSFER IN NUMBERS (2008-2012)

Turnover	€ 19,600,000
Invention disclosures	208
Priority patent applications	55
Software copyrights	29
New license agreements	937
Material transfer agreements (MTAs)	>1,000
Confidentiality agreements (CDAs)	>150
Inter-institutional agreements (IIAs)	69
Start-ups created	6

Training

In the past year, the EMBL International Centre for Advanced Training (EICAT) saw a steady growth in size and performance. The EMBL International PhD Programme (EIPP) consistently attracts high numbers of applications and took a step towards improving the overall ratio of member state applications. This development seems to be attributable to implementing a new advertisement strategy using direct e-mailing campaigns with a focus on member states. This new EIPP advertising campaign not only made the EIPP web page the third most popular EMBL page overall, but also generated peak traffic to the EMBL main page.

The project to further streamline the administrative work of the PhD Programme resulted in the successful introduction in 2012 of a new online application system, which makes the initial screening of incoming applications easier. In early 2013, the second module of the software, which will reduce the time and effort EMBL faculty spend on short-listing suitable candidates for interviews, has been launched.

THE EIPP IN 2012

Number of PhD students in 2012: 212

Number of applications in 2012: 1585

New PhD students joining EMBL in 2012: 59

Graduations in 2012: 44

Average first-author publications during PhD: 2

Like the EIPP, EMBL's Postdoctoral Programme experienced another year of progress in 2012. In particular, the EMBL Interdisciplinary Postdoctoral (EIPOD)

Programme was successful in securing a second FP7 grant (COFUND 2) that will support another 60 fellowships, each for three years, until mid-2017. In addition, the Programme managed to obtain a 12-month extension for its COFUND 1 grant, which takes it to the end of 2014 and is worth around 1.3 million Euros. Accordingly, a total of 22 new EIPODs were recruited in 2012, 14 of whom are women. A major effort in 2012/13 was dedicated to the communication and presentation of the Postdoctoral Programme. The website was completely restructured with a special focus on becoming a more useful tool for incoming postdocs. EMBL also offers a broad range of training initiatives to prepare its postdocs – of which there are currently 255 – for the next stage of their career. Flagship initiatives in this context are the annual 'Career Day', which provides insight into scientific careers outside of academia, and the personalised workshop entitled 'Preparing for the Academic Job Market' that takes place in Heidelberg at least twice a year. The format for the latter is now ready to be implemented at the EMBL outstations to make it even more easily accessible to all postdocs.

In March 2013, the EMBL Advanced Training Centre celebrated its third anniversary, following on from a year with the largest ever Course and Conference Programme hosted at EMBL Heidelberg. During 2012, EMBL Heidelberg organised 20 conferences, 18 courses, and many other seminars, training activities and meetings. In October 2012, EMBL welcomed 442 participants, speakers, organisers, exhibitors and members of the press to its largest conference to date – the EMBO|EMBL Symposium: the Complex Life of mRNA. Also in that month, Nobel Prize winner John Gurdon gave a keynote lecture at the EMBO|EMBL Symposium: Germline – Immortality through Totipotency just days after his award was announced. The EMBL Course Programme at EMBL Heidelberg continued its expansion by offering new courses on state-of-the-art research techniques such as single-cell gene expression analysis and advanced imaging. It was also an eventful year for EMBL Hamburg and Grenoble. Each of the outstations offered two prestigious EMBO courses, and EMBL Hamburg ran two additional courses.

In 2012, the EMBL Course and Conference Programme was generously supported by more than 85 organisations and particular thanks are extended to the 16 members of the Corporate Partnership Programme. Their contributions were used to support courses and conferences on topics ranging from 'Omics and Personalised Health' to 'Microfluidics' and 'Stem Cells in Cancer and Regenerative Medicine', as well as funding 169 fellowships for delegates from over 40 countries who would otherwise have been unable to attend.

The training programme at EMBL-EBI continued to evolve in response to demand from emerging research communities, and provided a curriculum covering the full spectrum of EMBL-EBI activities. 'Train online', the e-learning resource launched in 2011, served over 33 000 unique IP addresses in its first full year. The major data resources now have quick tours; ENA, IntAct and Reactome have long courses; and new course formats were developed based on videos and tutorials from the face-to-face courses. EMBL-EBI entered the realm of 'training informatics' through its involvement in the Innovative Medicines Initiative project EMTRAIN. These online resources complement dynamic face-to-face user-training and knowledge-exchange programmes, reaching almost 5000 scientists on-site and around the world.

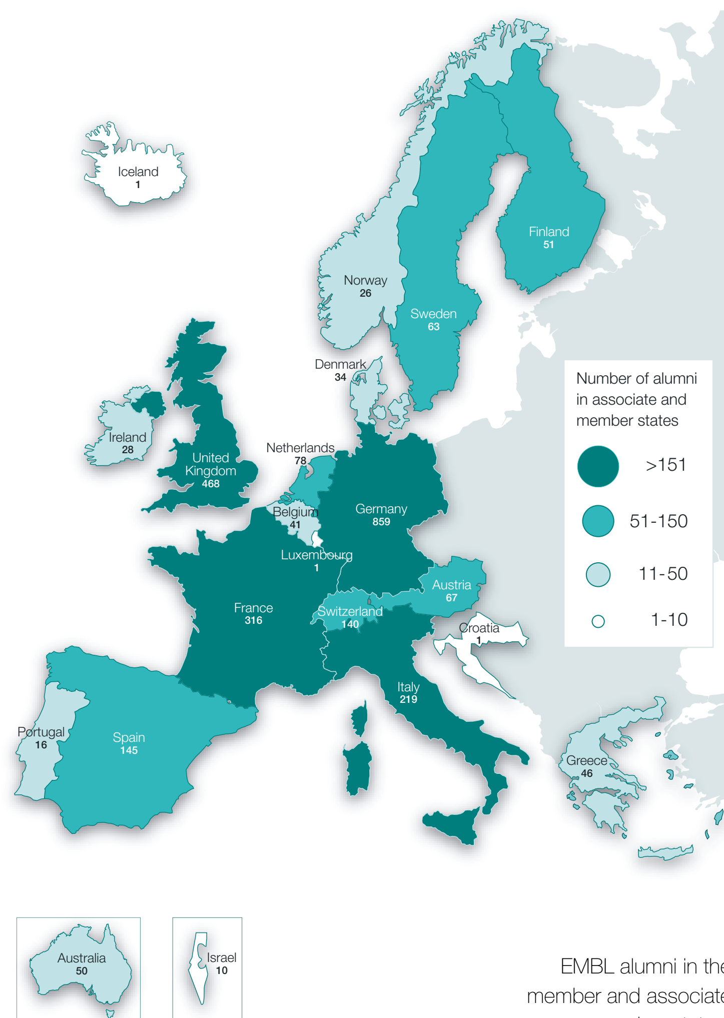
EMBL Alumni

EMBL's global network of alumni has grown to 5901 former scientific, technical and administrative staff, fellows and trainees. More than 80% reside in the EMBL member and associate member states, the majority working in academia (79%) with a growing proportion in industry (11%). 34% hold senior positions as professors, directors, group leaders and managers.

In the past year, 243 former staff joined the EMBL Alumni Association, bringing its membership to a total of 2281 members. From the alumni feedback survey it is clear that these members feel connected to EMBL. In response to the survey results, EMBL introduced two new resources – a searchable Members' Directory and an Alumni Google Map – offering more advanced and user-friendly ways of finding, sorting and reviewing statistics and profiles deposited and updated by alumni. The search categories on the EMBL website were also extended to include alumni, giving them a permanent presence on the EMBL web pages.

Hundreds of current and former staff participated in alumni events in eight countries around the world in 2012. Alumni met with EMBL Senior Scientists in

Worldwide distribution of EMBL alumni.



EMBL alumni in the member and associate member states.

Ireland, Spain, Greece, Germany, the UK, France, Argentina and Italy, exploring areas in which they could better support one another. The Alumni Association Board rewarded and supported the efforts of those local chapters who have met regularly by allocating funds to meetings in Spain and Greece.

EMBL Alumni Association Treasurer, Oscar Martin-Almendral, stepped down after five years of service. He will be succeeded by EMBL Heidelberg and EMBL-EBI alumna Annabel Goulding, who joined EMBL as an internal auditor in 2002 and left as Head of the Pay and Benefits section in Human Resources in 2011. EMBL and the Alumni Association Board welcome her, and thank Oscar for his outstanding service.

Katharina Ribbeck, a former postdoctoral fellow in both the Directors' Research and the Cell Biology and Biophysics Units, was selected as the 2013 recipient of the John Kendrew Young Scientist Award, acknowledging her insight and courage to start a new field of research and her creative and innovative science outreach. The portrait of Katharina, now Professor of Tissue Engineering at the Massachusetts Institute of Technology, Cambridge, USA, will be added to the Klaus-Tschira Auditorium portrait gallery for John Kendrew Award recipients.

Outreach

The EMBL outreach activities aim to promote public awareness and understanding of molecular life sciences. With the help and participation of many scientists, a large number of events were organised by EMBL's Office of Information and Public Affairs (OIPA), the EMBL-EBI Training and External Relations teams, the European Learning Lab for the Life Sciences (ELLS) and the Science and Society Programme. In 2012, regular programmes such as the popular lab visits, the Science and Society lecture series and science communication events were complemented by some exceptional activities.

On 27 April, EMBL Heidelberg opened its doors to the local public. Around 1500 guests came to marvel at the extraordinary architecture of the EMBL Advanced Training Centre, enjoy lectures on EMBL science, discuss hot topics in the rooftop science café, take part in visits to labs specialising in everything from advanced microscopy to zebrafish, and try some hands-on experiments, while children explored science in the kids' lab.

Media coverage of EMBL science reached a new peak with the publication implementing the use of DNA as a reliable data storage medium and the publications related to the ENCODE project, which were covered by a variety of media including TV channels. In November, EMBL Heidelberg hosted a group of 20 German journalists, who went on to report widely on the lab's activities in the national press. Two EMBL scientists, Eric Karsenti and Lars Steinmetz, were amongst the select keynote speakers at the popular European Science Open Forum (ESOF) in Dublin. Eric was also invited to speak at the annual meeting of the American Association for the Advancement of Science (AAAS) in Boston.

As part of the Genome Campus in Hinxton, EMBL-EBI has been enjoying a lively series of debates, film and poetry. Several public engagement events were organised, including an evening of lectures and discussions on the science of sporting success shortly before the 2012 Olympics and a learning lab for science teachers on making sense of biological data. The interactions with the Young Rewired State were quite memorable: four young coders joined EMBL-EBI scientists for a day and a half, during which they wrote a game based on genetic variation.

The 13th EMBL|EMBO Science and Society Conference, 'Biodiversity in the Balance: Causes and Consequences', addressed a number of topics ranging from 'Biodiversity: benefits and the risks of loss' to 'Human impact and visions of sustainability'.

The European Learning Lab for the Life Sciences (ELLS) organised three LearningLab events to train 57

high school teachers on different aspects of biology. In December 2012, the third lecture in the Insight Lectures series, which brings the latest EMBL research to secondary school teachers and students, was delivered by EMBL Heidelberg group leader Eric Karsenti. His presentation on 'TARA Oceans Expedition – a 3-year research cruise across the world's oceans' was followed live by 736 students and their teachers, split between the Klaus-Tschira Auditorium and 13 different European countries thanks to online live-streaming.

During the Open Day many visitors of all ages toured our labs, conducted experiments and discussed cutting-edge science.

Administration

For yet another year EMBL Administration strove to maintain and improve their smooth and streamlined provision of services to all staff. Highlights of administrative activities and projects undertaken include:

- The implementation in March 2013 of the Grant module of an internationally recognised integrated research management system called Converis, in order to be able to manage grants more effectively. In addition, SAP Business Warehouse/Business Objects, the engine behind a more customer-focused internal and external reporting system, has also been positively evaluated and targeted for launch before the end of 2013.
- EMBL Human Resources (HR) is in the process of implementing the European Charter for Researchers and the Code of Conduct for the Recruitment of Researchers. After a multi-layered consultation, EMBL HR is preparing a strategy and action plan, which will consider a broad spectrum of HR activities including ethical and professional aspects, recruitment, working conditions, social security and training.
- To welcome new staff and to provide them with the necessary information for a quick and efficient start in their new position, new initiatives have been put in place at both EMBL Heidelberg and EMBL-EBI. Such activities include improved employment information on the intranet, engaging and useful presentations by staff from different areas of the lab as well as the possibility to connect with other newcomers and existing staff.
- In close collaboration with the local authorities, a master plan was signed by the Mayor of the City of Heidelberg, Eckart Würzner, and EMBL Director General Iain Mattaj, to provide EMBL Heidelberg with some space to expand in the future. Additional land has been obtained from the city of Heidelberg with the agreement that EMBL will not construct any buildings in an area close to a nature preserve.

Integration of European Research

Member state relations

EMBL has continued to strengthen its relationship with its member states and engages in a continuous dialogue to gather feedback regarding the needs of their scientific communities. Meetings have been organised to make sure good relationships are maintained when leadership changes. For example, Baden-Württemberg Minister for Research Theresia Bauer visited EMBL Heidelberg, whereas the Spanish State Secretary for Research, Carmen Vela Olmo, visited EMBL Grenoble. In addition, Director General Iain Mattaj, new Head of EMBL Monterotondo Philip Avner and Director of International Relations Silke Schumacher had several meetings with the new leadership of the Italian Research Council (CNR), including its President Luigi Nicolais, and the leadership of the Italian Ministry of Research.

EMBL has been supporting science in the member states that face financial difficulties. The Director General Iain Mattaj visited Greece to meet with Vasilis Maglaris, Secretary General for Research and Technology. A mini-symposium was organised to offer opportunities for boosting research and collaborations. Several visits to Portugal took place, including a meeting with Miguel Seabra, President of the national funding agency Fundação para a Ciência e Tecnologia (FCT) and visits to several research centres. This resulted in a conference on 'Molecular Biology in Portugal and EMBL', that will be organised in July 2013 at the University of Lisbon.

Director General Iain Mattaj visited Iceland and Israel for the first time. In Iceland the EMBL delegation met, among others, the Minister for Science, Katrin Jakobsdottir, and the Rector of the University of Iceland,

The Baden-Württemberg Minister for Research, Theresia Bauer, visiting EMBL Heidelberg.

Kristin Ingolfsdottir. The possibility of Iceland joining the Nordic EMBL Partnership was discussed. In Israel, the EMBL delegation met with Minister for Science Rabbi Daniel Hershkowitz and Chief Scientist Ehud Gazit and visited several major universities and research centres in the country, delivering a number of scientific talks as well as talks about EMBL.

During the visit of the Director General and several EMBL group leaders to Luxembourg, an agreement on closer co-operation was achieved. In particular, the Luxembourg government has decided to financially support scientific collaborations with EMBL. A call for joint projects between EMBL faculty and national researchers was thus launched in December 2012.

Australia's associate membership is up for renewal in 2015, and will be preceded by a comprehensive evaluation process that started in the autumn of 2012 with a review conducted in Australia by an international expert panel. The panel commended the associate membership, its development so far, and its future scientific potential. The review process will be completed in May 2013 following the recommendations of EMBL's Scientific Advisory Committee to EMBL Council on the associate membership. To further capitalise on the relationship, EMBL Australia is exploring ways to transfer elements of the highly successful EMBL PhD Programme to Australia. For example, the newly established EMBL Australia Travel Grant allows Australian students to attend the annual EMBL PhD Symposium in Heidelberg.

One of the long-term goals that EMBL has set itself is to encourage all European countries to join EMBL. With this in mind, in November 2012 the EMBL Council adopted a Policy on Prospect Member States. The policy enables all European countries that are not yet EMBL member states to start engaging with EMBL by gaining access to services and facilities, with the possibility of becoming a member state after three years of reduced membership contributions.

EMBL has continued to strengthen its already close ties with the Czech Republic. As a result, the Czech Republic's Minister for Education, Petr Fiala, formally requested that the country join the lab's family of member states. EMBL has welcomed the request, with the proposal to be endorsed at this summer's Council meeting. In addition, a range of initiatives and visits have aimed to stimulate research connections with the country, including joint workshops and a Memorandum of Understanding (MoU) with the Central European Institute of Technology (CEITEC) in Brno and with the Biotechnology and Biomedicine Centre of the Academy of Sciences and Charles University (BIOCEV) in Vestec.

EMBL also reinforced co-operation with the Russian Federation. In October 2012, EMBL and the Kurchatov Institute agreed on the transfer of DORIS beamlines. EMBL and the Russian Federation for Basic Research jointly approved and supported six collaborative projects between researchers from EMBL and Russia.

During 2012 and 2013, EMBL also established contacts with the Slovakian scientific community and key decision makers in that country. Besides a scientific workshop at the leading Slovak Institute of Virology, meetings with the heads of several life science institutes and with the President of the Academy of Sciences, Jaromir Pastorek, took place. The Slovak scientific community is highly supportive of Slovakia becoming an EMBL prospect member state. The benefits of membership were also presented to the leadership of the Ministry of Science, and links with this country are expected to strengthen further.

During a visit of EMBL Director General Iain Mattaj to Argentina in 2012, he and the Minister of Science, Technology and Productive Innovation for the Argentine Republic, Lino Barañao, signed an MoU with the objective of enhancing long-term scientific co-operation between EMBL and Argentina. To further discuss the implementation of the MoU and Argentina's scientific ties to EMBL, Minister Barañao and Mrs. Águeda Menvielle, Director of International Relations in the Ministry of Science, visited EMBL Heidelberg in May 2013.

In January 2013, Director General Iain Mattaj, together with other representatives of EMBL and EMBO, visited the headquarters of the Scientific and Technological Research Council of Turkey (TÜBİTAK) to take part in an 'information day', with the goal of furthering links between the organisations and Turkey.

EMBL Partnerships

EMBL engages in a few selective, formal partnerships with scientific institutions in its member states. The scope of these partnerships is to promote excellence in life science research throughout Europe and to disseminate the successful EMBL organisational model, which is based on international recruitment, high staff

turnover and regular external reviews, throughout the member states.

Perspectives in Translational Medicine Conference

In order to leverage the successful partnership model and link its partner institutes working in the field of molecular medicine, EMBL and the Centre for Genomic Regulation (CRG) in Spain organised a joint conference on the topic of 'Perspectives in Translational Medicine' in September 2012 in Barcelona. The aim of the event was to offer a forum for bringing together the expertise of the different partners and to facilitate the building of a larger European network. The conference brought together 150 researchers from EMBL, the EMBL-CRG Partnership Unit, the Molecular Medicine Partnership Unit, and the Nordic EMBL Partnership Institutes. The programme featured more than 40 presentations by directors and group leaders from the institutes and included a range of networking and poster sessions.

Nordic EMBL Partnership for Molecular Medicine

On 5 March 2013 the Nordic EMBL Partnership for Molecular Medicine celebrated two important milestones: the renewal of the partnership agreement for an extended period of 10 years; and the expansion of the Nordic EMBL network with the official opening of the Danish Research Institute of Translational Neuroscience (DANDRITE) at Aarhus University, which, following a competitive national selection process, will become the Partnership's Danish node. This node further strengthens the Nordic EMBL Partnership for Molecular Medicine, which was initiated in 2007 and has since emerged as a strategic player in its field. The establishment of the Danish node was made possible through the Lundbeck Foundation, which funds the new centre jointly with Aarhus University. The DANDRITE Centre currently consists of three core groups, and has initiated a first international recruitment procedure to attract two additional young group leaders.

From left to right: Rector Lauritz B. Holm-Nielsen from Aarhus University, Iain Mattaj and Poul Nissen, Director of DANDRITE and Professor of Protein Biochemistry.

Molecular Medicine Partnership Unit

The Molecular Medicine Partnership Unit (MMPU) review took place in April 2012 and its research performance was commended by a panel of external experts. In addition, the outcome of the Excellence Initiative marks another success for the MMPU. As part of the University of Heidelberg's proposal for a 'Heidelberg Research Centre for Molecular Medicine', MMPU was awarded 1.5 million Euros over the next five years. MMPU will use the grant to fund MD and PhD positions in the context of a clinician-scientist training programme.

EMBL-CRG Partnership Unit for Systems Biology

The EMBL Scientific Advisory Committee (SAC) reviewed the proposal for establishing a new outstation in Spain, focussed on Tissue Biology and Disease Modelling. The SAC concluded that the EMBL-CRG Partnership Unit would significantly enhance and expand current EMBL efforts towards understanding human disease through the development of novel human model systems, new methodologies for three-dimensional mesoscale microscopy and the modelling of complex multicellular systems. In November 2012, the EMBL Council approved the development of a detailed financial plan and draft host site agreement for future assessment and final decision.

Partnership with EMBL Australia

In February 2013, EMBL signed a partnership agreement with Monash University, acting on behalf of the EMBL Australia Participants. The agreement formally frames the existing scientific collaboration between EMBL and research institutes in Australia, and aims at further enhancing co-operation ties and co-ordination between the partners. The EMBL Australia Partnership Laboratory Network opened a second node at the University of Queensland to underpin the EMBL Australia Bioinformatics Resource. The University of Queensland appointed Graham Cameron, the former EMBL-EBI Associate Director, as the Director of EMBL Australia Bioinformatics Resource (Braembl). In addition, a third node will be opened at the South Australian Health Research Institute (SAHMRI), with support from SAHMRI's parent institutions: the University of South Australia, the University of Adelaide and Flinders University. The building that will house the institute is currently under construction and three group leaders are being recruited in bioinformatics and computational biology with a focus on themes such as cancer, infection, immunity and brain research.

EBI-Wellcome Trust Sanger Institute Formal Collaboration Agreement

On 17 July 2012, EMBL-EBI and the Wellcome Trust Sanger Institute signed an agreement to formalise the scientific collaboration that the two institutes have enjoyed informally over many years in the areas of services, research, training, technology, public engagement and campus management. The co-operation agreement paves the way to help manage some complex interactions and appointments of joint personnel between the institutes.

European Research Infrastructures

In the past year, EIROforum, a partnership between eight of Europe's largest inter-governmental scientific research organisations, celebrated its 10th anniversary. At the anniversary General Assembly, the Directors General were joined by Anne Glover, Chief Scientific Advisor to the President of the European Commission (EC), and Anneli Pauli, Deputy Director General for Research and Innovation at the EC. As well as celebrating the formation of EIROforum, the members looked forward to the challenges facing European science and innovation. Ahead of the European budget summit in February 2013, the Directors General of the eight EIROforum member organisations wrote an open letter to the Heads of State of the EU Member States, the Presidents of the European Council, the European Parliament and the EC in which they stressed the importance of investments in our scientific resources – both human and technical – at a time when a return to growth is the most pressing policy priority across Europe. In addition, EIROforum published a position paper on Scientific Instrumentation for the EU Framework Programme Horizon 2020 elaborating on specific recommendations from a highly focussed perspective regarding the needs of research infrastructures based in Europe and their respective scientific user communities.

In parallel, considerable efforts were undertaken to raise awareness within EU institutions, on behalf of ELIXIR and the other ESFRI biological and medical science (BMS) projects. This included the preparation of a joint position paper for the ESFRI BMS Research Infrastructures and direct meetings with key opinion formers within the EU, with the aim of increasing the budget allocation for research infrastructure under Horizon 2020.

ELIXIR

ELIXIR, the pan-European research infrastructure for biological information, officially entered its implementation phase in January 2013. ELIXIR will provide the facilities necessary for life science researchers to make the most of our rapidly growing store of information about living systems. A total of 15 countries have signed the ELIXIR Memorandum of Understanding (MoU), which outlines the establishment of ELIXIR based on an international consortium agreement and using EMBL as a host of the new infrastructure. Based on the MoU, an interim governance structure, including a decision-making body (Interim ELIXIR Board) and a Scientific Advisory Board, was established. The Interim ELIXIR Board approved a three-year interim financial framework for the ELIXIR hub including annual interim budgets for 2012 and 2013. The Interim Board is currently working towards the ELIXIR Consortium Agreement, which will specify issues such as rules on membership, member obligations, ELIXIR's governance structure, financial issues and liability. The Interim ELIXIR Board appointed a founding Director, Niklas Blomberg, who took up his position in May 2013. Blomberg previously worked as Principal Scientist and Team Leader in Computational Chemistry and Computational Biology at AstraZeneca R&D in Mölndal, Sweden.

The current signatories of the ELIXIR MoU have submitted ELIXIR Node applications and these were considered by ELIXIR's Interim Scientific Advisory Board at its first meeting in December 2012. As ELIXIR transitions from the preparatory to the implementation phase, technical work is underway. Five technical pilot projects have been initiated to act as test beds for integrating services between the Nodes and the Hub.

ELIXIR was showcased at scientific conferences in Denmark, Switzerland, Spain, France and Ireland. Information was provided to scientific delegations from Russia and India, representatives of funding bodies, senior government officials, EU institutions and numerous journalists and film crews from Europe and beyond.

BioMedBridges

The FP7 project BioMedBridges, which started in January 2012, is co-ordinated by EMBL-EBI on behalf of ELIXIR. It will deliver data interoperability to bring together ten new BMS research infrastructures on the ESFRI roadmap ranging from marine biology to the clinical domain, via structural biology and translational research. Tools developed within the project will be tested in five use cases and will become shared e-infrastructure to allow interoperability between data and services in the biological, medical, translational and clinical domains. The first annual general meeting of the consortium took place in March 2013. More than 70 delegates discussed the project progress and issues of ethical and secure access to sensitive data, which will be regulated based on the project's

approved Ethical Governance Framework. One of the main pillars of the BioMedBridges project is training, initially for staff at participating research infrastructures and later for users of the new tools. The first training workshop on the use of ontologies in data integration and interoperability with semantic web frameworks took place at EMBL-EBI in March 2013.

Euro-BioImaging

The Euro-BioImaging research infrastructure project co-ordinated by EMBL is in the final year of its three-year FP7-funded preparatory phase. Euro-BioImaging will consist of a set of geographically distributed but strongly interlinked imaging facilities (Euro-BioImaging Nodes), which will provide open access to cutting-edge imaging technology for every European life scientist, flanked by appropriate services and training through a co-ordinating and supporting hub. During the first two years of its preparatory phase, Euro-BioImaging has had major successes:

- Euro-BioImaging drove the establishment of national imaging communities in 20 European countries, which are now co-ordinating imaging infrastructure activities at the national level to enable better use of funds and capacities by speaking with a single voice. Twelve of these countries have already put imaging infrastructure as a high priority on their national infrastructure roadmaps.
- The consortium carried out a Europe-wide survey, which exposed both a strong demand of researchers to access advanced imaging technologies and major gaps in the availability of such technologies.
- In 2012, the project performed a test-run of operating its open-access infrastructure. 110 researchers visited 41 European facilities, successfully carried out their imaging experiments and returned home with high-quality image data.
- To ensure close collaboration with the European imaging industry, Euro-BioImaging established its Industry Board, which already comprises more than 50 leading companies.
- Euro-BioImaging has established active international partnerships with the Australian and Indian national imaging infrastructure organisations, fostering training and exchange of staff, knowledge and best operational practice in imaging infrastructure.

Concrete steps towards the construction of the future Euro-BioImaging Infrastructure have now started, and the first Open Call for Euro-BioImaging Nodes launched in January 2013. These activities place Euro-BioImaging on the fast track to implementation, foreseen for 2014.

Relations with the European Commission

From left to right: Silke Schumacher, Iain Mattaj and Anne Glover.

EMBL and the European Commission (EC) held their annual meeting on 11 October 2012, as agreed in the MoU signed in 2011. The meeting, co-chaired by Director General Iain Mattaj and Rita Lecbychova, Head of Unit for Joint Programming of the Directorate-General for Research and Innovation, covered the exchange of recent developments in the EC and EMBL. A range of issues was addressed, from Horizon 2020 to the Charter and Code for Researchers. In addition, EMBL Director of International Relations, Silke Schumacher, met with Laurent Bochereau, Acting Director of the EC RTD Directorate D, International Cooperation. The meeting was an excellent opportunity to continue the discussion of EU international cooperation in research, and to provide input into the EC's communication on 'Enhancing and focusing EU international cooperation in research and innovation'. In addition, EMBL provided input into the European Research Area (ERA) survey concerning the ERA priorities in research funding and research organisations. EMBL Heidelberg also hosted visits from Ana Arana Antelo, Head of Unit for Research Infrastructures at the EC and Anne Glover, Chief Scientific Advisor to the President of the EC.

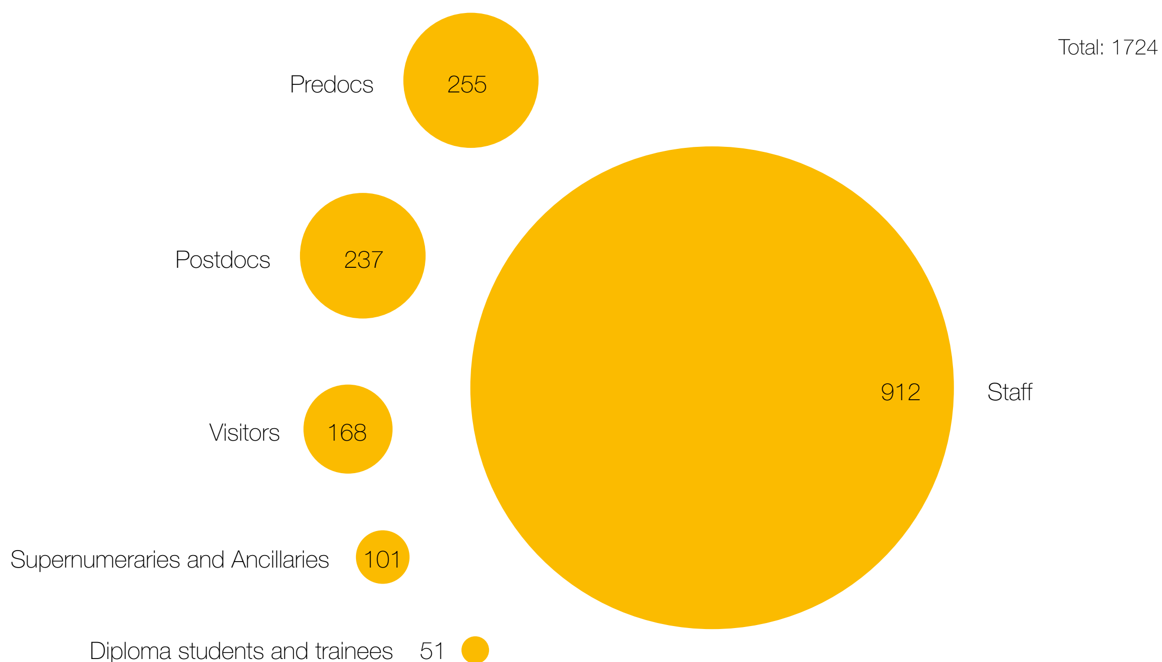
The EC has expressed its appreciation of EMBL's position paper on Horizon 2020 – the proposed framework for research and innovation that will run between 2014 and 2020. In a letter to Director General Iain Mattaj, Commissioner for Research, Innovation and Science, Máire Geoghegan-Quinn, applauded EMBL's call for an integrated approach, steps towards simplification and the need to enhance the proposed budget. Director General Iain Mattaj and Director of International Relations Silke Schumacher also met with rapporteurs of the European Parliament for Horizon 2020 to convey EMBL's main messages regarding the future EU funding programme.

EMBL also supported the petition, launched by the Initiative for Science in Europe (ISE), asking to protect the EU Horizon 2020 research and innovation budget from cuts. This resulted in meetings between a delegation of science representatives, including Nobel laureates, Herman van Rompuy, President of the European Council, Martin Schulz, President of the European Parliament, and José Manuel Barroso, President of the European Commission.

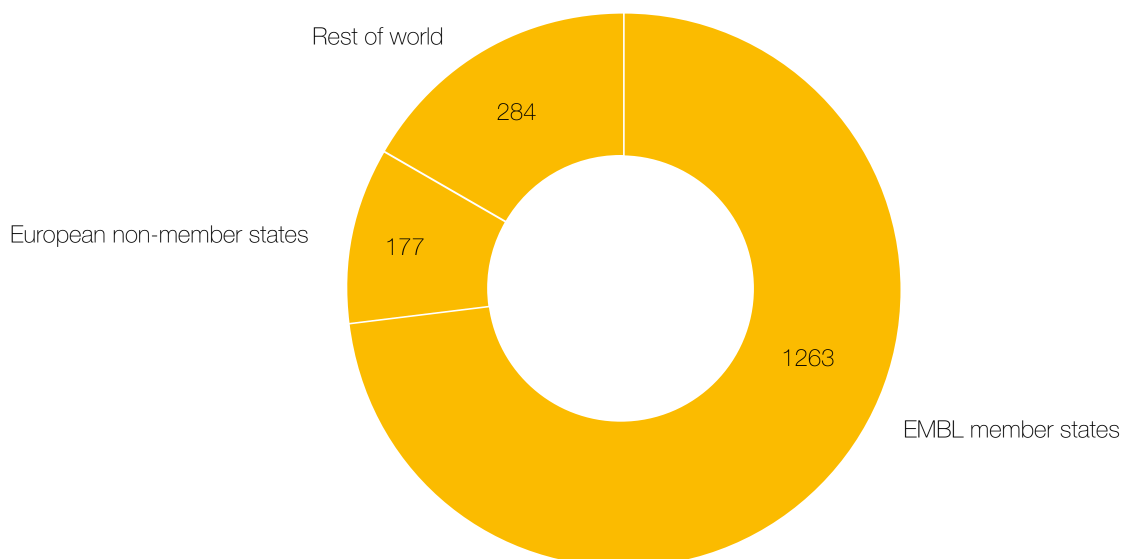
Personnel statistics

On 31 December 2012, 1724 people, including visitors, from more than 60 nations were employed by EMBL.

