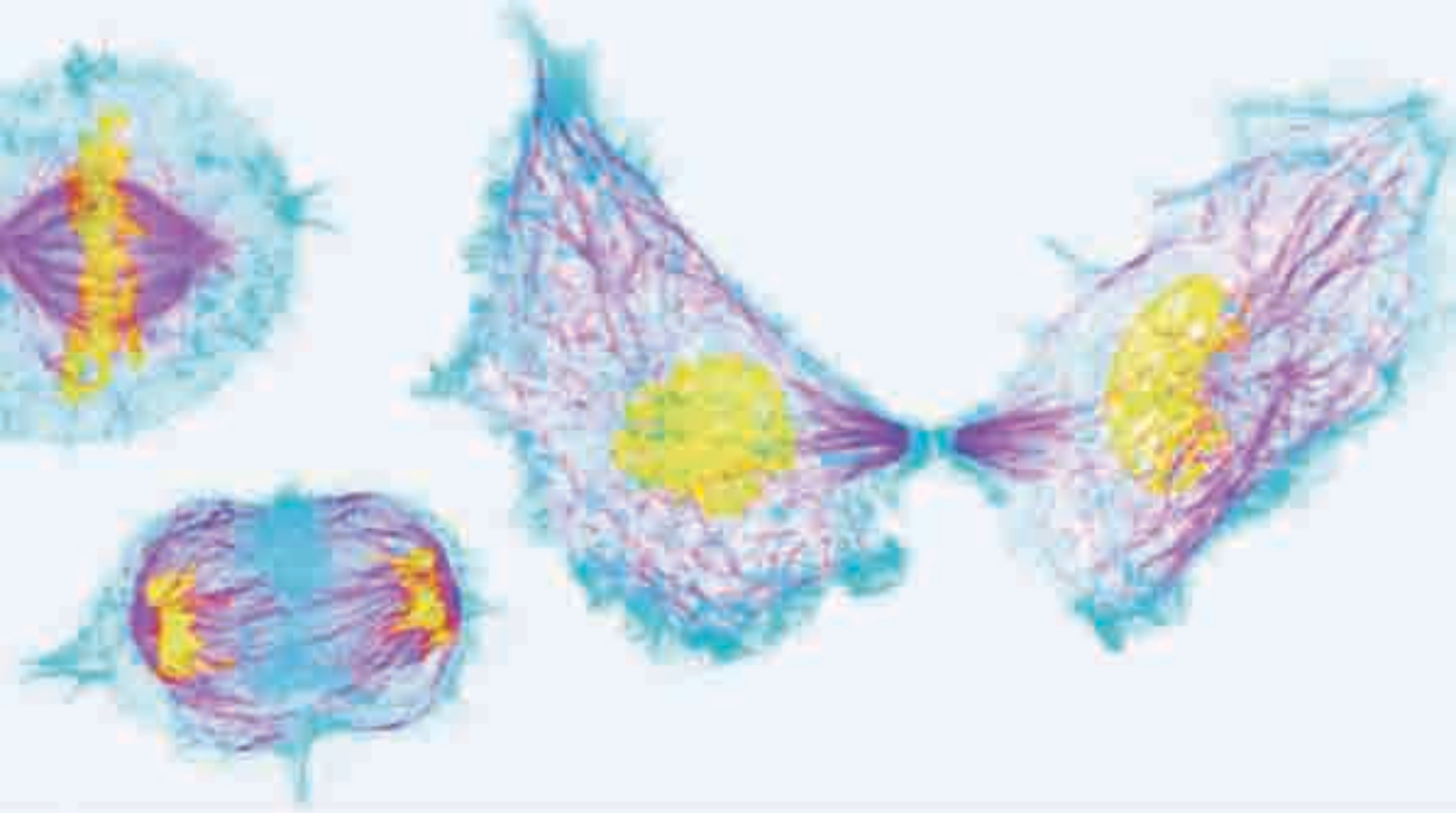


Annual Report

2009-2010

EMBL

Annual Report 2009-2010

European Molecular Biology Laboratory

Contents

The Director General's Report

Foreword

State of the Laboratory

vi Research
viii Services
xi Training
xiii Alumni
xiv Outreach
xiv Technology Transfer
xiv Administration
xv Facilities

Integration of European Research

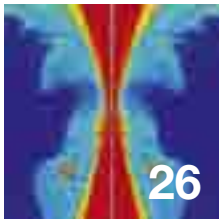
xvi Member state relations
xvi EMBL partnerships
xviii European Research Infrastructures
xviii EIROforum
xix Initiative for Science in Europe

xx Personnel statistics
xxii Financial report
xxiv Reviews of EMBL Scientific Units

Scientific Report

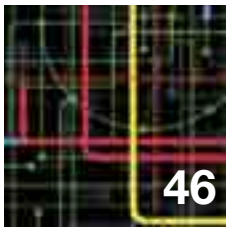
The grand scheme of things

4 Vital ingredients
11 Mapping the future
12 Ancient tweaking
16 Shaping up HIV
20 An accurate forecast of regulatory activity
21 It's a jungle in there
22 Young at heart



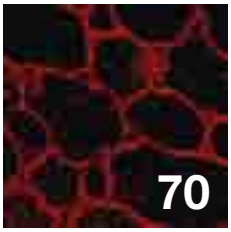
When disaster strikes

28 On call for DNA damage
32 Viral surveillance
36 The usual suspects
40 Poles apart
41 Take a deep breath
42 Surviving drought



Getting around and staying in touch

48 Built for speed
52 Getting on the right track
53 An elementary connection
54 Knowledge is strength
58 Biology gets the picture
62 The royal matchmaker
63 The grand scales of things
64 The EBI finds its wild side
68 Technology transfer at a glance



Upholding law and order

72 How females fight off their inner male
76 Fat chance for a cure
77 A nervous switch
78 Cue factors
80 Aberrant appendages
81 Tagging the tail on the histone
82 Movies for the human genome
87 When I grow up…
88 Waxing cutaneous
91 Decloaking the germ

92 A year in the life of EMBL

104 Index of group & team leaders

Imprint/Acknowledgements

When describing EMBL to people less familiar with it than I am, I often jokingly call it the last communist state. The reason is that at EMBL we still work according to five-year plans. The EMBL Programme specifies future strategies for all of our activities and is the basis on which our 20 member states decide on a five-year budget for the Laboratory. As the current Programme comes to an end in 2011 we have begun to develop our plans for 2012-2016. Looking so far into the future is never easy, because the ways in which science and technology evolve are heavily influenced by unexpected breakthroughs and are therefore hard to predict. At the same time it is a great opportunity to look both back and forward, to reflect and engage in critical self-evaluation.

EMBL's recent performance has been remarkably good. It consistently ranks as the top European research institute in molecular biology and genetics according to citations and holds rank four in the global comparison for the period 1999-2009. As Europe's only international research organisation in the life sciences EMBL's responsibility goes beyond excellent research. It provides cutting-edge infrastructure and services to the scientific community, offers advanced training for scientists, fulfills a crucial role in the integration of European research and acts as a role model for other research institutions.

Good performance, however, requires looking forward. The world of science is changing rapidly. Parts of biology are turning into 'big science'; big science in terms of data production, and in the requirement for international collaboration, infrastructure and technology, as well as interdisciplinary expertise and training.

Large-scale genome sequencing illustrates this new dimension. The same number of base pairs that took the Human Genome Project almost 10 years to produce, modern sequencing machines can now generate in hours at a fraction of the cost. The result: in 2009 DNA sequencing produced roughly 15 petabytes of data, the same amount the Large Hadron Collider at CERN will produce in a year when it is fully functional. Turning such quantities of data into useful information is a huge challenge.

To keep pace with this progress EMBL has to continuously evolve. The past year has seen many important developments that will not only prepare EMBL to meet the 'big science' challenges of the future but will also give European life sciences a headstart into the next five-year period. With the opening of the EMBL Advanced Training Centre the European biology community has a new meeting point and EMBL is equipped better than ever to address the growing need for interdisciplinary training in the life sciences. The inauguration of the beamlines at the PETRA-III synchrotron ring in Hamburg and the rapidly approaching upgrade of the ESRF beamlines in Grenoble herald a new era for integrated structural biology and service provision at both sites. On a broader scale EMBL has also been very active in the ESFRI (European Roadmap for Research Infrastructures) process, which will help to develop the next generation of international research infrastructures in Europe. Encouragingly, the first ESFRI projects are starting to take shape and have received initial financial commitments from their member states, but there is still a very long way to go.

These achievements and the many other accomplishments that are highlighted in the remainder of this report allow EMBL – and Europe's life sciences – to look confidently into the future. Our science is getting bigger and so are its challenges, but we are well prepared.

Iain W. Mattaj
Director General

State of the Laboratory

Research

This section features a few selected research highlights of the past year. More detail on EMBL's diverse research activities can be found in the following chapter, the Scientific Report.

Mouse Biology

The past year has been particularly eventful for EMBL's Mouse Biology Unit in Monterotondo, which celebrated its tenth anniversary in June 2009. Much of the success of the Unit can be attributed to its current Head, Nadia Rosenthal. She has directed EMBL Monterotondo since 2001, establishing it as a centre of excellence for mouse research and as a central hub in the international network of mouse biology. Nadia has now accepted the position of Scientific Head of the EMBL Australia partner laboratory network (see page xvi). She has been instrumental in building EMBL's relationship with Australia – EMBL's first associate member state – and will also hold the post of Director of the Australian Regenerative Medicine Institute at Monash University. She will continue to run EMBL's Mouse Biology Unit in Italy until her successor is found.

EMBL Monterotondo's site, the Adriano Buzzati-Traverso campus located in the Lazio region about 30 km north of Rome, has been purchased by the Consiglio Nazionale delle Ricerche (CNR). The 158,000m^2 international scientific campus was created in 1996 by a consortium of the CNR and international scientific organisations EMBL, EMMA (the European Mouse Mutant Archive) and the ICGEB (International Centre for Genetic Engineering and Biotechnology), and aimed to contribute to the development and internationalisation of Italian biological and biomedical research. The campus was built and until now owned by ENI SpA, a major Italian oil, gas and chemical conglomerate. ENI donated part of the existing research infrastructures for activities that were co-ordinated by CNR's Institute of Cell Biology (ICB/CNR) and overseen by campus co-ordinator Prof. Glauco P. Tocchini-Valentini.

Despite these ongoing administrative and organisational changes, EMBL Monterotondo's research activities have been highly successful during the past year. For example, Claus Nerlov and his group, in collaboration with colleagues in Madrid, discovered two proteins that control when and how stem cells at the base of the skin stop multiplying and switch to being skin cells. This work sheds light on the basic mechanisms involved not only in the formation of skin, but also on aspects of skin and other epithelial cancers (page 88).

Structural Biology

The past year saw a remarkable research highlight in the area of structural biology. What started in 2004 as a Unit-wide collaboration, spearheaded by the then joint Heads of the Structural and Computational Biology Unit Peer Bork and Luis Serrano, resulted in the publication of three papers back-to-back in *Science* in November 2009. The integrated structural biology study of the bacterium *Mycoplasma pneumoniae* also involved the groups of Anne-Claude Gavin, Rob Russell, Bettina Boettcher and Achilleas Frangakis, and support by the EMBL Core Facilities. Together, they have produced the most comprehensive picture of a "simple" bacterial cell to date (page 4).

EMBL's cryo-electron microscopy (cryo-EM) research possibilities are set to expand with the purchase of a state-of-the-art piece of equipment. The next-generation Titan KriosTM transmission electron microscope, made by FEI, will be delivered to EMBL Heidelberg in the autumn of 2010. This high-end instrument, of which there are only a few recently installed examples available to the scientific community, will allow several of EMBL's groups to enjoy more stability and higher throughput in their EM work. Refurbishments of the former nuclear magnetic resonance (NMR) area are underway to provide room for the new microscope, which will allow Structural and Computational Biology Unit groups to build on areas such as the automated study of the structural diversity of viral and eukaryotic coat proteins at the membrane, the molecular mechanisms of autophagy and the structure and function of large macromolecular assemblies.

EMBL Scientific Publications and Collaborations

- Total number of peer-reviewed publications: 343
- Internal collaborations: Publications co-authored by more than one EMBL group leader: 23
- External collaborations: 795 in total of which 84 resulted in publications

The EMBL Advanced Training Centre with the new canteen (left), under which the new training laboratories are located.

EMBL's structural biology outstations in Grenoble and Hamburg have also produced interesting research results during the past year. José Antonio Márquez and his group in Grenoble discovered that the key to a plant's response to drought lies in the structure of a protein receptor called PYR1 that interacts with the plant hormone abscissic acid (page 42). Meanwhile, the group of EMBL Hamburg Head Matthias Wilmanns has determined the structure of the signalling molecule Death-Associated Protein Kinase bound to calmodulin (page 54).

Cell Biology and Biophysics

The Cell Biology and Biophysics (CBB) Unit saw a change in leadership in 2010. Eric Karsenti, who led the Unit since 1998 and then ran it jointly with Jan Ellenberg in 2009, handed over sole responsibility to Jan in January 2010 so that he could co-ordinate the Tara Oceans expedition. This three-year, 150,000 km, marine journey was launched in September 2009 with Eric as the scientific co-ordinator and involves scientists from 50 laboratories in 15 countries. It will study questions of biodiversity and climate, the functioning of marine ecosystems and life's origin and evolution.

In his new role as sole Head of Unit, Jan Ellenberg, who led the Gene Expression Unit for three years before moving to CBB, aims to build on the Unit's existing strengths and its move towards a more systems-based, interdisciplinary approach. Jan also wants to develop CBB's interactions with other Units and add to its already powerful imaging and microscopy technology base. Together with Rainer Pepperkok, co-ordinator of EMBL's Advanced Light Microscopy Facility, and other collaborators in the European Commission-funded Mitocheck consortium, Jan's group

recently published the results of an unprecedented screen that identified many of the genes involved in mitosis in humans. Startlingly, they identified 600 new genes that play some role in mitotic cell division (page 82).

Developmental Biology

A special research highlight in the area of developmental biology came from Mathias Treier's group. Their discovery that switching off a specific gene located on a non-sex chromosome turns cells in the mature ovaries of female mice into cells typically found in testes overturned the dogma that the development of female traits is a default pathway. Instead, the study showed that the maintenance of the female state in cells is an active process (page 72).

Genome Biology and Bioinformatics

A new research focus emerging in both the Genome Biology Unit and at EMBL-EBI is the study of genetic variation data and particularly its link to phenotypic variation. This field of research has been revolutionised recently by ever-improving sequencing technology that is making the rapid sequencing of multiple individuals feasible for the first time. EMBL has taken the lead in analysing and curating sequence variation data in several different large-scale projects and in making the data available to the scientific community.

The 1000 Genomes project, launched in 2008 with the remit of producing and making publicly available the most detailed catalogue of human genetic variation, is now in full swing. The EMBL-EBI team led by Paul Flicek has been instrumental in ensuring the quality of the data, and in making it all – nearly eight terabases as of March 2010 – freely available online. In addition, EMBL-EBI's European Genome-phenome Archive (EGA) connects genetic and phenotypic information of individuals.

How this data can be put to use is exemplified by work by Jan Korbel's group in the Genome Biology Unit. They were the first to compare entire human genomes and determine that humans differ from each other mainly because of individual differences in gene regulation rather than in the sequences of the genes themselves (page 78).

With technology development playing an important role in the Genome Biology Unit's future activities, the end of 2009 saw the appointment of the Unit's first group leader in nanotechnology, Christoph Mertens. His group will focus on novel, droplet-based microfluidic approaches with applications in biology and biochemistry. With these novel techniques the Unit envisions developing customised 'lab-on-a-chip' applications that combine all steps of a genomics experiment in one device. This kind of technology would greatly speed up large-scale genomics research and decrease experimental variability and cost.

Services

Bioinformatics Services

EMBL-EBI continues to be Europe's hub for bionformatics, providing access to major core biomolecular resources including EMBL-Bank (nucleotide sequences), Ensembl (genomes), ArrayExpress (gene expression data), UniProt (proteins), the Protein Data Bank Europe (PDBe; macromolecular structures), and InterPro (protein motifs). EMBL-EBI's user community is increasing: on average, the common webportal to all resources currently receives four million hits a day.

From September 2008 to August 2009 all the core data resources grew significantly, for example receiving and processing more than 2.4×10^{10} bases of DNA sequence compared with 1.8×10^{9} bases in 2008. In total, the European Nucleotide Archive (ENA) now contains 8.9×10^{12} bases. In 2009, the EBI processed 3.6 million UniParc (Uniprot archive) entries (2.1 million in 2008), 60,646 microarray hybridisations (115,000 in 2008), 7174 macromolecular structures (5649 in 2008) and eight new eukaryotic genomes in Ensembl (12 in 2008).

In 2009 striking qualitative changes have accompanied the usual quantitative ones caused by the ever-increasing data flow rates that perpetuate exponential database growth curves. This has resulted in new data resources, restructuring of existing resources and the demise of some obsolete ones.

One highlight among the new databases is ChEMBL, a vast online database of information on the properties and activities of drugs and drug-like small molecules and their targets that was made freely available in January 2010. This drug-discovery resource is unique by virtue of its size: the number of small molecules is over 520,000, and it contains more than 2.4 million records of their effects on biological systems. The data could be a crucial bridge to help translate information from the human genome into innovative drug therapies.

Rationalisation of EMBL-EBI services has resulted in the restructuring of the nucleotide resources that are now organised into the Ensembl family of databases – which provide genomic data organised and annotated by the in-house Ensembl pipeline – and the ENA, the collection that stores, organises and makes available sequence data submitted by the scientific community. Owing to better integration of data in the Ensembl, ENA and UniProt databases, four databases have become redundant and will be discontinued: Integr8 (a portal for species with completely deciphered genomes), Genome Reviews (standardised annotation of non-vertebrate genomes), ASTD (Alternative Splicing and Transcript Diversity) and IPI (International Protein Index).

Despite these rationalisation efforts, the main data centre on the Hinxton campus is functioning at its maximum capacity of five petabytes and will not be able to sustain the data growth expected over

the coming years. To address this, EMBL-EBI has received funding from the UK Research Councils to lease two new state-of-the-art data centres in London. For the research community, the primary benefit will be improved access to the flood of biological information, and the new facilities will also provide the central hub for the emerging pan-European Life Science Infrastructure for Biological Information (ELIXIR, page xviii).

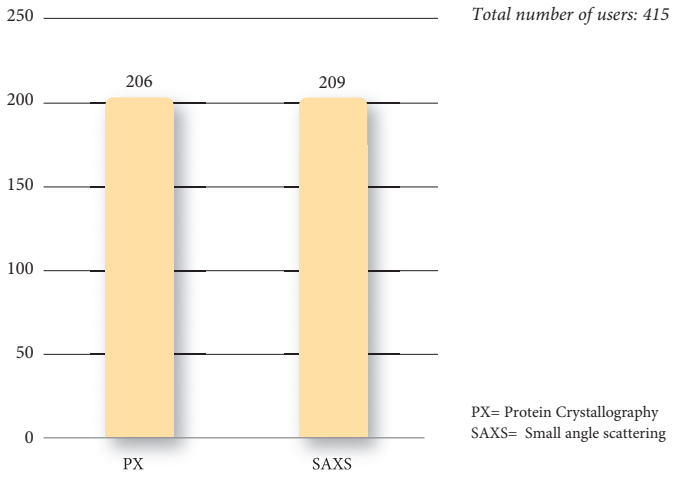
Structural Biology Services

EMBL's outstations in Hamburg and Grenoble jointly provide state-of-the-art synchrotron-based infrastructures and technologies in structural biology for the international scientific community

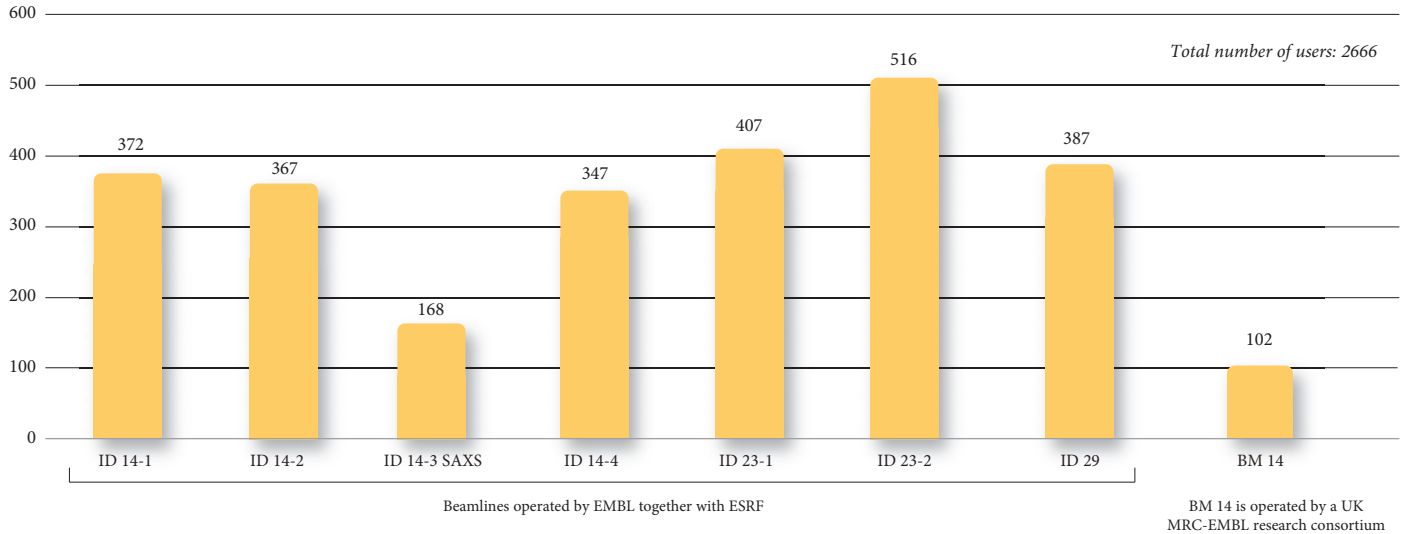
Conversion of the PETRA accelerator at Hamburg's German Synchrotron Research Centre (DESY) into the most brilliant storage-ring-based X-ray source in the world was completed in 2009. PETRA-III was officially inaugurated in the presence of Germany's Federal Minister for Education and Research, Annette Schavan, on 16 November 2009, with some user operation already having commenced in October. Of its 14 beamlines, three have been designed and will be built and run by EMBL for the structural biology community. Thomas Schneider and Stefan Fiedler's teams have constructed beamlines for small angle X-ray solution scattering and X-ray crystallography studies of biological macromolecules. These will operate at EMBL@PETRA3, the new integrated facility for structural biology. A new group leader recruited to run the sample characterisation facility, Rob Meijers, joined EMBL Hamburg in October 2009.

Construction work on DESY's new X-ray free-electron laser (XFEL), which will become operational in 2011, are ongoing. The signing ceremony between

Beamline Users EMBL Hamburg 2009



Beamline Users EMBL Grenoble 2009



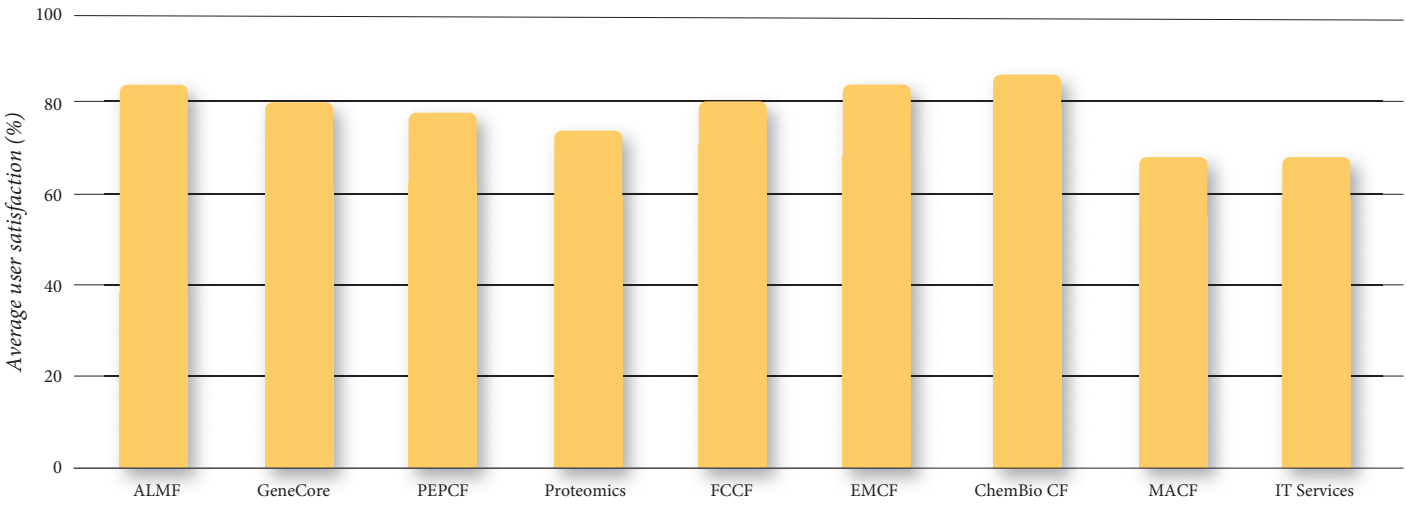
the partners from several countries took place on 30 November 2009. Once completed, the laser will allow researchers to analyse materials in atomic detail, film chemical reactions, generate three-dimensional images of the nanoworld and study processes under extreme conditions such as those occurring in the interior of planets. EMBL Hamburg scientists will be involved in exploring possible uses of the XFEL with biological samples.

For the beamline BM14 at the European Synchrotron Radiation Facility (ESRF), 2009 was a transition year in which the UK Medical Research Council (MRC), EMBL and India shared beam time. When the UK MRC contract finished as planned at the end of the year, a consortium comprising EMBL, campus partner ESRF and the Indian government took over the running of the beamline for a period of five years from 1 January 2010. The agreement guarantees the continuation of this state-of-the-art Multi-wavelength Anomalous Dispersion (MAD) beamline, and will strengthen links with the Indian crystallography community.

Core Facilities

EMBL's Core Facilities offer cutting-edge technology and expert support to researchers at EMBL and, when excess capacity allows, to external users. Like its research Units, EMBL also subjects its services to regular stringent quality control. In March 2010 the Core Facilities were reviewed by a panel of external experts and received a very positive evaluation. As well as the external evaluators, internal users are also very satisfied with the services offered by the Core Facilities. In December 2009 a large user survey was conducted to assess the quality of services, the general user satisfaction and the need for more or different services. The response was overwhelmingly favourable, with all eight facilities scoring good, very good or excellent for accessibility, comparison with other facilities elsewhere, staff competence, staff support and standard of results obtained.

User satisfaction with Core Facilities and IT Services



Internal users were asked to rank their satisfaction with the Core Facilities and the central IT Services that provide IT support for EMBL Heidelberg and Monterotondo. EMBL-EBI, Hamburg and Grenoble operate their IT infrastructure locally.

ALMF= Advanced Light Microscopy Facility, GeneCore= Genomics Core Facility, PEPCF= Protein Expression and Purification Core Facility, Proteomics= Proteomics Core Facility, FCCF= Flow Cytometry Core Facility, EMCF= Electron Microscopy Core Facility, ChemBio CF= Chemical Biology Core Facility, MACF= Monoclonal Antibodies Core Facility

Over the past year the Genomics Core Facility has seen major changes with next-generation sequencing having moved to the centre of activities. The new technology required significant investment both in terms of equipment – the facility currently operates three (four from June 2010) next-generation sequencing machines – but also in terms of staff training. In future it will be indispensible to upgrade the software and IT support and strengthen the bioinformatics expertise in the facility to ensure that the increased data production is matched with the necessary capacity for analysis.

Training

Advanced training has always been one of EMBL's core missions and the EMBL International Centre for Advanced Training (EICAT) organises a range of intramural and extramural training activities for EMBL staff and the scientific community. Over the period 2007-2009 130 courses, conferences and workshops were held at all five EMBL sites, reaching around 10,000 external participants.

In the future we will be able to further expand and develop our efforts in training, thanks to the new EMBL Advanced Training Centre, which officially opened on 9 March 2010 on the Heidelberg campus. The inauguration ceremony took place in the new 450-seat auditorium that will allow scientific conferences of unprecedented size to be held at EMBL. Among the 340 distinguished guests and friends of EMBL were the Ministers for Research of Germany, Professor Annette Schavan, and Israel, Professor Daniel Hershkowitz, Council delegates and representatives of ministries from many other EMBL member states, prominent scientists from all over the world and Klaus Tschira, whose ideas for the form of the building initiated the project and whose foundation provided generous support towards its realisation.

The higher capacity of the EMBL Advanced Training Centre will allow more scientists to benefit from EMBL's training activites, with the number of participants coming to the main Laboratory expected to more than double from the current level of roughly 2000 per year. Most importantly, however, the EMBL Advanced Training Centre will help us to explore new forms of training and build on the quality of our existing events. It will also strengthen EMBL's role as a central hub for advanced life science training in Europe and EMBO's role as a major funder of courses, conferences and workshops.

EMBL Director General Iain Mattaj and German Minister for Education and Research, Annette Schavan, at the EMBL Advanced Training Centre inauguration ceremony.

EMBL's Course and Conference Office has used the past year to gear up for the upcoming increase in activities. In close collaboration with EMBO, a new forward-looking meeting format covering important topics of the life sciences has been devised, the EMBO|EMBL Symposia. Furthermore, to celebrate the inaugural year of the EMBL Advanced Training Centre a series of lectures by Nobel Prize laureates, entitled 'Vision 2020', has been

The new EMBL Advanced Training Centre's 450-seat auditorium.

initiated. The practical course programme is also being expanded to include courses specially tailored to different levels of expertise, many of which will be organised in close collaboration with EMBL's corporate partners.

With the opening of the EMBL Advanced Training Centre, the Corporate Partnership Programme, which creates long-term relationships between EMBL and top-tier corporate partners that help sponsor training activities and conferences taking place in the new building, entered its active phase. As of April 2010 the programme attracted 15 leading companies from the life sciences and the pharmaceutical sector – Becton Dickinson, Boehringer Ingelheim, Eppendorf, GE Healthcare, Illumina, Leica Microsystems, Life Technologies, Merck Serono, Novartis, Olympus, PerkinElmer, Qiagen, Sanofi Aventis, Sigma Aldrich and Thermo Fisher Scientific – and 21 January 2010 saw top representatives from all the companies come together for their first official event. The Corporate Partnership Programme generates an annual income of €385,000, of which €100,000 will provide conference fellowships to young scientists who would otherwise be unable to attend. Further funds will support training activities in the EMBL Advanced Training Centre.

2009 was a very successful year for the EMBL International PhD Programme (EIPP) both in terms of graduations, which exceeded 50 for the first time since the programme's foundation in 1983, and new applications. The more than 1200 applications in 2009, of which one third came from EMBL member states, reflect the high reputation of the EIPP and prove that its attraction for European and global student communities continues to grow. With an annual intake of roughly 50 students, which keeps the student body at a steady state of 200, the admission rate to the EIPP was close to 1:25 in 2009, making the

EIPP one of the most competitive PhD programmes in biology in Europe. To stay abreast of this increasing popularity and the overall growth of the programme, we now run two similarly sized rounds of applications a year. This reorganisation is also a response to the Europe-wide harmonisation of studies according to the Bologna protocol and its impact on student schedules. To increase applications from currently under-represented member states and to attract students from other disciplines – most notably physics, chemistry and engineering – we are refining our advertising strategies and have initiated recruitment events targeted at specific countries and student communities. As part of this initiative we have also implemented a student ambassadors scheme to enhance the visibility of the EIPP among member state undergraduate communities.

Owing to the success of the EMBL Interdisciplinary Postdoc (EIPOD) initiative in obtaining external funding from the European Commission FP7 Marie-Curie Co-fund scheme, the EMBL Postdoctoral Programme is now managed by a full-time Postdoctoral Administrator together with a dedicated academic mentor, and has introduced several new developments in the past year. The 2009 EIPOD selection saw an increased number of nearly 200 applications, from which 19 were selected. A new postdoctoral 'second mentor' scheme encourages all EMBL postdocs to obtain additional advice and guidance from a mentor other than their dedicated academic supervisor. Together with other activities of the Postdoctoral Association and complementary training courses, EMBL postdocs are well prepared for a future career in academia or industry. In addition, with the beginning of the new Indicative Scheme in 2012, the programme will start to provide additional social benefits for all postdoctoral fellows following the recommendation of the European Charter for Researchers.

In the past year the European Learning Laboratory for the Life Sciences (ELLS) has organised five hands-on courses for secondary school teachers, three in Heidelberg and two in Monterotondo. In collaboration with EMBL-EBI, ELLS launched a new course dedicated to bioinformatics in the classroom at Hinxton in March 2010.

At EMBL-EBI the new IT suite has accommodated over 1400 trainees, hosting not only the EBI Hands-on Training Programme, but also offering a large number of courses and workshops to the scientific community. To complement these face-to-face training activities and to reach out to an even larger audience, a new e-learning pilot training programme was implemented and is now being expanded further.

To complement scientific training, EMBL started a formal programme of vocational training, the General Training and Development Programme, for all its scientific, administrative and support staff in 2008. We offer a variety of training courses from computer skills to language training to leadership and management expertise. This programme has been very popular with the staff, indeed it is enormously oversubscribed even though more than 100 courses were organised for the benefit of around 500 participants in 2009. The General Training and Development Programme helps the Laboratory to fulfill its obligation to the member states by ensuring that all categories of staff leave EMBL with a skill set that makes them attractive candidates for organisations and institutions in the national systems.

Alumni

The total number of EMBL alumni has grown to 4764 with 240 people leaving EMBL (including fellows, trainees and visitors) in the past year. 80% have taken up positions in the EMBL member states. The EMBL Alumni Association (EAA) acquired 85 new members, which brings the total up to 1649 members with a membership rate of 34%.

2009-2010 was a very eventful year for EMBL alumni. Local chapters met in Heidelberg, Porto, Helsinki, Dublin and Dilofo. The local chapter meeting in Dublin on 24 February 2010 was followed by an event organised by Science Foundation Ireland that showcased EMBL to the wider Irish scientific community and which was attended by the Irish Minister for Science, Technology and Innovation, Conor Lenihan TD.

The highlight of the year was the staff-alumni reunion in the EMBL Advanced Training Centre on 8 March 2010, the day before the official opening ceremony. 200 participants heard a variety of talks with lots of opportunities to network in between. The event culminated with the opening ceremony of the Matti Saraste Courtyard, which was made possible by donations of staff and alumni.

Former EMBL Hamburg predoc Jens Preben Morth, who is now an Associate Professor at Aarhus University, has been selected as the winner of the 2010 John Kendrew Award – which recognises excellence in science communication or academic achievement after leaving

EMBL. Jens Preben Morth receives the prize for his outstanding contribution to the structural biology of membrane proteins, and for his enthusiastic involvement in science education for school students.

The overwhelming response to the John Kendrew Award, as shown by the high number of outstanding applications and its enthusiastic reception by EMBL staff and alumni, motivated the EAA board to aim to offer the award indefinitely by launching the first EAA fundraising campaign in October 2009. To date €6000 has been raised, which will secure the award until 2014.

The EAA also launched an initiative in 2010 to establish a European Molecular Biology Archive (EMBA). The idea was inspired by the 2007 statement by Sydney Brenner and Richard Roberts in *Nature*: "Let's not wait until memories have faded and papers been discarded at the end of a career before deciding to save our heritage." Although the initiative gained instant support from key past and present EMBL figures, much work will be required to secure resources for establishing such an archive.

Outreach

The tremendous potential the life sciences holds for societal benefits endows scientists with the social responsibility to inform the public about advances in their research, as well as its potential applications, inherent risks and benefits and ethical implications. For this reason the Office of Information and Public Affairs (OIPA), EMBL-EBI's Outreach and Training team, the Science and Society Programme and EICAT, supported by many of EMBL's scientists, organise a range of outreach activities.

In the first half of 2010 alone 14 groups of students from various schools and universities have visited EMBL compared with 13 in the whole of the previous year. Most came from relatively close by – e.g. France, Spain and Belgium – but one group travelled from as far as Israel to learn about EMBL.

In 2009 EMBL relaunched its websites with a new look and feel based on a completely re-engineered content management system. A more recent addition to EMBL's internet presence is a new intranet portal – a dynamic 'one-stop-shop' for news and events that complements the EMBL Etcetera newsletter as a platform for information dissemination within the EMBL community. In a second phase, in 2010, further features of the content management system – including an optimised search facility, RSS feeds and a multimedia gallery – will be developed.

EMBL regularly communicates research highlights and news about other activities of the organisation to a broad range of international media. Particularly popular with the press in the past year were the first comprehensive picture of a minimal cell (*Mycoplasma pneumoniae*) and the screen conducted by the international Mitocheck consortium that identified 600 genes involved in mitosis in humans. Also the announcement of a £10 million investment in the ELIXIR project (page xviii) by the Biotechnology and Biological Sciences Research Council (BBSRC) received widespread attention by both local and European research media.

The Science and Society Programme's annual conference in November 2009 was organised under the lead of EMBO and enjoyed a record number of 270 participants. The two-day event, 'Food, Sustainability and Plant Science: A Global Challenge' focussed on our food supply which, next to climate change, is the greatest challenge that faces the world. Other Science and Society events included Forum Lectures by Jens Reich, James Mallory and Buddhist monk Matthieu Ricard and an EMBL-EBI Science & Society symposium in Cambridge, which tackled the question of 'Who owns science? Promises and pitfalls of public-private partnerships'.

Technology Transfer

One of EMBL's missions is to develop its discoveries to the benefit of society. When EMBL's technology-transfer subsidiary, EMBL Enterprise Management Technology Transfer GmbH – EMBLEM for short – was established in 1999, it was hoped it would break even within ten years. In fact, as the company celebrated the end of its first decade on 19 June 2009, there was even more cause for celebration. Exceeding all expectations, EMBLEM has actually been generating a profit for EMBL, its scientists and the member states since 2004, breaking even in less than half the time predicted.

As well as ensuring a modest but steady income for EMBL and benefiting the member states and society by translating basic research results into marketable tools and products, the creation of EMBLEM has allowed the protection and commercialisation of innovations to be streamlined. To date more than 400 EMBL staff are on record as inventors, over 250 commercial partners world-wide are licensing EMBL-derived technologies and EMBLEM has a portfolio of more than 260 granted patents and patent applications, over 90 copyrights and trademarks and 12 spin-out companies.

Administration

This year saw the departure of Bernd-Uwe Jahn after his eight-year leadership of the Administration, which produced significant improvements in staff relations and in the efficiency of the administration services. He was succeeded by Ralph Martens at the beginning of 2010. Ralph joins us from his previous position as Director of Common Administrative Services at the International Criminal Court in The Hague. Previously he has held positions in both public and private sectors in various countries in Europe and the US and was listed in the 'Who's Who in Leading Executives of America'.

During the course of the year we have made many improvements to our com-

EMBL's Administrative Director Ralph Martens

puter systems to address comments raised in a staff survey of administration services conducted in 2008. To help scientists make the best use of the flexibilities of the cash budgeting system, the financial systems – particularly the interface for non-accounting personnel – were further developed to produce a simple online system that allows budget holders to monitor the status of grants and budgets. A step change in the quality of information available to group and team leaders now enables them to make better-informed decisions about recruitment, staffing and purchasing. This was achieved by integrating the SAP Finance and HR modules to allow future personnel costs to be included in financial reports and to simulate developments up to six years into the future. Additionally, the in-house SAP team has developed reports based on the new system that allow a better oversight of the financial situation and better monitoring of budgets, particularly for externally funded projects.

The Personnel Unit also responded to suggestions raised in the administrative survey. Now dedicated teams of two people, always comprising one generalist and one recruiter, address the recruitment needs of every specific EMBL Unit. At the EMBL-EBI a Personnel Officer is responsible for both general Human Resources and Recruitment.

Facilities

With the completion of the EMBL Advanced Training Centre and the associated move of EICAT and large parts of EMBL's Administration in March 2010, space has become available in the main Laboratory in Heidelberg. This will be reorganised and refurbished over several years as budgets allow. Restructuring work in progress includes the redevelopment of space at the back of the building to house the central IT facilities and the Advanced Light Microscopy Core Facility, a redesign of the Directorate area and the refurbishment of building 3 (Containment), which will include new staff shower facilities that are expected to be operational by the end of August 2010. From July 2010 the roof of the connection building will be renovated and in August 2010 the complete refurbishment of the third floor will begin. In further steps the Cell Biology and Biophysics and the Structural Biology Units will be refurbished.

Integration of European Research

Member state relations

Australia joined EMBL in 2008 as the first associate member state. One of the goals was to establish EMBL Australia, EMBL's first non-European node. A joint venture supported by the Australian government and involving the universities of Sydney, Queensland, Western Australia and Monash and the Commonwealth Scientific and Industrial Research Organisation (CSIRO), EMBL Australia provides a direct link for Australian and European researchers that allows them to benefit from the world-leading science taking place on opposite sides of the globe. The first EMBL Australia Council meeting took place on 6 October 2009 under the chairmanship of Richard Larkins, the former Vice-Chancellor and President of Monash University. This was followed by the official launch of EMBL Australia in Melbourne, which was attended by Australia's Minister for Innovation, Industry, Science and Research, Kim Carr. During the event Senator Carr announced the appointment of EMBL Monterotondo Unit Head Nadia Rosenthal as Scientific Head of EMBL Australia (see page vi) and the establishment of an international PhD programme to allow Australian students to undertake their doctorate in Europe and gain a joint Australian-EMBL PhD degree. EMBL Australia will further comprise the creation of a new EBI mirror site at the University of Queensland, complemented by a national support network for bioinformaticians. The first group leader in the EMBL Australia Partner Laboratory Network, Edwina McGlinn, has already been appointed.

Kim Carr, David de Kretser, Nadia Rosenthal, Silke Schumacher, Iain Mattaj and Richard Larkins

EMBL Partnerships

Nordic Partnership for Molecular Medicine

2010 began with a meeting in Heidelberg with the scientists of the Nordic EMBL Partnership for Molecular Medicine. Established in 2007, the partnership includes the universities of Oslo, Umeå and Helsinki, all of which have established 'nodes' at the Centre for Molecular Medicine Norway (NCMM), the Laboratory for Molecular Infection Medicine Sweden (MIMS) and the Institute for Molecular Medicine Finland (FIMM). They combine complementary strengths and collaborate closely with EMBL to tackle challenging problems in biomedicine. Since the launch of the partnership, the nodes have hired 25 young group leaders, and the meeting was the first opportunity for them to get to know each other, visit EMBL Heidelberg and meet EMBL faculty who came from all five EMBL sites to participate.

One of the nodes, FIMM, was officially inaugurated on 16 March 2010. Operated by the University of Helsinki in collaboration with the Hospital District of Helsinki and Uusimaa, the National Institute for Health and Welfare, and the VTT Technical Research Centre of Finland, the joint research institute has 150 employees. They work on cancer, cardiovascular, neuro-psychiatric and viral diseases, and carry out translational research to explore new diagnostics and treatments and promote human health via research on personalised medicine.

Unit of Virus Host Cell Interactions

26 June 2009 saw the official signing ceremony of the Unit of Virus Host Cell Interactions (UVHCI) Unité Mixte Internationale (UMI) between EMBL, CNRS and the Université Joseph Fourier in Grenoble. The international unit is a unique structure in France in the areas of biology and health, and will facilitate interdisciplinary research in structural and molecular biology. Head of EMBL

Grenoble Stephen Cusack will direct the unit for the first five years, and the Deputy Head will be Rob Ruigrok, Professor at the Université Joseph Fourier. The international unit already underwent its first review by French reviewing body AERES in February 2010 and achieved favourable results.

EMBL Grenoble's campus, the Polygone Scientifique, is to be developed into a world-class science and technology park – an 'ecosystem' of innovation named GIANT – in an initiative supported by the French government. GIANT's first construction projects are underway and mark the beginning of a €500 million overhaul for the area, which has long boasted a top-class research infrastructure that includes the ESRF, the ILL, EMBL, CEA, CNRS and the Université Joseph Fourier, as well as three centres of technological excellence and several high-level university programmes. With the planned new teaching and research buildings, recreational facilities and meeting places for researchers, transport links and sustainable housing, it is projected that the site will welcome 20,000 scientists and students and 10,000 inhabitants by the year 2015. The entrance to the site will also feature a Visitors Centre, which will present the work of all the campus' scientific institutes to the growing numbers of interested members of the public using videos, models and interactive exhibits.

To exploit research results obtained at EMBL Grenoble a new EMBL spin-out company, Savira Pharmaceuticals, was created in September 2009 to focus on the development of drugs for the treatment of influenza. Co-founded by the Vienna-based biotech company Onepharm Research and Development GmbH, Savira builds on the breakthrough results from the groups of Stephen Cusack and Darren Hart at EMBL Grenoble, who produced high-resolution images of several crucial domains of the influenza virus polymerase, the enzyme that copies the virus's genetic material and allows it to multiply in human cells. These findings open new avenues for the structure-based development of anti-influenza drugs, which could target the viral polymerase to selectively stop its reproductive cycle. The spin-out company, which is based in Vienna, received €1 million support from the AustriaWirtschaftsservice (AWS) and was awarded third place by the City of Vienna Future Award in the category 'Newcomers and Start-ups'.

Collaborations with institutes in non-member states

Since the signing of an Agreement of Academic Exchange between EMBL and Japan's National Institute of Basic Biology (NIBB) in 2005, both institutes have worked together to span the distance with exchange visits for lectures, workshops, conferences and other academic activities. In addition, joint symposia on the topics of developmental biology, microscopy and imaging, epigenetics, systems biology, functional

Members of the Nordic EMBL Partnership for Molecular Medicine gathered in Heidelberg

genomics, mouse biology and structural biology have been held regularly in Japan and Europe to promote academic exchange between researchers of both institutes as well as other Japanese and European scientists. In October 2009 the first NIBB-EMBL PhD student mini-symposium took place in Heidelberg, with 20 students from both institutes presenting their research. The agreement was renewed during a visit by the NIBB Director General, Prof. Kiyotaka Okada, to EMBL in February 2010.

European Research Infrastructures

EMBL is involved in seven out of ten European Strategy Forum on Research Infrastructures (ESFRI) biomedical science projects. For those included on the ESFRI roadmap in 2006 the European Commission FP7-funded preparatory phase is now entering its third year. ELIXIR (European Life Science Infrastructure for Biological Information) is co-ordinated by the EMBL-EBI Director Janet Thornton and EMBL is participating in INSTRUCT, BBMRI and Infrafrontier. The projects from the 2008 update of the ESFRI roadmap will enter their preparatory phase later this year. Euro-BioImaging will be co-ordinated by Jan Ellenberg (Head of the Cell Biology and Biophysics Unit) together with a representative of the medical imaging community, and EMBL will also participate in EU-Openscreen and the European Marine Biology Resource Centre.

ELIXIR

The aim of ELIXIR is to construct a sustainable infrastructure for biomolecular data and related information in Europe. Over the past two years, an extensive stakeholder consultation has sought input into ELIXIR's scope and structure. One important conclusion from this stakeholder consultation is that ELIXIR's structure should comprise a hub, based at EMBL-EBI in Hinxton, Cambridge, UK, and several nodes located throughout Europe. The hub will be responsible for holding the core data collections and enabling the development and integration of nodes into a European-wide distributed infrastructure.

A request for suggestions to contribute to ELIXIR's construction was published in April 2010. Its purpose is: to consolidate the extensive stakeholder input that we have had to date; to generate ideas for further discussion with national funders as to how ELIXIR's stakeholders wish to contribute to ELIXIR on a pan-European level; and to gather information on the kinds of support that potential nodes expect from the hub as input to further discussions with research organisations and funders. It is intended that long-term infrastructure support will be paid for through national, European, EMBL and other funding streams. As part of this a centrally managed fund will be established that will both support the co-ordination activities (at the hub) and help leverage further investment. At the organisational level, ELIXIR may initially be established as an EMBL Special Project.

Euro-BioImaging

Euro-BioImaging aims to provide access and training to imaging technologies across the full scale of biological and medical applications, from molecule to patient. Euro-BioImaging is at an early stage. It has just been awarded funding for the preparatory phase by the European Commission and will enter its three-year preparatory phase in 2011. The first stakeholder meeting of Euro-BioImaging, which involves 23 countries across Europe, took place on 21-22 September 2009 at EMBL Heidelberg. The meeting gathered representatives of the scientific community, funding and governmental organisations and industry to discuss potential participation and the content and structure of the project.

EIROforum

From July 2009 EMBL took over for a year as the chair of EIROforum, a partnership of the seven largest intergovernmental research infrastructure organisations in Europe (CERN, EFDA-JET, EMBL, ESA, ESO, ESRF and ILL).

The main achievements during this year have been the renewal of the Statement of Intent between EIROforum and the European Commission, the publication of a policy paper on research infrastructures and the organisation of a technology-transfer conference.

The EC-EIROforum Statement of Intent was originally signed in 2003. Over the past year it was reviewed and updated. The new European Commissioner for Research and Innovation Máire Geoghegan-Quinn and the EIROforum Director Generals expressed their mutual interest in continuing the co-operation.

The seven EIROforum organisations are the largest providers of research infrastructures in Europe. To make their collective experience more widely available, especially to the new ESFRI projects, EIROforum published a position paper entitled 'Establishing New Research Infrastructures in Europe – The EIROforum Experience' in March 2010 just before the sixth European Conference on Research Infrastructures, ECRI 2010, took place in Barcelona. Representatives of European research across all disciplines and science policy makers gathered to discuss challenges and issues currently facing European research infrastructures, such as the prioritisation of research infrastructures, management and financial issues, governance structures and a general future strategy for European research.

In November 2009, EMBL and its EIROforum partners organised a conference on technology transfer in Heidelberg to exchange knowledge and best practices across disciplines. Around 100 participants attended the conference including representatives from European science and technology infrastructures,

The EIROforum Director General assembly took place at EMBL Heidelberg

European politicians, representatives of the European industry, representatives of the national governments of the European Union, representatives of the European Investment Bank, the European Patent Office, representatives of the institutions of the European Commission and other research communities. On the basis of the discussions at the conference a recommendation on the management of intellectual property and knowledge transfer was developed.

In July 2010 EMBL will hand over the EIROforum chairmanship to EFDA-Jet, the European Fusion Development Agreement.

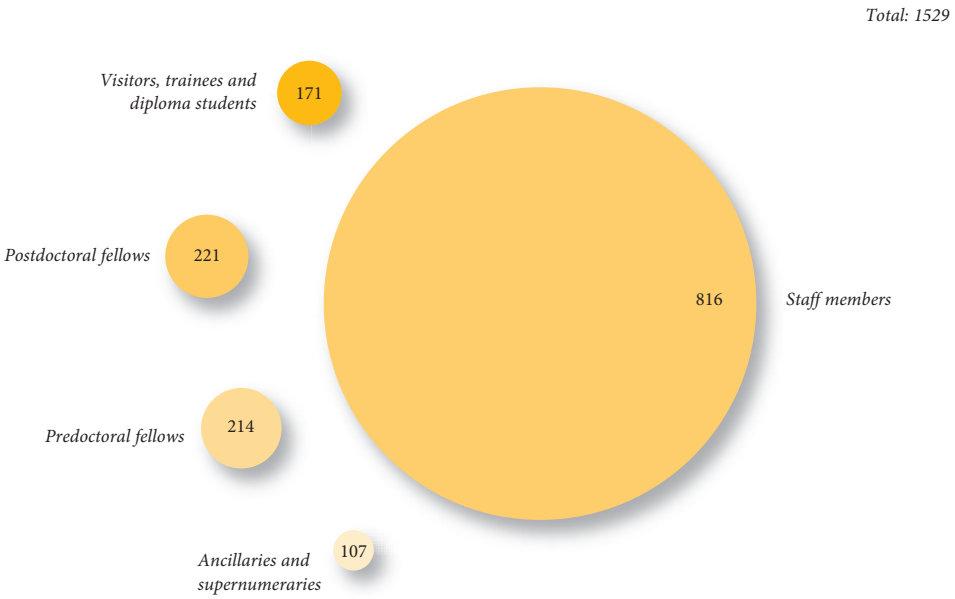
Initiative for Science in Europe

EMBL is a founder member of the Initiative for Science in Europe (ISE), an organisation of European scientific societies and organisations. ISE participated, with the Spanish EU presidency, the European Commission and ESFRI, in the organisation of ECRI 2010 and in May 2010, organised a conference on the future of the European Research Council (ERC) 'ERC – From Programme to Institution'. The event featured a discussion on the achievements, challenges and future of the ERC and was attended by many representatives of European science and science policy.

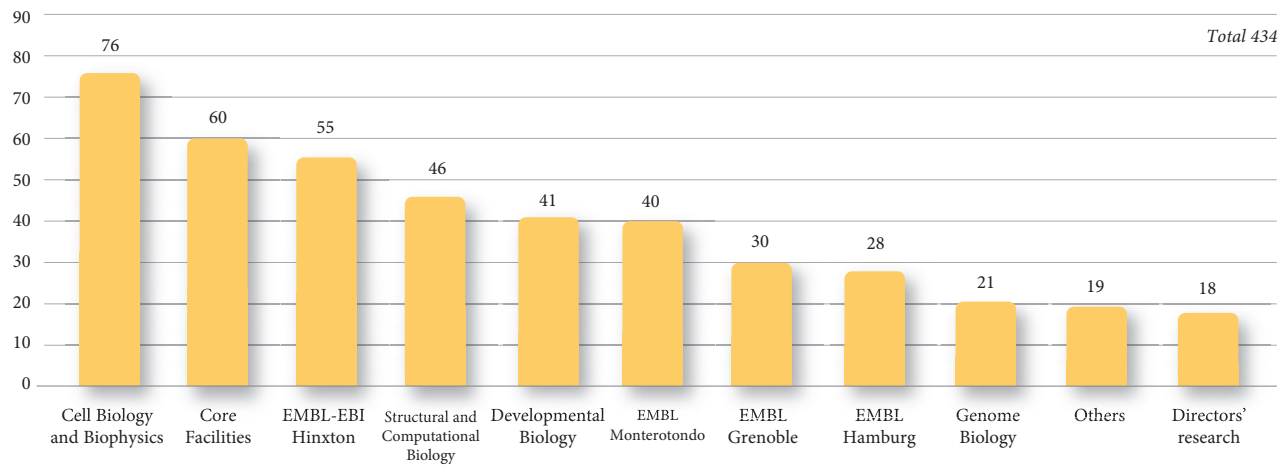
Personnel statistics

On 31 December 2009, 1529 people, including visitors, from more than 60 nations were employed by EMBL.

Personnel on 31 December 2009

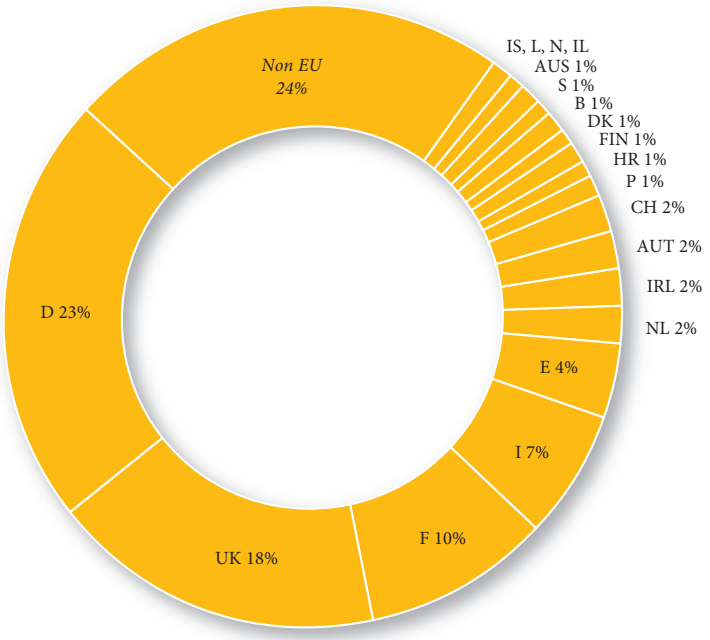


Visitors to EMBL Units during 2009

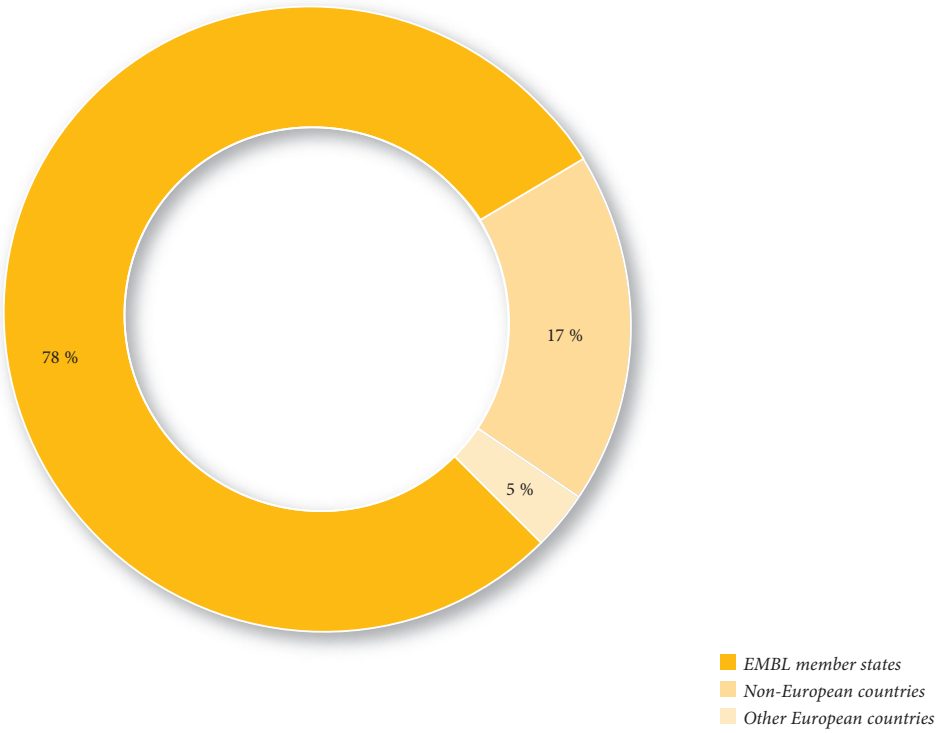


Staff Nationalities – Research

Please refer to CD for more information



Staff Nationalities – All



Financial report

External grant funding

	2009		2008	
	€ 000	%	€ 000	%
BBSRC	1,675	4.9	1,457	3.6
BMBF	4,144	12.1	3,257	8.1
DFG	1,404	4.1	1,461	3.7
EC	11,695	34.1	16,893	42.2
HFSPO	457	1.3	492	1.2
MRC	8	0	7	0
NIH	7,750	22.6	6,169	15.4
VW Foundation	243	0.7	294	0.7
Wellcome Trust	4,160	12.1	5,240	13.1
Others	2,747	8.0	4,735	11.8
TOTAL	**34,283**	**100**	**40,005**	**100**

Income/expenditure statement

INCOME	**2009**	**2008**
	€ 000	€ 000
Member states contributions		
Ordinary	86,436	82,947
One-off contribution from Germany	1,648	2,642
Internal Funding	18,609	18,734
External Funding	34,283	40,005
Other Receipts	15,479	18,867
Total Income	**156,455**	**163,195**
EXPENDITURE		
Staff Costs	82,933	81,255
Operating Costs	54,001	49,101
Capital Expenditure	23,555	25,942
Total Expenditure	**160,489**	**156,298**
Surplus (deficit) for the year	**(4,034)**	**6,897**

EMBL budget 2009: 156 € million

Please refer to CD for more information

Member states' contributions

	Ordinary and one-off contributions				Pension contribution	
	2009		2008		2009	2008
	€ 000	%	€ 000	%	€ 000	€ 000
Austria	1,859	2.2	1,739	2.2	27	27
Belgium	2,268	2.7	2,122	2.7	33	33
Croatia	119	0.1	56	0.1	2	1
Denmark	1,449	1.7	1,356	1.7	21	21
Finland	1,159	1.4	1,085	1.4	17	17
France	13,589	15.9	12,724	16.0	200	197
Germany	17,554	20.6	16,441	20.6	258	254
Greece	1,799	2.1	1,683	2.1	26	26
Iceland	85	0.1	80	0.1	1	1
Ireland	1,015	1.2	949	1.2	15	15
Israel	742	0.9	694	0.9	11	10
Italy	10,963	12.9	10,267	12.9	161	159
Netherlands	3,930	4.6	3,677	4.6	58	57
Norway	1,697	2.0	1,587	2.0	25	25
Portugal	1,040	1.2	973	1.2	15	15
Spain	6,590	7.7	6,174	7.7	97	95
Sweden	2,234	2.6	2,090	2.6	33	32
Switzerland	2,617	3.1	2,449	3.1	38	38
United Kingdom	14,544	17.1	13,625	17.1	214	211
SUBTOTAL	**85,253**	**100**	**79,771**	**100**	**1,252**	**1,234**
Special contribution Croatia	23		23		-	-
Luxembourg	119		112		2	2
Special contribution Luxembourg	41		41		-	-
Australia associate member	1,000		3,000		-	-
TOTAL CONTRIBUTIONS	**86,436**		**82,947**		**1,254**	**1,236**

Additional one-off contributions

	2009	2008
	€ 000	€ 000
Germany – to EMBL Advanced Training Centre	1,648	2,642
TOTAL	**1,648**	**2,642**

2009/2010 Reviews of EMBL Scientific Units

EMBL Units are reviewed in depth every four years by expert international panels organised by the Scientific Advisory Committee. To ensure openness, the review reports are submitted in confidence to EMBL Council and the Director General. The formal responses of the Director General to the reports are made public, to communicate the adjustments made by the Laboratory in response to the reviews, when needed.

Director General's Response to the Cell Biology and Biophysics Unit Review Report

Heidelberg · 6 and 7 May 2009

1. I wish to thank the review panel for their in-depth and constructive review report. Because of the circumstances at the time of review the number of group and team leaders in the Cell Biology and Biophysics Unit (CBB) was temporarily increased to fourteen, which meant that the panel had to work extremely effectively to stay within the time constraints. They did this without losing sight of the need to provide detailed evaluations in large part due to the excellent chairmanship of Sandra Schmid.

2. The panel was very positive about the broadening of the research focus of the Unit over the review period. In 2005, many of the groups in the Unit were working on aspects of the biology of microtubules and, although research in this area is still considered by the panel to be a major asset of the Unit and an important field of research, several new topics have been added by a broad recruitment policy that the panel viewed very favourably.

3. There was particular praise from the panel for Eric Karsenti's scientific vision and leadership throughout the review period as well as enthusiasm for the decision to appoint Jan Ellenberg as Eric's successor.

4. The panel appreciated the technology development efforts in the area of light microscopy, underlining the novelty and broad usefulness of both the Light Sheet Microscopy methods pioneered by the Stelzer lab and the high throughput, high content, cell-based phenotyping systems developed in the collaborative effort between Ellenberg and Pepperkok.

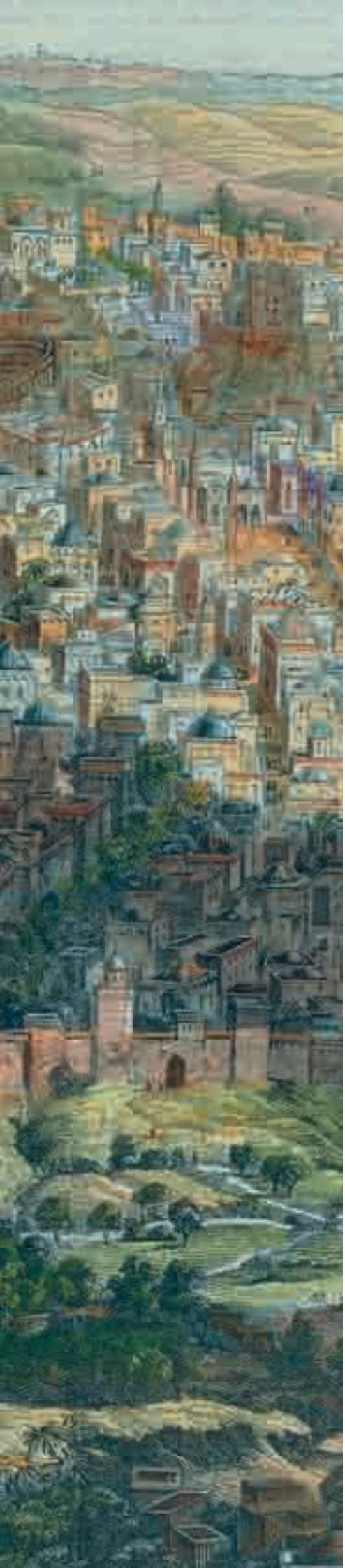
5. A particular feature of the Unit is the uniquely successful, in the view of the panel, blend of physicists and biologists that collaborate extensively in the analysis of complex biological functions by combining modelling and simulation with experiment. The catalytic role of Nédélec and Karsenti in generating and supporting this interdisciplinary environment was highlighted. The panel advised consideration of formal mechanisms that will help maintain and increase the exchange between the physicists, biologists and chemists in CBB at all levels of seniority in the future.

6. In view of the significant level of upcoming turnover, the panel endorsed Jan Ellenberg's plan to search widely, but advised to place emphasis on the possibility of synergy between new recruits and existing activities in order to avoid too much dispersal of activity. The panel also noted that there are currently no female faculty members in CBB and, although they were satisfied that this did not reflect an inherent bias, they recommended that CBB pay particular attention to attracting excellent female group or team leaders in future. They also advised Jan Ellenberg to ensure that the level of mentoring of new faculty recruits in CBB should be brought into line with that seen in other EMBL Units.

7. Although not explicitly part of the review, as these activities will be reviewed in 2010, the panel praised Pepperkok and Antony for their excellent performances in running the light microscopy and electron microscopy core facilities. The panel noted that these are both critical to the future success of not only CBB but also many other parts of EMBL. They warned of the need for ongoing investment in these core facilities to ensure that they remain state of the art, and commented positively on the plans to reorganise the light microscopy facility, that is currently scattered throughout much of EMBL Heidelberg, into one location. They also recommended that this location should be adjacent to the computational server cluster in order to avoid potential data transmission problems.

Iain W. Mattaj
Director General
28 May 2009

Scientific Report



The grand scheme of things

"The view of Jerusalem is the history of the world; it is more, it is the history of Earth and heaven," wrote Benjamin Disraeli. It's not hard to see what he means. As well as being the focus of three major world religions, at nearly 3000 years old, Jerusalem is one of the world's oldest continuously inhabited cities. The modern-day hustle and bustle of its winding streets echoes a daily rhythm of city life that stretches back into antiquity.

This diurnal pattern is both unique and universal: the individual dwellers and their lives are peculiar to Jerusalem, but the basic principles of how the city is built, organised and run is something it shares with other cities around the world. It may sound odd, but cities such as Jersualem also have a lot in common with how our cells and organs – and those of other creatures – are organised. For example, the cells in our bodies all have their own specialised functions, rather like professions, to perform. Yet even while they focus on their own particular niches, they must all co-operate to make the body as a whole function normally. Even within cells, molecules must follow certain rules to keep everything working properly.

So the organisational systems of biology are not so far removed from those of a busy town. Molecules need to travel to places, information must be exchanged, damage has to be detected and repaired, building plans drawn up and carried out, and law and order kept. Only when all these elements are in place can the cellular or molecular denizens of an organism come together to create a greater whole.

Of course, all cities must start somewhere. They originally began haphazardly, acquiring streets, buildings and districts according to the whims of the populace, with each new development building on what went before. By contrast, modern cities are planned in meticulous detail. The organisation of cells and bodies has a similar history – evolution has blindly produced ways of planning and building living systems, as biologists are now discovering. Teams at EMBL Heidelberg, for example, have dissected the biology of one of the simplest free-living bacteria to understand how its constituent parts work together and to pinpoint the bare essentials a cell needs to survive. At EMBL Monterotondo, scientists have looked at how the processes that build an embryo's heart could help adults to recover from heart attacks.

Other EMBL researchers have been studying how cells and embryos deploy the building plans laid down in DNA, to both understand these processes and look for ways of exploiting them in medicine. What's more, EMBL is finding ways to make these data accessible to researchers all over the world, welcoming their input, just as any truly cosmopolitan city should.

Vital ingredients

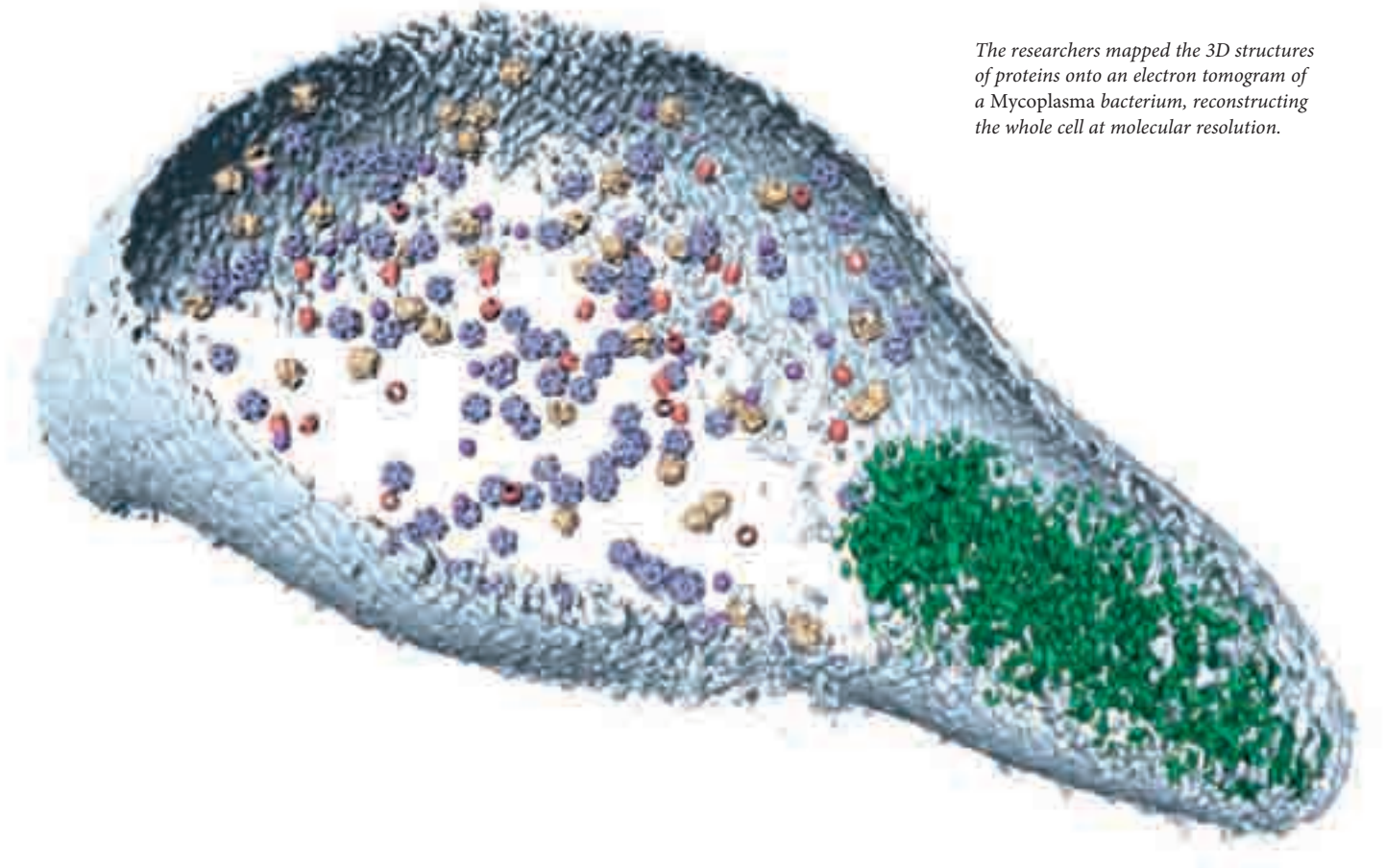
Sebastian Kühner, Anne-Claude Gavin, Peer Bork and Vera van Noort

What are the bare essentials of life? Which components and processes can an organism simply not live without? How are they organised in space and time in order to make a random collection of molecules come alive? These are mighty big questions that go well beyond what a single biologist can handle. But the answers are a lot more within reach since a group of EMBL scientists joined forces and pooled their skills and expertise to tackle the problem together.

It all started five years ago when Peer Bork and Luis Serrano, who were then joint coordinators of the Structural and Computational Biology Unit in Heidelberg, got the Unit's group leaders together to brainstorm about a common project that would put the complementary technologies and know-how of their groups to good use. "Within the Unit we have a unique combination of structural and computational methods that span a broad spectrum of scales. We can study everything from tiny individual molecules to the overall arrangement of the inner workings of a cell," says Peer. The idea: to combine these methods in an interdisciplinary approach to produce the first blueprint of a minimal cell, a cell that is stripped down to the absolute essentials.

The cell of choice was quickly found. It had to be small and simple to make a global analysis feasible, yet complex enough to be self-sufficient and viable on its own. *Mycoplasma pneumoniae*, a small bacterium that causes atypical pneumonia in humans and accounts for between 15-20% of pneumonia patients, is one of the smallest existing prokaryotes that has retained the ability to self-replicate. Unlike viruses and other pathogens, it does not depend on the cellular machinery of a host to survive and multiply. With a mere 689 genes, *Mycoplasma* lends itself well to genome-wide analyses and its thin width makes it amenable to whole-cell imaging at high resolution. In brief, it is the perfect organism to study molecular organisation at all levels.

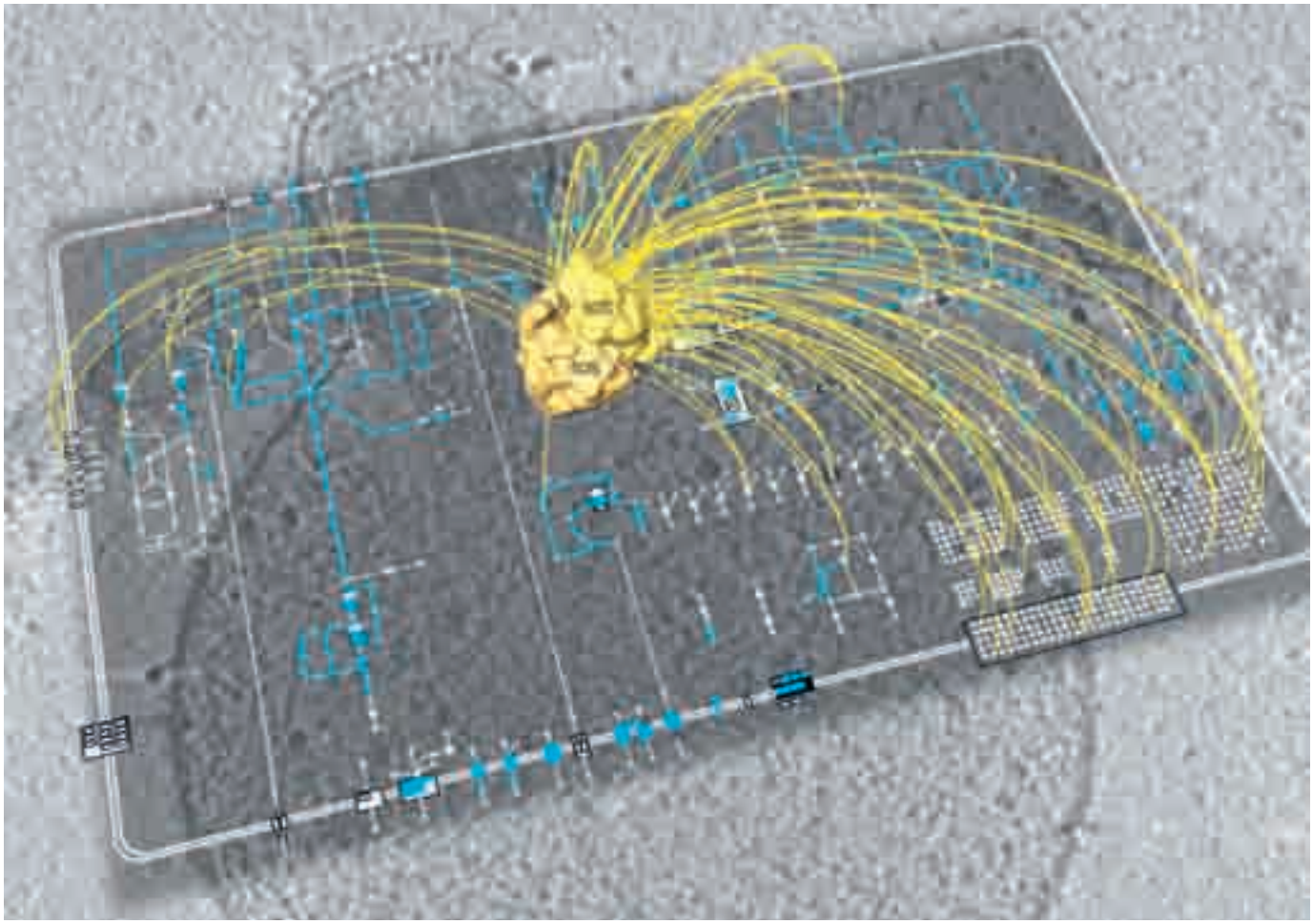
To generate a complete picture of this bacterium, the EMBL groups thoroughly investigated it along three dimensions: its proteome, its transcriptome and its metabolome. Sound Greek to you? Funnily enough, it is not. Whenever biologists study all representatives of a certain class of molecules, they add the suffix -ome to its name. But scholars insist there is no '-ome' root of Greek origin that refers to wholeness or completion: '-ome' seems to be an original invention of the molecular biology community. When scientists talk about a genome they refer to the complete genetic contents of a cell



The researchers mapped the 3D structures of proteins onto an electron tomogram of a Mycoplasma *bacterium, reconstructing the whole cell at molecular resolution.*

GRABBING THE LIMELIGHT

This in-depth study of *Mycoplasma pneumoniae* caught the attention of the top scientific journals and the media in general, throughout the world. *Nature News* and *Scientific American* focused on how this single-celled bacterium puts its genome to many uses, a sentiment the *Science signalling* editorial summed up in two words: "Simply *Mycoplasma*". The work merited a review by Craig Venter in *Molecular Systems Biology*, and was divulged to French, German and Dutch science enthusiasts by *Sciences & Avenir*, *Spektrum der Wissenschaft* and *Explore*, respectively. From *The New York Times*, *El País* and the *Frankfurter Allgemeine Zeitung* to websites like *Genome Web* and *Galileo*, everyone was keen to report on the unexpected complexity of this blueprint for life.



An image from an electron tomography scan (grey), overlaid with a scheme of the bacterium's metabolism, which shows interactions between proteins as blue lines and between proteins and the ribosome as yellow lines. The ribosome, also labelled yellow, was imaged using electron microscopy.

and the proteome is nothing more than all of its proteins. The RNAs make up a cell's transcriptome because they are produced through a process called transcription. Similarly, metabolome refers to the collection of small-molecule metabolites found in a cell or organism. All these -omes are very dynamic and often change from one second to the next, which is why they cannot be fully characterised using only one analytical method. This is where even more molecular biology jargon comes in: proteomics, transcriptomics and metabolomics are the newly created disciplines that are concerned with understanding the respective -omes, by applying a range of genetic, biochemical and computational techniques.

The proteome

Characterising *Mycoplasma*'s proteome alone demanded a combination of around ten different techniques. So, the teams of Anne-Claude Gavin, Rob Russell, Peer Bork, Bettina Boettcher and Achilleas Frangakis, with the support of the EMBL Core Facilities, pooled their skills to form a proteomics task force.

Anne-Claude's lab was first in line. Using a method called Tandem Affinity Purification, PhD student Sebastian Kühner first fished all soluble protein complexes from the *Mycoplasma* cell and then determined their components by mass spectrometry, a technique that identifies proteins based on their weight and charge. With the help of skilled bioinformatician Vera Van Noort and colleagues from Peer's and Rob's groups, the mass spectrometry data were then integrated with information from STRING, a database that stores information on protein interactions, to reconstruct which proteins are part of which complex.

They identified around 200 protein complexes, half of which had not been found in previous studies of the bacterium. This is the first time that an exhaustive proteome analysis has been carried out in a prokaryote. Prokaryotes are simple, mostly single-celled organisms, such as bacteria, that lack a nucleus and other membrane-bound organelles. Eukaryotes, on the other hand, are defined by cells organised into complex structures surrounded by a membrane and comprise all animals, plants and fungi. A few years earlier, Anne-Claude, Rob and

Peer had applied the same proteomics approach to a eukaryote, baker's yeast.

"We find many similarities between the two organisms," says Anne-Claude. Many of the proteins and complexes they found in yeast also exist in *Mycoplasma*, suggesting that they have essential functions that have been conserved during evolution. "What is even more interesting is that key principles of protein organisation also seem to be conserved." Just like in yeast, *Mycoplasma* proteins rarely act alone. Most cellular processes are carried out by molecular machines comprising several different proteins. These machines are often organised into higher order assemblies, like in a factory where machines completing different steps of a task are located side-by-side in pipelines. In addition, both yeast and *Mycoplasma* proteins tend to be promiscuous: they interact with different binding partners and take part in more than one complex. This suggests a great deal of multifunctionality among proteins and provides a mechanism by which cellular processes can be coupled in space and time. For example, the team found physical links between the bacterium's RNA polymerase, the enzyme that transcribes the DNA code into messenger RNAs, and the ribosome, the machine responsible for translating messenger RNA into protein.

To gain a better understanding of what the protein complexes look like and of their intricate spatial organisation inside the cell, Bettina and her team studied the purified complexes with an electron microscope. Electron microscopes have much greater resolving power than light microscopes and can magnify a specimen up to two million times. Although this is just about enough to visualise the overall shape of protein complexes, it does not tell you anything about how its components fit together and interact. Luckily, bioinformatics had a solution at hand. The atomic structure of most *Mycoplasma* proteins had been determined and deposited in databases. Using advanced computational methods Rob's team fitted the structures of 484 *Mycoplasma* proteins into the overall shapes Bettina had produced for the complexes. Next, the largest and most readily identifiable of these complexes, including ribosomes and the RNA polymerase, were mapped onto an electron tomogram of a *Mycoplasma* cell that was generated by Achilleas' group. Electron tomograms are three-dimensional reconstructions of parts of cells, or even whole cells in the case of small prokaryotes such as *Mycoplasma*. They have a sufficiently high resolution to see individual protein complexes in their natural context in the cell. "Electron tomography bridges the gap between high-resolution structural approaches and light microscopy," says Achilleas. "It allows us to explore the whole inner space of small prokaryotic cells and localise atomic structures of molecules inside those cells, in their native positions and environment."

Never before has there been such a complete and detailed, three-dimensional model of a prokaryote's proteome, making *Mycoplasma pneumoniae* one of the most structurally known organisms to date. But to obtain a full picture of the bacterium and understand all processes and interactions that happen in a minimal cell, the scientists needed to look at more than the cell's protein content. Next up: the transcriptome!

The transcriptome

Mycoplasma's genome contains less than 700 protein-coding genes, only 44 RNAs and uses no more than eight regulatory proteins to control gene expression. This simplicity allowed members of Peer's lab at EMBL and Luis's group, now based at the Centre for Genomic Regulation in Barcelona, to study all the parts of the genome that are transcribed into RNA. Supported by the staff at EMBL's Genomics Core Facility, postdoctoral fellows Marc Güell in Barcelona and Vera van Noort in Heidelberg analysed all RNA transcripts found in a *Mycoplasma* cell under different conditions. The result: *Mycoplasma*'s genome organisation and its transcription are not as simple as thought. Not only did the scientists identify 67 completely new transcripts, but they also found that the genome is transcribed differently depending on the environment of the bacterium.

Normally the genomes of bacteria are thought to be quite straightforward: they are divided into operons, which are units of transcription comprising one or more genes that are transcribed from the same promoter (transcription start site) and controlled by the same regulatory element (operator). Under normal conditions the researchers counted 340 such operons in *Mycoplasma*, but altering different factors such as the food source, temperature or pH changed the number of operons and the resulting transcripts. Some of these changes are because of internal promoters, found in the middle of known operons. These internal promoters are activated by environmental factors and break down known operons into smaller transcription units, producing so-called alternative transcripts of a DNA region. This dynamic structure allows *Mycoplasma* to react flexibly and rapidly to changing environments and to produce certain molecules only when required.

But alternative transcripts are not the only novelty revealed by the analysis. It also brought to light a second type of transcript that is equally surprising – so-called antisense transcripts. Antisense transcripts are generated when a protein-coding stretch of DNA is transcribed back-to-front rather than in the conventional direction. The role of such antisense RNAs is not clear. They do not encode proteins but they could help to regulate gene expression in various ways. Until quite recently, scientists did not even know they existed. It

was assumed that the transcription machinery could only move along one direction of a DNA strand, and thus at first the occasional 'backwards' RNAs that scientists stumbled across were dismissed as noise of the transcription process. New research shows, however, that bidirectional transcription leading to antisense RNA is quite common in many different organisms, especially eukaryotes. EMBL group leaders Lars Steinmetz and Wolfgang Huber, for example, found that bidirectional transcription is probably the rule rather than the exception in yeast and other studies have come to a similar conclusion for humans. So, as was found for its proteome, *Mycoplasma*'s transcriptome also shares many features with eukaryotic cells.

The metabolome

What is the only thing missing at this stage to complete the full picture of a minimal cell? An accurate account of its metabolism. As a first step Luis and his team in Barcelona developed a minimal, defined medium that supports the growth of *Mycoplasma*. "This was extremely challenging, because the human body, *Mycoplasma*'s natural environment, provides nutrients as higher order molecules. So, the bacterium has lost the ability to live on simple building blocks like amino acids and needs peptides to survive. This made it very difficult to find the simplest ingredients that would still sustain growth," explains Luis. Yet such a defined minimal medium is crucial for obtaining precise, quantitative measurements of an organism's metabolites and its exchanges with the environment. Roughly 1300 experiments later, the scientists arrived at a mixture of 19 essential nutrients that the bacterium would happily grow in and a comprehensive map of *Mycoplasma*'s metabolism, consisting of 180 reactions carried out by 129 enzymes.

The map identified 78 genes that are essential for the bacterium's metabolism and a bioinformatics scrutiny by Peer's group showed that the fraction of multifunctional enzymes is much higher than in more complex bacteria, even though the pathways are more linear, indicating a more streamlined metabolism. It also uncovered that, unlike other bacteria, *Mycoplasma*'s metabolism is not geared towards multiplying as fast as possible. With a duplication time of at least eight hours, the bacterium reproduces relatively slowly, which is a result of its pathogenic lifestyle and adaptation to its host. What it shares with larger bacteria such as *Escherichia coli*, however, is a great deal of flexibility and adaptability to changing environments. Surprisingly for such a simple bacterium with only eight regulatory proteins, *Mycoplasma* orchestrates complex changes in gene expression in response to different environmental stress factors. It seems to do so by assigning genes to one of four categories: catabolism, cell defence, biosynthesis and signal transduction. The genes in each of these categories are regulated together as a group and jointly bring about specific responses to different stress situations, such as sugar or amino-acid starvation or pH changes.

Integrating all the information collected on the three -omes – proteome, transcriptome and metabolome – the EMBL researchers have produced what is currently the most comprehensive picture of an entire organism. They have provided a complete parts list of all cellular components, an overview of their organisation and quantitative accounts of their dynamic interactions. "This is a unique resource for systems biology. We will make all our data available to the scientific community, who can use it to build and test mathematical models of individual processes," says Peer. He expects this approach to provide insights of unprecedented accuracy into whole-organism biology.

The overwhelming conclusion so far: in all aspects studied the bacterium turned out to be more complex and dynamic than previously assumed. The temporal and spatial organisation of its molecular machinery, the transcription process and its metabolism share many features with larger, more complex prokaryotes and even with eukaryotic cells. This complexity is unexpected considering *Mycoplasma*'s small genome and alleged simplicity, and it makes the bacterium a powerful model organism. It is simple enough for large-scale analysis, yet sufficiently complex for its cellular organisation to be representative of other prokaryotes, eukaryotes and even higher, multicellular organisms.

At the same time *Mycoplasma* is also representative of something else: minimal life. Owing to its unique evolution as a pathogen, the bacterium has been stripped of everything that is not absolutely essential. It has become adapted for life in the human body – a very special and rich biological niche that made many genes redundant so that they were lost over time. What is left today is the minimal structure and machinery required for survival, and thanks to the EMBL taskforce this has been captured in a blueprint. This blueprint and the universal organisational principles it highlights are probably the closest scientists have ever come to identifying the essentials of life. "The key lies in those features that *Mycoplasma* shares with all other organisms. These are the things that not even the simplest organism can do without and that have remained untouched by millions of years of evolution," Peer concludes.

Kühner S, et al (2009) Proteome Organization in a Genome-Reduced Bacterium. *Science* **326**

Güel M, et al (2009) Transcriptome Complexity in a Genome-Reduced Bacterium. *Science* **326**

Yus E, et al (2009) Impact of Genome Reduction on Bacterial Metabolism and Its Regulation. *Science* **326**

Misha Kapushesky and Alvis Brazma

Mapping the future

The huge wealth of data generated by the past decade's genome projects is now much easier for biologists to study thanks to a new online tool created at the EBI. Group leader Alvis Brazma and his team have made it possible for biologists to explore information about a particular gene's activity in much greater depth than ever before, without the need for specialist training in bioinformatics.

The new tool, the Gene Expression Atlas, was launched in March 2009, and is freely available to all. Biologists can search data for information about how a gene's activity, or expression, varies between different tissues, under different biological conditions (such as drug treatment), in disease and much more. Geneticists hunting down disease genes will be able to investigate possible leads much more quickly and extensively. "It will help biologists develop and test new hypotheses," says Alvis.

The seeds of the Atlas were sown when Alvis joined the EBI more than a decade ago. At the time, data from microarrays – experiments that allow scientists to study the activity of thousands of genes at once – were accumulating and Alvis wanted to find ways to get the most out of these results.

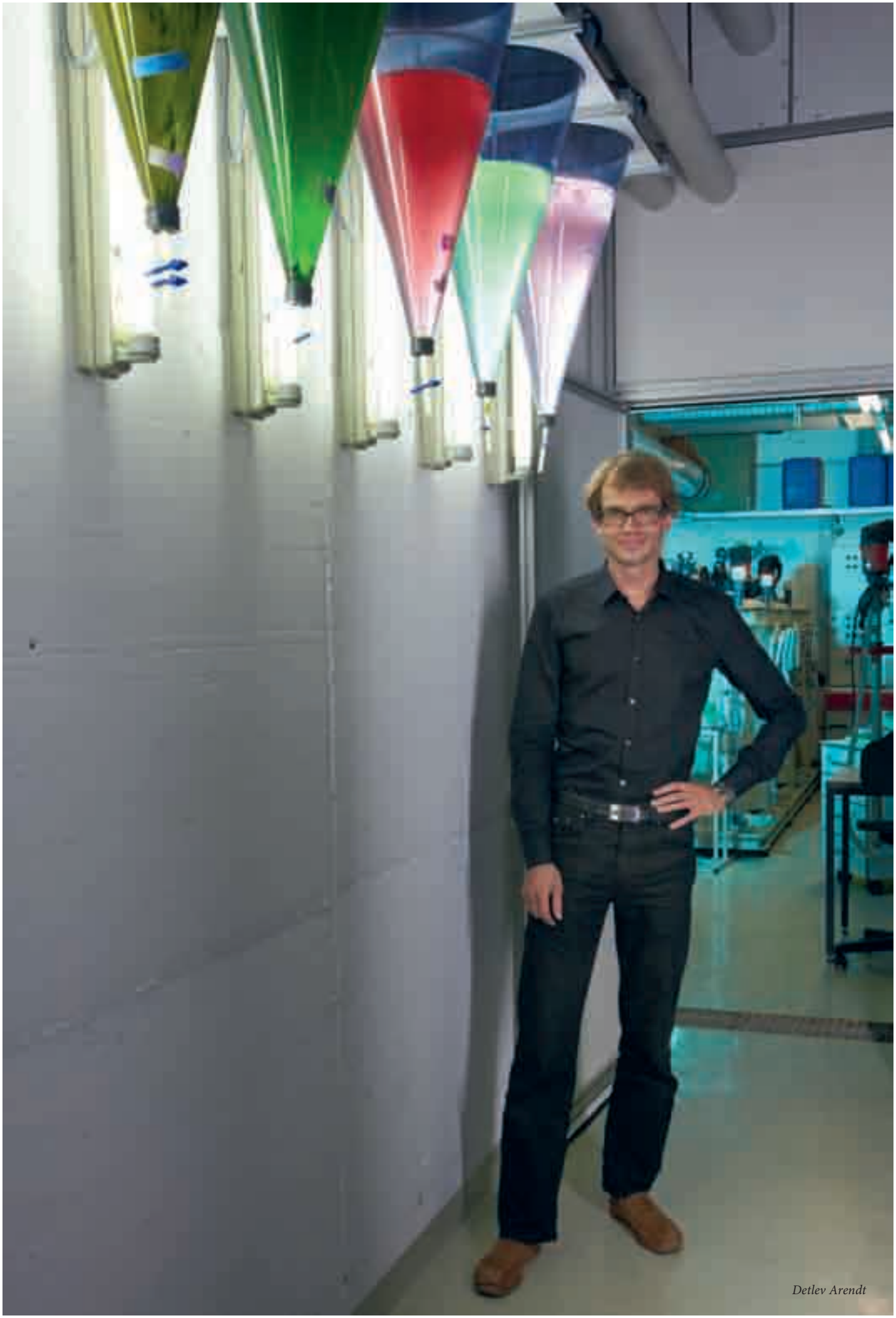
He and his colleagues developed a protocol that allows researchers to label the pieces of information in their microarray data in a uniform way. This led to an EBI database called ArrayExpress, in which these data could be deposited. Thanks to the labelling system, researchers could mine the data for new insights into the biology of genes and disease. "The plan was to summarise these data in a way that meant we could ask biological questions," says Alvis. "It was never intended just to be an archive."