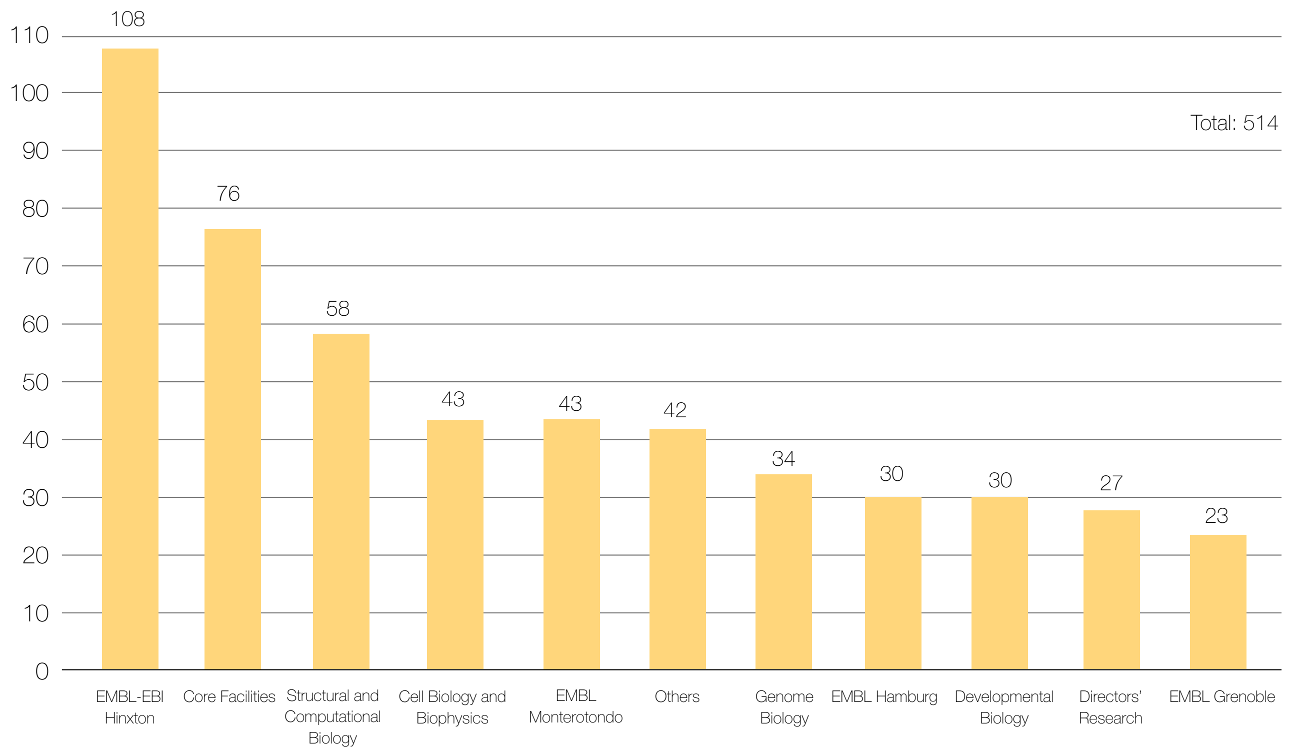
Personnel on 31 December 2012



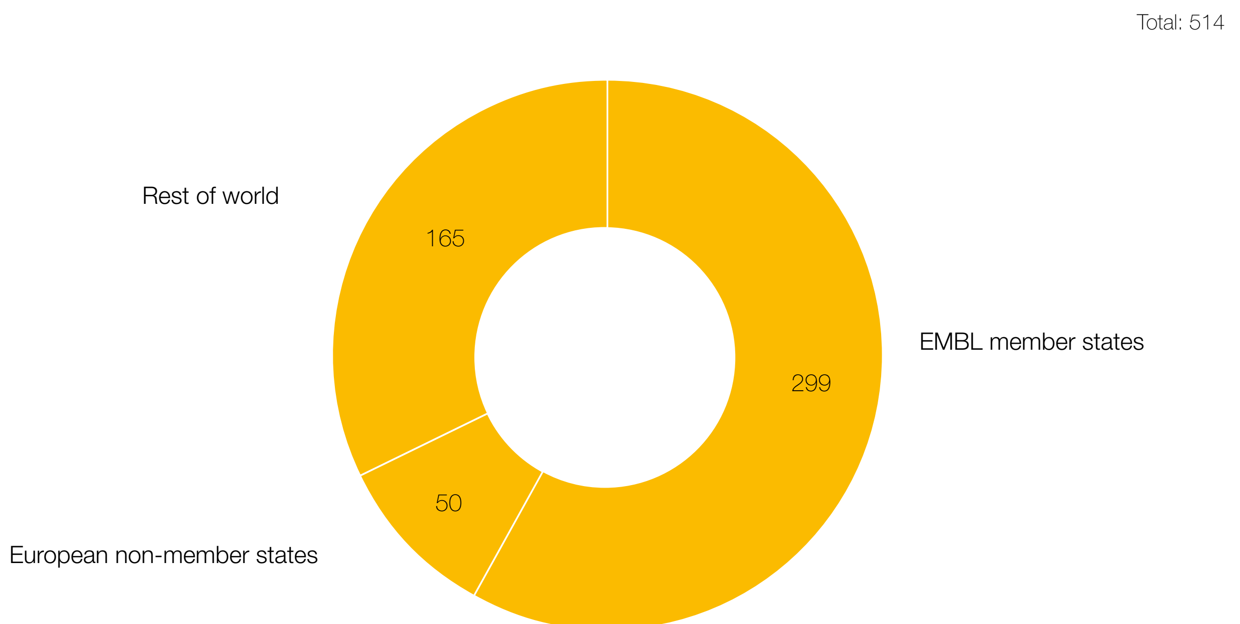
Staff Nationalities



Visitors to EMBL Units during 2012



Visitors' Nationalities



Financial report

Income/expenditure statement

Income	2012 €000	2012 %	2011 €000	2011 %
Member state contributions				
Ordinary contributions	94,842		92,927	
Special contributions	-		23	
Associate contributions	501		752	
Additional Contribution from the UK Government	7,052		-	
Germany – contribution to infrastructure work	680		600	
Internal tax	23,303		21,883	
External grant funding	36,613		37,849	
Other external funding	1,612		2,236	
Other income	16,226		15,068	
Total income	**180,830**		**171,338**	
Expenditure				
Staff costs	105,135		96,002	
Operating costs	56,884		58,969	
Equipment expenditure incl. Depreciation	16,686		13,058	
Total expenditure	**178,705**		**168,029**	
Surplus (deficit) for the year	2,125		3,309	

External grant funding

	2012 €000	2012 %	2011 €000	2011 %
ANR	363	1.0	223	0.6
BBSRC	2,156	5.9	1,468	3.9
BMBF	2,447	6.7	3,461	9.1
BW	103	0.3	57	0.2
DFG	1,362	3.7	1,816	4.8
EC	14,017	38.3	13,024	34.4
ERC	1,821	5.0	330	0.9
FINOVI	106	0.3	2	0.0
HFSPO	212	0.6	229	0.6
HUMBOLDT	65	0.2	17	0.0
MRC	153	0.4	173	0.5
NIH	7,192	19.6	8,797	23.2
VW Foundation	(1)	0.0	92	0.2
Wellcome Trust	4,455	12.2	5,014	13.2
Others	2,163	5.9	3,146	8.3
	36,613	**100.0**	**37,849**	**100**

Other external funding

	2012 €000	2012 %	2011 €000	2011 %
EMBL-EBI industry support	703	43.6	1,642	73.4
Elixir	558	34.6	-	0.0
Other external funding	352	21.8	594	26.6
	1,612	**100.0**	**2,236**	**100.0**

Member state contributions

Ordinary contributions	Contributions 2012 €000	Contributions 2012 %	Contributions 2011 €000	Contributions 2011 %	Pension contribution 2012 €000	Pension contribution 2011 €000
Austria	2,068	2.2	2,026	2.2	35	31
Belgium	2,551	2.7	2,500	2.7	43	38
Croatia	285	0.3	279	0.3	5	4
Denmark	1,650	1.7	1,617	1.7	28	24
Finland	1,318	1.4	1,292	1.4	22	19
France	15,070	15.9	14,766	15.9	255	223
Germany	19,338	20.4	18,948	20.4	327	286
Greece	1,717	1.8	1,682	1.8	29	25
Iceland	85	0.1	83	0.1	1	1
Ireland	1,224	1.3	1,199	1.3	21	18
Israel	986	1.0	966	1.0	17	15
Italy	11,599	12.2	11,365	12.2	196	171
Luxembourg	218	0.2	214	0.2	4	3
Netherlands	4,401	4.6	4,312	4.6	74	65
Norway	2,276	2.4	2,230	2.4	39	34
Portugal	1,100	1.2	1,078	1.2	19	16
Spain	7,815	8.2	7,657	8.2	132	116
Sweden	2,608	2.8	2,555	2.7	44	39
Switzerland	2,580	2.7	2,528	2.7	44	38
United Kingdom	15,952	16.8	15,630	16.8	270	236
	94,842	**100.0**	**92,927**	**100.0**	**1,605**	**1,402**

Special contributions

	2012 €000	2011 €000	Pension 2011 €000
Croatia	-	23	-

Associate contributions

	2012 €000	2011 €000
Australia	501	752

Additional contributions

	2012 €000	2011 €000
Additional contribution from the UK Government	7,052	-
Germany Contribution to Infrastructure work	680	600

EMBL budget 2012: € 181 million

2012/2013 Reviews of EMBL Scientific Units

To ensure that its research and service activities continue to operate at the cutting edge, EMBL regularly submits them to stringent external reviews. Research and Service Units are evaluated every four years by members of the Scientific Advisory Committee and additional international experts. The following section features summaries of the scientific reviews that have taken place in the past year and presents the Director General's responses to the review reports.

EMBL Genome Biology Unit Review

On 9 and 10 May 2012 the review of the Genome Biology Unit took place in Heidelberg. Tom Muir, from Princeton University, USA, chaired the panel of fourteen reviewers, six of whom, including Tom, were members of EMBL's Scientific Advisory Committee (SAC).

Evaluation Summary

The research being performed within the Genome Biology Unit is of the highest calibre, comparing favourably with the best work of this type being executed at other leading biomedical research institutes around the world. The Unit comprises an outstanding group of young scientists whose broad expertise collectively creates a fertile environment in which to carry out the interdisciplinary research now at the heart of the Unit's research endeavours.

The Unit has undergone substantial changes in leadership, composition and scientific focus since the last review. It has been renamed Genome Biology rather than Gene Expression to reflect the adoption of a more systems-wide approach to studying the link between genotype and phenotype, employing a variety of experimental and computational methods. Eileen Furlong and Lars Steinmetz were appointed joint heads of the Unit a year into the current review period. Although still early in their stewardship of the Unit, it is already evident that they make an effective managerial team; they have done an admirable job running the Unit and recruiting dynamic new Group Leaders, whilst continuing to drive their own research programmes to higher levels. They have deepened the Unit's expertise in bioinformatics and proteomics, as well as bringing in new capabilities in the area of high throughput biology/chemistry using state-of-the-art microfluidics devices. Their research vision for the Unit, embracing a multidisciplinary set of approaches for systems-wide genomic interrogations, was highly praised by the review committee.

The overall scientific output of the Unit over the last four years has been exceptional. The Unit has continued to make key contributions on how to optimally acquire, interrogate and exploit large-scale biological data sets for the purposes of discovering the mechanisms underlying phenotypic variation in biology. There have been a number of breakthroughs in areas such as transcriptome regulation and analysis, comparative genomics, epigenomics and posttranscriptional regulation.

As might be expected given the nature of the research, the Unit is exceedingly collaborative both within EMBL and beyond. Complementing their research with outreach, Genome Biology Group Leaders, and the Associate Director, have organized an impressively large number of international meetings, as well as practical training courses in all the technical areas represented in the Unit. These activities, along with a host of additional scientific service roles within Europe, further underlines the important leadership role the Unit plays in the biomedical community.

The overall performance of the Unit and its leadership were rated as outstanding.

Response to the Panel's Recommendations

The review panel stressed the Unit's dependence on computational infrastructure for the analysis and storage of the vast data quantities produced by 'omics' techniques. They recommend improvement in the local computational facility in terms of personnel and hardware as well as improved access to the EBI/Sanger, the site where much data presently resides. They further recommended an annual review of the usage and needs and the definition of a medium term strategy in this area. EMBL closely monitors and constantly adapts its computational infrastructure to the changing needs of researchers. To identify these needs regular meetings between representatives of the Heads of Units in Heidelberg and IT Services take place to discuss future plans for IT infrastructure. The recommendations will be discussed at the next meeting to which the two joint Heads of the Genome Biology Unit will be invited. It is however necessary to point out that an upgrade in connectivity of the EMBL Heidelberg site as suggested in the report will be very costly and therefore requires careful consideration.

The panel pointed out that several of the groups in the Unit have initiated programmes geared towards discovering small molecule modulators of cellular function. While the rationale behind these efforts is primarily the identification of pharmacological probes for dissecting basic biology, undoubtedly some of these agents will also be viewed as potentially therapeutic leads, raising the question of whether to pursue more translational research efforts. The review panel suggests it will be helpful for group leaders to have the opportunity to interact with representatives of the pharmaceutical/biotech industry before taking a decision on whether to embark on such projects. We welcome this suggestion and will develop suitable formats to increase the interaction with pharmaceutical and biotech industries.

The panel suggested metabolomics and evolution as two potential areas of future recruitment. We take note of this suggestion and will keep it in mind for future rounds of recruitment. However, the primary criteria for recruitment will always be to get the best possible candidates that complement the Unit.

The panel passed on several suggestions from the fellows in the Unit including specific suggestions that would aim to broaden training opportunities, particularly in computation, and prepare fellows better for the time of their departure from EMBL. These suggestions have been passed on to the faculty of the Unit and to those in charge of EMBL's internal training programmes for action.

Professor Iain W. Mattaj, FRS
Director General
4 June 2012

EMBL Monterotondo Review

The review of EMBL Monterotondo, took place on 25 and 26 October 2012. Eight international experts formed the panel and the Chair of EMBL's Scientific Advisory Committee (SAC), Sandra Schmid from the University of Texas Southwestern Medical Center in Dallas, US chaired the panel.

Evaluation Summary

The review took place following the Unit's first major wave of turnover. Three groups have departed since the last review in 2008, including that of the former Head of Unit Nadia Rosenthal. Two new group leaders were recruited, Martin Jechlinger in 2010, Christophe Lancrin in 2011, and a new Head of Unit, Philip Avner, took up his position in May 2012 and moved his laboratory to Monterotondo in September. Under the leadership of Nadia Rosenthal from 2001 to 2011, Monterotondo has matured to its present size and capacity (six group leaders and two staff scientists) and proven its ability to attract outstanding young scientists. I join the Panel in recognising the significant contributions of Nadia Rosenthal in developing a vibrant research unit, in recruiting outstanding young scientists and in the provision of opportunities for past group leaders to develop impactful programmes and to transition into leadership positions throughout Europe.

The productivity of the relatively small unit has been outstanding. Since the last review period roughly 100 peer-refereed publications have been published describing work done in the Unit, 30 of which are in high impact journals, including Cell, Nature and Science. Fourteen additional such high-impact papers were published by group leaders before their recruitment to EMBL-MR and attest to the Unit's ability to attract outstanding scientists. The success of Rocio Sotillo, Donal O'Carroll and Martin Jechlinger in securing prestigious external grant awards also attests to the quality of young scientists recently recruited. The Panel identified as research highlights in the review period the accomplishments of Cornelius Gross in demonstrating a role for microglial-mediated pruning in synaptic maturation and of Donal O'Carroll, in a collaboration with the group of Anton Enright at EMBL-EBI, in establishing the role of Piwi proteins and piRNA in silencing transposons in the mammalian germ line. Numerous collaborations between Monterotondo and other EMBL Units in Heidelberg, Hamburg and the European Bioinformatics Institute (EBI) were noted by the Panel and the positive role of EMBL Interdisciplinary Postdoctoral (EIPOD) fellows in linking the groups and research between EMBL Units was positively remarked on.

As for other EMBL units, productivity is enhanced by effective core facilities. Monterotondo's core facilities include transgenic mice, histology, gene expression, fluorescence-activated cell sorting (FACS) and mouse phenotyping. The particular strengths of the transgenic mouse and gene expression facilities were recognised by the Panel. These facilities also serve the broader research community. A new microscopy core has been added to meet the changing needs of groups, but the panel recommended that this core will need further expansion and added capabilities as multi-colour live cell microscopy becomes increasingly important. Similarly, the FACS core will need improvements to keep up with critical technologies needed on site.

Dr Avner discussed his future plans for the Unit with the Panel. His plans include the need to more tightly focus the Unit's research efforts around two areas, neurobiology of behaviour and epigenetic regulation of developmental biology, which are existing strengths. The increased focus aims to create greater critical mass and improve the visibility and impact of the Unit in the international community and will in the longer term require an increase in the total number of groups.

Although much progress has been made, the Unit has yet to fully accomplish its goal of being recognised as an international centre of excellence for mouse genetics and mammalian biology. Nevertheless, the overall performance of the Unit and its leadership were rated as excellent.

Response to the Panel's Recommendations

The Head of Unit's assessment of the need to further focus the scientific directions of the Unit around the areas of neurobiology and epigenetic regulation of development was supported by the Panel, which also recommended considering recruitment in these areas to link them and provide further opportunities for synergy. This is an issue that I have discussed in detail with Philip Avner on numerous occasions and developing a tighter research focus is a goal we share.

The Panel raised issues concerning the level of independence of staff scientist appointments at the Unit. Staff scientists at EMBL are normally hired as part of a research group headed by a group leader. Furthermore, EMBL units are rather small and their composition is a matter of the scientific strategy of the Head of Unit. The Panel commented that funding availability should not be allowed to unbalance the research of an EMBL Unit in one particular direction. I agree with this recommendation and, while continuing to provide agreed support to the existing staff scientists at Monterotondo, will work with all EMBL Heads of Unit to ensure uniform treatment of these positions across the Laboratory.

A shift towards a culture of more external funding was recommended by the Panel. For new group leaders, this should be accompanied by a structured mentorship in the preparation for and timing of grant applications. Group leaders should be encouraged to take advantage of ERC funding opportunities at all stages, including early career, consolidation and senior awards. Growth in group size beyond five or six people for more senior group leaders should be dependent on external funding, as it is in other parts of EMBL. Again, I agree with the principles behind this recommendation and fully support it in relation to more senior group leaders. I however note that, because of EMBL's fixed term contract system, care needs to be taken to avoid overloading recently-recruited group and team leaders too early in their EMBL careers.

In order to increase the visibility of the Unit, the Panel recommended that group leaders and junior colleagues (postdoctoral fellows and PhD students) should participate in international meetings and it emphasised that it is the responsibility of the individual group leaders to provide the resources to facilitate this within their own budgets. One possibility suggested by the postdoctoral fellows for increasing the Unit's visibility in the local scientific community are joint meetings/career days co-organised with and held at a local university. These are valuable recommendations that are being or will be implemented throughout the Unit.

The Panel recommended that access to journals, especially in the area of neuroscience, needs to be available. The problem with access arose through the termination of the subscriptions of an individual researcher in the Unit, which meant a series of journals were no longer available. The individual concerned has been asked to discuss the requirements with the EMBL librarian.

Finally, the Panel met with the fellows in the Unit and passed on suggestions from them related to administrative aspects of their training and for enhancing training opportunities, particularly in increasing interactions with the local and international scientific community. These suggestions have been passed on to the faculty of the Unit and to those in charge of EMBL's internal training programmes for action.

Professor Iain W. Mattaj, FRS
Director General
21 November 2012

EMBL Grenoble Review

From 25 to 27 February 2013 the review of EMBL Grenoble took place. Ten international experts, including two members of EMBL's Scientific Advisory Committee (SAC), formed the Review Panel. The panel was chaired by Andrea Musacchio from the Max Planck Institute of Molecular Physiology in Dortmund, Germany.

Evaluation Summary

The research, service and technology development activities of EMBL Grenoble are rated as outstanding. All groups and teams at the Outstation work on problems of great scientific interest and have been able to bring about remarkable advances in structural and molecular cell biology. The panel singled out some of the research and beamline environment development work for special praise.

Under the leadership of the Head of Unit Stephen Cusack the Outstation continues to operate at the cutting edge of structural biology, continuing its long track record of making enabling technical contributions that benefit the entire European structural biology community and beyond. The panel highlighted the key role of different groups and teams at EMBL Grenoble in the development of new technologies such as the MultiBac recombinant expression system; the ESPRIT construct screening technology; the new CrystalDirect system for crystal harvesting and mounting; a gripper for the EMBL MD2 and MD3 diffractometers to handle crystallisation plates for in situ diffraction; and the new MD3 diffractometer. The panel was highly impressed with the level of the services that are currently being offered and by plans for future developments. The services are in part self-sustained, for instance through incorporation in EU-funded programs like I3 P-Cube and in its follow-up BioStruct-X. Such programmes are successful and oversubscribed among external users, and significantly increase the visibility of the Outstation.

The Outstation has strong and durable liaisons with local institutions and the current relationship with ESRF is very strong and effective. The last three years have seen significant changes in the beamlines. A rapidly changing environment at the ESRF has provided multiple opportunities for growth and development of the EMBL-associated resources. More changes are upcoming as the ESRF is preparing for a second consecutive major upgrade in 2020, which will allow it to keep pace with developments at other synchrotron sources. Another potentially exciting development is the creation of an EMBL-CNRS-CEA-UJF joint Unité Mixte de Service, with founding director Darren Hart, to structure the EMBL and local French platforms for local, national and EU (INSTRUCT) access.

Since the last review there have been only minor changes in the composition of the senior scientific and technical staff, with no turnover of group or team leaders. This has resulted in a very significant increase in the overall count of externally-funded staff members, with three investigators having been awarded an ERC grant. However, in the next few years there will be considerable change due to staff turnover as a result of retirements and ends of contract. In order to remain competitive, the requirements posed by recent developments in the application of a complex array of hybrid approaches will have to be reflected in adequate future recruitment. In addition, the two critical senior members of the Outstation are approaching retirement age and adequate plans for their replacement will have to be prepared with care and well ahead of time.

Response to the Panel's Recommendations

As the panel points out, the last review period has been unusual in the EMBL context because of the lack of turnover of group and team leaders at EMBL Grenoble. As a consequence, the next four years will see significant change. The panel notes that this provides an opportunity for the adoption of new research and service directions. They recommend consideration of researchers engaged in hybrid methods and other computational aspects of structural biology. This is useful advice, which we will take into account during the next round of hiring.

Related to the above, the Head of Outstation, Stephen Cusack, and the second senior scientist in Grenoble, Florent Cipriani, who leads the instrumentation group, are approaching the age of 65. Although this is not yet imminent in either case, the panel notes that both are unusually creative and successful in their activities and that succession planning for both will be important in order to maintain the status and impetus of the Outstation. I agree with this assessment and very much appreciate their advice on both the timing of recruitments and possible new directions for technology development in relation to sample environment on the beamlines.

The panel is appreciative of the individual and collective efforts of the three Team Leaders most closely involved in protein sample handling, Cipriani, Marquez and McCarthy. I agree with these evaluations and with the panel's positive opinion of the promise of the new CrystalDirect system for 'crystal to beam' handling.

The above collaboration has been hugely facilitated by the beamtime made available to the instrumentation developers for testing their new developments on beamline BM14. BM14 is run collaboratively by EMBL, ESRF and India. The panel is very positive about the benefits that have accrued to all partners and to continuation of this collaboration. I will support the prolongation of the three-way agreement and hope that the other partners are also positive about continuing our collaboration.

The panel points out that, despite its modest size, several of the research groups at the Outstation are carrying out research of the highest calibre, singling out projects by Cusack, Berger and Pillai for special praise. I am gratified by the recognition of the quality of the research work being pursued at EMBL Grenoble and by the variety of the activity that is being supported. It will be important to maintain research strength and focus during the upcoming round of recruitment.

The panel notes that the EMBL Grenoble group and team leaders have been successful in external fundraising over the course of the last review period and acknowledges that this demonstrates the competitiveness of the work they are carrying out and, in the case of the service activities, their usefulness to the community. They caution however that mechanisms should be in place to prevent the availability of external funds from unbalancing the activity of the Outstation as a whole. I agree on the need for action to allay this concern.

In conclusion, I thank the panel for their detailed evaluation of the activities of EMBL Grenoble and congratulate the staff of the Outstation for this very positive report.

Professor Iain W. Mattaj, FRS
Director General
20 March 2013

EMBL-EBI Research Activities Review

From 18 to 20 March 2013 the review of the European Bioinformatics Institute's (EMBL-EBI) research activities took place. Thirteen international experts, including four members of EMBL's Scientific Advisory Committee (SAC), formed the Review Panel. The panel was chaired by Reinhard Jahn from the Max Planck Institute for Biophysical Chemistry in Göttingen, Germany.

Evaluation Summary

The panel rates the overall performance of EMBL-EBI as outstanding. In the past years, under the outstanding leadership of Janet Thornton, the EMBL-EBI has firmly established itself as one of the leading bioinformatics institutes worldwide, providing services and carrying out research of highest calibre. Janet Thornton was unanimously praised by the panel for her inspiring leadership and the sharing of some responsibilities by the recently appointed Joint Associate Directors Ewan Birney and Rolf Apweiler was also appreciated.

The research conducted was felt to be of outstanding quality and increasing international visibility. The combination of technology development in bioinformatics for the community as a whole with the application of these tools for outstanding research is unique and constitutes an enormous asset to the scientific community in Europe and beyond. While it is not always easy to differentiate between the service and research components of EMBL-EBI's work, there are about 100 research staff working predominantly or full-time in research, with approximately 40% of them being funded by external grants. The panel views this ratio as healthy, particularly when considering that several group leaders have just arrived and not yet been able to attract outside funds. The success in raising external funding is yet another indicator for the quality of the research conducted at the EMBL-EBI.

The publication record of the unit was evaluated as outstanding and includes many high-profile papers together with other groups or with large consortia, with the EMBL-EBI scientists frequently playing a key role. The quality of the research carried out by service team leaders was viewed as comparable to that of the research group leaders, an impressive achievement when considering that only 20% of the team leaders' time can be dedicated to curiosity-driven research. Among the numerous highlights during the past review period three were mentioned, with two of them being widely publicised by the popular press around the globe, thus increasing the international visibility of the institute: (a) the ENCODE Project; (b) a project showing that DNA can be used to as a stable and efficient storage medium for large amounts of data; and (c) an analysis of 120 000 SNPs in more than 30 strains of E. coli revealing a novel mechanism by which bacteria protect important genes from random mutations. The integration into the European research landscape is evaluated as exceptional, with EMBL-EBI scientists participating in numerous consortia, thus making it a valuable, indeed indispensable, asset to life science research in Europe and its translation into commercially valuable products.

The EMBL-EBI is also a role model in the training of young scientists, both through its predoctoral and postdoctoral programs, coordinated locally by Nick Goldman as part of the EMBL training activities, and through the turnover principle among the group leaders, thus ensuring both the influx of 'new blood' and wide dissemination and networking through its alumni. The EMBL-EBI was considered to have superbly managed the first major turnover in its group leaders, with the leaving scientists all receiving excellent outside positions and the new recruits being viewed as excellent and harbouring a lot of potential for the future.

Response to the Panel's Recommendations

I am delighted that the panel reviewing the research activity of EMBL-EBI formed such a positive impression of the past four-year period and the changes in EMBL-EBI research status in that time. It is easy to forget that the research activity at EMBL-EBI is still relatively young, with this review coming at a time when the first major turnover of group leaders has just taken place. The start of the programme was hindered by the precarious financial situation of the EMBL-EBI in its first years, when it was necessary to devote all the available resources to service activities.

I concur with the extremely positive evaluation of the leadership of Janet Thornton, who has in my opinion been an outstanding leader of EMBL-EBI since her appointment in 2001. Her commitment to research remains total, and this is reflected in her own successful work and the environment she has created for other group and team leaders to engage in research. She has been helped by other senior staff at EMBL-EBI, notably by Nick Goldman who has responsibility for both mentoring incoming group leaders and overall leadership of training activities for fellows at EMBL-EBI.

The panel recommends that we consider adding research on computational aspects, including novel algorithm and database type development, of the science going on at EBI to the portfolio. While this would in principle be very useful to EMBL-EBI, and therefore to bioinformaticians and other life scientists in general, and is pursued in the service teams at EMBL-EBI, it is not obvious to me or my senior colleagues that we can offer researchers in these fields a suitable environment to carry out cutting edge research, and we are thus reluctant to implement this suggestion.

The panel notes that while those carrying out statistical, biological and biomedical research at EMBL-EBI are well integrated, the cheminformatics research teams are currently more isolated. I feel that this reflects the fact that these groups are still very new, this is clearly an area where effort is required.

The panel points out that clinical researchers and others working in medicine stand to benefit enormously from interaction with the research teams at EMBL-EBI, and suggest we move more in the direction of medical research. I agree with both these recommendations and in fact many EMBL-EBI research groups and teams engage in both informal and formal collaborative efforts with clinicians. However, this is clearly an area that can and should grow in future.

I am grateful to the panel for their acknowledgement of the value of the participation of EMBL-EBI research groups and teams in large-scale projects with multiple collaborators. As they correctly point out, these efforts are invaluable to the whole scientific community and combine with the service activities of EMBL-EBI, which were not reviewed here, to make EMBL-EBI an extremely valuable asset for Europe.

I agree with the panel that the tight linkage between research and services at EMBL-EBI is a unique strength of the institute and we will continue fostering connections and collaboration between research and services at all levels.

Professor Iain W. Mattaj, FRS
Director General
19 April 2013

Science Highlights

Ewan Birney

ENCODE-ing the future

In three back-to-back papers totalling five pages of the 25 April 1953 issue of *Nature*, seven authors laid out the evidence and interpretation that established the double helical structure of DNA. Since then, our understanding of DNA has become much more complex as we unravel the behaviour of the genes encoded in that elegant molecule.

In 2003, the Encylopedia of DNA Elements (ENCODE) project set out to identify the functional elements encoded in the human genome. The 6 September 2012 issue of *Nature* dedicated 57 pages to six papers reporting the outcome of the project's second phase, and these were published in parallel with further open-access reports in the pages of *Genome Research*, *Genome Biology*, *BMC Genomics*, *Science* and *Cell*. The main *Nature* paper listed more than 450 authors.

Ewan Birney, Joint Associate Director of EMBL-EBI, led the analysis of the data generated by ENCODE, a project coordinated by the National Genome Research Institute (NHGRI) in the US. The central challenge was to enlist a huge number of scientists to work systematically through the human genome, using exactly the same wet-lab methods and reagents and performing the same computational methods on their results.

The outcome is a body of knowledge that has already begun to accelerate biomedical research.

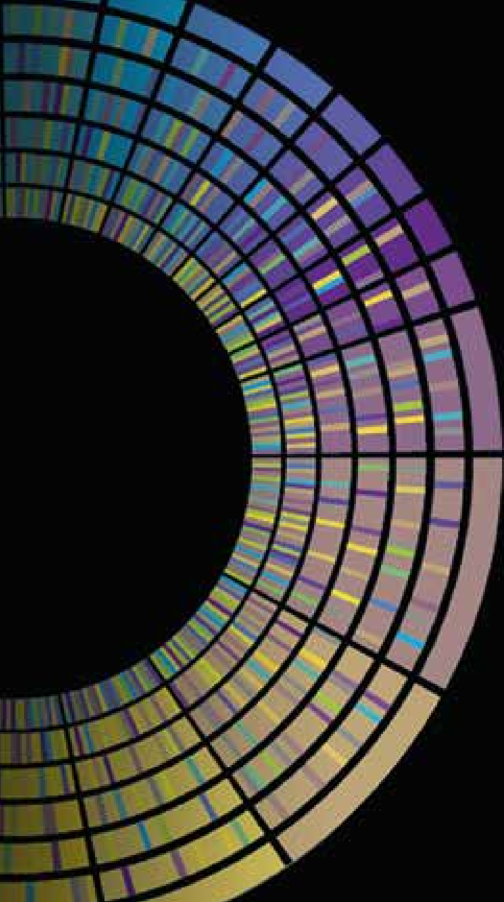
ENCODE in a nutshell

With the falling cost and rising uptake of DNA sequencing, data analysis has become the major bottleneck in genome research. The sheer scale of the data collected by the participants of the ENCODE project is daunting, even to those working on the project: more than 1800 genome-wide experiments were completed, comprising the equivalent of sequencing the human genome 1700 times over.