The trouble was, this mining was far from straightforward, and so was really only an option for computational data-analysis experts. Alvis and his team decided to create something that was much more accessible. The ArrayExpress curators revisited the data and created new ways of enriching and labelling it. Led by Misha Kapushesky, a project leader within Alvis's group, researchers developed and implemented statistical software that could exploit this new labelling to analyse possible relationships between all the items of data more effectively than before.

Biologists can now analyse these data simply by typing a few details – the name of a gene, a species or a disease – into a box on the EBI website. The Atlas displays detailed results, ranked in order of relevance, and users can click through to relevant information on other EBI databases.

Thousands of academic researchers are already using the Atlas and the EBI's industry partners are also interested in taking the technology further, says Alvis. Currently, the Atlas contains about 15% of the data from ArrayExpress, and describes more than 200 000 genes in more than 6000 different biological states. Alvis' team plans to boost this to 50% in the near future, and has already started work on a similar Atlas for protein data.

Kapushesky M, Emam I, Holloway E, Kurnosov P, Zorin A, Malone J, Rustici G, Williams E, Parkinson H, Brazma A (2010) Gene Expression Atlas at the European Bioinformatics Institute. *Nucleic Acids Research* **38**, Database issue



Detlev Arendt

Ancient tweaking

Twenty years ago, scientists knew nothing of the scraps of RNA that are now known to influence just about every process in our bodies. Back then, the textbooks were simpler: genes code for proteins via the intermediate of RNA, and proteins called transcription factors regulate other proteins. This recipe was so entrenched in the basic orthodoxy of molecular biology that it was even given the name 'the central dogma' by the co-discoverer of DNA, Francis Crick.

Scientists now know, however, that this classic view of protein regulation is far too blunderingly inefficient for evolution to settle for. At some point hundreds of millions of years ago, the generation of a small stretch of RNA that could tweak this process gave an individual the edge over everyone else. And so regulatory RNA was born. These scraps of RNA – on average only 22 nucleotides long and now dubbed microRNAs, or miRNAs for short – bind to some messenger RNAs and label them for inactivation or destruction.

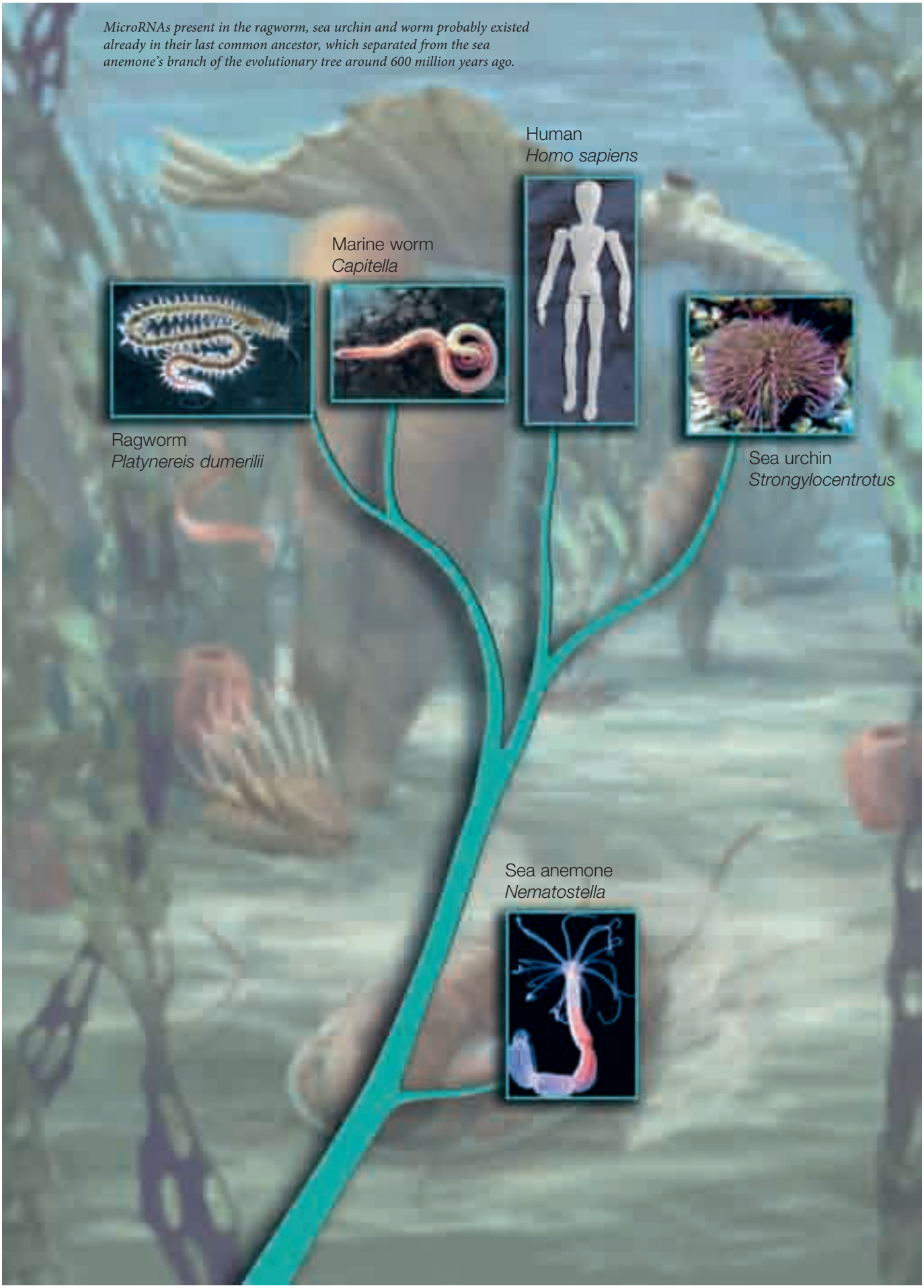
So far thousands of miRNAs have been identified in animals. These superintendents of protein regulation are involved in the earliest stages of an animal's development, determining which cell types grow where and when, and how these cells differentiate into the different body parts. However, since the discovery of miRNAs, many scientists have wondered whether the same miRNAs govern specific tissues in different animals. Knowing this would not only give clues to the age of these different miRNAs, but also to the age of the cells in which they are found.

Developmental biologist Detlev Arendt from EMBL Heidelberg, whose work has recently traced the evolutionary origins of the brain and other organs, wondered whether he could study the oldest known animal miRNAs in a group of animals that at least look old themselves. In doing so, he hoped to investigate whether miRNA expression is conserved across the animal kingdom.

Studying ancient-looking animals is crucial for this process because "all animals evolve, but the speed with which they change differs," explains Detlev. "Animals living near the coastline today exist in a similar environment to that which their ancestors thrived in, and so some of these animals haven't been forced to change their body plans, or the genes that control these plans, because they are already well-adapted to this ancient environment." By studying these animals, biologists can glean clues to ancient morphologies.

Detlev and Fay Christodoulou, who was a PhD student in Detlev's group, used one such 'living fossil', the marine ragworm *Platynereis dumerilii*, which is thought to have changed little over the past 600 million years. In collaboration with an American team from Cold Spring Harbor Laboratory who sequenced the ragworm's miRNAs, and Peer Bork's group at EMBL who contributed to the bioinformatic analysis, they probed for 34 ancient miRNAs in the bodies of young worms.

By tracing a blue dye that becomes trapped when specific probes bind to the miRNA, Fay was able to see where the miRNAs were expressed in the worm embryos. "Initially the high temperatures required for the probes to bind



MicroRNAs present in the ragworm, sea urchin and worm probably existed already in their last common ancestor, which separated from the sea anemone's branch of the evolutionary tree around 600 million years ago.

destroyed the embryo's intricate body structures, but then I realised that I could reduce the temperature and still see binding," notes Fay.

After four months of developing this approach, and over a year of analysing the tissue samples, Fay showed that many miRNAs are highly specific for certain tissues and cell types in the worms. She then looked at the expression of these same miRNAs in three other marine species – a sea anemone, a worm and a sea urchin. Fay explains that sea anemones are radially symmetrical, whereas worms, like humans, are bilaterally symmetrical because they have a back and a belly, and their right side is the mirror image of their left. Sea urchins are more complex but they are still thought to stem from animals with bilateral symmetry. By comparing miRNA expression in these animals with that in the ragworm, the researchers hoped to resolve in which tissue particular miRNAs were first active when they evolved more than half a billion years ago.

For many of the miRNAs included in the study, Fay saw similar patterns of miRNA tissue specificity in the other marine species as she had seen in *Platynereis*. This suggests that these regulatory RNAs fulfill a similar role today as they did hundreds of millions of years ago in the common ancestor of these animals. "It seems that the evolution of miRNA and tissue identities are closely coupled," remarks Detlev.

Overall, miRNA expression patterns in the musculature, gut and nervous system are the most conserved, the researchers discovered. "In all bilaterians we found the same set of three miRNAs in cells that form the hair-like cilia that propel the animals forward," says Fay. "And we found the oldest known family of animal miRNAs – miR-100, miR-125 and let-7 – encircling the gut of each of these animals, where they are thought to have a role in developmental timing," she notes.

Detlev adds that "let-7 expression is perfectly timed with the transition of the worm from its immature free-swimming larval stage into its sedentary adult stage." Let-7 could therefore be important for regulating the timing of transitions towards the later stages of development, a suggestion that is supported by other studies of let-7's role in the development of molluscs, flies and zebrafish. "This is an important find as it demonstrates how miRNAs, which are responsible for regulating the developmental timing of one tissue, can keep their ancient roles even when body plans get more complex," remarks Detlev.

This gives clues to how tissues specialise and build complex structures within animals' bodies that help them to adapt to new environments, especially if they differ considerably from those of their ancestors. "What is interesting is that animals on different branches of the evolutionary tree adapt in very different ways, but keep an ancient set of tissues and cell types that they re-shape according to the needs of their new environment. In some animals like insects these old components are sometimes modified beyond recognition, while in others, like the ancient-looking marine species, they are quite easy to distinguish," explains Detlev.

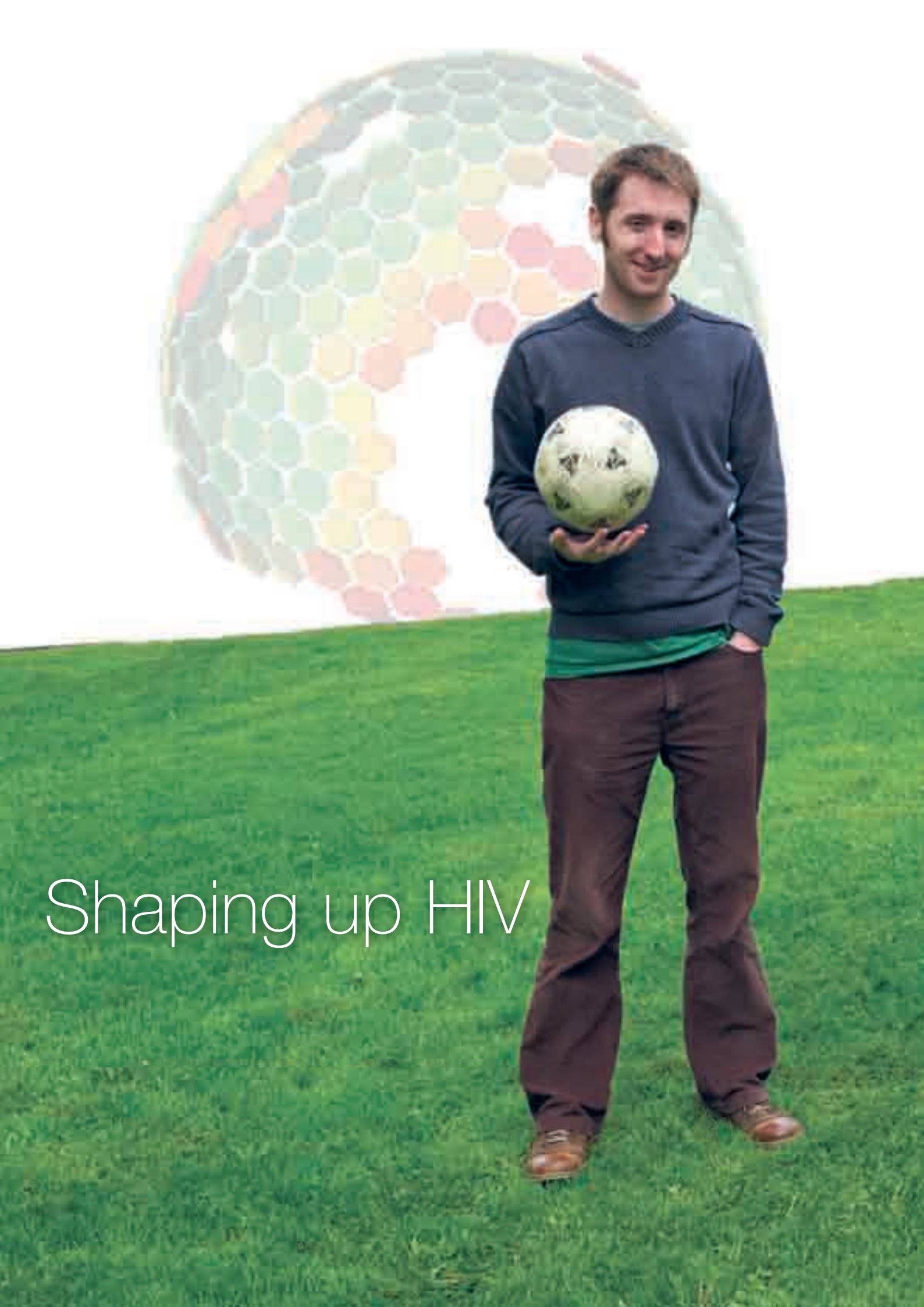
In addition, the pattern of miRNAs in the brain allowed the researchers to deduce that worms and humans share some miRNAs that are specific to the ancient parts of the central nervous system that secrete hormones into the blood. This offers a solution to the ongoing dispute over whether the last common ancestor of all bilaterians had a brain at all.

Detlev explains that the position of the central nervous system differs between invertebrates and vertebrates. "The brain and nerve cord are on the underside of invertebrates, while in vertebrates they are found towards the back," he explains. In the past, these observations convinced many scientists that the common ancestor of vertebrates and invertebrates was brainless, and that the nervous systems had evolved independently in both lineages. "But if it had, you would not expect the same miRNA in identical regions within the brains and nerve cords of descendents from different branches," explains Detlev. And this is exactly what the EMBL researchers found.

This example demonstrates how these findings are helping to establish miRNAs as an important new tool for reconstructing ancient animal body plans at important evolutionary junctures.

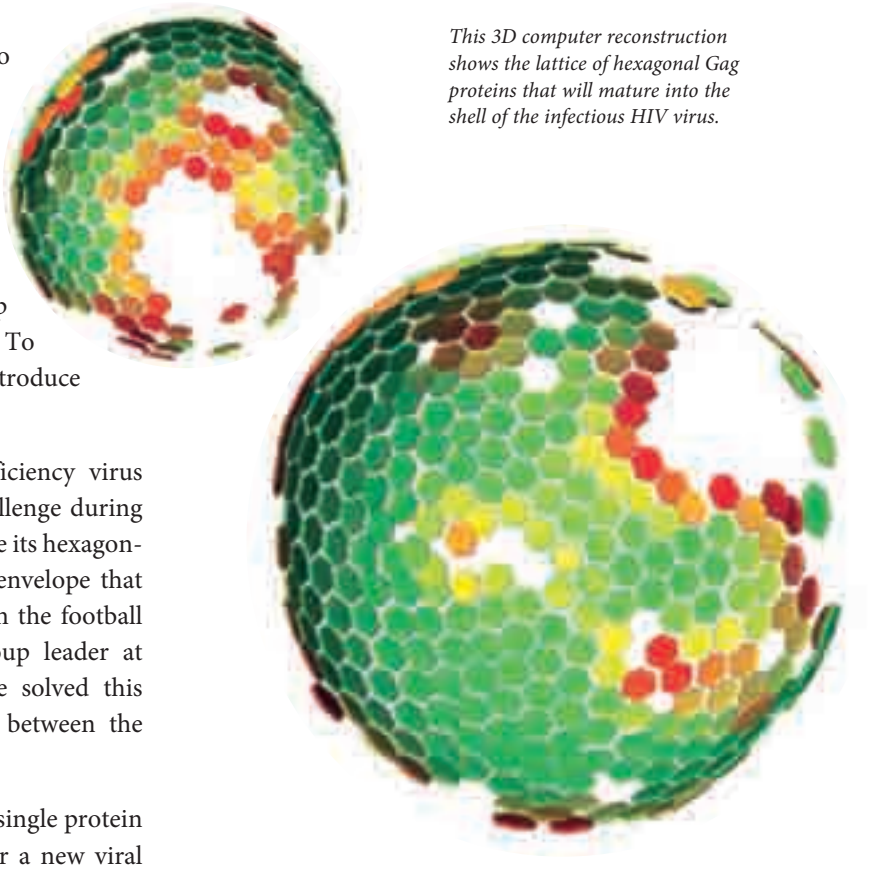
In future work Detlev's group would like to investigate the regulatory role of each of the conserved miRNAs by interfering with their expression. He explains that when you consider that within the human genome there are almost 700 miRNAs, which affect the expression of around 30% of the human genome, it is clear that we need to know what they are doing. If scientists can learn to block these tiny RNAs, or mimic their effects, they could use them to develop new treatments for cancer, help repair damaged organs and slow the process of ageing. But for now, the contribution of a handful of marine-dwelling creatures to scientists' understanding of how miRNA regulate gene expression should at the very least instil intrigue into the already colourful world of rock pools.

Christodoulou F, Raible F, Tomer R, Simakov, O, Trachana K, Klaus S, Snyman H, Hannon GJ, Bork P, Arendt D (2010) Ancient animal microRNAs and the evolution of tissue identity. *Nature* **463**: 1084-1088



Shaping up HIV

John Briggs, James Riches and Alex de Marco



This 3D computer reconstruction shows the lattice of hexagonal Gag proteins that will mature into the shell of the infectious HIV virus.

In the 1960s, a Danish company, seeking to improve on the traditional football made from the bladder and stomach of animals, invented the modern football. The designers realised that to form a perfect ball they needed to combine 20 leather hexagons with 12 pentagons, and in so doing demonstrated one of the basic laws of shape – that you cannot wrap a sheet of six-sided hexagons around a sphere. To induce the sheet to bend, the company had to introduce five-sided pentagons alongside the hexagons.

On the micro scale, the human immunodeficiency virus (HIV), which causes AIDS, faces a similar challenge during the assembly of new viral particles: how to coerce its hexagon-shaped building blocks to form the spherical envelope that surrounds its viral innards. Lifting a page from the football manual, structural biologist John Briggs, group leader at EMBL Heidelberg, wondered if HIV likewise solved this shape conundrum by introducing pentagons between the hexagons.

At the early stages of assembly, HIV relies on a single protein to bring together everything that is needed for a new viral particle. This protein, called gag, is shaped like a string of diamond-shaped beads and is produced when the virus is copied by the host cell's machinery. These copies are exported from the cell's nucleus, and are turned into long chains of amino acids that fold into the gag protein. Hundreds of these gag proteins aggregate just beneath the cell's membrane where they team up newly copied viral DNA with proteins and enzymes that are vital for the virus to infect new cells. While everything is being corralled into position, the gag proteins bundle together – six at a time – to form hexagon-shaped plates.

It is the arrangement of these plates that John and his team wanted to inspect. But to do this, they needed to halt the virus' development while preserving its structure very close to its natural state. In collaboration with Hans-Georg Kräusslich's team at the University Clinic Heidelberg, John's group took advantage of a technique called cryo-electron tomography – cryo-ET for short – which combines flash-freezing with electron beam scanning to create three-dimensional images of the internal structures of the virus.

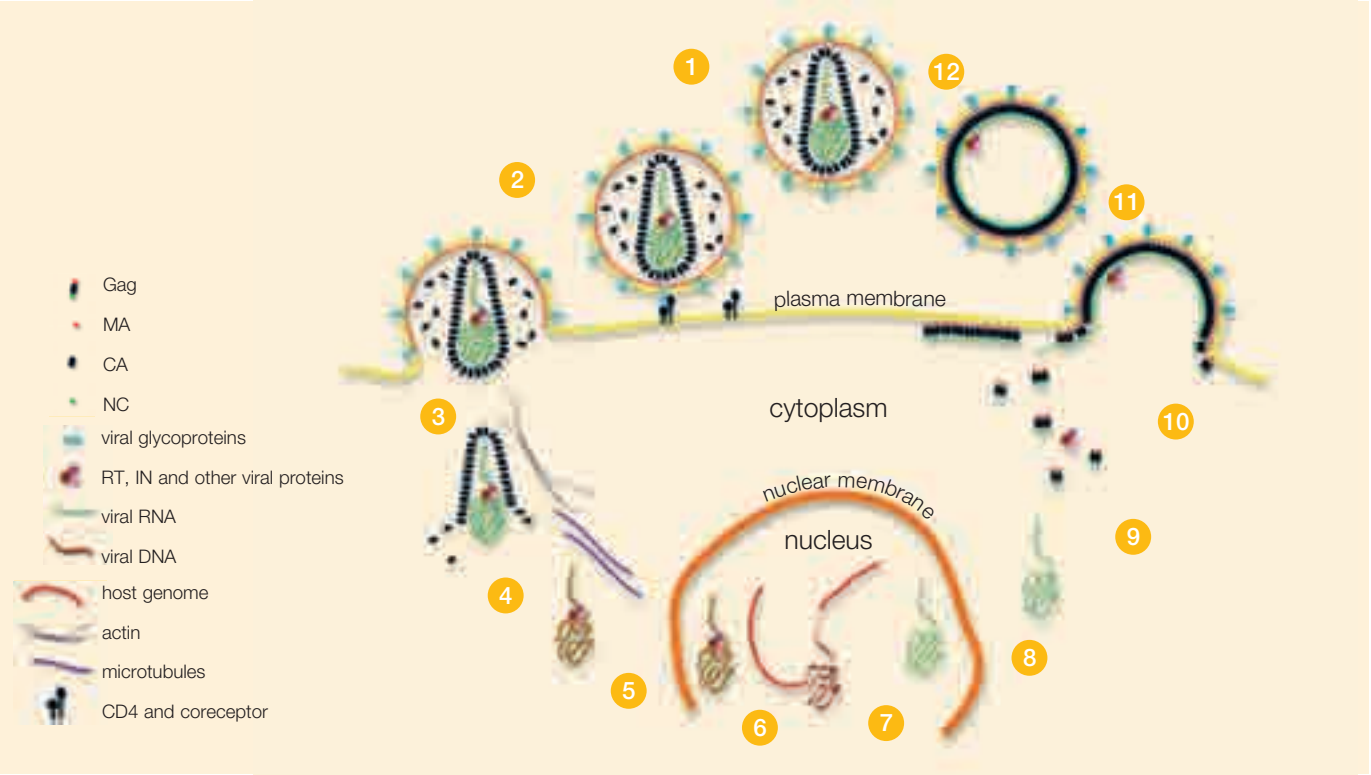
"What is nice about this approach is that it allows us to look inside samples of biological material without breaking them apart or staining them, as we had to do in the past, which often led to unexplainable artefacts," says John.

This work requires "swift, agile hands" explains James Riches, an engineer in the Briggs group, and whose fiddly job it is to place the 3mm copper disks containing frozen viral particles on to a platform that automatically moves through a range of tilt angles while beams of electrons illuminate the sample. At each angle, a thin region of the sample is exposed to the negatively charged particles. When the electron beam collides with this solid structure, it casts a shadow onto a camera. These two-dimensional projections are then fed to a computer, which assembles the data into a composite three-dimensional image.

Inspecting these images, the researchers found that the hexagon-shaped plates were arranged into a lattice, but that this lattice was incomplete; it had gaps within it – areas where the gag proteins were missing altogether. "We were surprised that these gaps were not pentagon-shaped, but were instead irregularly shaped, variable in size and that there was always more than one of them," says John.

He thinks that these gaps prevent the hexagons from becoming too closely packed together when the lattice curves to form the viral envelope. He goes on to explain that during the formation of this envelope, nicks in the gag protein transform the immature virus into its infectious form. An enzyme called protease performs this function, and preventing its cutting action has been the target of some of the most successful HIV drugs of the past 10 years.

Knowing more about this maturation process could help scientists develop new ways to stop the virus from maturing and spreading to other cells. With this in mind, John's group got



HIV's lifecycle. Once the virus has infected a host cell and harnessed its machinery to produce copies of itself (1-9), the structural protein Gag is essential for directing the assembly of the viral coat (10). Once the virus has left the host cell, Gag must be cleaved so that it can mature (12) and become capable of fusing with a new susceptible cell.

in closer to the HIV virus, to have a more detailed look at how the gag proteins bundle together to form each hexagon-shaped plate.

Understanding the role of the gag protein requires not only knowing how the envelope is laid down, but also identifying the structure of the individual gag proteins within the envelope. "You can solve the structure of the brick, but if you want to know the structure of the house then you need to know how the bricks pack together," reckons John.

To zoom in on the three-dimensional image of the virus' structure, his team sampled small segments of the viral envelope and assembled them one on top of the other. This was essential because each close-up is missing some information; at this focus, parts of the image will be fuzzy or missing. John explains that a small, but different, part of each image is concealed and so by overlaying the images you should get to see the whole thing.

"It's a bit like if you were to take a picture of a person standing in a field in a snow storm," explains John. "Now if you take one picture, some of the person will be obscured by the falling snow, but if you took 50 pictures, each with a different pattern of snowflakes, and average all those images together then you would probably get a clear image of the person."

Up this close, the researchers discovered that each hexagon-shaped plate consists of six gag proteins arranged with their flat side facing inwards towards a central hole. And they recognised that this configuration is very different from how the gag proteins arrange themselves in mature viruses, in which they turn their flat side outwards.

At some point during maturation the proteins must rotate, reckons John, whose team is now working on producing an even higher resolution structure of the gag proteins to gain a more detailed understanding of the virus' assembly and maturation.

"We know that the cutting of the gag protein in five different positions is essential to this maturation process," explains John. "But what happens if you only allow cuts one, two and three but not four and five, what kinds of structures do you get out?" To investigate this, Alex de Marco, a PhD student in John's group, is studying mutant lines in which some of the cutting by the protease is prevented, and the team will then examine the changes they see in the viral structures. Imaging these structures could give scientists clues on how to throw a spanner in the viral works, and stop the formation of new viral particles and their spread to other cells. New ways to tackle the spread of HIV are needed because the virus, being small and fast to mutate, always has the upper hand in this match.

Briggs J, Riches J, Glass B, Bartonova V, Zanettia G, Kräusslich H-G (2009) Structure and assembly of immature HIV. *PNAS* **106:** 11090–11095

An accurate forecast of regulatory activity

As predicted, the same CRM is active (red/pink) early in Drosophila *development (left) in the tissue that will give rise to all muscle types, but is active only in the embryo's body wall muscle (blue) at a later stage (middle), when a different CRM drives development in the gut muscle (green, right).*

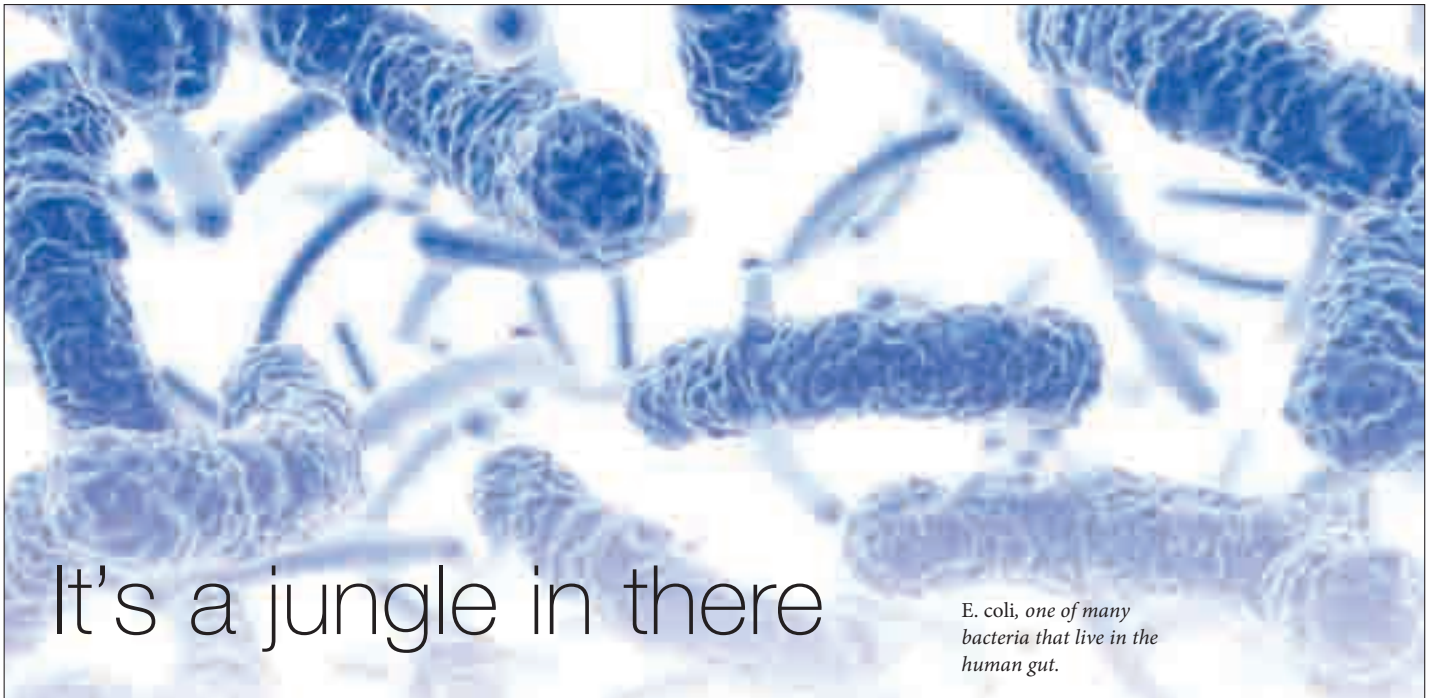
To build an apartment block, a construction company must have the blueprint for the entire building. But workers on the ground only need to refer to small sections of the document that contain the instructions that are relevant to them. Similarly, each cell in an embryo carries the genetic blueprint for the whole organism but, by referring to different parts of this developmental plan, cells can perform different functions depending on where they are and how far the embryo has developed. The foremen who control what part of the genetic blueprint the cellular machinery will act on are called transcription factors. They bind to stretches of DNA known as cis-regulatory modules (CRMs) and, once bound, they can switch genes on or off. "We wondered: would it be possible to use information on when and where this binding happens to predict CRM activity?" says Eileen Furlong, joint head of the Genome Biology Unit at EMBL Heidelberg.

Biologist Robert P. Zinzen, computer scientist Charles Girardot and statistician Julien Gagneur teamed up to take a novel, integrated approach to answering this question. They identified and recorded the binding profiles – i.e. the combinations of transcription factors binding at different times and places – of approximately 8000 CRMs involved in regulating muscle development in the fruit fly *Drosophila*. The activity of a number of these CRMs had been previously studied, and the EMBL team used this information to group them into classes according to the type of muscle and the developmental stages in which they are active. The scientists then trained a computer to unravel the binding profiles for each of these groups, and to then search the 8000 newly identified CRMs for ones whose binding profiles fitted one of the pictures. Such CRMs were thereby predicted to have similar activity patterns, implying that they are involved in regulating the development of the same muscle type.

"When we tested the predictions experimentally, the results were not only accurate but also enlightening," says Eileen. She explains: "It turns out that the regulatory code, where one binding profile leads to one pattern of CRM activity, is actually not that straightforward." They found that CRMs with strikingly different binding profiles can have similar patterns of activity. Although it may seem unnecessarily redundant to have several different work teams capable of building the same part of the house, this unexpected plasticity does make sense in evolutionary terms, the researchers say. The fact that different combinations of transcription factors, or binding codes, can regulate the same developmental process means that even if some transcription factors or CRMs change or are lost during an organism's evolution, it can still develop a gut muscle, for instance. In other words, having different work teams that use different tactics to build the same room provides a back-up system: if one team has an accident, another can step in and do the job.

"What's exciting for me is that this study shows that it is possible to predict when and where genes are expressed, so we can infer what functions those genes may perform," concludes Eileen. "It is an important first step towards understanding how regulatory networks drive development."

Zinzen RP, Girardot C, Gagneur J, Braun M & Furlong EEM (2009) Combinatorial binding predicts spatio-temporal cis-regulatory activity. *Nature* **462**: 65-70



It's a jungle in there

E. coli*, one of many bacteria that live in the human gut.*

Etiquette dictates that there are some conversation topics one should avoid at a polite dinner party. Religion and politics are the obvious ones. Faeces, you might think, would be another. But that might be about to change. EMBL researchers have created a detailed catalogue of the microbial genes found in human faeces and in so doing have built a fascinating picture of the intricate ecosystem flourishing inside our intestines. The catalogue will act as an invaluable reference for researchers trying to understand how this ecosystem contributes to human health and disease – surely the stuff of dignified dinnertime discourse.

It's a slightly uncomfortable thought, but the microbes that live in or on your body outnumber your own cells by at least a factor of ten. Most of these 100 trillion or so organisms live inside your gut, and are thought to play a role in many aspects of our biology, such as modulating our immune system and contributing to our nutrition. Changes in the balance of this ecosystem might be involved in diseases such as inflammatory bowel disease or obesity.

But many of these microbes can't be grown in the lab, making it hard to determine what species are present. Instead, Peer Bork, joint Head of the Structural and Computational Biology Unit at EMBL Heidelberg, together with an international team of scientists, turned to metagenomics, a technique that involves sequencing as many genes as possible straight from an environmental sample. In this case, it was samples of faeces taken from 124 European volunteers as part of an EU project called Metagenomics of the Human Intestinal Tract, or MetaHIT. The approach allowed the team to construct a "metagenome", a virtually comprehensive catalogue of all the genes present in the gut's microbial ecosystem.

This was made possible thanks to a new kind of technology called Illumina sequencing, which yields very short stretches of DNA sequences. Although this dramatically cuts costs, it presents a considerable challenge for bioinformaticians, who have to find a way of piecing these fragments of sequence together correctly into entire genes. Until now, researchers didn't think it was possible to do this for metagenomics work. But Peer and his colleagues have shown that if they sequence a large number of samples, it can indeed be done. "It's a dramatic advance," says Peer. "We can now really tackle ecology."

Although the team hasn't sequenced enough to find every single gene, the new dataset is 200 times bigger than any other metagenomic study done so far. "There will certainly be more genes to come," says Peer of the gut metagenome, "but we don't expect many more."

The team used the data to construct a "minimal genome", the bare essential set of genes needed for a bacterium to survive in the gut, as well as a "minimal metagenome", the set of genes needed for the entire ecosystem to function. Intriguingly, the team also found that, contrary to previous suggestions, people seem to share most of their microbial species. They also showed that people with bowel diseases such as Crohn's disease harbour distinctive microbial populations.

Although the creation of this catalogue is a significant milestone, Peer emphasises that it is just the beginning of a new area of research. "We intend to follow up on the disease findings and the ecology," he says. Who knows, years from now, it might not only be polite, but *de rigeur* to discuss the composition of one's gut ecosystem at the dinner table, over a dessert of yogurt microbiologically customised to increase your well-being.

Qin J *et al* (2010) A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**: 59-64

Nadia Rosenthal

Young at heart

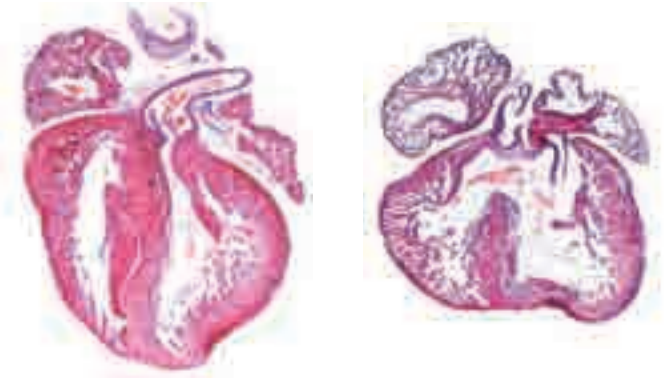
Pity Tithonus! Lover of Eos, the Titan of dawn, he was granted the gift of eternal life by Zeus, king of the gods. Trouble was, the absent-minded Eos forgot to ask Zeus to grant him eternal youth as well. As the years passed, the poor immortal Tithonus withered and aged so grievously that he implored daily for death. Eos took pity on him and, unable to restore his youth, turned him into a grasshopper to end his suffering.

Although this fable is thousands of years old, it has particular resonance today. Thanks to modern medicine, we can expect to live far longer now than ever before. But much of this increased lifespan is plagued by diseases of ageing and biologists working in the field of regenerative medicine are trying to find ways to replace ailing and damaged tissue. Some are hoping to do this by reawakening the processes behind embryonic development in adult tissues, so that old organs might renew themselves. Now, a team led by Nadia Rosenthal at EMBL Monterotondo has discovered a possible way of tapping into the heart's inner youth so that it can repair itself. In doing so, they have also gained new insights into how heart development can go awry in the womb. And like Tithonus, the solution lies in an insect – not in the chirping legs of a grasshopper, but in the gauzy wings of a fruit fly.

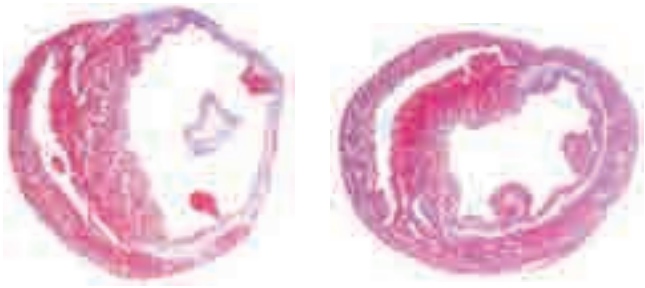
Heart disease remains the world's biggest killer of children and adults alike. Congenital heart problems are the most common form of birth defect, and heart failure caused by long-term high blood pressure and heart attacks are prevalent causes of heart problems in adults. So researchers are keen to both dissect the biological processes governing heart development and devise new treatments for heart damage.

Although Nadia and her team have ultimately found some answers to these problems, they originally started work on a different project. PhD student Paschalis Kratsios had joined the lab, and was studying a gene called NFκB. The protein made by this gene is involved in a signalling system that induces cells to divide or to die. Paschalis was using a genetic technique to turn off NFκB signalling in the hearts of adult mice to study its role in keeping the heart healthy. But the work progressed slowly, despite assistance from another lab member, Foteini Mourkioti. "I am eternally grateful to her," says Paschalis, "she was the postdoc who took me by the hand." The experiments were working, but Paschalis was frustrated by having to wait for several months before trying to identify any physical defects in the hearts of the mice. And even at the end of the wait, it was hard to spot anything wrong.

So in the meantime, he started looking for another project, one that he could develop himself and call his own. "I wanted my own baby!" he says. "I'm really grateful to Nadia for the opportunity to join the EMBL family and have the intellectual freedom to explore." Reading up on the literature about heart disease, he came across a family of genes called the *Notch* family. Unlike the NFκB pathway, there were a number of human heart malformations associated with mutations in these genes. One of these, Alagille syndrome, results in congenital heart defects including holes between chambers, an enlarged right ventricle and narrowing of certain arteries. Research in mice



In a normal newborn mouse's heart (left), the two major heart chambers are clearly separated by tissue, but when Notch signalling was inactivated in an embryo's heart muscle cells (right), this septum between the ventricles was incomplete.



In these microscopy images of adult hearts, healthy tissue is shown in red and damaged tissue in blue. Normally (left), a heart attack causes extensive tissue damage to the left ventricle (right-hand cavity), but this tissue damage was reduced in mice in which Notch was re-activated after the heart attack (right).

had also confirmed that *Notch* played an important role in heart development. Other work, meanwhile, hinted that *Notch* could be involved in regenerating damaged tissues, and was therefore worth investigating as a potential focus for therapies to treat heart disease. "From the existing literature on several model organisms, it looked like *Notch* would be a good target," says Paschalis.

Notch genes take their name from the first member of the family to be discovered in the fruit fly *Drosophila melanogaster*. Fly geneticists had been studying mutant flies that had little nicks, or notches, in the edges of their wings. When they isolated the associated gene, they found that it coded for a receptor – the cellular equivalent of a TV satellite dish – that sits in the outer membrane of cells to receive signals from the environment and other cells.

But dissecting *Notch* function in mice wasn't going to be easy. There are a total of four *Notch* genes in the mouse, all of which are active in the developing heart. It was quite possible that one gene could compensate for the loss of another, making traditional genetic experiments to switch off the activity of genes extremely hard to do. So Paschalis and his colleagues took the opposite approach – they used a genetic trick to make *Notch* signalling overactive in their chosen cell types.

They began by looking at overactive *Notch* in developing mouse hearts and saw that the mice developed heart defects. "We saw that the mice's hearts became bigger," says Paschalis. Further analysis showed this was because the developing heart muscle cells had undergone too much cell division and had not developed their specialised functions properly. The team then used another genetic trick that allowed them to switch off all *Notch* signalling in the developing hearts. Sure enough, this had the opposite effect to overactivating *Notch*: one of the heart chambers was abnormally small and its walls thin – just like human Alagille syndrome patients. This all points to the levels and timing of *Notch* signalling as being crucial for controlling developing heart muscle cell division, and gives a valuable new insight into its role in congenital heart disease.

The role of *Notch* in heart cell division was also an exciting finding for research into treatments for acquired heart disease – conditions that develop later in life, such as heart failure following heart muscle damage caused by a heart attack. There is a great deal of interest in developing therapies that make the heart repair itself, which include boosting cell division. "Our findings lend support to the notion that, in certain situations, redeployment of embryonic signalling pathways could prove beneficial for tissue regeneration in the adult," says Nadia.

With this in mind, the team decided to switch on *Notch* signalling in adult mice whose hearts had been surgically damaged to mimic the effects of a heart attack. The mice with extra, genetically induced, *Notch* signalling had improved heart function and were more likely to survive than those without. "It turned out to be beneficial," says Paschalis.

Obviously, it would be impossible to perform the same kind of genetic tricks in human patients, so the team investigated another – more clinically relevant – tactic to raise *Notch* signalling in the mice. This time, they injected the damaged hearts with an antibody that switches on *Notch* signalling and again found significant improvements compared with the untreated mice. As well as having reduced scarring, the treated hearts grew new blood vessels. And in contrast to the findings in the embryonic heart, there was little extra cell division in the adults. Instead, the *Notch* signalling seemed to help the existing muscle cells survive. "Overall, these results highlight the importance of timing and context in biological communication mechanisms," says Nadia. To complete the experiments, they then switched off *Notch* function in mice with damaged hearts. As expected, these mice were much more likely to die and exhibit severe heart failure.

In theory at least, this suggests that researchers could indeed redeploy a developmental process to repair damaged hearts and exploit *Notch* as a possible way to treat congenital heart disease. Nadia's team will, however, be taking the work for-

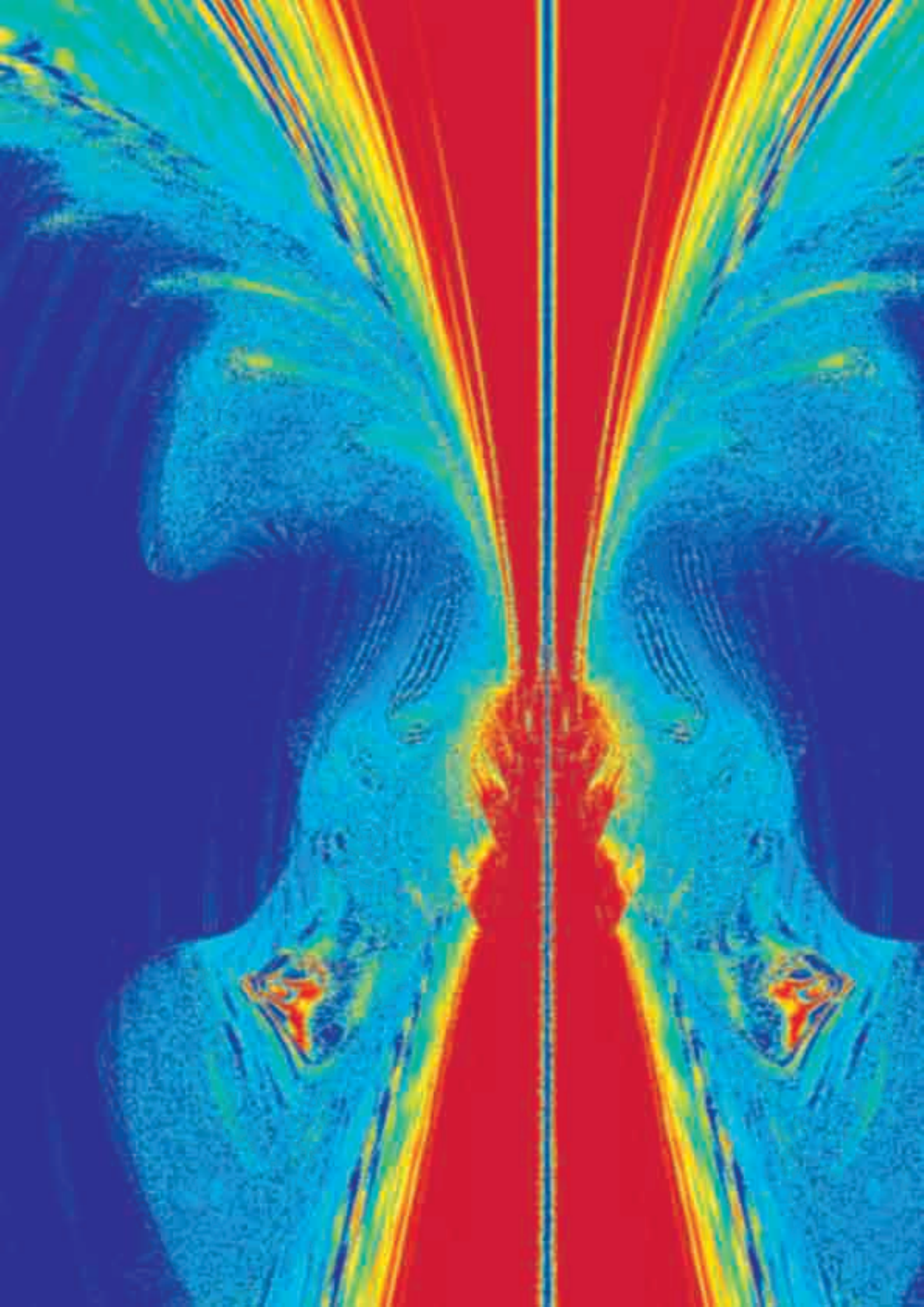
Artist's representation of how looking at embryonic hearts can help repair adult hearts.

ward without Paschalis, who is now pursuing his postdoctoral studies in a neurobiology lab in New York. But the story doesn't end there. It turns out that Paschalis did indeed finish his NFκB experiments, and uncovered yet more intriguing results for Nadia's team to follow up.

When Paschalis disabled NFκB signalling in the mouse heart, he found the mice developed heart muscle wasting, inflammation and heart failure as they aged. These symptoms weren't obvious in younger mice, which is why he'd had a hard time spotting them. But when he increased the blood pressure of these young mice, the symptoms appeared much sooner. In other tissues, NFκB is known to play a role in reducing oxidative stress – damage caused by chemically reactive by-products of metabolism. To see whether a lack of oxidative stress protection could explain the mice's symptoms, Paschalis fed them a diet rich in anti-oxidants. This partly resolved the symptoms, but didn't reverse them completely, suggesting that other unknown factors were coming into play.

Overall, the findings pointed to NFκB signalling being needed to protect the adult heart from mechanical and age-related stress although it's not yet clear how this might work in human heart patients. "So far, we don't know of any human NFκB pathway mutations that are associated with heart disease," says Paschalis. Even so, the work done by Nadia and her team raise the intriguing possibility that we might one day be able to enjoy our later years minus at least some of the ravages of age – and we won't need to be magically transformed into insects to do so.

Kratsios P, Catela C, Salimova E, Huth M, Berno V, Rosenthal N, Mourkioti F (2009) Distinct roles for cell-autonomous *Notch* signalling in cardiomyocytes of the embryonic and adult heart. *Circulation Research* **106**: 559 - 572



When disaster strikes

A little after midnight on Sunday 2nd September 1666, a fire broke out in a bakery in Pudding Lane in the city of London. Fanned by the wind, flames quickly spread to neighbouring houses and streets. The dithering Lord Mayor delayed demolishing houses to create firebreaks, the only major fire-fighting technique available at the time. As a result, the blaze turned into a raging firestorm that engulfed the city for three days, consuming thousands of homes, destroying St Paul's Cathedral and turning hundreds of thousands of Londoners into destitute refugees. At last, on the Wednesday, the fire died down, leaving what contemporary diarist Samuel Pepys described as "the saddest sight of desolation that I ever saw."

Had London had a dedicated firefighting organisation – and a more decisive Lord Mayor – much of this destruction could have been averted. Modern firefighting techniques and services mean that small fires no longer turn into major conflagrations. What's more, other emergency services, such as ambulance crews, can help to limit the damage and loss of life, and minimise the rebuilding work that needs to be done. Of course, preventing disasters, both big and small, is far preferable, which is why governments dedicate time and money to monitoring, gathering intelligence and defence.

Organisms and cells have their own equivalents of emergency services: immune surveillance systems to warn of microbial attack, proteins that detect and minimise damage, and cells that trigger the repair of damaged tissue. EMBL researchers have been investigating these mechanisms in a range of creatures, from plants to humans. A team at EMBL Grenoble, for example, has determined the structure of a hormone receptor that helps plants to respond to water shortages, giving scientists a new avenue for the development of more drought-tolerant crops. Other EMBL researchers, meanwhile, have uncovered a gene that makes mosquitoes resist infection by the malaria parasite – a possible step towards restricting the spread of this deadly disease.

Within our own cells, molecular emergency workers are constantly fending off calamity. Our DNA is under endless chemical and physical attack, and it is thanks to dedicated DNA-repair systems – being investigated by a team at EMBL Heidelberg – that our genes sustain only a fraction of the damage that hits them. Of course, no system is perfect, and a collaboration between Grenoble and Heidelberg is revealing how the influenza virus slips through our defences, how it forms virulent new strains and, ultimately, how it triggers lethal pandemics. Understanding how such mishaps unfold at the cellular level will hopefully allow researchers to find ways of erecting the equivalent of firebreaks and other timely measures to stop them becoming full-scale catastrophes like the Great Fire of 1666.

On call for DNA damage

Gyula Timinszky and Andreas Ladurner

Our DNA is under attack. Ultraviolet light, cigarette smoke, toxins, even harmful by-products of our metabolism constantly bombard our DNA, disrupting its delicately ordered structure. The scale of the problem is enormous: each of our bodies' trillions of cells receives tens of thousands of DNA-damaging hits per day. If unrepaired, or repaired incorrectly, this damage leads to mutations that can ultimately cause diseases including cancer. As the sole source of information for the construction and function of our bodies, our genomes have to be continuously repaired. To do this, our cells use an impressive fleet of DNA-repair machines – a crack squad ready to jump in, correct the mistakes, and rejoin nicked, severed or damaged strands. But how do these machines know where to go?

Each cell has an early warning system to detect DNA damage and this is handled by proteins of the PARP family. Like cops on the beat, PARP proteins patrol the genome, tracking along DNA strands and looking for trouble. With a special affinity for the ends of cut DNA strands, their speciality is homing in on breaks in the DNA. When they find such damage, they set cellular alarm bells ringing by marking that region as defective. They do this by attaching chemical tags to neighbouring proteins associated with the DNA in that region. These include histones – around which DNA strands are wound – and other proteins that package our huge genomes into a manageable system called chromatin. The chemical tags take the form of long, occasionally branching chains of a molecule called poly-(ADP-ribose), or PAR, which was first identified by scientists around 50 years ago and from which the PARP family gets its name: PAR polymerase, i.e. an enzyme that can make chains of PAR. Responding to this emergency signal, it's not long before repair machines arrive on the scene and the damage is put right. But quite how these machines respond to the PAR signal or, more precisely, what binds to PAR and mediates the response, was an open question for 50 years until Andreas Ladurner and his group at EMBL Heidelberg recently helped to find the answer.

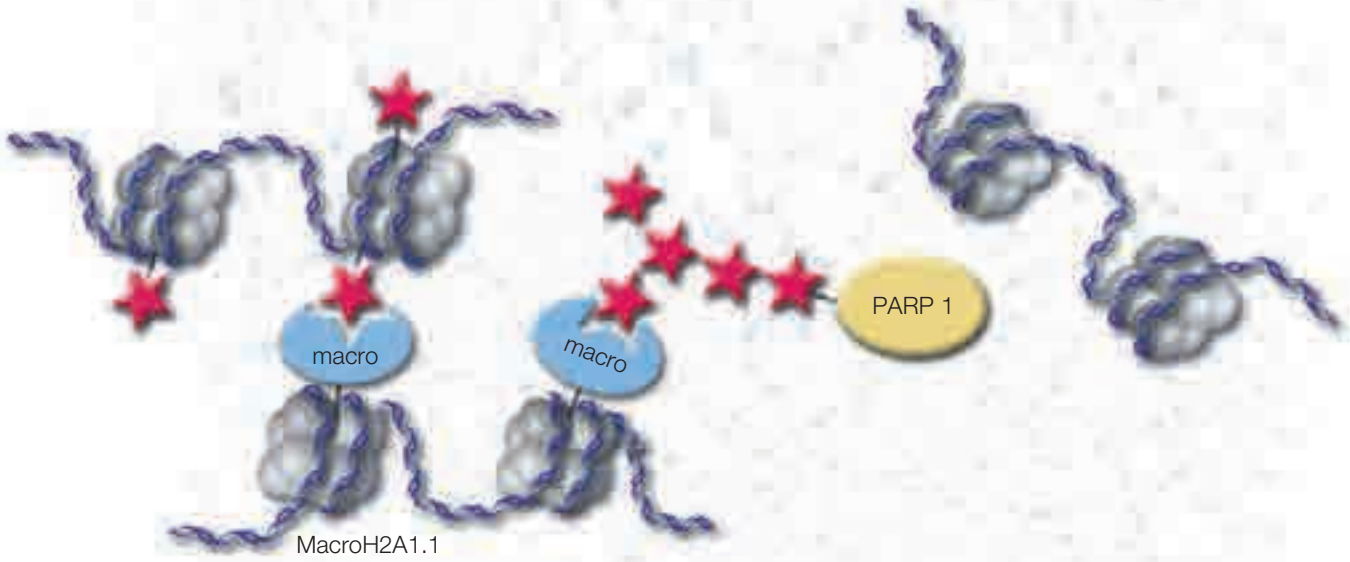
The story began back in 2003 when Andreas first came to EMBL. Pulling together scant evidence from distant corners of the literature and drawing on his experience as a structural biologist, he came up with a possible PAR-binding candidate: the macrodomain. This is a chunk of protein (or protein domain) of unknown function that is common to many different proteins and is found in species ranging from humans to single-celled bacteria. Andreas suspected that this segment of protein might bind to PAR because it has a structure similar to that of other protein domains that bind to PAR-like molecules and is often found in proteins that perform other tasks involving PAR, such as PAR synthesis. Could these domains be the elusive piece in the jigsaw puzzle that had lain unfinished for 50 years? All that remained was to test the hypothesis.

Two Diploma students proved Andreas' hunch to be correct and showed that macrodomains could indeed bind to PAR – a groundbreaking discovery. In a follow-up paper they showed that the same was true for a specific human macrodomain-containing histone, macroH2A1.1. So having shown, at least in the test-tube, that macrodomains do bind to PAR specifically, the question was whether they could also do this in real life, inside a cell. Could macrodomain proteins really be the first emergency response to the PAR alarm signal, homing in solely on damaged DNA? It was around this time that postdoc Gyula Timinszky started in the lab. Assigned with the task of designing an experiment to test whether macrodomains could respond to DNA damage and drawing on EMBL's expertise in microscopy, he came up with an ingenious solution. He refined a technique together with Julien Colombelli in Ernst Stelzer's group at EMBL that meant he could place a dish of live cells on a microscope, then zap them in the nucleus with a laser to cause localised damage to their DNA. Peering down the microscope Gyula could

"The first time I saw it I was amazed: it's incredible how quickly it happens. The damaged area begins to light up within a few hundred milliseconds."

then watch in real time as macrodomains, labelled fluorescent green, swarmed across the nucleus, homing in on the damaged DNA and lighting up the area that had been hit by the laser. "The first time I saw it I was amazed," explains Gyula, "it's incredible how quickly it happens. The damaged area begins to light up within a few hundred milliseconds." This means that PARP finds the damage and sets off the emergency alarm by labelling neighbouring histones and proteins with PAR, which is then immediately detected and bound by macrodomains, all in under a second. Response times like these would be the envy of any emergency service. Life at the subcellular level really can happen incredibly fast.

With the assay up and running, Gyula set about systematically testing different macrodomain proteins to see how they might respond. In the meantime, Andreas happened to give a talk at the Stowers Institute in Kansas, USA, where he presented Gyula's assay and some of their latest data. Afterwards, he was approached by a member of the audience, Joan Conaway, who together with her husband Ronald ran a molecular biology lab interested in chromatin remodelling. "It was an amazing coincidence," says Andreas. "She explained to me how they had been working on a family of



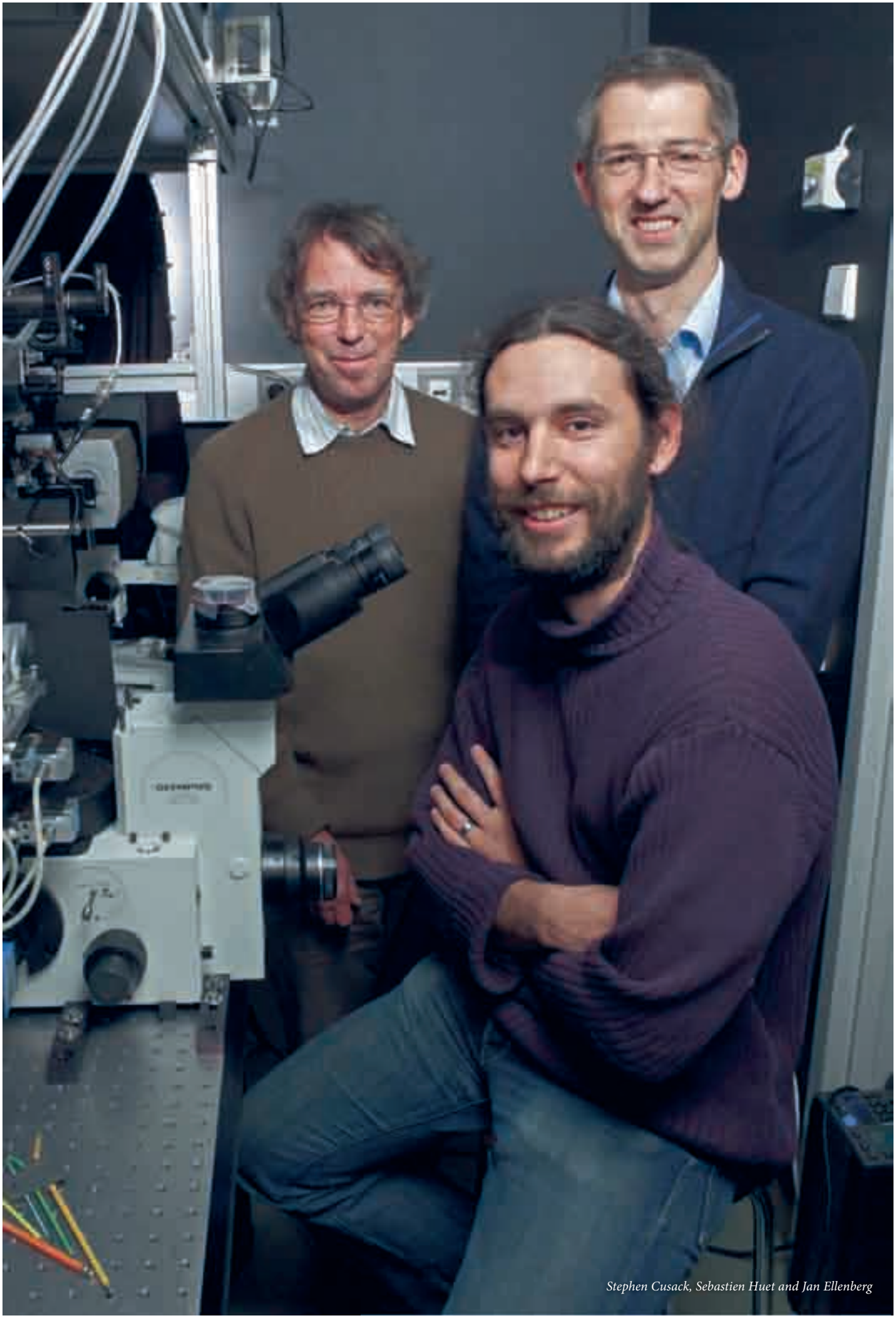
Chromatin containing the macroH2A1.1 macrodomains responds to the emergency signal (PARP1) and helps mediate the DNA repair machinery.

chromatin remodellers including one called CHD1-like. This was the alternative name for Alc1 – one of the macrodomain proteins we had in our pipeline. She basically told me that her PhD student had already done all the biochemistry on this protein and asked whether we would like to team up together for publication!" What they had shown was that Alc1, in addition to its macrodomain, contains a domain involved in remodelling chromatin. Their data showed that, by creating the PAR alarm signal, PARP promotes chromatin remodelling in damaged regions of the genome by recruiting Alc1. Relaxing the chromatin might help to repair the damage by allowing access for the repair machinery. With Gyula's assay showing that Alc1 is rapidly recruited to damaged DNA – and is actually the fastest macrodomain protein that Gyula tested – their data combined to make a really interesting story.

Gyula, Andreas and the lab then focussed their attention on a second case looking again at the macrodomain-containing histone, macroH2A1.1. At first glance it seems bizarre that a histone should possess a domain that makes it respond to DNA damage in this way. As Andreas explains, "We tend to think of chromatin in architectural terms – as a relatively static infrastructure, to which modifications are made. Why would a histone, wrapped up in DNA and embedded in the chromatin, be able to dash across the nucleus to bind PAR groups at sites of DNA damage?" But their assays showed that it can. "Histone macroH2A1.1 is slower than Alc1," admits Gyula, "which isn't so surprising – you can imagine that they have quite a fight to get there, dragging all the chromatin they're wound up in with them. But within five minutes of DNA damage, these histones do arrive at the site of the damage." "We think that what this does is to rearrange the chromatin around the damaged area," says Andreas. "This may help keep the ends of severed DNA strands closer together until the breaks can be repaired. The bottom line is that you have histones out there that can sense when PARP rings the alarm bell." This comprehensive study shows for the first time that cells have at their disposal a whole family of proteins ready to respond to the PAR signal. Determining the details of what happens next – for example, precisely how macrodomain proteins bring about changes in chromatin structure and how they relay information to the DNA-repair machinery – will undoubtedly keep the group busy solving puzzles for many years to come.

Gottschalk AJ, Timinszky G, Kong SE, Jin J, Cai Y, Swanson SK, Washburn MP, Florens L, Ladurner AG, Conaway JW, Conaway RC (2009) Poly(ADP-ribosyl)ation directs recruitment and activation of an ATP-dependent chromatin remodeler. *PNAS* **106:** 13770-13774

Timinszky G, Till S, Hassa PO, Hothorn M, Kustatscher G, Nijmeijer B, Colombelli J, Altmeyer M, Stelzer EHK, Scheffzek K, Hottiger MO & Ladurner AG (2009) A macrodomain-containing histone rearranges chromatin upon sensing PARP1 activation. *Nat Struct Mol Biol* **16**: 923-929



Stephen Cusack, Sebastien Huet and Jan Ellenberg

Viral surveillance

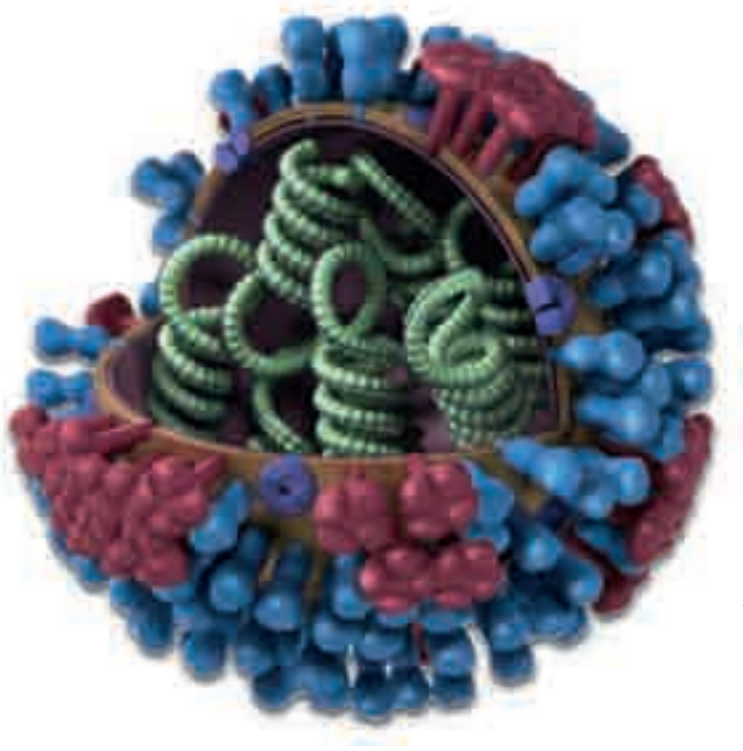
State-of-the-art gadgets. Aerial surveillance of a suspicious target. An adversary with a penchant for gambling. While these might sound like the ingredients of a detective novel or a spy thriller, they are in fact aspects of a successful collaboration between two EMBL teams, which has revealed new insights into how the influenza virus develops pandemic strains. The collaboration, between Stephen Cusack's team at EMBL Grenoble and Jan Ellenberg's group at EMBL Heidelberg, has also highlighted the benefits of the EMBL Interdisciplinary Postdoctoral Programme, EIPOD. By combining their expertise, the teams have shown how the flu virus can win at a kind of genetic card game and create deadly viruses that are able to jump from animals to humans.

For humanity, the consequences of this card game can be dire. The 1918 Spanish flu pandemic killed around 40 million people – more than the Great War. Last year, a flu strain that originated in pigs developed the ability infect and spread between humans. And researchers have long been keeping an eye on one highly pathogenic bird influenza virus, called H5N1, which seems poised to trigger another lethal pandemic. Now, Jan and Stephen have managed to track vital parts of the virus as the particles assemble themselves in living cells. "As well as giving new leads for much-needed new drugs against flu, the work offers valuable new insights into how the virus can jump from one species to another and become highly virulent," says Stephen.

The secret of influenza's ability to seemingly ambush us out of nowhere lies in its predilection for gambling. Normally, the viral strains that cause seasonal flu outbreaks each winter differ only slightly from the previous year's strains. This is because the virus's genes have been altered only slightly through minor mutations. It's a bit like playing poker with a similar hand of cards from one game to the next. As a result, most people manage to build up an immunity that lasts from one flu season to the next and vaccines based on past strains work well.

The trouble really starts when two very different flu viruses infect the same cell at the same time. The virus's genetic material is divided into eight segments, each carrying instructions for building different parts of the virus. When a single virus infects a cell, it makes copies of these segments, which are then packaged with proteins to make new viruses. If, however, two different flu viruses are copying themselves in the same cell, segments from one virus can get packaged with segments from the other. This results in a "reassortant" virus with a completely new combination of genes – rather like shuffling and dealing a pack of cards.

The outcome of this shuffling depends on the genetic hand each new virus gets. Usually, the new combinations of genes are incompatible and result in weaker or defunct viruses. Sometimes, however, a virus is dealt a winning hand: a set of compatible genes that result in a virus that is completely new to the human immune system. This is strike one in the virus's favour. Strike two comes when the new virus – particularly one that previously only infected animals or birds – acquires the ability to infect



Model of the influenza virus.

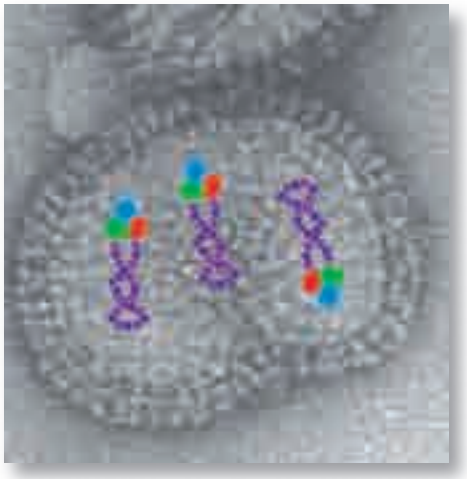
humans and to spread easily between them. This is how a pandemic strain is born.

Stephen's team has long been interested in one particular part of the virus, its polymerase – the enzyme that both copies the virus's genome and turns the instructions encoded in it into proteins needed for survival and replication. Several mutations known to help the virus adapt to life in a human host arise in the genes for the polymerase. "What's more, the way the enzyme is assembled in the cell is thought to play a key role in determining which reassortant viruses are viable," says Stephen.

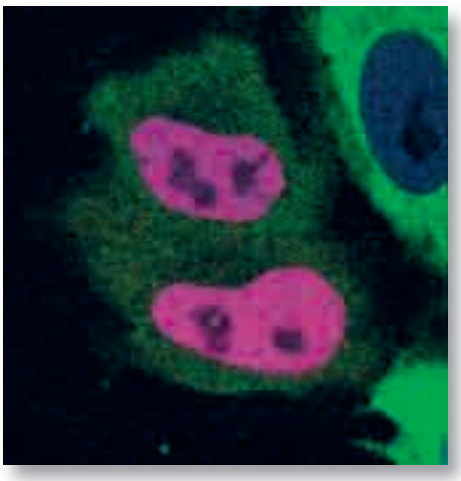
"As well as giving new leads for much-needed new drugs against flu, the work offers valuable new insights into how the virus can jump from one species to another and become highly virulent."

The polymerase is actually made of three separate parts, or subunits – PA, PB1 and PB2 – which are each encoded by a different genome segment. Thanks to the Grenoble team's work, scientists now have a better picture of the structure of these subunits and how they might interact with proteins in the host cell. But until now, no-one knew for sure how the three subunits made their way into the cell's nucleus, where they produce new viruses, nor how they assembled together to form a single enzyme.

So Jan and his team decided to take a closer look. Jan's postdoc Sebastien Huet and Sergeyi Avilov, an EIPOD postdoc shared between the Ellenberg and Cusack labs, turned to state-of-the-art microscopy techniques – with the help of Heidelberg's Advanced Light Microscopy Facility – that allowed them to track proteins and watch how they interact inside living cells. One technique, called Fluorescence Cross Correlation Spectroscopy, or FCCS, relies on tagging proteins in a cell with fluorescent molecules. With the help of a specialised light microscope, researchers can measure the concentration and follow the movement of these tagged proteins by seeing how the fluorescence fluctuates as these molecules diffuse in and out of a tiny laser spot that they can point at different sections of the cell. By labelling different proteins with different colours of fluorescent molecules, researchers can also track interactions between these proteins: if both fluorescent signals travel through the laser spot in the same way, it shows that the two proteins are indeed sticking together and interacting. It's almost like having spy-satellite surveillance of an enemy's ordnance factory at work. "This really allows us to probe the predictions of how these subunits behave under physiological conditions," says Jan.



Schematic representation of the particles containing the viral RNA, the nucleoprotein and the polymerase complex superimposed to an electron micrograph of the virus.



Fluorescence Cross Correlation Spectroscopy image of the three different influenza polymerase subunits: PA is shown in purple, PB1 in green and PB2 in blue.

In the first study to apply this technique to viral proteins, Jan and his team tracked the behaviour of the three polymerase subunits in cells. They found that PB1 and PA bind together in the cell's cytoplasm before moving into the nucleus. PB2, meanwhile, hangs around by itself in the cytoplasm and enters the nucleus separately, before associating with the PB1/PA duo to form the complete polymerase.

Further work allowed the team to translate the Grenoble group's earlier findings about the polymerase's structure into what is happening inside living cells. "This is why Stephen and I are so keen on this collaboration," says Jan. PB2 hitches a ride into the nucleus on a host protein called importin by using a kind of molecular grappling hook called a nuclear localisation sequence. "We know this based on the structural work of Stephen's group," says Jan. Jan's team altered the amino acids in PB2's nuclear localisation sequence and found, as expected, that PB2 stayed in the cytoplasm. Surprisingly, however, the other subunits also stayed in the cytoplasm, and bound to PB2 to form a polymerase enzyme that was intact but in the wrong location.

This suggests that PB2 helps to keep the other subunits in the right place – the nucleus – until they can assemble into the polymerase. "If you could interfere with that process with drugs," says Jan, "that would be a good way of trapping the polymerase in the cytoplasm and stopping it from getting into the nucleus." It might also explain why some reassortant viruses fail to replicate properly, he adds. If the polymerase subunits don't fit together well, they won't stay in the nucleus for long enough to make a working enzyme.

By measuring the brightness of the fluorescence given off by labelled subunits that were normal and non-mutated, Jan's team determined that a small number of these polymerases were sticking together. This suggests that polymerases already working on the virus's genetic material help attract other polymerases entering the nucleus to the right location, so boosting the rate of viral replication. As for the rest of the subunits, FCCS revealed that they were moving around the cell far more slowly than expected. This suggests that they are bound to a range of host proteins, which may be important for polymerase entry and assembly within the nucleus.

In future, Jan and Stephen plan to study these interactions between the host cell proteins and the influenza polymerase in more detail. Because FCCS gives such precise measurements, it will allow Jan and his team to study how mutations in the polymerase subunits affect how efficiently the subunits interact with host cell proteins – and so how they allow the virus to adapt to new host species such as humans.

Huet S, Avilov S, Ferbitz L, Daigle N, Cusack S, Ellenberg J (2010) Nuclear import and assembly of the influenza A virus RNA polymerase studied in live cells by Fluorescence Cross Correlation Spectroscopy. *J Virol.* **84**: 1254-1264


2.00
1.90
1.80
1.70
1.60
1.50
TEP1*S3
TEP1*R1

The usual suspects

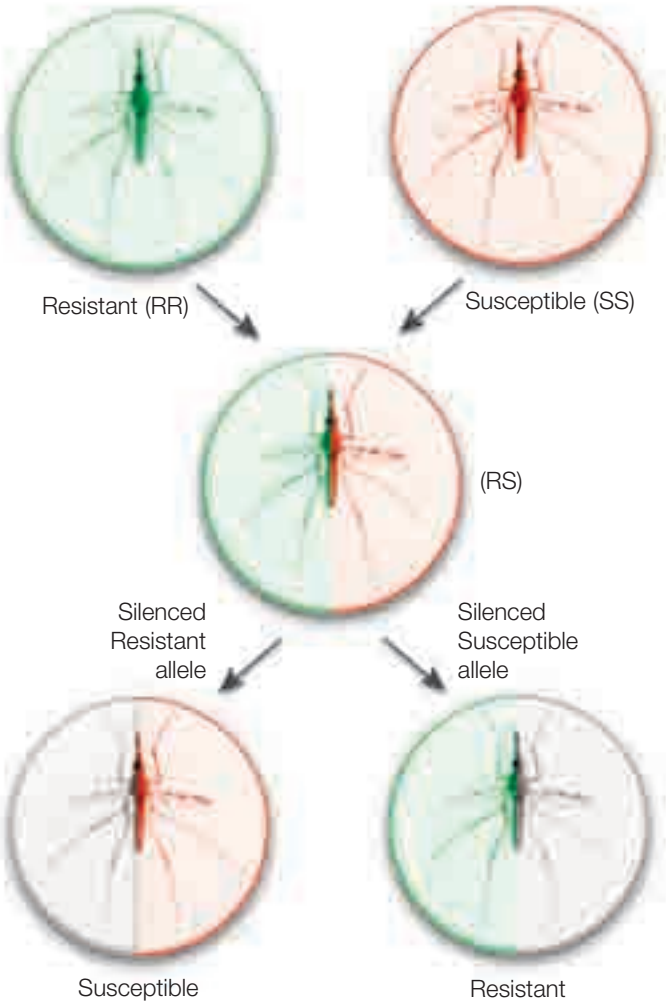
From Hercule Poirot to CSI, most detective stories and TV police shows revolve around the investigators narrowing down their list of suspects. At the start of the story, there are often as many suspects as there are characters, but thanks to varying measures of deductive power and luck, the good guys systematically strike suspects off their list until finally they are left with only one – the culprit.

Although they use different interrogation techniques, geneticists often face the same task of whittling down a list of suspects – and in some cases, they too are dealing with killers. In this particular story, the killer's identity is known from the start: malaria, a disease that claims almost one million human lives a year. The parasites that cause malaria are also well known to scientists. Belonging to the genus *Plasmodium*, they spend part of their life cycle inside humans or another mammalian host – depending on the exact species of parasite – and another part inside mosquitoes. And just as we suffer from malaria, mosquitoes infected with *Plasmodium* parasites do to. However, this is where the mystery begins: not all mosquitoes that are infected with malaria parasites fall ill. Instead, some are resistant to the parasite, meaning that their immune system is able to eliminate the disease-causing agent from their bodies. Why can some individual mosquitoes fight off the parasite better than others? Enter our lead investigators: Lars Steinmetz, joint head of the Genome Biology Unit at EMBL Heidelberg, and EMBL alumni Rui Wang-Sattler, now at the Helmholtz Zentrum in Munich, and Stephanie Blandin, now at the Institut National de la Santé et de la Recherche Médicale (INSERM) in Strasbourg, France.

To get to the bottom of this enigmatic variation in individual resistance, Lars, Rui and Stephanie turned to *Anopheles gambiae* mosquitoes, the main carriers of the parasite that causes the most severe form of human malaria in Africa, and to *Plasmodium berghei*, which causes malaria in rodents. The scientists compared the complete DNA sequences, or genomes, of mosquitoes that are resistant to malaria with those of mosquitoes that are susceptible to the disease. This comparison suggested that a section of one of the mosquitoes' chromosomes is linked to the insects' ability to fend off *Plasmodium*. But as this section contains 975 genes, this still left our investigators plenty of suspects.

"To really understand this matter, we needed to go further," says Rui. "But the question was how to go from this section of DNA to the single gene that makes mosquitoes resistant to malaria," Stephanie adds. This is when Lars' expertise was called upon.

Lars and his group had previously developed a way of doing just this in baker's yeast. To go from a large DNA section, or genetic interval, to the gene that causes a particular trait, they took two yeast strains with different versions, or alleles, of a particular gene and crossed them, obtaining one strain of yeast with both alleles of that gene. They could then take these yeast cells, divide them into two sets and delete – or knockout –one allele in one set and the other allele in the other. As both sets of cells were genetically identical for that one gene, any differences in

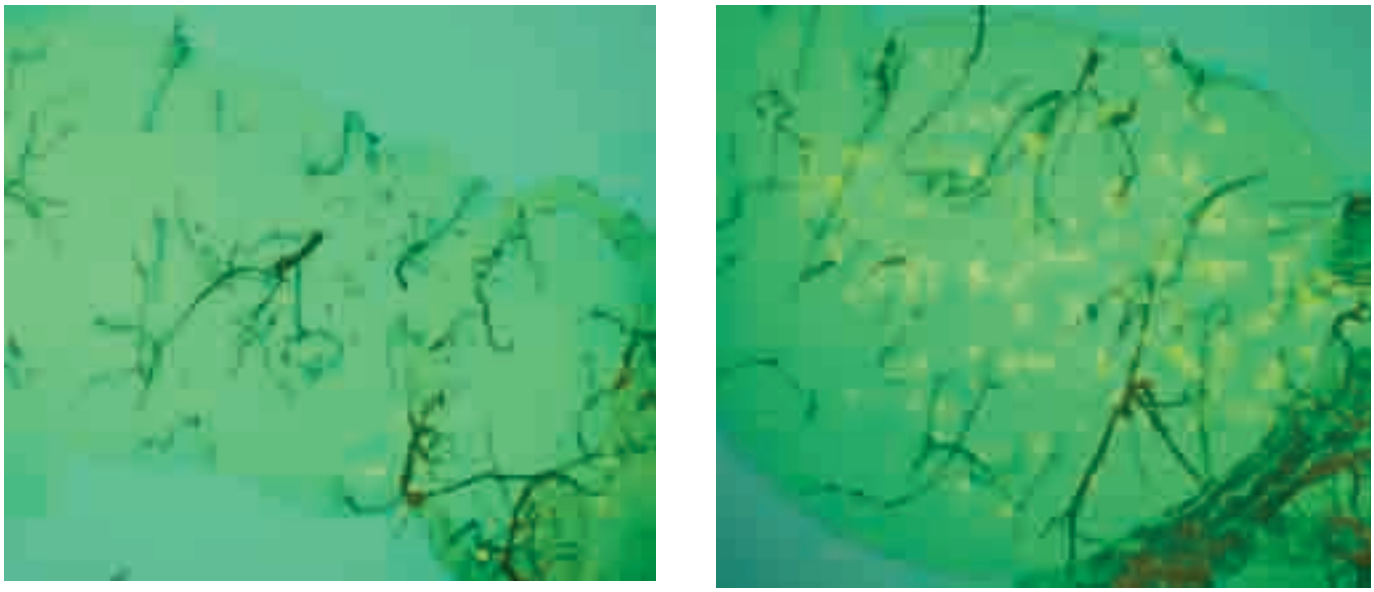


Reciprocal allele-specific RNAi entails obtaining individuals each carrying the 2 different alleles of interest, silencing one or other of those alleles, and comparing the results.

the traits of the two sets of cells must stem from the fact that they had different alleles of that gene. By doing this systematically for all the genes in the interval, Lars and colleagues were able to narrow down their suspects and determine exactly which genes played a role in a particular trait. "Unfortunately, we can delete genes in yeast, but not in mosquitoes," Stephanie explains, "so we couldn't just apply that same method to mosquitoes."

Nevertheless, the fundamental problem was the same, so it made sense to try a similar approach – it just had to be one that would work in mosquitoes. Although you can't knock genes out in mosquitoes, you can knock them down. Rather than altering an organism's DNA sequence as they would in a gene knockout, scientists knock a gene down by using specially engineered RNA molecules to interfere with the gene's expression, i.e. to hinder the production of the protein that the gene encodes. Lars, Stephanie and Rui took advantage of technological advances in this field of RNA interference and created a new method that essentially uses this knockdown approach in the same way that the previous method used gene deletions. This meant devising a way to knock down not just a specific gene, but a specific allele of that gene.

To test their new technique – called reciprocal allele-specific RNA interference – Stephanie and her colleagues at INSERM produced mosquitoes that each carried two different alleles of a gene called TEP1: one allele from the strain of malaria-resistant mosquitoes, and the other from a strain of susceptible mosquitoes. They then individually knocked down each of the alleles, creating mosquitoes whose only functional version of TEP1 was either the 'resistance' or the 'susceptibility' allele.



More parasites survived (fluorescent green dots) in the midgut of a mosquito with only the 'susceptibility' allele turned on (right) than in a genetically identical mosquito with only the 'resistance' allele turned on (left), which contained mainly dead parasites (black dots).

When they compared the ability of these otherwise identical mosquitoes to fight off *Plasmodium*, the scientists found that the immune systems of mosquitoes whose only functional TEP1 allele came from the resistant strain were able to kill *Plasmodium* parasites, whereas in mosquitoes with only the 'susceptibility' allele, parasite survival was much higher. In short, silencing different alleles produced mosquitoes with different degrees of resistance to malaria, meaning that an individual mosquito's resistance to the *Plasmodium* parasite depends largely on which form(s) of this one gene it carries.

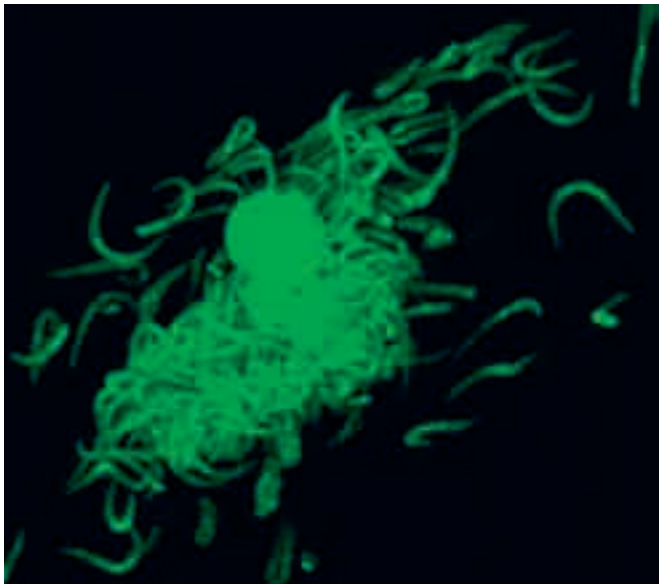
The team didn't choose TEP1 at random. Previous work by Stephanie and others in the lab of former EMBL Director General Fotis Kafatos had shown that this gene encodes a protein that attaches itself to malaria parasites in the mosquito's gut, so that they are then eliminated by the mosquito's immune system. Nevertheless, Lars points out, "If you consider all the complexity that could have been going on, we were lucky – not only did our new technique work, but the biology fell into place too: we'd picked the right gene!"

Our investigators did what all those detectives in mystery stories dream of: they devised a way to efficiently narrow down the suspect list and identify the culprit. Crucially, scientists will now be able to do the same in many other situations, as this method is applicable to many different species and cell types, including human cell lines in culture. "Along with other advances like the ability to induce pluripotent stem cells, this kind of technique really opens up the door to understanding the function of individual alleles," Lars says, "and to do this in a single individual." In this respect, the new technique is complementary to pre-existing approaches, in which scientists had to scour genetic data from many individuals. "And you can apply it to any organism in which you can do RNA interference, which is a huge expansion of the method we developed for yeast," Lars emphasises.

In the meantime, as far as mosquitoes' resistance is concerned, Stephanie and colleagues in the malaria field have a culprit: TEP1. This culprit may, however, have accomplices, so the scientists would like to investigate the roles of other mosquito genes in increasing – or decreasing – an individual mosquitoes' ability to defeat the malaria parasite.

"Not only did our new technique work, but the biology fell into place too: we'd picked the right gene!"

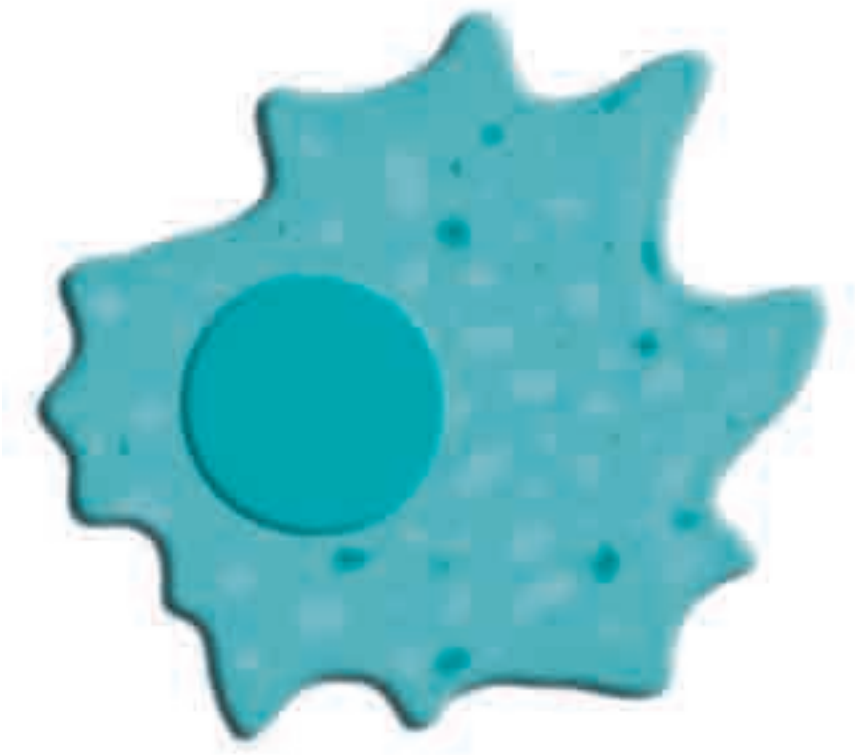
They are also investigating human malaria, for which evidence seems to point to the same suspect. If TEP1 enables mosquitoes to fight the parasites that cause human malaria as it does for the rodent malaria parasite, this could prove to be a boon to malaria eradication programmes, a large part of which focus on eliminating the mosquitoes that transmit the disease. "Knowing which allele or alleles confer resistance to malaria could help to make mosquito eradication programmes more effective," Stephanie posits. If scientists can identify the alleles responsible, and distinguish between resistant and susceptible mosquitoes in the wild, such programmes could be restricted to areas where they are most necessary: areas where most mosquitoes are susceptible, and therefore likely to carry the disease. Thus, in a plot twist that's reminiscent of the best detective stories, this winged enemy may, in fact, turn out to be an ally.



Approximately 12 days after infection, a Plasmodium *oocyst is bursting open, releasing thousands of developing parasites (labeled green) into the mosquito's bloodstream.*

Blandin SA, Wang-Sattler R, Lamacchia M, Gagneur J, Lycett G, Ning Y, Levashina EA, Steinmetz LM (2009) Dissecting the genetic basis of resistance to malaria parasites in *Anopheles gambiae*. *Science* **326**: 147-150

Poles apart



The discovery of how certain immune cells make a Jekyll and Hyde-style switch in their behaviour is offering new insights into how the body repairs damaged muscle tissue, say researchers at EMBL Monterotondo. The teams, led by Claus Nerlov and Nadia Rosenthal, have unearthed the genetic mechanism that allows immune cells to change from a state in which they boost the body's defence mechanisms to another that promotes healing and repair. The researchers hope the work paves the way for new treatments for muscles damaged by disease, injury or surgery.

When a tissue is damaged, the body has to strike a delicate balance between cleaning up the debris and fending off infection, and fostering regeneration and repair. If there is too much cleaning and defence, the tissue can't regenerate well and will tend to scar; if this process is impaired the tissue can get infected and cells die. Both cleaning/defence and regeneration are needed, in sequence and under strict control. But some disease conditions disrupt this regulation, which delays healing and can cause further damage.

One of the key players in tissue healing is an immune cell called a macrophage. Macrophages crawl around tissues like an army of biological Pac-men, gobbling up microbes and fragments of dead and dying cells. They also release chemical signals that promote inflammation, a condition that attracts other immune cells to the injury and primes them to attack invaders.

In recent years, scientists have found evidence to suggest that macrophages can also promote tissue healing by dramatically switching their characteristics to become a cell type called an M2 macrophage. M2 cells release chemical signals that damp down inflammation, thus promoting regeneration. This shift is called macrophage polarisation but until now, scientists knew little about how it was triggered. Thanks to a stroke of serendipity, the EMBL teams now have an answer.