The experiments were carried out on 147 different cell types, using a specific set of methods. One method might bring to light many different activities happening across the genome, whereas others might reveal a single, important detail. The first phase of ENCODE was dedicated to hammering out these methodologies; in the second phase, everything was scaled up.

More than they'd bargained for

ENCODE revealed an unexpected amount of hustle and bustle in the human genome: some form of recognisable biochemical activity seems to be happening just about everywhere. They



What is **ENCODE?**

could clearly see that, at a minimum, the packaging of the genome – and its readiness for transcription – is actively controlled.

"ENCODE data are changing the way we think about genomes," says Ian Dunham from Ewan's research group, who played an important part in the analysis. "It is fascinating to see where the data intersect with human variation studies, and what that can tell us about the information coming from genome-wide association studies. With ENCODE, we now have a very rich source of information to explore."

Open data and genomics

ENCODE is a community resource, and from its inception has shared its data as soon as they are shown to be reproducible. This is not a new concept: since the advent of large-scale genome sequencing, the field of genomics has been at the forefront of initiatives to make research data freely available on a timely basis. The Human Genome Project routinely released assembled sequences of clones within 24 hours – a practice often met with disbelief by potential collaborators.

"Widespread data-sharing in science is essential because it allows different groups to use each other's data. Because of the sheer scale of the ENCODE project, we were able to deliver a staggering volume of standardised analyses that were made available to everyone," says Ewan. "We had the depth of expertise needed to tackle this large-scale work systematically, and this is the distinction of our 'catalogue' approach. What we delivered is already allowing many smaller-scale, hypothesis-driven studies to pursue new avenues in their analyses – and it is these studies that will lead us to real understanding."

Early data release certainly proved an effective policy for ENCODE: before the big suite of ENCODE papers was published in 2012, more than 200 papers using ENCODE data had already been published by scientists outside the consortium.

Turning the page

Generating and releasing ENCODE's experimental data was a time-consuming task, and one that involved several late-night, marathon teleconferences. But beyond delivering great science, Ewan and Ian put a lot of thought into how ENCODE's results should be delivered.

Ewan's driving motivation was to encourage readers to explore the data, to interact with it and to find their way to information that might lie at the borders of their comfort zone. ENCODE needed more from the publishing process in this modern age than an online PDF version of the paper – they wanted to secure resources that would let people reproduce the analysis and easily navigate the ideas in 30 different papers.

Nature took up the challenge, and developed an app that features interactive graphics and electronic 'threads' that knit together the related ideas interwoven throughout the many papers. According to Ian, these threads constitute the authors' own 'meta-review' of the papers.

The 'Methods' section grows up

ENCODE's introduction of 'virtual machines' into the Methods section could, quite possibly, transform the traditional publishing model into a more interactive experience. These are working versions of the computer code, which constitutes a major part of the experimental method – they provide a transportable, functioning distillation of the analysis. Anyone with a reasonably powerful computer can directly reproduce the results and, furthermore, can work with the code to develop their own analyses.

"I can imagine a future in which one doesn't just publish and forget, but each publication element contributes to a highly interconnected, worldwide research resource," says Ian. "Many people are already working towards this goal in different areas but a push must come from the research community to really make it happen."

What's next for ENCODE?

Ewan and Ian have moved on to new projects, but ENCODE marches on, extending the analyses to cover hundreds of cell types and even other species.

"ENCODE was inspiring and challenging by turns," reflects Ewan. "The real reward is seeing this constant stream of papers that are based on the data. ENCODE as a project has been refining the art of consortium science, and in the process has produced high-quality results that can benefit everyone. Looking back on my ten years with ENCODE, it has been a lot of hard work and excellent science – I've met and interacted with so many great scientists and have honestly had a lot of fun."

ONLINE EXTRAS

ENCODE app: http://itunes.apple.com/app/id553487333

ENCODE virtual machine: http://scofield.bx.psu.edu/~dannon/encodevm/

Watch Ewan and co-authors talk about ENCODE on the EMBL YouTube Channel: youtu.be/KiwXtHRfBC8.

All the papers can be explored on the ENCODE website: www.nature.com/encode

Turning back the clock

Scientists have known for about half a century that it's possible to effectively put cells in a time machine, 'winding back' their development to the point when they were stem cells. These stem cells can then be coaxed into becoming a different type of cell altogether, so the process – called reprogramming to induced pluripotency – holds such promise that its pioneers were awarded the 2012 Nobel Prize in Physiology or Medicine. But despite these tremendous advances, researchers still don't have a full answer to a crucial question: What exactly happens inside a cell as it changes from a specific type to one that can turn into virtually anything?

"We saw some changes that there was no hint of before"

"People have been looking at this a lot, primarily at the level of how genes are expressed and controlled. No-one's really looked at what happens to proteins," says Jeroen Krijgsveld from EMBL Heidelberg. As team leader in the Genome Biology Unit and Head of the Proteomics Core Facility at EMBL Heidelberg, Jeroen was ideally poised to launch an investigation into the matter, and the results were published in *Cell Reports* in 2012.

Jenny Hansson, a postdoctoral scientist in Jeroen's group, used mass spectrometry to monitor how the production of 8000 different proteins changed over the two weeks it took to reprogramme skin cells back into stem cells. But first, the EMBL scientists had to contend with a shortcoming of reprogramming. Even the best-honed methods for turning cells back into stem cells are very inefficient: only about 1% of cells in each experiment undergo the transformation. Under such conditions, attempting to follow protein changes in all the cells in a pluripotency experiment would be wasteful. Working with Konrad Hochedlinger from Massachussets General Hospital in Boston, USA, who had developed a sorting method for such a purpose, the scientists were able to fish out the cells that were likely to become stem cells, and thus study only those in which they were truly interested.

In general, they found that the amount of most proteins the cells produce varies

tremendously between the start and end of the reprogramming process: either going from being produced at high levels at the beginning to hardly being produced at all, or vice versa. Thus, scientists now have more precise information about the molecular processes that take place at different stages of reprogramming.

The study also turned up some surprises. "We saw some changes that there was no hint of when people looked at gene expression," says Jenny. "The genes for these proteins are not getting turned up or down, but the amounts of proteins are changing in unexpected ways."

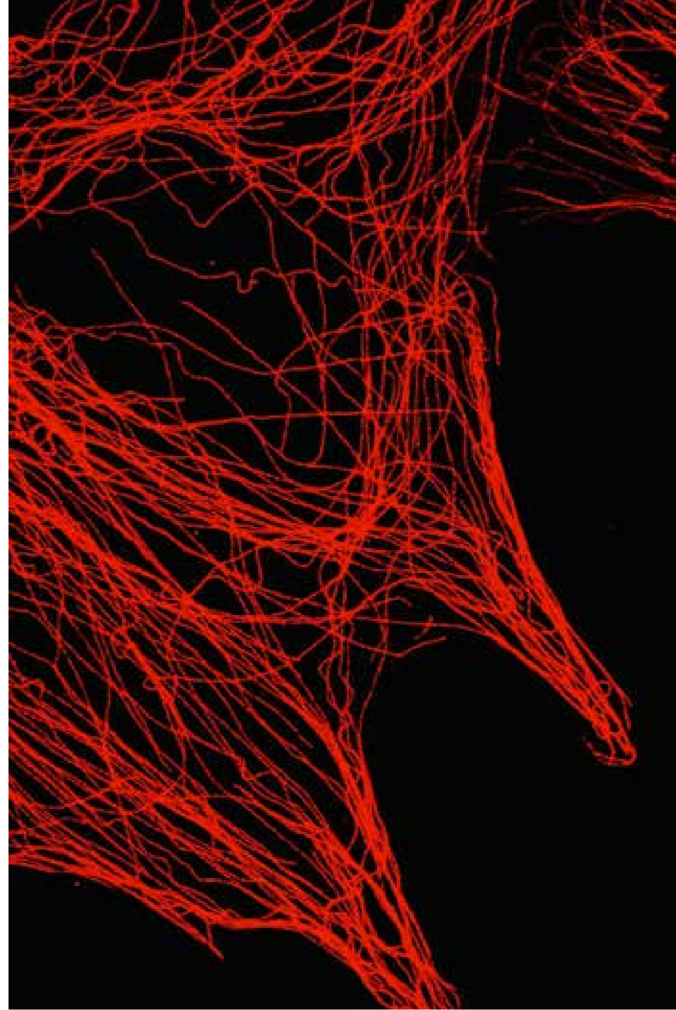
Jenny and Jeroen are keen to point out that this is just a starting point. They and others can now mine this data for patterns of protein change, and for proteins that inexplicably deviate from those patterns. For instance, together with Sina Rafiee in Jeroen's group, they have already discovered that a protein called Nup210, a nuclear pore protein that was recently shown to be essential for stem cells to differentiate into neurons, is also crucial for the transition back to stem cells.

Of the multitude of questions to explore, Jeroen would love for someone to answer: Why does reprogramming take 15 days? What happens in those middle days when nothing much seems to be going on? "And ultimately, can we speed up the process? When you're talking about taking cells from patients, having to wait two to three weeks is wasted time," he muses.

Hansson J, Reiland S, Rafiee MR, Polo J, Gehring J, Okawa S, Huber W, Hochedlinger K, Krijgsveld J (2012) Highly coordinated proteome dynamics during reprogramming of somatic cells to pluripotency. *Cell Rep.* **2**: 1579-92

DOI: 10.1016/j.celrep.2012.10.014

Jenny Hansson and Jeroen Krijgsveld



A touch of complexity

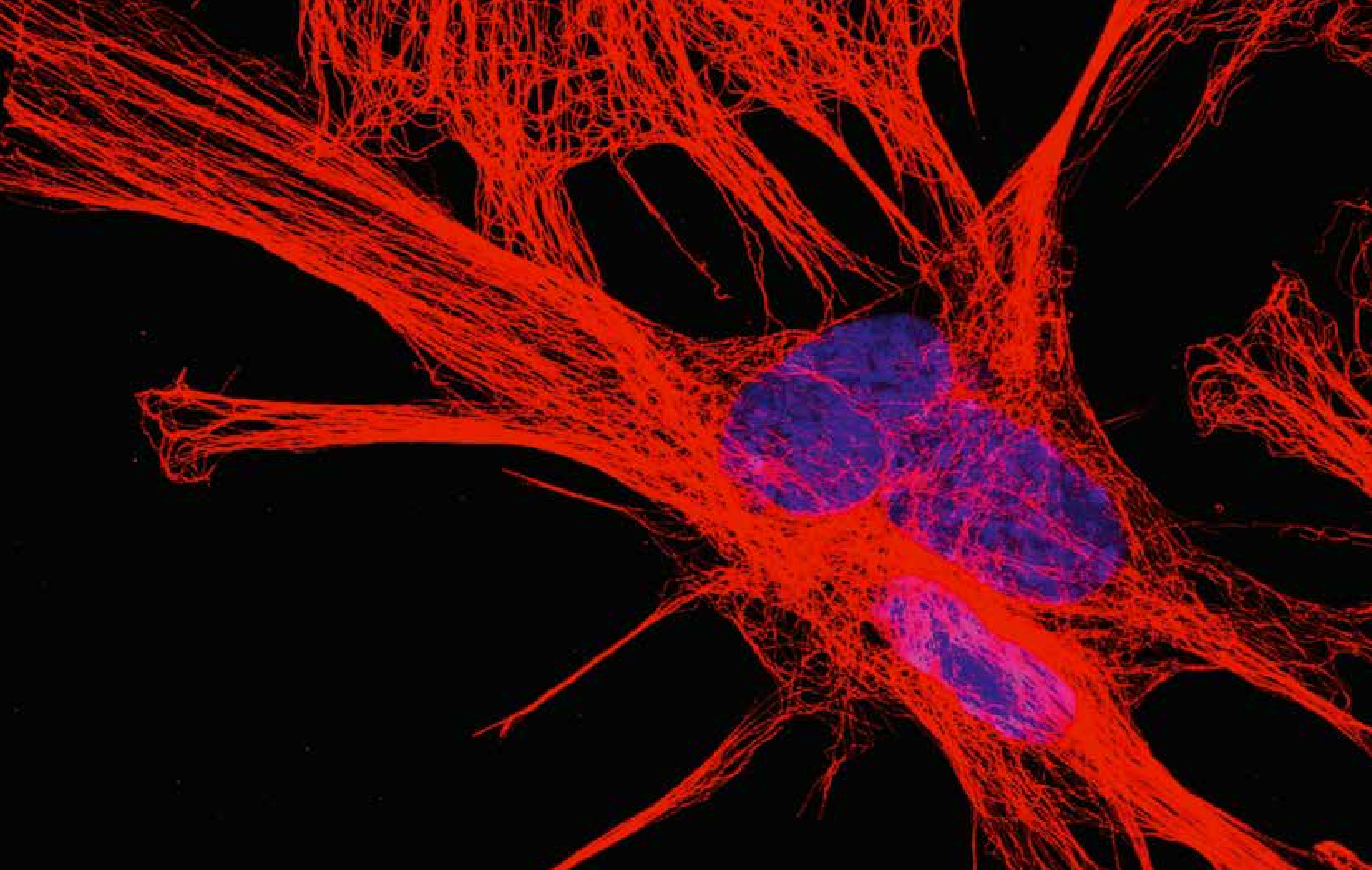
It is often not the expected but the unexpected findings that help researchers unearth the root of a scientific mystery. Paul Heppenstall and his team at EMBL Monterotondo recently experienced two such discoveries, which promise to yield intriguing new insights into the biology of nerve cells.

"Acetylation is probably not very important at the whole-organism level in a non-stressed environment"

Paul's team is interested in how touch sensation works, focusing on the function of the molecules in touch-sensitive nerves. Many genes involved in this process have been found by studying the tiny worm *Caenorhabditis elegans*, which is easy to breed in the lab. Paul had previously studied the mouse equivalent of some of these genes and when he moved to EMBL Monterotondo, he continued with this approach by investigating a protein encoded by one of them, an enzyme called MEC-17.

MEC-17 adds or removes a particular chemical mark, called an acetyl group, to or from other proteins. This is known as acetylation and cells use this process to alter a molecule's behaviour. MEC-17 acetylates a protein that helps to form the cell's internal scaffolding, the cytoskeleton. This is made up of a stiff web of filaments called microtubules that plays a key role in many vital processes, such as cell-shape maintenance and cell division. It also acts rather like a rail freight network by allowing packages of proteins and other molecules to be sent along it from one part of the cell to another, a process of key importance in long or intricate cells such as nerve cells.

Microtubules are made from building blocks of proteins known as tubulins and are highly



Unexpected finding: whether mouse microtubules are acetylated (red) or not doesn't affect their stability.

dynamic. They can abruptly switch between growing longer and becoming shorter by adding or removing tubulin blocks at their ends. Although this allows them to respond quickly to the cell's needs, they must also remain stable if required. So a key goal has been to understand what controls this stability.

Biologists have known for about 30 years that stable microtubules are acetylated, but didn't know how or why. More recent research suggested that cells use acetylation by MEC-17-like enzymes to make microtubules more stable.

To find out more, Paul and his team looked at the mammalian equivalent of MEC-17, an enzyme called αTAT1. To their surprise, they found that αTAT1 activity seemed to make microtubules less, rather than more, stable. Yet it clearly added acetyl groups to α-tubulin. When the team disabled αTAT1's acetylation activity in cells in the lab, they found that it still made microtubules less stable. This suggested that αTAT1 had two separate effects: to control microtubule stability and another, unrelated function to acetylate microtubules, and also that acetylation wasn't involved in controlling microtubule stability after all.

To investigate αTAT1's role in whole animals, Paul's team created mice that lacked the gene for the enzyme. This yielded another surprise: mice lacking αTAT1 were almost completely normal, suggesting that lab-grown animals could survive with non-acetylated and less stable microtubules. "Acetylation is probably not very important at the whole-organism level in a non-stressed environment," says Paul, adding that the mice may show signs of problems as they age or experience stress.

So what's going on? That question is the target of the team's next efforts, says Paul. "We are going to go after the mechanism."

Kalebic N, Martinez C, Perlas E, Hublitz P, Bilbao-Cortes D, Fiedorczuk K, Andolfo A, Heppenstall PA (2013) Tubulin acetyltransferase αTAT1 destabilizes microtubules independently of its acetylation activity. *Mol. Cell Biol*. **33**: 1114–23

DOI: 10.1128/MCB.01044-12.

How viruses come of age

Viruses are the simplest life forms known, consisting largely of a protein shell wrapped around a small genome. In fact, it's not clear whether they should be thought of as living at all. Yet even these simple entities go through something akin to human adolescence. Like humans, viruses need to grow up before going out into the world to do what they're supposed to: infect as many cells as possible. This involves a big change, from an immature form into a mature and infectious viral particle.

Like humans, viruses need to grow up before going out into the world

Human immunodeficiency virus (HIV) goes through such a shift, and previous studies have worked out the structure of the mature viral particle in great detail. However, a comparable understanding of the immature form has lagged behind, so it has not been clear precisely what changes occur during this transition to viral adulthood.

This gap in our knowledge is something that John Briggs, group leader in the Structural and Computational Biology Unit in Heidelberg, wants to fill in. His lab has now taken a major step towards that goal by obtaining a high-resolution structure of the immature form of a virus very similar to HIV. A few years earlier, John's lab applied cryo-electron tomography to obtain three-dimensional images of the immature form of HIV, but not at a very high resolution (see EMBL Annual Report 2009/10, page 16). "What we wanted was a sufficiently detailed view so we could see the structural proteins in some detail, and work out how they stick together to create the virus," says John.

In 2012 John and his colleagues managed to produce such an image, and have described their findings in *Nature*. The team studied Mason-Pfizer monkey virus (MPMV), which, like HIV, is a retrovirus that has an outer shell, or capsid, made out of similar repeating protein units. These capsid proteins are derived from a large molecule called Gag that is cut by a viral enzyme during the maturation process. This releases the individual capsid proteins and allows them to form the mature virus's shell.

Tanmay Bharat and John Briggs

Using cryo-electron microscopy and tomography, John's group has now been able to show that the packing of capsid proteins, and the connections made between them to stabilise the capsid shell, are very different in the immature and mature forms of MPMV.

The next challenge is to gain a clearer picture of how viruses make the shift from adolescence to adulthood. John's lab has studied mutated viruses that are stuck in adolescence because their Gag protein cannot be cut, but these viruses in 'arrested development' are unlikely to represent natural intermediates in the maturation pathway. "We still need to work out the steps between immature and mature virus in the natural state," says John.

Such an understanding could eventually open the door to novel therapies for HIV and other virus-caused diseases. "There's an interest in assembly and maturation inhibitors as therapies," says John. "Using our data for this is some way off, but there are already some candidate compounds in the pharmaceutical companies."

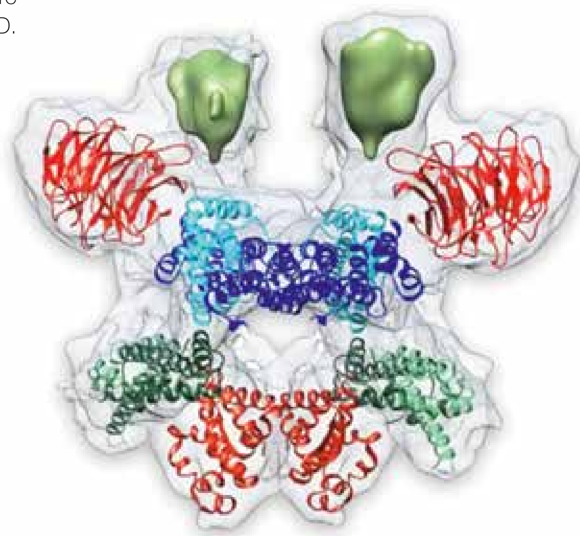
ONLINE EXTRA

Watch the HIV shell protein change from the immature to the mature form of the virus: youtu.be/57CtUqVDhcE.

Bharat TAM, Davey NE, Ulbrich P, Riches JD, Marco A, Rumlova M, Sachse C, Ruml T, Briggs JA (2012) Structure of the immature retroviral capsid at 8Å resolution by cryo-electron microscopy. *Nature* **487**: 385-9

DOI: 10.1038/nature11169.

The docking station-like core of TFIID.



Plug-in puzzle

Like the docking station of a laptop, the TFIID protein complex enables all the important elements involved in gene transcription to be plugged in together, at the right place in the genome. It is central in determining which genes are transcribed into messenger RNA (mRNA), the intermediate molecule that is translated into all the proteins the cell needs to function. For the first time, Imre Berger and his team at EMBL Grenoble have described in molecular detail the internal parts and wiring of the TFIID 'docking station' that are known as the core complex.

The core complex of TFIID is made up of 10 parts, or subunits. Imre's analysis, published in *Nature*, shows that some of these adopt a very defined structure whereas others have more intricate, extended geometries, winding like worms through the complex and holding it together.

The complete TFIID complex – the full 'docking station' made of all its internal parts and wiring plus the case around them – comprises 20 subunits. The EMBL scientists also describe how the structure of the initial, symmetrical, core complex of 10 subunits needs to become asymmetric in order to bind to the remaining 10 parts and form the complete TFIID complex. This organisation and mechanism are much more complicated than what previous theories, based on partial data, anticipated.

"We know now in some detail what the core of TFIID looks like, and what happens when further subunits are bound. We believe that we have opened the door to determining the architecture of the entire human TFIID complex in the near future, and likewise of other large multiprotein assemblies involved in gene regulation, and to explain their roles in catalysing biological function," explains Imre.

TFIID is a very large protein complex that is present at very low levels in cells, which has hampered previous attempts to purify it and decipher its structure and function in molecular detail. Even the most advanced methods met their limits when trying to produce its various subunits in the right proportions. The solution to this bottleneck came from studying the strategy certain viruses, such as Coronaviruses, use when they replicate: they produce very long protein chains that are then divided into individual proteins. Imre's group inserted the genes coding for the various parts of TFIID into such a Coronavirus. The modified virus was then inserted into an insect cell that became able to produce large quantities of correctly assembled complexes of the 10 core subunits of TFIID. After purification, the complexes were analysed at high resolution by combining electron microscopy and data from X-ray crystallography.

Many important biological functions are carried out by protein complexes that are often too big to be produced on a large scale by traditional means. This ingenious technique, combining a modified Coronavirus with the insect cells-based MultiBac system previously engineered by Imre's group, opens new avenues for the study of the structure and mechanism of other large multi-protein assemblies, both in the academic and the industrial settings.

Bieniossek C, Papai G, Schaffitzel C, Garzoni F, Chaillet M, Scheer E, Papadopoulos P, Tora L, Schultz P, Berger I (2013) The architecture of human general transcription factor TFIID core complex. *Nature* **493:** 699-702

DOI: 10.1038/nature11791.

Perfect timing

Synchronise watches!

Some parts of our genome do the molecular equivalent of synchronising their wristwatches, according to recent research findings by scientists at EMBL Monterotondo. Sara Buonomo and her colleagues have identified a protein that ensures that different sections of DNA all copy themselves at the right time, to maintain the correct structure and function of chromosomes. Published in *The EMBO Journal*, the work could shed new light on how a faulty genome structure can contribute to a cell becoming cancerous.

Before a cell divides, it has to copy its DNA accurately, and then package the copy neatly into chromosomes. Mistakes in either process can be disastrous, triggering mutations that alter cell behaviour or structural changes that make chromosomes more prone to breaking or ending up in the wrong cell during division.

DNA copying, or replication, takes place during a stage of the cell's life cycle known as S-phase and, given its importance, the whole process is tightly controlled. Copying starts at a range of specific points along each chromosome known as replication origins. To ensure accurate DNA replication, the behaviour of these replication origins needs to be coordinated, but the underlying mechanism has been somewhat mysterious. One particular puzzle is that not all the replication origins start copying at the same time. Broadly speaking, there seem to be two kinds of chromosome areas: those that all replicate early in S-phase, and those that replicate late. "If they all replicated at once, DNA replication would take one hour instead of 10 hours," says Sara. So why does the cell seem to do things the hard way?

One possible explanation relates to how DNA is packaged up with proteins to form a structure known as chromatin. Chromatin comes in two forms: a loosely packed form called euchromatin, where most of the cell's active genes lie, and a tightly packed form called heterochromatin, which contains inactive genes. Some regions of heterochromatin also play a vital mechanical role in chromosome structure.

Intriguingly, euchromatin usually replicates early, whereas heterochromatin usually does so late. One idea is that different replication times allow the cell to make sure that these different chromatin structures are preserved from one cell generation to the next. This would require some sort of mechanism that could control the timing of replication for the whole genome.

In collaboration with Rachel Santarella-Mellwig and Claude Antony at the Electron Microscopy Core Facility at EMBL Heidelberg, Sara and her team studied a protein called Rif1 and showed that it attached itself to areas of DNA that were about to replicate. When they disabled Rif1 in mouse cells in the lab, they found that the timing of replication was disrupted, with some heterochromatin replicating early, and some euchromatin replicating late. This meant that the team had found the first protein to control replication timing across the whole genome.

Further work with Vladimir Benes from the Genomics Core Facility at EMBL Heidelberg revealed that the chromatin in cells lacking Rif1 was more loosely packed than normal. "This says that it is important to segregate different types of chromatin in time to repackage them correctly," explains Sara. So it seems as though timing replication could indeed help cells pass the correct chromatin structures down through cellular generations. Exactly how Rif1 coordinates this will be revealed by the team's next round of experiments – all in good time, of course.

Cornacchia D et al. (2012) Mouse Rif1 is a key regulator of the replication-timing programme in mammalian cells. *EMBO J.* **31**: 3678-90

DOI: 10.1038/emboj.2012.214.

Mysterious beginnings

In the classic children's book, Winnie the Pooh came up against a conundrum every time he had to tell his left paw from his right: "He knew one of them was the right, and he knew that when you had decided which one of them was the right, then the other one was the left. But he never could remember how to begin." Scientists studying small regulatory RNA molecules called piRNAs have been grappling with a similar issue for decades. Once the cell has some piRNAs, it can use them to produce more – but how is that very first batch created? Orsolya Barabas' group at EMBL Heidelberg have now come a step closer to the answer.

Once the cell has some piRNAs, it can use them to produce more – but how is that very first batch created?

Franka Voigt, a PhD student in Orsolya's lab, determined the three-dimensional structure of a protein called Zucchini, and by comparing it to other proteins, was able to solve the mystery of how it acts. Scientists already knew that Zucchini was crucial for producing piRNAs – mice and flies that lack Zucchini are unable to generate them – but there was an ongoing debate as to why this protein was so important. Most of the proteins in Zucchini's family act on lipids – the fatty molecules cells use to build membranes – but some cut up DNA and RNA molecules instead. Accordingly, some scientists thought Zucchini might produce the small piRNAs itself by cutting up longer stretches of RNA, whereas others suggested that it processed the lipids needed to create the niche where piRNAs are formed.

"When you look at the pictures, it's pretty obvious even for the uneducated eye that Zucchini is much more like this," Orsolya says, pointing at the structure of an RNA-cutting protein, "than like that," pointing at a lipid-cutting family member.

Orsolya Barabas and Franka Voigt

The labs of Mikiko Siomi at Keio University, Japan, and Gregory Hannon and Leemor Joshua-Tor at Cold Spring Harbor, USA, independently reached the same conclusion as Orsolya, so this part of the puzzle seems solved: Zucchini is essential for producing piRNAs because it cuts up long RNA molecules. But the overall conundrum of piRNA production is by no means untangled. Among other things, scientists now need to work out whether Zucchini randomly cuts the RNA or targets specific points, and whether it works alone or with other proteins.

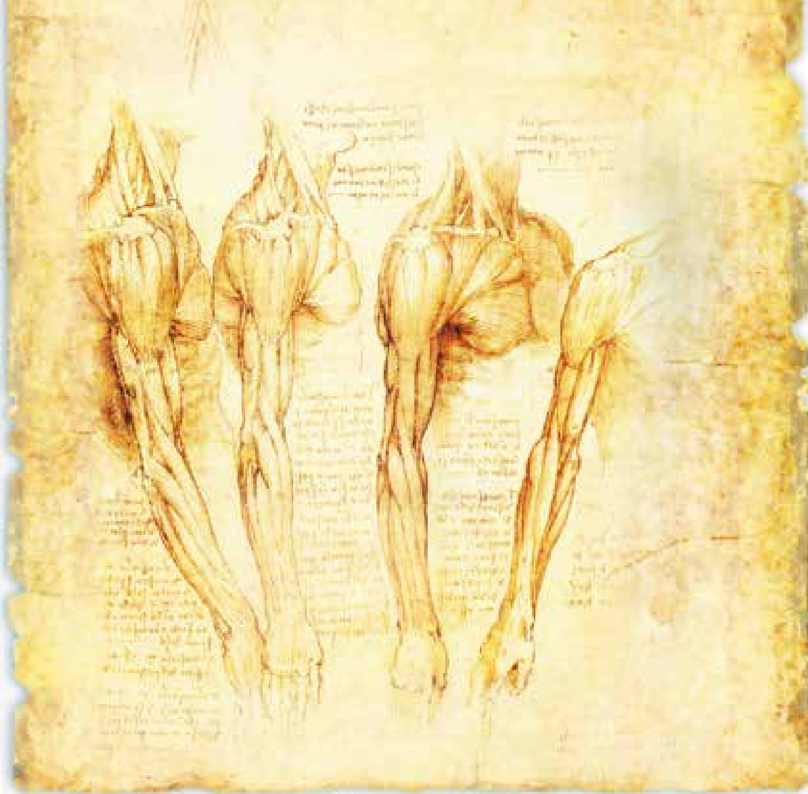
The study performed by Orsolya's group, which was published in *RNA*, is part of a wider collaboration with Ramesh Pillai's group at EMBL Grenoble, as both are interested in the role of piRNAs and other molecules in controlling what could be a rampant threat to our genomes: transposons, or jumping genes (see page 62 for Ramesh's latest work on this topic).

ONLINE EXTRA

Listen to Orsolya and Ramesh talk about transposons on the EMBL Explore Podcast: www.embl.org/explore.

Voigt F, Reuter M, Kasaruho A, Schulz EC, Pillai RS, Barabas O (2012) Crystal structure of the primary piRNA biogenesis factor Zucchini reveals similarity to the bacterial PLD endonuclease Nuc. *RNA* **18**: 2128-34

DOI: 10.1261/rna.034967.112.



What controls arm growth, at the molecular level?

Twisted communication

Transforming one fertilized egg into a healthy and well-proportioned embryo with four limbs and all the necessary organs: embryonic development is a very complex and precise process that requires a tight control of when and where genes are turned on. One of the genes at the heart of the action is called *Fgf8,* and it controls the growth of the limbs and the formation of the different regions of the brain. The team of François Spitz at EMBL Heidelberg has discovered how *Fgf8* is, itself, fine-tuned by a series of interdependent regulatory elements. Their findings, published in *Developmental Cell*, shed new light on the importance of the genome's structure for gene regulation.

Fgf8 and its many regulatory elements lie in a long stretch of DNA where they are interspersed with other, unrelated genes. By selectively removing portions of this region of the genome, and thus changing the relative positioning of the regulatory elements, François and his group were able to show that these many regulatory elements work in combination to control the expression of *Fgf8*. The absence of some of them is sufficient to drastically affect the development and length of the limbs in the embryo.

"What is really striking in this genomic region is that the high degree of intertwining between unrelated genes is nevertheless translated into very specific gene expression patterns in the embryo. We showed that this specificity is not imposed by a local code that distinguishes each gene: it arises from the overall organisation of the region," explains François. "*Fgf8* responds to the input of specific regulatory elements, and not to others, because it sits at a special place, not because it is a special gene."

It is likely that the way DNA folds in three dimensions could, under certain circumstances, bring different sets of regulatory elements in contact with each other and with *Fgf8*, to turn the gene on or off. These findings highlight a level of complexity of gene regulation that is often overlooked. They show that some regulatory elements may be engaged in a one-to-one relationship with their target gene not because it has the appropriate DNA sequence, but rather because of the order in which they appear on the chromosome. This local genomic organisation may lead to a specific 3D folding of the DNA that could put the regulatory elements into contact with one gene rather than another, and thus modulate its action.

François' group is still looking into the molecular details of this regulatory mechanism and, more globally, into the impact of the 3D structure of DNA on the communication between the various elements of the genome. Further down the line, this work could also increase our understanding of how genomic rearrangements might disrupt these 3D-regulatory networks and lead to diseases and malformations.

Marinić M, Aktas T, Ruf S, Spitz F (2013) An integrated holo-enhancer unit defines tissue- and gene specificity of the Fgf8 regulatory landscape. *Dev. Cell* **24**: 53-42
DOI: 10.1016/j.devcel.2013.01.025.