Claus and his group were studying a protein called C/EBPβ, which is produced by tissues in response to inflammation. They had altered the gene for C/EBPβ in mice so that it could no longer respond to inflammation, to see what effect the lack of C/EBPβ would have on immune cell development. Disappointingly, the answer seemed to be very little: the mice developed normal immune cells. Intriguingly, however, their macrophages failed to polarise when challenged with inflammatory signals.

This finding caught the attention of Nadia, who wondered whether it might be relevant to her group's work on muscle regeneration. She and Claus teamed up to see how his mice responded to muscle injury. Sure enough, the mice's macrophages could clear up the damage debris, but they couldn't make the all-important M2 switch and their muscles failed to regenerate properly. "If macrophages don't make this switch, then the muscle won't repair itself," says Nadia. "You just end up with scar, instead of new tissue."

Blocking the C/EBPβ gene in muscles only, but not in macrophages, had no effect, showing that macrophage polarisation is indeed essential for regeneration. Just like the potion Dr Jekyll mixed to transform into Mr Hyde, a cocktail of chemical signals in inflamed tissue acts on the C/EBPβ gene, triggering it to switch the cell into an M2 macrophage. This, say the teams, makes the gene and the other genes it controls attractive targets for therapies. "From a medical point of view, it would seem that the trick to improve muscle repair is finding a way to increase C/EBPβ production and keep it high," says Claus.

Ruffell D, Mourkioti F, Gambardella A, Kirstetter P, Lopez R, Rosenthal N, Nerlov C (2009) A CREB-C/EBPβ cascade induces M2 macrophage-specific gene expression and promotes muscle injury repair. *Proc Natl Acad Sci USA* **106**: 17475-17480

Take a deep breath

Emphysema is a very nasty disease. It causes the walls of alveoli – the tiny air sacks of the lungs responsible for taking up oxygen from the air – to irreversibly break down. Without this support, parts of the lung begin to collapse. Sufferers find it increasingly hard to breathe, until eventually they just can't take up enough oxygen to survive. Most patients are smokers. Indeed, smoking results in a 90% increase in the risk of developing emphysema and other kinds of chronic inflammatory lung diseases, which together constitute the fourth leading cause of death and disability in developed countries.

Our body responds to cigarette smoke and other irritants in the lungs by activating immune cells, such as macrophages, to seek out and destroy the foreign material. To move around, these macrophages secrete a protein-digesting enzyme called MMP12 to break down the complex network of proteins and fibres that surround and support the cells of the body. However, continued exposure to irritants over-stimulates the macrophages causing them to produce too much MMP12, which builds up over time and damages the delicate structure of the lungs. Being able to measure the activity of MMP12, therefore, would give an early indication of the amount of damage to be expected.

With this in mind, Carsten Schultz and his group at EMBL Heidelberg, in collaboration with researchers of Marcus Mall's group from the University Clinic Heidelberg as part of the Molecular Medicine Partnership Unit (MMPU), have developed a tool that could be used for the diagnosis of emphysema and similar diseases. They designed a special fluorescent probe that measures MMP12 activity in macrophages: if there is MMP12 activity, the probe is taken up by the macrophages causing them to fluoresce brightly. They demonstrated that the method could work by testing samples of lung tissue from mice with acute lung inflammation, which showed a distinct rise in macrophage MMP12 activity compared with tissue from healthy mice. The design of new functional molecules from scratch is one of the main goals in Chemical Biology, which draws together expertise in chemistry, molecular biology and imaging techniques to make new tools for studying cells and how they function.

As Carsten explains, "It's a very powerful test, because we can actually measure MMP12 activity rather than just abundance of the enzyme. Therefore the readout really reflects what's actually going on inside the lungs. At the moment, the test can be used on samples isolated by lavage, an invasive process where cells are washed out of patients' lungs. However, it should one day be possible simply to use sputum samples. This way we could easily use MMP12 and similar enzymes as biomarkers to monitor the risk of emphysema formation, the progression of the disease, and to check on the response to therapeutic interventions."

This is particularly important because, although the damage caused by emphysema cannot be repaired, careful management of the condition improves the quality of life of these patients. It is therefore vital to monitor the disease so that treatment can be better tailored to patients' needs.

The researchers hope that the new technology they have developed can also be applied to test other enzymes involved in lung inflammation. And with a clearer picture of the processes underlying these diseases, future treatments should be more specific and side-effects reduced.

Cobos-Correa A, Trojanek J, Diemer S, Mall M, Schultz C (2009) Membrane-bound FRET probe visualizes MMP12 activity in pulmonary inflammation. *Nat Chem Biol* **5:** 628-630


José Márquez

Surviving drought

We've all felt it: a quickening of the heart and a slight shortness of breath as you walk into an exam room. Most of us recognise that the hormone adrenaline is responsible for this reaction, but we're not unique in responding to stress with a release of hormones. Plants do this too – but unlike you and I, they don't have the option to flee; rooted to the spot, they can only stay and fight it out. To do this, plants release the hormone abscisic acid (ABA), which coordinates their response to stresses such as drought, extreme temperature and high salt levels.

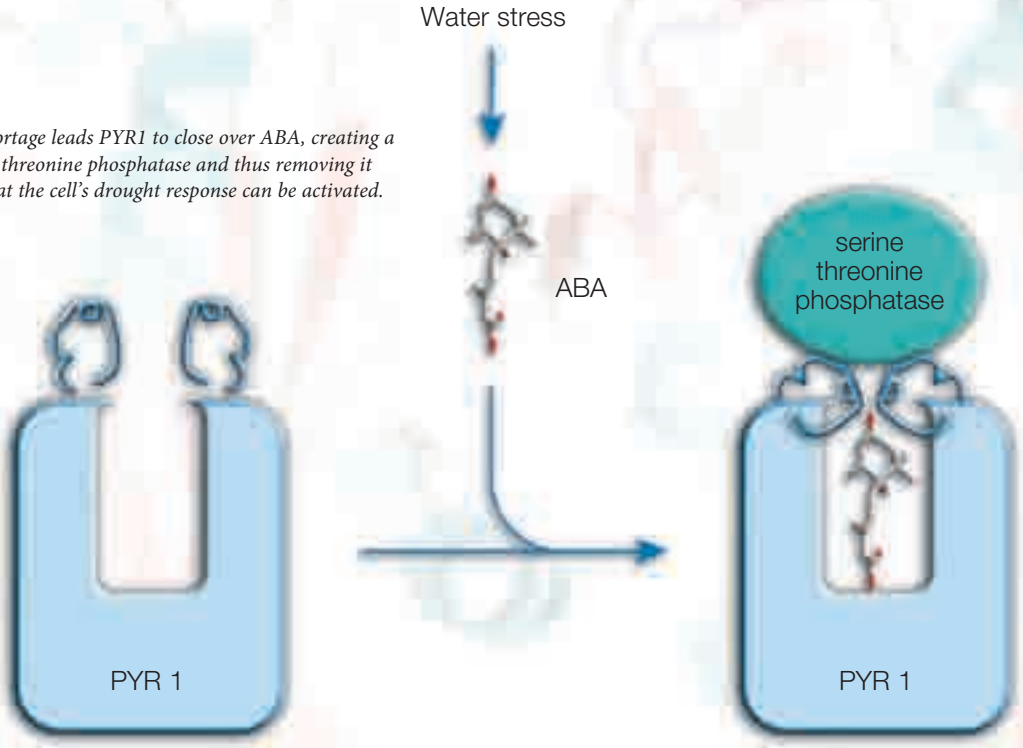
ABA acts as a chemical courier, relaying messages from one cell to another. Cells respond to the hormone if they possess a receptor, which, once bound to the hormone, signals to the cell to go on the offensive. For plants, this means closing the tiny holes in their leaves to avoid water loss, diverting resources to their roots to increase water uptake and switching on the production of proteins that protect cells from dehydration.

Understanding how plants transmit these water-saving measures to their roots and leaves could help farmers ward off the effects of drought in crops, and has prompted many scientists to search for the receptor that binds ABA. But identifying this receptor has proved unusually challenging.

Since 2006, several proteins have been suggested but, because of conflicting findings, their exact role remained controversial. Then in 2009, two groups independently homed in on yet another potential ABA receptor , PYR1, and a family of proteins that bound ABA was identified, although it was still not clear how these proteins interacted with the hormone. Given the struggles of the past three years, some plant biologists reserved final judgment on whether the hunt for the ABA receptor was over until they could visualise the receptor physically interacting with the hormone. This is where plant biologist Pedro Luis Rodriguez from the Universidad Politecnica de Valencia in Spain stepped in, approaching structural biologist José Márquez from EMBL Grenoble with a question: would it be possible to obtain the structure of the ABA receptor?

"For such a contentious field, knowing the structure is the final proof, because you not only see the receptors but you also see how they bind the hormone," explains José, who, together with his team, examines protein structures to understand how they transmit messages. "Because of our interest in signalling, we were very keen to collaborate on this project."

But José and Pedro were not alone in finding this an interesting puzzle to solve. Four other groups – two in Asia and two in the US – simultaneously set out to define the structure of the receptor. And so the race began. The first step for José's group was to produce vast quantities of the ABA receptor in bacteria before adding the protein to a cocktail of chemicals to try to induce it to form a crystal. José explains that the proteins usually fall to the bottom of the flask in an amorphous blob, but occasionally they stack together in a regular fashion. This is what they were searching for — crystals.

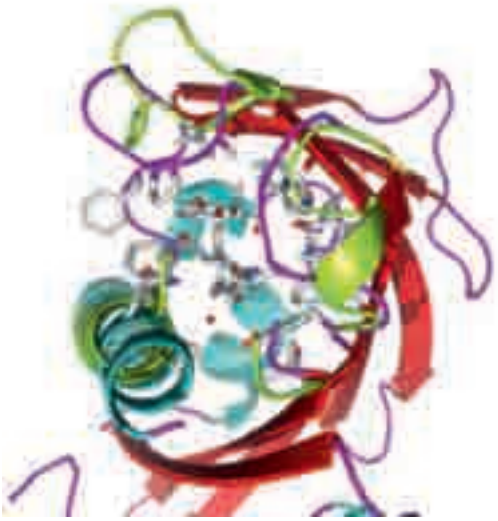


The stress of water shortage leads PYR1 to close over ABA, creating a docking site for serine threonine phosphatase and thus removing it from circulation so that the cell's drought response can be activated.

To increase the likelihood of crystal formation, they set up hundreds of experiments, each one differing slightly in the chemicals added. "Growing crystals can take days, weeks, or even months," says José, adding that each flask needs to be checked every few days. "So as you can see, with all these samples you would quickly run into problems."

Fortunately, José's high-throughput crystallisation facility was on hand to help solve the problem. Using robots to dispense and regularly check close to 3000 samples, it was only a matter of weeks before the researchers found a crystal. "This was exceptionally fast and shows the power of this type of collaboration," says Pedro. Another advantage of automating the crystal-forming process is that less of the protein is wasted because of the accuracy with which the robots can dispense the samples. "They can dispense one tenth of a millionth of a litre – a quantity that humans would find impossible to repeatedly pipette," explains Pedro. By reducing protein waste, the researchers could set up more experiments and so were more likely to secure the perfect cocktail of chemicals in which to grow the crystal.

The structure of PYR1 (coloured ribbons) in its open, unbound state (light green loops) and how it folds around ABA (white rods) when it binds to this hormone (turquoise and purple loops).



With the crystal in hand, José's group used the powerful X-ray beams of the European Synchrotron Radiation Facility in Grenoble to determine the structure of the crystalline ABA receptor (for an overview of this technique, see page 56). Ahead of the other international teams, and in collaboration

After 15 days of drought, an Arabidopsis thaliana *plant will normally be withered and dry (far left), unless it has been genetically engineered to enhance its response to ABA (centre left, centre right and right).*

with Adam Round from EMBL Grenoble who is involved in the Partnership for Structural Biology, they revealed that the ABA receptor acts in pairs: two copies of the protein bind together, each side creating a pocket into which ABA slots. Guarding the entrance to these pockets, they found long flexible loops, which close like a lid over the hormone trapped inside.

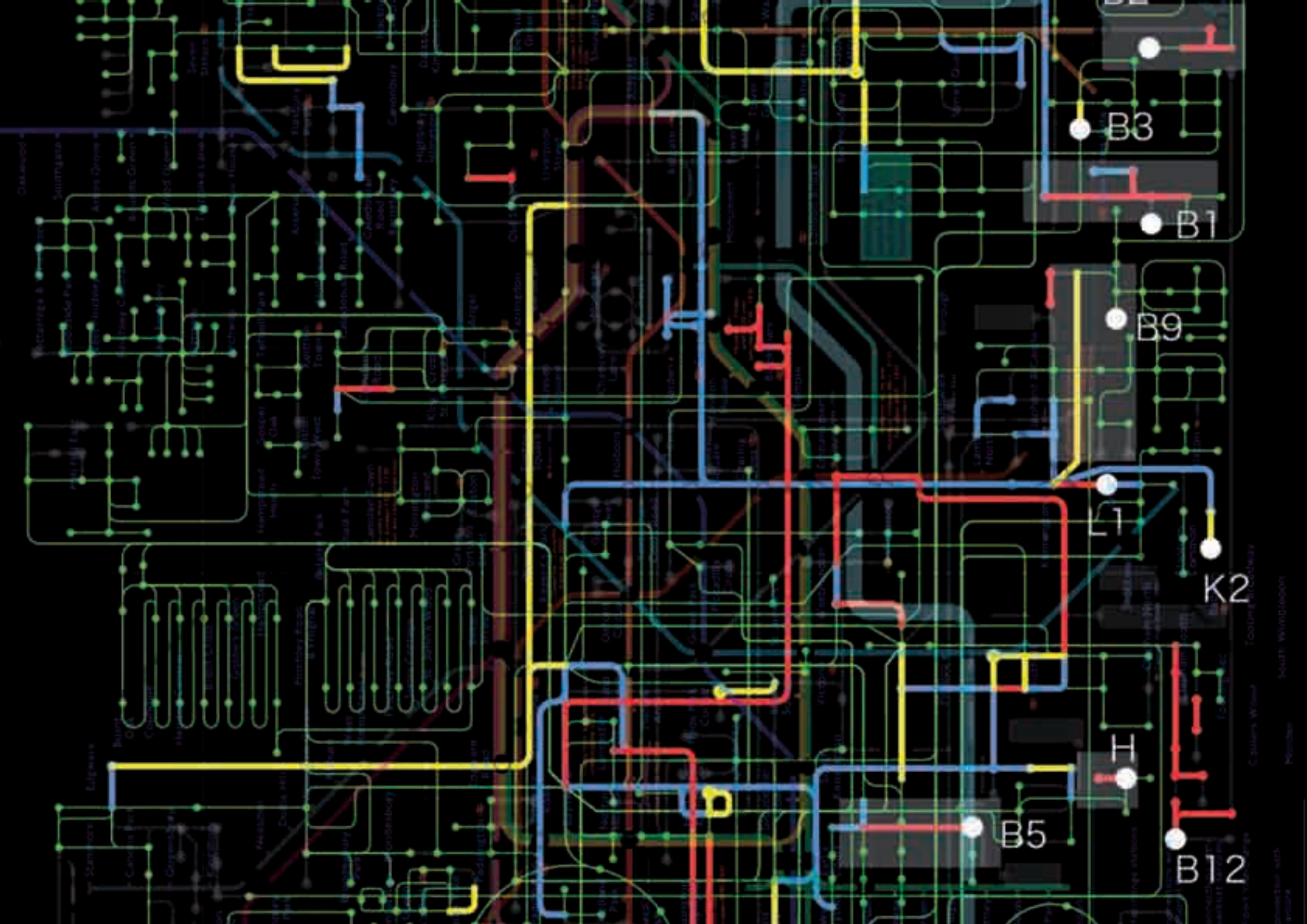
This closing mechanism led to an unexpected discovery. José's group realised that the 'closing lid' could create a docking site for an enzyme called serine threonine phosphatase, which is known for having the capacity to suppress the stress-signalling pathway. But when locked in place on the docking site, it can no longer prevent the series of signals cascading through the plant that prepare it for the stress of losing water. "The revelation came when we superimposed the receptor bound to ABA with the unbound form, and saw the difference in the loops," says José. He thinks that when the enzyme binds at the docking site, this inactivates it and prevents it from circulating freely. This then allows another enzyme, a protein kinase, to activate transcription factors that turn on the stress-response genes.

José hopes that this finding could help alleviate crop damage in drought-prone areas. If scientists find a molecule that mimics ABA – which is too expensive to manufacture in large quantities – they could simulate the stress response in plants to help protect them against water loss. The new molecule could be sprayed over crops before a period of dry weather, stimulating them to save water for the harsh times to come.

This is something that José's group has already started to do using their structural models. José explains: "We're trying to fit millions of different molecules into the receptor's pocket in the hope of getting one to fit."

If this works, it could have far-reaching effects for all of us, because most of our food is grown in semi-arid regions such as North Africa and South Asia, which suffer recursive droughts. These regions see huge quantities of food go to waste because of damage caused by unusually dry periods within the growing season. Reducing this waste would particularly benefit people in developing countries who typically sustain themselves on cheaper, more traditional crop varieties that are less drought-resistant. Who'd have thought stress could be a good thing?

Santiago J, Dupeux F, Round A, Antoni R, Park S-Y, Jamin M, Cutler S, Rodriguez P, Márquez J (2009) The abscisic acid receptor PYR1 in complex with abscisic acid. *Nature* **462**: 665-668


B3
B1
B9
L1
K2
H
B5
B12

Getting around and staying in touch

Had anyone been tempted of late to underestimate the importance of a working transport network, the events of April 2010 soon put them straight. Icelandic volcano Eyjafjallajökull belched countless tonnes of ash into the atmosphere, shutting down air travel across most of Europe for a week. Millions of passengers were stranded, airline phone systems jammed by frustrated customers, and trains, coaches and ferries overwhelmed by the sheer number of people struggling to get back home. Airlines lost millions in revenue daily, businesses were forced to close. The events brought home just how much our own livelihoods, and those of our countries and cities, depends on being able to travel and keep in touch.

Like any city or economy, our cells depend on effective transport and communications systems to survive. Proteins and other molecules need to be manufactured and freighted to their destinations. Cells need to communicate with each other to monitor their external environments and work out how to respond. They also need to co-ordinate their responses to the outside world with what is happening inside them. If a cell experiences the biochemical equivalent of Eyjafjallajökull – perhaps a mutation that paralyses its signalling network – serious diseases such as cancer can result.

For example, the internal scaffolding of a cell, its cytoskeleton, acts rather like London's Underground train network. Researchers at EMBL Heidelberg have shown how this allows molecules to commute to the appropriate cellular location to perform their jobs correctly. Just as Paris is divided into districts such as the business centre La Défense or the artistic area of Montmartre, so the cell is divided up into compartments with different functions. EMBL Director General Iain Mattaj and his team are tracing the earliest evolutionary origins of these compartments, like archaeologists sifting through clues at the ancient city of Ur.

Communication within and between cells is also vital to keep hearts beating, create memories, control growth and for development. EMBL scientists have uncovered the detailed workings of one such communication system, opening the way for new drugs to treat disease. Other EMBL teams, meanwhile, have applied the city analogy one step further – to biologists themselves. Thanks to this work, life scientists now have their own communication framework, so they can clearly and effectively unpick the daily hubbub of a busy living system.


Christoph Müller

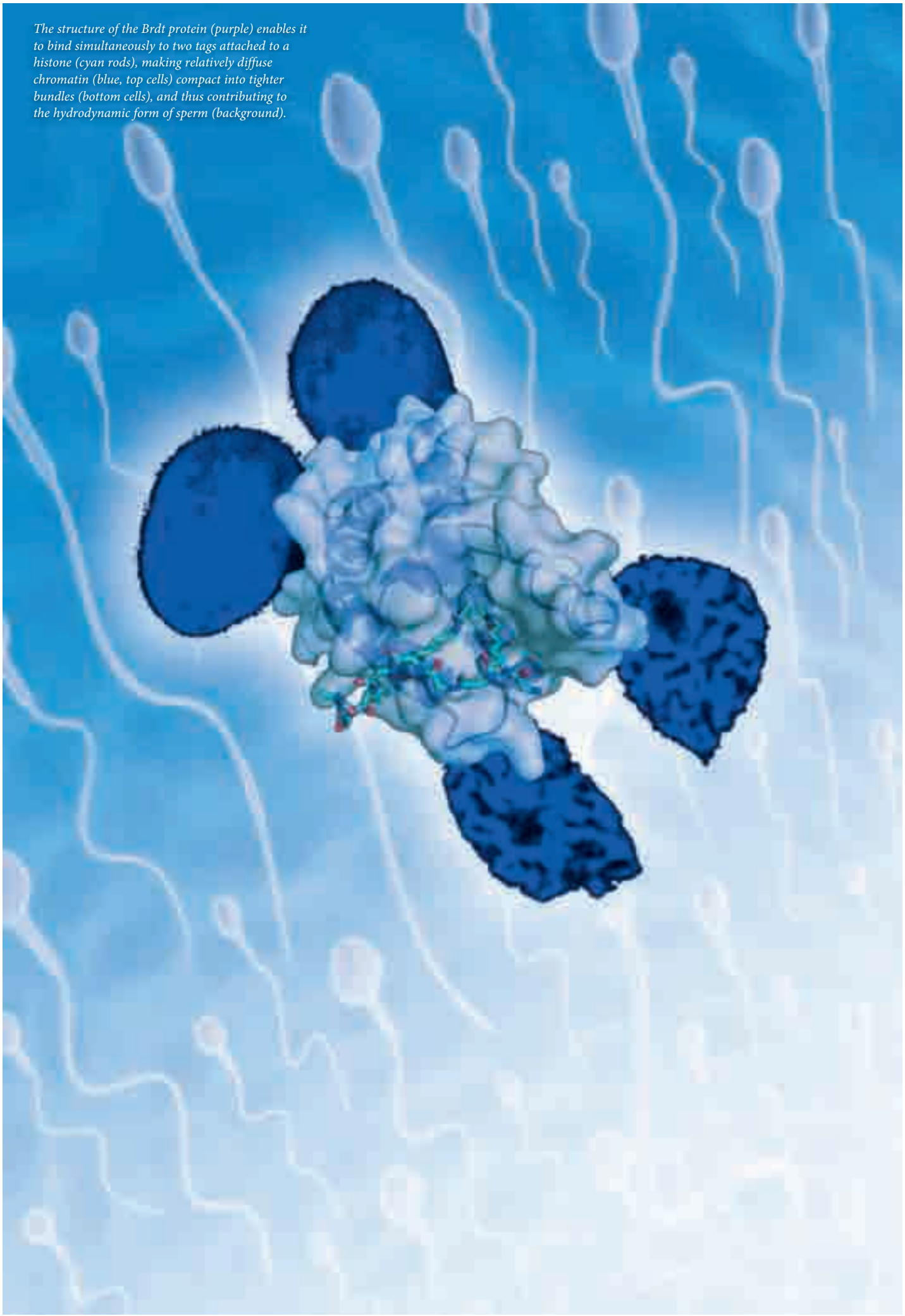
Built for speed

The launch of the Speedo LZR Racer swimsuit in February 2008 started a revolution in the swimming pool. With swimmers squeezed into figure-hugging full-length bodysuits in gun-metal black, world records were toppled across the board. Two years later, criticised as 'technological doping' and accused of devaluing honest competition, the controversial polyurethane supersuits have been banned. So extreme was the difference they made to the performance of the athletes that, if they are allowed to stand, it is unclear whether the suit-assisted records will ever be beaten. Aside from the special low-drag, water-repellent properties of their polyurethane panels, these suits worked by compressing the swimmer's body, making as compact and streamlined a silhouette as possible. Evidently, the key to speed in the water is compaction.

In the body, sperm are the only cells that can swim. Only the fastest wins the chance to pass on their genes to the next generation, and so sperm have evolved to become exceptionally streamlined. For them too, this is a matter of compaction. As sperm have no other purpose than carrying the father's DNA, the sperm head contains very little else: all non-essential cellular components are squeezed out during production, leaving only the DNA-containing nucleus. In the pursuit of hydrodynamics, the only thing left to do is to ensure that the DNA itself takes up as little space as possible.

To do this, sperm DNA undergoes drastic re-packaging during spermatogenesis. Under normal circumstances, DNA – an incredibly long and unwieldy molecule – is wound around proteins called histones to form nucleosomes that are then organised into an even more complex structure called chromatin. However, in developing sperm cells, most of these histones are removed and replaced by smaller proteins that wrap the DNA up more tightly, shrinking and reshaping the nucleus and streamlining the sperm head. This is an incredibly dramatic process that must be tightly controlled. It remains unclear how this is orchestrated, but like other changes in chromatin structure, it is directed by histone modifications. These modifications take the form of molecular tags, which are chemically attached to the tail portion of histone proteins and which flag up regions of the genome for the attention of chromatin remodelling machines. During spermatogenesis, histones throughout the genome are modified in several huge waves, each wave adding a different kind of tag. The very last of these is a wave of acetylation during which many acetyl groups are attached specifically to the tails of one type of histone, histone H4. Specific proteins recognise these tags, bind to them, and trigger the massive changes that bring about compaction of the DNA.

This process is the reason why Christoph Müller, a structural biologist who studies transcriptional regulation, became involved in a story about sperm development. At the time, he and his lab were based at EMBL Grenoble where they were working to understand how chromatin and DNA-binding proteins influence patterns of gene expression. Christoph was approached by Saadi Khochbin, a group leader at the Institut Albert Bonniot also based in Grenoble, who is primarily interested in the



The structure of the Brdt protein (purple) enables it to bind simultaneously to two tags attached to a histone (cyan rods), making relatively diffuse chromatin (blue, top cells) compact into tighter bundles (bottom cells), and thus contributing to the hydrodynamic form of sperm (background).

role that chromatin plays in sperm development and fertility. Presuming that the wave of histone H4 acetylation prior to sperm DNA compaction must be a crucial step in the process, Saadi's group had gone on a hunt for sperm proteins that contain so-called 'bromodomains' – protein modules that bind to acetyl tags. After testing a few candidates, they hit on Brdt, a unique protein found exclusively in developing sperm cells that seemed to be crucial for compaction. Saadi proposed a collaboration to study the structure of Brdt's two bromodomains and Christoph agreed, initially incorporating them into a larger screen to study bromodomains from many different histone-binding proteins. However, through the work of Jeanne Morinière, a pre-doc in Christoph's lab, it quickly became clear that the Brdt bromodomains were the most interesting in the screen and so they became the focus of project.

"The tags are like letters in an alphabet. Used individually they don't convey much meaning. But, combined and read out as words, you suddenly find there's a lot more you can say."

With some interesting preliminary structural data the groups realised that to really understand how the Brdt bromodomains worked, they would need to determine their preferred targets. So they started testing the binding of histone tails with different combinations of acetyl tags attached to the Brdt bromodomains. To their surprise, it seemed that one of the Brdt bromodomains favoured a histone tail with two acetyl tags rather than just one.

The path of a good project never runs as smoothly as its final publication would suggest, so with Jeanne leaving the lab having completed her PhD and Christoph relocating the rest of the lab to Heidelberg, structural analysis of the Brdt bromodomain was handed over to Carlo Petosa. Formerly a staff scientist in Christoph's group, Carlo stayed on in Grenoble, becoming an independent group leader at the Institut de Biologie Structurale (IBS). After a lot of hard work and analysis, Carlo was able to piece together a complete picture, showing that the Brdt bromodomain forms a pocket that binds to both tags at once.

But why was this so unexpected? The explanation leads us into the history of the field of chromatin research. The concept that histone tags might act as a code to which different 'interpreter' proteins bind to induce changes in chromatin structure was only proposed about 10 years ago. Since then, the details have grown increasingly more elaborate. Different protein domains – functional protein sub-units preserved by evolution and re-used in different proteins – have been identified that bind to different kinds of tags: for example, bromodomains for acetyl tags and chromodomains for methyl tags.

"We knew that individual proteins could bind either to single tags via one tag-binging domain, or sometimes to several tags at once via multiple domains. But it was just assumed that each domain would only bind to one tag," explains Carlo. "So finding a single bromodomain able to bind two tags simultaneously was quite a revelation."

On re-examining the structures of other bromodomains in the protein databases they now think that the ability to bind two tags is not restricted to just Brdt – several other bromodomain proteins probably work the same way.

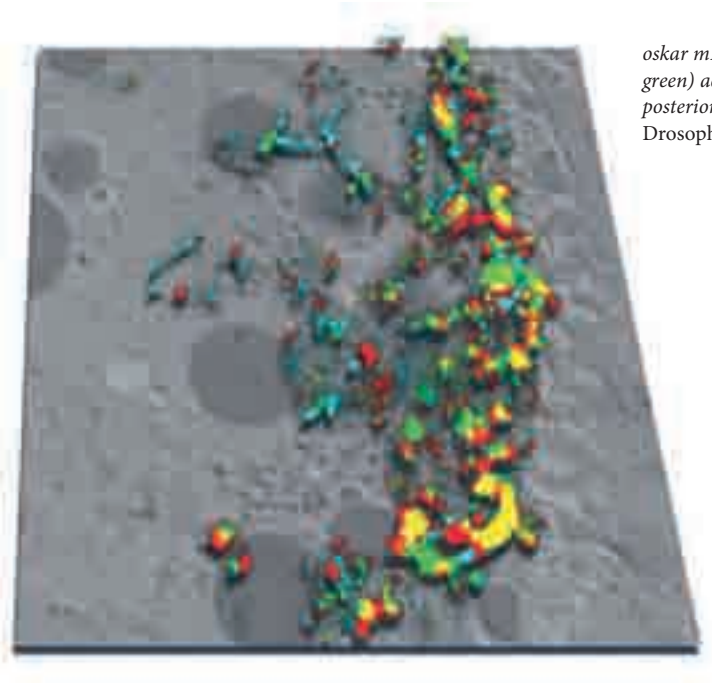
"This finding adds an extra level of complexity to the interpretation of the histone code," explains Christoph. "In a way, it's like any other language. The tags are like letters in an alphabet. Used individually they don't convey much meaning. But, combined and read out as words, you suddenly find there's a lot more you can say. If single proteins or even single protein domains can read out combinations of tags, the histone code can direct a greater variety of specific chromatin remodelling responses."

But what does all this have to do with sperm DNA compaction? How could the special way in which Brdt binds to histones be important for this unique role? "Well, we can speculate," says Christoph. "Compaction of the DNA only begins when the wave of acetylation is complete and H4 histones are fully tagged. The acetyl groups are added to each histone sequentially and Brdt binds to the last two tags in this sequence, making Brdt binding the very last step in the process – perhaps the final signal for compaction to begin."

From this point on the compaction process seems to unfold automatically. The DNA throws off its bulky histones, replacing them with smaller proteins, and tightly winds itself into a dense, compact package. Completing the rest of their development, the slimmed-down and streamlined sperm are now race-fit and ready. A little wiser to the secrets of speedy sperm, the researchers are now trying to find out whether defects in Brdt might be a cause of male infertility.

Morinière J, Rousseaux S, Steuerwald U, Soler-López M, Curtet S, Vitte A-L, Govin J, Gaucher J, Sadoul K, Hart D, Krijgsveld J, Khochbin S, Müller C, Petosa C (2009) Cooperative binding of two acetylation marks on a histone tail by a single bromodomain. *Nature* **461**: 664-668

Getting on the right track



oskar mRNA (fluorescent green) accumulates in the posterior pole of a Drosophila *oocyte.*

Travelling on Paris' metro system can be confusing at the best of times. Each of its lines has two tracks that run in opposite directions and it is all too easy to hurriedly jump on the wrong train and end up at say, the Bastille, when you really wanted to be at the Opéra. Now, Anne Ephrussi and her team at EMBL Heidelberg have discovered how certain molecules perform the cellular equivalent of picking the right metro track to reach the right destination inside the cell.

A protein's cellular location is vital: if it ends up in the wrong place, its misplaced activity can cause developmental defects or disease. Localised proteins have prominent roles in a number of cell functions, such as in the connections between neurons that enable learning and memory.

Embryos also rely on proteins being in the right place for their proper development. A developing fruit fly embryo, for example, relies on proteins localised at either extremity of the egg from which it develops to determine where its head and tail should be. One of these proteins, Oskar, is the focus of Anne and her team, who have been working to understand the cellular mechanisms that ensure that the protein ends up only at the hind end of the egg, where it directs the development of the embryo's posterior structures.

The process begins in the nucleus, which houses the cell's DNA, the instructions for making proteins. First, the cell transcribes the DNA into its chemical relative, called RNA. The RNA then leaves the nucleus and enters the surrounding cytoplasm. Here, molecular machines called ribosomes read the RNA instructions to make the protein in a process called translation.

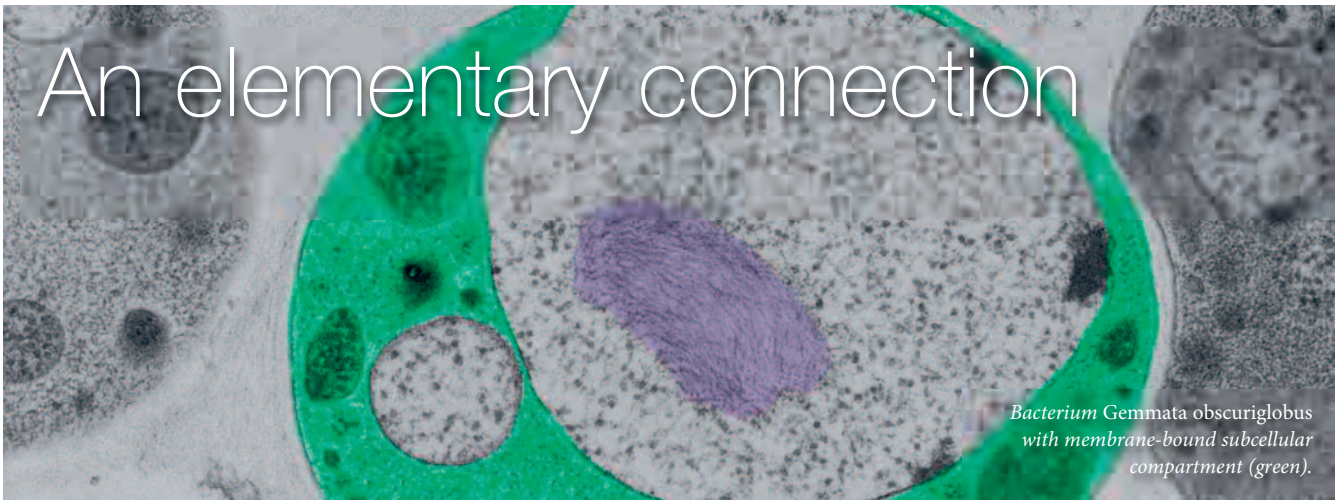
Biologists now know that many RNAs are carried to where they are needed in the cell before they are translated. To understand how *oskar* RNA molecules are transported, Anne and her team turned to a new technique that allowed them to use an electron microscope to analyse the transport of the RNA in unprecedented detail.

oskar RNA isn't actually made in the egg, but in a small group of connected cells called nurse cells. Anne's team showed that *oskar* RNA becomes associated with certain proteins in the nurse cell to form particles called ribonucleoproteins, or RNPs. The RNPs are then hooked up with two "motor" proteins called kinesin and dynein that have the ability to move up or down the cell's internal scaffolding, the cytoskeleton. The struts of the cytoskeleton span the length of the cell and like metro train tracks, they have a directionality to them. This directionality is read by the motor proteins: kinesin moves along the struts in one direction, dynein in the other. So the balance of motor protein types attached to an RNP will dictate its final location in the cell.

Anne's team found that the *oskar* RNPs use dynein to ride the nurse cell cytoskeleton to enter the developing egg cell and, once there, they mainly use kinesin to track along the egg cytoskeleton to reach its tail end. However, the RNPs have both dynein and kinesin on them at all times. "A key question now is how these opposing motors are regulated, such that the RNA can reach its correct location," says Anne. The team now hopes to build on this work to find out how these molecular commuters' journeys are integrated with their professions – in other words, how *oskar* transport is linked to its translation.

Trucco A, Gaspar I, Ephrussi A (2009) Assembly of endogenous *oskar* mRNA particles for motor-dependent transport in the *Drosophila* oocyte. *Cell* **139**: 983-998

An elementary connection



Bacterium Gemmata obscuriglobus *with membrane-bound subcellular compartment (green).*

Is it like anyone you know?' asked Holmes. He stood upon a chair, and, holding up the light in his left hand, he curved his right arm over the broad hat and round the long ringlets. 'Good heavens!' I cried in amazement. The face of Stapleton had sprung out of the canvas."

Family relationships can be hard to spot, as Sherlock Holmes and Dr Watson discovered in their famous case, *The Hound of the Baskervilles*. Only by looking at a portrait of a family's infamous ancestor, Sir Hugo Baskerville, in a new way were they able to spot a hidden family link to their prime suspect, Jack Stapleton.

EMBL Director General Iain Mattaj and postdoc Damien Devos recently performed a Holmes and Watson-style unmasking of another distant ancestral relationship, this time between bacteria and more complex eukaryotic cells such as ours. By taking a new approach to looking for similarities between proteins, they and their colleagues at EMBL Heidelberg confirmed that certain bacteria contain features previously thought to be unique to eukaryotes. This gives a surprising new insight into how eukaryotes may have originated.

A key difference between bacteria and eukaryotes is the presence of membranes inside eukaryotic cells. These "endomembranes" divide the cell into compartments with the help of membrane-coat proteins, which bend the membranes into the right shape. When researchers looked in bacteria for amino-acid sequences resembling those of eukaryotic membrane-coat proteins, they couldn't find any, suggesting that the genes for these proteins did not evolve before the origin of the eukaryotes.

But straightforward sequence comparisons don't tell the whole story. Thanks to evolution, sequences change over time, and similar three-dimensional protein structures can be encoded by a range of DNA sequences. So the retention of structural features between proteins from distant species can reveal their relatedness even if the DNA or amino-acid sequences have changed such that no relationship is detectable. "It's only through the structure that you can detect the similarities," says Damien.

So Damien decided to look for structures instead. Having previously found structural features characteristic of eukaryotic membrane-coat proteins, he used these as a starting point to search all known "proteomes" – the total collection of proteins produced by an organism – for signs of these structures. He was stunned to find signs of the proteins in a little-known group of bacterial species called Planctomycetes-Verrucomicrobia-Chlamydiae, or PVC, bacteria. "This is a very interesting group of organisms in terms of their relationship to other bacteria and to eukaryotes," says Iain. Unlike most other bacteria, some PVC bacteria have extensive internal membranes, hinting at some relationship with the eukaryote lineage.

To find out more, Rachel Mellwig, also from Iain's group, studied the new proteins in the bacteria and showed that they did indeed localise like membrane-coat proteins. This helps answer a long-standing mystery about the origin of eukaryotes: did they arise suddenly because bacteria fused with another kind of cell, the Archaea, or did they evolve their special features gradually, from a bacterium-like ancestor? The team's findings make the latter more likely. "Our results show that the gradual evolution of an internal membrane system is possible," says Damien. "This is an important but not the only part of eukaryotic evolution."

Indeed, it may be that membrane proteins are just the start of the story, and that using structural searches to unmask hidden relationships could yield many more surprises about our cells' origins. "Not many people have looked for relatedness to other eukaryotic characteristics using a similar approach," says Iain. "There may be a lot more evidence waiting out there."

Santarella-Mellwig R, Franke J, Jaedicke A, Gorjanacz M, Bauer U, Budd A, Mattaj IW, Devos DP (2010) The Compartmentalized Bacteria of the Planctomycetes-Verrucomicrobia-Chlamydiae Superphylum Have Membrane Coat-Like Proteins. *PLoS Biology* **8**

Matthias Wilmanns

Knowledge is strength

Big Brother is watching you. This famous, ominous phrase was coined by novelist George Orwell more than 60 years ago in his novel, *Nineteen Eighty-Four*. It told the story of a future dystopia where the government's figurehead, Big Brother, constantly monitored everyone's conversations via TV-like "telescreens". Worse, Big Brother dictated the vocabulary that citizens, or comrades, could use. By controlling their language, he controlled their thoughts and behaviour. It is a cautionary tale about the power of words.

So it might come as a surprise to learn that some biologists are trying to follow in Big Brother's footsteps. But this is not as sinister as it sounds. Instead of spying on people's conversations, they are trying to tap into the communication between the molecules in our cells. This communication is essential to life, and by working out what these molecules are saying to each other, these researchers hope to learn more about diseases such as cancer and neurodegeneration. And by controlling this communication with drugs – rather like Big Brother's control over conversations – they hope to find new treatments for such illnesses.

One such biologist is Matthias Wilmanns at EMBL Hamburg. Matthias and his team recently performed the molecular equivalent of installing a telescreen in the home of one particular protein called death-associated protein kinase, or DAPK. They have uncovered the structure of DAPK as it interacts with another key protein involved in cell regulation. This has allowed them to find out how these proteins 'talk' to each other, and how scientists might be able to intervene. "These kinases are a pretty hot target," says Matthias. The discovery is all the more important because DAPK is similar to several other vital signalling proteins in the cell. So the insights the team have gained into DAPK's function might also be applicable to these other molecules.

The idea of manipulating cellular signalling with drugs is tried and tested. The problem is that, although scientists know a lot about which proteins are involved in cell regulation, and which proteins interact with each other, they know relatively little about how most of these interactions occur. Furthermore, molecules in cells use many different mechanisms to communicate, in the same way as we may use face-to-face speech, mobiles, landline phones or e-mail to talk to each other. Scientists know more about some of these mechanisms than they do about others.

Matthias and his team decided to look at one of the lesser understood mechanisms, called calcium-dependent signalling. This mechanism is involved in many vital parts of our biology, such as controlling our heartbeat, muscle contractions and the communication between neurons in the brain. When an appropriate signal arrives at a cell – say a chemical signal from one neuron to another – it triggers a flood of calcium ions within the cell. This flood is detected by a protein called calmodulin, which becomes active when calcium ions bind to it. Calmodulin then turns this calcium signal into action by switching various parts of the cell's machinery on or off, thus altering the cell's behaviour.

X-RAY VISION

To understand the 3-dimensional structure of biological molecules and complexes, structural biologists have their own equivalent to Superman's X-ray vision: X-ray crystallography. This technique allows scientists to determine how molecules or atoms are arranged within a crystal, by exposing the crystal to X-rays and studying how it scatters them.

The atoms that make up a crystal are located at regular intervals, and when X-rays strike them, they tremble. This trembling eventually makes each atom emit X-rays with the same energy in all directions, in a process called elastic scattering. The waves of energy travelling outward from each atom interfere with those coming from other atoms, reinforcing or cancelling each other out depending on the angle at which they meet. The resulting pattern of diffracted waves, called a diffraction pattern (see image), can be recorded by placing a detector in front of the crystal. Every substance has its own unique diffraction pattern, depending on the identity and the position of the atoms within its crystals, so by analysing that pattern scientists can determine what atoms make up a particular crystal, and how they are positioned in relation to each other.

X-ray crystallography enables scientists to use this property of crystals to determine the 3 D structure of a molecule, or of a complex formed when several molecules bind to each other. To do so, they first induce the molecule to form crystals, then expose those crystals to an X-ray beam and record the angles and intensities of the X-rays that emanate from that sample. This diffraction pattern allows the researchers to produce a 3-dimensional model of the arrangement of atoms within the crystal. Finally, they use computer software to refine that model: the software calculates the diffraction pattern which the model structure would generate, compares it to the one the scientists obtained in their experiments, and alters the model until it fits those observations.

Diffraction pattern (middle), and ribbon and globular representations of protein structures (top and bottom, respectively).

This technique was initially developed by geologists interested in the arrangement of minerals, but ever since it was first employed to determine the structure of a protein in the late 1950s, molecular biologists have been turning to it regularly in order to peer into the intricacies of how life's molecules are arranged, and how they interact. The number of protein structures solved by X-ray crystallography is now over 50,000, and counting. Whether they are represented as globular masses or as coloured loops and ribbons, the models generated by this biological equivalent of Superman removing his glasses are still bringing scientists closer to solving such diverse mysteries as how cells communicate (see page 54), how plants survive drought (see page 42), and how bacteria evade our immune system (see page 91).

One of the main ways calmodulin does this is via a team of cellular errand boys called kinases, each of which controls the activity of specific parts of the cell's vastly complicated machinery. Kinases are an extremely important part of the cell's control mechanisms and there are about 600 of them encoded in the human genome. Some have already been successfully targeted with drugs such as Gleevec® (Imatinib), which is used to treat chronic myeloid leukaemia. Calmodulin is predicted to control about 10% of all the cell's kinases, but how it does so was unknown.

Matthias and his team wanted to find out how calmodulin interacted with its kinases. This meant that they needed to produce a three-dimensional molecular structure of calmodulin bound to one of its targets. To do this, they turned to a technique called X-ray crystallography, which involves mixing calmodulin and one of its kinases together to allow the proteins to interact, and then crystallising the resulting mixture. Shining X-rays on to the crystal allows the structure of the proteins to be deduced by the way the X-rays are scattered.

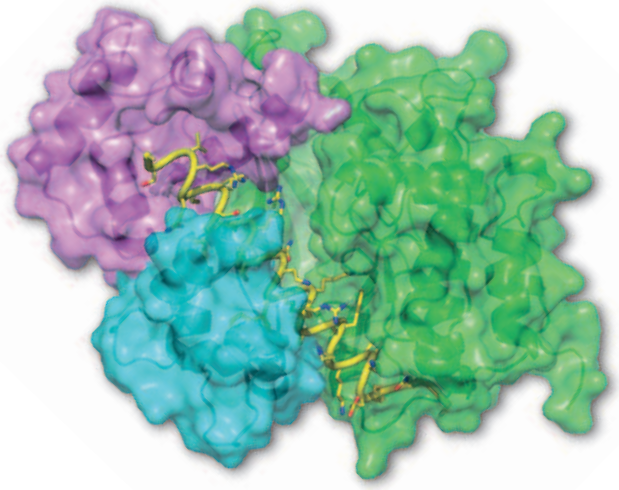
They began by working on a calmodulin-controlled kinase called titin. But try as they might, the team could not get titin and calmodulin to form crystals. Their luck changed when they were contacted by Adi Kimchi, a biologist at the Weizmann Institute in Israel. Adi's team was working on DAPK and they wanted to understand how calmodulin controlled its activity. Adi had heard about Matthias's work, and so offered to collaborate with him.