Bending borders

We like the walls of our houses to be solid and strong so that they create a safe and warm environment while protecting us from the outside world. But this means that we have to get in and out of the sanctuary of home through special portals commonly known as doors.

Cells take a different approach. Rather than enclosing themselves in an inflexible border that can only be traversed at one specific point – the equivalent of a cellular doorway – cells are wrapped in a fluid, flexible membrane that allows material to pass through at many points. This involves bending the flat surface of the membrane towards the inside of the cell, forming a bud that is eventually pinched off to create a membrane-bound parcel, or vesicle, that carries its cargo to wherever it's needed in the cell.

This process is called endocytosis and cells use it for a variety of tasks, such as taking up nutrients from the surrounding environment, or recycling receptors sitting in the membrane. It is also exploited by viruses to gain unauthorised entry to the cell by binding to the cell surface, triggering endocytosis and sneaking in.

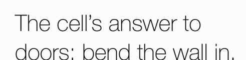
The lab of Marko Kaksonen, group leader within the Cell Biology and Biophysics Unit at EMBL Heidelberg, specialises in the study of endocytosis. In recent years, Marko has worked with John Briggs, group leader in the Structural and Computational Biology Unit, to develop a new method for studying dynamic biological processes, which they call correlated light and electron microscopy. Last year, they reported proof-of-principle results showing that the technique works (see EMBL Annual Report 2010/11, page 4). Now they've applied it to endocytosis. "This study represents the first large-scale application of this method," says John.

The basic logic of correlated light and electron microscopy is straightforward: to bring together the benefits of light and electron microscopy to overcome the limitations of each. Light microscopy is great for studying biological processes in living cells, but the resolution is limited so it cannot provide detailed data on biological structures. Electron microscopy, on the other hand, produces very high-resolution images, but can only be applied to samples that have been fixed or frozen in their tracks.

In research reported in *Cell,* Marko and John were able to create the first three-dimensional movie of how the cell membrane bends and buds, and identify which proteins are involved in this process. Light microscopy experiments had previously revealed that the roughly 50 proteins that assemble at the membrane during endocytosis arrive in a specific order: the proteins could be tagged with a fluorescent probe, and their arrival at the membrane visualised. This approach, however, wasn't able to connect the presence of the labelled proteins with the physical membrane changes; this required the high resolution of the electron microscope.

Marko and John's studies show that the membrane only begins to bend once molecules of a crucial protein called actin bind to each other at the site of endocytosis. "There are many different proteins that could potentially bend the membrane, but we showed that even though these proteins are already there the membrane remains flat until actin arrives," says Marko.

While buildings have a fixed number of entry and exit points, the cell's border cleverly creates new portals on the spot. But just as you can't open a door without a key, these new studies show that you can't enter the cell without actin.



The cell's answer to doors: bend the wall in.

ONLINE EXTRA

Watch how the cell swallows molecules: youtu.be/50eIv8FN-lk.

Kukulski W, Schorb M, Kaksonen M, Briggs JA (2012) Time-resolved electron tomography reveals how the plasma membrane is reshaped during endocytosis. *Cell* **150**: 508-20

DOI: 10.1016/j.cell.2012.05.046.

Peer Bork and Shinichi Sunagawa

I am the one and only

New parents can spend hours proudly gazing at their bundle of joy, trying to decide whether Baby has Dad's nose, Mum's eyes or Grandad's eyebrows. Of course, infants are not just carbon copies of their parents: each has a combination of DNA variations that makes him or her a unique individual. But it now seems that as they grow, babies become even more special. Researchers at EMBL Heidelberg have revealed the existence of another kind of genetic variation, one that could affect our health and certain aspects of our physiology. But you won't see evidence of it in a baby's adorable brown eyes or fluffy blonde hair. Instead, you will have to inspect the contents of his or her nappy.

The genetic variation in question belongs to the trillions of microbes that inhabit our guts. This teeming ecosystem, known as the gut microbiome, is known to play a pivotal role in helping us to digest food, break down toxins, resist certain infections and develop a healthy immune system. It comprises many different species of bacteria, and other microbes such as viruses and yeasts. Alterations in the balance between these species have been associated with a number of diseases, such as inflammatory bowel disease, obesity and diabetes.

Thanks to recent advances in DNA sequencing and bioinformatics, researchers have been able to learn a great deal about this ecosystem and how it works. They have identified many new species that live alongside well-known gut bacteria such as *Escherischia coli*, or *E. coli*. Now, Peer Bork and his colleagues at EMBL Heidelberg have gone one step further and have developed a way to identify different bacterial 'strains' – genetically distinct variants within species – in the gut microbiome. Their findings show that we all harbour a unique set of strains in our gut, and that these particular bacteria probably stay with us for a long time. "We could show that each person has a unique microbiome, at the strain level," says Peer. "We both have *E.coli* in us, but I have a different one from you."

The work is groundbreaking because this is the first time scientists have been able to extract this kind of detail from such a complex microbial ecosystem. This is invaluable, because small genetic changes can transform a harmless bacterium into one that causes disease, or allow it to become resistant to antibiotics. Being able to study the gut microbiome at this level of detail could in future allow scientists to develop more personalised medical therapies, trace the origins of some diseases and develop a greater understanding of how our gut microbes interact with us.

Until now, most of the work to study species diversity in microbial samples has relied on using a single gene. This gene encodes part of the cell's protein-producing machinery called 16s rRNA and shows small changes in its sequence from one species to the next. By identifying all the different versions of this gene in a sample, researchers can estimate how many different species are present, and also identify particular species if the sequence of their 16s rRNA gene is already known.

Although undoubtedly useful, this approach has its limitations – it only looks at one gene among many in each microbe. It also limits studies to looking at the species level: the equivalent of studying the inhabitants of a town and determining whether they own a dog or a cat. Being able to study the sequences of all the genes in the sample would yield far more information, such as what sorts of functions each microbe can perform. It would also let researchers identify different genetic variants, or strains, within each species – in other words, being able to work out the breeds of those cats and dogs: Siamese, Ragdolls, Labradors, Yorkshire Terriers or members of a completely new breed. What's more, genetic variants could give researchers an idea of how environmental pressures, such as the host's immune system or antibiotic use, favour the evolution of different strains. "This information was completely missing," says Peer.

But getting this information was not going to be easy. The vast majority of microbes cannot be grown in a lab, which means that scientists have to sequence directly from microbial samples – in this case, samples of stools from volunteers. But each microbial genome contains thousands of DNA letters, far too many to be sequenced as a whole.

"We both have *E.coli* in us, but I have a different one from you."

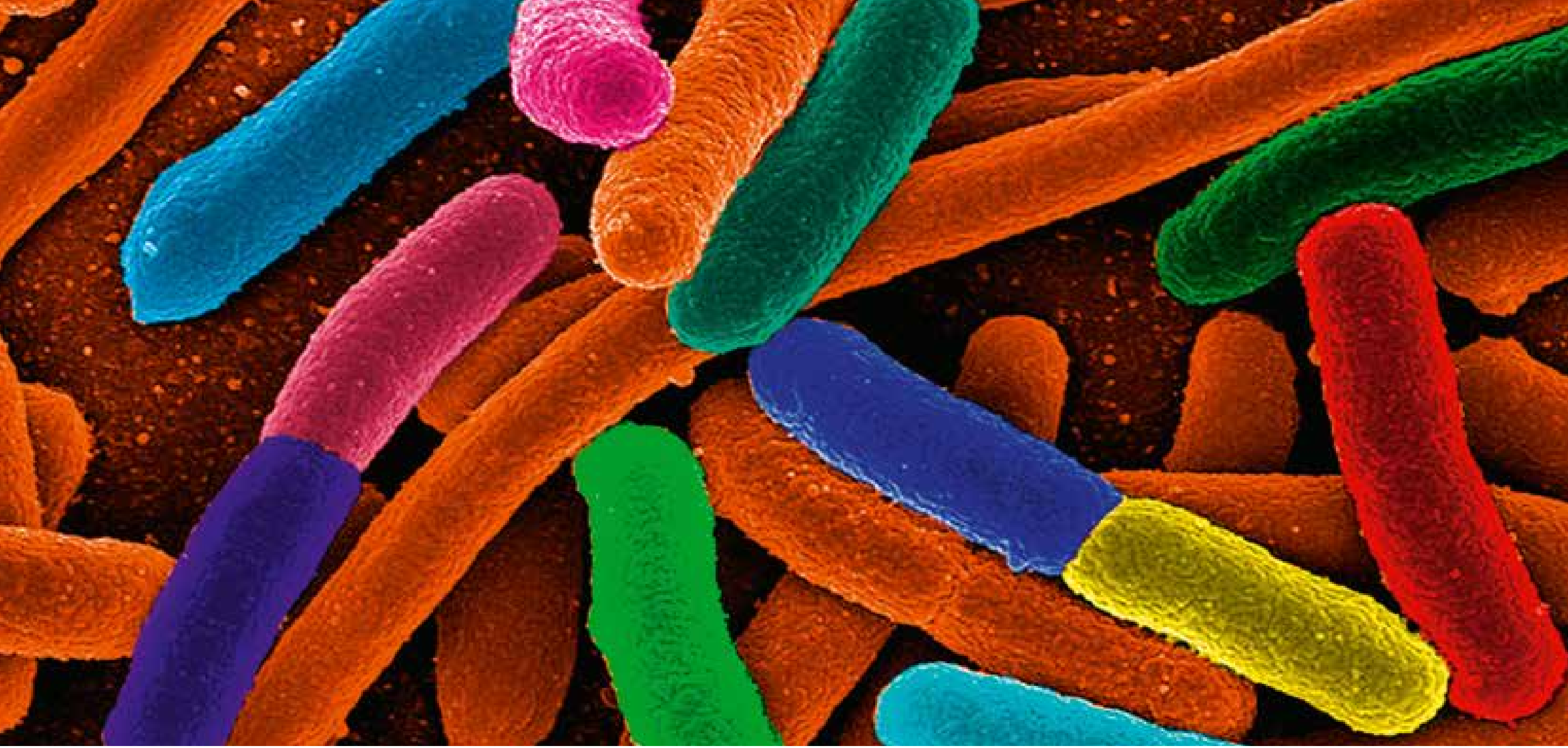
When studying large, single genomes such as the human genome, researchers had solved this problem by breaking the genome up into millions of tiny pieces. With the help of some biochemistry they could deduce the DNA sequences of the pieces, before using sophisticated computer programs to piece the genomes back together again.

Just over a decade ago, a DNA-sequencing method called Illumina sequencing became available, which made the process much faster and more efficient. In theory, this meant that sequencing microbial ecosystems was more feasible than before. But Illumina sequencing generates vast amounts of information, so researchers believed that it could not be applied to samples containing thousands of different microbial species. The computer programs would be overwhelmed by the complexity of trying to match so many tiny fragments of DNA correctly, they argued. Peer thought otherwise: "I love to go against dogmas," he says. "So we tried to crack it."

And crack it he and his colleagues did. Together with scientists from an EU-funded consortium of researchers called MetaHIT, Peer's group used the method to create a catalogue of genes present in the gut microbiome. The work was well received by other scientists, but the prevailing assumption was that it would not be possible to also catalogue the genetic variation present in these genes. "This was another dogma!" says Peer. As well as being unable to resist the challenge, Peer was intrigued by the range of new insights the data could provide. "If you have this resolution, you can do lots of things with it," he says.

In collaboration with George Weinstock and his colleagues at Washington State University in the USA, Peer and his team studied faecal samples from 207 people in Europe and North America. They were looking for a range of variations, including those that added or removed chunks of DNA to a gene, and changes to individual letters in a gene's DNA sequence, known as single nucleotide polymorphisms, or SNPs. The results were spectacular: the team found more than 10 million SNPs, and tens of thousands of other kinds of mutations.

"We found an astonishing amount of variation," says Peer. Although this was not in itself surprising – microbes are renowned for their ability to mutate and evolve rapidly – it reflects the enormity of the challenge that faced Peer's team when they tried to make sense of all the data. The team analysed more than a trillion (a million million) letters of microbial DNA sequence and found that each person harbours a unique set of microbial strains. The abundance of these



Your gut harbours a unique set of bacterial strains.

strains can change with diet, but the strains themselves remain constant. “You have an *E. coli* and it stays yours,” says Peer.

The success of the project arose from a collection of new developments rather than one big innovation, says Peer. The data analysis demanded that the team devise a range of new computer programs, customise other software originally designed for use in human genome analysis, and then connect everything up so that the programs worked effectively together. They also had to refine various statistical techniques to mine the sequence data. “It wasn’t one single cool idea – there were lots of little things needed to get there.”

It’s early days yet, but there are several potential applications of the work. One is screening patients for antibiotic resistance. This would allow doctors to pick a suitable antibiotic to use straight away, instead of using trial and error as they normally do at the moment. “We could even say how many genes in you might be resistant to a particular antibiotic,” says Peer.

Another is to trace the development of disease. Researchers believe that most disease-causing microbes start off being harmless, but then acquire new genes or mutations that allow them to invade and harm their hosts. “Sometimes there are just a few point mutations that turn a normal bacterium into a nasty one,” explains Peer, “and nobody really knows where they come from.”

The ultimate dream is the development of truly ‘personalised’ medicine: treatments that are specifically tailored to your genetics and physiology. In the case of your gut bacteria, this could involve a tailor-made diet or probiotic yoghurt designed to promote health or treat disease. “This is far in the future,” says Peer, “but this at least opens up those possibilities.”

This all raises an important question: where do all these strains come from? Biologists know that babies born via natural delivery pick up bacteria from the birth canal, whereas those born via Caesarean section acquire them from the people who handle them after they are lifted from the womb. But the contents of a young baby’s nappy are probably only the start of the story: it is likely that we pick up most of our bacteria in the first few months or years of our lives, before finally arriving at our stable, unique microbiome at a later stage. But if this is true, where do these other microbes come from? The sandpits we play in as children; the places we visit; the people we meet? How do they all conspire to turn an already unique new baby into an even more distinctive adult? “These are basic questions that we can address now,” says Peer.

Schloissnig S et al. (2013) Genomic variation landscape of the human gut microbiome. *Nature* **493**: 45-50
DOI: 10.1038/nature11711.

Knotty problems

Inside each of our cells' nuclei is a tightly wound 'string' that contains the instructions for making us and keeping us functional. The cellular machinery that reads these instructions moves along the string like a bead on a rope, in a process known as transcription. Lars Steinmetz's group at EMBL Heidelberg has now discovered that the way our genetic rope is bent determines the direction in which the transcription machinery tends to move.

"What kind of a knot is formed there?"

The 'string' of our genetic information is formed by two intertwining strands of DNA and the transcription machinery can move along either of these, but in opposite directions, like the two lanes of traffic on a road. Three years ago, Lars' group discovered that when the transcription machinery lands on most promoters – the DNA sequences that mark where transcription should start – it moves not only along the gene, but also along the other 'lane' of DNA (see EMBL Annual Report 2008/09, page 52). To their surprise, however, this didn't happen everywhere: some genes, it seemed, had promoters that worked only in one direction.

While Lars and his team were trying to untangle this mystery, their collaborator Nick Proudfoot and his group at Oxford University in the UK discovered that genes can bend into a loop, so that when the transcription machinery reaches the end of the gene, it finds itself back at the beginning and starts again. Nick's group also found that inactivating a specific protein prevented these gene loops from forming.

Wondering whether there could be a connection between gene loops and transcription direction, Lars, Nick and Judith Zaugg from Nick Luscombe's group at EMBL-EBI, turned off that protein, and looked at what happened to those promoters from which transcription seemed to happen only in one direction. Looking throughout the whole genome of yeast cells, they found that when genes with 'one-way' promoters were not able to form loops, transcription from those promoters became bi-directional. "Where before there was

Lars Steinmetz

no transcription, now you saw transcription happening," says Lars.

It seems that with no transcription 'beads' trapped in a loop, more are free to move in the opposite direction. And because both strands of DNA can contain information, as the transcription machinery moves along the other strand it can read out sequences of instructions that affect how other genes are read.

"So by forming a gene loop, cells can drive transcription mainly in one direction, increasing output from that gene," says Lars. And if, as the scientists expect, cells can undo gene loops – for instance, in response to changes in their environment – then the cell could damp down the output from that gene and, at the same time, influence other genes. "This could spread regulation throughout the genome," Lars concludes.

Additionally, the scientists found that undoing gene loops also increases transcription starting in the middle of genes. So it could be that looping a gene makes it difficult for the transcription machinery to hop on 'midway', guaranteeing that the whole gene is transcribed.

To have any such effect, however, the gene loop itself has to be formed by a first pass of the transcription machinery – a step Lars and colleagues would like to know more about. And of course, these new findings beg the question: What happens in places the scientists had found before, where transcription *is* bidirectional, with two genes being heavily transcribed at the same time? "Is each gene forming a loop? What kind of a knot is formed there?" Lars posits, as he ponders the questions still left to untangle.

Tan-Wong SM, Zaugg JB, Camblong J, Xu Z, Zhang DW, Mischo HE, Ansari AZ, Luscombe NM, Steinmetz LM, Proudfoot NJ (2012) Gene loops enhance transcriptional directionality. *Science* **338**: 671-5.

DOI: 10.1126/science.1224350.

Imre Gáspár, Anne Ephrussi and Ivo Telley

A force for change

It's often the sudden, accidental discoveries that grab the limelight in the history of science. But many more significant advances in research are the result of months, or even years, of painstaking work and dogged perseverance. Tenacity like this recently paid off for Ivo Telley, Imre Gáspár, Anne Ephrussi and Thomas Surrey at EMBL Heidelberg, who spent many months developing a new method that will allow scientists to probe the forces at work inside cells in much greater detail than before. What's more, they have already used the technique to solve a long-standing cell biology puzzle: how very young embryos 'know' how many cells to make. "It really needed a stubborn person to do this," says Ivo.

The work relates to a cellular structure known as the mitotic spindle, which is essential for cell division. When cells reproduce, they copy their chromosomes and then divide in two. Crucially, they must ensure that each new cell gets one copy of each chromosome – if this goes awry, conditions such as Down Syndrome and cancer can result.

The mitotic spindle consists of rope-like structures called microtubules. In the final stages of division, the copied chromosomes line up in the middle of the spindle along the

midline of the cell. One end of each spindle microtubule attaches to a chromosome, while the other is pegged down by one of the two structures called centrosomes that are on opposite sides of the cell. Radiating from the other side of the centrosome is a starburst of more microtubules, collectively known as asters. When the time comes, the spindle microtubules yank the copied chromosomes apart and into the correct daughter cell. The trouble is, no-one knows exactly how they do it.

"The spindle is really one huge machine," says Ivo. Although scientists have gained many groundbreaking insights into the spindle by studying whole cells, they were limited in their ability to directly probe the forces being generated by microtubules. "We wanted to have a system in which we can perform mechanical manipulations of the spindle," explains Thomas, who now works at the London Research Institute in the UK.

Previously, biologists had studied spindles by extracting them from frog eggs, which are large and widely used in cell biology. A key drawback of this approach, however, is that it is extremely difficult to do any kind of genetic experiments in frogs. Such experiments would allow scientists to further test their hypotheses by, for example, using genetics to disable a key protein involved in spindle formation and then observing its effects on cell division.

Fruit flies, in contrast, are famed for their use in genetics, and while he was working on spindles from frog eggs, Ivo got talking to Imre, who was working as a joint postdoc between Anne's and Thomas's labs. Anne's team studies fly oocytes (developing egg cells) and early embryos, and so had some of the tools needed to start developing individual fly embryo extracts.

Biologists already make extracts from fly embryos to study their contents, but usually do so by mashing up thousands of embryos, all at slightly different stages of development. "The main problem is that it destroys the fine architecture of the cell," says Imre, who had developed a method for extracting the contents of single oocytes. Assuming that a similar approach might work with a single embryo, after months of painstaking work they eventually developed a method that involved using a coarse needle to suck the material out of individual embryos.

A key innovation was recognising a vital, but fleeting point in time during which the nuclei were at the correct stage to be removed. "You have about 2 minutes to do the extract," explains Ivo. "It's a very tight time window." Once extracted, the nuclei continued to cycle through their normal divisions and could be studied in a drop of fluid on a microscope slide. As well as pinpointing the right time, Ivo had to customise his equipment for the task, with invaluable input from the EMBL mechanical workshop and the electronics department.

> "It really needed a stubborn person to do this"

Although the method is tricky, "it's something you can learn," says Ivo, who has been inundated by enquiries from interested cell biologists around the world. "I was over-run!" he says. The team has now published full details of the method for the benefit of the biology community.

To test the technique further, the EMBL scientists used it to investigate a problem that has long intrigued developmental biologists: how the nuclei in a developing insect embryo space themselves correctly. Insect embryos are unusual in that during the earliest stages of development, their nuclei aren›t separated from each other by cell membranes, as they are in mammalian embryos. Instead, after the nuclei of the egg and sperm fuse during fertilisation, the resulting embryo nucleus divides rapidly. The nuclei then move to the perimeter of the egg, the cortex, where they somehow space themselves at a fixed distance from one another. Cell membranes form between them at a later stage.

Exactly how the nuclei 'know' where to go and how far apart from one another they should be has puzzled biologists for years.

"It's an interesting problem, because the embryo is huge in comparison with the nuclei," says Thomas. Although researchers have suggested various explanations, many of which involve microtubules, it has been impossible to test these properly because it is so difficult to see what is going on in whole embryos.

Thanks to the new technique, Ivo was able to track the nuclei under the microscope and precisely measure the distance between them. He found that the nuclei hold themselves 28 micrometres apart, about a third of the width of a human hair. To see whether this distance is fixed, or whether it varies depending on the size of the embryo, Ivo placed a single nucleus in a tiny chamber that measured roughly 30 micrometres across, and waited for the nucleus to divide. Instead of altering themselves to fit the space, the spindles buckled, showing that their size is indeed fixed.

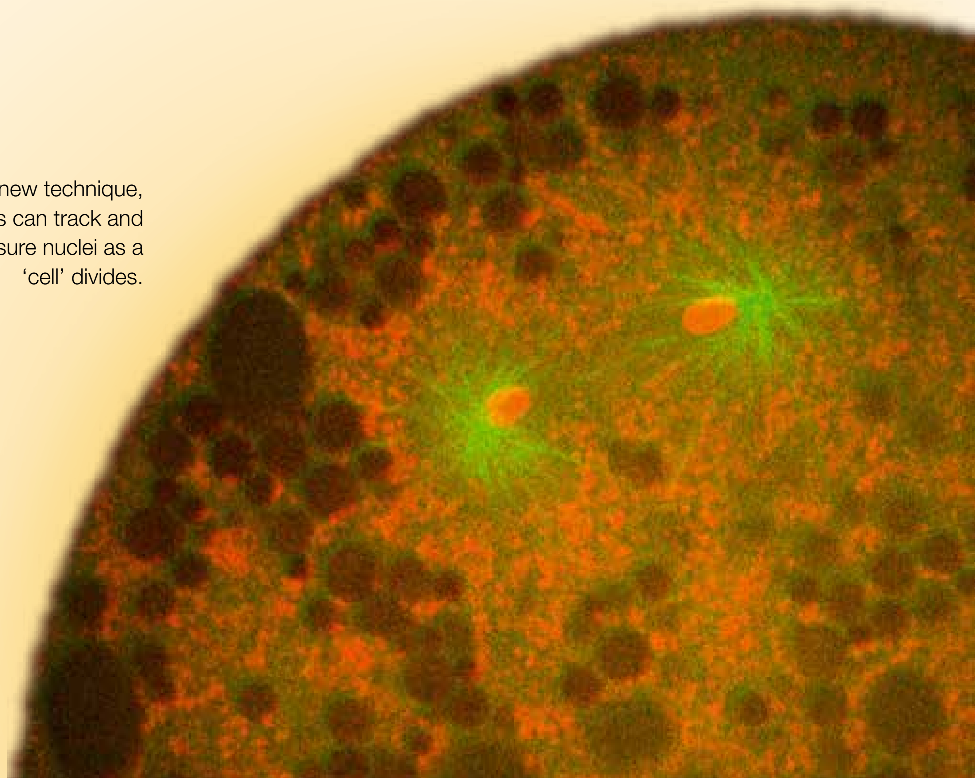
Using a variety of chemical treatments and by cutting parts of the spindle with a laser, Ivo found that it wasn't the central part of the spindle that was responsible for spacing the nuclei, but the asters. "This was quite surprising, because asters are typically thought to be linked to the cell cortex, which was not present in the extract," says Anne. Further work suggested that the asters might be pulling the nuclei by hanging onto part of the cell's internal scaffolding, a meshwork of a protein called actin.

What does this all mean? Ivo wondered whether this spacing would allow insect embryos to make the right number of cells for their size: small embryos would make fewer cells, whereas large ones, like those of the house fly, would keep dividing their nuclei and space them at about 28 micrometers until the embryo was filled with the correct number. A search through the literature showed his hunch was correct: nuclei in the early embryos of many insect species are spaced by about this distance. "It's a fantastic result," says Ivo. "This study links a cell biological problem to an evolutionarily conserved mechanism."

ONLINE EXTRA

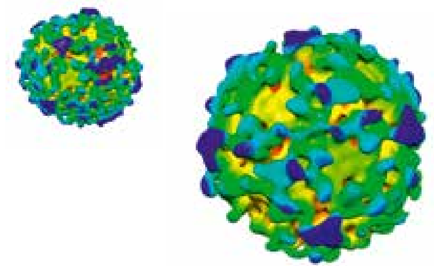
Watch a membrane-less 'cell' divide on the EMBL YouTube Channel: youtu.be/7krCUBuHvpw.

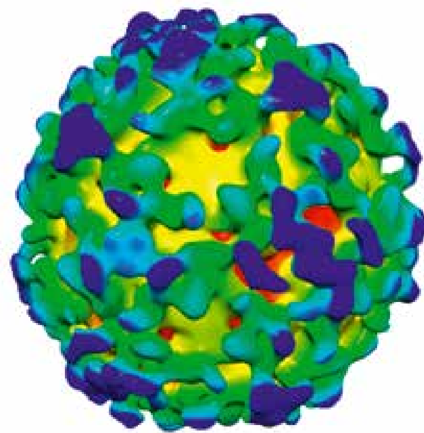
Telley IA, Gáspár I, Ephrussi A, Surrey T (2013) A single Drosophila embryo extract for the study of mitosis ex vivo. *Nat. Protoc*. **8**: 310-24
DOI: 10.1038/nprot.2013.003.

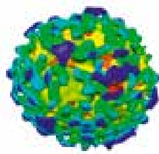


With the new technique, scientists can track and measure nuclei as a 'cell' divides.

Precious parcels







The cell's transport vesicles change shape for different cargoes.

When you want to send a present to a friend, you don't just stick a stamp on it with a note about where it has to be delivered. You package it up into a parcel, which not only protects the goods inside but also allows a label to be added so that your gift can be tracked as it wends its way to its lucky recipient.

The internal postal system of the cell follows a similar logic. Proteins that have to be moved from one place to another, for example, are bundled up into tiny cellular parcels called vesicles. These packages are usually made up of a membrane similar to the one that encloses the cell, combined with a coat of proteins that helps ensure that the vesicle reaches its correct destination.

There are a number of these protein coats, the most well-known of which are clathrin, COPI and COPII. Clathrin-coated vesicles ferry cargo to and from the plasma membrane that forms the outer surface layer of the cell. Other vesicles are specialised in moving cargo between different structures inside the cell, for example the endoplasmic reticulum (ER) and the Golgi apparatus, both of which are involved in protein synthesis. COPI-coated vesicles deliver material from the Golgi to the ER whereas vesicles coated in COPII move in the opposite direction ferrying cargo from the ER to the Golgi.

In the past, biologists have used electron microscopy to study the cage-like structures formed by these vesicle-coating proteins, especially clathrin and COPII. However, these studies have often looked at simplified cellular parcels – consisting of the outer protein cage but lacking the inside membrane – which behave differently from their real biological counterparts.

So John Briggs, group leader in the Structural and Computational Biology Unit at EMBL Heidelberg, recently set out to determine the structure of a more realistic vesicle. Using cryo-electron tomography, his team looked at individual COPI-coated, membrane-bound vesicles, and were able to work out how the COPI protein complex arranges itself into a coating lattice. And whereas previous models of vesicle coats based on studies with clathrin and COPII suggest that they form regular cages – with each component of the cage making the same connections with other components throughout the coat – John's studies, published in *Science*, paint a very different picture.

They revealed that the components of COPI, called coatomers, assemble into three-part building blocks called triads, which form the basic structural unit of the protein cage. Importantly, they also showed that the subsequent assembly of these triads is not regular. "The triads can link up with different numbers of neighbours, and this causes the coatomers to adopt different conformations and connect with each other in distinct ways," says John.

The arrangement of these coatomer triads is important because it affects the shape and size of the resulting vesicle. So regulating the way triads link up provides the cell with a means to optimise vesicles for carrying different cargoes – just as parcels come in a range of dimensions and shapes for shipping books, clothes or laptops. The next step for John's group is to work out just how the cell ensures that it makes the right parcels for the specific goods that need shipping.

Faini M, Prinz S, Beck R, Schorb M, Riches JD, Bacia K, Brügger B, Wieland FT, Briggs JA (2012) The Structures of COPI-Coated Vesicles Reveal Alternate Coatomer Conformations and Interactions. *Science* **336**: 1451-4

DOI: 10.1126/science.1221443.

Stephen Cusack

Faultless fit

It's a classic childhood prank: jamming a potato up a car's exhaust pipe to make the engine stall. But a prankster can't use any old potato – if it's the wrong size or shape, it won't fit the pipe properly and the engine will keep running. Stephen Cusack and his group at EMBL Grenoble have now brought us a step closer to pulling off a similar trick on the influenza virus, after determining the structures of a key viral enzyme bound to drugs designed to jam it. The work is already helping them to design drugs that fit the enzyme more tightly, and which could therefore make it stall completely. The findings could bring the dream of a new universal anti-flu drug closer to reality.

Most of us will have experienced the feverish misery of an infection with a 'seasonal' flu strain – one of the variants of the influenza virus that regularly circulate in the community. These viruses remain similar from year to year, meaning that most people have some degree of immunity against them. Even so, seasonal flu can be deadly for babies, the elderly and people with underlying health problems and it kills up to half a million people around the world each year.

Once in a while, however, a new, dramatically different flu virus emerges and spreads quickly around the world. Such 'pandemic' viruses can quickly kill young, healthy individuals as well as the more vulnerable. The infamous flu pandemic of 1918 killed between 30 and 50 million people. More recently, a new pandemic virus, called H1N1, emerged in 2009. Nicknamed 'swine flu' because it originated in pigs, this strain was far less deadly than the 1918 virus. Nonetheless, it was a chilling reminder of the ability of such strains to appear unpredictably and sweep around the world with alarming rapidity. Of particular concern is a highly virulent strain called H5N1, commonly known as bird flu. Although mainly confined to birds, if it transfers to humans, the mortality rate is around 60%. Indeed, another bird flu strain, H7N9, has broken out sporadically in China in the spring of 2013 again causing several human fatalities. If either of these strains becomes transmissible between humans there could be devastating consequences.

Thanks to decades of medical research, however, we are now no longer entirely defenceless against flu. Vaccinations against seasonal flu viruses are generally effective and there are a number of drugs, such as osteltamivir (Tamiflu), which can slow the spread of the virus in the body.

But vaccines that work against one strain are usually ineffective against others, and take months to produce. And thanks to their ability to change and evolve rapidly, flu viruses are becoming resistant to the drugs currently used against them. So it's no surprise that new flu treatments are the subject of intense research.

A key focus of this research is the enzyme that the virus uses to replicate, its polymerase. Blocking this enzyme with drugs could halt or slow the course of infection. The polymerase is made

up of three main parts, or subunits, and two of these – the PB2 and PA subunits – are responsible for the virus's ability to hijack the host cell's protein-production machinery, duping it into producing viral proteins.

Stephen and his colleagues at the international Unit for Virus Host Cell Interactions (UVHCI) in Grenoble first identified the role of the PA subunit in 2009, when they determined the structure of part of the subunit. PA acts as an endonuclease and this work showed that, with the help of PB2, it clips a small section from host molecules called mRNAs that carry information between the cell's genome and its protein-production machinery. This clipped section is rather like a password, telling the cell that the mRNA is a bona fide part of the protein-production process. By adding this 'password' to its own mRNAs, the virus fools the host cell into producing viral proteins. This process is known as 'cap-snatching' and is unique to the family of viruses to which influenza belongs. At the time, the team noted that knowing the structure of this part of PA could aid the development of drugs designed to block its activity (for more on this work, see EMBL Annual Report 2008/2009, page 28).

Chemical compounds that inhibit the flu endonuclease have been under development since the 1990s, but with limited success: none has yet reached the clinic. To find out more, Stephen and his team decided to study the structure of the PA endonuclease while it was bound to some of these known compounds. The activity of the endonuclease depends on its 'active site', a deep pocket in its structure that binds and then cleaves the mRNAs. Compounds that block the nuclease likely bind to this pocket, but little was known about how well they fitted into it – the closer the fit, the better the blocking activity of the inhibitor.

Stephen and his team crystallised the endonuclease from the 2009 H1N1 swine flu virus in the presence of several known inhibitors and determined the three-dimensional atomic structure of the complexes by the powerful technique of X-ray crystallography. Key to the success of the work were the High Throughput Crystallisation Facility at EMBL Grenoble and the high-intensity X-ray beamlines at the European Synchrotron Radiation Facility (ESRF).

ATTACK ON ALL FRONTS

EMBL has joined forces with industry to turn its discoveries into new drugs to tackle flu. The collaboration with industry partners Savira and Roche has resulted in several lead compounds, which are in the early stages of the development pipeline. As well as receiving funding from the two companies, the research is being supported by FluPharm, an FP7 project funded by the European Commission.

The story of the collaboration began about 5 years ago when Stephen Cusack, Darren Hart, Rob Ruigrok and their UVHCI colleagues determined the structures of two potential influenza drug targets (see main story). They submitted the findings as invention records to EMBL's technology transfer organisation, EMBLEM. The EMBLEM team then identified a Vienna-based company called onepharm as a partner to develop drugs based on these structures.

EMBLEM and onepharm jointly founded a new company, Savira, to take the work forward. The pharmaceutical giant Roche, which markets the anti-flu drug Tamiflu, then became interested and joined the collaboration. Savira not only has in-house expertise in testing compounds in biochemical and cell-based anti-viral assays but also manages the overall project. This involves outsourcing different strands of the research, such as synthetic chemistry, toxicology studies and animal trials, to expert teams who are partners in the FluPharm project or at Roche.

Stephen's team continues to crystallise the fragments of the viral proteins bound to the potential drug compounds produced by FluPharm to help develop them further. "It's still a few years before anything enters clinical trials," says Stephen. "But it's very important, so it's worth waiting to come up with the right compound to move forward with."

Striving to avoid a repetition of the 1918 flu pandemic.

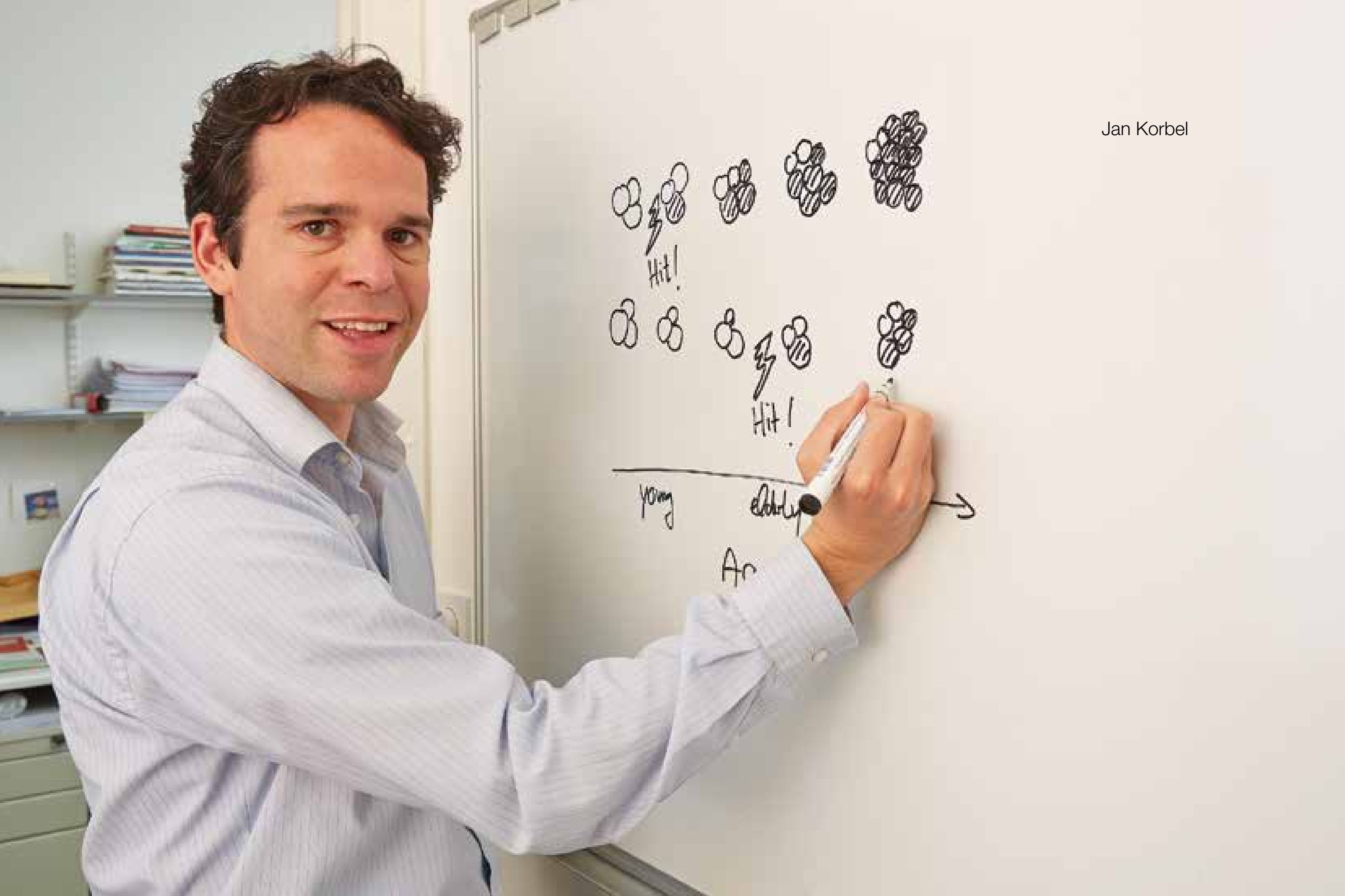
Stephen's team found that different inhibitors bound to different parts of the active-site area. The innermost part, the enzyme's active centre itself, has to be a very specific shape because it is the part that binds and then cuts the mRNAs. In fact, this is achieved by two metal ions that form an integral part of the enzyme. Thus there is very little scope for the enzyme to change and yet still work effectively, so the virus is less likely to be able to evolve resistance to any drugs that bind directly to the active centre. What's more, such a drug would probably be effective against all flu strains, including any new ones. The areas surrounding the site, however, have more scope to change. "One of the lessons from the study was to be aware of where these were," says Stephen, "and when you are trying to improve the compounds, avoid the places where there may be changes."

None of the drugs bound snugly into the active site, meaning they were less likely to be effective inhibitors of the enzyme than compounds that bind more tightly. Even so, the findings were revealing. All of the drugs bound to the two metal ions lying at the heart of the active site. "This is obviously a key feature which has to be built into all endonuclease inhibitors," says Stephen. More clues as to how to design compounds that fit well came from a molecule called EGCG, which is found in green tea and was recently shown to block the endonuclease activity. Stephen's team found that EGCG does indeed bind the endonuclease's active site, although the molecule's chemistry makes it an unsuitable candidate for further development, says Stephen. "It clearly does fit rather well, and gave interesting insights into how to design other molecules to do so."

Stephen and his team are now working with pharmaceutical companies (see box) to develop their findings into a drug that can tackle any flu virus. "That's the goal, but I think one has to be a bit humble," cautions Stephen. "The virus generally finds a way around things." Nonetheless, such a drug could be used in combination with other drugs, making it harder for the virus to evolve resistance to all of them at once. "The important thing is to have several options," says Stephen.

Kowalinski E, Zubieta C, Wolkerstorfer A, Szolar OHJ, Ruigrok RWH, Cusack S (2012) Structural analysis of specific metal chelating inhibitor binding to the endonuclease domain of influenza pH1N1 (2009) polymerase. *PLoS Pathog.* **8**: e1002831

DOI: 10.1371/journal.ppat.1002831.



Jan Korbel

Age matters

The prostate, a walnut-sized gland in a man's reproductive system, plays a crucial role in the creation of human life by making the fluid in which sperm swims. But in cases of extreme malfunction, it may also hasten death. Despite being one of the most common types of cancer in men, the causes of prostate cancer are largely unknown. Equally puzzling are its effects: some patients experience only mild symptoms whereas others suffer from tumours that are highly aggressive and life-threatening. In February, Jan Korbel's team at EMBL Heidelberg and colleagues announced in *Cancer Cell* that they had uncovered yet another fascinating aspect to the disease: the mechanisms underlying prostate cancer cases in younger patients are strikingly different from those in older men.

While better known as a disease of the elderly, around 2% of prostate cancer cases occur in patients under the age of 50. The study revealed that such early-onset prostate cancer seems to be driven by 'male' hormones leading to very specific DNA rearrangements in the cells' genome, whereas a different mechanism appears to operate in prostate cancer occurring in elderly men.

The work was carried out as part of a German contribution to the International Cancer Genome Consortium on Prostate Cancer, a consortium that involves scientists and clinicians from several institutions in Germany in Heidelberg, Hamburg and Berlin. The main goals of the consortium are to unravel mechanisms leading to prostate cancer formation, and also to pinpoint genetic markers that could indicate people most at risk of acquiring the highly aggressive disease. For this purpose, the team aims to sequence and analyse the genomes of 200 prostate cancers occurring in young patients.

The idea may appear straightforward: find the genetic changes that turn normal cells cancerous, and then work towards developing new diagnostic tests, or even develop new cancer treatments. "But this is easier said than done," Jan explains. "The genome of a cancer cell is highly complex. Countless sections of DNA can be deleted, duplicated, flipped or relocated in a cancer genome. Some of the DNA changes we observe seem to drive cancer cells to reproduce uncontrollably, whereas others appear to be so-called passenger alterations that just tag along. To potentially enable new clinical breakthroughs in the future, we are trying to understand the basic mechanisms by which prostate cancer can arise."

Is has long been known that this chaotic mess in the genome can involve both mutations of individual DNA bases (so-called point mutations) and larger changes in the structure or number of chromosomes. Research investigating the specific role of such DNA alterations in cancer has been boosted by the recent advancement of technologies that enable the rapid sequencing of whole cancer genomes.

In the study led by Jan's team, the genomes of tumour cells isolated from the cancer cells of 11 early-onset patients were sequenced and compared side-by-side with existing data from seven elderly-onset patients. They found that the DNA of cells from tumours in younger patients harboured a relatively small number of changes compared with that from older patients. However, this small number of changes led to far-reaching structural alterations of chromosomes.

"We observed that sections of the DNA coding for genes that are normally present on different chromosomes can become broken and subsequently joined together, thereby altering the code of both genes in such a way that this results in a new gene – known as a fusion gene," Jan explains. "The majority of fusion genes in early-onset prostate cancer patients seem to be formed when androgens, such as testosterone, bind to their receptor protein. The androgen receptor can then bind multiple pieces of DNA at the same time, and thereby bring DNA sections from different chromosomes closer together. While this process also happens in healthy cells, errors in the process lead to broken chromosomes that can become joined together in the wrong order. The resulting gene fusions can contribute to the growth and spread of cancer, because genes with the potential to cause cancer can be switched on through these rearrangements," says Jan. "It is as if you swap words or sentences in a book: the meaning changes. Fusion genes in prostate cells can lead to uncontrollable cell growth and cancer."

Working with partners at the University Clinic Hamburg-Eppendorf, the team was able to verify what they saw in tumour tissue from more than 10 000 patients. "We found that the androgen receptor is more active in tumour samples from younger individuals than older ones, and that gene fusions are more common in young than in elderly patients. What is more, based on earlier research we already knew that testosterone levels are particularly high in young men," Jan explains. "Our research results suggest that young men with particularly high testosterone levels may perhaps be at specific risk of developing prostate cancer, although we were not yet able to investigate this question as information on testosterone levels in the group of patients that we investigated was not available."

"Young men with particularly high testosterone levels may be at risk"

Crucially, their findings could help researchers to stratify approaches to a formidable challenge. "As life expectancies rise in the Western world, cases of prostate cancer are likely to increase and it is urgent that better diagnostic methods are developed," says Jan. "Our findings of age-dependent genetic rearrangements in prostate cancer suggest that clinicians could specifically search for gene fusions to screen for prostate cancer in young men in the future. Although it is clear that further research will be needed for this, our recent findings may even help to improve the future treatment of prostate cancer."

Weischenfeldt J *et al.* (2013) Integrative genomic analyses reveal androgen-driven somatic alteration landscape in early-onset prostate cancer. *Cancer Cell* **23**: 159-70

DOI: 10.1016/j.ccr.2013.01.002.

Let's get together

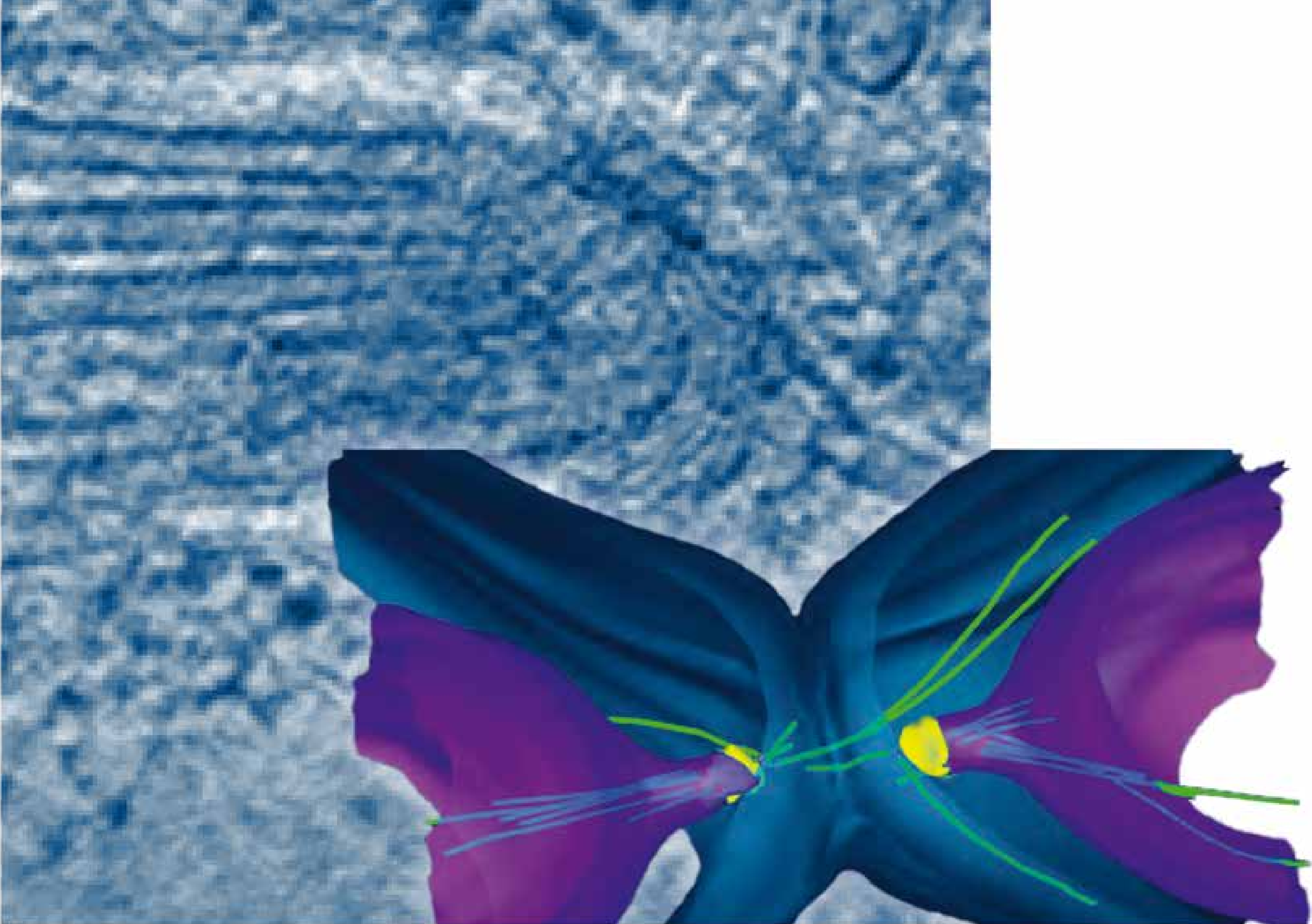
Many couples, when their relationship gets serious, merge their financial resources into a joint account as a signal of long-term commitment. Sexually reproducing species do something similar with their genetic material when their sex cells – sperm and eggs in humans – come together. Each sex cell has a diminished 'genetic account' and carries only half the usual amount of genetic material – known as the 'haploid' state – but the genetic bank balance is restored to the full 'diploid' state when a sperm fuses with an egg during fertilisation.

The microtubules reach out directly to the partner nucleus

The situation is inverted in our evolutionarily distant cousins, the single-celled yeast. These organisms normally exist in a haploid state, but under the right conditions they can fuse with another haploid yeast cell to create a diploid organism containing twice as much 'genetic capital' as usual. This diploid cell is usually transient and soon divides to generate two daughter cells in which the 'genetic bank balance' is restored to the normal haploid state.

A key step in this process is the fusing of the haploid cells' nuclei, which house their genetic material. To achieve this, the two nuclei have to be pulled together once the mating cells have fused into a single cell. Last year, research carried out by EMBL postdoctoral fellows Romain Gibeaux and Antonio Politi under the supervision of Claude Antony and François Nédélec, group leaders in the Cell Biology and Biophysics Unit at EMBL Heidelberg, and EMBL alumnus Michael Knop, now at Heidelberg University, helped to uncover exactly how those two nuclei come together.

Nuclei are not the only cellular components that have to be moved around the cell. A common way to do this takes advantage of the cellular scaffolding formed from protein filaments called microtubules (MTs). In addition to providing structural support for the cell, MTs can also be used as tracks along which protein motors can pull their cargo. Before Romain and Antonio's work, it was known that MTs emanate from the nuclei after yeast cells fuse, so it seemed likely that the nuclei also exploited this system to move towards each other. This led to a model of nuclear migration in which the MTs grow out of each nucleus and extend towards the other until they overlap, at which point they latch on to each other and pull themselves together. Romain and Antonio, however, found that this is not quite how things work. Using a number



Mating yeast cells pull on each other's nuclei.

of different techniques – including electron tomography, molecular genetics, and imaging of live cells – they found that MTs do not interact with each other directly at all during nuclear migration. Instead, they found that the MTs growing out of one nucleus reach out directly towards the partner nucleus, where they bind to a structure called the spindle pole body (SPB) through another protein, a molecular motor called Kar3.

The SPB is anchored in the nuclear membrane, and Kar3 is in turn anchored to the SPB. Kar3 functions like a steam engine and pulls the SPB (and the attached nucleus) along the MT track towards the other nucleus. This continues until the SPBs of the two nuclei come into contact, at which point nuclear fusion begins.

Based on their biological data, the EMBL scientists created computational models of this process, which revealed that it's not necessary to have MTs growing from both nuclei – just as a transfer of money into a joint account only requires one person to have the relevant bank details. "We showed that even if microtubules extend from only one nucleus, the two nuclei can be bought together so long as they bind to Kar3," says Romain.

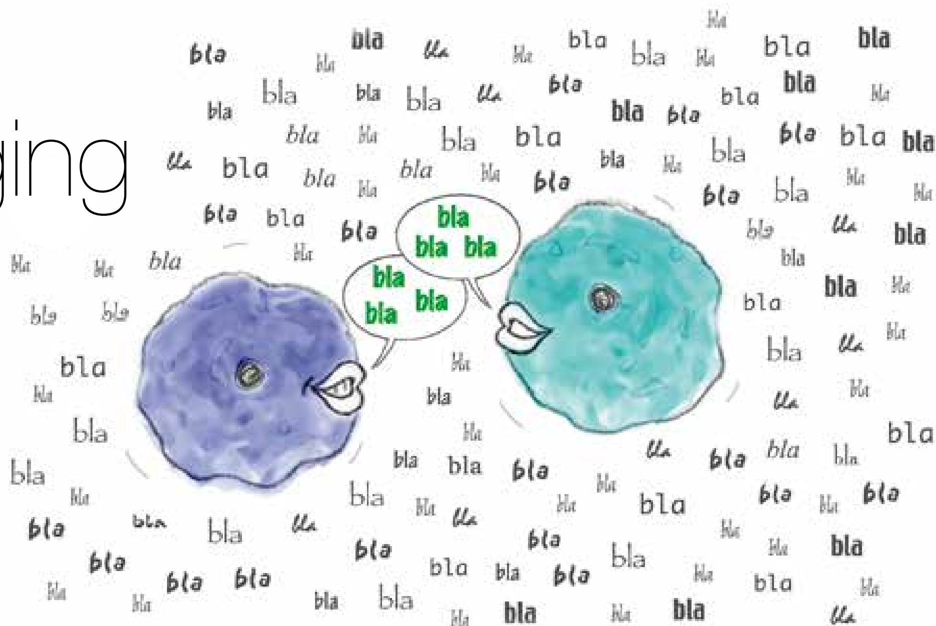
ONLINE EXTRA

Watch nuclei pulling on each other's microtubules on the EMBL YouTube Channel: youtu.be/uC1GjTvAjzw.

Gibeaux R, Politi AZ, Nédélec F, Antony C, Knop M (2013) Spindle pole body-anchored Kar3 drives the nucleus along microtubules from another nucleus in preparation for nuclear fusion during yeast karyogamy. *Genes Dev.* **27**: 335-49

DOI: 10.1101/gad.206318.112.

Better bugging



New technique makes cells 'talk with an accent'.

From James Bond to The Avengers, when Hollywood spies walk into a room that may be bugged, they know to turn on a radio so that any bad guys attempting to listen in are unable to distinguish the spies' words from the radio's babble. Scientists wanting to 'listen in' on conversations between cells may not use microphones dipped into the dishes and tubes in which cells are kept, but they are often in the same situation as the super-villains: struggling to hear through the noise. Now, Jeroen Krijgsveld and his group at EMBL Heidelberg have developed the equivalent of a high-tech sound-filtering gadget that solves the problem.

While humans talk by making and detecting vibrations in the air, cells communicate with each other by releasing – or secreting – and detecting proteins in their surroundings. But cells in the lab have to be fed, and the serum used to feed them also contains proteins – many more proteins than the cells themselves secrete. So scientists attempting to eavesdrop on cells' conversations are faced with a conundrum: either they are stuck in the super-villain's situation of not being able to tell real information from background noise, or they have to eliminate the noise – which means starving the cells.

A new method developed by Jeroen and colleagues offers a third option. Katrin Eichelbaum, a PhD student in Jeroen's group, essentially made the cells talk with an accent that could be easily filtered out from the surrounding noise. She coaxed cells into building their proteins with an artificial amino acid instead of the methionine they would normally use. Thus, proteins produced by the cell stand out from those in the feeding serum, and researchers can then fish them out using a technique called click chemistry (for more on this technique, see EMBL Annual Report 2011/12, page 84). Eliminating the need to starve cells is an important development, as the scientists showed that starving cells, even just for a few hours, affects secretion. Katrin and Jeroen were able to pinpoint the effects of starvation thanks to another advantage of their new method. It allows them to measure exactly how much of each protein the cells have released at 2-hour intervals, so they can follow how secretion changes over time.

Working with Stephan Herzig at the German Cancer Research Centre (DKFZ) in Heidelberg, Katrin and Jeroen also found grounds for caution when using cell lines grown in the lab as a proxy for primary cells – cells taken directly from an organ. "If you take a liver cell line, you assume that you're looking at liver cells," says Katrin, "but we saw that the secreted proteins were really different between the liver cell line and primary liver cells."

Additionally, the double advantage of not having to starve cells and being able to follow changes over time enabled the EMBL scientists to be the first to follow how white blood cells called macrophages – which can't be grown without serum – react to bacteria to kick off a rapid immune response. As well as continuing to spy on cell communication, Jeroen's lab now plan to use their new approach to study how cancer cells respond to drugs.

"There's much more for the community to explore," Jeroen says. "Our method could be used to watch how cells react to drug treatments; or to search for biomarkers, like the proteins cancer cells release that help them invade tissues; or to see how secretion changes if cells are grown in 3D instead of on a regular Petri dish. We've really seen a great deal of interest already."

Eichelbaum K et al. (2012) Selective enrichment of newly synthesized proteins for quantitative secretome analysis. *Nature Biotechnology.*
DOI: 10.1038/nbt.2356.

Just one drop

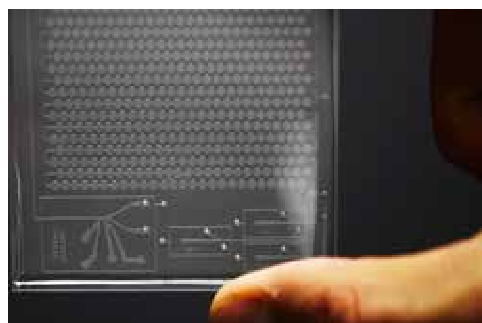
Detergent companies like to boast that a single drop of their product is enough to clean a whole sink full of dishes. Now Christoph Merten's group at EMBL Heidelberg have accomplished an even more impressive feat of drop-based efficiency. They have developed a new technique for producing monoclonal antibodies that requires a million times less sample than conventional methods, dramatically increasing the efficiency of producing these molecules for use as drugs, and potentially enabling people who survive certain diseases to help others fight them.

Monoclonal antibodies are substances produced in the lab that bind to a specific target molecule. This means they can be widely used in medicine, for everything from diagnostics – by 'sticking to' and highlighting a feature that only diseased cells or infectious agents carry – to treatment – for example, by directly inhibiting disease-causing entities. They are also commonly used in biomedical research to determine whether a given sample contains a particular molecule. However, all these applications require scientists to be able to identify the cells that produce the antibody they're looking for. Until now, this meant that cells had to be grown in a lab dish until there were enough to allow the antibody and/or its effects to be detected. This in turn meant that the process took weeks, and only a limited number of cells could be screened in each experiment.

Bachir El Debs, a postdoctoral fellow in Christoph's group, has circumvented these hurdles by scaling down the whole process. Christoph's lab specialises in miniaturisation, using an approach called microfluidics to perform experiments on tiny droplets of liquid (see EMBL Annual Report 2011/12, page 38). They have designed a microfluidic chip that is essentially a screening lab that fits in the palm of your hand. Barely bigger than a coin, this plastic chip houses a maze of tiny channels. Bachir, Christoph and colleagues can use this device to generate droplets containing single cells, and run these droplets through the channels. As they move through the maze, droplets that contain a specific antibody are separated from those that don't. In work published in 2012 in *PNAS*, the EMBL scientists demonstrated that they could reliably detect antibodies produced by single cells, and applied their technique to search for cells producing antibodies against ACE-1, an enzyme that can cause high blood pressure.

Using droplets that are one millionth the size of a conventional test tube not only allows scientists to screen many more cells at a time – the team showed that around 300 000 cells can be screened per day – but also enables work on much smaller samples and abolishes the need to grow cells in the lab, opening up new possibilities.

"Because the number of cells we need in the assay is very low and because it can be performed at a single-cell level, eventually it may be possible to screen antibody-releasing cells from human donors," says Christoph. With this technique, researchers could look for the antibodies that enable some people to successfully fight off a disease, by analysing cells from those people directly. For instance, "Two per cent of all HIV-infected individuals are estimated to develop neutralising antibodies," says Bachir. "You could take antibody-producing cells from these people and use our technique to try to identify cells that neutralise HIV." So in biomedical research labs, 'just one drop' could help to develop new treatments or vaccinations, helping our immune systems to do the cleaning much more efficiently.



A tiny maze separates cells with and without antibodies.

El Debs B, Utharala R, Balyasnikova IV, Griffiths AD, Merten CA (2012) Functional single-cell hybridoma screening using droplet-based microfluidics. *Proc. Natl. Acad. Sci. USA* **109**: 11570–5

DOI: 10.1073/pnas.1204514109.

Martin Beck, Sigrid Milles, Edward Lemke and Alessandro Ori

Unlocking the pore

Deep inside each of our cells lies a precious dataset: our genome, the gigabytes of genetic information encoded in our unique sequence of the four letters of DNA. Eukaryotic cells – that not only make up multicellular plants and animals like us, but also single-celled organisms like yeast – sequester this precious cargo in a membrane-bound nucleus, which keeps the genome stored safely away from the rest of the cell.

But wrapping the genome in a membrane poses a problem: how do the encased genes interact with the rest of the cell? For a gene to produce the protein it encodes, specialised protein assemblies have to gain access to the nucleus. There, they transcribe the DNA and create the intermediary molecule messenger RNA (mRNA), which has to exit the nucleus to the cytoplasm where it is then translated into protein.

To deal with this transport problem, eukaryotes have evolved nuclear pores through which material can pass. But these are not simply holes: the pores are lined with a highly complex arrangement of proteins that determine what can enter and exit the nucleus. Yet precisely how the proteins of the eukaryotic nuclear pore complex (NPC) are arranged, and how they carry out their job, remains unknown.

Enter Edward Lemke and Martin Beck, both group leaders in the Structural and Computational Biology Unit at EMBL Heidelberg. Martin's lab focuses on studying assemblies of large molecules, such as proteins, using electron microscopy and mass spectrometry; Edward's specialises in using light microscopes, high-resolution spectroscopy and chemical biology to study proteins that, in contrast to most proteins, are naturally unfolded or 'intrinsically disordered'. "We use different techniques, but we have a common interest in the NPC," says Edward.

"We use different techniques, but we have a common interest"

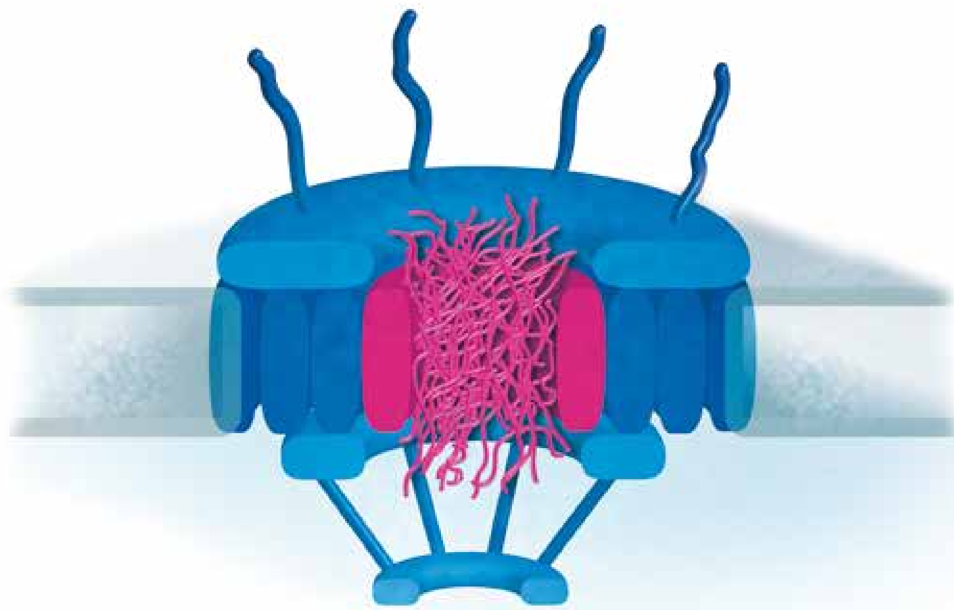
Together with Peer Bork, Joint Head of the Unit, Edward and Martin recently combined their expertise to shed light on the complex dynamics of the NPC. Their findings are described in two papers published in *EMBO Reports* and *Molecular Systems Biology*.

The proteins that make up the pore are known as nucleoporins (Nups), and there are 30 of these proteins encoded by the human genome. They come in two broad classes: 20 are architectural scaffold Nups, which give the pore its shape; and the other 10 are so-called FG Nups, which form a permeability barrier in the pore's central channel that controls what passes in and out. In the *EMBO Reports* paper, the team looked at these FG Nups in more detail.

It was known that all 10 barrier FG Nups are intrinsically disordered, but the team aimed to reveal their molecular interactions and show how they might form a selective permeability barrier. Their first step was to use light microscopy combined with fluorescence tags on the barrier proteins to see how they behaved. "We found that these proteins, in isolation, have a collapsed structure, which leads them to aggregate into tangles called amyloid fibres," says Edward. "We also found that when you have a high concentration of these fibres they can form a structure known as a hydrogel, which is essentially like a gummy bear."

These hydrogels are notable because they have properties similar to the permeability barrier: selected molecules can move through them more easily than others. In fact, some researchers have suggested that the permeability barrier may actually be a hydrogel, and so the next step was to produce a detailed picture of its structure.

This is not possible with a light microscope, but electron microscopy is perfect for the job. Using this tool, the team showed that the hydrogel consists of a web of interlaced amyloid fibres. Light microscopy and electron microscopy together allowed them to create a link between initially uncorrelated observations from the single molecule to the supramolecular level. Martin and Edward believe that the cell must be able to control which form FG Nups adopt – the collapsed or fibrillar state – probably by altering them through post-translational modifications that affect their behaviour.



The nuclear pore's scaffold (blue) and barrier proteins (pink).