Matthias jumped at the chance. DAPK is interesting for two main reasons. One is that it is known to malfunction in a number of cancers, and so has potential uses in developing new therapies and diagnostic techniques. The other is that many calmodulin-controlled kinases share similar structural features. So finding out how calmodulin controls DAPK could give useful insights into how it controls these other targets. "Our motivation was to identify a prototype protein kinase," says Matthias. "Many things that are true of the prototype will also be true of other members of this kinase family."

No-one had ever produced a structure like this before, and with good reason. DAPK and calmodulin are large proteins, which makes it difficult to crystallise them both together. But the work was made possible thanks to Hamburg's high-throughput crystallisation facility run by team leader Jochen Müller-Dieckmann.. "It was critical for us," says Matthias. With the help of high-energy X-rays produced by the European Synchrotron Radiation Facility (ESRF) in Grenoble and the German Synchrotron Radiation Facility (DESY) in Hamburg, the team was able to deduce a structure.

This showed how calmodulin bound to a particular section of DAPK, switching the kinase on so that it could activate its targets. Intriguingly, the team's finding is different from earlier work in this area. Previous studies by other groups have looked at very short, synthetic sections, or peptides, of DAPK bound to calmodulin. Matthias's team compared one such peptide with the much larger DAPK fragment that they used, and showed that it behaved very differently. Without wanting to disparage this previous work, Matthias says his findings mean scientists may have to reconsider the results of this and similar experiments. They also show the value of being able to crystallise larger, more life-like protein fragments for study.

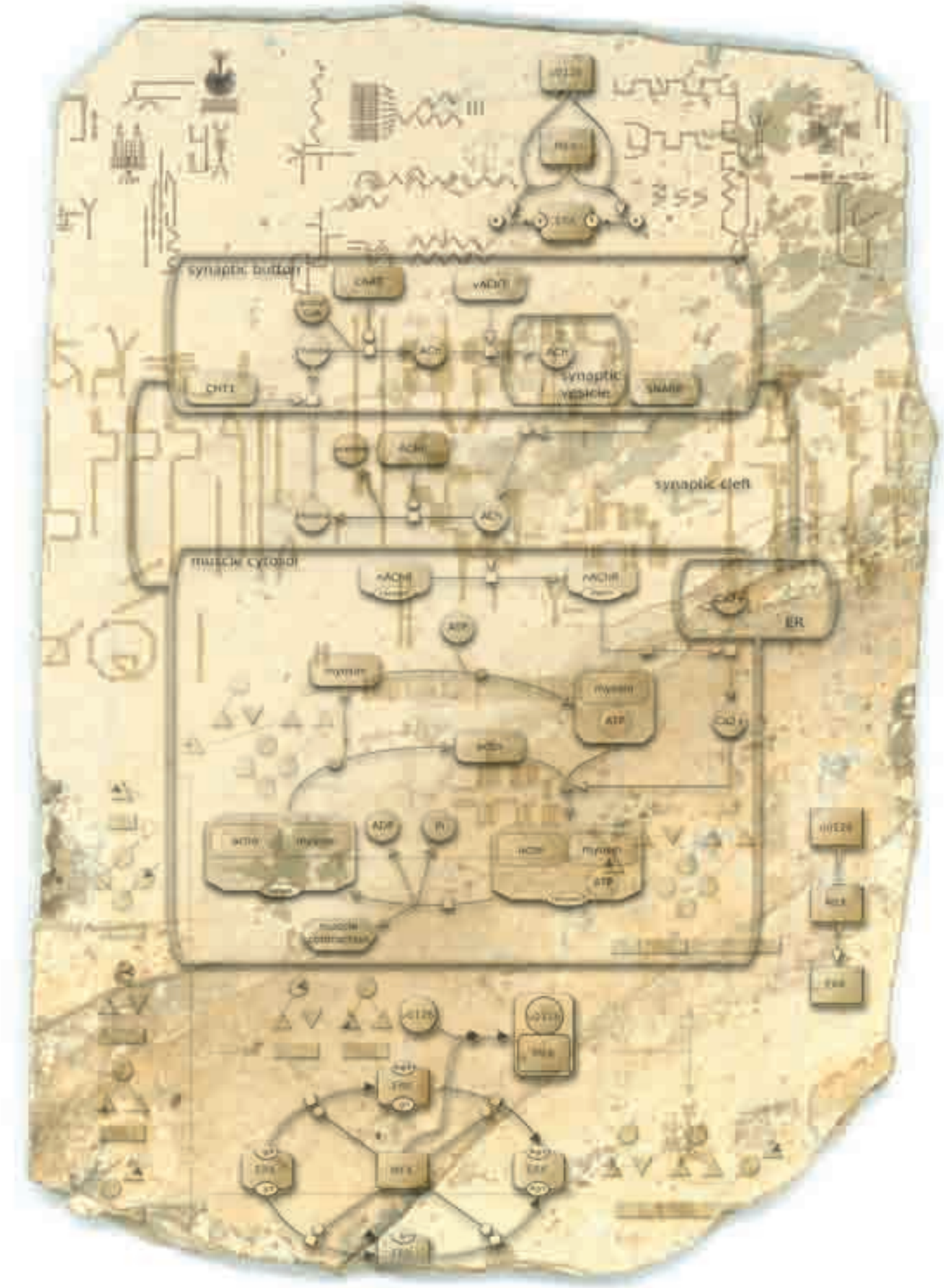
The structure of DAPK (green and yellow) bound to calmodulin (violet and blue).



Next, the team altered some of the building blocks, or amino acids, in their DAPK fragment to see which ones were needed for calmodulin to bind. They found one in particular that was vital for the interaction. "That will provide a platform to get into drug discovery," says Matthias. The work also gives biologists a starting point to study how calmodulin controls other kinases, although Matthias cautions that some gaps remain to be filled. For example, can DAPK adopt other structures to become active? And what other control mechanisms for these kinases remain to be found? "All these kinases have multiple layers of regulation," says Matthias.

He and his team are planning to explore these questions with the help of chemist Carsten Schultz at EMBL Heidelberg, as well as continuing the collaboration with Adi Kimchi. If all goes to plan, scientists may be able to both listen in on cellular conversations and control what is being said to help treat disease. Who knows? One day, the phrase "Big Brother is watching you" may be a comforting thought – from a cellular perspective at least.

de Diego I, Kuper J, Bakalova N, Kursula P, Wilmanns M (2010) Molecular Basis of the Death-Associated Protein Kinase–Calcium/Calmodulin Regulator Complex. *Science Sinaling* **3:** 106



From a multitude of graphical languages to a standard way of depicting biological processes.

Biology gets the picture

Imagine that you wanted to witness the creation of exciting new scientific ideas. Where would you go? A lab, perhaps? Possibly. There's no doubt that a huge amount of fantastic work gets done there. But if you really want to see the raw creative sparks flying then the best advice is to head for the research institution tea room. It's here, among the coffee stains and chocolate biscuit wrappers that animated arguments take place. Ideas are mooted, discussed and challenged. Diagrams are scribbled on paper napkins, cardboard plates and, occasionally, pieces of fruit. If you're in the tea room of an engineering department, chances are you could take these scribbles to any engineer anywhere and he or she would be able to understand them. The tea time scrawls of biologists are, however, much harder to decipher.

The problem is that, unlike engineers, biologists have no universal standardised way of depicting biological processes in a graphical form. So a set of symbols that make sense to one biologist could well be gibberish to another. Now, Nicolas Le Novère of EMBL-EBI and his colleagues are creating a universal notation for biologists to explain what they mean clearly and succinctly – not only to each other but to computers as well. "People won't have to read the legend in figures," says Nicolas. "When they see the diagram they will know what it is in the same way an engineer will recognise symbols for batteries and switches and so on."

Until now, different groups of biologists, such as biochemists or geneticists, had their own peculiar conventions for drawing pictures of processes such as gene interactions or metabolic pathways. Many times, if they had to explain something in a paper, they would invent their own symbols and furnish their diagrams with detailed figure legends that were often bigger than the diagram itself. And while biology remained a largely academic discipline, this hardly mattered. Over the past decade, however, various biological disciplines have been transformed into applied sciences, more similar to engineering than blue-skies exploration. New disciplines such as synthetic biology, cell reprogramming and biological engineering are using or manipulating biological material to build devices or new tissues. 'Pure' biologists are also uncovering ever more complexity in the metabolic pathways and networks of gene interactions within a cell. "People need standardisation when representing these networks and pathways," says Nicolas. As well as helping researchers communicate more accurately and efficiently with each other, a standardised notation would make it much easier to build computer software that could process information presented as diagrams.

Work on the new notation started back in 2005, when Nicolas was collaborating with Japanese researcher Hiroaki Kitano from Tokyo's Systems Biology Institute, to develop a software language for systems biology called Systems Biology Markup Language, or SBML. Kitano had already tried to develop a standardised notation and together with Nicolas, had secured a grant from the Japanese government to really take the idea forward. They convened an international team of collaborators, including Kurt Kohn, a researcher from the National Cancer Institute in Maryland, United States, who had also

Nicolas Le Novère

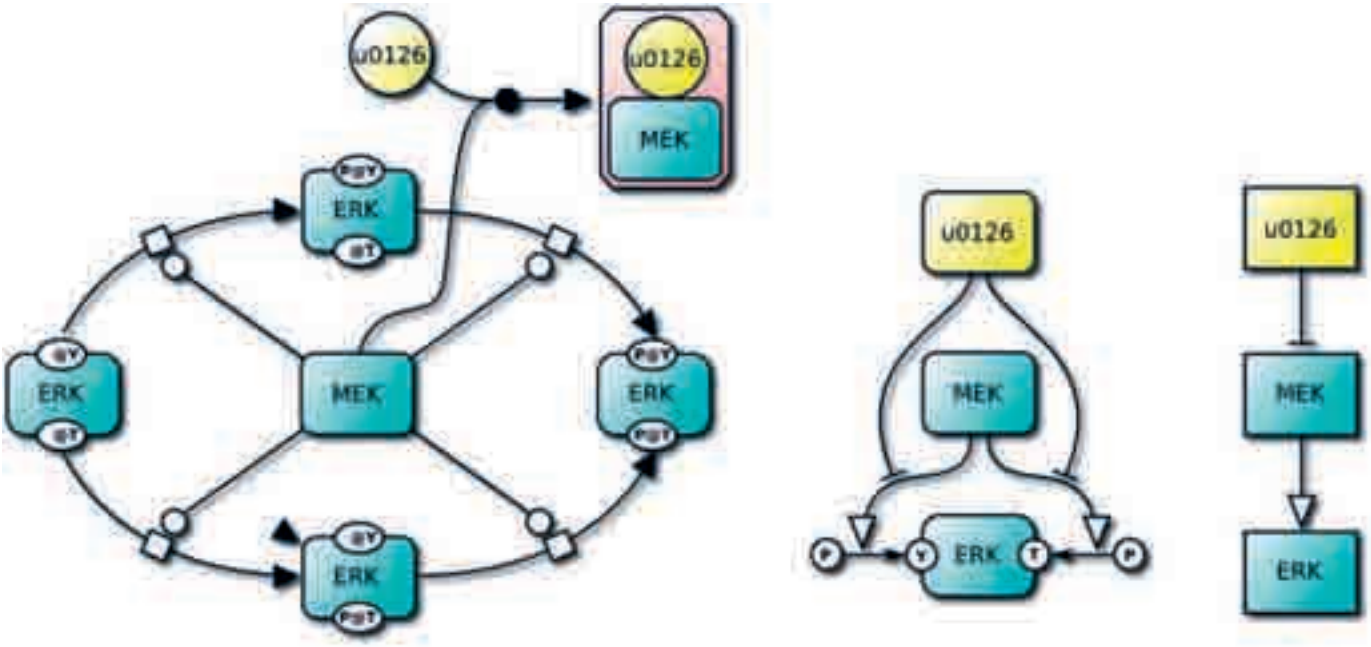
attempted to develop a universal notation. Given the importance of standardisation, it might seem surprising that it has taken so long for work like this to start in earnest. But Nicolas points out that until recently, there has been very little demand. It is only in the past 10 years that certain areas of biology have become more like engineering disciplines and that industry has developed more of an interest in them. What's more, funding bodies have only just begun to perceive the need for this kind of work.

The team started work on the new system, which they named The Systems Biology Graphical Notation, or SBGN. Although the system has one name, the team actually developed three different symbolic methods of describing biological processes from three different perspectives. But these methods are far more than sets of symbols. They have the properties of languages: symbols have specific meanings, or semantics, and they have a defined set of rules, or syntax, about how they should be used to construct communication. Importantly, says Nicolas, the relationships, or ontology, these symbols have to each other follow biological rules. "SBGN is not just another graphical language," he says. Putting something so specific together wasn't easy, but it appealed to his way of working. "I am a very picky person," he says. "I liked the challenge of developing something that was semantically very neat."

Nicolas became the main co-ordinator and driver of the project, which published full details of the three languages in the journal *Nature Biotechnology* in August 2009. Each one is used for a different purpose. "There are three different representations of an underlying biological reality," explains Nicolas. "You choose your language depending on your data and the question you are asking."

The first, Process Descriptions, is used when researchers want to describe how something happens, such as how a molecule is altered as it proceeds through a series of biochemical reactions as part of the cell's metabolism. This allows researchers to describe the mechanics of the process in a sequential, unambiguous manner. It's a great way of depicting processes, but it has one big drawback. If you try to represent all the interactions a molecule has in this kind of diagram, it becomes overwhelmingly complex. "You end up with billions of billions of combinations," says Nicolas.

Clearly, another way of representing every one of a molecule's interactions is needed, and this is where the second language, called Entity Relationships comes in. This allows you to draw a map of all the possible relationships between items or "entities" in the diagram clearly and simply, but it does not show the order in which interactions occur. The third language, called Activity Flows, is more abstract. Here, the entities in



The same biological process (the phosphorylation of MAP kinase) represented in the three SBGN languages.

the diagram aren't molecules themselves, but molecular activity, such as a particular kind of enzyme reaction. This kind of diagram is useful for biologists studying the chemical signalling systems cells use to communicate. Often, their experiments will reveal the presence of a particular kind of enzyme activity, without revealing what that enzyme is. So having a way to express these relationships in an abstract, if ambiguous, way is very useful.

But it's not just molecular biologists who stand to benefit. Nicolas has been working with clinicians to describe disease processes using entity relationship diagrams and is developing new ways of describing other biological phenomena, such as blood flow, using SBGN. Ultimately, he wants to roll SBGN out to cover all of the biological sciences. "We are trying to develop something robust and to cover everyone's needs," he says. To this end, any biology researcher can participate in SBGN development via the project's website: http://sbgn.org/. Decisions about whether new symbols and ideas should become part of SBGN are democratic and a steering committee oversees the development of the project.

As well as continuing to develop SBGN, the next big challenge is getting biologists everywhere to use it. Nicolas is confident that this will come in time as researchers begin to see the benefits, including software that uses the language. What's more, SBGN is not hard to learn. "It is really easy," says Nicolas. "It just takes a bit of repetition. After all, we are all scientists, juggling with gene and protein names; we must be clever enough to learn a few symbols!"

Certainly, SBGN received an enthusiastic reaction from the research community and industry when it was published, with biologists keen to learn more and collaborate. This willingness to share is good news, says Nicolas, as SBGN is part of a larger movement to scale up standardised notation and computational methods in biology. At the moment, much of this work is divided into specialised ghettos, with small communities of researchers focussed exclusively on one topic and not really communicating with the others. "We have to bring all these people together," says Nicolas. "This is the only way to build something coherent and transparent to all researchers." Ultimately, the hope is to transcend geographical and discipline barriers in biology – to create a kind of world-wide tea room where scribbles will need no more lengthy explanations.

Le Novère N et al (2009) The Systems Biology Graphical Notation. *Nature Biotechnology* **27**: 735-741

The royal matchmaker

Ramesh Pillai

In the 21st century, the marriages of Europe's kings, queens and princesses may no longer be arranged by their elders and councillors, but inside cells a royal family is still playing matchmaker. Named after a dynasty of British rulers, the Tudors are, to biochemists such as Ramesh Pillai, a group of proteins that all contain a section called a Tudor domain. And, as Ramesh and his group at EMBL Grenoble discovered, these agony aunts of the protein world are experts at putting the right molecules in contact with each other to ensure that germ cells such as sperm are correctly formed.

The founding member of this royal family is a fruit fly protein called Tudor, which when mutated results in germ-cell formation defects, producing mothers incapable of having grandchildren. Michael Reuter, a postdoctoral fellow in Ramesh's group, was purifying another group of proteins, called piwis, from mouse testes, when he found that they were attached to a Tudor protein. Called Tudor domain-containing protein-1, or Tdrd1 for short, this protein has four Tudor domains, and Michael showed that these domains can bind to specific chemical marks on the piwi proteins. This caught the scientists' eye, because piwi proteins are known to be involved in stopping so-called 'jumping genes' from moving around. Jumping genes – or transposons, to give them their scientific name – are sequences of DNA that can move from place to place within a cell's genome, either by popping out of one place in a cell's DNA and slotting in again somewhere else or by enlisting the cell's machinery to copy and paste them. This can have detrimental effects on the cell if, for example, this extra genetic material is inserted into an important gene, so the cell uses piwi proteins to clamp down on transposons and silence them.

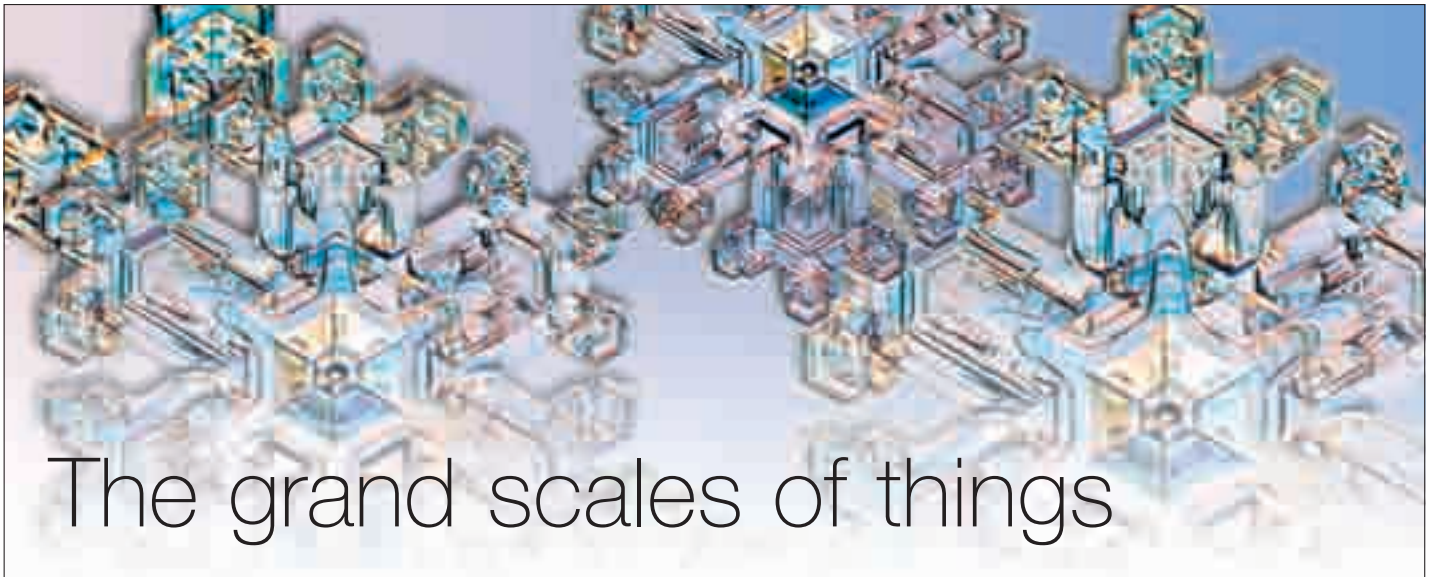
Scientists suspected that piwi proteins are guided to the transposons they silence by molecules called piwi-interacting RNAs – piRNAs – but how these guides are recruited was not known.

"The answer, we found, is that Tdrd1 acts like a hub, bringing people – or in this case piwi proteins – together," Ramesh explains: "Different piwi proteins can bind to the Tudor domains in one Tdrd1 molecule, and that brings them into close contact, allowing them to interact and recruit piRNAs."

To confirm Tdrd1's role, Ramesh and colleagues turned to mice that Shinichiro Chuma and his team at the University of Kyoto in Japan had genetically engineered to have no Tdrd1. "It was as if these mice had no piwi proteins," Ramesh remarks. Michael, who carried out the work in Ramesh's lab, explains: "The piwi proteins are there, they're chemically modified the right way, but if Tdrd1 isn't there, the piRNA guides the proteins have are different, so they cannot function properly." As a result, male mice with no Tdrd1 are sterile.

Since Ramesh and Michael's discovery, scientists have found other proteins with Tudor domains that bind to piwi proteins, and even the Tudor protein found in fruit flies is now known to function in a similar manner. On the whole, this growing royal family of Tudors seems to rule over the recruitment of piRNAs. These matchmaking royals also exist in humans, as do piwi proteins and piRNAs, so Ramesh and Michael's findings could give scientists studying human infertility a new target gene to investigate.

Reuter M, Chuma S, Tanaka T, Franz T, Stark A, Pillai R (2009) Loss of the Mili-interacting Tudor domain–containing protein-1 activates transposons and alters the Mili-associated small RNA profile. *Nat Struct Mol Biol* **16**: 639-646



The grand scales of things

What do you have in common with a snowflake, a cauliflower and Sweden's coastline? More than you might think. According to Jan Ellenberg and his team at EMBL Heidelberg, the mathematics behind these intricate structures also helps our cells run smoothly. Fractals – rough shapes that look the same at all scales – could explain how the cell's nucleus keeps the molecules that drive the biology of our DNA in the right places. "This finding provides a better physical way of understanding or modelling our cells' biochemistry," says Jan.

The nucleus, the brain centre of the cell that houses nearly all of its DNA, is divided into distinctive neighbourhoods, each with different biochemical properties. Some contain "heterochromatin", which are sections of chromosome where DNA is packed tightly around proteins called histones. Genes contained in heterochromatin are mostly inactive. By contrast, other areas of the nucleus contain "euchromatin", where the DNA and histones are more loosely packed. Genes here are easily activated and read, or transcribed, by proteins that find and bind to specific regions of DNA. But there are no obvious physical barriers, such as the membranes that wall off other parts of the cell, to account for how these areas are maintained – all the more puzzling because proteins in these areas are highly dynamic and can freely move between the different neighbourhoods.

One possible explanation is that the high density of DNA and histones crowded together in the nucleus could make the proteins behave in unexpected ways: it could make proteins seem more diluted and obstruct their movements, and encourage them to bind to DNA and stay bound for longer. This could stabilise existing biochemical interactions and act as a positive-feedback mechanism to create and maintain nuclear neighbourhoods. Jan and his team wanted to see whether this phenomenon, called molecular crowding, was really at work in living cells.

Using the state-of-the-art live cell imaging facilities at EMBL Heidelberg, the team tracked the behaviour of a range of molecules injected into nuclei. They found that the molecules were indeed being crowded in the nucleus: they moved as though they were navigating obstacles. But surprisingly, both large and small molecules were obstructed in the same manner. Further experiments confirmed the idea that molecules in the nucleus were "seeing" the same crowded environment, regardless of their size. "It's not just crowded, it's a fractal signature," says Jan.

The researchers then tracked the movement of three kinds of proteins that interact with chromatin. They found that euchromatin seems to have a higher fractal dimension, which means that it exposes a larger and rougher surface. By contrast, heterochromatin's low fractal dimension makes its surface flatter and smoother.

This would affect the target search strategies used by the proteins to find their binding partner on DNA, says Jan. In the rougher euchromatin, proteins would use large-scale exploration, allowing them to scan large regions of chromatin when looking for rare binding sites – such as those used for transcribing genes. On the other hand, the structure of heterochromatin drives proteins to perform local and systematic searches. Proteins that interact with histones, which are involved in keeping genes silenced, might need to do this.

So the nucleus might be able to switch the function of different areas of DNA between silent and active by altering the fractal structure of chromatin, although how it might do this is the subject of Jan's next set of experiments. It's an intriguing thought that the mathematics behind complex, but inanimate, objects like snowflakes should play such an important role in keeping living ones in order.

Bancaud A, Huet S, Daigle N, Mozziconacci J, Beaudouin J and Ellenberg J (2009) Molecular crowding affects diffusion and binding of nuclear proteins in heterochromatin and reveals the fractal organization of chromatin. *EMBO J* **28**: 3785-3798

Paul Kersey

The EBI finds its wild side

There's nowhere quite like London's Natural History Museum. Its imposing, beautiful architecture dominates the streets of South Kensington, like a great temple to the natural world. Inside, countless specimens of plants and animals, collected by great naturalists such as Darwin, colonise its galleries, displays and hidden storerooms. Thousands of visitors come to marvel about Earth's history and the formidable diversity of living things, even as scientists are hard at work behind the scenes, using its resources to understand more about how these myriad forms so strange and wonderful came to be, and how they live.

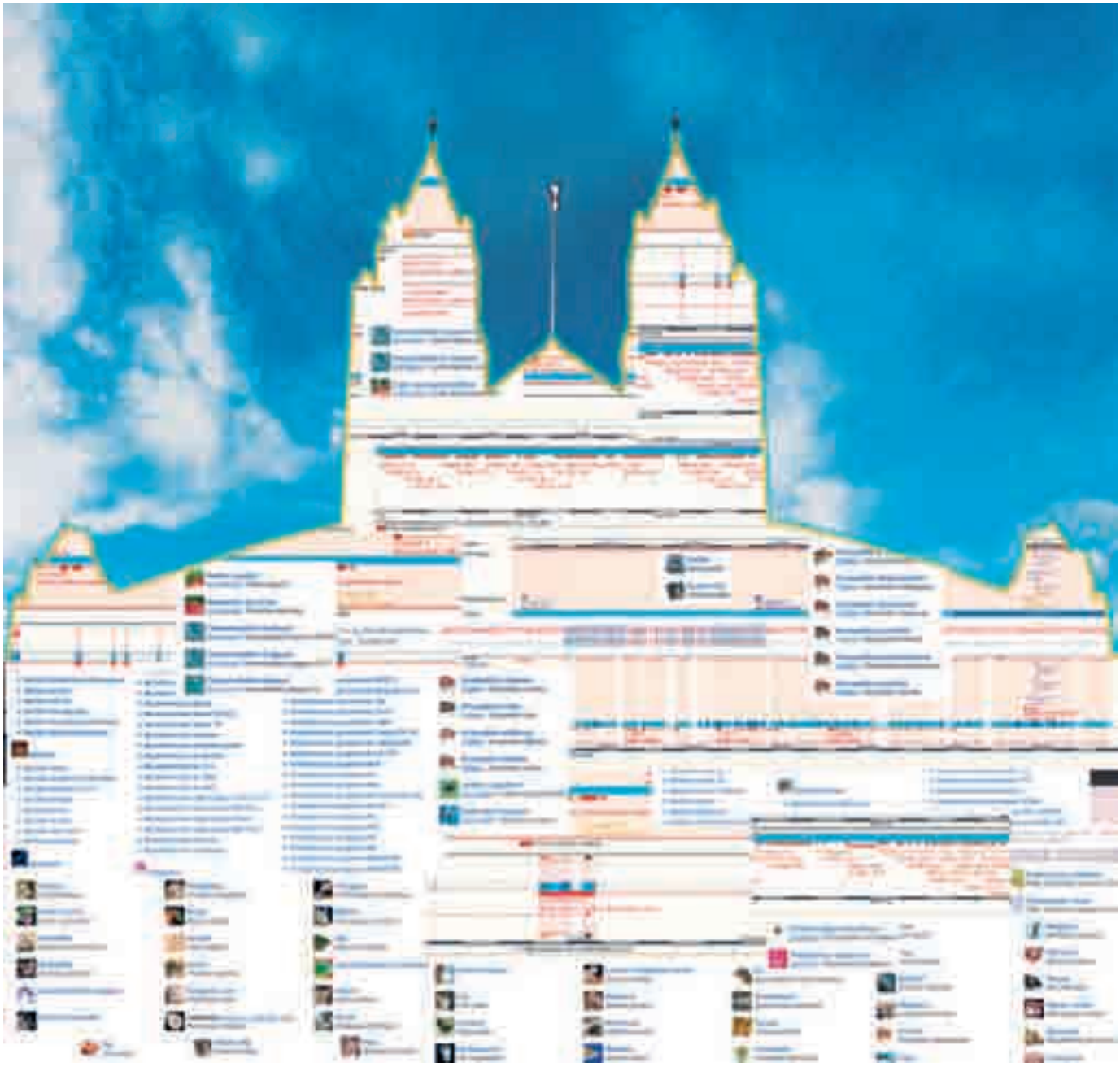
It's not as grand and doesn't yet have as many specimens, but EMBL-EBI has recently launched its own natural history museum of sorts. This one doesn't have display cabinets, dinosaur skeletons or lofty marbled halls. It doesn't have a building as such, or even any intact animals. This is a museum of DNA, which is housed in silicon circuits and is home to a growing throng of genome sequences that can be visited remotely by thousands of scientists around the world.

At first, this analogy might seem a little far-fetched. But it's not really. Like a museum, the new resource aims to enable as many people as possible to access it, make discoveries and inspire them to create new ideas and connections. And it's designed to let researchers do the computational equivalent of wandering through galleries and comparing different species with ease. "We aim to offer an integrated service delivery across taxonomic space," says Paul Kersey, the team leader at EMBL-EBI who heads the project.

This new resource is a database called Ensembl Genomes, and it builds on the existing, highly successful Ensembl genome database that is 10 years old this year. Ensembl was originally designed to analyse and interpret the data from the Human Genome Project, but its scope was later widened to cover all chordates – creatures that possess a stiffened cord of tissue called the notochord, the precursor to the backbone. Ensembl's great success, says Paul, lies in its interpretative power and also in the high quality of its data. "It's been an excellent place to come for scientific data as well as for visualising it," he says.

The idea of creating Ensembl Genomes was first suggested about three years ago during an horizon-scanning meeting at the EBI, at which the team discussed a forthcoming sea-change in how sequence data would be collected and used. Although the EBI has always had to face the exponential growth of data being produced by genome projects, recent changes in sequencing technology have taken this to a whole new level. "There has been a revolution in the past two years," says Paul. As a result, the volume of new data has been immense.

Older sequencing technology produced relatively long stretches of sequence, which could then be stitched together using bioinformatics software to assemble the sequence of entire genomes. This "long-read" technology was, however, a relatively expensive way of doing things. Recently "short-read" technology, which generates much shorter lengths of sequence, entered the scene. It is much



The ENSEMBL databases – a digital museum of sorts

harder to stitch the pieces together, but it is vastly cheaper than the older methods. The sudden drop in costs – sequencing an entire bacterial genome is now about 50 000 times cheaper than it was before – has transformed the rate and scale of data production. "The cost that remains is making sense of the data," says Paul.

And there is a lot of making sense to do. 1200 bacterial genomes have been sequenced over the past 15 years and several projects aiming to greatly expand that number are underway. Many labs are also now going back and "re-sequencing" species whose genomes have already been read. Originally, only one member of each species had their genomes read to produce a single "reference" sequence characteristic of that species. Now, however, researchers want to study the variation in the genome sequences that exists between individuals of the same species, and to work out how this relates to variations in individual physical attributes and physiology. The 1000 Genomes project – in which the EBI is directly involved – has attracted most attention with its focus on human genomes, but similar projects relevant to Ensembl Genomes are also underway, for example the project to sequence 1001 strains of the plant *Arabidopsis*.

The ready availability of relatively cheap whole-genome sequencing is also changing the way biological research is

done, says Paul. Not so long ago, a biologist encountering a new species might study its physiology, ecology or biochemistry, and hope that one day a consortium of researchers would be awarded a large grant to sequence its genome. The day is not far off, Paul says, when researchers will sequence a new species' genome first, and ask biological questions later. "Whole-genome sequencing is being used in a diagnostic capacity to find out what a new species is," he points out. "The technologies are being used on a very wide scale."

Cheaper, faster sequencing is also allowing biologists to tackle biologically and economically important genomes that until now have been difficult to attempt. Chief among these are the crop plants such as cereals, the genomes of which are huge and riddled with sections of sequence that are technically hard to read and reconstruct. "I think 2010 will see a number of very important plants being sequenced," says Paul. "The wheat genome, which is both large and complex, is the last big challenge. That will be the last hurdle for primary genomics."

"The driver is the interest of the scientific community," says Paul. "If they love their genome, then we want to care for it too."

While things might be getting easier for the DNA sequencers, the opposite is true for bioinformatics service providers such as the EBI – and for their users. "Historically, genome data were rare and valuable things. They have now become commonplace," says Paul. "It's very hard for the EBI, with its service mission, to deliver that data to the public in a usable way." His team realised that it would be unfeasible for the EBI to become directly involved in the work on so many genomes, and yet there was a clear need for a resource that would help biologists to find and use the data, without the need for specialised bioinformatics training.

So they had the idea of working with communities to offer help with the bioinformatics where needed, but to also provide a user-friendly, centralised interface, via the EBI, through which users can access data from all these different resources and projects in an integrated manner. Given Ensembl's success, it made sense to follow its model: to use the genome as an index system for all the different kinds of biological data about a gene or organism. "We have this huge flood of many new types of data," says Paul. "The genome is a natural way of organising access to it."

But it's more than just organisation. The data in Ensembl Genomes are presented in a way that is both meaningful and easy to access – users come to interpreted information first, rather than just raw data. "You put the raw data in a biological context," explains Paul. "One of the key aims is to allow users to see their data in the context of other genomes." Users can then click through to see the raw evidence if they wish. "You want to guide users to references, interpret data and then let them sensibly navigate through the informatics without drowning in it," says Paul.

Importantly, the EBI is working in partnership with different genomics communities to complement, rather than duplicate, their bioinformatics efforts. "We aim for robustness of infrastructure and breadth of scope," says Paul. For example, his team is a collaborator in VectorBase, a community database funded by the National Institutes of Health in the USA that focusses on the genomics of mosquitoes and other insects implicated in disease. The EBI contributes to the annotation of the genomes of these species, then provides a portal through which this information can be accessed in a broader taxonomic context.

To reflect its breadth of scope, Ensembl Genomes is divided into five main parts, each representing one of the five kingdoms of life: bacteria, protists, fungi, plants and invertebrate metazoa (multicellular animals). This complements the vertebrate focus of the existing Ensembl site. The bacteria, protists and invertebrate sections were launched in April 2009, and the plants and fungi sections followed in October 2009. The number of users is already increasing. "It's still a young resource," says Paul. He and his colleagues at the EBI on the Ensembl and Ensembl Genomes projects, Paul Flicek and Ewan Birney, are hard at work expanding the resources: forming collaborations with specialist communities and adding more species, and improving the power of user interfaces. "We want to make sure we have all the important model organisms in the database," says Paul.

In future, the team also plans to find ways of handling data from bacterial genomes more effectively, given the sheer numbers of species being sequenced. But, like any exacting museum curators, the researchers are being choosy about which organisms they collect on the database, ensuring that only the ones that are most relevant and useful to their visitors make the grade. "The driver is the interest of the scientific community," says Paul. "If they love their genome, then we want to care for it too."

www.ensembl.org

www.ensemblgenomes.org

Technology transfer at a glance

Technology transfer at EMBL, managed by wholly owned subsidiary EMBL Enterprise Management Technology Transfer GmbH (EMBLEM, established in 1999), has become an integral component of EMBL. EMBLEM's proactive technology transfer approach ensures the rapid commercial development of promising innovations while concomitantly securing the free dissemination of knowledge for basic research purposes. This allows for the development of medicines, diagnostic tools, devices and spin-out companies to the benefit of the member states.

EMBLEM's technology portfolio

EMBLEM's technology portfolio broadly spans the life sciences and includes enabling technologies, molecular tools and techniques, animal models, instruments and devices and software programmes and databases.

SPIM/DSLM

Single Plane Illumination Microscopy (SPIM) was invented by EMBL scientists and it revolutionised the field of light microscopy and live imaging. It enables the study of large, living specimens from different angles, under physiological conditions and with minimal harm. Images obtained along different axes and at varying time points are assembled into three-dimensional images or movies, which provide insights into the dynamic cellular processes that are occurring in living organisms. A license agreement was concluded with Zeiss and first prototypes were developed. SPIM underwent a major upgrade with the introduction of the Digital Scanned Laser Light Sheet Microscope (DSLM), which makes it possible to illuminate a specimen with a thin laser beam and scan it line by line, horizontally and vertically. DSLM was crucial in obtaining the first complete developmental blueprint of a vertebrate in 2009.

The need for technology transfer

What is technology transfer, and what does this success story actually mean for EMBL and its scientists? With innovation being driven by excellent basic research, it is essential to have mechanisms that ensure that scientific results and discoveries can properly and rapidly be translated into practical applications and marketable products – be it equipment, medicines, diagnostic tools or even entire companies – for the benefit of society as a whole. This has always been one of EMBL's missions and with the creation of EMBLEM, the previously ad hoc protection and commercialisation of intellectual property and inventions has been professionalised and

ARP/wARP

The transformation of electron density maps from crystallography experiments into three-dimensional structures requires complex mathematical modelling. In the 1990s, 'off-the-shelf' software solutions were rare and insufficient for research at the cutting edge of structural biology. To address this, scientists at EMBL Hamburg developed ARP/wARP, a software suite for automatic structure determination from crystallographic data. It is now used by several thousand users in academic and industrial research and has been continuously updated during the past ten years to meet changing user needs and environments. The latest version was launched in January 2010.

PROcellcare

Advanced Light Microscopy Facility scientists and staff from EMBL's mechanical workshop built a demonstration model of an automatic dispenser system for microscopy. The system assures the stability of cell-culture conditions with an incubation chamber positioned in the optical axis of a microscope and facilitates live-cell imaging under changing conditions. A commercial prototype – developed in collaboration with the external engineering company PROdesign and patented and licensed through EMBLEM – has been available to researchers under the name 'PROcellcare' since October 2009.

streamlined, benefiting the member states and society with the rapid commercial development of promising innovations.

Successful technology transfer builds on the renown of the institute and boosts public trust, and the scientist inventors can enjoy the recognition and remuneration that marketing an invention brings. More than 400 EMBL scientists are on record as inventors; statistically, every third EMBL scientist has been actively engaged in technology transfer.

EMBLEM now has a portfolio of more than 260 granted patents and patent applications, in excess of 90 copyrights and trademarks and 12 spin-out companies. Of the more than 250 satisfied commercial licensees of EMBL technologies, over half are recurring customers interested in establishing a long-term relationship with EMBL and EMBLEM.

ESPRIT

ESPRIT, the library-based construct screening technology, was developed at EMBL Grenoble to identify soluble constructs of 'difficult-to-express' protein targets that resist the classical approach of bioinformatics and PCR cloning. A diverse random library of DNA constructs is generated and screened to identify rare clones of interest. All unidirectional truncations of the target gene are synthesised by exonuclease degradation to generate potential expression constructs, and a 'scanning' version identifies internal domains. In each experiment, up to 28 000 individual constructs are assayed in parallel for yield and solubility using colony picking and liquid-handling robotics.

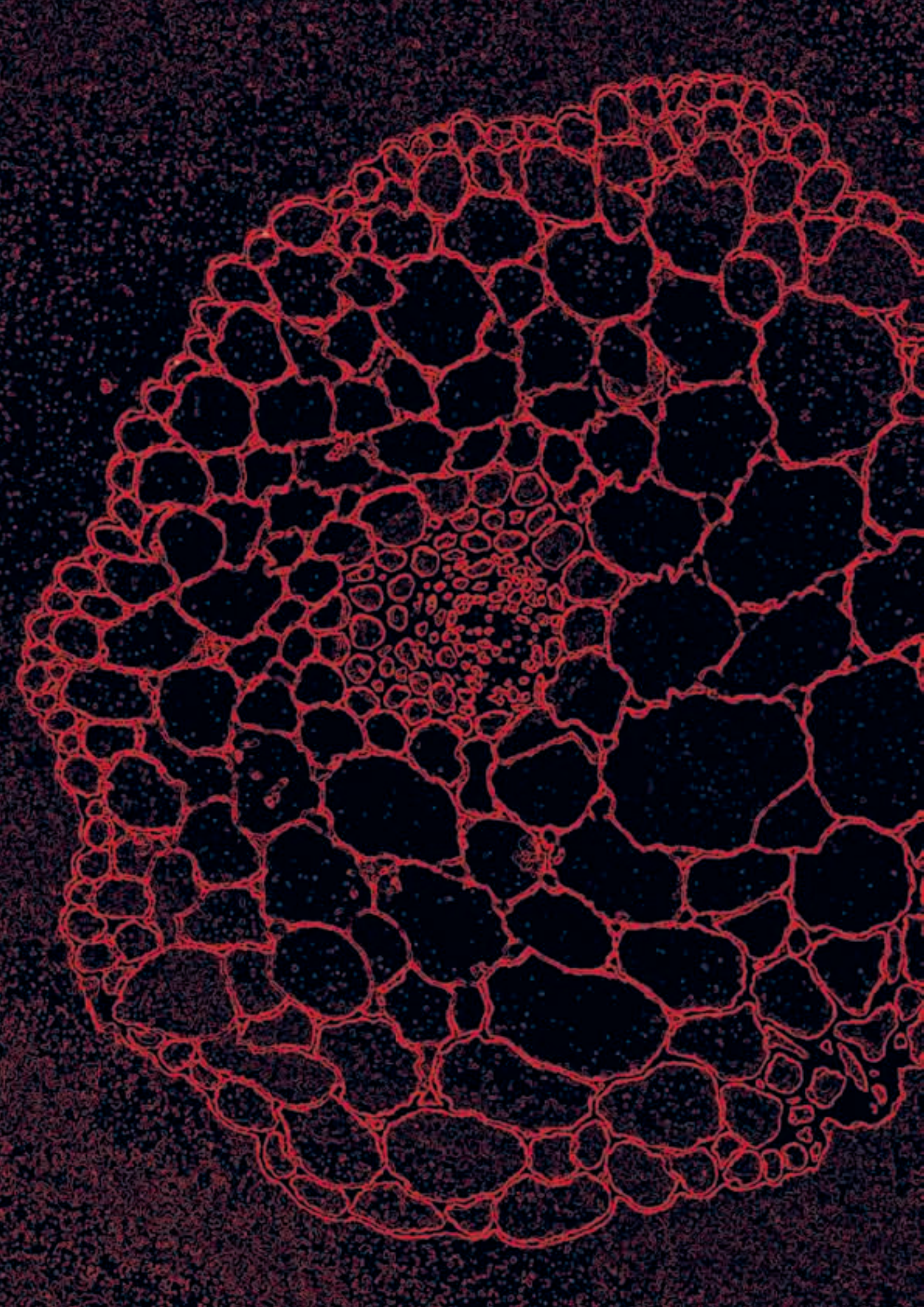
Portfolio companies

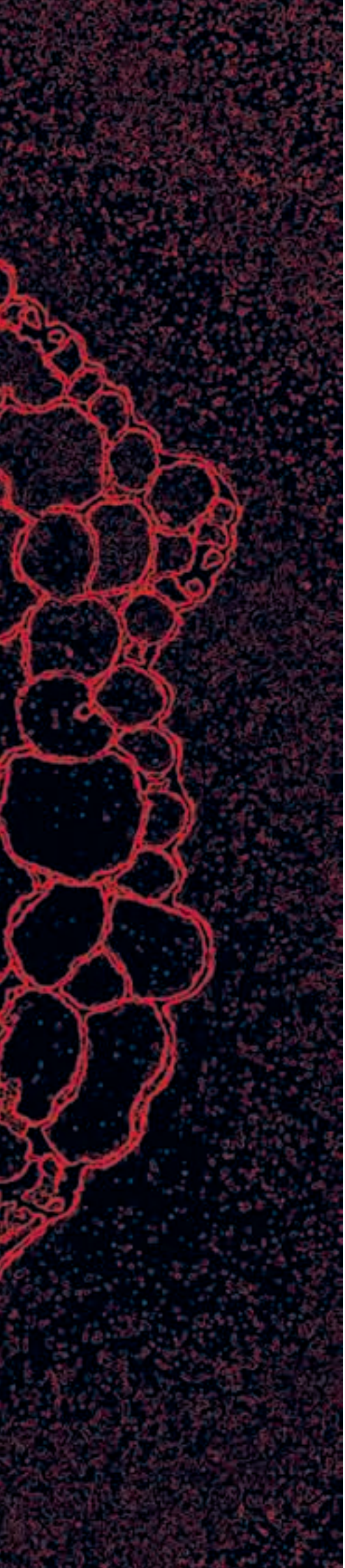
EMBL holds a reservoir of unique opportunities suitable for venture capital financing, and EMBLEM invests in core technologies that can be foundations for building successful, profitable businesses over the long term. Current spin-out companies are dedicated to such areas as the discovery and development of drugs for the treatment of influenza; disorders of the central nervous system and cancer; DNA engineering; and research services into genome-based applications of RNA-mediated interference.

- Sygnis Pharma AG, formerly Lion Bioscience AG (1997)
- Cenix Bioscience GmbH (1999)
- Cellzome AG (2000)
- Anadys Pharmaceuticals, Inc. (2000)
- Gene Bridges GmbH (2000)
- ENVIVO Pharmaceuticals, Inc. (2001)
- Triskel Therapeutics Ltd. (2006)
- Elara Pharmaceuticals GmbH (2006)
- BioByte Solutions GmbH (2008)
- Savira Pharmaceuticals GmbH (2009)

Outreach and training

Through education and teaching initiatives, site visits to member states and regular contributions to the events of the Association of European Science & Technology Transfer Professionals, EMBLEM acts as a source of best practice and is considered the benchmark in Europe for technology transfer from an international academic setting. As well as providing inventors and founders with all the tools and support required to rapidly develop and deploy their ideas, EMBLEM provides training for scientists in what sort of results are commercialisable, patentability criteria and in the components of a successful start-up company. EMBLEM has been instrumental in creating a pan-European technology transfer network by collaborating with the technology transfer divisions of various national and international research institutes and universities, and continues to extend its services to EMBL alumni across Europe and has entered into partnerships with other EIROforum members such as the European Space Agency (ESA). Since the beginning of 2007, in a consortium with the technology transfer division of the German Cancer Research Centre (DKFZ), it has been responsible for the technology transfer activities of the University of Heidelberg Medical Faculty and associated clinics.