Whereas the molecular interactions between the FG Nups that form the permeability barrier were the focus of the *EMBO Reports* paper, the research reported in Molecular Systems Biology looked at both barrier and scaffold proteins to shed light on the overall architecture and organisation of the NPC.

"We're very interested in determining the structure of this complex, but it's extremely intricate," says Martin – although there are only 30 distinct Nups in humans, they are present in multiple copies per NPC, each of which contains about 1,000 proteins. To begin with, the team wanted to work out just how many types of each protein are present per NPC, relative to each other. For example, whereas there may be 60 copies of protein A there may only be 20 copies of protein B. This question was tackled with a tool used in Martin's lab, targeted mass spectrometry, which revealed the relative abundance of the various Nups that comprise the NPC.

With these ratios in hand, the next step was to find out how many copies of each protein there are in the complex. This doesn't require working out the absolute amount of each and every protein, but just one: if you know there are 60 copies of protein A, and there's a 3:1 ratio of A:B, then you know there must be 20 copies of B.

To obtain these absolute numbers the team turned to Edward's speciality, fluorescence microscopy. One of the Nups was labelled with a fluorescent tag and the NPCs were then viewed under a microscope, enabling the number of fluorescent signals to be counted and the copy number of the labelled Nup to be established. Combining this with the protein ratios that were derived earlier gave the copy number of all the NPC Nups.

Martin and Edward wondered whether these ratios and copy numbers would be the same across different biological conditions and cell types. They solicited the bioinformatics expertise of Peer's group for a large-scale analysis and found that whereas the scaffold proteins are always found in the same ratios and copy number, this is not true of those that form the permeability barrier.

This makes sense, say Martin and Edward, because the permeability proteins affect what can pass through the NPC – and this may differ depending on the job the cell has to do. Just how the cell changes the composition of the permeability barrier to adapt to the demands it faces remains a question for future research.

Ori, A *et al.* (2013) Cell type-specific nuclear pores: a case in point for context-dependent stoichiometry of molecular machines. *Mol. Syst. Biol.* **9**: 648,

DOi: 10.1038/msb.2013.4

Milles S, Huy Bui K, Koehler C, Eltsov M, Beck M, Lemke EA (2012) Facilitated aggregation of FG nucleoporins under molecular crowding conditions. *EMBO Rep.* **14**: 178–83

DOI: 10.1038/embor.2012.204.

Making waves

What does a developing embryo have in common with the audience of a football match? More than you might think: some of an embryo's cells can perform the molecular equivalent of a Mexican wave, and doing so ensures that certain body parts end up being the correct size, say Alexander Aulehla and his team at EMBL Heidelberg. "This is a fundamental problem that biologists have been interested in for decades," says Alexander.

> The remaining tissue seemed to 'know' how big it was – but how?

Alexander's team has been studying how embryos form somites, a series of tissue blocks running down an animal's back that gives rise to structures such as vertebrae. A key question is how two differently sized animal embryos scale their somites so that they both end up with the same number.

In the 1970s, biologists found that if they reduced a frog embryo's size by removing cells at an early stage of development, rather than forming fewer somites of the usual size, it formed the normal number but they were smaller. More recent work has revealed waves of gene activity travelling along the developing spine and tail, like a series of Mexican waves in a sports stadium audience. This raised the question of whether the waves were involved in making the somites grow to scale.

The difficulty facing researchers was that live, developing mouse or frog embryos are too bulky to study in great detail under a microscope. So Alexander and his team took samples of tissue destined to become somites from developing mouse embryos and grew them as flat sheets of cells in a lab dish.

Remarkably, the flat sheets showed comparable waves of gene activity, with concentric rings of waves radiating out from the centre of the dish. Rings of somites formed progressively at the perimeter. Intriguingly, the gaps between each ring became smaller as successive bands of somites formed and left an ever-shrinking area of non-somite tissue at the centre of the dish. Clearly, the remaining non-somite tissue seemed to 'know' how big it was – but how?

Alexander Aulehla doing the Mexican wave.

Alexander's team found that the secret lay in how the wave passes through the tissue. Each cell in the tissue experiences a pulse of gene activity that cycles from low to high to low again. The starting point of each cycle is staggered from one cell to the next, with the innermost cell starting first, and the outermost starting last. Intriguingly, the difference between the innermost and the outermost cell is always one complete turn of the cycle, irrespective of the tissue size.

This suggested how embryos could scale their somites: if they have to fit one full turn of the cycle across a tissue, they will have to increase or decrease the staggering – the degree to which the point in the cycle reached by one cell differs from that of its neighbour – according to the size of the tissue. In small areas of tissue, the difference is large; in large tissues, the difference is small. So no matter how big the tissue, the wave completes its pass in the same amount of time.

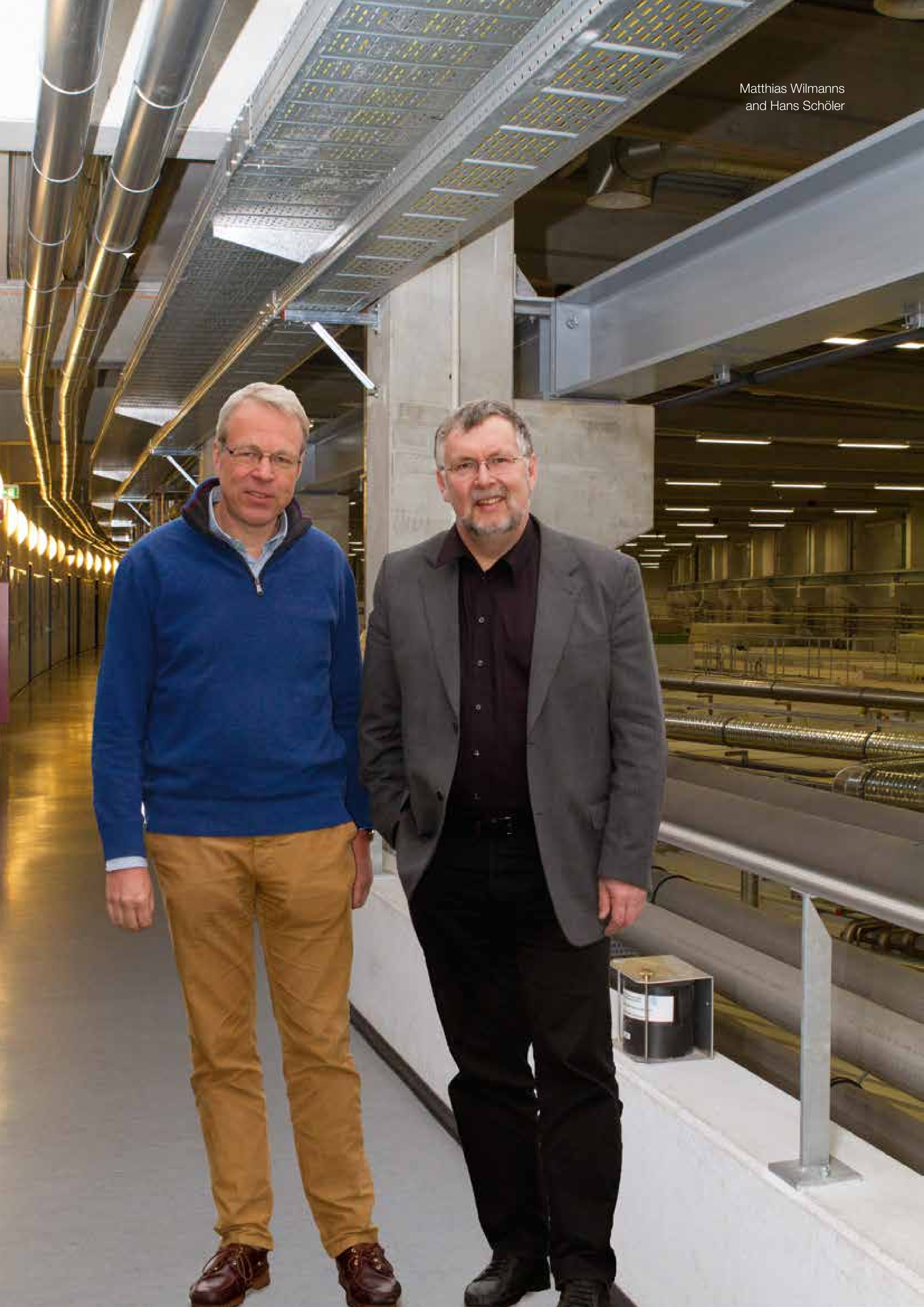
To understand how this works, imagine an individual person in a crowd performing their Mexican wave. In a larger stadium, the next person in the wave is only one stage behind their neighbour: when they are raising their arms, their immediate neighbour is lowering theirs. But in a smaller stadium, the difference is greater: when one person is raising their arms, they are two or more stages behind their neighbour, who has already sat down again.

It's a beautiful but complex system, so Alexander's team is collaborating with theoretical physicist Paul François at McGill University in Québec to build more detailed mathematical models of their findings. They plan to see whether other animals scale themselves in the same way. "This is what is so exciting about it," says Alexander, "it could be a general mechanism."

ONLINE EXTRA

Watch waves of gene activity pulse through a mouse embryo as its vertebrae form: youtu.be/T-fa1gYfwJk.

Lauschke VM, Tsiairis CD, François P, Aulehla A (2013) Scaling of embryonic patterning based on phase-gradient encoding. *Nature* **493**: 101-5
DOI: 10.1038/nature11804.



Matthias Wilmanns
and Hans Schöler

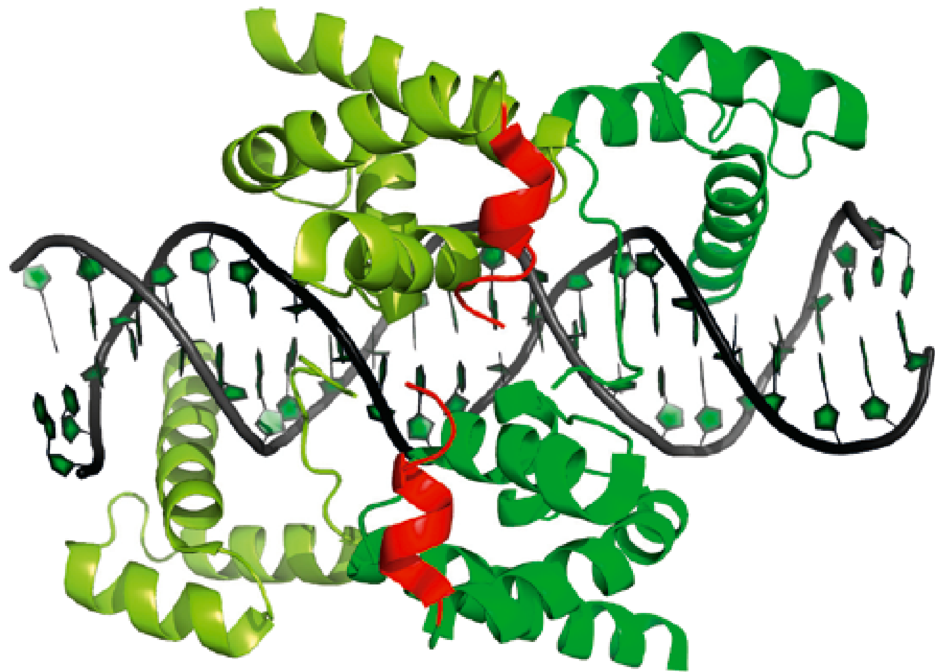
Bridging the gap

Not long after they met while working at EMBL in Heidelberg, Matthias Wilmanns, now group leader and Head of EMBL Hamburg, and Hans Schöler, now director at the Max Planck Institute for Molecular Biomedicine in Münster, started discussing the transcription factor Oct4 and its role in induced pluripotent stem cells. A recent publication in *Nature Cell Biology* describing key details about the structure and function of Oct4 is not only an important milestone in the understanding of how cells are reprogrammed, but also a celebration of 20 years of scientific collaboration and friendship.

"Correct me if I am wrong Matthias," says Hans, looking across the table at his friend, "but although we had seen each other around EMBL a number of times, it was actually our wives who met first." Matthias nods and adds, "Yes, our children were in the same Kindergarten and our wives got talking and became friends and eventually we also got to know each other." The two men then discovered they were neighbours and would regularly ride their bikes together to work at EMBL. "It was during one of these bike rides that we started talking about Oct4," says Hans. The transcription factor Oct4 is a protein that binds to DNA and controls the genes involved in reprogramming terminally differentiated cells. These so-called induced pluripotent cells (iPS cells) can then divide and differentiate into other cell types, but the precise role of Oct4 in this process is still not completely understood. "This was one of my pet topics - Gert Vriend, at the time also at the EMBL and now professor at the University of Nijmegen in the Netherlands, had just done some modelling work on Oct4, and I was eager to see what it actually looks like," Hans adds. "We had started doing some work with mammalian cells, cross-linking proteins with DNA and targeting different genes, and had identified a new type of binding site for Oct4. At the time, my lab was one of the very few worldwide interested in understanding pluripotency. When we spoke about it, however, Matthias immediately saw the potential of what could be done, and how we could complement each other." Matthias smiles, "Hans' enthusiasm was infectious; it was easy to get excited about the project too. Hans' idea also coincided with my move to EMBL Hamburg and represented an ideal project for my new research portfolio."

At the time, EMBL had also just introduced the idea of two scientists sharing a research student. Hans recalls how Matthias rang him up to tell him that he had found a suitable candidate. "Matthias rang and said, 'Hans, there's a student here and he wants to work on Oct4!' Matthias always showed that extra dedication," adds Hans. The student, Attila Remenyi, now at the Eotvos Lorand University in Budapest, became the driving force of the project. "We gave Attila basic wet-lab training in Heidelberg," says Hans, "but he was also travelling regularly up to Hamburg to harvest the first crystals." When Hans moved to the University of Pennsylvania in 1999, Attila and Matthias visited him and his group regularly and the collaboration continued. There, Alexey Tomilin, who had moved with Hans to the States, also played an important

A previously unexplored linker sequence (red) between two DNA-binding elements of Oct4.



role in the project and worked closely with Attila. "Attila worked more on the crystallographic side of things," explains Hans, "while Alexey was interested in the cell developmental process." "Attila was a very pragmatic research scientist," recalls Matthias. "For reasons we did not understand, Oct4 did just not want to crystallise." Therefore, Attila took the decision to do all the initial work on the structurally very similar transcription factor Oct1, which was a lot easier to handle in the lab. The first batch of papers from the project were published in 2000-2004, but it has taken another ten years to make significant progress on the structure of Oct4 itself. The paper published in *Nature Cell Biology* in February 2013 describes important details about the structure of Oct4, in particular a previously unexplored linker sequence between two DNA-binding elements of the protein. "We have been very intrigued with this linker sequence for the last decade or so, so we are really pleased to have finally solved its structure," says Matthias. "This will really help us understand how it works in reprogramming cells to pluripotency." The 3D structure shows that the linker sequence is very exposed and conserved, and mutations have revealed just how vital the linker is for reprogramming. "The linker is so exposed, it just has to be used for binding – but the question is what does it actually bind to?" explains Hans. "We think it binds important factors which would then attract other factors to the Oct4 target genes which would then trigger the reprogramming process. Without this initial binding to the linker, nothing happens."

So what does the future hold? "We are really interested in in vivo programming," says Hans. "Once we understand how reprogramming of cells works, we want to work on actively moving cells back through their development in order to gain more stem cells. This would be crucial for counteracting ageing since the loss of stem cells is one important factor in the ageing process – obviously a very ambitious project but we will try." "My vision for the future is making sure that structural biology tools play an important role in complementary hybrid projects to capture the bigger picture," adds Matthias. "In Hamburg we are currently building the Centre of Structural Systems Biology, where although the focus will be on infection biology, other avenues of biomedical research would also fit into the concept of the centre, and we hope to be able to push structural biology techniques into new fields. And I would of course like to convince Hans to join us in this venture, especially in terms of exploring multicomponent complexes, unlocking the interactome and using this knowledge in technology-orientated approaches. How do things work and how can this knowledge be applied to screening and other technologies?"

Esch D *et al.* (2103) A unique Oct4 interface is crucial for reprogramming to pluripotency. *Nat. Cell Biol.* **15**: 295-301

DOI: 10.1038/ncb2680.

Context and variation

In a DNA version of 'spot-the-difference,' EMBL scientists and their colleagues in the 1000 Genomes Project compared the DNA profiles of 1092 healthy people from Europe, the Americas and East Asia, systematically tracking what makes us different from each other. Their findings, published in *Nature*, provide a powerful resource for people working in all areas of human health and biomedical research.

"The 1000 Genomes Project has achieved something truly exceptional," says Paul Flicek, group leader at EMBL-EBI and co-chair of the project's Data Management Group. "We have clarified which DNA sequences are common – and which are rare – in people from different areas and ethnic backgrounds. This baseline knowledge about human variation is vital for the ongoing search for genetic links to diseases."

Jan Korbel at EMBL Heidelberg co-leads the project's study of variation in large sections of chromosomes. He explains that the map of variation is extremely useful for understanding cause and effect in living systems. "When we find a very small change in a person's DNA sequence that is associated with a disease – something we call a SNP – we now have enough information to look at everything happening around it. We can see whether there is a larger change in that person's genome, also termed structural variation, that could be a contributing factor."

Which DNA sequences are common and which are rare?

"We've solved all the straightforward stuff – now it's time to tackle the really hard part," says Laura Clarke, who sits at the centre of the 1000 Genomes data team at EMBL-EBI. If you want to find something amongst the project's 500 000 files, Laura's the first person you'd ask. She fields questions from just about everyone: PhD students, undergraduates, small start-up pharma companies, postdocs – even project officers from other large-scale data collection activities who want help getting started.

"The study we published was done by aligning all the genome sequences and comparing them, so we could identify the most obvious differences," says Laura. "Now, the analysis team is tackling more complex questions, zooming in on very specific areas to come up with a more in-depth view of what is going on."

The final dataset includes the genomes of people in other regions of Asia, and offers many new avenues to explore. "We have moved on from collecting data to tackling the more complex questions about biological function," says Laura. "It's all very well to say that there is this or that variation, but now we have to establish what that really means. Our team looks for ways to make the data more accessible, so that people can have what they need to find answers."

"One of the best things about the 1000 Genomes community is that it is meritocratic," Laura reflects. "If anyone suggested a method and could prove that it was better than what we had, it would be taken on board. And that is a very nice way to do science. If we could share only one lesson we've learned with the next big projects, that would be it."

1000 Genomes Project Consortium (2012) An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**: 56-65

DOI: 10.1038/nature11632.

Biodiversity in the Balance

Causes and Consequences

Natural capital

The general principle for conserving biodiversity – the variety of life at all levels of biological organisation on Earth – has inarguable value. But there is much more debate about how best to do it. This year's Science and Society Conference brought scientists, journalists, policy makers, historians, philosophers, economists, clinicians and members of the public to EMBL Heidelberg to discuss one of the most formidable challenges facing humankind: the prevention of a disastrous mass extinction of species.

Robert May, professor of zoology at the University of Oxford, set the tone for 'Biodiversity in the Balance: Causes and Consequences' in his keynote lecture. "As human numbers and ecological footprints continue to grow, there are increasingly serious questions about maintaining the biological diversity upon which ecosystem services ultimately depend. We need to better understand how many distinct species inhabit the planet, current extinction rates, and, most importantly, the likely consequences of these extinctions," he warned.

The event, held in November 2012, was the 13th in an annual series of two-day conferences and provided a platform for more than 400 scientists and members of the public to come together to debate the impact of research on society and to engage in the development of a shared understanding of science. Jointly organised by EMBL and EMBO, the conference aimed to raise consciousness and facilitate dialogue on four distinct themes relating to the global biodiversity crisis.

Evolving web of life

Tragically, biodiversity loss is increasing at an unprecedented rate due to factors such as habitat destruction, land clearing, climate change, illegal wildlife trade, and the spread of invasive species. But while extinctions are possible, this outcome is by no means certain. One way to narrow down these uncertainties is to study the past: "We cannot run an experiment on the earth, so there is a benefit from studying ancient events," said Michael Benton, a palaeontologist at the University of Bristol. "Even if global scale problems cannot be resolved by a simple test, the interplay between phylogenetic and paleontological data offers rich potential for studying intrinsic and extrinsic factors."

For events going back millions of years, evidence is scarce, and mystery immense. But new techniques could have important implications for understanding puzzles from the past. "Only in the last five years have we begun to understand some aspects of the genetic basis of adaptations and the nature of genomic differences between species," said Axel Meyer, an evolutionary biologist at the University of Konstanz.

Rita Colwell, a renowned microbiologist at the University of Maryland, spoke about how interdisciplinary studies are shedding light on disease epidemics. Focussing on research into the

Rebecca Miller, Eric Karsenti, Colomban de Vargas and Ilkka Hanski at their panel discussion.

evolution of the *Vibrio cholerae* genome, she said: "It is important to understand the bacteria as part of the natural environment – it carries out critical functions in the carbon cycle and nitrogen fixation. Because of this, it is highly unlikely that it will ever be eradicated, but it can clearly be controlled."

On the other hand, Ilkka Hanski, a zoologist at the University of Helsinki, warned that we could already be in the midst of a sixth mass extinction that began with the expansion of human populations 50 000 years ago. "The number of surviving species is much smaller than predicted by the species-area relationship – habitats become fragmented and encroached upon by external forces, while many species have been driven beyond an 'extinction threshold'," he said. "A realistic approach to halting the loss of biodiversity in human-dominated landscapes is to set aside clusters of protected areas."

Biodiversity in decline

If ideas to tackle the decline are to work, one priority must be to increase our knowledge about what is actually happening in the natural world, participants heard. Detailing the International Union for the Conservation of Nature's (the IUCN's) Red List of Threatened Species, programme officer Rebecca Miller emphasised the importance of determining what is being lost, where, why and, crucially, how species and their habitats can be protected. "As most conservation action takes place on a local level, national red lists can be especially influential in the protection and recovery of threatened species," she said. "It is going to be a difficult journey, but we can get there. We need to make people realise that something can be done, to take the difficult decisions, and deliver the funding needed to address the problems."

One initiative that has sought to deliver answers is the Tara Oceans project, a two-and-a-half-year expedition to collect and analyse biological samples from the world's oceans. EMBL's Eric Karsenti and Colomban de Vargas, a biologist from the Roscoff Biological Station, presented some of the initial findings from the voyage. "For humankind, it is very embarrassing: there are very few people studying the critical role played by arguably the species-richest domain of life – protists," said Colomban. "Through our voyage, we aimed to change this, and raise awareness of the challenges the world's oceans face from human impacts," added Eric.

Jesse Ausubel, an environmental researcher at the Rockefeller University, reflected on the paradoxical challenge of developing technologies to bring together the immense flood of genetic and other information to help conserve species. "To understand the biodiversity question, we must zoom out as well as in," he said. "We need macroscopes."

Benefits and risks

Such approaches could help societies to understand more about the specific roles biodiversity has in enhancing lives and livelihoods, or as Shahid Naeem, Professor of Ecology at Columbia University put it: "why all these millions of species matter". Naeem called on societies to embrace what biodiversity has to offer in terms of regulating air, soil and water. "Rather than being the passive epiphenomenon of ecological and evolutionary processes, biodiversity plays an active role in governing ecosystem processes and the environments they generate," he said. "Our research shows that if you take rare species and look at them closely, they have a lot of properties in common with the ones we value: to lose them is like losing biological insurance."

One example discussed was human health. Tari Haahtela, who is Head of the Allergy Department at Helsinki University Hospital, presented a study that he said demonstrated strong connections between the diversity of microorganisms with that of plants and animals. "Our research suggests that contact with natural environments is a prerequisite to balanced immune function," he said.

"We need to make people realise that something can be done"

Heidi Wittmer, who leads the scientific coordination team for the UNEP-hosted Economics of Ecosystems and Biodiversity (TEEB) considered the striking economic benefits of biodiversity. "Despite the fact that the international convention of biological diversity has been active for two decades, political responses are still not adequate to address the problem," she said. "Our current economic setting systematically undervalues nature's contribution to human wellbeing – we can do better and need to act now."

Visions of sustainability

Exactly what that action should be has been the subject of intense debate. Harold Mooney, Professor of Environmental Biology at Stanford University, reflected on the recent development of an international agreement on biological diversity and ecosystem services. "In essence, it does for biological diversity what the Intergovernmental Panel on Climate Change has done for climate science and policy – and more," he said. "There will be periodic assessments of the current state of knowledge, but bottom up from regions to the international arena, with efforts to close gaps in expertise and promote capacity building. There is still a lot to do, but I am optimistic it could lead to change."

The final session sparked conversation that flowed out of the plenary hall and into the coffee sessions, particularly in regard to how potential benefits can be weighed against unexpected side effects. Volker Mosbrugger, a historian from the Senkenberg Society for Nature Research asked: "Might 'bio-geoengineering' offer a solution? The challenge is to develop a 'wise earth system management' that is done in full understanding and command of the consequences."

For Camille Parmesan, Professor of Integrative Biology at the University of Texas, getting approaches right will involve transferring observations of real-life responses of plants and animals to recent climate change to inform 'creative conservation'. "We have a lot of options: restoring and expanding natural systems, sequestering carbon, but what we are short on is labour, where citizen science could play a big role," she said. "Creative conservation solutions are not without risk, but successful conservation in a time of rapid environmental change will acknowledge that doing nothing carries risk as well. This hardly means we can ignore it."

ONLINE EXTRA

Watch some of the conference's talks at:
www.embl.de/training/events/2012/SNS12-01/Recorded_talks.

Shimmering crystals and bright beams

At the synchrotrons in Hamburg and Grenoble, scientists shine bright X-ray beams on crystals formed by large molecules – macromolecules – to determine their three-dimensional (3D) structure and understand how they carry out crucial roles in health and disease. Researchers and instrumentation developers from both EMBL outstations work together to provide top-notch technological services and expert advice to both in-house researchers and the scientific community at large, designing innovative tools and techniques to push the boundaries of such macromolecular crystallography (MX) studies. They have now developed a high-precision microdiffractometer (known as MD3), an automated crystal-harvesting technology (CrystalDirect™) and data management resources (CRIMS), which together promise to revolutionise how scientists use X-ray crystallography to decipher the 3D structure of biological macromolecules.

MD3 – TINY CRYSTALS IN FOCUS

Small and thin but very demanding: proteins that carry out important tasks in our bodies often form small crystals that cannot be analysed using conventional instrumentation. Studying them requires intense beams, high crystal stability and fast detectors. Florent Cipriani and his team from EMBL Grenoble and Gleb Bourenkov from Thomas Schneider's group at EMBL Hamburg have worked in close collaboration to overcome these hurdles and develop a third-generation, high-precision microdiffractometer (MD3), which is now associated with the new micro-focus beamline at the PETRA III synchrotron ring in Hamburg.

A microdiffractometer is a device that enables scientists to analyse crystals at synchrotron beamlines. The new MD3 microdiffractometer allows scientists to shine high-intensity micro-focused X-ray beams on crystals less than 1 mm in size – hundred times thinner than a human hair – and to rotate the crystals, which are attached to a spindle, while data is being collected. Florent and Gleb were able to attain this unprecedented degree of precision by using an air-bearing spindle oriented in an upright rather than the usual horizontal position, which helps to avoid the negative impact of gravity. The horizontal placing of the spindle is a historical legacy. In the 1980s, only large X-ray beams were available, limiting the minimum size of crystals that could be studied to a few tens of a micrometer. Since then, technology has advanced and scientists can now study much smaller crystals, or micro-crystals. At the same time, the brilliance of synchrotron beams has increased dramatically whereas beam sizes have constantly decreased – a reason why Florent and Gleb designed a microdiffractometer for analyzing so-called micro-crystals. In their new microdiffractometer, Florent and Gleb have also implemented a fast 4D scanning feature that allows the crystal to be inspected by the X-ray beam while it rotates, making data collection much faster and less laborious than ever. A typical example is the helical 4D scan used to collect data from needle-shaped crystals. Altogether, the EMBL scientists have been able to gain precision, stability and reduce the time required for processing micro-crystals using the new MD3 prototype installed at the integrated EMBL@PETRA3 facility in Hamburg.

This project was conceived back in 2007 and has reached its current form as a result of joint efforts between the facilities at Grenoble and Hamburg. As Gleb says, "The achievement was a combination of the best beams, which we are trying to create at our beamline, and the diffraction instrumentation, in which Grenoble has the best experts. These have been drawn together into something that is unique and better than the rest. It was a very close collaboration in many aspects, from all sites – and something of mutual enrichment for us and colleagues at Grenoble."

CRYSTALDIRECT™ – A CRYSTALLOGRAPHER'S DELIGHT

Bid adieu to the era when fishing out crystals for X-ray analysis was a labour-intensive and time-consuming process, and say hello to CrystalDirect™: an automated crystal-harvesting technology developed by Florent and José Márquez from EMBL Grenoble. Florent and José have developed special crystallisation plates covered with a thin layer of film, on which they can grow crystals directly. They then use a laser to recover the crystal along with the film, mount it on a robotically controlled pin and shine the X-ray beam on it. Although this new technique seems simple, it took Florent and José three years to optimise the film, crystallisation plates and the laser conditions. As Florent explains, "This project integrates knowledge from different fields; we travelled to many institutes across Europe to meet experts and to select and put together the right technologies."

The prototype version of this crystal harvester is now fully functional and can be used to collect and process hundreds of crystals with minimal user intervention – a big boon for projects such as the search for novel drug targets, in which researchers have to screen large numbers of crystals. The EMBL scientists aim to work with users to develop software to help control the whole process remotely. "We want users to be able to locate crystals in crystallisation plates remotely and analyse them with X-ray without delay. This system should also be more reproducible and reliable than the standard manual handling method," says José. He and Florent hope that this technology will one day become routine in all structural biology facilities around the globe.

CRIMS – COPING WITH DATA

Keeping track of experimental parameters and data generated from a million experiments annually is not an easy task. José and his team at the High Throughput Crystallization Facility in EMBL Grenoble have developed a computerised platform, called Crystallization Information and Management System (CRIMS), that tracks all of an experiment's details and makes the results available to users online, in real time. "We developed CRIMS to cover all aspects of crystallisation, from the time we receive a purified protein to displaying the results," José says. CRIMS provides an interface for researchers to design an experiment and control all aspects of it remotely. This makes data collection and management simple, allowing facility operators to focus on the experiments and enabling scientists to efficiently screen for the best crystals in their samples – and to easily track the conditions in which those crystals formed, in case they need to produce more.

CRIMS is a user-friendly software that can be operated and readily modified even by researchers who are not IT specialists, and has evolved into a real EMBL-wide project, with labs from Heidelberg, Hamburg and Grenoble jointly supporting the software's development. In addition, it is currently licensed to seven other laboratories in Europe. Working together with José, Rob Meijers from EMBL Hamburg has already managed to develop a similar interface for the experiments needed to prepare and analyse samples before crystallisation. Rob says, "Without CRIMS, the data management process is manual, or you use a commercial software that is not as well adapted as CRIMS."

More information on instrumentation development can be found here: www.embl.fr/instrumentation

The greatest show on earth

The plumage of a peacock, the iridescent eye spots on a butterfly's wings, and the social organisation of a termite colony are all wonders of the biological world. Yet perhaps the greatest biological display is performed without an audience, remaining hidden from us most of the time: the process of building a new organism through embryonic development.

Historically, biologists have tried to piece together the steps of embryonic growth — most notably, using the fruit fly *Drosophila* — by looking at development at successive points in time, and then piecing these snapshots together. Recently, however, the laboratory of Lars Hufnagel, group leader in the Cell Biology and Biophysics Unit at EMBL Heidelberg, has created a new kind of light microscope that allows biologists to film the process in unprecedented detail.

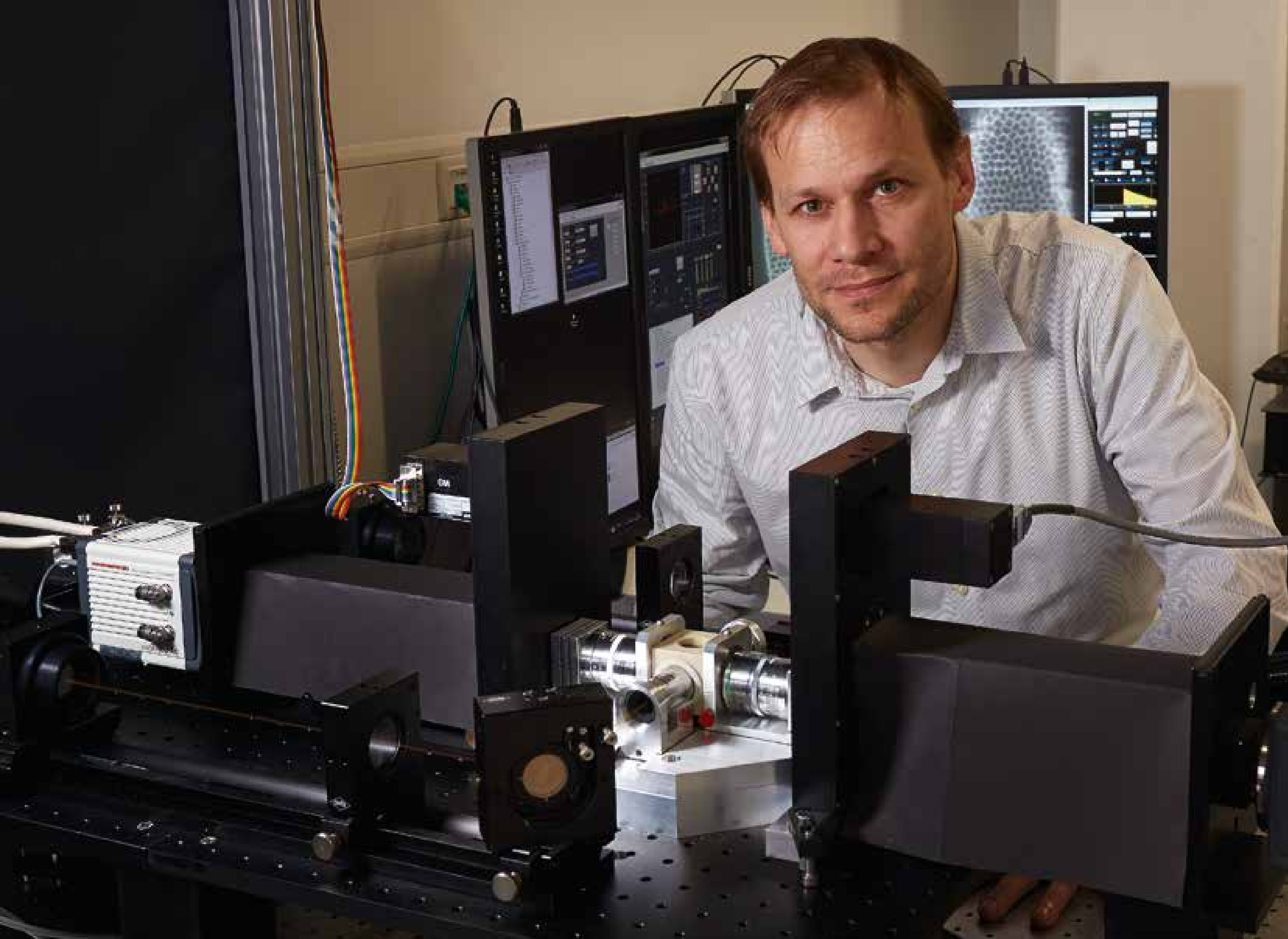
Their movie reveals the processes that sculpt the growing embryo.

The new microscope, called MultiView Selective-Plane Illumination Microscopy (MuVi-SPIM), builds on previous work with SPIM. In traditional light microscopes, biological specimens are illuminated with the equivalent of floodlights that spread light throughout the sample. By contrast, SPIM uses focused light beams that illuminate the sample only one slice at a time (for more on SPIM, see EMBL Annual Report 2007/08, page 70).

MuVi-SPIM takes this basic idea, but shines light from multiple directions at once, while a number of sophisticated cameras simultaneously capture two views of the sample over an extended period of time. The resulting images are then combined into a three-dimensional (3D) movie, which can be rotated around any axis so that the sample can be viewed from any angle.

A key advantage of MuVi-SPIM over other light microscopes is the speed with which images of 3D images of samples can be captured. With the widely used confocal microscope acquiring images of an entire fruit fly embryo can take around 10 minutes. This means that a lot of biological action can be missed between shots. "In 10 minutes, cells can divide and move, and even large-scale morphological changes can occur," says Lars.

MuVi-SPIM, however, captures the entire embryo within 3 seconds, giving the technique a much greater time resolution than confocal approaches. "MuVi-SPIM is 200 times faster



Lars Hufnagel

than confocal microscopy, and so we can even follow rapid biological process within the cell, while still being able to monitor the dynamics on the organismal scale," says Lars.

Lars and colleagues have applied MuVi-SPIM to create a detailed movie capturing the embryonic development of *Drosophila*. The movie begins with a 2.5-hour-old embryo, and follows it over the next 20 hours, at which point the larva is able to walk away. Their movie dramatically reveals the dynamic processes of cell movement that shape and sculpt the growing embryo.

"This has great potential as an educational resource," says Lars. "Lots of people I've spoken with in the *Drosophila* field are already using this movie to show others the hallmarks of *Drosophila* development." However, one obstacle to sharing such movies as widely as possible is the enormous amount of data generated by MuVi-SPIM. "We generated terabytes of data, which is hard to distribute over the internet," says Lars. "We're now working on compressing the raw data so that it can be viewed in a web browser."

ONLINE EXTRA

Watch a fruit fly develop on the EMBL YouTube Channel: http://youtu.be/MefTPoeVQ3w.

Krzic U, Gunther S, Saunders TE, Streichan SJ, Hufnagel L (2012) Multiview light-sheet microscope for rapid in toto imaging. *Nat. Methods* **9**: 730-3
DOI: 10.1038/nmeth.2064.

A record catch

Naturalists in the 19th century uncovered a whole new world of biology when they cast nets over the sides of their ships and studied the tiny sea creatures that they had caught under their microscopes. But in the 21st century, it seems you don't have to go so far afield to make new discoveries. Two EMBL Heidelberg teams, led by Matthias Hentze and Jeroen Krijgsveld, have cast the molecular equivalent of baited fishing lines into human cells and caught a haul of proteins that will shed new light on a hitherto under-appreciated aspect of cell biology. The findings could help yield a greater understanding of certain genetic diseases and cancer.

A haul of proteins will shed new light on under-appreciated aspect of cell biology.

The work relates to a cellular molecule called RNA, which the cell uses in a broad range of functions. One of these concerns how the cell turns the genetic information encoded in DNA into proteins. To do this, the cell makes an RNA copy of the DNA, which the cell's protein-making machinery then reads to assemble a particular protein.

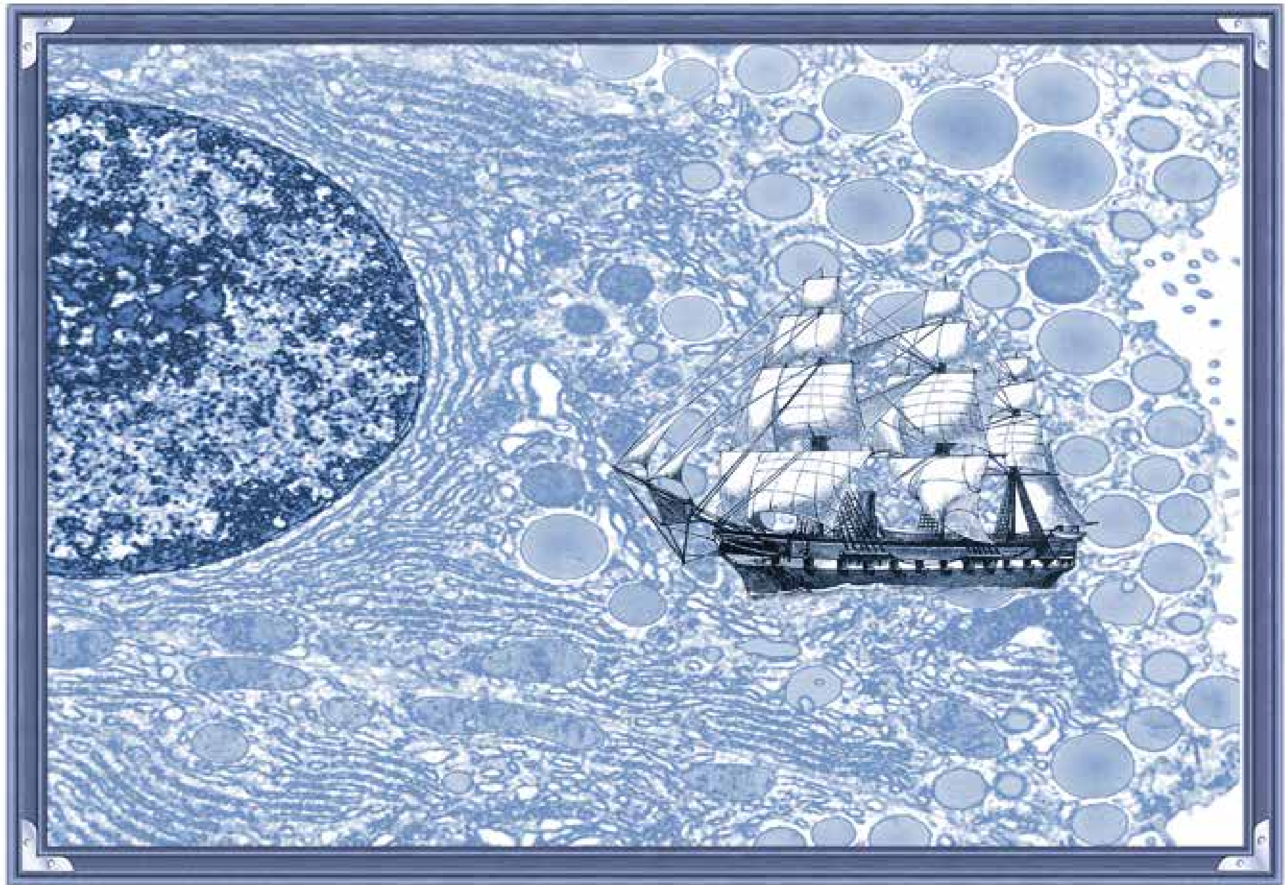
For a long time, this form of RNA, called messenger RNA or mRNA for short, was thought to be a mere carrier of information. In recent years, however, work from a number of labs, including Matthias', has revealed that mRNA plays a much more complex role in the cell, such as influencing how, when and where proteins are made.

Key to these functions are proteins called RNA-binding proteins, or RBPs. As their name suggests, these proteins stick to RNA, including mRNA, altering its behaviour. At the same time, they also interact with other proteins, allowing the cell to integrate RNA into its responses to changes in its physiology or environment.

Intriguingly, some cellular proteins, including enzymes involved in controlling metabolism, have been found to 'moonlight' as RBPs. What's more, several human diseases such as cancer and neurological conditions have been linked to faulty RBP activity. Until now, however, scientists only had a sketchy picture of the number of RBPs in cells, and what they might be doing.

Previous techniques used by researchers to identify RBPs have several drawbacks. Some mistakenly count interactions between proteins and other proteins as between proteins and RNA. Others cannot distinguish between an RNA and protein interaction that serves a genuine function within the cell and instances where an RNA and a protein just happen to have bumped into one another. What's more, many of these techniques study interactions between molecules in a test tube, rather than in living tissue.

So Alfredo Castello and his colleagues in Matthias's lab developed a new method to systematically fish for RBPs in live cells in a lab dish. Normally, proteins bind to and are then released from RNA all the time, but the team used a technique to change the biochemistry



Fishing for RNA-binding proteins

of the mRNAs to stop the proteins from being released. By shining ultraviolet light onto cells, the team triggered the formation of permanent chemical links between the mRNAs and the proteins, effectively turning the mRNAs into fishing lines studded with hooks to trap the proteins.

They then developed a new suite of methods that allowed them to identify the proteins that they had caught. Key to the success of the project was a collaboration with Jeroen and his team, who work a few metres down the corridor from Matthias's lab. Jeroen and his colleagues develop techniques that allow scientists to rapidly identify proteins and precisely measure their quantities in a sample. One challenge they faced in this project was to develop an approach that allowed the team to distinguish 'real' results from background noise in their large numbers of samples. Therefore, they teamed up with Bernd Fischer, a member of Wolfgang Huber's group at EMBL Heidelberg, who brought in the statistical expertise needed to refine the technique.

"We identified 860 RBPs, more than 300 of which were new and entirely unexpected," says Matthias. "We will now be able to learn about new ways how RNAs and proteins can interact. Many of these new RBPs are proteins that are known to be involved in genetic diseases and cancer, uncovering connections of RNA biology with these diseases." The next challenge is to verify and study the RBP 'parts list' identified in the study, and to apply the method to different cell types. The findings have generated quite a stir among biologists, and the team has published details of their method so that other researchers around the world can get involved, says Jeroen. "There has been a lot of interest from many areas."

Castello A et al. (2012) Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* **149**: 1393-406

DOI: 10.1016/j.cell.2012.04.031.

Castello A, Horos R, Strein C, Fischer B, Eichelbaum K, Steinmetz LM, Krijgsveld J, Hentze MW (2013) System-wide identification of RNA-binding proteins by interactome capture. *Nat. Protoc.* **8**: 491-500

DOI: 10.1038/nprot.2013.020.

Nick Goldman

Perfect storage

Holding a small test tube up to the light, Nick Goldman's face breaks into a broad smile as the realisation of what he is looking at begins to sink in. At first glance, the tube seems empty – you might think he is peering at a grain of dust, or a crystal of salt. But this unremarkable container could hold the beginnings of a wondrous story: proof that DNA could be used to store the world's ever-growing stores of digital data.

Encoded in the tiny specks Nick clasps between his fingers are all of Shakespeare's sonnets, an excerpt from Martin Luther King Jr.'s "I have a dream" speech, a PDF of Watson and Crick's paper announcing the structure of DNA, a colour photo of EMBL-EBI and a copy of the cipher used to encode and decode it all. At just under a megabyte, this is the largest amount of digital information that has been successfully stored using DNA molecules.

"Genomes have an impressive track record for complex data storage," explains Nick, a group leader at EMBL-EBI, who came up with the idea together with his colleague Ewan Birney. "They hold the genetic information that makes a cat, a zebra, a tree, and you. DNA also lasts a long time: we can extract it from the hair of woolly mammoths dating back tens of thousands of years and make sense of it – and those weren't even carefully prepared samples."

Indeed, if such samples had been more meticulously prepared we could still make sense of them 100 000 years later or more, Nick says. "Moreover, DNA is incredibly small, dense, and does not need any power for storage like your hard drive does, so shipping and keeping it is easy."

The study, published in *Nature* by Nick, Ewan and their colleagues at EMBL andAgilent, stands out from previous attempts to store data in DNA because it demonstrates an error-free, working method that scales up to global levels. The group's successful trial has raised the tantalising prospect that the huge quantities of data produced by film makers, particle physicists, biologists, climate researchers, governments and astronomers might one day be stored in a device no bigger than a tea cup.

Next-generation storage

Like many of the best ideas, Nick and Ewan's solution was formulated over a pint of beer. After a marathon meeting in Hamburg with their colleagues at EMBL about how to manage the flood of data being submitted to EMBL-EBI, they retreated to the hotel bar and started tackling the problem a bit more creatively.

"I can't really think without a pen in my hand so I'm sure we probably started drawing out ideas on some napkins," says Nick. Playful speculation quickly gave way to more spirited napkin scribbling and, of course, more beer.

"We sat down and Nick asked me, 'Well, look, DNA is a really efficient way of storing information. Is there something we can do?'" says Ewan. "As we wrote down our ideas, we realised that we could actually do all of the component parts of something that would, at least in principle, scale to something that might be valuable."

The process is not as straightforward as you might think: writing and reading DNA is prone to errors, particularly when the same DNA letter (A, G, T or C) is repeated. The 1s and 0s used by computers have previously been mapped directly onto DNA's four bases, but this approach creates long strings of repeating single bases, and these can be misread by DNA-sequencing machines. Imagine Shakespeare's sonnets without the romance, or Martin Luther King's speech without the dream.

Nick and Ewan therefore added a step, turning the binary code (0s and 1s) into a ternary code (0s, 1s and 2s). From there, rather than directly linking a number to one of four letters, each digit is represented by one of three letters – the rule being that it must not be a repeat of the letter preceding it. The long string of letters is broken up into smaller fragments, which overlap in a staggered fashion so that if there's a mistake in one fragment, it's covered by the next.

"We figured, let's break up the code into lots of overlapping fragments going in both directions, with indexing information showing where each fragment belongs in the overall code, and make a coding scheme that doesn't allow repeats," Ewan explains. "That way, you would have to have the same error on four different fragments for it to fail – and that would be very rare."

Putting the code into action

Having made what appeared to be an error-free codec, the group set about finding a partner to turn those As, Gs, Ts and Cs into guanine, adenine, thymine, and cytosine molecules. California-based Agilent Technologies agreed to synthesise the DNA, saving the project tens of thousands of dollars.

"We realised that, in principle, we could actually do it"

After downloading the encoded files, Agilent experts printed chemical reagents onto a glass slide with tiny reaction chambers, painstakingly adding the chemical solutions holding the nucleotides, spot by spot, to deliver a string of DNA. They produced thousands of copies of the 'files', freeze-dried them and sent the tubes to EMBL-EBI via FedEx.

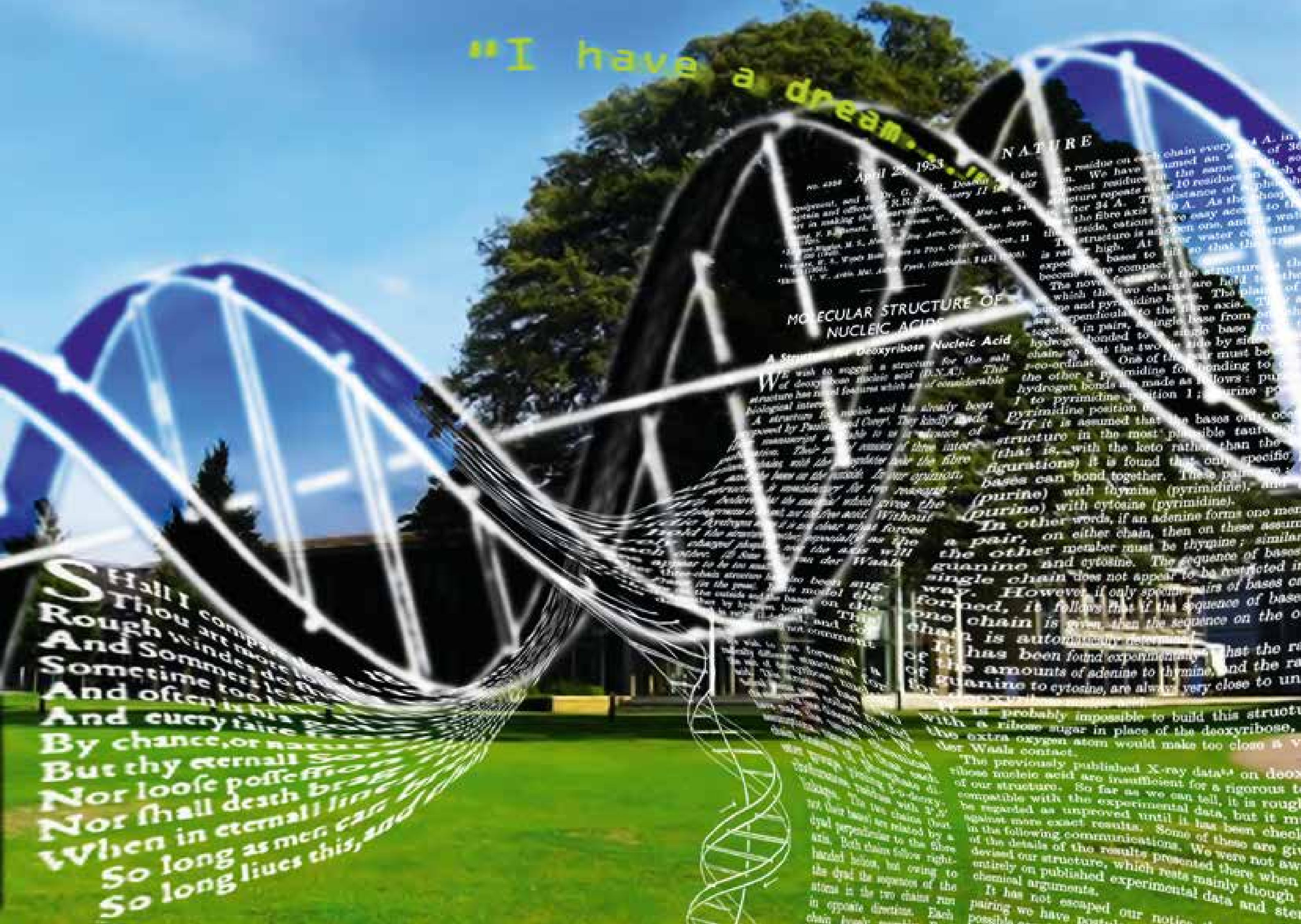
When they arrived, Nick at first refused to believe there was anything in the test tubes. "I'm a mathematician," he says proudly. "I had never seen this before – empty-looking test tubes that are packed with the things you need to run experiments."

Paul Bertone at EMBL-EBI cleaned up the samples and prepared them for sequencing, then took them to Vladimir Benes and his team at the Genomics Core Facility in Heidelberg. Vladimir's team sequenced them and Paul brought those files back in his pocket (on a hard drive) to EMBL-EBI. Nick uploaded the files to his laptop, ran the codec and, with the click of a mouse, revealed the original files in their full (and error-free) glory.

"We've created a code that's error-tolerant using a molecular form we know will last in the right conditions for many thousands of years, or possibly longer," says Nick. "As long as someone knows what the code is, you will be able to read it back if you have a machine that can read DNA."

Scaling up

Information is currently archived on, for example, magnetic tape, hard drives or film. All of these require carefully controlled environmental conditions and must be rewritten or replaced periodically. DNA is different: as long as it is stored somewhere cold, dry and dark, it should survive just fine for thousands of years. Using Nick and Ewan's method, one gram of DNA would hold 17.6 million gigabits of information – the equivalent of around half a million DVDs.



However, manufacturing (or 'writing') DNA and then sequencing it to read it back is expensive and time consuming – mainly because DNA synthesis is a fairly new technology. But as with DNA sequencing, prices are expected to fall rapidly over the coming years.

"The main use for this right now is for permanently storing high-value information that you want to keep safe, but if you really needed to go and get it you'd be prepared to wait a little while," Nick explains. "We estimate that in 10 years from now, DNA storage might be useful for information that is needed in 50 years' time – like keeping your wedding videos safe for your future grandchildren to see or national historical records."

Fast retrieval of digital information from DNA is still a way off, so you might need to wait a while before you can listen to the latest boy band on your DNA player. But for Ewan and Nick, the potential benefits are closer to home as they look for solutions for storing the ever-increasing amounts of biological data being produced and shared by scientists around the world.

"Science is becoming more and more data driven, and a considerable amount of that is concentrated in the life sciences," says Ewan. "We are getting to the point where we might consider losing information. There's been a lot of debate about what and how much we can afford to lose – and those are very serious questions for each scientific community. By thinking creatively about storage technologies, we can save more information to be used by future generations."

Goldman N, Bertone P, Chen S, Dessimoz C, LeProust EM, Sipos B, Birney E (2013) Towards practical, high-capacity, low-maintenance information storage in synthesized DNA. *Nature* **494**: 77-80
DOI: 10.1038/nature11875.

Complex loading

It's natural to think of DNA as a static code, like the printed words in a book. But for decades biologists have known about the existence of mobile genetic elements, or 'jumping genes', which are bits of DNA that can copy themselves and then paste this copy into a new location.

If this were to occur in a book – copying a passage from page 50 and pasting it in the middle of page 25 – the result would be gibberish. The same logic applies to our genome so when jumping genes land in the middle of another gene, they can render it meaningless. Cells have evolved ways to deal with this internal threat but when these safeguards fail, infertility can result, showing just how damaging marauding genetic elements can be.

The DNA of transposons is covered with chemical tags, like locking up a criminal.

Ramesh Pillai, a group leader at EMBL Grenoble, studies these jumping genes, also called transposons. Previous research by many groups has revealed a key strategy that cells use to police the genetic vandalism caused by transposons. After it has copied itself from the genomic DNA into an RNA molecule and is floating around the cell, specialised proteins latch on to the transposon and chop it up. These RNA fragments are then loaded onto a protein called Miwi2, which travels to the cell's nucleus and uses the loaded RNA to guide it to regions of the genome harbouring the offending transposon. Once found, the DNA of these transposons is covered with chemical tags called methyl groups, after which they can no longer be read or copied – like locking up a criminal so they can't commit more crimes.

However, the details of exactly how the cell achieves this are still being unravelled. In research published in *Molecular Cell* last year, Ramesh and colleagues filled in another piece of this puzzle: how transposon fragments are initially delivered and loaded onto Miwi2.

"We found a new protein called Fkbp6 that works with the protein Mili to facilitate the loading of Miwi2," says Ramesh. Fkbp6 belongs to a class of proteins called prolyl isomerases that can change the shape and structure of other proteins, so at first they thought that it must affect the structure of Miwi2 so that it was easier to load with RNA transposon fragments. "But to our surprise we found that Fkbp6 is inactive as an isomerase," says Ramesh.

Further examination of Fkbp6 led to the realisation that another part of the protein can recruit a so-called heat-shock protein (HSP), part of the family of 'chaperones' that help newly synthesised proteins fold properly. "So now it looks like Fkbp6 recruits HSP, which in turn alters the conformation of Miwi2 to facilitate the loading process," says Ramesh.

Ramesh Pillai on a visit to Heidelberg.

In addition to loading the RNA transposon fragments onto Miwi2, HSP also plays a role in getting rid of the by-products of this process, which also depends on Mili. After Miwi2 has been loaded with RNA, a small fragment remains that blocks the ability of Mili to pick up subsequent pieces of RNA. "The Fkbp6–HSP complex allows Mili to release these fragments and continue to function," says Ramesh.

Now the big challenge is to work out how loading Miwi2 with transposon RNA causes it to travel to the nucleus, find its target in the genome, and silence it through methylation.

ONLINE EXTRA

Listen to Ramesh and Orsolya Barabas talk about transposons on the EMBL Explore Podcast: www.embl.org/explore.

Xiol J, Cora E, Koglgruber R, Chuma S, Subramanian S, Hosokawa M, Reuter M, Yang Z, Berninger P, Palencia A, Benes V, Penninger J, Sachidanandam R & Pillai RS (2012) A Role for Fkbp6 and the Chaperone Machinery in piRNA Amplification and Transposon Silencing. *Mol Cell*. **47**:970-9.

DOI: 10.1016/j.molcel.2012.07.019.

Cataloguing the genes involved in the cell's exports.

Export economy

Just as a country's economy has to maintain a balance between imports and exports, so too do cells. Material is ingested, or recycled from the cell membrane, and delivered to where it's needed or to where it will be broken down. At the same time, newly manufactured molecules, especially proteins, are shipped from their production sites to other cellular locations, or secreted at the cell's membrane.

Biologists focussing on a small number of players in these transport networks have provided important insights into these processes, but this approach does not produce a global view of how these transport or secretory pathways are regulated. So Rainer Pepperkok, team leader and Head of the Advanced Light Microscopy Core Facility, and Jan Ellenberg, Head of the Cell Biology and Biophysics Unit, both at EMBL Heidelberg, set out to remedy this with a large-scale analysis of the entire protein secretory pathway in human cells.

The team used a genetic technique called RNA interference (RNAi) to shut down each of our 22 000 genes, one by one, to see the effects this had on the secretory pathway. This was achieved with a high-throughput RNAi screening platform developed by Rainer and Jan that uses microarrays consisting of hundreds of tiny spots, each of which contains human cells with a given gene silenced. Crucially, this screening platform was adapted so that the cells in each spot can be viewed with a microscope and the images automatically analysed by computer software to identify changes in protein transport.

"The amount of detailed information we've obtained goes far beyond what's been done before," says Rainer. This systems-level analysis revealed that 15% of all human genes have some effect on the cell's ability to transport proteins. Their global view of protein transport pathways also identified distinct networks that control and regulate protein transport in previously unappreciated ways.

"We have found connections between processes that are usually studied by different camps of biologists," says Rainer. One such link is between cellular signalling and the transport of newly synthesised proteins through the secretory pathway. Signalling molecules bind to receptors at the cell surface, which triggers a cascade of events within the cell that eventually leads to a change in the cell's behaviour; often, these receptors are engulfed into the cell, and targeted for either degradation or recycling. Rainer's team found that this process and the secretory pathway come together through a molecular network related to the epidermal growth factor receptor (EGFR). It's well known that activation of EGFR leads to cell proliferation, after which EGFR is internalised and degraded by the cell. However, Rainer's lab also showed that stimulation by EGF activates early events in the protein secretory pathway as well, as newly synthesised EGFR needs to be transported back to the cell surface.

In addition to these findings, the RNAi screen identified several hundreds of other genes involved in the regulation of the secretory pathway. Surprisingly, around 80 of these are implicated in regulating gene expression, and so may play a role in coordinating the creation and export of proteins to replace those that have been degraded within the cell – an issue now being explored in Rainer's lab. "Our results are giving us a completely new perspective on how the secretory pathway is regulated and allow us to focus on the most interesting aspects in our future work," says Rainer.

Simpson JC *et al.* (2012) Genome-wide RNAi screening identifies human proteins with a regulatory function in the early secretory pathway. *Nature Cell Biology* **14**:764-74.

DOI: 10.1038/ncb2510.

Picky partner

"It takes two to tango" so the saying goes, meaning that for any partnership to succeed, both parties must work together and ideally complement each other. Matthias Wilmanns, Head of EMBL Hamburg, and Eiríkur Steingrímsson from the University of Iceland, Reykjavik and also the EMBL Council delegate for Iceland, have proved that this expression can hold true for molecules as well as people.

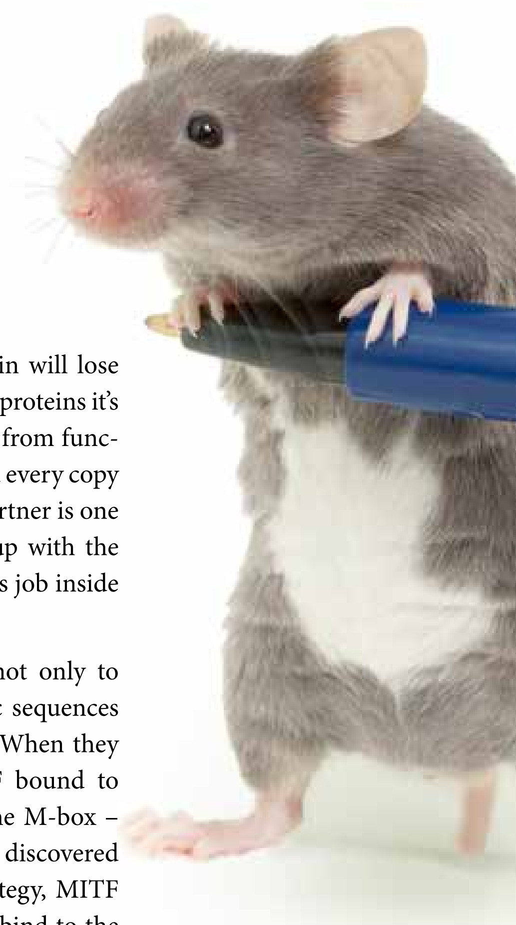
At first glance, the subject of Eiríkur's research seems far removed from that of Matthias. Whereas the Hamburg group determine the 3D structure of molecules, the Icelandic lab study mice with unusual patterns on their skin. The mice carry faulty versions of a protein called microphthalmia-associated transcription factor (MITF), so named because it has been linked to microphthalmia – a condition in which one or both of a person's eyes is smaller than the eye socket and usually results in blindness. MITF is also known to be involved in melanoma and other skin disorders, probably because it plays a key role in controlling genes that drive the development of melanocytes, the skin's pigment-producing cells.

At a Council meeting in 2006, Eiríkur approached Matthias and proposed they team up to see whether the structure of MITF could help to explain its role in these diseases and, more specifically, why changes to different parts of the protein cause varying degrees of pigmentation disorder in both mice and humans. The key, they discovered, is that MITF is picky about its partners. The protein has a so-called zipper structure that allows it to pair up with other molecules. But Vivian Pogenberg in Matthias' group found that, unlike similar proteins, MITF's zipper has a kink in it. "This kink limits MITF's binding specificity," Matthias explains. So if a mutation destroys that kink, the protein will lose its specificity. It could then bind to proteins it's not supposed to, preventing them from functioning correctly, for instance. And every copy of MITF that binds to a 'wrong' partner is one less copy that's available to pair up with the 'right' molecule so that it can do its job inside the cell.

To fulfill that job, MITF binds not only to other proteins but also to specific sequences of DNA to control certain genes. When they looked at the structure of MITF bound to one of those DNA sequences – the M-box – Vivian, Matthias and colleagues discovered another quirk. In an unusual strategy, MITF uses the amino acid isoleucine to bind to the M-box. And just as eliminating the zipper kink enables MITF to bind to a greater variety of proteins, replacing its isoleucine with another amino acid, asparagine, allows MITF to bind to other DNA sequences. But interestingly, a mouse can overcome this by taking advantage of the fact that it carries two copies of each gene, one inherited from its father and the other from its mother. For example, if one copy of the MITF gene has the scope-expanding asparagine-for-isoleucine swap, and the other copy has a mutation in a different part of its sequence that would normally stop MITF from binding to DNA at all, the mouse's cells can somehow produce fully functional MITF proteins. "I find this neat from a molecular point of view," Matthias enthuses. "It's a fascinating piece of genetics."