Upholding law and order

By day, he was the mild-mannered, rich philanthropist Bruce Wayne. By night, he was the sinister, slightly unhinged superhero Batman, the Dark Knight who stalked the underworld of Gotham city, rounding up thieves, swindlers and other evil-doers to dole out justice. And it's just as well he did. Gotham is perhaps the most dysfunctional of all fictitious cities. Not only was it ridden with crime and plagued by bizarre villains like the Joker and the Penguin, but its police force and judiciary were hopelessly corrupt. It's a good metaphor for what happens in our bodies and cells when order breaks down, and shows why biology needs its own kind of police force and superheroes to keep chaos and disease at bay.

Living systems, be they cells, organs, bodies or ecosystems, are astoundingly intricate and complex. To function, their components must all follow strict sets of biological rules: enzymes need to manage reactions in the right place, at the right time and at the right speed; genes need to co-ordinate their activities to build a developing embryo; cells must acquire their specialist functions and remember them; tissues must control how their constituent cells are renewed. If cells disobey these rules, or if the biological policing of them becomes corrupt, the system can rapidly break down. Cells 'forgetting' to follow their specialist programmes, for example, can result in diseases such as cancer.

Not surprisingly, biology has come up with mechanisms to try to make sure this doesn't happen. Several teams at EMBL, for example, are studying how the cell controls the activity of its genes. Some are looking at how cells use proteins that bind to DNA and switch genes on or off. Others are studying how the protein scaffolding around which DNA is wound controls gene behaviour. As well as being controlled, certain genes do a lot of rule enforcement themselves. The proteins they make direct how a cell acquires its specialist function. A mutation that disables one of these genes can have a dramatic effect: a team at EMBL Heidelberg, for example, has found that the loss of just one gene can cause certain cells in the ovary to switch sex and become testis cells. Other researchers have been studying how genetic mutations can result in abnormal cell division, generating movies that show the effects of turning off each of the 23 000 human genes.

As well as giving new insights into how our cells police their behaviour, findings such as these offer biologists the chance to develop drugs to correct faulty biological law enforcement – and so, in their own way, take on the role of cellular superheroes.


Mathias Treier

How females fight off their inner male

"The difference between the sexes is, happily, one of great profundity." So wrote the author Virginia Woolf, in her famous novel *Orlando: A biography*. It told the tale of a man who lived for 400 years and awoke one day to find himself transformed into a woman. Most people would probably agree with the quote about sex differences and find the idea of one sex transforming into another absurd. But as researchers at EMBL Heidelberg recently showed, transforming female cells into male cells may also be possible.

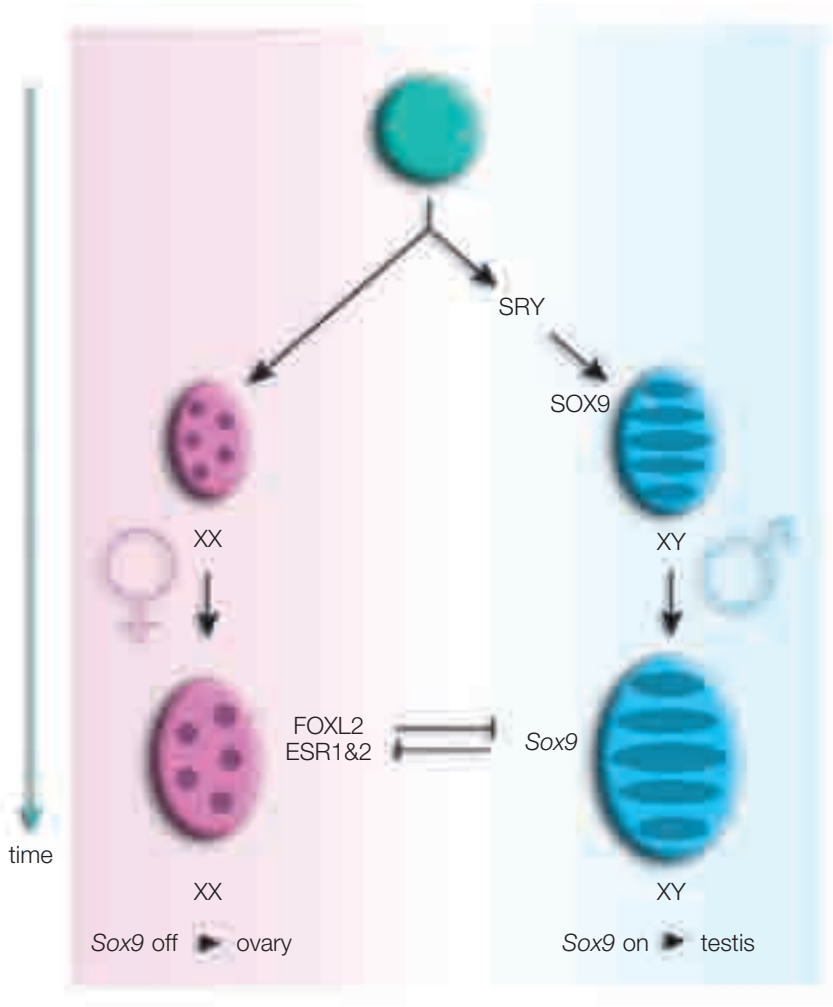
The first question anyone asks when given news of a birth is the sex, often so that they know what colour gifts to buy: blue for a boy or pink for a girl. But biology is never as simple as society might like it to be. Rather than being divided into two discrete camps – pink and blue – human sex can also be ill-defined, running from pink through purple to blue. Some people, for example, have some male and some female traits, a condition known as intersex. In extreme cases, individuals can carry the chromosomes typical of one sex, yet physically be of the other, a phenomenon known as sex reversal.

Now, Mathias Treier and his team have made a fascinating discovery that confounds more of our preconceptions about sex. It seems that sexual identity might not be as set in stone as we once supposed. And the celebrated *différence* between male and female is in fact surprisingly fragile – from a biochemical perspective at least. The team's work has also uncovered other unexpected similarities between the way mammals and other vertebrates determine their sex, and this has raised new questions about how sex determination works and how it evolved.

About 15 years ago, the model of sex determination in most mammals went something like this: the developing gonad initially has the potential to form either an ovary or a testis. The default development setting is female – in the absence of any other instructions it will develop as an ovary. This is what normally happens if an embryo is carrying two X chromosomes. In embryos with an X and a Y chromosome, a gene located on the Y called *Sry* triggers the male pathway of development resulting in a testis. It does so by switching on the activity of another gene called *Sox9*, which in turn switches on genes that direct male development. Once the development of the ovary or the testis is complete, there is no further need for these pathways: once an ovary or testis, always an ovary or testis.

Hints began to emerge, however, that things are rather more complicated than this. Some of these hints came from studies of people and animals whose chromosomal sex didn't match their physical gender, such as XX males. These studies revealed the presence of genes that seemed to be actively repressing *Sox9* – and thus testis development – in females, which raised the intriguing prospect that female development wasn't just the default pathway, although the role of these newly found genes still wasn't clear.

About a decade ago, Mathias and his team first isolated one of these genes, called *Foxl2*. The FOXL2 protein was already known to be important for ovary development – mutations in the gene in humans trigger a syndrome that includes premature menopause, and disturbances in the equivalent gene in goats causes XX kids to develop as males. This suggested that FOXL2 is important for the



The presence or absence of SRY and SOX9 defines the sex of mouse embryos but, for females, maintaining that sex in adulthood depends on FOXL2 and oestrogen receptors (ESR1&2) countering the effects of SOX9.

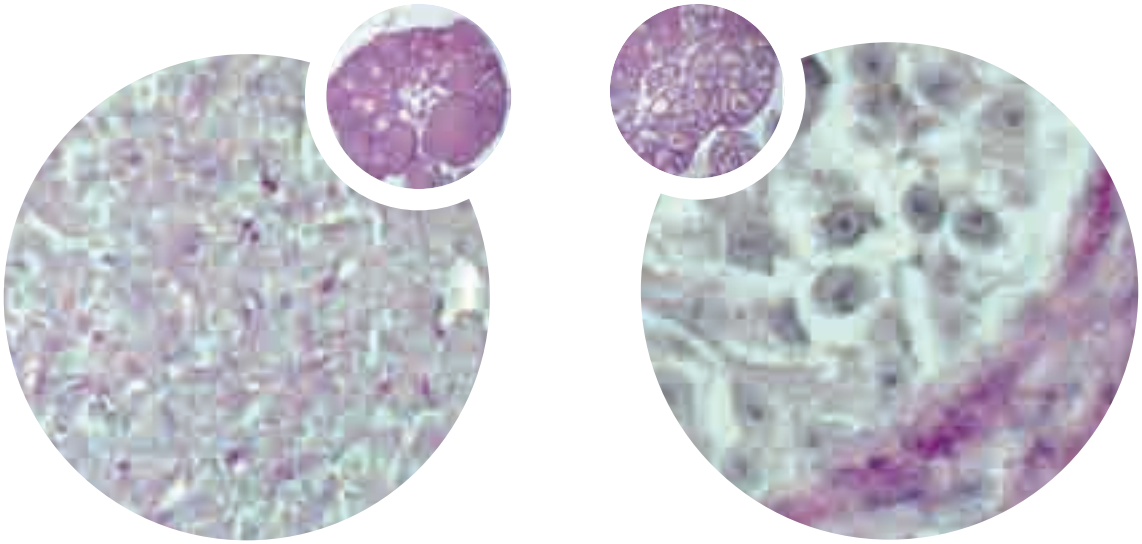
cells in the ovary to develop their specialised functions – to 'differentiate' – and that the protein might suppress parts of the male development programme.

To find out more, Mathias's team inactivated FOXL2 in female mice from the point of conception onwards, to see whether this would cause the ovaries to develop more like testes. The experiment worked, but the results were hard to interpret: the ovaries began to develop normally, but then withered after birth, so the team could not see whether they had developed any features characteristic of testes. So they turned to a newer technology that allows biologists to inactivate a gene at any point that they choose during the lifespan of an animal. "We were interested in what FOXL2 was actually doing in the adult," says Mathias. "And we got a surprising result."

When the team turned FOXL2 off in adult female mice for only a few days, they found that the somatic cells – the cells that are not egg cells – of the ovaries had, effectively, undergone a sex change. Instead of containing cells that would normally nourish developing eggs, the ovaries now housed tube-like structures typical of testes. What's more, these cells had activated genes typical of the Sertoli and Leydig cells of the testis, cell types that aid sperm development. They even made the male hormone testosterone. The few developing eggs that remained in these ovaries were in the process of dying. "It was very clear cut that there was a full somatic sex transformation," says Mathias. Further investigation showed that the transformation did indeed result from the loss of FOXL2 activity in the ovarian cells themselves, and was not dependent on any interactions with the developing eggs in the ovary.

This suggested something remarkable – that one kind of adult tissue was capable of switching completely to another kind, something that developmental biologists had long believed was impossible, or at least very rare. This phenomenon, called transdifferentiation, had been documented in only a handful of adult cell types in the body. "Our result may be one of less than ten examples," says Mathias. For many years, this kind of flexibility, or 'plasticity' was believed to be the preserve of immature cells and stem cells. "I think we will see many more examples like this as more molecular switches like FOXL2 are identified," says Mathias.

The work could give a boost to researchers looking for alternatives to human embryos as sources of stem cells to use as therapies. It lends weight to the idea that researchers could develop drugs to make existing tissues switch to another



Ovary of a normal adult female mouse, with close-up showing the typical female granulosa cells (left), which, when Foxl2 *was silenced (right), took on characteristics of cells normally found in testes.*

desired type. If cancer cells harbour similar kinds of flexibility, perhaps they too could be coaxed into forming harmless cell types. "It tells us that there may be alternative routes circumventing the use of stem cells, which has huge implications for regenerative medicine," says Mathias.

The other surprise is that the ovary needs FOXL2 activity throughout adulthood to repress *Sox9* and prevent the development of male characteristics. Further experiments revealed that FOXL2 is essential for the production of the female hormone oestrogen to keep the ovary in the pink. Mathias envisages a balance of power between female factors (oestrogen and FOXL2) and male factors (SOX9 activity) that dictates the sexual identity of the adult gonad. He hopes this model will give new insights into premature ovarian failure, a condition that causes infertility and early menopause in young women. But it also raises the intriguing possibility that declining FOXL2 activity levels could underlie the normal onset of the menopause. Could it be that a woman's ovaries shut down in later life because the balance shifts more towards testis differentiation? This could explain some of the masculine traits, such as facial hair, often seen in older women, says Mathias.

This might sound preposterous, but many 'lower' vertebrates can switch the sex of their gonads during their lifetimes. They often do so in response to sex hormones such as oestrogen, and many are known to harbour both *Sox9* and *Foxl2* genes. And although mammals are not known to completely switch their gonadal sex under normal circumstances, Mathias's experiments show that it might be possible – when forced to at least. "It shows that the plasticity has been preserved in mammals and that switching somatic sex is still possible," says Mathias.

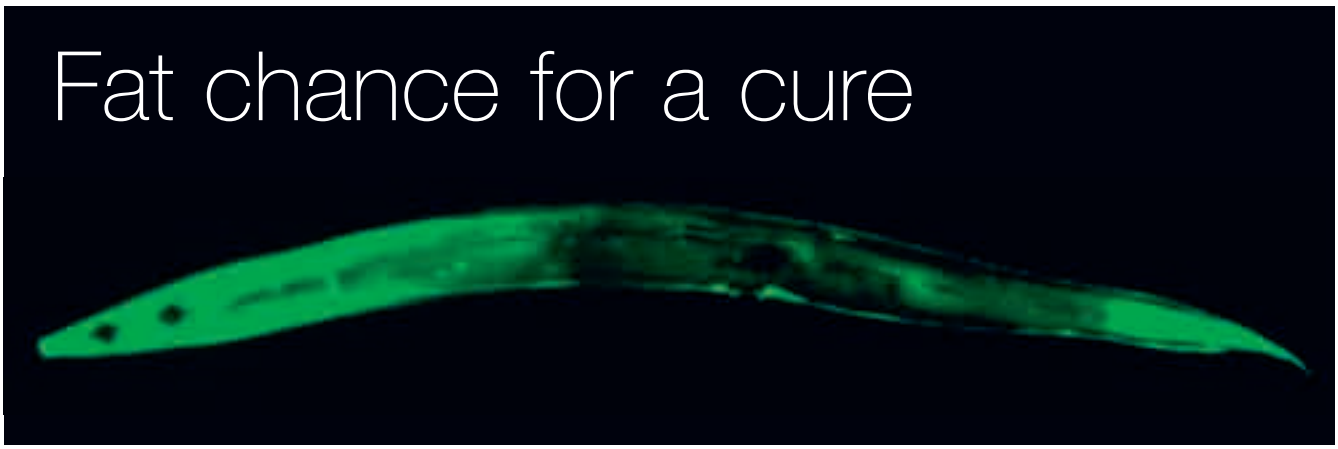
Indeed, perhaps the most salient outcome of the team's findings is how they change biologists' understanding of the mechanism and evolution of sex determination. On the face of it, vertebrates have a smorgasbord of sex determination strategies. Some reptiles use the temperature of the environment, mammals use a method by which the sex with two different sex chromosomes (XY) is male, whereas birds do the exact opposite: the female has a ZW chromosome set. And the duck-billed platypus has a bizarre system of ten sex chromosomes, which seem to have both mammal- and bird-like characteristics.

But the upshot of Mathias's work is that all of these diverse methods may ultimately feed into a universal mechanism that has at its heart FOXL2 and SOX9 opposing each other's action. All that is needed is some kind of cue to tip this Yin and Yang-style balance in favour of one sex or the other. "What we see with SRY on the Y chromosome is another layer of regulation," says Mathias. "How you trigger the switch – why should it matter?"

Mathias and his team now plan to look at the universal mechanism in more detail to determine its components and how they work. "We want to dissect what the core module is," he says. Whatever they find, it is likely to further question society's simplistic notions of what it means to be a boy or a girl. Perhaps it is time we ditched our two-tone attitude and embraced the kaleidoscopic nature of life.

Uhlenhaut H, Jakob S, Anlag K, Eisenberger T, Sekido R, Klugmann C, Treier A-C, Kress J, Klasen C, Holter NI, Riethmacher D, Schütz G, Cooney AJ, Lovell-Badge R, Treier M (2009) Somatic Sex Reprogramming of Adult Ovaries to Testes by FOXL2 Ablation. *Cell* **139**: 1130-1142

Fat chance for a cure



C. elegans *worm with a FAR protein tagged fluorescent green.*

Parasitic worms can no longer rely on secrecy to steal vital nutrients from their hosts, thanks to work by structural biologists at EMBL Hamburg. The team, led by Paul Tucker, has deduced the molecular structure of a protein like the ones such worms use to commandeer their rations. With this blueprint revealed, researchers can now learn more about the biology of these worms and explore ways of starving them out by blocking the proteins with drugs.

The parasites in question are roundworms, or nematodes, and they are a source of immense human suffering and economic loss worldwide. Researchers estimate that about a sixth of the world's population is infected with nematodes, which cause severe diseases such as river blindness and elephantiasis. For some, there is no cure. Worms also infect livestock and crop plants, decimating agricultural revenue.

So finding a weak spot that could be targeted by better drugs or pesticides could make a big difference, and this is just what Paul and his team may have found. Parasitic nematodes cannot make certain molecules, called fatty acids and retinoids, that are vital for their survival. Instead, they secrete proteins called fatty acid- and retinoid-binding proteins, or FAR proteins, that grab these molecules from the host's gut or bloodstream and deliver them back to the worm.

Paul first became interested in FAR proteins when Rositsa Jordanova, a Marie-Curie Trainee, joined the lab to study them in *Ascaris galli*, a nematode that was devastating poultry farming in her native Bulgaria. But she hit a big problem: these worms cannot be grown in the lab and this made the experiments she wanted to do extremely difficult. So the team turned to another nematode called *Caenorhabditis elegans*, which can be grown in the lab and which is commonly used for biomedical research.

Although it is not a parasite, *C. elegans* has several FAR proteins that seem to be similar to those in parasitic species. "It made a great deal of sense to look for similar proteins in *C. elegans*," says Paul. It still wasn't easy, though – the biochemistry of the proteins makes it hard to get them to form the crystals needed for X-ray studies – and Hamburg's high-throughput crystallisation facility was key to the success of the project. "It was a long and hard process to get useful crystals," says Paul.

The team successfully deduced the structure and found that the protein had a structural feature, or fold, that had not been seen anywhere else. "It's a unique structure," says Paul. This makes it an ideal drug target: drugs that block fatty-acid binding will be unlikely to affect other proteins and so this minimises the risk of harmful side effects.

Further experiments showed where the protein was active in the worm and that its fold can bind a range of fat-like molecules. Although useful for the worm, molecules with fat-like properties aren't normally thought of as drugs because they cannot enter human cells. But in the case of anti-FAR drugs, this would be a good thing. "You don't want the drug to enter the host cell, just the nematode," says Paul. So tackling FARs will require a new approach based on state-of-the-art methods, says Paul. "This highlights the importance of chemical biology," he concludes.

Jordanova R, Groves MR, Kostova E, Woltersdorf C, Liebau E, Tucker PA (2009) Fatty Acid- and Retinoid-binding Proteins Have Distinct Binding Pockets for the Two Types of Cargo. *J Biol Chem* **284**: 35818 - 35826

Paul Tucker

A nervous switch

In 1863 a Heidelberg doctor described a devastating neurodegenerative condition that causes children to forget how to walk and talk before their teens. The symptoms begin with muscle weakness, poor balance and a slurring of speech, and develop into a gradual breakdown in all motor control. Mysteriously, physicians sometimes noticed similar, although milder, symptoms in the parents of afflicted children, and even milder effects in the grandparents, suggesting that the disease might result from a genetic defect that grows more and more severe each generation. Only in the 1990s did geneticists discover long tracks of the amino acid glutamine within certain proteins of children suffering from these strange neurological symptoms. Substantially lengthening these glutamine tracks caused the proteins to clump together in the nuclei of cells – usually nerve cells – leading to the cell's demise. At least nine inherited diseases, which can affect both adults and children and include Huntington's disease, are known to result from these repeat mutations.

But these glutamine repeats, it seems, are not solely to blame for the deteriorating nerve control. For clues to other contributing factors to the development of these diseases, scientists are trying to identify the roles of the affected proteins.

One such protein, which causes a rare but devastating neurological condition known as spinocerebellar ataxia type 1, is the focus of chemist Annalisa Pastore, a former EMBL group leader who now works at the National Institute for Medical Research in London. Annalisa hopes to gain clues about the function of this protein – called Ataxin-1 – by investigating its structural form to predict with which

Toby Gibson

molecules it interacts. However, Ataxin-1, unlike most proteins, does not fold into three-dimensional shapes, but instead remains floppy and mostly unfolded. Thus, the structural models most researchers usually use to predict protein–protein interactions were of no use for this protein.

So Annalisa teamed up with computational biologist Toby Gibson from EMBL Heidelberg, who has developed a computer program to search for interaction sites on the non-folded regions of proteins. The researchers applied this bioinformatic resource, called the Eukaryotic Linear Motif Resource, to compare the amino-acid sequence of Ataxin-1 across many species, ranging from mosquitoes to zebrafish to humans. In this way, they sought to predict which short regions, or motifs, of Ataxin-1's amino-acid sequence might interact with other proteins.

"Annalisa and I were sitting together when my database threw up a promising candidate motif that suggested that Ataxin-1 might interact with a protein, called U2AF65, that regulates alternative splicing – a means by which a single gene can produce many variants," says Toby. "But what got Annalisa even more excited was that this motif overlapped with an existing motif that was already known to interact with two other proteins," says Toby.

In the lab, Annalisa and her group confirmed Ataxin-1's role in alternative splicing – an important find for scientists examining proteins involved in these neurological diseases for clues about why they are the targets of glutamine repeats. Annalisa also verified that she and Toby had spotted in the database a three-way molecular switch, which involves three proteins binding at one interaction site. "What is key in this case is that Ataxin-1 can only bind one of its interacting proteins at a time," Toby explains.

Toby's group is keen to find more motifs that could be important for the development of many other diseases. "We suspect that there are over a million of these regulatory motifs in human proteins, and you could say that it is my team's holy grail to learn how to find them all!" Toby says.

---

de Chiara C, Menon R, Strom M, Gibson T, Pastore A (2009) Phosphorylation of S776 and 14-3-3 Binding Modulate Ataxin-1 Interaction with Splicing Factors. *PLoS ONE*: **4**

Cue factors

Our genes were once thought to be responsible for shaping who we are. But now scientists are having a rethink. Thanks to a glut of data from new sequencing projects, researchers are beginning to recognise that the regions of the human genome that encode proteins are unlikely to be behind the millions of differences between people. So the question remains: what accounts for these differences?

Searching for an answer, biologists have pored over the few individual genome sequences that have been completed so far. And these researchers have asked: if the rare stretches of DNA that code for proteins are not responsible for many of the differences found between humans, then what about the remaining 98% of the genome that does not encode proteins – the so-called non-coding DNA?

Small regions of non-coding DNA are known to serve as docking sites for regulatory proteins called transcription factors, which are responsible for cueing when and where genes are turned on and off. These tiny transcription factor-binding sites are usually found just upstream of genes – although they can occur anywhere within the genome – and they recruit other proteins that are essential for transforming genes into their respective proteins.

Scientists now suspect that variation in the efficiency of transcription factor binding, and the effect of this binding on how genes are turned on, could be responsible for many of the little quirks that make each of us unique. Jan Korbel, a genome biologist from EMBL Heidelberg, is particularly interested in variation among people. So he teamed up with Stanford University geneticist Michael Snyder to investigate the effect of differences in transcription factor binding on the variation in gene expression in 10 humans – five Europeans, three Africans, and two East Asians – by looking at their complete genome sequences.

By chemically gluing transcription factors to their DNA-binding sites, the researchers identified the binding patterns of these regulatory proteins and looked to see whether they bound every site equally strongly in all 10 people. The proteins they studied were NFkB, which is involved in the immune response, and Pol-II, a protein that helps convert DNA to RNA. For NFkB, less than 10% of the 15 000 binding sites varied between people. Yet for Pol-II, about a quarter of its 19 000 binding sites were altered across individuals – considerably more variation than is seen in coding regions, which vary, on average, by only a fraction of a percent among people.

Jan's group was able to trace some of this variation back to either small differences or larger rearrangements in the DNA comprising these binding sites. Still, the majority of binding sites showed no difference in their DNA sequences, despite their variable binding affinity among people. In these cases, Jan explains, the researchers frequently found variations in nearby regions that could explain the differences. "This is a very interesting finding, as it suggests that these regulatory proteins are not working alone, but instead co-operating with nearby transcription factors to regulate how genes are expressed," he says. In fact, Jan's team has just developed a new test for assessing the extent of this transcription factor co-operation.

Jan and his team went on to show that the variation at these binding sites had a profound impact on an individual's gene expression. "Many genes' expression differed by greater than two orders of magnitude between people as a result of variation in transcription factor binding," says Jan.

"What makes this study novel is that it determined how differences in gene expression among people are associated with the genome-wide variation in transcription factor binding, and also with sequence variation in the factor's binding sites," remarks Jan. "Our findings suggest that variation in non-coding DNA may be responsible for many of the differences we see between people."

Lars Steinmetz and Jan Korbel

What's more, these findings also offer an explanation for the differences between humans and their closest cousins. When Jan's team looked in both humans and chimpanzees, they found that nearly a third of the Pol-II-binding sites differed between the two species. Jan explains that there seems to be almost as much variation among humans as between humans and chimpanzees.

The researchers' results were buoyed by a second study in which genome biologist Lars Steinmetz of EMBL Heidelberg, also in collaboration with the Snyder group, used yeast to map genome regions bound by a transcription factor called Ste12 that is important in yeast mating.

Lars and his team compared binding sites between two strains of the common baker's yeast *Saccharomyces cerevisiae*, one of which descended from cells collected from a rotten fig in Merced, California, and the other derived from cells isolated from the lung of a person suffering from AIDS-related immune problems in San Francisco.

The yeast's small genome and short generation time offered Lars and his group the opportunity to fine-map these sites across many more individuals than had previously been possible using human cells. This gave the researchers the power to detect very small differences in transcription factor binding. Owing to the sensitivity of this approach, Lars' group found even higher rates of variation in Ste12 transcription factor binding between the yeasts than was seen in the human study, and because of yeast's frequent genome shuffling, they were able to ascribe this variation to very narrow regions of the genome.

"Our results, and Jan's study involving humans cell lines, suggest that the bulk of the differences among individuals are not found in the genes themselves, but in the non-coding portions of the genome and in particular in the regulatory regions, which we know relatively little about," says Lars.

When mapping the DNA responsible for the variation in transcription factor binding, Lars and his team found that almost 90% of the binding differences are influenced by sequence variation in nearby regions. This offers tantalising evidence that variation in the regulation of genes by transcription factors is often controlled by sequence differences close to the genes. However, in several cases the data revealed variations far from the genes that also influence gene expression. These regions can encode the transcription factors themselves or signalling molecules that influence the efficiency of a certain transcription factor's binding ability.

Knowing whether genes are regulated by nearby regions or regions on entirely different chromosomes will determine how scientists study variation in gene regulation. If gene expression variation is indeed often controlled locally – by so called cis-regulatory elements – then researchers are safe to continue to take a reductionist approach and look to ascribe the differences between individuals to the regulatory regions around genes. However, if variation in gene regulation turns out to be under the control of regions scattered far and wide throughout the genome – often termed trans-regulatory variation – then researchers need to change tack and study the whole genome simultaneously. Such a broad approach requires the skills of system biologists who specialise in studying entire, intact biological systems.

Either way, once scientists have mapped these regulatory regions, they can look at how individual variations in gene regulation affect how we look, how we think, whether we are susceptible to certain diseases and how we will react to specific treatments. This will open doors to personalised medicine, which seeks to make use of an individual's differences to provide better medical care.

Kasowski M *et al* (2010) Variation in transcription factor binding among humans. *Science* **328**: 232-235

Zheng W, Zhao H, Mancera E, Steinmetz L, Snyder M (2010) Genetic analysis of variation in transcription factor binding in yeast. *Nature* **464**: 1187-1191

Katarzyna Oktaba Sosin, Maria-Cristina Gambetta and Jürg Müller

Aberrant appendages

Having a second pair of hands might seem like an advantage but animals born with extra limbs, because of changes in their DNA, generally do not fair well. For more than 25 years, scientists have known about the existence of a mutation in a fruit fly gene that causes just such aberrant appendages, yet the identity of this gene remained a mystery.

That is until developmental biologist Jürg Müller and his team at EMBL Heidelberg set out to find the gene responsible. By comparing the DNA of mutant and normal flies, Jürg's group pinpointed the mutation and found that it disrupts the genetic code for the protein Ogt, an enzyme that sticks sugar molecules to the outside of proteins.

"Ogt is atypical because, unlike other enzymes that add sugars to proteins that eventually reside on the cell's surface or are secreted by the cell, Ogt adds a sugar to proteins in the cell's nucleus and cytoplasm," says Jürg.

"The sugar modification added by Ogt has been found on hundreds of other proteins, and so it was a little unexpected that removing this one protein would cause such profound developmental defects," explains Maria Cristina Gambetta, the PhD student who carried out this work in Jürg's lab. For this reason, Ogt was the last gene on a shortlist of 11 candidate genes that the group examined.

Because flies lacking Ogt show dramatic changes in their head-to-tail body patterning, the researchers suspected that the Ogt protein adds a sugar to a family of proteins that regulate the genes required for normal patterning. These proteins – called Polycomb proteins – do this by remodelling the DNA into compact structures called chromatin. By sitting on the chromatin, the proteins condense the DNA, making it inaccessible to the cellular machinery that would transform the genetic code into active proteins.

To confirm Ogt's role in adding a sugar to the Polycomb proteins, the next step for Jürg's group was to look whether the sugar and Polycomb proteins typically sit together on the chromatin. By tagging the proteins and sugars, they found that they were bound together in regions that are important for controlling whether genes are turned on or off. These findings hinted that the sugar was somehow needed for Polycomb proteins to silence the genes involved in building the fruit fly's body plan.

Jürg's team went on to show that the sugar was attached to just one member of the Polycomb protein family – a protein called Ph. "We don't know why adding a sugar to Ph is important for its function, but this is what we hope to investigate next," says Jürg. "But what we have found so far is a novel and unexpected role for the sugar added by Ogt; the fact that fruit flies that lack this sugar show compromised Polycomb silencing but no other obvious defects is remarkable." Ogt is likely to fulfil a similar role in humans, and so it could determine the positioning and number of our own arms, legs, ribs and, well, everything.

Gambetta MC, Oktaba K, Müller J (2009) Essential Role of the Glycosyltransferase Sxc/Ogt in Polycomb Repression. *Science* **325**: 93 - 96

Tagging the tail on the histone

Nearly 60 years ago, Pamela Lewis, a geneticist at the California Institute of Technology in Pasadena, noticed that some of the flies she was experimenting on had tiny comb-like structures on their second and third pairs of legs, and not just the first pair as is usual. Lewis called these structures 'sex-combs' because males use them to grasp females during mating and she went on to discover the first *Polycomb* gene, one of many such genes now known to encode proteins that disrupt head-to-tail body patterning in a variety of animals, ranging from humans to fruit flies to worms.

Christoph Müller and Jürg Müller

When these genes are mutated, structures in one part of the animal are transformed into structures normally found in another part of the body. Developmental biologist Jürg Müller, a group leader who studies Polycomb group proteins (see page 80) in EMBL Heidelberg's Genome Biology Unit, explains that these proteins are responsible for turning off genes in the body regions where their products don't belong. This suppression is usually permanent: the genes remain silenced even after the cell divides. Polycomb proteins do this by tagging histones – spool-like structures around which DNA is wound – causing them to become compacted to the point that genes on the DNA itself become inaccessible to the cellular machinery.

But how these silencing proteins choose which histones to tag and how this tagging leads to the silencing of nearby genes remains a mystery. To investigate this puzzle, Jürg teamed up with structural biologist Christoph Müller, joint head of the Structural and Computational Biology Unit at EMBL Heidelberg, to study how one Polycomb group protein complex known as PhoRC binds to the flexible tails of histones.

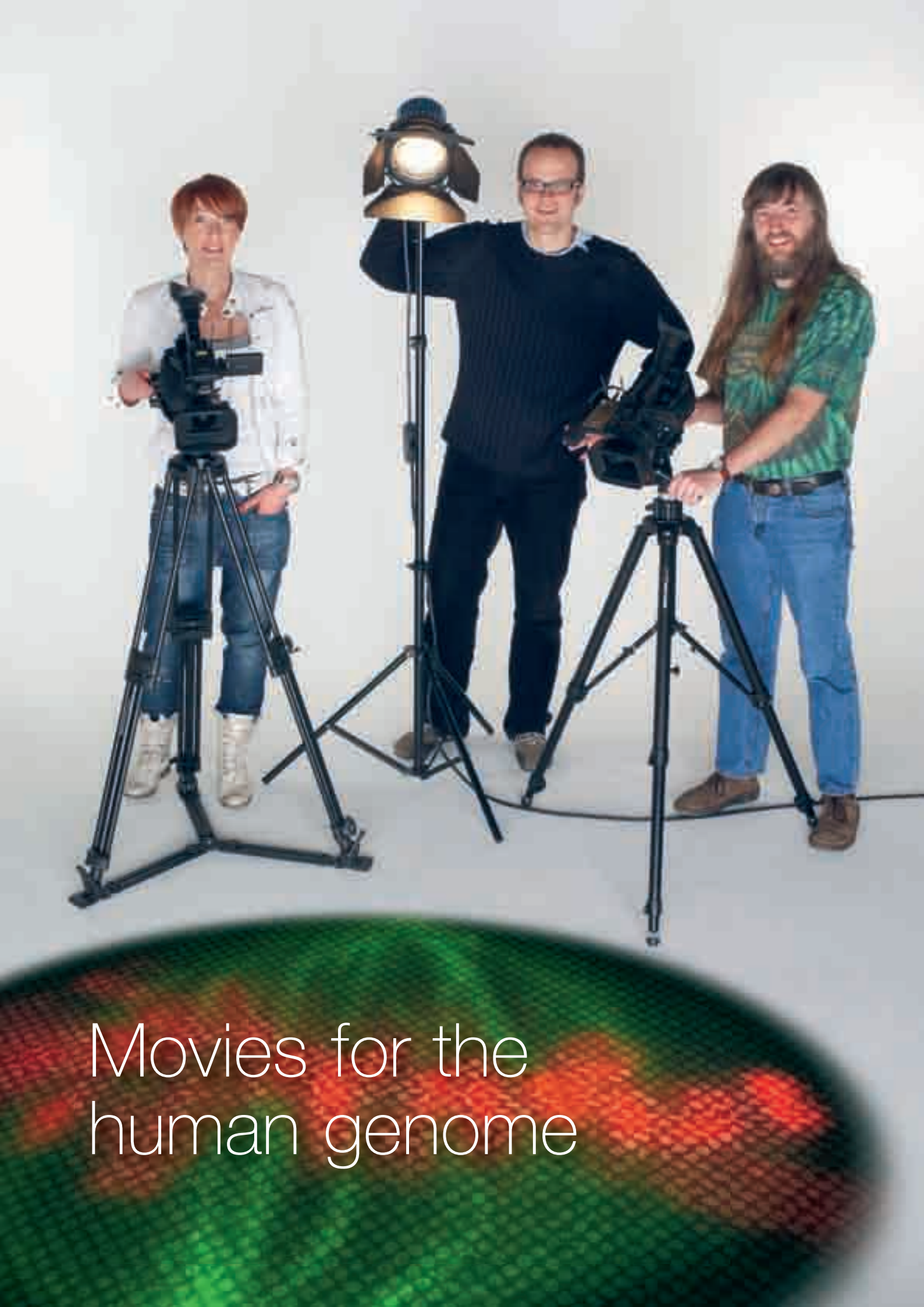
The researchers had previously found that PhoRC is made up of two parts: the first part, Pho, contains a DNA-binding site, whereas the second, dSfmbt, recognises a molecular tag that is attached to a region of the histone tail. This arrangement helps dSfmbt find the DNA regions to be silenced, Christoph explains. "dSfmbt is very selective about which tag it binds to, preferring to bind only to a certain kind of tag, while other, very similar, tags are not recognised," he says. "But we were surprised that dSfmbt is less specific about which region of the histone tail it binds." Combining biophysical analysis and crystallography, the researchers examined the structure of the bound dSfmbt at high resolution and to their surprise, they saw that this region of the dSfmbt protein possessed four similar cage-like structures, only one of which secured the histone tag.

Jürg, Christoph and their groups also found that dSfmbt binds to another protein, called Scm. The researchers found that Scm uses similar cage-like structures to bind the same tags on histones. The researchers suspect that when these two histone-tethered proteins bind to each other they act as a bridge between neighbouring histones that helps to keep the DNA tightly coiled. Through their combined efforts, these two proteins switch off the genes that are meant to be inactive in a particular part of the developing fruit fly.

But many questions still remain. "What we still don't fully understand is how PhoRC helps to recruit other Polycomb proteins that go on to ensure that the nearby genes are silenced," explains Jürg. He and his group would also like to pinpoint what stops Polycomb proteins from silencing genes in all cells, allowing these genes to remain 'on' in regions of the body where their activity is needed to produce the fly's normal body pattern.

Grimm C, Matos R, Ly-Hartig N, Steuerwald N, Lindner D, Rybin V, Müller J, Müller C (2009) Molecular recognition of histone lysine methylation by the Polycomb group repressor dSfmbt. *EMBO J* **28**: 1965-1977

Christian Conrad, Beate Neumann, Jan Ellenberg, Jutta Bulkescher, Thomas Walter, Jean-Karim Hériché



Movies for the human genome

As biology moves into the high-throughput age, the processing of unimaginably large datasets is becoming an integral part of life science research. Hence major projects rely increasingly on the expertise of mathematicians, bioinformaticians and computer scientists. Nowhere is this better illustrated than in the EU-funded project MitoCheck, which began in 2004 with the goal of identifying all the proteins involved in mitosis, and of working out how they are regulated.

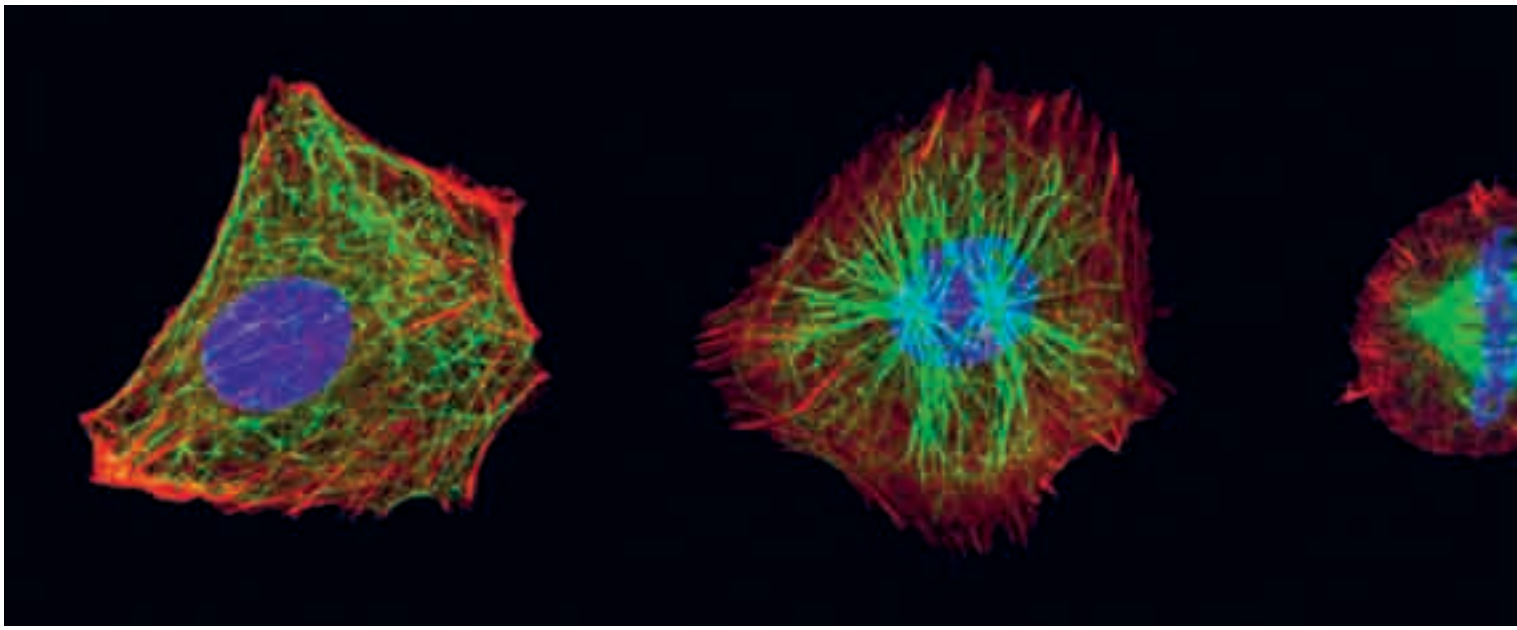
Mitosis, the division of a cell's nucleus, is one of the most fundamental processes in nature. "Every single-celled or multicelled organism, to proliferate and reproduce, has to divide," says Jan Ellenberg, a member of the MitoCheck consortium and the newly appointed Head of the Cell Biology and Biophysics Unit at EMBL. "Without mitosis, there is no life, and if mitosis goes wrong you get problems including cancer and defects of tissue regeneration."

Yet, although scientists have been studying mitosis for well over a century, they still don't know the identity of all the human genes that regulate it. The part of MitoCheck that was carried out at EMBL, under Jan's leadership, was designed to identify those genes in living cells. This genome-wide screen represented the first step in the MitoCheck functional genomics "pipeline" to build a more complete picture of mitosis and to notch up an impressive number of methodological firsts.

The systematic identification of all the genes that are involved in mitosis required the group at EMBL to conduct a screen in which each of the 23 000 genes in the human genome was suppressed one by one, and the effects of that suppression on the observable characteristics of the cell were then recorded by live cell microscopy. The technology behind the screen was developed from scratch by Jan and Rainer Pepperkok, head of EMBL's Advanced Light Microscopy Facility.

In mitosis, microtubules – cellular "bamboo sticks" – form a spindle that invades the nuclear space and pulls the duplicated halves, or chromatids, of each condensed chromosome in opposite directions to form two genetically identical sister cells. To track this process, Jan and Rainer began by fluorescently labelling a cell's chromosomes. Using a powerful tool called RNA interference (RNAi), which selectively disrupts gene expression, they then silenced the genes one by one. This involved spotting short pieces of RNA – small interfering RNAs, or siRNAs – onto an array and laying the fluorescently labelled cells over them, so that they would be absorbed by the cells. A different gene was silenced in the cells in each spot and, using an automated microscopy system, Jan and Rainer were able to film the cells growing and dividing over a 48-hour period.

The completion of this screen in 2008 represented a major achievement for MitoCheck, and it generated a dataset so vast – 40 terabytes, or two million digital images in all – that according to Jan, the cost of storing it alone amounts to about €100 000. Before MitoCheck, the technology for efficiently producing or analysing such a large amount of image data simply did not exist. So during the first two years of the project, two mathematicians, Michael Held (now at the Swiss Federal Institute of Technology, or ETH, in Zürich) and Thomas Walter in Jan's group, devoted all their time to developing new computational tools that would be up to the task.



These tools included pattern recognition algorithms that allow the automatic measurement of the biological consequences of gene silencing on cell division, or more specifically on the morphology of chromosomes. On the basis of this quantitative cell division signature extracted from the microscopy movies, the genes could be grouped into clusters according to the similarity of their effects.

Having identified the genes involved in mitosis, the MitoCheck consortium was presented with a bioinformatic challenge: whereas it was interested in the systems view of mitosis – that is, in all the proteins encoded by the genes Jan's group had identified, and their complex interactions – databases tend to be structured to store information about single biological entities, such as a gene or protein sequence.

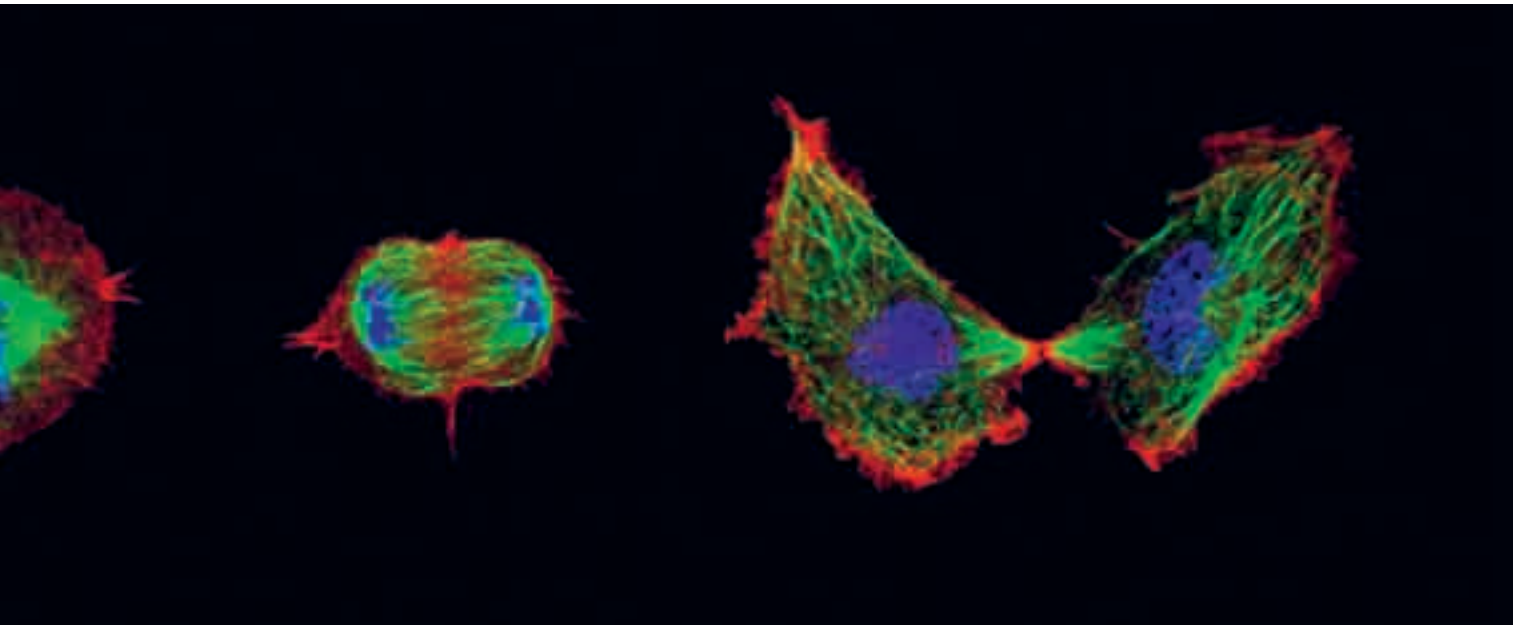
"We now have sophisticated database systems which are able to give us a wealth of information, but they are not very well designed to give us the bigger picture," says EMBL bioinformatician Reinhard Schneider who, with group member Venkata Satagopam and others, came up with new data-mining tools fit for probing the functions of these proteins. These tools have now been fed into the www.mitocheck.org database, which disseminates all the consortium's data. The database was set up by Jean-Karim Hériché, who worked previously with MitoCheck member Richard Durbin at the Sanger Institute in Cambridge, UK, and who is now part of Jan's group at EMBL.

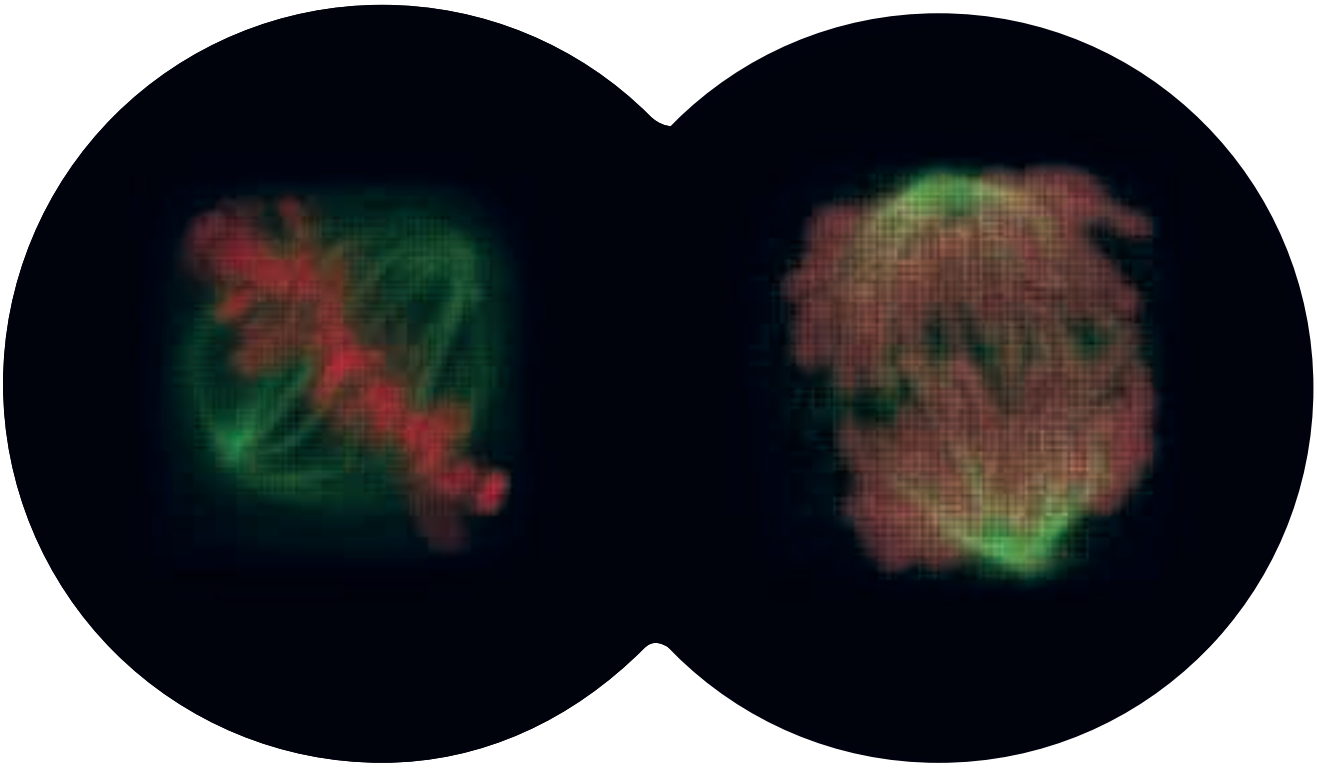
Given the problems their dataset presented, why did Jan and Rainer choose two days as an appropriate running time for their cellular movies? "There is always a trade-off in high-throughput experiments," Jan says. They had to weigh up technical, biological and economic considerations. The cells in question – human cancer cells in culture – divide about once a day, so two days gave them data on two full division cycles. A day less, and they may have missed some effects of gene suppression that only kicked in in the second cycle. A day more, and they would have been dealing with a dataset half as big again, with all the extra processing time and cost that would incur.

In many cases, Jan's screen implicated novel genes, about which little was known, in important aspects of mitosis, such as the assembly of the spindle. In other cases, the gene clusters revealed a role, either direct or indirect, for whole cellular processes that had not previously been linked with mitosis, such as RNA splicing – the modification of RNA that takes place once DNA has been transcribed but before it can be translated into a protein. "We have some very intriguing results that suggest that there is a specific splicing regulation of mitotic proteins that control how they work in different phases of the cell cycle," says Jan.

They found that around 5% (approximately 1200) of human genes are involved in mitosis and for about half of these they could confirm this function in a second validation screen with independent silencing reagents. Moreover, the cell movies identified gene effects on many other cellular functions such as migration and survival that, when combined with what they were learning about mitotic genes, allowed the EMBL researchers to annotate 13% of the human genome with a potential cellular function. As Jan emphasises, however, the time-lapse movies, although highly informative, show only indirectly that a gene is required for a particular cellular

Normal cell division, with DNA labelled blue, microtubules green and actin filaments red.





Two large images of dividing cells, each composed of several microscopy images of human cells in which different individual genes were silenced. The smaller images are placed according to genes' effects: images for genes that affect chromosomes make up the chromosomes (red/pink), while the mitotic spindle (green) is composed of images for genes that affect it.

function. For direct proof, they have to take the approximately 600 validated mitotic genes to the next stages of the functional genomics pipeline of the MitoCheck consortium.

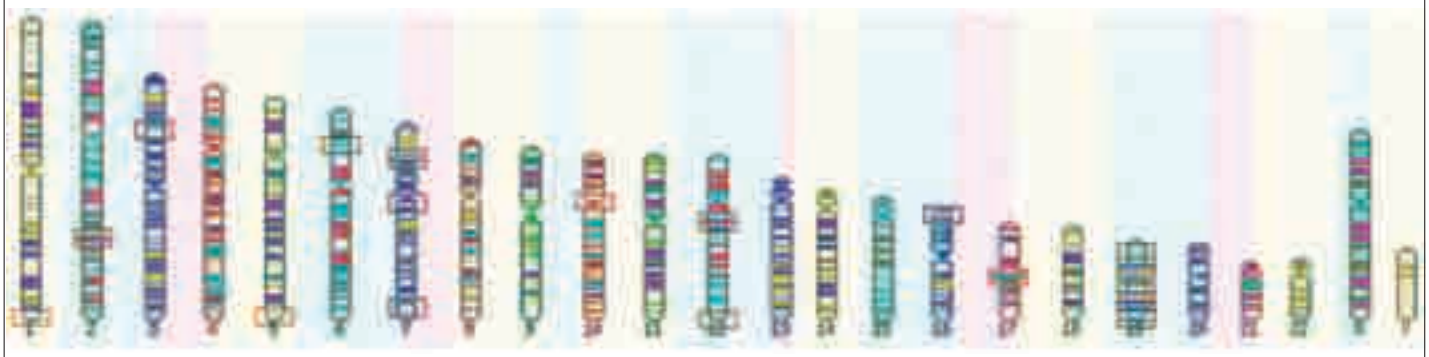
At the Max Planck Institute for Molecular Cell Biology and Genetics in Dresden, former EMBL group leader Tony Hyman and his team are now complementing Jan's screen by making important tools for a systematic functional analysis of human genes. Their approach involves inserting a tagged copy of a mouse transgene into cultured human cells in which the equivalent human gene has been knocked down using RNAi, to see whether they can rescue the effects of that knock-down. Again, the cells' chromosomes are fluorescently labelled, and Jan and his group look for changes using a high-resolution confocal microscope. The fluorescent tag on the transgene also allows the two teams to study the localisation of its protein product in the cell during mitosis. This part of the pipeline is not yet automated, however, and is therefore not high-throughput, though work is continuing at a steady pace.

Further along the pipeline at the Institute for Molecular Pathology in Vienna, the coordinator of the MitoCheck consortium, biochemist Jan-Michael Peters, and colleagues are taking the transgenes generated by Tony's group and are investigating the proteins they encode. Using state-of-the-art proteomics technology, they are studying both the binding partners of these proteins and the modifications that they must undergo to function in mitosis. Armed with this knowledge, researchers will be able to start mapping out the protein interactions and the regulatory networks that contribute to both normal and aberrant cell division.

But even if all the localisations and interactions of the mitotic protein network are known, says Jan, mitosis won't be solved. To really dissect the molecular mechanisms involved the combined expertise of biochemists, biophysicists, cell biologists and computer modellers will be required. This molecular systems biology pipeline still has a lot of work ahead of it, he says, but, "fortunately, this pioneering future systems biology project has been funded by a new EU project, termed MitoSys, and I think within the next five years we have the exciting possibility to solve many questions in cell division down to the molecular level."

Neumann B, et al (2010) Phenotypic profiling of the human genome by time-lapse microscopy reveals cell division genes. *Nature* **464**: 721-727

When I grow up…



If you had a hard time deciding what you wanted to be when you grew up, spare a thought for the cells in your body. If they make the wrong career decision – if a brain cell, say, accidentally turns into a muscle cell – disaster can ensue. Nick Luscombe and his team at the EBI have now completed a census of the proteins that make these decisions, and in so doing have created a valuable resource for research into human health and disease.

One cell type differs from another because different sets of genes are active in each. But every one of your cells contains all the genes in the human genome, and so needs to be told which genes to switch on or off so that the correct sets are active. This job falls to proteins called transcription factors, which bind to DNA to control gene activity. Until now, no-one had a clear picture of how many transcription factors existed in humans, or in which cell types they each worked.

Previously, researchers had produced rough head counts of these proteins using computer programs that automatically search the human genome for DNA sequences commonly found in genes that code for transcription factors. But this approach results in a lot of genes being wrongly labelled, tells researchers very little about the ones that are correctly identified and can miss those transcription factors that don't contain the common sequences.

So Nick and his postdoc Juanma Vaquerizas studied each of the possible transcription factors suggested by the automated searches in more detail. They used additional information from the EBI's Ensembl and InterPro databases to look for more evidence that the candidates did indeed function as transcription factors, as well as to look for candidates that the automated searches had missed.

The team found a total of 1391 confirmed transcription factors. "The most striking thing is that most of them haven't been identified in terms of what they do," says Nick.

To find out more, the team searched another database to see in which tissues the genes were active. They were surprised to find that the transcription factors were largely divided into two camps: those that were active everywhere and those that were active in just one or two tissues.

Nick's team is now working with Ewan Birney and other EBI colleagues to use this data for ENCODE, a project that aims to determine the function of every gene in the human genome.

The researchers also compared DNA sequences between different species to study the evolution of transcription factors. They found sudden bursts in the numbers and diversity of transcription factors whenever living things became significantly more complex, for example when single cells clubbed together to form multicellular animals. Understanding more about the history and biology of these proteins should yield fascinating new insights into how evolution itself works.

Nick's group now plans to investigate how different combinations of transcription factors work together to define different cell types. One day, it might even be possible to treat disease by taking the control of a cell's career decisions into our own hands.

Vaquerizas J, Kummerfeld S, Teichmann S, Luscombe N (2009). A census of human transcription factors: function, expression and Evolution. *Nature Reviews Genetics* **10**, 253

Claus Nerlov

Waxing cutaneous

When we're in the bath, our skin prevents both water from moving into our bodies and essential nutrients from leaching into the tub. But because most of us don't spend our entire lives submerged underwater, our skin's chief role is to control how much water evaporates from our bodies. In fact, the skin's role as a semi-impermeable barrier to fluid loss is so important that people suffering from serious burns often die, not as a direct consequence of their injuries, but from dehydration.

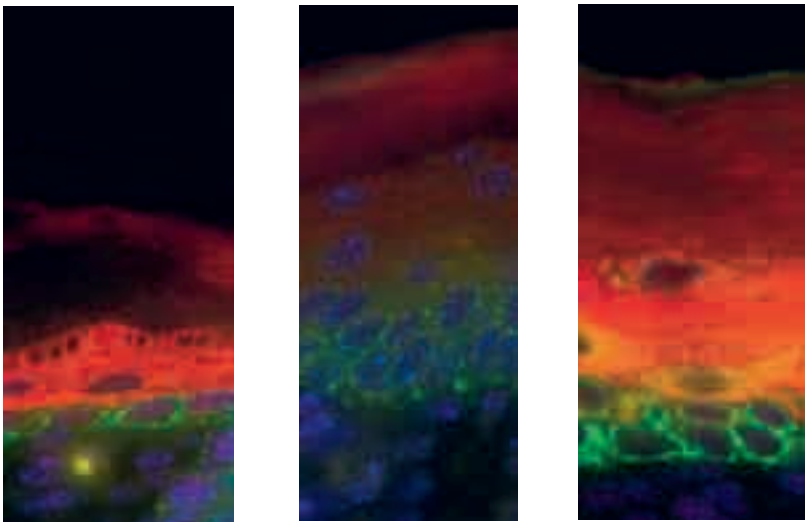
Each of us is covered by about 2 square meters of skin – about the area of a queen-size bed. For this waterproof suit to do its job, stem cells at the base of the skin replenish the layers above by producing a continuous stream of new cells initially like themselves and then a variety of specialised cell types. As a result of this continuous production, the specialised cells – which are destined to become the different layers of skin – move outwards until they are finally shed.

The reason why stem cells stop producing more of themselves and start producing other types of cells had flummoxed researchers ever since the existence of skin stem cells was uncovered in the 1970s. So cell biologist Claus Nerlov and his colleagues at EMBL Monterotondo set about identifying the proteins responsible for this switch. As the processes of cell proliferation and differentiation are so tightly coupled, the researchers suspected that a single protein acted as the toggle between them. From the start, Claus' team looked to a group of proteins called C/EBPs, because they are known to regulate this shift in other parts of the body.

To examine whether C/EBPs form the potential stem cell switch in skin, Claus and his group, in collaboration with colleagues at the Centro de Investigaciones Energéticas, Medioambientales y Tecnologicas (CIEMAT) in Madrid, looked at how skin forms in mice in their absence. Eliminating C/EBPs from all cells prevented mouse embryo development at an early stage – too early for the researchers to observe the proteins' role in skin formation. And so Rodolphe Lopez, a postdoc working in Claus' lab, created a mouse strain in which these proteins could be deactivated in skin cell tissue while continuing to function in all other tissues. This allowed the mouse embryo to develop more or less normally.

The researchers found that mice without C/EBPs had taut, shiny skin that didn't act as a barrier to water, and that they died of dehydration shortly after birth. When they looked more closely at the skin of these young mice, they saw that the cells at the base of the skin were not differentiating into mature skin cells; instead, they remained in their immature state and continued to multiply.

"We saw what was happening at the cellular level, but we didn't know what was happening at the molecular level," explains Claus. To investigate, he and his team introduced mutations into different stretches of fully working copies of the genes that encode the C/EBP proteins. They found that some mutations prevented the C/EBPs from stifling the activity of proteins that halt cell proliferation, whereas other mutations blocked C/EBPs from binding



In normal skin (left), the stem cells at the base (green) differentiate into skin cells (red). In mice whose skin has neither C/EBPα nor C/EBPβ (middle), stem cells appear in upper layers of skin, and there are no differentiated skin cells. In skin where C/EBPα is present but has lost its capacity to interact with E2F (right), skin cells start differentiating abnormally, before they have properly exited the stem cell programme (yellow/orange).

to DNA and activating genes that signal to the cells to start differentiating. "Essentially, we found that C/EBPs are doing two things – stopping proliferation and starting differentiation," says Claus. "C/EBPs tie together these two processes by doing both of them simultaneously, which makes a lot of sense when you consider how closely coupled these processes are in the cell."

In addition to confirming their suspicions, this experiment hinted at something more unexpected, admits Claus. In adult mice with no working copies of the C/EBP proteins, his team discovered a slew of genes that are usually expressed only in embryonic stem cells but were still active in the fully mature skin. These same genes are expressed in highly aggressive skin cancers, suggesting that C/EBPs are actively suppressing genes usually associated with malignancy – an important find for anyone interested in understanding how epithelial cancers like skin, breast, and oral cancers develop.

But the role of C/EBPs in cancer goes more than skin deep. Since 2001, Claus has been studying a form of blood cancer called acute myeloid leukaemia, and his pursuit has taken him deep within our bones. He explains that in the bone marrow of a healthy individual, blood stem cells churn out more cells like themselves and generate others that are destined to become the medley of cells that form our blood – for example, red blood cells, white blood cells and platelets.

Such cellular differentiation is hierarchical, with blood stem cells at the top giving rise to immature progenitor cells, which then become more and more differentiated to form many types of specialist cell. But when a person develops leukaemia, there is an explosion of one type of immature progenitor cell that overtakes the production of red blood cells in the bone marrow and eventually results in anaemia and death. "Fifty years ago, before chemotherapy, this type of leukaemia killed you within weeks," says Claus.

When doctors looked at the DNA of patients suffering from this cancer they found that 90% of the tumours harbour two types of mutations in the C/EBP gene. Claus wondered how these mutations affect the functioning of the resulting C/EBP protein so Oxana Bereshchenko, a postdoc in Claus' lab, created mice with each mutation individually, and then another strain with both mutations. From these mice they extracted blood, which they injected into the bone marrow of a healthy mouse, and watched to see how the cancers developed.

They saw that mice with both mutations became sick and died more rapidly than the mice with only one. They therefore concluded that these mutations have complementary roles: one of them allows the expansion of the malignant population of stem cells, whereas the other helps the stem cells to differentiate into the progenitors, which ultimately grow to such high densities that they kill the mouse.

"This has important implications for the treatment of these cancers," Claus says. If clinicians use drugs to kill the population of progenitors – which seem to be the problem when a patient's blood is examined – then they are aiming at the wrong target, he explains. This is because the mutations are occurring in the stem cells that seed the cancer and not the progenitor cells that maintain it, and so even if the progenitors are eliminated, the stem cells will just give rise to more.

Lopez R, Garcia-Silva S, Moore S, Bereshchenko O, Martinez-Cruz A, Ermakova O, Kurz E, Paramio J & Nerlov C (2009) C/EBPα and β couple interfollicular keratinocyte proliferation arrest to commitment and terminal differentiation. *Nat Cell Biol* **11**: 1181-1190

Bereshchenko O, Mancini E, Moore S, Bilbao D, Mansson R, Luc S, Grover A, Jacobsen S, Bryder D & Nerlov C (2009) Hematopoietic Stem Cell Expansion Precedes the Generation of Committed Myeloid Leukemia-Initiating Cells in C/EBPα Mutant AML. *Cancer Cell* **16**: 390-400

Decloaking the germ

The bacterium *Listeria* infects humans through contaminated food. Once in the gut, this pathogen can be life-threatening if contracted during pregnancy or by newborns and those with weakened immune systems. But for most people, an encounter with *Listeria* causes nothing more than vomiting and diarrhoea because our immune system recognises the long, propeller-like projections on the bacterial surface – called flagella – and mounts an assault on *Listeria* until it is wiped out. *Listeria*, however, has evolved a way to dodge this fate.

To anyone who has ever tried to cross enemy lines, this bacterium has an enviable ruse. After detecting the warmth of the human body, *Listeria* shuts down the production of flagella – the equivalent of enveloping itself in an invisible cloak. It does this by activating a protein called motility gene repressor, or MogR for short, which binds to DNA close to the flagella gene and suppresses it.

While thinking about how to rid humans of this irksome and sometimes serious threat, researchers Aimee Shen and Darren Higgins at Harvard Medical School in Boston, USA had the idea of stopping MogR from doing its job. This strategy would allow them to de-cloak the bacteria, leaving it exposed to the wrath of the immune system. But to do this, they needed to understand how MogR recognises its binding site.

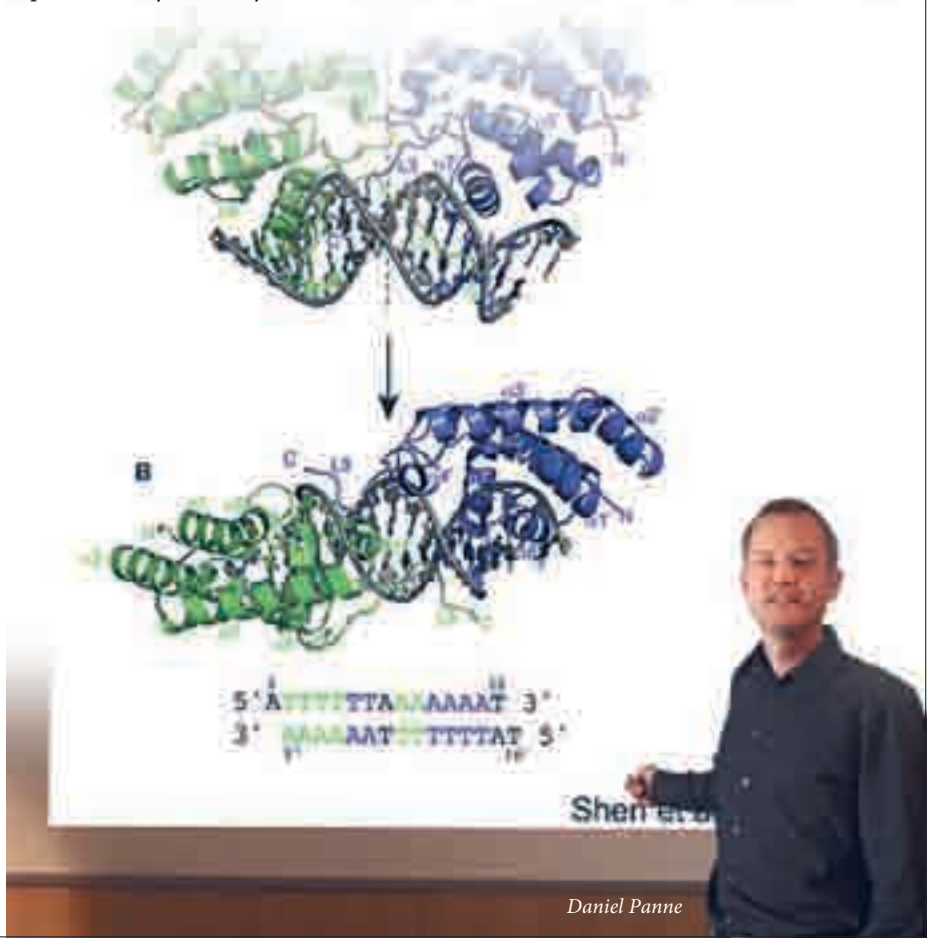
The scientists knew that the MogR protein regulates about 25 genes in the bacteria's genome. These binding sites are very rich in As and Ts, two of the four nucleotide 'letters' that make up the DNA alphabet, and are palindromic – that is, they read the same forwards as they do backwards. Knowing this, the researchers looked for where this particular motif occurred in the bacteria's DNA; to their dismay, they found more than 500 matches. This suggested that the DNA sequence of this regulatory motif was not sufficient to explain MogR's binding pattern. To understand how MogR selects for specific binding sites, the researchers turned to structural biologist Daniel Panne at EMBL Grenoble.

Daniel's lab was able to determine the crystal structure of the MogR protein bound to its DNA-docking site. From this, he saw that MogR recognises both the DNA's sequence and also its overall warped shape, which is created by the particular chemical properties of the AT-rich region that cause the DNA to bend to a 52-degree angle. This provided the first example of DNA shape-dependent recognition.

But this was not the whole picture: Daniel also noticed that MogR binding relied on specific electrostatic forces – similar to the static forces experienced by clothes just out of the dryer – that are generated by the different charges of the chemical bases that make up the docking site. Importantly, only the 25 MogR-regulated genes boast docking sites that rely on DNA sequence, shape and electrostatic forces for binding.

Scientists have since found more examples of proteins that use similar modes of DNA recognition to MogR. Daniel explains that these findings not only provide insight into a new mode of DNA recognition that is seemingly more common than suspected, but could also help bacteriologists develop a drug that offers protection against *Listeria* infection by disrupting MogR binding and thus stopping the bacterium from donning its invisibility cloak.

Shen A, Higgins D, Panne D (2009) Recognition of AT-Rich DNA Binding Sites by the MogR Repressor. *Structure* **17**: 769–777



Daniel Panne

A Year in the Life of EMBL

April

50th Birthday Newsletter

The first edition of the EMBL etcetera newsletter was published in 1999, had eight pages, and its stories included the opening of the Monterotondo outstation and the construction of the new Kinderhaus. So it was definitely time for a substantial redesign when the newsletter celebrated its 50th issue in April 2009. It is now printed in four colours and has an attractive new layout – the first revamp since the 25th issue. Originally conceived to communicate news and to keep alumni informed about EMBL, it now covers a broad range of topics. As Sarah Sherwood, the first editor of the newsletter puts it: "Reading EMBL etcetera now is like poking your head into an open window at the lab."

First ever European Learning Laboratory for the Life Sciences in Spain

Together with Rossana De Lorenzi, Education Officer in Monterotondo, EMBL alumna Teresa Alonso organised a course for biology teachers at the Instituto de Biología y Genética Molecular in Valladolid in April, with seminars in Spanish and practical activities in English. The first LearningLAB in Spain was a huge success, with a great atmosphere among the teachers and scientists. As a result, the teachers have set up a local network with a virtual platform to exchange useful ideas and teaching materials.

2009

May

New web pages

The newsletter was not the only medium to be completely revamped. The EMBL web pages also received a fresh, modern design with their re-launch in mid-May. The site is now managed via a Content Management System, which allows the pages to be updated more easily. Visitors can choose their EMBL site from the new 'portal' page and rotating banners promote conferences and other events. The immediately recognisable intranet pages contain the day's events, news, announcements, and quick links to downloads, the library and all the externally available sections such as research, services, training, press releases and jobs.

Spanish students visit the Laboratory

A group of biotechnology students from the Universidad Francisco de Vitoria in Madrid visited EMBL Heidelberg on 8 May as part of a tour of universities and research institutes in Europe. The visitors, who were particularly interested in the PhD Programme, learnt about the way EMBL works from the Head of Communications Lena Raditsch and the Dean of Graduate Studies Helke Hillebrand. The visit was rounded off with two presentations by EMBL predocs on their research projects and a look inside the laboratories.

2009

Communicators of science

They could not have chosen a better time for their stay – when Nicla Panciera and Adam Gristwood came to EMBL for their placement as part of a course run by the European Initiative for Communicators of Science (EICOS) their schedule included Career Day, Lab Day and EMBLEM's 10th birthday celebrations as well as interviews with scientists from all Units and career stages. The EICOS programme aims to improve communication between scientists and journalists to make research more intelligible to the public and is open to journalists from all over Europe and from all backgrounds. Both journalists were particularly impressed by the good communication skills of EMBL scientists and the institute's interdisciplinarity.

June

EMBLEM's 10th anniversary

On 19 June, to celebrate 10 years of successful technology transfer, EMBLEM invited both their clients and EMBL staff to a Mediterranean buffet, complete with a cocktail bar, firedancers and a magician. A wonderful array of dishes were prepared by Claus Himburg and his canteen staff, musical entertainment came from the live band 'Stage Diva' and attendees enjoyed the party until the small hours.

And another 10th anniversary

On 22 June the EMBL Mouse Biology Unit in Monterotondo celebrated its 10th anniversary with a scientific meeting featuring talks from various speakers including internationally renowned immunologist Klaus Rajewsky, who led the Unit before Nadia Rosenthal took over in 2001. The outstation is located on the 'Adriano Buzzati-Traverso' International Scientific Campus in Monterotondo, which facilitates collaborations with the European Mouse Mutant Archive (EMMA), the International Centre for Genetic Engineering and Biotechnology and the Institute of Cell Biology operated by the Italian National Research Council (CNR). Together, these institutes form a European centre of excellence and innovation in mouse biology. During the past 10 years scientists have generated mouse models of over 20 different human diseases, including heart failure, Alzheimer's disease and multiple sclerosis, as well as mental and behavioural disorders such as depression and anxiety.

July

EMBL takes over as EIROforum chair

In July 2009 EMBL took over as the chair of EIROforum, a partnership of the seven largest intergovernmental research organisations in Europe (CERN, EFDA-JET, EMBL, ESA, ESO, ESRF and ILL). Activities during the one-year tenure included the renewal of the Statement of Intent, which was signed with the European Commission in 2003 and which outlines the joint activities and plans for the ongoing collaboration and the biannual Assembly of the EIROforum Director Generals in November and May. In November 2009 EMBL and its EIROforum partners organised a conference on technology transfer in Heidelberg to exchange knowledge and best practices across disciplines. The mission of EIROforum is to support European science in reaching its full potential by facilitating interactions with the European Commission and the European Union, national governments, industry, science teachers, students and journalists.

RISE students visit EMBL

For the fourth year running, a group of 30 North American undergraduate biology students visited EMBL Heidelberg on 10 July as part of the RISE (Research Internship for Science and Engineering) programme funded by the DAAD (German Academic Exchange Service), through which scientists-to-be visit some of the most interesting research institutes on the continent. The students showed particular interest in internships for undergraduates and PhD opportunities at EMBL.

Heidelberg's International Summer School

From 21-23 July EMBL Heidelberg was again host to the 23 participants of the 13th International Summer Science School Heidelberg, a chance for students from Heidelberg's twin cities Montpellier, Rehovot and Simferopol to get a real insight into how scientific research is conducted and to see professional facilities close up. Education officers Philipp Gebhardt and Julia Willingale-Theune organised the programme, which involved seminars, practicals and a chance to interview a panel of EMBL scientists. Highlights were Francesco Pampaloni's 3D microscopy seminar and the visit to the EMBL frog facility.

New canteen opens in Heidelberg

It was the beginning of a new era when on 31 August the canteen opened its brand new ‘casino’ in the EMBL Advanced Training Centre at EMBL Heidelberg. With its greater capacity, it will also cater for the expanded Courses and Conferences Programme. The move was an incredible feat of organisation on the part of the kitchen team. Starting on a Friday after close of service, they packed up and shipped their entire operation, seamlessly opening again for lunch on Monday at the new site. Comments on the opening day ranged from "very elegant" to "incredibly spacious" and diners were all suitably impressed by the new facilities.

The first EMBO Meeting in Amsterdam

2009 saw the launch of The EMBO Meeting, a new platform for life scientists to meet and enjoy outstanding lectures, discussions, workshops and networking opportunities. As in previous years at the ELSO Meeting, EMBL set up its stand to promote its activities and career opportunities. At the first meeting in Amsterdam on 29 August to 1 September, the focus was on the dynamics, maintenance and evolution of chromosomes, signalling pathways in development and cancer, and stem cells. In a special lecture, the UK’s Astronomer Royal Martin Rees gave a glimpse into the cosmos, explaining what happens at the fringes of our galaxy and how likely we are to find other thinking creatures like ourselves.

August

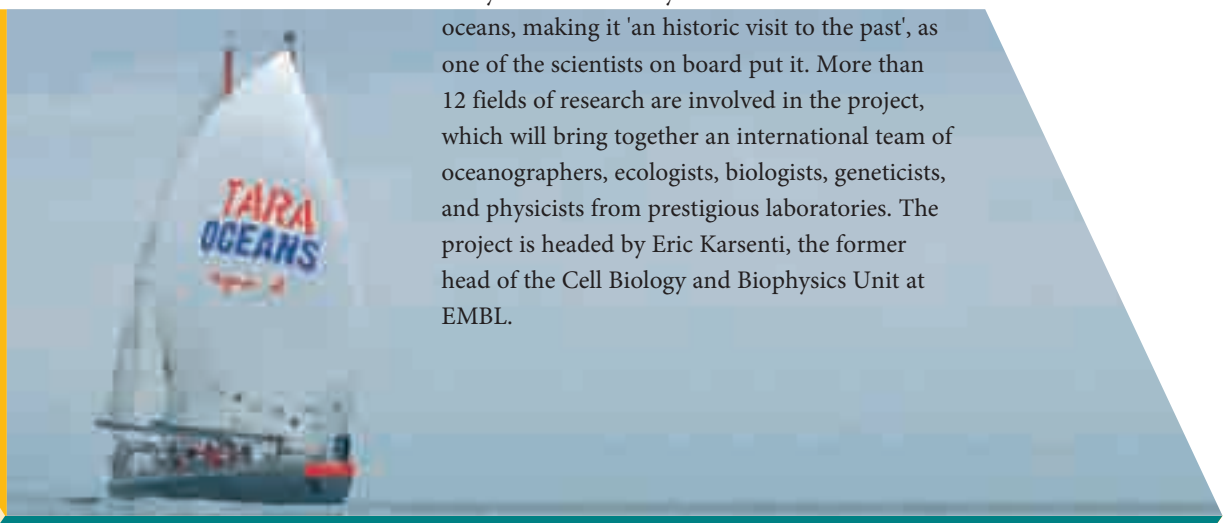
European Union Contest for Young Scientists

Cecilia Thomas was one of last year's winners of the European Union Contest for Young Scientists (EUCYS) with her discussion entitled ‘Antimicrobial peptides – the new weapon against multiple drug resistance?’ Cecilia is from Denmark and won a placement at EMBL Heidelberg at the contest in September 2008. She spent one week at the institute meeting scientists from various Units who gave her a unique insight into the day-to-day life of a scientist. After her stay at EMBL she went back to Copenhagen to begin her studies in human life science engineering.

September

Tara sets sail

Three years to sail around the world, 150,000 km to cover, 60 ports to visit and 50 participating laboratories – these are only a few of the figures behind the enormous project called the Tara Oceans Expedition. The vessel left Brittany on 5 September for a three-year scientific adventure to study the marine ecosystems across the world's oceans, making it 'an historic visit to the past', as one of the scientists on board put it. More than 12 fields of research are involved in the project, which will bring together an international team of oceanographers, ecologists, biologists, geneticists, and physicists from prestigious laboratories. The project is headed by Eric Karsenti, the former head of the Cell Biology and Biophysics Unit at EMBL.



2009

Forum for Young Life Scientists at the DKFZ

On 10-11 September EMBL predocs got together with PhD students from various research institutes in Heidelberg to hold the first Heidelberg Forum for Young Life Scientists at the DKFZ. The programme contained six sessions, each of which was organised by one of the institutes and EMBL's segment, 'Through the Looking Glass: Understanding Molecular Biology through Evolution', featured Nobuhiko Tokuriki from Israel'sWeizmann Institute of Science as the keynote speaker. The Forum provides an ideal opportunity to bring together the diverse know-how of scientists from the Heidelberg institutes and expertise of renowned researchers from abroad. With PhD students of different nationalities and diverse educational backgrounds, the atmosphere was highly interdisciplinary and interactive with plenty of networking among students from the different institutes.

A packed Operon

October saw a couple of highlights in this year's EMBL Science & Society Programme when Jorge Cham and Mattieu Ricard – two speakers who couldn't be more different – both packed the Operon to the rafters with their EMBL Forum lectures. Jorge Cham, creator and artist of 'Piled Higher and Deeper'– the comic strip about the trials and tribulations of doing a PhD – is something of a folk hero among predocs all over the world and his followers flocked to hear his talk on 'The Power of Procrastination' whereas Matthieu Ricard is probably the first Buddhist monk to set foot on EMBL premises. His talk entitled 'Train your Mind, Change your Brain – Cultivating the Inner Condition for Genuine Happiness' not only impressed the general public who turned up in large numbers to hear the Dalai Lama's French interpreter but also the hard-core scientists at EMBL.

2009

October

Puzzles in Biology

More than 200 attendees – including students and visitors from as far afield as Armenia, Israel and India – learnt about 'Puzzles in Biology' at the 11th EMBL PhD Student Symposium on 29 October. Distinguished speakers included Stefan Hell, the inventor of Stimulated Emission Depletion microscopy (STED), Ari Helenius, winner of the 2007 Marcel Benoist prize, the highest research award in Switzerland, Kim Nasmyth and Pierre Chambon. Highlights included the presentation of the PhD Symposium Writing Prize, which this year went to EMBL-EBI's Diva Tommei for 'The Dark Side of Stem Cells'.

Science Days in Rust

Each year Science Days, a major science festival hosted by the Europa Park in the south of Germany, attracts thousands of visitors of all ages. As a mini celebration of the Year of Darwin, this year the stand, organised by EMBL's European Learning Laboratory for the Life Sciences, featured marine worms (Platynereis dumerilii) an evolutionary fossil to explain eye development linked to studies performed at EMBL, an Axolotl couple, and various other activities including "genome tape measures" aimed at bringing the modern concepts of evolution and evolutionary research nearer to the public.

The big switch on

On 16 November Prof. Annette Schavan, Germany's Federal Minister for Education and Research, helped push PETRA-III's 'on' button to inaugurate the world's most advanced synchrotron radiation source. With PETRA-III, scientists should gain fundamental new insights into the structure of matter as it opens up completely new opportunities in the field of structural biology, for example in the research of protein structures. For the past two-and-a-half years, EMBL Hamburg's campus partner, the German Synchrotron Research Centre (DESY), has been upgrading its beamline facilities to provide modern, world-leading services. Out of PETRA-III's 14 beamlines, three are being designed and built by EMBL.

November

EMBL-EBI open day

EMBL-EBI held its second Masters Open Day of the year on 3 November. The day was mainly targeted at Masters students with a view to introducing and promoting the PhD scheme at EMBL-EBI. It was the perfect opportunity for young scientists to find out more about opportunities at Europe's main centre for bioinformatics and Manchester University sent a whole group of students from their Bio-informatics course for the day. The popular lunchtime demonstration session also gave visitors the chance to quiz the experts face-to-face on the EMBL-EBI's core resources.

A GIANT undertaking

EMBL Grenoble's campus, the Polygone Scientifique, is to be developed into a world-class science and technology park named GIANT – an 'ecosystem' of innovation – in an initiative supported by the French government. GIANT's first construction projects are already underway and mark the beginning of a €500 million overhaul for the area, which has long boasted a top-class research infrastructure that includes the ESRF, the ILL, EMBL and the CNRS, as well as three centres of technological excellence and several high-level university programmes. With the planned new teaching and research buildings, recreational facilities and meeting places for researchers, transport links and sustainable housing, it is projected that the site will welcome 20,000 scientists and students and 10,000 inhabitants by the year 2015. The entrance to the site will also feature a Visitors' Centre, which will present the work of all the campus' scientific institutes to the growing numbers of interested members of the public using videos, models and interactive exhibits.

December

A fond farewell

After eight years at EMBL, Administrative Director Bernd-Uwe Jahn retired at the end of 2009. The man described by Chair of Council Eero Vuorio as "a warm-hearted connoisseur of wine and good food with a great sense of humour", as well as someone "with a thorough knowledge and understanding of the rules and procedures of Council and the Laboratory, who could be depended on for immediate and correct answers" handed over his responsibilities to Ralph Martens. On 8 March 2010 a newly planted area of the EMBL Heidelberg campus was designated 'Uwe's Orchard' and staff from all EMBL sites said goodbye to Bernd-Uwe Jahn with a farewell dinner in the new canteen, followed by a party in the EMBL Advanced Training Centre lobby.

2010

PSB students meeting in Grenoble

Grenoble's Partnership for Structural Biology (PSB) has 80 students from widely diverse backgrounds, but the biologists, chemists, physicists and computer scientists all share a common interest: structural biology. The annual PSB Student Day, organised by and for the students of the partnership, is a unique opportunity for the smorgasbord of interests in the different communities – EMBL Grenoble, UVHCI, the Institut de Biologie Structurale, the Institut Laue Langevin (ILL) and the ESRF – to come together. This year's meeting on 26 January at the ILL began with a poster session, after which senior students presented their results followed by shorter presentations by the first-year students. The goal of the PSB Student Day is to give all young scientists in the PSB a chance to interact with each other and to present and discuss their research projects in a relaxed and informal atmosphere.

January

EMBL Corporate Partnership Programme

Managing directors, vice presidents and other top representatives from all 15 partnership companies gathered on 21 January for their first official event, which included scientific talks and a dinner in EMBL Heidelberg's new canteen. The EMBL Corporate Partnership Programme was established in 2008 with Leica Microsystems, GE Healthcare, Life Technologies and Olympus as founding partners to provide support for EMBL's scientific conferences, courses and other events. According to Jörg Fleckenstein, EMBL's Senior Manager of Resource Development, the funds will be used to support events in the new EMBL Advanced Training Centre and to keep registration fees at a reasonable level. Among the benefits for the companies are the association with the EMBL name, preferred access to conferences and events and an annual round-table discussion with EMBL Director General Iain Mattaj.

Alumni meeting

Around 200 participants attended the third EMBL staff-alumni reunion on Monday 8 March in the new EMBL Advanced Training Centre, a day before its official opening ceremony. Alumni proudly presented their career paths after leaving EMBL, highlighting the role of collaborations and interdisciplinary research – key elements to the success of EMBL – in their own achievements. One example was a PhD Programme recently founded by former predoc Giuseppe Testa, which brings together scholars from the life sciences and humanities to address science and society issues. Another was Eric Karsenti's Tara Oceans Expedition, which enthralled the audience and demonstrated how to bring science to the public effectively.

2010

March

The EMBL Advanced Training Centre opens its doors

On 9 March the German Minister for Education and Research, Prof. Annette Schavan, officially opened the new EMBL Advanced Training Centre. With its auditorium for an audience of 450 people and a large display area for the presentation of scientific posters – as well as a permanent exhibition of scientific achievements at EMBL – the EMBL Advanced Training Centre offers unique conditions for scientific conferences and events. With a total space of around 17,000 square metres, the building offers an excellent infrastructure for courses and conferences and convenient rooms and facilities for public events, teacher training and visitors. About 80 staff members from Administration and scientific management use the office space.

FIMM launch

On 16 March the Nordic EMBL Partnership for Molecular Medicine officially inaugurated the Institute for Molecular Medicine Finland (FIMM) in Helsinki. Together with the Centre for Molecular Medicine Norway and the Laboratory for Molecular Infection Medicine Sweden, FIMM's research will be dedicated to molecular medicine, the investigation of the basis of disease and the discovery of new treatments. FIMM has 150 employees who work on cancer, cardiovascular, neuro-psychiatric and viral diseases, carry out translational research to explore new diagnostics and treatments and promote human health via research on personalised medicine.

EMBL goes Down Under

On Monday 29 March EMBL Australia was launched in the presence of Kim Carr, Australia's Minister for Innovation, Industry, Science and Research, making Australia the first associate member state of EMBL. As a joint venture supported by the Australian government and involving the universities of Sydney, Queensland, Western Australia and Monash and the Commonwealth Scientific and Industrial Research Organisation (CSIRO), EMBL Australia was initiated in 2008. During the day's programme, Senator Carr announced the appointment of Nadia Rosenthal as the Scientific Head of EMBL Australia, the establishment of the EMBL Australia PhD Programme, and the appointment of Marcus Heisler, who is currently at EMBL Heidelberg, as the first group leader within the faculty development programme.

March

Science & Society minisymposium in Hinxton

Is there a risk that private interests undermine the quality and reliability of research? And if so, what can scientists do to safeguard against this? These were among the issues addressed in a series of four talks followed by a panel discussion at this year's EMBL-EBI Science and Society symposium held during the Cambridge Science Festival in March 2010. Speakers included Tim Hubbard from the Wellcome Trust Sanger Institute, David Searls, who serves on several advisory panels for university-based research institutes in the US, Stuart Parkinson, Executive Director of Scientists for Global Responsibility, UK, and Gábor Lamm, Managing Director of EMBLEM, Heidelberg, Germany. The symposium was chaired by Sir John Sulston, Nobel Prize winner for Medicine in 2002.

Index of group & team leaders

A

Arendt, Detlev 13

B

Birney, Ewan 67, 87
Boettcher, Bettina 8
Bork, Peer 6, 13, 21
Brazma, Alvis 11
Briggs, John 18

C

Cusack, Stephen 33

E

Ellenberg, Jan 33, 63, 84
Ephrussi, Anne 52

F

Flicek, Paul 67
Frangakis, Achilleas 8
Furlong, Eileen 20

G

Gavin, Anne-Claude 8
Gibson, Toby 77

H

Huber, Wolfgang 10

K

Kersey, Paul 65
Korbel, Jan 78

L

Ladurner, Andreas 30
Le Novère, Nicolas 59
Luscombe, Nicholas 87

M

Márquez, José 43
Mattaj, Iain 47, 53
Müller-Dieckmann, Jochen 57
Müller, Christoph 49, 81
Müller, Jürg 80, 81

N

Nerlov, Claus 40, 89

P

Panne, Daniel 91
Pepperkok, Rainer 84
Pillai, Ramesh 62

R

Rosenthal, Nadia 23, 40
Round, Adam 45
Russel, Rob 8

S

Schneider, Reinhard 85
Schultz, Carsten 41, 57
Steinmetz, Lars 10, 37, 79
Stelzer, Ernst 30

T

Treier, Mathias 73
Tucker, Paul 76

W

Wilmanns, Matthias 55

Annual Report 2009-2010

Publishe dby the EMBL Office of Information and Public Affairs

Lena Raditsch, Sonia Furtado

DG Report

Iain Mattaj

Science writers

Claire Ainsworth, Jennifer Carpenter, Sonia Furtado, Lucy Patterson, Anna-Lynn Wegener

Graphics and cover design

Petra Riedinger

Photography

EMBL Photolab (Hugo Neves, Udo Ringeisen, Marietta Schupp)

Layout and composition

Nicola Graf

Printed by

ColorDruck Leimen, Germany

Acknowledgements

Page 2: The National Library of Israel, Eran Laor Cartographic Collection, Shapell Family Digitization Project and The Hebrew University of Jerusalem, Department of Geography – Historic Cities Research Project

Page 14: Nüchter and Holstein, Universität Heidelberg (*Nematostella*), Kirt L. Onthank (*Strongylocentrotus*), The Field Museum, Chicago (Cambrian sea)

Page 34: Douglas Jordan/CDC

Page 38: Marina Lamacchia/INSERM

Page 45: (top) Pedro Luís Rodriguez

Page 50: Carlo Petosa/IBS

A special thank you to the EMBL group leaders and Heads of Units for their cooperation.

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