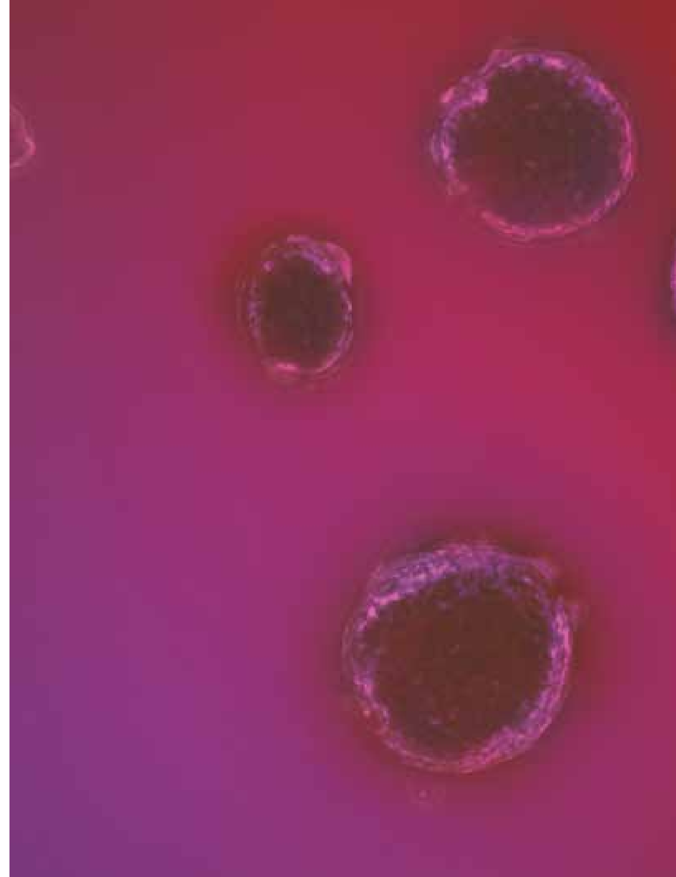
Pogenberg V, Ogmundsdóttir M, Bergsteinsdóttir K, Schepsky A, Phung B, Deineko V, Milewski M, Steingrímsson E, Wilmanns M (2012) Restricted leucine zipper dimerization and specificity of DNA recognition of the melanocyte master regulator MITF. *Genes Dev.* **26**: 2647–58

DOI: 10.1101/gad.198192.112.



Mice with a faulty version of MITF have a striking white belly.

Back to normal



Paul Bertone, a group leader at EMBL-EBI, patiently tries to explain why some people might not be happy about the findings of his latest research on cancer genomics.

"Cancer is primarily a disease of the genome," he says. "Cancer genomes are deeply flawed and very powerful. They drive the production of cells that proliferate out of control, migrate to where they don't belong and activate or suppress inappropriate pathways." But it's not just mistakes in the sequence of a cell's DNA that determines its behaviour – the activation or repression of key genes is influenced by the sum of all the biochemical modifications to the DNA, the so-called epigenome.

As cells age and divide, some epigenetic factors, or marks, are not reproduced faithfully. These errors can accumulate over time and have been associated with the development of cancer.

"We dropped a bomb on the whole epigenome"

"If you could revert some of these epigenetic marks back to a normal state, you might be able to control the expression of cancer-related genes and make the cells less malignant," reasons Paul.

It's a promising idea, but there was no system in place to test it out. Working with fellow developmental biologist Steve Pollard at University College London, Paul and his colleagues decided not to test the epigenetic factors one by one, but to reconfigure the entire epigenome by reprogramming the cells.

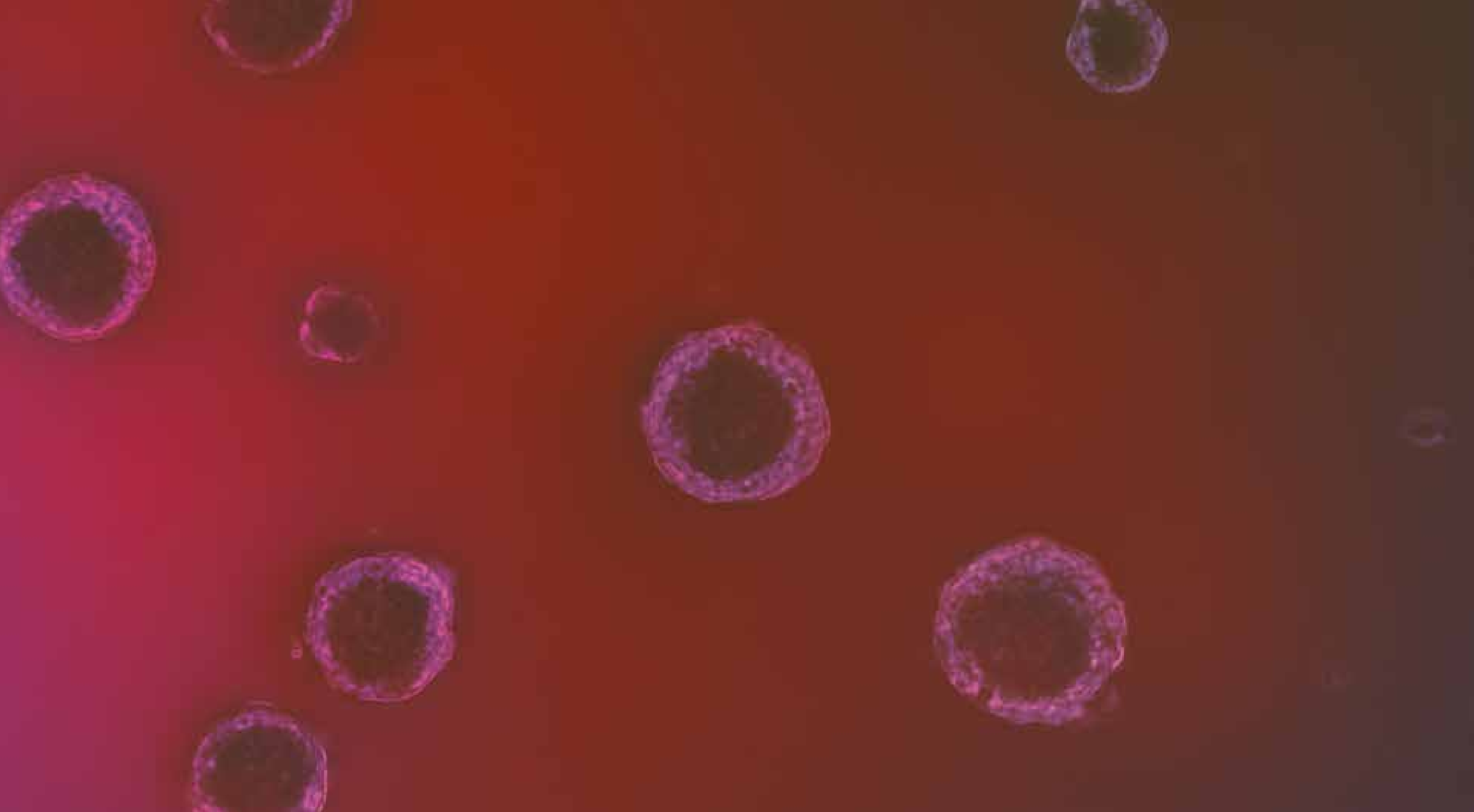
"The developmental state is a reflection of the epigenome," Paul explains. As development progresses, cells become more specialised, and each cell type has specific epigenetic marks that stabilise these changes.

"We wanted to see if completely resetting epigenetic marks had an effect on the behaviour of cancerous neural stem cells – specifically, glioblastoma stem cells, which are very aggressive." The idea was to reprogramme them into induced-pluripotent stem (iPS) cells, and in so doing erase most of the epigenome.

To replace the defective epigenetic marks, the team reverted the cancer cells back into an embryonic stem-cell state. Then they brought them forward again to the neural stem-cell state. Steve compares the process to rebooting a computer to remove any accumulated memory errors: the hardware (the cancerous genome) doesn't change, but the additional defects (the epigenetic flaws) are wiped clean.

"We effectively erased the marks that had been established during development, as well as potentially defective changes that had accumulated over the lifetime of the patient – you might say we dropped a bomb on the whole epigenome," Steve says. "These changes are much, much more extensive than what is being proposed for targeted epigenetic therapies, so if this approach doesn't work it's not likely that fixing individual regions would make much of a difference."

The researchers measured changes to the transcriptome and epigenome of the cells at each stage of the process, finally transplanting the



Cancer cells were reverted to a stem-cell state.

re-differentiated neural cells into brain tissue. They waited to see whether the cells would still develop tumours and, if so, whether these would be less malignant than those produced from the original cancer stem cells.

"It doesn't work – not in a way that would be realistic in a patient," says Paul. "The cells with restored epigenetic marks are just as lethal as those from the original brain tumours."

In a previous study the group had identified a number of genes specific to this cancer, and they expected that when the epigenetic status of thousands of genes had been restored they would also see some change in malignancy. But they didn't. So they went a step further.

"We considered whether more dramatic epigenetic changes might produce an effect. Instead of re-differentiating the cancer iPS cells back to their original type, we steered them towards an entirely different lineage."

The reprogrammed cells were manipulated to become a different type of cell, then transplanted to see whether there was any change. This time, the cells made benign tumours.

"The chromosomes in a cancer cell are structurally damaged, so at a very profound level they are always going to have problems. In our experiment, the cells still proliferate out of control but changing them into another type of cell made the tumour less lethal. Because they were no longer brain cells, they were unable to invade the surrounding brain tissue."

The group discovered that it is indeed possible to affect cell behaviour using epigenetic reprogramming, by making cancer cells incompatible with their environment.

"Experiments like these are important, but we were working in unchartered territory and it was impossible to secure external funding," Paul reflects. "Each of the labs involved put forward the necessary resources to complete the project. If I didn't have the freedom to conduct independent research, we never would have been able to pursue this line of inquiry. Places like EMBL are vital to keeping the scientific process moving forward, especially when new approaches might challenge popular ideas."

Stricker SH *et al*. (2013) Widespread resetting of DNA methylation in glioblastoma-initiating cells suppresses malignant cellular behavior in a lineage-dependent manner. *Genes Dev*. **27**: 654-69

DOI: 10.1101/gad.212662.112.

Engström PG, Tommei D, Stricker SH, Ender C, Pollard SM, Bertone P (2012) Digital transcriptome profiling of normal and glioblastoma-derived neural stem cells identifies genes associated with patient survival. *Genome Med*. **4**: 76

DOI: 10.1186/gm377

A year in the life of EMBL

April

Girls' Day

In the past few years, EMBL Heidelberg has celebrated Girls' Day with great enthusiasm. This year was no exception: on 26 April, the main Laboratory welcomed 15 young women (and men) to spend the day with professionals of their choice from either scientific Units or administrative departments. This event exposes schoolgirls to career paths that are traditionally considered 'male-dominated' such as science, engineering, mathematics and computing. The pupils had fun shadowing people whom they one day aspire to be and got a taste of life in EMBL that ranged from imaging at the microscope to managing the sumptuous menu at the cafeteria.

2012

Buzzing HUB

On 26 April, more than 40 bioinformaticians from EMBL Heidelberg and other institutes around the city convened at the University of Heidelberg for the first HUB (Heidelberg Unseminars in Bioinformatics) event titled 'Challenges in Systems Biology'. HUB takes the standard seminar format and turns it inside out, enabling active audience participation, exchange of expertise and open discussion and debate on the latest trends and challenges facing the fields of bioinformatics and computational biology. The meeting began with an introduction by Thomas Lemberger, Chief Editor of *Molecular Systems Biology* and was followed by moderated discussions and flash talks by scientists. Topics at other HUB events this year have included 'Funding Bioinformatics' and 'Biggest Challenges in Bioinformatics', and the latter enabled participants to jointly contribute to a paper published in *EMBO Reports.*

May

Great initiative

What happens when leading scientists from EMBL and EMBO meet with policy makers from the European Parliament, European Commission and EMBL's member states? Scientists get the golden opportunity to participate in the decision-making process on an issue that concerns them directly – funding for research. This was the agenda of the conference organised by The Initiative for Science in Europe (ISE), on 3-4 May in Barcelona. ISE is an independent platform of learned societies and scientific organisations, including EMBL, which supports all fields of science and bridges the gap between individual researchers, the funding system and political leaders.

Immortality – truth or fiction

Stephen Cave, a well-known British author and philosopher, delivered an EMBL Forum lecture on 10 May as part of EMBL's Science and Society Programme. In his talk, Stephen focused on one of mankind's greatest desires – the quest to live forever. He explained that often rhetoric in modern biology promises cures for all diseases and much more; however, such far-reaching ideas should be treated with caution. Striving for immortality is not only a false promise, but a dangerous one, he warned. "Ultimately we are not going to achieve immortality – and that is for the best," he said.

Ready, set, GOLD

It was not bioinformatics this time, but the recipe for making the strongest athlete that grabbed the attention of participants at the conference organised by staff at EMBL-EBI on 10 May. At the event, entitled 'Going for Gold: the science behind sporting success', Alun Williams from the University of Manchester and Robert Gray from the University of Birmingham explained that the right proportion of genetic, physiological, psychological and nutritional factors makes the best sportsmen. EMBL-EBI's Paul Flicek presided over the discussions at the event where David Fletcher from Loughborough University also shared a personal story of near Olympic success.

June

Digging for data

EMBL-EBI's Director, Janet Thornton, got behind the wheel of a digger to officially kick off the construction of a new building at the Genome Campus on 13 June. Staff, invited speakers and guests from the UK's Biotechnology and Biological Sciences Research Council (BBSRC), Capital Works and Oxford Archeology East, witnessed the 'ground-breaking' event. The new facility at EMBL-EBI has been generously funded with a grant of £75m by the UK government with the aim of expanding EMBL-EBI and supporting ELIXIR, the nascent pan-European research infrastructure for life-science data.

2012

Celebrating brilliance

June was a month of celebrations for EMBL scientists at both Hamburg and Grenoble outstations who strive to provide the brightest beams for groundbreaking structural biology studies. The BioSAXS beamline reopened at the European Synchrotron Radiation Facility (ESRF) in Grenoble on 13 June with improved features for users such as sharper focus and shorter data collection times. Whereas in Hamburg, the P12 BioSAXS beamline was commissioned on 15 June with the first remote measurement, implementation of a new chromatography system and development of a novel microfluidic system.

July

The Nobel meeting

EMBL Heidelberg predocs Sigrid Milles and Felix Klein joined hundreds of young researchers and 27 Nobel laureates at the prestigious Lindau Nobel Laureate meeting held between 1-6 July against the backdrop of Lake Constance. This year's meeting focused on physics and featured a series of lectures, panel discussions and social events. Participants had fun discussing diverse topics with the Nobel laureates including quasicrystals, energy security and funding challenges. The meeting concluded with a boat trip to the flower island of Mainau and participants headed home with the spirit and inspiration of Lindau.

Favourite day of the year: Lab Day!

Science and lab life were celebrated with great splendour on 5 July as more than 200 staff members from all EMBL sites convened at the main Laboratory on the occasion of Lab Day. This year featured the most extensive programme of scientific talks to date, presented by EMBL PhD students and postdocs. For the first time, participants were invited to submit scientific posters alongside the now traditional fun representations of their labs. The highlights of the event were the graduation ceremony and presentation of the prestigious John Kendrew Award to two alumni this year: Simone Weyand and Gáspár Jékely were rewarded for their significant contributions to structural and developmental biology, respectively. In the evening, participants enjoyed music from the EMBL choir and the newly formed percussion band Macumba.

EMBL shines through at ESOF 2012

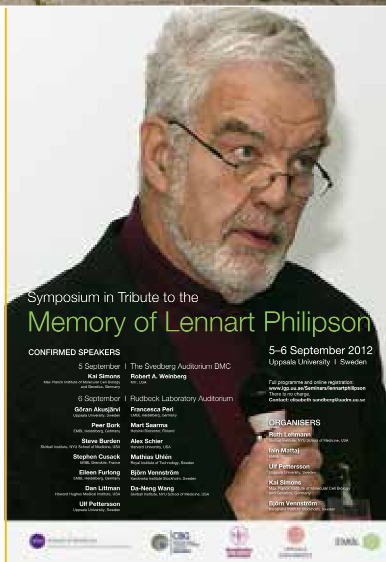
Europe's largest general science event held in July in Dublin had a strong EMBL presence alongside more than 4500 other participants including leading scientists, funders, journalists, policy makers and members of the public. Lars Steinmetz and Eric Karsenti from EMBL Heidelberg delivered keynote lectures at the event. Halldor Stefansson, head of EMBL's Science and Society Programme, participated in a panel discussion whereas EMBL-EBI Director Janet Thornton took part in a podium discussion focused on the importance of research infrastructures. Lena Raditsch, Head of Communications at EMBL, also gave a presentation addressing the question "Can outreach make you a better scientist?" ESOF is a pan-European meeting aimed at showcasing the latest advances in science and technology and stimulating public interest in research activities.

August

Undergraduate reception

Helke Hillebrand and the graduate office at EMBL welcomed 16 undergraduates on 14 August at the main Laboratory. Helke introduced them to all the scientific and non-scientific training offered as part of the International PhD programme at EMBL. The event also included visits to several labs to get a flavour of laboratory life and the varied topics on which the scientists work. The students were highly impressed with what EMBL had to offer and, having successfully made it through the PhD selection process, one of them will soon be joining EMBL as a predoc!

2012



September

In reverence and respect

A conference titled "The molecules of life" was held at Uppsala University, Sweden on 5-6 September in memory of the second Director General of EMBL, Lennart Philipson, who served the institute between 1982-1993. EMBL staff and alumni gave scientific presentations covering a range of topics from cancer research to work on the human gut microbiome, reflecting Lennart's own diverse research interests. Lennart's great organisational skills and clear vision for the advancement of science truly made him a respected leader and a revered scientist.

Faculty retreat

12-13 September saw the coming together of faculty members from all EMBL sites at Stromberg, Germany to discuss science, connect with colleagues, and establish new collaborations. Highlights of the event included talks from scientists, the presentation of new EMBL-EBI service clusters, the announcement of the fully operable PETRA III beamlines in Hamburg, and an introduction to new EMBL thematic centres. Speaking at the retreat, Iain Mattaj, Director General of EMBL, emphasised the importance of support activity at EMBL, including networking with member states, industry relations and outreach activities that form a key part of the success of EMBL as a research institute, in addition to the exceptional science performed in its laboratories.

2012

Sound of science in the city of music

'Science and networking' was the motto of this year's PhD students' retreat in the beautiful city of Salzburg, Austria on 21-23 September. Participants presented their work in a relaxed peer-to-peer setting and listened to career-oriented lectures from external speakers – all EMBL alumni, who have made their mark in careers beyond academia. Besides scientific sessions, the attendees explored the stunning sights of Salzburg and took part in local cultural activity.

October

A true Dame

EMBL-EBI Director, Janet Thornton was awarded a most prestigious honour – Dame Commander of the Order of the British Empire, by Prince Charles on 12 October in a ceremony at Windsor Castle. The title was awarded in recognition of her outstanding contribution to science and for efforts to promote life sciences in Europe. The medal was displayed at EBI for weeks after the ceremony as she insisted that the award was for all at EBI to cherish and treasure.

A noble talk

Crowds poured into the Advanced Training Centre at EMBL Heidelberg on 8 October for the keynote lecture delivered by John Gurdon at the EMBL/EMBO symposium 'Germline – Immortality through Totipotency'. Just days earlier, Gurdon had received the Nobel Prize for Physiology or Medicine together with Shinya Yamanaka for 'the discovery that mature cells can be reprogrammed to become pluripotent'.

14th International PhD Symposium

The 14th International PhD symposium 'Networks in Life Sciences', organised by EMBL PhD students from all sites, was held at EMBL Heidelberg from 25-27 October. Participants included PhD students from several European and non-European universities with training in different scientific fields, from biology to mathematics. Well-known scientists gave talks on topics covering areas of genomics, proteomics and systems biology. The event also included poster presentations and moderated panel discussions. Prizes were given for the best 'poster' and 'student talk' and the meeting concluded with the famous EMBL party.

2012

Flu at the café

Fifty people gathered downtown in Grenoble on 10 October for the first Café Scientifique event organised jointly by EMBL and Café Sciences et Citoyens de l'Agglomération Grenobloise, as part of the city's science festival. The audience engaged in a lively discussion with medical and research professionals working with the influenza virus and debated almost everything concerning the flu, from vaccination prospects to public health policies.

November

EMBL and Argentina

EMBL Director General Iain Mattaj signed a memorandum of understanding with Lino Barañao, Minister of Science, Technology and Productive Innovation for the Argentine Republic on 7 November. The aim of the agreement is to encourage cooperation and scientific collaboration between Argentina and EMBL. During his visit, Iain also attended a symposium in Argentina organised by EMBL alumni.

Making an impression

"What is molecular biology?" wondered 15 journalists visiting EMBL Heidelberg for a one-day tour in November. Matthias Hentze promptly explained the basics of the science and so impressed was the audience that EMBL's research featured in several radio interviews and broadsheet articles in the weeks following the event. The visit, jointly organised by Germany's National Academy of Sciences, Leopoldina, and EMBL's communication team, aimed to introduce molecular biology to journalists with no scientific background.

December

2012

Graduation ceremony

The graduation ceremony for EMBL's 23 latest doctorates was held at the Klaus Tschira Auditorium in the Advanced Training Centre at EMBL Heidelberg on 13 December. Helke Hillebrand, Dean of Graduate Studies, congratulated the new doctorates and welcomed their mentors and proud friends and families to the ceremony. The event included light-hearted talks by Nenad Bartonicek and Tim Wiegels who gave a fun overview of EMBL PhD life, and celebrations concluded with a wine and cheese session.

January

2013

High praise

On 9 January, Anne Glover, Chief Scientific Advisor to the European Commission, was welcomed to Heidelberg by EMBL's Director General, Iain Mattaj and Director of International Relations, Silke Schumacher. During her visit, she interacted with EMBL group leaders, PhD students and postdocs as well as representatives from EMBO. Anne, a molecular biologist by training, discussed the challenges of her job and explained that strategic communication is key to addressing issues related to policy making and funding. She was so impressed with the quality of science performed at EMBL that she tweeted that evening: "It is molecular biology at its very best."

Bilateral benefits

Scientists and instrumentation groups from the structural biology facilities at EMBL Hamburg and Grenoble gathered in Hamburg from 14-16 January for the bilateral meeting to work together and discuss new projects. The meeting started off with a workshop on CRIMS (Crystallization Information Management System). Participants also attended talks on recent developments and challenges at the BioSAXS and MX beamlines, and toured the EMBL beamlines at PETRA III.

Corporate ties

The 4th EMBL Corporate Partnership Programme annual gala event was held at the EMBL Advanced Training Centre on 26 January. The theme of this year's event was personalised medicine. EMBL-EBI Director Janet Thornton addressed the participants on the importance of bioinformatics in the life sciences. This was followed by talks from group leaders Paul Flicek (EMBL-EBI) and Peer Bork (EMBL Heidelberg). Through the Corporate Partnership Programme, companies such as GE Healthcare, Leica Microsystems, Life Technologies and Olympus support EMBL's mission to provide excellent training to scientists the world over.

February

Burns night

It was all about great food, high spirits and dancing at the annual Burns night celebrations at EMBL Heidelberg on 2 February. The culinary start to the evening involved malt whisky tasting and digging into haggis and trifle. This was followed by live music by the Heidelberg and District Pipes and Drums band and traditional Scottish country dancing into the wee hours of the morning.

2013

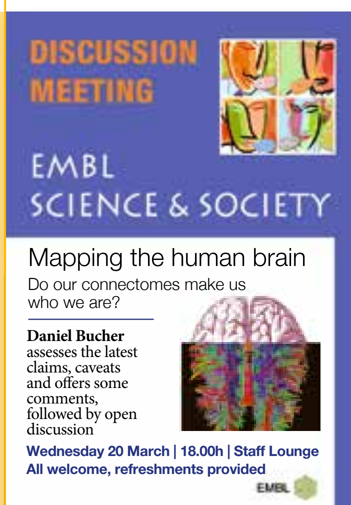
Art in EMBL

Works of art inspired by the research conducted at EMBL and created by 22 students aged between 16 and 18 from Paul-von-Denis Gymnasium, Schifferstadt, Germany, were displayed in the EMBL Advanced Training Centre foyer from 21 February - 6 March. The students drew inspiration from their visit to the main laboratory in the summer of 2012. Around 80 pieces were on display, ranging from paintings and photographs to ceramics and 3D sculptures. The group ART@EMBL ATC, which aims to connect staff and visitors with artwork to inspire and enjoy, supported the exhibition.

March

Crucial connections?

On 20 March, EMBL Heidelberg hosted the first Science and Society Discussion Meeting. This inaugural session focused on connectomics – the modern approach of studying the connections between neurons in the brain. Daniel Bucher from Detlev Arendt's group started off the meeting with a short presentation assessing the latest claims and caveats of the field. Later, participants debated whether understanding the brain's circuitry would one day help in 'reading the human mind' and cure diseases or whether it is all a false promise.



April

Open Day

EMBL Heidelberg opened its doors to the public on the 27 April, attracting more than 1500 guests. The event was made possible by over 180 employees including scientists, technicians, caretakers, canteen and administrative staff, who combined their efforts to offer visitors a variety of activities. The guided tours of 12 different labs were all fully booked, with extra tours added and also 'sold out'. Guests big and small got the chance to do their own experiments in the 'Research Stations' and the 'Kids' lab', and attend talks ranging from cell division to evolution. And for a break with a view, participants could talk to EMBL scientists about hot topics such as the use of animals in research or the ethics of genome research, over coffee and cake, at the 'science café'. The day was a resounding success, enjoyed by guests and organisers alike, in true EMBL spirit.

Index of group & team leaders

A

Antony, Claude 13, 34
Aulehla, Alexander............................... 42

B

Barabas, Orsolya 14, 63
Beck, Martin 38
Benes, Vladimir..............................13, 61
Berger, Imre...12
Bertone, Paul.................................61, 66
Birney, Ewan....................................2, 58
Bork, Peer 18, 40
Briggs, John 10, 17, 27

C

Cipriani, Florent................................... 52
Cusack, Stephen................................ 28

E

Ellenberg, Jan..................................... 64
Ephrussi, Anne 24

F

Flicek, Paul 47

G

Goldman, Nick.....................................58

H

Hart, Darren.. 30
Hentze, Matthias................................. 56
Heppenstall, Paul................................... 8
Huber, Wolfgang................................. 56
Hufnagel, Lars 54

K

Kaksonen, Marko.................................17
Karsenti, Eric 50
Korbel, Jan................................... 32, 47
Krijgsveld, Jeroen...................... 6, 36, 56

L

Lemke, Edward................................... 38
Luscombe, Nick 22

M

Márquez, José.................................... 53
Meijers, Rob 53
Merten, Christoph 37

N

Nédélec, François............................... 34

P

Pepperkok, Rainer 64
Pillai, Ramesh 15, 62

S

Schneider, Thomas...............................52
Spitz, François.....................................16
Steinmetz, Lars................................... 22

W

Wilmanns, Matthias....................... 44, 65

Published by the EMBL Office of Information and Public Affairs

Lena Raditsch, Sonia Furtado Neves

DG Report

Iain Mattaj

Contributors

Claire Ainsworth, Chloe Balharry, Proteeti Bhattacharjee, Adam Gristwood, Dan Jones, Isabelle Kling, Nancy Podevin, Mary Todd-Bergman, Rosemary Wilson

Photography

EMBL Photolab (Hugo Neves, Udo Ringeisen, Marietta Schupp)
p.viii: hammeskrause architekten; p. xvii: Lars Kruse; p. 2: Jon Mold; p. 44: Marta Mayer; p. 65: Guðmundur Ingólfsson; p. 66 top: Christine Panagiotidis

Artwork

Petra Riedinger

Layout

Nicola Graf

Printed by

ColorDruck Leimen, Germany

The paper this report is printed on is 60% recycled and 40% from certified managed forests.

EMBL Heidelberg
Meyerhofstraße 1
69117 Heidelberg
Germany

Tel. +49 (0)6221 387 0
Fax +49 (0)6221 387 8306

EMBL Grenoble
6, rue Jules Horowitz, BP 181
38042 Grenoble, Cedex 9
France

Tel. +33 (0)4 76 20 72 69
Fax +33 (0)4 76 20 71 99

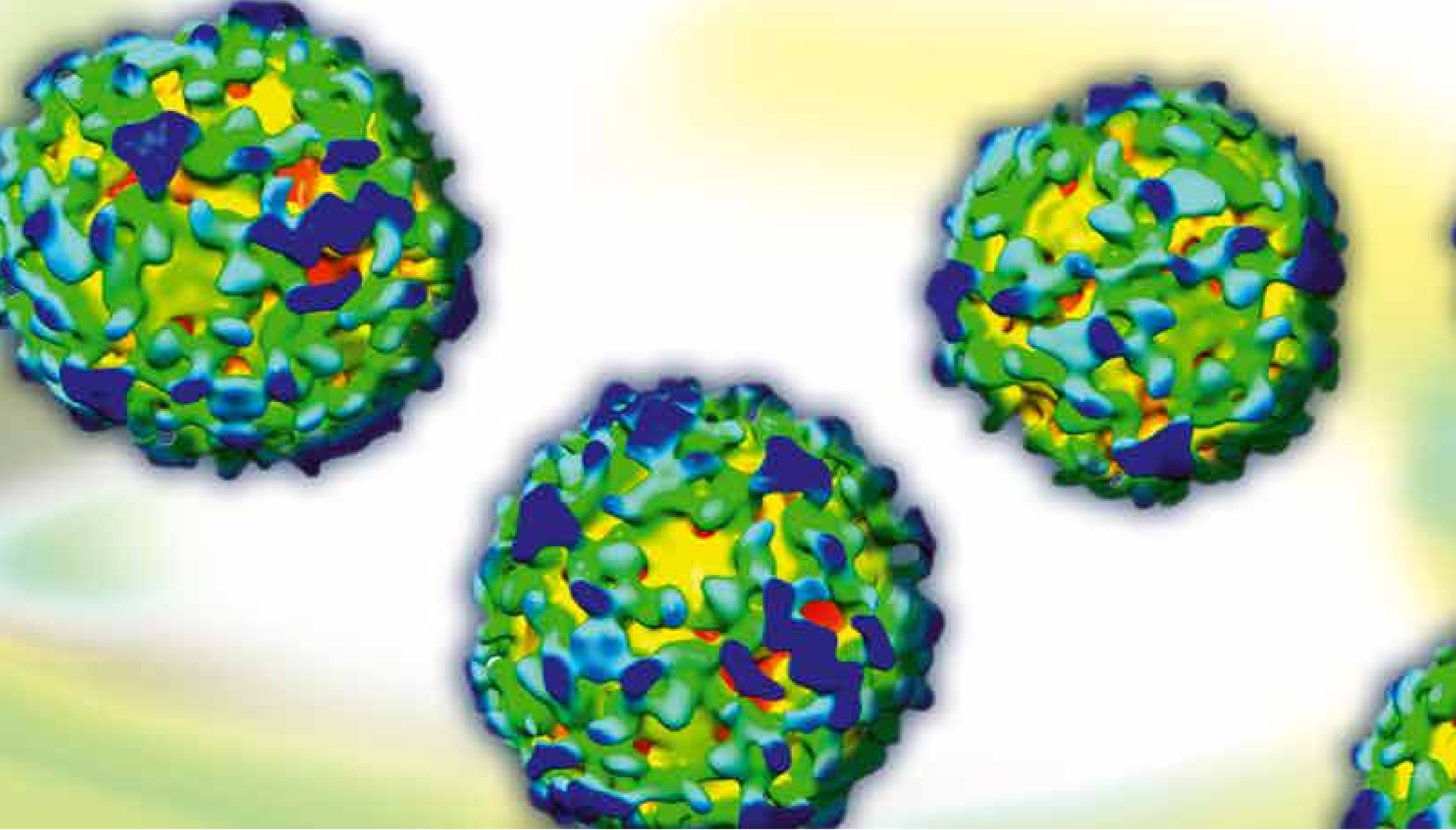
EMBL Hamburg
c/o DESY
Notkestraße 85
22603 Hamburg
Germany
Tel. +49 (0)40 89 902 0
Fax +49 (0)40 89 902 104

EMBL-EBI
Wellcome Trust Genome
Campus, Hinxton
Cambridge CB10 1SD
United Kingdom
Tel. +44 (0)1223 494 444
Fax +44 (0)1223 494 468

EMBL Monterotondo
Adriano Buzzati-Traverso Campus
Via Ramarini, 32
00016 Monterotondo (Rome)
Italy
Tel. +39 06 900 912 85
Fax +39 06 900 912 72

www.embl.org

© Copyright European Molecular Biology Laboratory 2013



EMBL member states:
Austria, Belgium, Croatia, Denmark, Finland, France, Germany, Greece, Iceland, Ireland, Israel, Italy, Luxembourg, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom. Associate member state: Australia